Improving the synthetic biology toolbox of Synechocystis sp. PCC 6803

Caicedo Burbano, S.P.

Link to publication

Improving the synthetic biology toolbox of *Synechocystis* sp. PCC 6803

Patricia Caicedo

Improving the synthetic biology toolbox of *Synechocystis* sp. PCC 6803

Improving the synthetic biology toolbox of *Synechocystis* sp. PCC 6803

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen

op woensdag 7 april 2021, te 13.00 uur

door

Sandra Patricia Caicedo Burbano

geboren te Popayan

Promotiecommissie

Promotor:

    prof. dr. L.W. Hamoen   Universiteit van Amsterdam

Copromotor:

    dr. F. Branco dos Santos   Universiteit van Amsterdam

Overige leden:

    prof. dr. S. Brul   Universiteit van Amsterdam

    prof. dr. K.J. Hellingwerf   Universiteit van Amsterdam

    prof. dr. J. Hugenholtz   Universiteit van Amsterdam

    prof. dr. M.J. Barbosa   Wageningen University & Research

    dr. M.T. Wortel   Universiteit van Amsterdam

    dr. A.D. van der Woude   Photanol B.V.

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Table of contents

Preface

In recent years there has been an increasing awareness of the detrimental ecological consequences caused by our dependence on fossil fuels. It is of major importance the need of developing sustainable alternatives that allow the reduction of CO_2 emissions and the diminution and eventual replacement of the use of fossil fuels. A strategy that was initially explored consisted on the production of biofuels using plants to generate substrates (sugars) that can be subsequently used in microbial fermentation processes. More recently, a more efficient alternative has emerged, the use of cyanobacteria to directly transform atmospheric CO_2 into biofuels and other useful products.

Cyanobacteria are photosynthetic prokaryotic organisms with the ability of using the energy in light to convert water and CO_2 into biomass and O_2. They pose several advantages over plants as cell factories. Namely, cyanobacteria tend to have a faster growth rate than plants; they do not compete for arable land; and their genetic content can be easily modifiable. *Synechocystis* sp. PCC 6803, a model organism of cyanobacteria, has been extensively investigated as a suitable option for the sustainable production of (several) different compounds. This host can be engineered, acting as a chassis in which components or pathways can be added or removed, resulting in the direct conversion of CO_2 into products of interest, fueled by solar light and generating O_2 as the only by product.

The use of microorganisms as cell factories frequently requires extensive molecular manipulation, but in comparison to the tools available for other model organisms such as *Escherichia coli* and yeast, the tools available for *Synechocystis* are still rather limited. In order to up-scale their usage into fully developed industrial processes, it is crucial that some aspects can be improved by increasing the synthetic biology toolbox of cyanobacteria, especially in the model organism *Synechocystis*.

In this thesis, we wanted to contribute to the development of synthetic biology tools that can facilitate the engineering of cyanobacteria. This thesis starts with an overview of cyanobacterial cell factories. **Chapter 1** includes the recent advances in the development of synthetic biology tools in cyanobacteria, the challenges that some of their biological characteristics represent and the description of biological parts, regulatory elements and other mechanisms commonly used to engineer them.

Chapter 2 describes the development of a new method to create fully segregated integrative genomic libraries in polyploid organisms. This was accomplished through the use of a counter selection system, which only allows the survival of cells that have integrated an element of the library in all the copies of its chromosome. The functionality of this method was tested in the model organism *Synechocystis* to find native genomic sequences capable of acting as promoter sequences. This method can be applied in other polyploid organisms to screen for genetic elements to control gene expression.

In **Chapter 3** we constructed a genomic library to find a suitable promoter able to express the glyoxylate shunt enzymes from *Chlorogloeopsis fritschii* in *Synechocystis* in a timely manner. *Synechocystis* can be engineered to produce fumarate during the day and night by removing both the *fumC* and *zwf* genes. The incorporation of the glyoxylate shunt enzymes can in theory increase the night-time production of fumarate even further, only if the enzymes are active exclusively during the night. The library was screened by natural selection under continuous cultivation in a fluctuating day/night regime. It was not possible to achieve the stable timely expression of the shunt enzymes over time, resulting in the cells selecting against the expression of the enzymes. Nevertheless, the

results validate the potential of the method to find genetic elements performing under specific conditions.

Chapter 4 describes the construction of different variants of a fumarate biosensor. This biosensor was made fusing two native *E. coli* sensor histidine kinases, sensor DcuS and osmolarity sensor EnvZ (DcuSZ), and coupling the two-component system to the expression of a fluorescent protein in response to extracellular fumarate. Nine different DcuSZ were created, which differ in the length of the linker region between the DcuS and EnvZ functional domain. Three of them (DcuSZ 1, 6, 9) were functional biosensors with a combined signal dynamic range of 0.1-20 mM fumarate in *E. coli*. The variant DcuSZ 6 was the best candidate found to be used as fumarate biosensor in *Synechocystis*.

In **Chapter 5**, we wanted to analyze the effect of the insertion site on the expression of heterologous genes in *Synechocystis*, due to factors such as the surrounding nucleotides to the gene of interest (which may have a regulatory role), the chromosomal three-dimensional structure, or the proximity to the origin of replication. This chapter describes the expression of a reporter cassette when inserted in six different loci previously identified as neutral sites in the chromosome of *Synechocystis*. The results did not show a big variation in the expression between these different locations. This supports the idea that in this relatively slow-growing organism, and for at least these six neutral sites, the levels of expression are not very much affected by chromosomal location *per se*.

Finally, in **Chapter 6** the results of this thesis are placed in a broader context and the current challenges of this research field and potential solutions are discussed.

General Introduction

Patricia Caicedo Burbano

*Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences,
University of Amsterdam, Netherlands*

Cyanobacteria

Cyanobacteria are an ancient and extremely diverse group of prokaryotic organisms, widespread in all kind of habits and ecological niches. They are distributed from equatorial to polar regions, including brackish, fresh and salt water and many terrestrial habits including deserts[1,2]. Cyanobacteria are gram-negative photosynthetic bacteria that can tolerate extreme conditions such as alkaline, saline, freezing cold or scalding hot environments. They play important roles in the global nitrogen and carbon cycles[2].

Cyanobacteria evolved billions of years ago. They are considered the ancestors of algae and higher plants[3], and they are responsible for around half of the primary production (production of organic compounds by autotrophs) on earth[4]. Cyanobacteria are characterized for being the only known prokaryotes able to perform oxygenic photosynthesis, converting solar energy into chemical energy by reducing CO_2 to sugars, using water as the source of electrons and producing O_2 as a by-product. Cyanobacteria can also engage in other forms of metabolism, such as fermentation, to adapt to different environments and conditions, for example when sun light is not available[5].

Cyanobacteria as cell factories

Due to their ability of efficiently harvest atmospheric CO_2 as a carbon source, getting energy by absorption of sunlight (the most abundant form of renewable energy), photosynthetic cyanobacteria have been studied as promising cell factories for the sustainable production of desired chemicals[6]. Their unicellular physiology, photosynthetic growth, and relative rapid cell division, make cyanobacteria an ideal organism for the production of a broad range of compounds such as pharmaceutical, chemicals, food supplements and biofuels[6,7]. Furthermore, the use of cyanobacteria can contribute to the generation of sustainable energy, mitigating the effects of global warming.

Another characteristic that makes cyanobacteria attractive from a genetic engineering perspective is that some species are naturally competent and able to incorporate foreign DNA in their genome by recombination, as well as receive conjugative plasmids. Using strategies of synthetic biology and genetics and metabolic engineering, the photosynthetic mechanism of cyanobacteria has been modified to synthesize valuable products such as ethanol[8], ethylene[9], isopropanol[10], butanol[11] and other alcohols, sugars, organic acids, fatty acids, hydrocarbons, biofuel and bioplastic precursors[12,13,14,15,16,17].

Synthetic biology in cyanobacteria

Cyanobacterial potential has been exploited through the development of molecular tools[18,19] characterized for diverse cyanobacterial species, such as the widely studied *Synechocystis* sp. PCC 6803[20], *Synechococcus elongatus* PCC 7942[21], *Nostoc* sp. PCC 7120[22] and *Anabaena* sp. PCC 7120[23]. Also, for some faster growing strains, such as *Synechococcus* sp. PCC 7002[24] and particularly *Synechococcus elongatus* UTEX 2973[25]. In addition to those molecular tools, the increasing capacity to manipulate genetic sequences and the sequencing of the genome of more than 270 cyanobacterial species[26] have facilitated the development of refined tools in the fields of metabolic engineering and synthetic biology[27].

Synthetic biology aims to construct and/or restructure biological systems with programmable and predictable behavior, using characterized, standardized and interchangeable biological parts[28]. Some examples of these parts are promoters, ribosome binding sites (RBS), coding sequences, terminators and backbones. These biological parts, together with standardized methods for the assembly and manipulation of genetic elements, and predictive computer models, form the so called "toolboxes" of components that can be used to modify an organism[29] (Figure 1.1). Similar to other engineering processes, these parts must be characterized, and if necessary modified, for the context in which they work. Due to the complexity of biological systems, development and characterization of biological parts in the desired host is crucial[30].
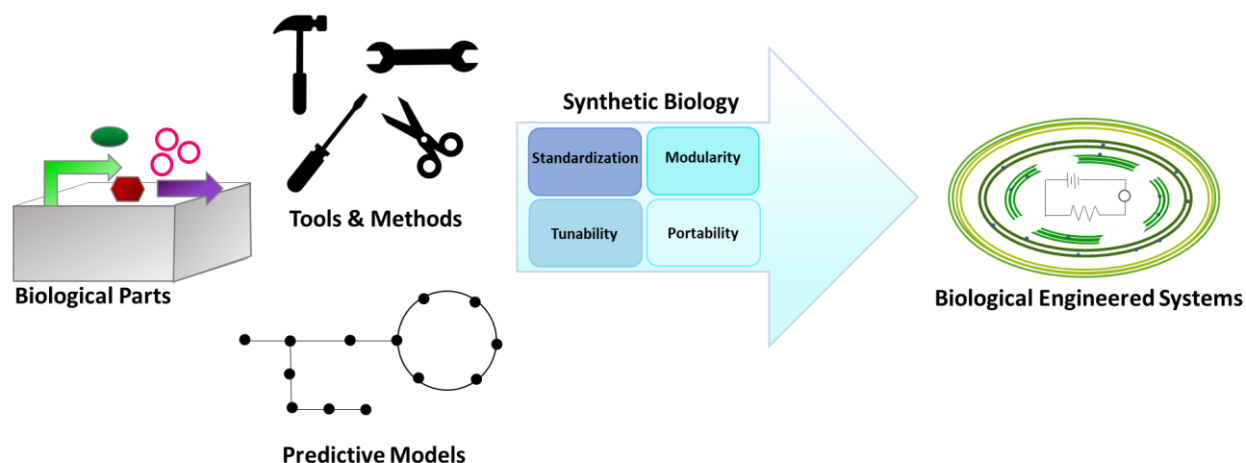
Figure 1.1 Scheme of the process to develop a synthetic biological system with programmable and predictable behavior using engineering principles. These systems can be constructed using a toolbox, composed by a catalog of characterized, standardized and interchangeable biological parts; tools and methods to edit the genome and assemble the parts; and predictive models to minimize the trial and error process (Adapted from Sengupta *et al.*[31]).

The advances of synthetic biology in cyanobacteria have been extensively reviewed[18,22,27,30,31,32,33,34,35,36]. However, despite the increasing progress and great potential of cyanobacteria to be used as cell factories, the tools available to modify cyanobacteria lies behind the tools developed for other model organisms like *Escherichia coli* and *Saccharomyces cerevisiae*. This is in part due to the challenges that particular cyanobacterial biological characteristics present. These include traits as slow growth, the presence of a circadian rhythm, polyploidy and differences in polymerases.

Biological differences with other bacteria

In comparison with other model organisms such as *E. coli* and yeast, cyanobacteria generally have a low maximal growth rate, which can limit the speed of biotechnological processes. The doubling time reported for *Synechocystis* is of 12-16 hours[37]. There are nevertheless, other cyanobacteria with faster growth rate, for example *Synechococcus* sp. PCC 7002, with a reported doubling time of 2.5-3.5 hours[38,39] and *Synechococcus* UTEX 2973, with reported doubling time of less than 2 hours[40], which are being consider as alternatives to overcome this limitation.

An interesting feature of cyanobacteria is the presence of a circadian rhythm. They were the first prokaryotes reported to have the circadian clock. This is an essential mechanism present also in eukaryotes for timing cellular processes to respond to changes in the environment and the daily variations in energy and nutrient availability[41]. The regulator effects of this mechanism make of it an important factor to take into account for cyanobacterial biotechnology, because almost the entire cyanobacterial cellular machinery, including metabolism, cell division and gene expression, experiences circadian control[42,43].

A characteristic that presents a major challenge is the polyploidy of some cyanobacteria. This requires that genetic modifications be integrated in all the chromosomes of the cell. One of the most widely used species, *Synechocystis*, has multiple chromosome copies per cell, that can vary depending on growth phase and cultivation conditions[44,45]. In order to obtain a stable mutant strain, it is necessary that the inserted cassette achieve full chromosomal segregation. The mutants must be selected (usually with an increasing concentration of antibiotic) to avoid the cells to revert to the original phenotype. Selection can take weeks (or longer), making this a time-consuming procedure.

Polyploidy is also challenging for high throughput screening. Every genetic modification has to be tracked on a chromosome by chromosome basis, until finding a clone that has the gene of interest fully segregated. Traditional screening methods like PCR rapidly become labor intensive and not entirely reliable. For this reason, the development of new techniques to speed up the segregation process and improve the screening methods is highly relevant.

Another particularity of cyanobacteria is the difference in their RNA polymerase. In other bacteria, the RNA polymerase consists of an apoenzyme of five subunits, $\beta\beta'\alpha_2\omega$, which binds a sigma factor forming $\beta\beta'\alpha_2\omega\sigma$, the holoenzyme that can bind a promoter and start transcription[46]. However, in the cyanobacterial RNA polymerase the β' subunit is split into two parts: the γ and the β' subunits. Cyanobacterial RNA polymerase holoenzyme only works with sigma (σ) factors from the σ^{70} family. Besides, cyanobacteria have different consensus promoter sequences than other bacteria[34,42].

These differences in the transcriptional machinery complicate the use of biological parts across different organisms, because the parts designed for other organisms may not be directly applied to cyanobacteria, or they might perform differently in them. Hence, the importance of the characterization of the different parts within the environmental context of cyanobacteria for the design of tools specific to engineer them in a stable and controlled way[32].

Biological parts in cyanobacteria

The majority of the studies on biological parts have been done on cis-acting regulatory elements (CRE). Although many parts have been characterized, only a fraction of parts derived from cyanobacteria have been added to the BioBricks registry, which have been summarized by Sengupta *et al.*[31] Even though most of the biological parts characterized in cyanobacteria are used to modify gene expression through transcriptional regulation, other parts have been developed to perform post transcriptional, translational or post translational regulation[30] (Figure 1.2).
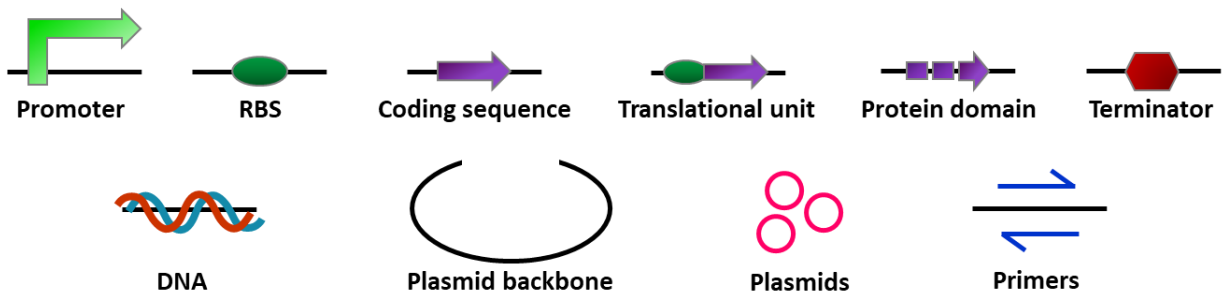
Figure 1.2 Standardized representation of the types of biological parts annotated in the registry of standardized biological parts (Adapted from http://parts.igem.org/Catalog[47]). Biological parts (biobricks) are used to construct engineered biological systems with programmable and predictable behavior.

Transcriptional regulation

Promoters are regions of DNA where the RNA polymerase and transcription factors bind to initiate transcription. They are essential elements involved in the transcriptional regulation of protein coding and RNA genes. For this reason, engineering of these parts is a promising way to control the cyanobacterial gene expression. The number of promoters (constitutive and inducible) well characterized in cyanobacteria has been rather small thus far. They have been summarized in several reviews published in the last years[18,19,32,33,42].

Cyanobacterial native promoters characterized do not represent yet the dynamic range required, and they can interact with other components of the regulatory network of the host, causing undesired results. Foreign promoters have been used to avoid cross talk, but they tend to function less efficiently in cyanobacteria, because of the differences in the transcriptional machinery. Synthetic promoters, whose DNA have been altered to modulate the transcriptional capacity, have also been developed as an alternative[48]. More studies and systematic characterization of the behavior of foreign and native promoters in cyanobacteria are necessary to improve the controlled expression of genes of interest[27,33].

A substantial number of cyanobacterial promoters reported respond to light intensity, the phase of the day and CO_2 or metal concentrations[31]. Native promoters used to engineer cyanobacteria usually come from photosynthesis-related essential genes such as photosystem I (P*psaA*, P*psaD*)[49,50], photosystem II (P*psbA1*, P*psbA2*)[51,52], photosynthetic antenna protein phycocyanin (P*cpc*)[53], and CO_2 fixation (P*rbcL*, P*cmp*, P*sbt*)[54,55]. Others come from genes coding the RNAseP (P*rnpB*)[56], plastocyanin (P*petE*)[57] and metal-ions responsive genes like P*nrsB*[58]. The foreign promoters used in cyanobacteria are generally derived from *E. coli* operons. Some examples of them are Plac (lactose operon), Ptrc/Ptac (from the combination of promoters from the tryptophan and lactose operons)[11], Ptet (tetracyclin resistance)[56], PBAD (*araBAD* promoter-arabinose operon)[59] and PrhaBAD (L-rhamnose operon)[60].

Some native or orthogonal promoters can drive strong expression of heterologous genes and they appear constitutive under certain conditions, or due to the absence of their regulators[42]. Some promoters are responsive to environmental signals (chemical or physical factors such as metabolites, macronutrients, metals, among others). These inducible promoters allow adjustable control over gene expression in terms of expression levels and timing of the expression of a particular gene[60].

13

Inducible promoters are especially valuable when the introduced gene(s) or pathways put a heavy metabolic burden on the host cells or when the products are toxic for them. Nevertheless, some limiting factors like the toxicity of inducers, leakiness in the absence of the inducer and inducer photolability have hindered the development of cyanobacterial inducible systems and a wide range of expression capacities is still needed for optimal gene expression[18,27]. Recent efforts to enrich the cyanobacterial promoter toolboxes have been published by various authors[24,25,61,62,63].

Transcription factors and **Transcription regulators** are also important elements in the transcriptional regulation. They are able to enhance or inhibit the expression of a gene by binding to specific DNA motifs in its promoter sequences. Transcription factors are composed by a DNA binding domain (DBD), an oligomerization domain (to interact with other transcription factors and regulators) and a transcription regulation domain (to control gene expression). Transcription regulators interact with transcription factors or chromatin, modifying chromatin structure to make the genes accessible, or facilitating the recruitment of the transcriptional machinery, allowing the transcription of the genes[64,65].

Transcriptional terminators are additional genetic elements used to control gene expression at the transcriptional level. Terminators isolate transcriptional units and prevent the transcription of genes adjacent to an inserted gene. They can reduce the effect of genetic context on gene expression, improving modularity and predictability[30]. There are two types of terminators: Rho-dependent terminators, which requires a homohexameric Rho protein to uncoil the RNA-DNA hybrid preventing the elongation of the nascent RNA strands, and the Rho-independent (intrinsic) terminators, which are composed of a GC-rich hairpin-loop motif within the nascent RNA transcript followed by an U-rich tail, that causes dissociation of the transcription elongation complex.

In contrast with the studies on *E. coli*, few studies have been done to characterize terminators in cyanobacteria, especially in the model organism *Synechocystis* sp. PCC 6803 which only possess Rho-independent terminators (genes coding for homologues of Rho proteins have not been found in cyanobacterial genomes)[34,61,66]. Some of the terminator sequences characterized in *E. coli* have been successfully used in cyanobacteria[67]. In addition to them, the terminator of RuBisCO has been optimized for expression in cyanobacteria[33].

Translational regulation

Ribosome binding sites (RBS) are crucial control elements for translation initiation that determine the expression level of genes. In prokaryotes, the translation process initiates by the binding of a ribosome to the mRNA at the ribosome binding site, which usually contains a core sequence (Shine-Dalgarno, 5′-GGAGG-3′). The efficiency of an RBS to recruit ribosomes for translation depends on multiple factors. For instance, the RBS position, sequence and its potential to bind the anti-Shine-Dalgarno sequence in the 16S rRNA; the surrounding nucleotides, that might result in secondary structures regulating the binding to the ribosome; and the spacing between the RBS and the translation start codon[32,33].

In silico RBS calculation tools have been used to predict the efficiency for specific ribosome-mRNA pairs, and the effect of altering an RBS sequence. This is done based on the statistical thermodynamic modeling of the mRNA-ribosome interaction, along with criteria such as the sequence itself, the spacer length, the level of homology to the Shine-Dalgarno sequence and the predicted mRNA

folding[34]. This information has been applied to develop RBS libraries for cyanobacteria. Native, foreign and synthetic RBSs have been characterized in cyanobacteria by several authors[20,24,32,61,68,69].

The effect of genetic context on gene expression is an important factor to take into account when using biological parts like promoters or RBS. Context specific characteristics of an expression construct can alter the performance of a promoter-RBS combination, ultimately reducing the modularity of biological parts. Secondary structures that can interfere with transcription and/or translation, can be formed between a heterologous gene, causing variability when using identical promoter-RBS cassette to express different genes. In a similar way, the surrounding sequences to the gene of interest may have uncharacterized regulatory roles and the location of insertion of the cassette may influence its copy number, affecting the gene expression. For this reason, it is often difficult to predict the likely expression of a construct during the design process[27].

In addition to the commonly used transcriptional and translational regulatory parts, another way to control gene and/or protein expression is by altering the RNA conformation or RNA base pairing with DNA or RNA sequences. These non-coding RNAs called RNA regulators, such as small RNAs (sRNAs), antisense RNA (asRNA), Riboregulators and Riboswitches can be used as translational regulatory tools to engineering cyanobacteria[34].

Small RNAs are non-coding molecules (300-500 nt) able to regulate the target mRNAs through base pairing (perfect or imperfect)[70]. sRNAs have been functionally characterized in cyanobacteria, showing great potential to develop genetic regulatory tools[19,71,72]. They have the advantage that these regulatory tools hardly impose metabolic burden on the host cells[73]. **Antisense RNAs** play an important regulatory role during translation and have been explored as an alternative for gene silencing and downregulation in cyanobacteria. This, since the traditional strategy to delete essential genes or pathways is usually lethal to host cells. Instead, these genes or pathways could be knocked down using antisense RNAs without causing cellular death[74,75,76].

Another type of regulatory RNA tools in cyanobacteria are the **Riboswitches.** They are versatile, easy to design (regulation mechanism is mediated by base pairing against the target mRNA) and allow fine tune control of the expression of a target gene through the manipulation of a secondary structure within a mRNA. Riboswitches are non-coding RNA molecules, commonly located in the 5´untranslated region (5´UTR) with the ability to bind with small molecules (ligands). They possess a sensing (aptamer) and a regulating domain. When the aptamer binds to a specific ligand, a conformational change occurs, making RBS inaccessible and mediating a transcription or translation switch (on/off), affecting the expression of the genes that are adjacent.

Some advantages of the riboswitches are (i) a tight regulation to induce expression in the presence of the ligand and to repress translation in the absence of it, minimizing leakiness; (ii) the levels of protein expression of genes controlled by riboswitches can be modulated by varying the concentration of the ligand; and (iii) they show a high degree of modularity, have fast response times, and can be used to regulate multiple genes at the same time[27,34,77,78].

Degradation tags are post translational mechanisms that can regulate protein stability, controlling the protein abundance. They can be useful to alter the cellular mechanism in cyanobacteria, as described by Laundry *et al.*[79] In bacteria, ssrA degradation tags are used as quality control mechanisms. When the ribosomes stall during translation, the tmRNA or SsrA ribosome rescue system facilitates the addition of the sequence AANDENYALAA (ssrA tag) to the C-terminus of the

polypeptide. This sequence targets the protein for degradation, thus preventing the accumulation of aberrant, unfinished proteins[80].

Some *E. coli* and cyanobacterial degradation tags have been tested in *Synechocystis*. Seventy-one cyanobacterial genomes were screened to find *ssrA* tag sequences to determine a consensus sequence. This was used to design degradation tag sequences covering a broad range or strengths, creating a library able to regulate protein degradation[79]. This tool has been used to fine-tuning gene expression such that it permits the expression of a toxic protein by controlling the rate of protein degradation to minimize toxicity[81,82].

Gene editing tools

In addition to the need for standardized parts derived from cyanobacteria, there is a strong need for efficient genome editing tools to modify these organisms. This, because genetic modifications require a very tight controlled expression of target genes or pathways, sometimes difficult to achieve using commonly used tools. Recent alternatives have risen as an alternative to improve this process.

A very promising tool that has been implemented in cyanobacteria is the **CRISPR-Cas system** (clustered regularly interspaced short palindromic repeats – CRISPR-associated proteins). This prokaryotic immune system provides adaptive resistance to foreign plasmids or phages[83]. CRISPR-Cas systems have been classified in three types, based on the presence of three different signature *cas* genes: type I system (Cas3), type II system (Cas9), and type III system (Cas10)[84]. The type II CRISPR-Cas9 system from *Streptococcus pyrogenes* has been proved the most accurate and robust to perform genome edition in eukaryotic and prokaryotic cells[85]. In cyanobacteria, RNA guided genome editing has been achieved using the type II CRISPR-Cas from *Streptococcus pyrogenes*[86,87], but due to the apparent toxicity that Cas9 nuclease may have in cyanobacteria, the RNA directed dsDNA nuclease Cpf1 from *Francisella novicida*, nontoxic for cyanobacteria, has been used as an alternative[88].

In contrast with other genome-editing techniques such as zing finger nucleases and TALENs (transcription activator like effector nucleases) which depend on DNA-protein interaction, CRISPR-Cas work with RNA–DNA complementarity rules. This, avoid the need of protein engineering, simplifies the protocol, improves the efficiency and fidelity of genome editing, and reduces the time needed to produce mutants or mutant libraries; overcoming the limitations associated with multiple chromosomes in cyanobacteria. This is possible because with CRISPR/Cas9 all the copies of the chromosome are edited simultaneously[87]. This method enables chromosomal integration, allowing the edition of multiple sites in a short period of time[34]. This system has been successfully used to edit the genome of *Synechococcus elongatus* PCC 7942[87], *Synechococcus elongatus* UTEX 2973[86], *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120[88].

The **CRISPRi** (clustered regularly interspaced short palindromic repeats interference - CRISPR interference) can be used to modulate repression of gene expression in several cyanobacterial species including *Synechococcus* sp. PCC 7002[29], *Synechococcus elongatus* UTEX 2973[89], *Synechocystis* sp. PCC 6803[90,91] and *Anabaena* sp. PCC 7120[92]. In this system, the catalytical activity of the Cas9 protein has been eliminated through active site point mutations, resulting in dCas9 (dead Cas9 protein without nuclease activity)[93]. This dCas9 can only bind to the DNA instead of causing a double stranded break, blocking the RNA polymerase´s progress by stopping the initiation or the elongation of transcription. CRISPRi allows multiple gene repression (heterologous and native genes), using the dCas9 in combination with a sgRNA (single guide RNA) expressed under the control of an inducible

promoter. This results in a tight repression that can be tunable by varying the concentration of inducer[94].

Libraries in cyanobacteria

Libraries are practical genetic tools that facilitate the development of cyanobacteria as production organisms. They have been utilized to have a better understanding over complex genetic phenotypes and tolerance mechanisms. They have also been used to characterize biological parts in cyanobacteria. Some types of libraries that have been created in cyanobacteria are genomic, expression and knockout libraries[34].

Genomic libraries are a collection of fragments of genomic DNA. They have been extensively used in many genetic studies, or for the isolation and cloning of genes of interest. Screening of genomic libraries are useful to identify genes or other DNA sequences performing a specific function (for example promoters, RBS´s and other genetic elements), genomic mapping or even to complete genome sequencing[20,24,34,95,96].

Expression libraries are a collection of genomic fragments or a representative sample of cDNA, constructed in a way that they can be transcribed and translated by the host organism. They can be used to screen for a property of interest or to analyze expression patterns. Expression libraries using inducible promoters and optimized RBSs can be effective to increase the generation of products and the characterization of tolerance mechanisms[34,97,98].

Knockout libraries are made of gene disruption mutants, and they provide a unique opportunity for functional analysis of an organism genome. These libraries can be generated in cyanobacteria by multiple methods including homologous recombination of selectable markers, markerless knockout, or transposon mutagenesis[18,99,100,101].

A library can be constructed to be integrated in a replicative plasmid or into the host´s chromosome(s). Even though it is easier to construct a library using a replicative plasmid, the copy number in a plasmid might vary more than the copy number of chromosomes, compromising the quantitatively characterization of a genetic element. Besides that, a genomic library integrated in the chromosome will be more stable, reducing the use of selection markers such as antibiotics to keep it integrated[34]. On the other hand, the polyploidy (multiple chromosome copies per cell) of some cyanobacteria creates a limitation in terms of creating integrative libraries and tracking them on a chromosome-by-chromosome basis. If full integration is not achieved, the performance of the system is not optimal, and the cultures can become genetically unstable[45,102]. Even though the design and construction of libraries in cyanobacteria has become easier, the lack of efficient techniques to achieve full segregation in polyploidy strains makes large-scale modifications still challenging.

Detection systems in cyanobacteria

One of the challenges when creating cell factories is the design of a fast and easy strategy to evaluate the performance and production of engineered strains. An intrinsic property of bacteria can be utilized for that purpose. Biological systems can respond to various environmental stimuli. That response to environmental signals can be connected to protein production or can be coupled with a reporter system to construct a sensor. These biological systems are promising candidates to be used as detection systems with a wide range of detection, which is relevant for industrial applications.

Initial work on the developing of biosensors has been done mostly in model organisms such as *E. coli* and *S. cerevisiae*. However, cyanobacteria have been explored as a new chassis with metabolic plasticity, able to perform oxygenic photosynthesis and able to easily adapt and modify fluxes in response to stimuli from the environment[103]. Cis or trans-acting regulatory elements can be arranged to construct biological circuits and then integrated in cyanobacteria to function as a biosensor, that senses an input and produces an output that can be observable and quantifiable[30]. These biosensors can be applied to several fields such as agriculture, metabolic engineering, diagnosis and pollutant survey[104,105].

In cyanobacteria, several **Biosensors** have been developed in response to physical (light, temperature, CO_2 and O_2), chemical or metabolic signals, and they have been reviewed in detail by Sengupta *et al*.[31], and Immethun and Moon[30]. These sensor systems can be used to facilitate the screening and quantification of production strains and also to construct more complex genetic circuits in cyanobacteria. Other parts and endogenous response pathways reacting to stress conditions could be also explored to develop other biosensors.

Final remarks

Cyanobacteria have attained relevance for their potential as promising cell factories for the production of biofuels and high value chemicals[17]. Synthetic biology has played an essential role in the advances to engineer these cells and re-route the carbon flux to enhance the production of desired compounds. This has been achieved through the use of cis and trans-acting elements, genome editing techniques and predictive tools[31].

Efforts to continue expanding the toolbox of characterized biological parts in cyanobacteria are fundamental to design more elaborated and robust circuits, to achieve a more predictable expression output. With this work we aim to contribute by developing tools and techniques to enrich the toolbox of cyanobacteria, especially in the model organism *Synechocystis*.

Construction of fully segregated genomic libraries in polyploid organisms such as *Synechocystis* sp. PCC 6803

Patricia Caicedo Burbano, Tycho Smit, Hugo Pineda Hernández, Wei Du, Filipe Branco dos Santos

Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Netherlands

Abstract: Several microbes are polyploid, meaning they contain several copies of their chromosome. Cyanobacteria, while holding great potential as photosynthetic cell factories of various products, are found amongst them. In these clades the diversity of genetic elements that serve within the basic molecular toolbox is often limiting. To assist mining for the latter, we present here a method for the generation of fully segregated genomic libraries, specifically designed for polyploids. We provide proof-of-principle for this method by generating a fully segregated genomic promoter library in the cyanobacterium *Synechocystis* sp. PCC 6803. This new tool was first analyzed through fluorescence activated cell sorting (FACS) and then a fraction was further characterized regarding promoter sequence. The location of libraries on the chromosome provides a better reflection of the behavior of its elements. Our work presents the first method for constructing fully segregated genomic libraries in polyploids, which may facilitate their usage in synthetic biology applications.

Keywords: polyploid microorganism, genomic library, fully segregated, counter-selection, cyanobacteria, promoters

Introduction

A genomic library is a collection of DNA fragments, ideally representing the entire DNA content of the genome from which the library was derived[106]. Genomic libraries are useful to screen for target DNA fragments contributing to desired complex phenotypes (e.g. chemical tolerance), which are normally difficult to be rationally engineered[107,108]. Furthermore, genetic elements, such as promoters, that are indispensable in genetic engineering, can also be identified. These include promoters not only with different strengths but also with different inducible properties[109]. Thus far, genomic libraries were mostly constructed either on a replicative plasmid, or integrated in the chromosome of monoploid microorganisms (i.e. organisms containing a single chromosome copy)[110]. While for polyploid microorganisms, large genomic libraries have only been developed to explore native DNA fragments, such as promoters through a reporter (e.g. fluorescence), on a replicative plasmid[109]. Because the copy number of the replicative plasmid may vary under the cultivation conditions tested[111], this might compromise the correlation between the activity of the reporter and the promoter strength. For polyploid microorganisms as well, genomic libraries directly located on the chromosomes would have a relatively stable copy number in comparison to the ones located on replicative plasmids. But clearly, integrating such a library on all the chromosomes of a polyploid microorganism is a major challenge, without going through the regular time-consuming segregation steps for a limited number of representatives[112]. How can one ensure that the genomic library integrated on the chromosomes of polyploid microorganisms is fully segregated? In other words, how can one assure that all existing copies of the chromosome contain the intended fragment?

We decided to tackle this challenge by constructing a fully segregated chromosomal genomic library in polyploid microorganisms using a counter-selection approach (Figure 2.1). This approach includes a positive selection (via antibiotic resistance genes) and a counter (negative) selection (via conditional expression of toxic genes). Specifically, both selection cassettes need firstly to be fully integrated into the target genomic locus of the targeted polyploid microorganism. This is achieved by the positive selection with increasing dosage of the antibiotic through possibly several rounds of segregation. During the positive selection process, the toxic gene for counter-selection is regulated as non-functional (e.g. non-induced or no substrate available). This polyploid microorganism, with fully integrated chromosomal selection cassette, is served as the background strain ready for transformation. Then, the genomic DNA fragments are prepared either through partial enzymatic digestion or mechanical forces to a target size range. Those fragments will later be inserted into a non-replicative vector, flanking with upstream and downstream homologous regions of the target genomic locus of the background strain, resulting in a plasmid library. After introducing the plasmid library into the background strain, the selection cassette will be replaced with a random DNA fragment through double crossover at the target genomic locus, by homologous recombination. With controlled functional expression of the toxic gene (e.g. adding an inducer or substrate) for counter-selection, only the cells with fully replaced chromosomes can survive. This would lead to fully segregated chromosomal genomic libraries created in polyploid microorganisms.
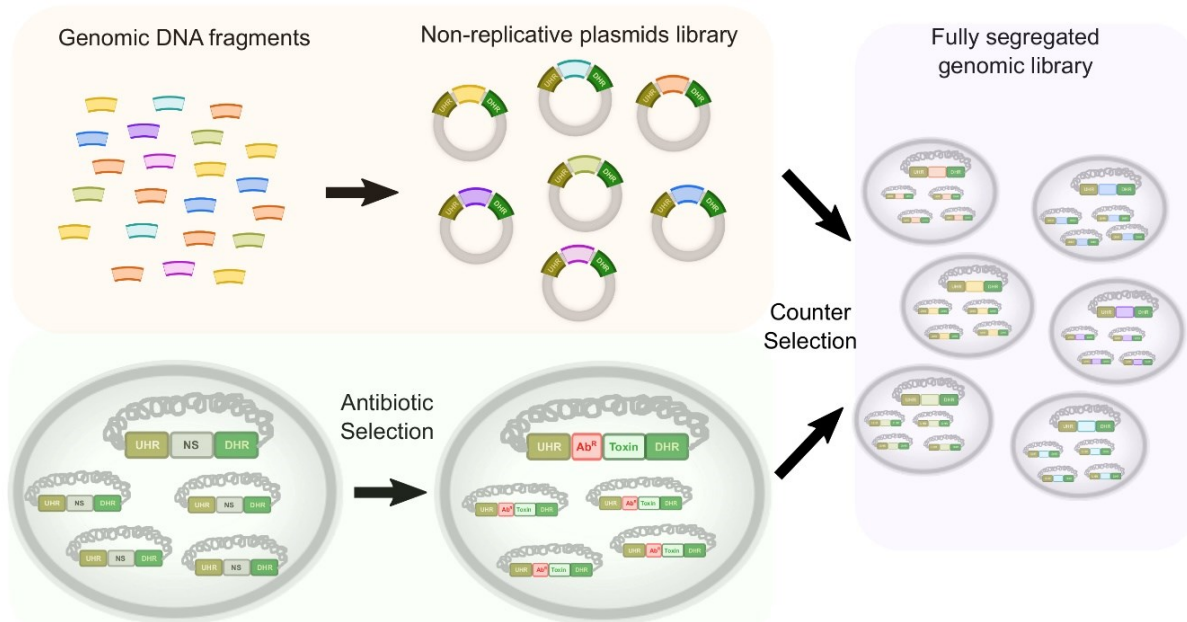
Figure 2.1 A schematic drawing of the construction of fully segregated genomic libraries in polyploid microorganisms. The full segregation of the genomic libraries is enabled through a counter-selection strategy, where a positive selection (via antibiotic resistance genes) and a counter (negative) selection (via conditional toxic genes) are implemented. Specifically, target genomic locus of the microorganism needs firstly to be fully integrated with both positive and counter selection genes via antibiotic selection and segregation (light green background). Then the prepared genomic DNA fragments are ligated to the non-replicative vector flanking with upstream and downstream homologous regions of the target genomic locus, resulting in a plasmid library (light yellow background). After introducing the plasmid library into the microorganism, double cross-over occurred at the target genomic locus will replace the selection genes with the DNA fragments. Cells with only fully replaced chromosomes can survive under counter-selection. Hence, the fully segregated genomic libraries are created in polyploid microorganisms (light purple background). NS, neutral site; UHR, upstream homologous region; DHR, downstream homologous region; Ab^R, antibiotic resistance gene; Toxin, conditional toxic gene. Five chromosomes inside the cells are depicted only to mimic the polyploidy.

To test this approach, polyploid cyanobacteria – promising photosynthetic microbial hosts, that can be employed to directly convert atmospheric CO_2 to biochemical compounds[6] – were chosen. As one of the model cyanobacterial species, *Synechocystis* sp. PCC 6803 (hereafter, *Synechocystis*) that has multiple chromosome copies, was selected as a case study. The chromosome copy number in *Synechocystis* is variable and can be affected by growth phase and by physical and chemical factors[44,45,113]. Nevertheless, despite the variation, the copy number of chromosomes in polyploid cyanobacteria is reported to be much smaller and less variable than the copy number of exogenous plasmids. This has been observed in the expression levels obtained through plasmid expression in comparison with chromosomal expression in *Synechocystis* sp. PCC 6803[56,114,115], *Synechococcus* sp. PCC 7002[116] and *Synechococcus elongatus* PCC 7942[117].
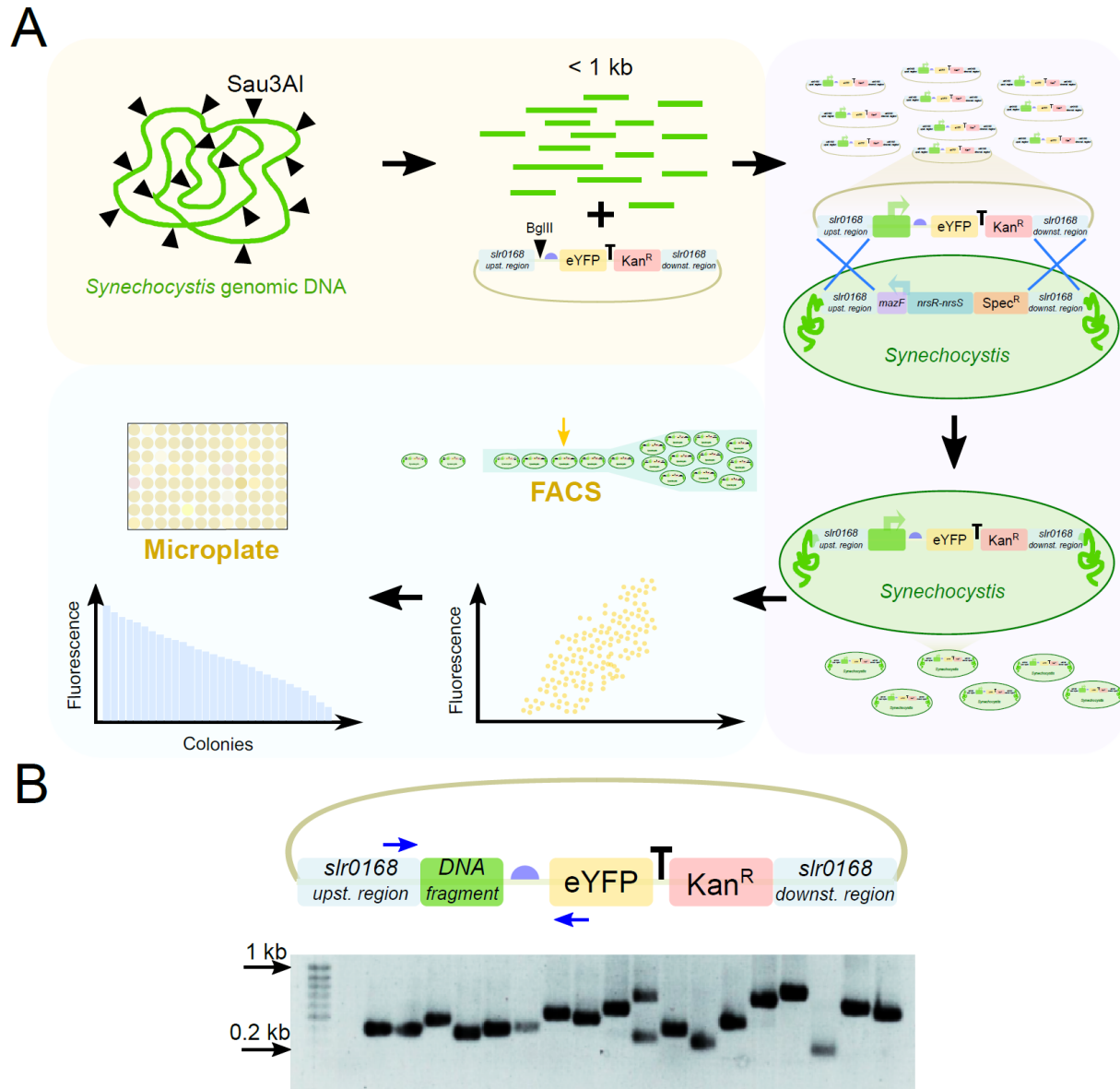
Figure 2.2 Construction of a fully segregated genomic promoter library in *Synechocystis*. (A) The whole process can be divided as three main modules (as indicated with different colors background): (1) inserting the enzymatic digested genomic DNA fragments to the integrative vector targeting the *slr0168* genomic locus of *Synechocystis*; (2) transforming *Synechocystis* with the plasmids containing DNA fragments to obtain the fully segregated promoter library on the *Synechocystis* chromosome; (3) characterizing the promoter library in *Synechocystis* via the fluorescence intensity. (B) PCR verification of a few representative *E. coli* colonies harboring the genomic library. The gel picture shows that the size of genomic DNA fragments generated to create the promoter library are ranging roughly from 200 bp to 1000 bp. Blue arrows indicate the position of the primers.

Despite the efforts to identify native promoters in cyanobacteria, the number of well characterized constitutive and inducible native promoters is still considered to be somewhat small[18,19,27,33]. Because a large number of the promoters that have been applied in the cyanobacterial studies come directly

from other model organisms such as *Escherichia coli*[18], we deemed that a fully segregated chromosomal genomic library would be useful in order to mine the native cyanobacterial promoters.

Overall, in this work we successfully constructed a fully segregated genomic library on the chromosomes of *Synechocystis* (Figure 2.2 A). As a conceptual framework, we screened and characterized in total 72 native promoters with their sequences analyzed. Our work marks the first report of construction of a fully segregated chromosomal genomic library in polyploid microorganisms, without going through the time-consuming segregation step of only a small number of representatives[112]. This library could in principle be exploited either for fundamental research (e.g. detailing complex phenotypes) or synthetic biology applications (e.g. mining genetic elements).

Results and Discussion

To initialize, *Synechocystis* genomic DNA was isolated and partially digested by Sau3AI to a size of approximately 200 bp to 1 kb (Figure 2.2 B). Those DNA fragments were inserted on a suicide plasmid (at the BglII restriction site) that integrates at the *slr0168* genomic locus (neutral site) of *Synechocystis*. Between the homologous regions, and downstream of the DNA fragments inserted, a ribosome binding site, a fluorescence reporter (eYFP) and a terminator were attached (Figure 2.2 A). This enables the screening and quantification of the promoter strength using the fluorescence intensity as a proxy. Additionally, a kanamycin resistance gene is included in this vector for propagation purpose. To introduce this plasmid library containing random DNA fragments into *Synechocystis*, and ensure their complete segregation, we adopted a counter-selection strategy that has been developed for markerless genetic modification[118]. To implement this strategy, we first build a *Synechocystis* strain, that has a fully segregated *mazF* (toxic) cassette, expressed under the strict control of a nickel-inducible promoter, at the *slr0168* genomic locus. This was done by selection using spectinomycin. Cultures of this *Synechocystis* strain were then incubated with the plasmid library in a shake flask without antibiotic for 24 hours. Then, kanamycin was added to ensure the integration of the promoter library targeting the *slr0168* genomic locus (replacing the *mazF* cassette) and pushing the cells towards the direction of complete segregation. After a few days, nickel was added in the shake flask to select the cells with fully segregated inserts. In order to have a representative promoter library, it is critical to determine the timing of adding nickel. This should not be too soon to avoid killing the cells that were not yet fully segregated. But should also not be too late, which would favor fast growers to take over the population, reducing library representativeness. We tested the effect of adding nickel at different days after adding kanamycin, and more varied colonies with fully segregated chromosomes (verified by PCR in Fig S2.1) were obtained when nickel was added after 4 days (data not shown here).

The strength of each randomly inserted DNA fragment with promoter activity, was then checked by fluorescent activated cell sorting (FACS) (Fig 2.3 A). The *Synechocystis* strain without any DNA fragment insertion serves as the negative control (important to set the lower threshold of fluorescence). While the strain carrying a strong constitutive promoter (P*cpcBA*) was used as the positive control[119]. Our results indicated that a small proportion (~12.2%) of the total cells analyzed (50.000) sparked fluorescence at different intensities. For the rest of the cells the fluorescence was below the threshold, indicating the DNA fragments inserted did not drive eYFP expression. For the cells displaying significant fluorescence, it is of course interesting to characterize the underlying sequence responsible for its intensity. Therefore, cells that displayed fluorescence above the set threshold were collected by FACS and spread on BG-11 plates to isolate single colonies. To firstly have

an overview of the colony heterogeneity, we performed a colony PCR to specifically amplify the DNA fragment inserted. Our results (Fig S2.2) showed that for a sample of 18 colonies tested, the size of the DNA fragment varies, but always falls into the expected range (200 bp to 1 kb). This again corroborates the idea that our approach is feasible and can lead to the successful construction of a genomic promoter library directly on the chromosomes.
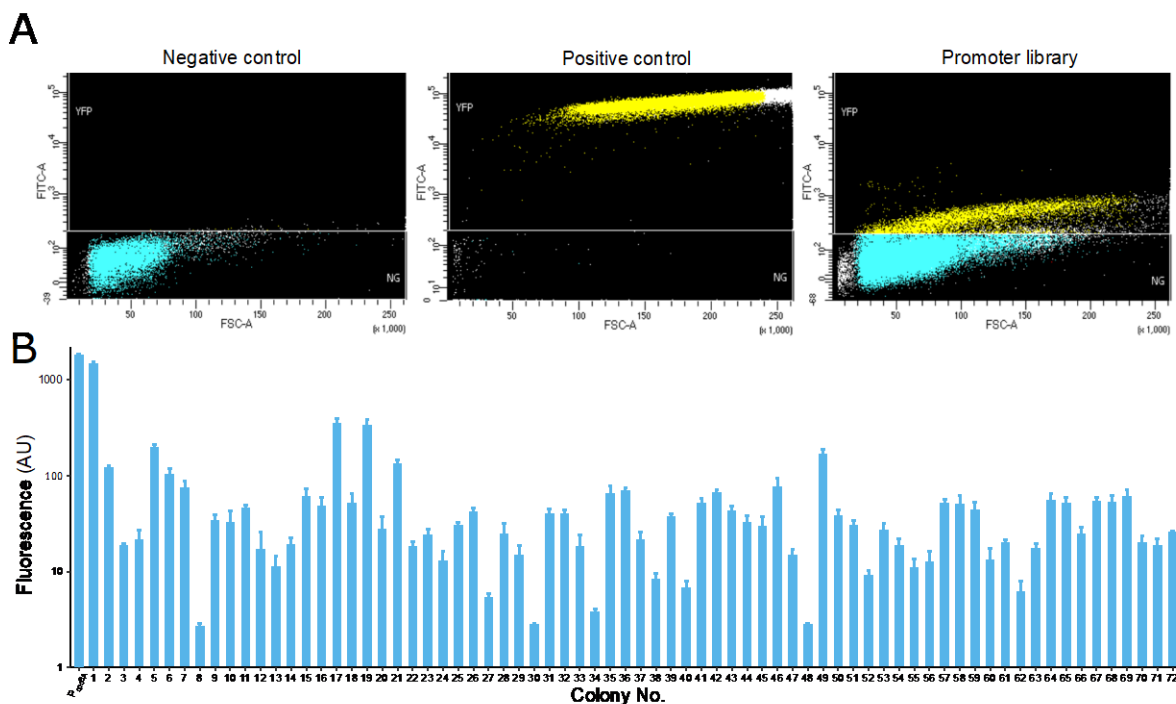

Figure 2.3 Characterization of the *Synechocystis* genomic promoter library via fluorescence intensity measurement either by FACS (A) or microplate (B). For FACS, the promoterless strain that was unable to express eYFP was used as negative control and also to determine the fluorescence threshold (indicated by the white line). The strain expressing eYFP driven by the strong constitutive P*cpcBA* promoter was used as positive control. Individual colonies that display fluorescence below the threshold were assigned the color blue, while the ones above the threshold were shown in yellow color. The *Synechocystis* cells transformed with the genomic promoter library showed heterogeneity in the fluorescence intensity, with low or mild expression in most of the population. In the microplate, the fluorescence intensity represents the transcriptional strength of the 72 promoters sequenced, in the order of Table S2.1, compared with the strength of P*cpcBA*. Fluorescence was normalized by OD_{730} of the cultures. The error bars indicate the standard deviation of three replicates.

As a proof-of-principle validation of our approach in terms of mining native promoters, we picked in total 72 colonies from the plate for further characterization. To measure the fluorescence of each colony, we applied a 96-well plate method[120]. The results showed a wide distribution of the fluorescence from all the colonies, which reflects the diverse promoter strength of the respective DNA fragments (Figure 2.3 B). We next sequenced the inserted DNA fragment acting as promoters in those 72 colonies (Table S2.1). The results indicated that all the fragments with promoter activity had unique DNA sequences, though some colonies have overlapping sequences that cover the same region of the chromosomal locus (specifically, colonies 11, 13, 22, 35 and 39 in Table S2.1). Some sequences showed the typical promoter location (5' UTR region), other sequences were part of a gene, and others overlap two neighbor genes. The ones located as part of a gene, suggest that some

regions of the annotated genes could also be regulatory elements of other flanking coding sequences, or the promoters driving the expression of sRNAs[75,121]. Some of the promoters found, belong to regions of the genes coding for proteins involved in replication and transcription such as ligase, DNA gyrase and transcriptional regulators. Others were part of genes coding for important metabolic enzymes such as ATP synthase, NADH dehydrogenase, phosphorylases, transferases, synthases and dehydrogenases. The promoters that were related with photosynthesis were part of genes coding for phycocyanin, plastocyanin and thioredoxin. Furthermore, some sequences found are part of the genes encoding for hypothetical proteins. This can be an indication of the regulatory role of the regions that are not coding for known proteins, or that are still not well characterized. Overall, these findings validate the potential of constructing a promoter library on the chromosome to identify native promoters that can modulate different levels of target gene expression for genetic engineering purposes.

Genomic libraries serve as an important tool that can be utilized to map important DNA fragments/genes for both fundamental and applied research[122]. In terms of exploring native promoters in polyploid microorganisms, a genomic library integrated on the chromosome is more preferred (although challenging) than on a replicative plasmid. This is because the copy number of the chromosome is relatively stable, while the replicative plasmid copy number may vary more. This may be due mostly to the independent replication system of the replicative plasmid such that its replication is loosely controlled by the cells. Furthermore, it is also difficult for cells to equally distribute the replicative plasmids to the daughter cells via the plasmid partition mechanism during cell division[123,124]. Therefore, when cells have different copy numbers of the replicative plasmid, the reporter gene dosage is expected to be different. Under this situation, it would be problematic to directly estimate the promoter strength based on the activity of the reporter (e.g. fluorescence). In this work, we tackled this challenge using a counter-selection approach and successfully constructed a fully segregated genomic library on the chromosomes of the cyanobacterium *Synechocystis* as a case study. As a conceptual validation, we screened in total 72 native promoters that showed a wide distribution of promoter strength. Because cyanobacteria hold great promise to be developed as photosynthetic cell factories, the native promoters mined here through a genomic library, may greatly facilitate various synthetic biology implementations. This approach to construct a fully segregated genomic library is also applicable to other polyploid microorganisms, where desired DNA fragments contributing to target phenotypes can be reliably mapped for both fundamental and applied research.

Materials and Methods

General cultivation conditions. *E. coli* strain DH5α was used as the host for plasmid molecular cloning. It was grown at 37 °C in Lysogeny Broth medium (LB) in an incubator with a shaking speed of 200 rpm or on LB plates containing 1.5% (w/v) agar. The concentration of antibiotics used were 50 µg mL^{-1} for kanamycin and 50 µg mL^{-1} for spectinomycin. *Synechocystis* sp. PCC 6803 (glucose tolerant, obtained from D. Bhaya, Stanford University, USA) was cultivated at 30 °C in liquid BG-11 medium supplemented with 10 mM TES-KOH (pH = 8), in a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific) under constant moderate red light illumination ($\sim$ 30 µmol photons $\cdot$ m$^{-2} \cdot$ s^{-1}), or on BG-11 agar plates supplemented with 10 mM TES-KOH and 0.3% (w/v) sodium thiosulfate, in an incubator with white light ($\sim$30 µmol photons $\cdot$ m$^{-2} \cdot$ s^{-1}) and 1% CO$_2$. For *Synechocystis* mutant construction, kanamycin, spectinomycin or nickel sulphate were added to the medium with a final concentration of 50 µg mL^{-1}, 20 µg mL^{-1} and 15 µM, respectively. Biomass concentration in the

cultures was measured by optical density at 730 nm (OD_{730}) in a spectrophotometer (Lightwave II, Biochrom).

Plasmid and strain construction. All plasmids, strains, and primers are listed in the Table S2.2. To construct the required plasmids, both homologous regions of *slr0168* were amplified from the genomic DNA of *Synechocystis*, using Herculase II polymerase (Agilent) and primers designed to introduce restriction sites for KpnI in the 5′ of the upstream region, HindIII in the 3′ of the downstream region, and XbaI in between the two regions. The fragments were fused together and completely amplified using *Pfu* DNA Polymerase (Thermo Scientific). After gel extraction and purification (Thermo Scientific), the fused fragment was inserted in the vector (prepared by digestion of the plasmid pFL-XN with KpnI and HindIII). This resulted in the new plasmid pFLXN, serving as the background plasmid for the following constructs. To construct the plasmid pFLXN-MAZF, kanamycin resistant fragment from the selection cassette of the pWD42 plasmid was firstly replaced with the spectinomycin resistant fragment, resulting in the plasmid pWD007. Then, the new selection cassette was cut from pWD007 by double digestion with SpeI and AvrII and inserted into the pFLXN plasmid digested by XbaI.

To generate the CPC-YFP cassette, the region containing the P*cpcBA* promoter (including the RBS), eYFP, terminator BB0014 and kanamycin resistant cassette (Kan^R) was firstly PCR amplified from pHKH-cpcBA-YFP. Then, this fragment was further prepared by XbaI and NheI digestion and inserted into the pFLXN plasmid digested by XbaI, resulting in the plasmid pFLXN-CPC-YFP. To generate the RBS-YFP cassette, only the region containing the ribosome binding site from the P*cpcBA* promoter (excluding the promoter region), eYFP, terminator BB0014 and Kan^R was PCR amplified from pHKH-cpcBA-YFP and then inserted into the pFLXN plasmid following the same above approach. That resulted in the plasmid pFLXN-RBS-YFP. For the construction of the pFLXN-LIB-YFP plasmid, a BglII site was additionally added before the RBS during the PCR amplification of the RBS-YFP cassette, to allow later introduction of the genomic DNA fragments of the library. All the fragments amplified in this study were confirmed by Sanger sequencing at Macrogen Europe (The Netherlands).

Synechocystis mutants Syn-MAZF, Syn-RBS-YFP, and Syn-CPC-YFP were created by natural transformation of the corresponding plasmid to *Synechocystis* wild type as described previously[125]. Full segregation of each mutant was verified by PCR using the genomic DNA as the template for 35 cycles. To obtain a fully segregated genomic library integrated in the *Synechocystis* chromosomes, we adopted a liquid transformation approach to introduce the promoter library to Syn-MAZF. Firstly, fresh Syn-MAZF cells were collected from 20 mL liquid culture ($OD_{730} \approx 1$). After being washed twice with fresh BG11 medium through centrifugation (3900 rpm, 10 min), cells were further concentrated to a total volume of 200 μL. The pFLXN-LIB-YFP plasmids were mixed with these cells to reach a final plasmid concentration of 30 μg mL^{-1}, and then incubated at 30 °C with moderate light intensity (white light, 30 μmol photons $\cdot$ m^{-2} $\cdot$ s^{-1}) for 5 hours. Next, the mixture was inoculated in 20 mL of BG11 (without antibiotic) and incubated at 30 °C in a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific) under constant moderate red-light illumination (~30 μmol photons $\cdot$ m^{-2} $\cdot$ s^{-1}). After further incubation for about 24 hours, kanamycin was added to the cultures at a final concentration of 50 μg mL^{-1}. After another 4 days of incubation, nickel sulphate was added to the cultures at a final concentration of 15 μM. When the cultures reached $OD_{730} \approx 1 - 1.5$, a sample was collected to perform FACS analysis and the rest was concentrated to make glycerol stocks.

Genomic DNA promoter library construction in *E. coli*. Genomic DNA of *Synechocystis* was extracted using pelleted cells from 1 mL culture ($OD_{730} \approx 1$). The pellet was suspended in 200 µL TE buffer and mixed with 200 µL of phenol/chloroform/isoamylalcohol (25:24:1). After the addition of 100 mg of glass beads (0.1 mm, diameter), the suspension was vortexed for 5 min and then centrifugated at 14.000 rpm for 5 min. The supernatant was carefully collected and mixed with 200 µL of chlorophorm/isoamylalcohol (24:1). After centrifugation at 12.000 rpm for 5 min, the supernatant was transferred and precipitated with 1/5 volume of 5 M NaCl and 2 volumes of absolute ethanol, keeping it at -20 °C for at least 30 min. Then, DNA was pelleted after centrifugation at 12.000 rpm for 10 min, washed with 500 µL 70% (V/V) ethanol, and dried at room temperature. The DNA pellet was dissolved in 20 µL MQ water. RNA in the genomic DNA sample was removed by incubation with RNAse (final concentration 100 µg mL^{-1}) at 37 ˚C for 30 min. The genomic DNA was then precipitate with cold ethanol, dried and re-dissolved in MQ water.

Posteriorly, 10 µg of the genomic DNA was partially digested with 5 µL Sau3AI (New England) at 37 °C for 5 hours. DNA fragments with a size ranging from 200 bp to 1000 bp were generated and purified using MSB spin PCRapace kit (Invitek molecular). At the meanwhile, 4 µg of the vector pFLXN-RBS-YFP was digested with 2.5 µL of BglII (Thermo Scientific) at 37 °C for 4 hours. After 1 hour of digestion, the phosphate groups at each end of the vector were removed via the addition of alkaline phosphatase to the mix for 3 hours at 37 °C, to prevent vector self-ligation. Gel purification was used to purify the vector using Gen JET gel extraction kit (Thermo Scientific). The genomic DNA fragments and the vector prepared were ligated using T4 ligase (Thermo Scientific) and the ligation was transformed into DH5 alpha competent *E. coli*. The transformation in *E. coli* was done in a large scale, plating the resulting colonies in square petri dishes, until reaching an approximate number of 200.000 colonies. All of them were collected in LB medium and incubated at 37 ˚C in an incubator with a shaking speed of 200 rpm. After that, a maxiprep was done to extract the plasmidic DNA that was later transformed into *Synechocystis*.

Fluorescence analysis. A sample of the *Synechocystis* strain Syn-LIB-YFP containing the genomic library was used for analysis using Fluorescence Activated Cell Sorting (FACS ACS ARIA III). To determine the eYFP fluorescence, the blue laser 488 nm (filter: 530/30, mirror: 502LP) was used, with an excitation wavelength of 494 nm and an emission wavelength of 519 nm. FACS analysis was performed in a total of 50.000 cells. A gate was set to define the eYFP expression, using the fluorescence of the strain Syn-CPC-YFP as positive control (high expression) and the strain Syn-RBS-YFP as negative control (no expression). Cell sorting of the clones expressing eYFP was carried for individual characterization. These cells were inoculated on plate, and 72 clones were selected and inoculated on 48 well plates with BG-11 medium supplemented with 10 mM TES and 50 mM NaHCO$_3$. The cells were re-inoculated to an OD_{730} of 0.25 to let the cells grow for three days ($OD_{730} \approx 1$). Then 200 µL of each culture was used to measure the fluorescence on plate reader (FLUOstar optima, BMG labtech) using a 470 nm filter for excitation and a 510 nm for emission. The measurement was done on a 96 well plate (Greiner-F bottom) with three replicates.

Supporting information
Table S2.1 List of sequences from the selected promoter library

Colony No.	Sequence	Annotation
01	GATCCCGACTTCGTTATAAAATAAACTTAACAAATCTATACCCACCTGTAGAGAAGAGTCCCTGAATATCAAAATGGTGGGATAAAAAGCTCAAAAAGGGAAAGTAGGCTGTGGTTCCCTAGGCAACAGTCTTCCCTACCCCACTGGAAACTAAAAAAACGAGAAAAGTTCGCACCGAACATCAATTGCATAATTTTAGCCCTAAAACATAAGCTGAACGAAACTGGTTGTCTTCCCTTCCCAATCCAGGACAATCTGAGAATCCCCTGCAACATTACTTAACAAAAAAGCAGGAATAAAATTAACAAGATGTAACAGACATAAGTCCCATCACCGTTGTATAAAGTTAACTGTGGGATTGCAAAAGCATTCAAGCCTAGGCGCTGAGCTGTTTGAGCATCCCGGTGGCCCTTGTCGCTGCCTCCGTGTTTCTCCCTGGATTTATTTAGGTAATATCTCTCATAAATCCCCGGGTAGTTAACGAAAGTTAATGGAGATCAGTAACAATAACTCTAGGGTCATTACTTTGGACTCCCTCAGTTTATCCGGGGGAATTGTGTTTAAGAAAATC	UTR of *cpcB*
02	GATCAGCCAAAGCCGACGGAGAACTGATACCATGGT	part of *sll1885*
03	GATCGACGCCAAGGACAGCTTCATTACCATAGACCAAGCAGAACATCGTTTCCGGGGAAGCAAGCAACAGGAACTGTATCGTTTTACCGAC	part of *slr1788*
04	GATCGCCAATTTCCAATAATCCAGGGCCGTTGCGATAGAGCGTTCCCTGGAGGTCTGGGGGTATGCTCCCTTCCACGTCTTCGACCCAATAATCCCACTCCTGCGGCTGGGACTGGTAACCTCTCAGCCAATCCTGCGGACTGTAGGAGCGCTGCGATGGGGAACTGGTTGGGGGGGGAAAGTGACCATGGTCAATGGGGGAAGTTTGTTACATTTCTTTACTTTTACCATTGTAACGTTCTCCGCTCCGGGGAACTGGGGGGTTAAAATCTATACGTAAAGAATCTGAAGTTTTTCCACTTTCGCCCAGGGAGGTAAATGATGGCGACCGAAGTTTAAACAAGCCCAGCAATAGCACAATCCGCAAACACGCCCCCCCGTTACCGGGGTCTTACTGCACAACGATGACTTCAACTCCATGGAAGCATGTGGGTGCAAACTTT	UTR of *ssr2723*
05	GATCGCCTGTGATAGTGTGTGGTTTGGGCCATATCACTGGTTGTTGTCGGCGCTGGGGTGATAATCTTATAGGTGGCGTATTTGACTGAGTCTCAGTAGTGATGGTGTGGTGTAGGTATAAGGGAGTTCTTAGACAA	part of *sll0169*
06	GATCGCCGACAATTTTGGCGAATTGTCCGAGTCTATCCATCTCGAACTGTACAATTA	part of *ssr1238*
07	GATCGCCGGAGATAGGGTGGTGACGGCCAAACCACTGGAATTGTCGGCGCTGGGGAAAACATCGTCTCGGGAACGTCTTGGACTGCGTCCCCGTCGTGAGGGAACGGTGGAGGTCAAAGGGC	part of *sll0169*
08	GATCGCCGCAAAAACCTCAAATTGGTGGAGGGTGGCATTTTTCATAGTGGATTAATTTTATAGACTACTTCTATAAATTACATAAAAAATTGGATATTTATCTATTTTTAGCTTTCCTTAATAATC	part of *cmpR*
09	GATCGCATTTTGCTTAAGGGGATAGTTATGTATCATTTCAATGGCTTTGACTAAATTTTCCCAGTCCTGACCCTGGGCTAAAATTGGTGTCTTTTCCTGCACATACTGGCAATAGGACATAAGAT	part of *slr6047*
10	GATCGCCTATCCTCCTAGGATTTGACCCATCTCACCAGTTGGGGACAGCCCCAAAGTGATACCCTTAAAGATGGCAAATTTGACCAAAACTCAACAGTAATCCTGTAGTCTAACTATAAACAAATTCCTAGACAATCAGCATGGTAGCCGTAAAAGAAGCAACTAACGTTGGCAAAATTACCCAGGTCATCGGGCCTGTAATTGACGCCCAGTTCCCCAGTGGTAAATTGCCCCGTATTTATAATGCCCTTAAAGTCCAAGGCAGAAACTCTGCTGGTAACGAAGTAGCTGTTACCTGTGAAGTGCAGCAGCTTCTCGGCGATAACCAAGTCCGAGCCGTAGCCATGAGTTCCACCGACGGTCTCGTCCGGGGCATGGACGTGGTAGACACCGGGGCCCCCATCAGCGTTCCCGTCGGCACCGGCACCCTGGGTCGTATTTTTAACGTTCTTGGTGAGCCTGTTGACAACAAAGGCCCCGTGCCCGCTGGTGAAACTTTCCCCATTCACCGTCCCGCTCCCAAATTGGTGGATTTGGAAACCAAGCCCCAAGTATTTGAAACCGGCATTAAGGTAATTGACCTGCTTACTCCCTACCGTCAGGGTGGCAAAATCGGTCTCTTCGGTGGTGCTGGTGTGGGCAAAACCGTAATCATGATGGAATTGATTAACAACATCGCCATCCAACATGGTGGTGTATCTGTATTTGGTGGCGTAGGGGAACGGACCCGGGAAGGGAATGACCTCTACAACGAAAT	UTR and part of *atpB*
11	GATCGCCCCTTTCGGAAATGTGCCCGGGTGGTGGGGGAAGTGCTGGGGAAATATCATCCCCACGGTGATACGGCGGTGTACGATGCCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCCGTGGATAAT	part of *gyrA*
12	GATCCCGGTAGTGGCCGAAAAAACACCCCCATGGGACTACTCGAAACCCAAGTTCACCACCTGCTCAAAGCCCAACAGCGCCAACAAACCATTCCCCCCTGGCCCCATTCGTTAACCATGGGTCGTTTGGTATCCCGGGCCCTGCGACTACGGC	part of *slr1495*
13	GATCAGCGTAGTTTTGATTTTTTGATTTGATATTCCCTGGGAGTAACTTAAACCTATGTCAATTGCAGTTGGGATGATCGCCCCTTTCGGAAATGTGCCCGGGTGGTGGGGGAAGTGCTGGGGAAATATCATCCCCACGGTGATACGGCGGTGTACGATGCCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCCGTGGATAAT	part of *gyrA*
14	GATCGCCGTGGTTTCCTTTCCATAGTTATGGTCACGGATAGCTTTAATGC	part of plasmid pCC5.2_M
15	GATCGCCAAAGGGTTTCAGCCGGTGTCCCTTGGTAAACGCATTTTGCGGGCGGTGACGGCCCCCACGGTGGCCCTAGCCCA	part of *slr0722*
16	GATCTGCCAATCATTGCGGCCAACGTTTAAGATGGGGGCATTGTTTTAGTCCACCATTGTAATTACCAACGGAGGGAAAATGTTACAACTACTTATCGTCTGGATTGTCACCTCAGTCAGCCTGCT	UTR and part of *slr0284*
17	GATCGCCTTTTTGGGCACGGAGTAGGGCGTTACCCCGGCCCGTTCAACCACAAGTCCCTATAGATACAATCGCCAAGAAGTATGTCTAAAAAGTTTTTAACAATCCTCGCTGGCCTTCTGCTGGTGGTCTCCAGTTTCTTTTATCCGTGAGCCCCGCTGCCGCTGCCAATGCAACAGTGAAAATGGGTTCTGACAGTGGTGCCCTAGTTTTTGAACCCAGCACCGTCACCATCAAAGCTGGAGAGGAAGTCAAATGGGTTAACAACAAACTCTCCCCTCATAACATTGTGTTTGCCGCCGATGGTGTGGATGCTGACACCGCTGCTAAGCTCTCCCACAAAGGCTTGGCCTTTGCCGCTGGTGAAAGCTTCACCTCCACCTTCACTGAGCCCGGCACCTACACCTACTACTGTGAACCCCACCGGGGTGCTGGCATGGTAGGCAAAGTTGTCGTTGAGTAATCAGCCAGCTCAATCTGTGTGTCGTTGATTTAAGCTTAATGCTACGGGGTCTGTCTCCAACTCCCTCAGCTTCTCGCAATGGCAAGGCAAATAATGTTTCTCTTGCTGAGTAGATGTTCAGGAGGACG	UTR and part of *petE*
18	GATCGCCTCTTGCAGTCCCCCGCCGATACTCCCCACCTGCTCTCCTTGGCAAGGCTGAGAATCCGTTACAAAAATTTTCCCGGGGCCAGGAGTTTACAGCGGAGATTTGGACACCGTGCTTCAACAATGGCATCTGACTGAAGAAGAACTCTTTGCCCAAACTAGAGCCCTCTATGCCAGCGGTGCGGCCCATGCCAAACAGAGTACCGAGGAGCAACAGGACTGGAGCTAGCTCTCGTTAATCCATTGCCCCAACTAGGCTGAGATGATTTTCCGGGTGGAACCCTCGAGTGTGCCGATGTCGACCTCCACCATAGGCATTGAAGTTAATCTATGGTCAAGCTTTTTAGCAAGTATTAATCCCTACTATTGGTGCGATAAGCAATGCCAAAGCCAAAACCAGAGATTTTCCCCTATGCTTGAAGCAACATCTCCCGTTACAAAAAATTAACTAAATGTCTAATTTGATTGCTGTTGCTTTCTGTGGCTACTTTGCCCTGTT	part of *ssr2551*
19	GATCAAAAATTGGGGCACAAACTCAATGCTAAATTAGAA	part of *slr6016*
20	GATCATCAATGAACTATGCTCATACAACAGCAACAATTCACCGCCAATGATAAGGTG	UTR of *slr0789*

Colony No.	Sequence	Annotation
21	GATCGCCCAGAATGGAGTTTCCCCTTTGACGGTGGACGTTTTAAAAAAACCTGAGCTACACCATTCCCGACCAAGGTACATATTCCCTTAACAATGGTGTGTTCGCTGGGGGAACTTTTAGCCTTAATCT	part of *slr1127*
22	GATCGCCCCTTTCGGAAATGTGCCCGGGTGGTGGGGGGAAGTGCTGGGGGAAATATCATCCCCACGGTGATACGGCGGTGTACGATGCCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCCGTGGATAAT	part of *gyrA*
23	GATCCCGATGCAAAATACACTGGATGTTTAACATTTCAGGAATTTCCTCTGAACAATAGTAAGTTCCATGAGAATTTTCCCGGTACAGCAATTCGCTTTTCTTACAAGGGAAACTCTACTAACATGGGAACGATGGTTATTCAGTCATCTTTTTCTCCCACAGTAAGTTATTGTTTGGCCATGTCTGGCATGATTGGTGTAATGAGAGCAGGGGTTTACGAAGGA	part of *slr1930*
24	AACAATAATTATTTCCGCGTTGGTGGGGTAGCTGTGGATTTGCCCTACGGTTGGAATGATAAATGCGAGGACTTTTGTGACTACTTTTTGCCCAAGGTTGACGAGTACGAAAAACTAATTACCAACAACCCCATTTTCCGTCGCCGGGTGGAGGGGGTGGGCACTGTGACCAGGGAAGAAGCCATTAACTGGGGATTATCGGGCCCCATGCTCCGAGGTTCCGGGGTGAAATGGGACCTGCGTAAGGTG	part of *ndhH*
25	GATCTAGCTTATCAACCTTGTTGTAAAATTTAGTTAATTCATCCAAAACATGAGCAAATGGAAATACTTTTAACCACTTTTTGATAAGATTTATTGCATTAAATACAGA	part of *slr1684*
26	GATCTCCAACTCATAAAGTCAAGTAGGAGATTAATTCCATGGGCCATCATCATCATCATCATCATCATCACAGCAGCGGCCATATCGAAGGTCGTCA	UTR *cpcB*
27	GATCGGACAAACCTTCACCGAGGAGGGCTATTTTTGGAGTCGCCCCAGTGCCATTAACTACAGTGAAGGGGCAGATGCTTCCCCCACGGGCATTTCCGGTGCAAGCAACCTAGCTCCCAGTAATCCTGACCTTTTATCCCGCATTGAAGCAGAGGCCCAACGCTTAGAGGACAATGCCGTGCAACCCACTGCCGATTTGCTTTACAGCTCTGGGTCAGGGCTT	part of *kdpC*
28	GATCGCCCCCTTAACCATGGATGTGATTAACGATGCTGATGGAGGTAAAATTTGTTATCAGATTTATCAG	part of *sll0509*
29	GATCACCCTACCAGTCAAAGACATGGTTAAAATTACAAAAAAGGTTAATACCGACATCAAGACATCAATCATCGGTACTAAATTAATCCCTGGAGTCCTTGCCCCCCTGACTCGGCTTTCATCGGTTATGGCTCAGCAATTCTAGCCGCCATTAATAAATTGGGCATTGGCTGATGAGGTTCGTACCAGACCTCACGGTAAATAAGTTCTAATTCACTGCCCACTGCGGCAAAATAATCCATTTGTTGGGACTGTAAACTGACCAAAACCCTAAATACTAACAGGGCAAAATCGCCACCATCATCCCCGCTGCCGTGGTAATCAACGCTTCTCCAATGCCCGAGGCGGCCTGGGTAGCCTCCGCACTGGAGCCTCCTCCCCGATATTGAGATTGTTAAAAGTGCGAATTAAGCCCGTTACGGTGCCTAACAAGCCCAAGAGAGGAGCTAGGGCAATGATAGTTTCCAGGAGTTTATCCCCCCGTCGCATTCGGGCAATTCTTTATCGGCGGTGGCTTCCATGGCCAAACGGAAAGTTCTGGACTAGGAAGGCCGTGGCGGGCCTCTGCCCAGTCCCGCA	overlapping *sll0478* and *sll0477*
30	GATCAGCTCAGCGTGACCCAGTTCCTCATCCCCTTCTAGGGCATCAAAATTAATCACTTCCCCATATTGTTCTAGTTCTCCTTCATCAGT	part of *slr6004*
31	GATCGCCAGCGAAGCCTACAAATATATGCAATTCCAGCCGGGACAACCGCTGAAAGGCACCAAGGTCGATGTTTGTTTCATTGGTAGTTGCACCAATGGTCGTTTGAGCGACCTGGAGGAAGCCGCTAAGGTAGCCAAGGGTCACCAGGTAGCGGCCGGGGTAAAAGCTTTTGTGGTGCCCGGTTCCGAACAAGTCAAGCAACAGGCCGAAGCAGCAGGGCTA	part of *leuC*
32	GATAGCCTATCCCCCCAAGAATTGACCCATCTCACCAACCGGTGACACCCCCAAAGTGATACCCTAAAATACAACAAACAAAACCAAAACTCAACAGTAAACCTGTAGTCTAACCAAAAACAAATTCCTAGACAATCAGCATGTAACCCAAAAAAGAAACAACTAACGTTGGCAAAATTACCCAGGTCATCGGGCCTGTAATTGACGCCCAGTTCCCCAGTGGTAAATTGCCCCGTATTTATAATGCCCTTAAAGTCCAAGGCAGAAACTCTGCTGGTAACGAAGTAGCTGTTACCTGTGAAGTGCAGCAGCTTCTCGGCGATAACCAAGTCCGAGCCGTAGCCATGAGTTCCACCGACGGTCTCGTCCGGGGCATGGACGTGGTAGACACCGGGGCCCCCATCAGCGTTCCCGTCGGCACCGGCACCCTGGGGTCGTATTTTTAACGTTCTTGGTGAACCTGTTGACAACAAAGGCCCCGTGCCCGCTGGTGAAACTTTCCCCCATTCACCGTCCCGCTCCCAAATTGGTGGATTTGGAAACCAAGCCCCAAGTATTTGAAACCGGCATTAAGGTAATTGACCTGCTTACTCCCTACCGTCAGGGTGGCAAAATCCGTCTCTTCGGTGGTGCTGGTGTGGGCAAAACCCTAATCATGATGGAATTAATAAACAACATCGCCATCCAACATGGTGGTGTATCTGTAATTTGTGGCGTAGGGGAACGGACCCGGAAAGATAACGACCTCTACAACGAAAT	UTR and part of *atpB*
33	GATCGCCTTTTTGGGCACGGAGTAGGGCGTTACCCCGGCCCGTTCAACCACAAGTTCCCTATAGATACATTGGCAAAGAAGTATGTCTAAAAAGTTTTTAACAATCCTCGCTGGCCTTCTGCTGGTGGTCTCCAGTTTCTTTTTATCCGTGAGCCCCGCTGCCGCTGCCAATGCAACAGTGAAAATGGGTTCTGACAGTGGTGCCCTAGTTTTTGAACCCAGCACCGTCACCATCAAAGCTGGAGAGGAAGTCAAATGGGTTAACAACAAAATTTCTCCCTTCATAACATTGTGTTTGCCGCCGATGGTGTGGATGCTGACACCGCTGCTAAGCTCTCCCACAAAGGCTTGGCCTTTGCCGCTGGGTGAAAGCTTCACCTCCACCTTCACTGAGCCCGGCCACCTACACCTACTACTGTGAACCCCACCGGGGTGCTGGCATGGTAGGCAAAGTTGTCGTTGAGTAATCAGCCAGCTCAATCTGTGTGTCCGTTGATTTAAGCTTAATGCTACCGGGGTCGTCTCCAACTCCCTCAGCTTCTCGCAATGGCAAGGCAAATAATGTTTCTCTTGCTTGAGTAGATGTTCAGGAGGACG	UTR and part of *petE*
34	GATCTCGAATATCTTTTGACAACTCAGTGTAACAGTGGCTTCACGCTCCGTAACAGGTGGCTTTGTAGATTTCACACTTTTAAAGACGGGGTTTTAGACCGGGTTTTACGATAAACTTAGTAATTATTAGAGGAGCGTGGTGCTTTTCGCCCCCAGCACTGTTTACAGGAAGATAAGACGTGAATCAAACCAACCTCGACTTTTTAGCCACCAGCGACCCCGCCCTTGCGGCCATCATTGACCGGGAATTACAACGGCAACGCACCCACATTGAATTGATTGCCAGCGAAAATTTCACTTCAGCAGCAGTAATGGCGGCCCAGGGTTCCGTCTTAACCAATAAGTATGCCGAAGGTCTGCCCGGGTAAACGCTATTACGGCGGCTGTGAATTTGTGGACCAAGCGGAAACGCTAGCCATCAGTCGAGTCAAAGAGTTATTTGGGGCTGCCCACGCCAACGTCCAACCCCATTCCGGGGCCCAGGCCAATTTTGCCGTATTTTTAACCCTGCTCCAGCCCGGCGACACCATTATGGGCATGGACCTTTCCCACGGCGGCCATTTAACCCACGGTTCCCTTTTCAACGTTTCTGGTAAATGGTTCGAGTTGGCCCACTACGGTGTGGAAAAGGAAACGGGTCGGTTAGATTACGACAAATTCGGCCAACAGGCTCTGGAAGTTAAACCCAAACTGTTGATTTGTGGCTACTCTGCCTTTCCGCGACAAATCGAGTTTGATAAATTCCGGGC	UTR and part of *gylA*
35	GATCGCCTGATTGCGACTGACTTCAGAAACGAAATGTCACAGTCATACCTGGAATATGCCATGAGTGTGATTGTTGGGCGGGCGTTACCCGATGCACGGGATGGGTTGAAGCCAGTCCATCGGCGCATTCTCTATGCCATGTACGAGCTGGGGCTGACCCCCGATCGCCCCTTTCGGAAATGTGCCCGGGTGGTGGGGGGAAGTGCTGGGGGAAATATCATCCCCACGGTGATACGGCGGTGTACGATGCCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCCGTGGATAAT	pat of *gyrA*
36	GATCGCCTGGAGGGTTACATCAACCCCCCATGACTTTTATCAGTGGCCAGCTCACCATTGATGACCCCACTATGCTCCGCCTGGAACAGTCCGCCTTTGCCTTTGCCCCCATCCGCAATGCCGGAGGATTTGTCACCCTCGACCATGCCCCCATCGAAACGGC	part of *slr6006*
37	GATCCCTTTCTGGAAAAATCCTTAATTGAAGCTTGCTTGGAAGCCTTTAAAAGTGGGGCAGTGGTGGCGGCCCAGGACATGGGAGCGGCGGGGCATTACCTGTTCCACGGCGGAAATGGCAGCCAAGGGAGGCGTGGGCATTGAGTTGGATTTGGATAAAATTCCAGTGCGGGAAACGGGCATGGTGCCCTACGAATATCTCCTATCCGAATCCCAAGAACGGATGTTATTTGTGGCTCAACCGGGGCGGGAACAGGAATTAATCGATATTTTCCATCGTTGGGGTTTACAGGCGGTGGTGGCGGGGCAGGTGATTGCG	part of *purL*

Colony No.	Sequence	Annotation
38	GATCCCCTGTTACTGACGGGAAAGGTTAGCTTTGGGGGCATCATTAAAACCATTAGCCTTGCCTACGTACCCGAGGTTAAG GTGGGGGATTACGTGATTGTCCATGTGGGCTTTGCCATTAGCATTGTGGACGAAGAGGCGGCCCAGGAAACTTTGATAGA CTTGGCAGAAATGGGAGTTTAATTCCTAGAGTTAAGAATAGAGAAATAGACACGACTCCAGAACCGTTACGGGTGACAAA TAATGTACGGCCCA	UTR and part of *act*
39	GATCGCCTGATTGCGACTGACCTCAGAAACGAAATGTCACAGTCATACCTGGAATATGCCATGAGTGTGATTGTTGGGCG GGCGTTACCCGATGCACGGGATGGGTTGAAGCCAGTCCATCGGCGCATTCTCTATGCCATGTACGAGCTGGGGCTGACCC CCGATCGCCCCTTTCGGAAATGTGCCCGGGTGGTGGGGGGGAAGTGCTGGGGGAAATATCATCCCCACGGTGATACGGCG GTGTACGATGCCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCC GTGGATAAT	part of *gyrA*
40	GATCGCCTATCCTCCTACGATTTGACCCATCTCACCAGTTGGGGACAGCCCCAAAGTGATACCCTTAAAGATGGCAAATTTG ACCAAAACTCAACAGTAATCCTGTAGTCTAACTATAAACAAATTCCTAGACAATCAGCATGTAACCCAAAAAAGAAACAAC TAACGTTGGCAAAATTACCCAGGTCATCGGGCCTGTAATTGACGCCCAGTTCCCCAGTGGTAAATTGCCCCGTATTTATAAT GCCCTTAAACTCCAAGGCAGAAACTCTGCTGGTAACCAACTAGCTGTTACCTGTGAATTGCAGCAGGCTTCTCAGCAATAA CCAAAATCCGACCCGTAGCCATGAGTTCCCACCGACACTCTCATCCGGGGCACGGACTCGACCAAAACACATGCGCCCCAA TAAACCTTCCCGTCCAAACCATCACCACGGCTGTCACATTATTAACGACCATAATCAGCCTAGCTTCACAACAAACACCCAC TTCCCACTAGCAAAACTTCCCCCAATCACCTTAACGGCATCCCAAAACGGTGCAATTGAAAACCAAGAACCAATCATTAAAA ACCGGCATTAAGGTAATTGACCTGCTTACTCCCTACCGTCAGGGTGGCAAAATCGGTCTCTTCGGTGGTGCTGGTGTGGGC AAAACCGTAATCATGATGGAATTGATTAACAACATCGCCATCCAACATGGTGGTGTATCTGTATTTGGTGGCGTAGGGGAA CGGACCCGGGAAGCCAATGACCTCTACAACGAAAT	UTR and part of *atpB*
41	GATCGCCAGCCCAGACCATTTGCAAAGTTTTCACTAGCCCGAGAATAATACACAGCACGGAAACACTTTCCAGCACAAACC TGGTAATGCTGACTAAATTGCCCAGATAAT	part of *sll0939*
42	GATCGTAACGGTAGATATTGGGCGAGTAATCAGAAAATAGGTGTTGCCATGGGAGCTAAAACCAAGCTCAGGGCCGATG GTCACTCGAACGGTGTTGCCCGCAGTGGAAGATACCTGGGGAGGAAGTGGTGCACAGGTTTCGGCGGCGGCTGTACCAG CAAATAATACAGCGCA	part of *sll0172*
43	GATCCCTTGGCCGGAGCGGAAGCGGTCATTTCCCATTTAATTGTGCGGCAATTTCAAGTGCCCTGTGCCCATGCCCCAGCG TTGAGTCCTTTGCCTTTAGACCCAGACATTTCCCCCCGGTCAGCGGCGGGAGGAATTGGGCTACACTTTTTTACCCTGTGTAT TGATGGGACTGAGTCGGGCTCCCCATTACCAAAGATTTCCCAATGCCCAGGGCCAAACCCTCTGGAGTCGGGATGTGAAT GCTGTGGTTGTGCCCGCCAGCGCCTGCGGTGGTTCCGCTTTGATGAGTTTTAGCCA	part of *slr1287*
44	GATCTCCGCGGCGGCAATTTCCCCCGTGTTTAGCATTAGCTGAAAATTAGGGGGAGTGAGGGGGGGGTGGGAAAAAGATG CCATGCATTATCAAAGGGAATGTTATCATTGTTTTCCAGATAAAACAAAGAAACACTCTTCACTTTGTTGTGTTTTAATAAAC GAAGGGGTGGACCCATTAGTTGGTAATGGAGAAATAAACCCAGGTTTCTATGACTTGTTTAAATCTGCCATCTCGAAAATA TTTCCCTCATGGTGGGGGAAAGATTCGGGAAAGGTTTTCCCCACGATGATGCGGCGAGGTATGATGCCGAGGTGCGCTAG AAGCTGGACTTTTTCAGGCGGGAAGTGGGTAGTGACGCCCACGGGAACTTCGGCTCCGAGGATAATAATGT	part of *slr0168*
45	GATCGGGTAAGTAACAATAATCACAATCTAAATTGCAATAGGAAGTGGGTTGAATAATAATTAAATTGGTGGGGCCAAAA TTAGTTATGTCCACATCCCTTATACTTTCATTAATGGGTTGGGTTTGGTTGACTAAAGCTTGCATAAGT	part of *aslB*
46	GATCATCATGGAGCGGGCTTGCCATTTGGCCACGTTCCACAGGGGCTTGACGGCCAATAAACCATTGACGACCTGGCTCAT GGCGGTGGGGGCAGGGGTGGCGGTGCTGACAGTCATCATGGAATGCTTTAATGGGACTAGGGAAAATTGTCGCTGAAAAATC CAAACCTGTCTAGGCATGGGTATCAAGCGGC	UTR and part of *sll0564*
47	GATCGCCGCTTGATACCCATGCCTAGACAGGTTTGGATTTTTCAGCGACAATTTTCCCTAGTCCCATTAAAGCATTCCATGA CTGTCAGCACCGCCACCCCTGCCCCCACCGCCATGAGCCAGGTCGTCAATGGTTTATTGGCCGTCAAGCCCCTGTGGAACG TGGCCAAATGGCAAGCCCGCTCCATGAT	UTR and part of *sll0564*
48	GATCGCCTATCCTCCTAGGATTTGACCCATCTCACCAGTTGGGGACAGCCCCAAAGTGATACCCTTAAAGATGGCAAATTT GACCAAAACTCAACAGTAATCCTGTAGTCTAACTATAAACAAATTCCTAGACAATCAGCATGGTAGCAGCAAAAAAAACAA ATAACGTTGGCAAAATTACCCAGGTCATCGGGCCTGTAATTGACGCCCAGTTCCCCAGTGGTAAATTGCCCCGTATTTATAA TGCCCTTAAAGTCCAAGGCAGAAACTCTGCTGGTAACGAAGTAGCTGTTACCTGTGAAGTGCAGCAGCTTCTCGGCGATAA CCAAGTCCGAGCCGTAGCCATGAGTTCCACCGACGGTCTCGTCCGGGGCATGGACGTGGTAGACACCGGGGCCCCCATCA GCGTTCCCGTCGGCACCGGCACCCTGGGTCGTATTTTTAACGTTCTTGGTGAGCCTGTTGACAACAAAGGCCCCGTGCCCG CTGGTGAAACTTTCCCCCATTCACCGTCCCGCTCCCAAATTGGTGGATTTGGAAACCAAGCCCCAAGTATTTGAAACCGGCA TTAAGGTAATTGACCTGCTCAACCCCTACCGTCAGGGTGGCAAAATCCGTCTCTTCGGTGGTGCTGGTGTGGGCAAAACCC TAATCATGATGGAATTGATAAACAACATCGCCATCCAACATGGTGGTGTATCTGTATTTGGTGGCGTAGGGGAACGGACCC GGGAAGGGAATGACCTCTACAACGAAAT	UTR and part of *atpB*
49	GATCGCCTTCCATGGGAATAGATTGGGGGTGAAAAATTTTTGACAAATTACTAAAAAATTCATTAACATTCAGCTCCCATAG CACTTACCTTTTTCCACTATACATCAGTGCCTAAGAGCCTGTTGGTCGAGTTTTAGCTTGTCCCTAAAACCGAGCCGTTAGA AATCC	UTR and part of *sll0274*
50	GATCACTGTTTGCCGATGCCAGTAAAAGATTAGAAAAAGCCCTGAAATATGTGGCCATTTCCGACGATGCCGGGGAAAGG CTGAAATATCCCAAAACCAGTCTCAGTGTTTCCATTCCGGTGCGGATGGACGATGGCTCTTTAAAAAATCTTTCCCGGTTACC GAGTGCGCTACGACGACACCAGGGGGCAGGCAAAGGGGGCGTGCGCTACCATCCCAACGTCACCATGGATGAGGTGCA ATCCTTAGCCTTTTGGATGACCTTTAAATGTGCCCTGTTAAATTTGCCCTTCGGAGGAGCTAAGGGGGGCATCACCCTCAAT CCCAAAGAACTCAGCCGAGCAGAATTGGAGCGTTTAAGTCGGGGTTACATCGAGGC	part of *gdhA*
51	GATCTAGAAATAGGACAAAAAATAGTAAATTCCCATGGCGGGGAATTTTAGAGAGTGGGCTAAGGCTCCTGGCCTCCTGC CTCGATTAAGTTTTGCGACTATTTCGCCCCGTCGTCCCATAAAAGGTTCCTAAACTTGCTATAAGAG	part of *sll1582*
52	GATCAAGGACAACCAATGCTTCTGATGGTGAAAGCCCTAAAGTTGATTCTAGGTTTGATTTTAAT	part of plasmid pSYSM_M
53	GATCAGCGTAGTTTTGATTTTTTGATTTGATATTCCCTGGGAGTAACTTAAACCTATGTCAATTGCAGTTGGGATGATCGCC CCTTTCGGAAATGTGCCCGGGTGGGTGGGGGAAGTGCTGGGGAAATATCATCCCCACGGTGATACGGCGGTGTACGATG CCTTGGTGCGCATGGCCCAGGACTTTTCCATGCGGGAACCGCTTATTGACGGCCACGGGAACTTTGGCTCCGTGGATAAT	part of *ccmK1* and *slr0513*
54	GATCAGCGTAGTTTTGATTTTTTGATTTGATATTCCCTGGGAGTAACTTAAACCTATGTCAATTGCAGTTGGGAT	UTR and part of *ccmK1*

Colony No.	Sequence	Annotation
55	GATCCTGCCAAAACCATATCGTGTTCTAACCCCTATATGCAATTTATATGCAATAATCAGAGGCCAAACTCCGGCAATTGTC CAATTTTTTTTGCATTACTGCCAGCCAACTTCATTCATTAAACGAGCTGAATCGGGTTGGTAACGTCCCAAGTTAGAAACAT TGAGGGGGTCTCCTTTTAGTTGGCCATGGGCGGCTAGCACAG	overlapping *slr0513* and *slr0514*
56	GATCCTGCCAAAACCATATCGTGTTCTAACCCCTATATGCAATTTATATGCAATAATCAGAGGCCAAACTCCGGCAATTGTC CAATTTTTTTTGCATTACTGCCAGCCAACTTCATTCATTAAACGAGCTGAATCGGGTTGGTAACGTCCCAAGTTAGAAACAT TGAGGGGGTCTCCTTTTAGTTGGCCATGGGCGGCTAGCACAG	part of *slr1403*
57	GATCGTCCTTCTGCCGGGGGAGGGGGCGGAAATTTTAACGCCTCTGGTAATGCCAACAATGGCGGTGCCGCAGGGTCTGT TGGTGATGATAAAGCTGGGGGAACTGGGGGAACTGGGGGAACTGGAGGTTTCGG	part of *slr6063*
58	GATCAGCTCAGCGTGACCCAGTTCCTCATCCCCTTCTAGGGCATCAAAATTAATCACTTCCCCATATTGTTCTAGTTCTCCTT CATCAGT	part of *slr0806*
59	GATCTACCTGGCGTTTAAAAATTAATTGGGATTCTGGCAGAGGTTGTACCGACCCTCCGGCTAAACTATAGGTGTGGGGAT TGGCGATTCTGGCCCCTAGATTTTGGTGGGCGGCCATGATTTCTTCCAGGC	UTR and part of *guaB*
60	GATCCCCACTGGATTAGGGTTATTAGGCTGTGCTGACGTCCAAAGGGTCAGATTATTTGGTACGCTAATTTAAACAGTTTG GGAATAGTAAGTAAGCGTGAACATTACGATTGGACGGGGAAAAACAGCCCGTCGAGCCTATGGTATCGACGAAATTGCA CTGGTTCCCGGAGTGCGGACCTTAGACCCAGCCCTGGCCGATACCCGCTGGAAAGTGGGCGCCATTGAGCGGGAAATCCC CATTATTGCCAGCGCCATGGATGGGGTTGTGGACAGCCGCATGGCAGTACTACTATCCGAACTAGGCGCTTTGGGGGTAG TCAATTTAGAGGGCATTCAAACCCGCTACGAA	part of *sll8049*
61	GATCGACCTAGCAGATGAGGATGCCGAGATTAAACCTGTGCCAACCGCCGGCGGGGGCAGTAAACCGGAGGCAGACCTA GACCAGTTAAGTAACATCATCAAGGCTTTCAAT	part of *slr1951*
62	GATCGCCCGATGGTGTAGAAGGTTTCAGTGAGTAAAACTTCTCGACGGCCAGCGTCATCTAAGAT	overlapping *slr2024* and *slr2025*
63	GATCATTTTCCCCTTTCTCCGGTAAGCTGTACTAGCACTTCCACCACTGCCATAGGCAAATTTCCAGATGCTGCCCTGGTCGT TATCTTGCACAAAGGCGCTATTTTCCTGGGCCATGCCGGAAATAACCGTTTCCACTACATCGTGATGACTAGCAGGGGTCA TTTCATCCATGGGTGGGGTTGGCTAACTCTGGTGTAGGGGATGCCATAGTCATACTCTCAACTTCTCAACACAAAGGTTTCA CCCCAAGTCTAACGGTCAACAGGCCCAATAGCATAGCCTAGGAAAGCCGTTGCTTGATGAAAGTTACTAATTGCCCCAACT GTTTTTTTCAACAGGTCAATTTCGCCGTGCATTTGGGCTACCTGTTGTTTGAGTTGTTGCACTTCAGCGGCCAAATCTCCGGC TGGGGCAGCGCCACTGGCAACGGCGGCCACAGGAACGTGTTTTTTAGCTTT	part of *pyrG*
64	GATCAGTGCCGAGGAGATTGAAGCCCAAGGGGCAGCCACCTTTCTGAAAGATGTGGACGGCGTGTTAGTGCCCGGCGGT TTTGGCATTCGGGGGGTGGACGGCAAAGTCCAGGC	part of *petE*
65	GATCGCCTTTTTGGGCACGGAGTAGGGCGTTACCCCGGCCCGTTCAACCACAAGTCCCTATAGATACAATCGCCAAGAAGT ATGTCTAAAAAGTTTTTAACAATCCTCGCTGGCCTTCTGCTGGTGGTCTCCAGTTTCTTTTTATCCGTGAGCCCCGCTGCCGC TGCCAATGCAACAGTGAAAATGGGTTCTGACAGTGGTGCCCTAGTTTTTGAACCCAGCACCGTCACCATCAAAGCTGGAGA GGAAGTCAAATGGGTTAACAACAAACTCTCCCCTCATAACATTGTGTTTGCCGCCGATGGTGTGGATGCTGACACCGCTGC TAAGCTCTCCCACAAAGGCTTGGCCTTTGCCGCTGGTGAAAGCTTCACCTCCACCTTCACTGAGCCCGGCACCTACACCTAC TACTGTGAACCCCACCGGGGTGCTGGCATGGTAGGCAAAGTTGTCGTTGAGTAATCAGCCAGCTCAATCTGTGTGGTCGTT GATTTAAGCTTAATGCTACCGGGGTCTGTCTCCAACTCCCTCAGCTTCTCGCAATGGCAAGGCAAATAAAGTCTTCTCTTCC TGAGTAGATGTTCAGGAGGACG	part of *putA*
66	GATCGCCCCGGTGATGGTGCGGTTGATGTAAAGATTGCCTACTTCAAACTCCGCCGCGGCACGATTAATGT	part of *accD*
67	GATCGGTGGGCCGCAAATGTTCATTGATGGGTTGCCAGGTCTTGGCGTCAATCAACTGGCGAATCCGCTCGTCACTGTCCA CCCGGAGATGGTGCCCACATTCGACACACACCATCCAATTGCCCTGCA	part of *slr1982*
68	GATCGGCAGACTCATGTCCATTAGAATAAGCTGGGGAGACTCCGAGATAGCCATGGTCACAGCCTGTTCACCGTCAACGG CTATCACCACCTCGTATCCTTTGCG	UTR and part of *atpB*
69	GATCGCCTATCCTCCTAGAATTTGACCCATCCCACCAACCGGTGACACCCCCAAAGTGAAACCCTAAAAGACAACAAACTT GACCAAAAATCAACAGTAACCCTGTATACTAACTATAAAAAAAAATCCTAGACAATCAGCATGGTAGCCGTAAAAAAAGCAA ATAACGTTGGCAAAAATACCCAGGTCATCGGGCCTGTAATTGACGCCCAGTTCCCCAGTGGTAAATTGCCCCGTATTTATA ATGCCCTTAAAGTCCAAGGCAGAAACTCTGCTGGTAACGAAGTAGCTGTTACCTGTGAAGTGCAGCAGCTTCTCGGCGATA ACCAAGTCCGAGCCGTAGCCATGAGTTCCACCGACGGTCTCGTCCGGGGCATGGACGTGGTAGACACCGGGGCCCCCATC AGCGTTCCCGTCGGCACCGGCACCCTGGGGTCGTATTTTTAACGTTCTTGGTGAGCCTGTTGACAACAAAGGCCCCGTGCC CGCTGGTGAAACTTTCCCCCATTCACCGTCCCGCTCCCAAATTGGTGGATTTGGAAACCAAGCCCCAAGTATTTGAAACCGG CATTAAGGTAATTGACCTGCTTACTCCCTACCGTCAGGGTGGCAAAATCCGTCTCTTCGGTGGTGCTGGTGTGGGCAAAAC CGTAATCATGATGGAATTGATTAACAACACATCCATCCAACATGGTGGTGTATCTGTATTTGTTGCCGTAAGGGAACGAAC CCGGAAAGAAAAACACCACAACAACGAAAT	UTR and part of *ahcY*
70	GATCGCCATTCCAGACCGATAAAATGTTAAGTTATATCACATTCAAGCAAAGTTCTTAGGAATTACCCCGTATGGTAGCAAC GCCCGTTAAACAGAAATACGATATTAAAGATATTAGCCTCGCTCCCCAAGGTCGTCAGCGCATCGAATGGGCGGCCCGGG AAATGCCCGTGTTAAAACAAATCCGGGAACGCTTTGCCCAGGAAAAACCCTTCGCCGGTATCCGCTTGGTGGCCTGCTGTC ACGTTACCACCGAAACCGCTAACTTGGC	UTR and part of *petE*
71	GATCGCCTTTTTGGGCACGGAGTAGGGCGTTACCCCGGCCCGTTCAACCACAAGTCCCTATAGATACAATCGCCAAGAAGT ATGTCTAAAAAGTTTTTAACAATCCTCGCTGGCCTTCTGCTGGTGGTCTCCAGTTTCTTTTTATCCGTGAGCCCCGCTGCCGC TGCCAATGCAACAGTGAAAATGGGTTCTGACAGTGGTGCCCTAGTTTTTGAACCCAGCACCGTCACCATCAAAGCTGGAGA GGAAGTCAAATGGGTTAACAACAAACTCTCCCCTCATAACATTGTGTTTGCCGCCGATGGTGTGGATGCTGACACCGCTGC TAAGCTCTCCCACAAAGGCTTGGCCTTTGCCGCTGGTGAAAGCTTCACCTCCACCTTCACTGAGCCCGGCACCTACACCTAC TACTGTGAACCCCACCGGGGTGCTGGCATGGTAGGCAAAGTTGTCGTTGAGTAATCAGCCAGCTCAATCTGTGTGTCGTTG ATTTAAGCTTAATGCTACGGGGTCTGTCTCCACTCCCTCAGCTTCTCGCATGGCAAGCAAATAATGTTCTCTGCTGAGTAGA TTTCAGGAGGACG	part of *gcvP*
72	GATCCGTCAGTATCACCTCAGTCGGGGTGAGGAACAACGCAATATTTGTTTAATCCCTGAATCGGCCCACGGCACCAATCC AGCCAGTGCGGTGATGTGTGGCATGCAGGTGGTCCCGGTTAAGTGCGACGGGGAAGGCAACATCGATGTGGAAGATTTG ACCAGTAAAGCAGAAAAGTACGGC	part of *trxA*

Table S2.2 Plasmids, strains and primers used in this study

Plasmids, strains, and primers*	Relevant characteristics	Reference
pFL-XN	BioBrick "T" vector with XbaI and NheI on each side	126
pWD42	Amp^{R}Kan^R, containing selection cassette (a Kan^R fragment and a nickel inducible *mazF* expression fragment)	127
pHKH-cpcBA-YFP	Vector containing the CPC-YFP cassette	119
pWD007	pWD42 derivate, replaced Kan^R with Sp^R for the selection cassette	This study
pFLXN	pFL-XN derivate, Amp^R, containing *slr0168* gene upstream homologous region and downstream homologous region	This study
pFLXN-MAZF	pFLXN derivate, Amp^{R}Sp^R, containing *slr0168* gene upstream homologous region, selection MAZF cassette, and downstream homologous region	This study
pFLXN-RBS-YFP	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, RBS-YFP cassette, and downstream homologous region	This study
pFLXN-CPC-YFP	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, CPC-YFP cassette, and downstream homologous region	This study
pFLXN-LIB-YFP	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, promoter library-YFP cassette, and downstream homologous region	This study
Synechocystis sp. PCC6803	*Synechocystis* sp. PCC6803 glucose tolerant wild type	D. Bhaya
Syn-MAZF	*Synechocystis* sp. PCC6803 with selection MAZF cassette integrated at *slr0168* locus, Sp^R	This study
Syn-RBS-YFP	*Synechocystis* sp. PCC6803 with RBS-YFP cassette integrated at *slr0168* locus, Kan^R	This study
Syn-CPC-YFP	*Synechocystis* sp. PCC6803 with CPC-YFP cassette integrated at *slr0168* locus, Kan^R	This study
Syn-LIB-YFP	*Synechocystis* sp. PCC6803 with promoter library-YFP cassette integrated at *slr0168* locus, Kan^R	This study
H1 0168 KpnI forw	GCAGGTACCAAGTGGGGCACATTGAACGC	This study
H1 0168 XbaI rev	GTGGACAAATCCCCCAGGTTTCTAGAGGCCACATTGTTGTCAAAGG	This study
H2 0168 XbaI forw	CCTTTGACAACAATGTGGCCTCTAGAAACCTGGGGGATTTGTCCAC	This study
H2 0168 HindIII rev	CGGAAGCTTAGGAGACTTTGGTGGGCTG	This study
BglII cpcRBS forw	AGATCTCCAACTCATAAAGTCAAGTAGGAG	This study
Pcpc forw	CGATCCCGACTTCGTTATAA	This study
KanTn rev	CGCTGAGGTCTGCCTC	This study
H2 0168 down forw	CACTCTGCACTGTGTCTGTGC	This study
YFP up reverse	CTGCTGCCTGCATATGACGAC	This study
H2 0168 middle forw	GTCACTGAAGCGGTCTAACT	This study
YFP middle rev	TTCAGCTCGATGCGGTTCAC	This study
Omega seq forw	GCTCACAGCCAAACTATCAGGTC	This study
NrsS N term seq rev	GCGTAATACCGCTTCCACCGTGG	This study

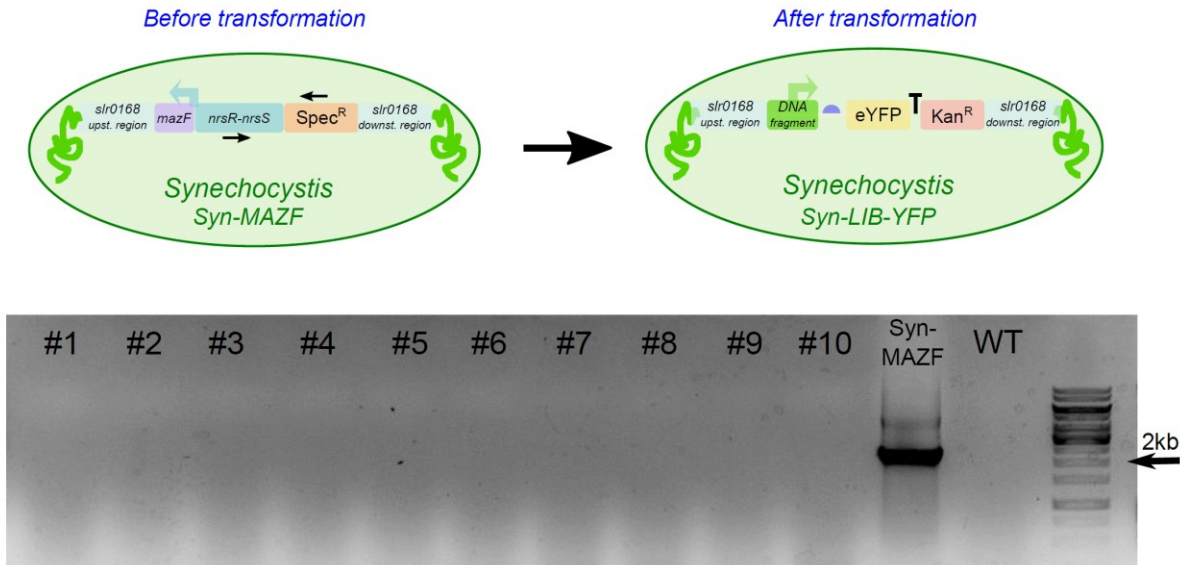* primer sequences are given from 5'→3'

Figure S2.1 Colony PCR verification of 10 representative *Synechocystis* colonies has shown that the constructed genomic DNA library in *Synechocystis* were fully segregated. Black arrows indicate the position of the primers. WT indicates *Synechocystis* wild type. This set of primers were chosen due to the small fragment size differences between the *slr0168* homologous regions.
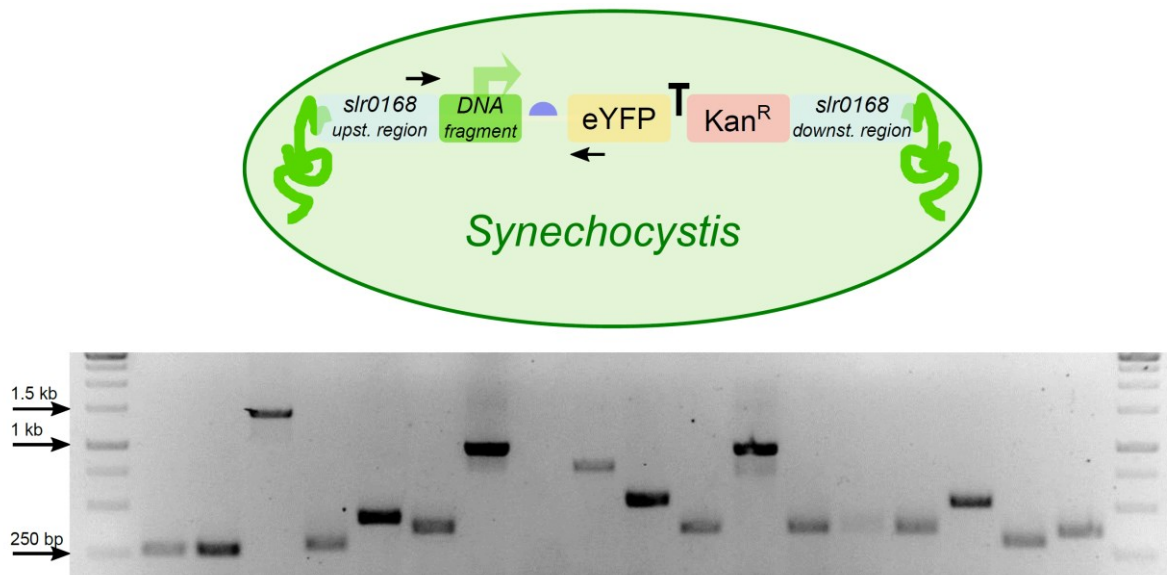

Figure S2.2 Colony PCR verification of a few representative *Synechocystis* colonies from the constructed fully segregated *Synechocystis* genomic DNA library. Black arrows indicate the position of the primers.

Construction of a promoter library in *Synechocystis* sp. PCC 6803 to increase the night production of fumarate

Patricia Caicedo Burbano, Yuki Esser, Wei Du, Filipe Branco dos Santos

Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Netherlands

Abstract: *Synechocystis* can be engineered to produce fumarate, a compound used to make certain plastics, food additives, and medicines. It has been probed that applying a growth couple strategy, the deletion of a single gene (*fumC*) involved in the TCA cycle of *Synechocystis,* results in a stable cell factory that produces fumarate as it grows during the day. However, at night cells don´t grow and are unable to produce fumarate. Diverging flux from the pentose phosphate pathway to the TCA cycle leads to an increase in the production of fumarate during the night. This can be achieved by removing the z*wf* gene. The incorporation of the glyoxylate shunt enzymes from *Chlorogloeopsis fritschii* can increase the night-time production of fumarate even further, if they are expressed only during the night. However, a promoter that can activate the glyoxylate shunt enzymes in a timely manner is not yet reported. In this study, we have constructed a promoter library to look for a suitable promoter to increase the night production of fumarate in *Synechocystis*. Even though an increment in the fumarate production was observed, suggesting the expression of the enzymes during the night, we could not find a native promoter in the genome of *Synechocystis* with the right timing and strength required to stabilize the night-time expression of the glyoxylate shunt enzymes, which affected fitness and resulted in the loss of enzyme expression over time.

Keywords: promoter library, fumarate production, glyoxylate shunt, *Synechocystis*

Introduction

Cyanobacteria cell factories have been studied and exploited for their capacity of sustainable production of carbon compounds directly from CO_2 and water[6]. Nevertheless, a limitation of large-scale applications of these cell factories is the tendency to lose their production capability[128]. This occurs because the pathways used for production directly compete with biomass formation, imposing a high fitness burden when biomass formation is preferred over production[129].

A stable production can be achieved using a "growth-coupled production strategy", i.e. coupling the synthesis of a specific target product to the growth of that cell factory, stabilizing the production[130,131]. Newly developed algorithms allow the identification of anabolic side products that can be coupled to cell growth by deletion of their re-utilization pathways[129]. One of these products is fumarate, a compound with very interesting applications in the food industry as an acidic regulator[132], in the plastic production as a suitable building block for (biodegradable) polymers[133] and as a medicine for psoriasis[134].

Synechocystis can be engineered to produce fumarate if the re-assimilation step in the TCA cycle, mediated by the enzyme fumarase, is removed by knocking out the *fumC* gene. Yet, a drawback of the growth-couple strategy, is that cells do not produce during the night, since during that time there is no growth. To survive the night, *Synechocystis* degrades its glycogen storage to produce energy[135], but instead of using the TCA cycle to provide the electron carriers for ATP production via respiration, *Synechocystis* prefers the Pentose Phosphate Pathway (PPP)[136]. Previous work showed that the deletion of the *zwf* gene, which catalyzes the first reaction of the oxidative pentose phosphate pathway, redirects the night-time metabolic flux towards the TCA cycle. This results in the accumulation of fumarate during the night as well[129].

The deletion of the *fumC* and *zwf* genes enables the production of fumarate during the day and night; however, by knocking out fumarase, the cyclic nature of the TCA is disrupted. In an attempt to reconnect the cycle and improve the night-time production efficiency in terms of carbon usage, an additional strategy was tested. As response of change in nutrient availability, bacteria have the capacity to use dissolved carboxylic acids, such as acetate, lactate, pyruvate and succinate from the surrounding environment[137,138]. To do so, many organisms including chlorophototropic bacteria, purple sulfur bacteria, and some cyanobacteria[139,140,141,142] possess the glyoxylate cycle[143,144].

This cycle, also known as glyoxylate shunt, is a modified TCA cycle in which two key enzymes, the isocitrate lyase (ICL, encoded by *aceA*) and the malate synthase (MS, encoded by *aceB*), convert isocitrate and acetylCoA into succinate and malate (Figure 3.1). Isocitrate lyase splits isocitrate into succinate and glyoxylate, after which malate synthase combines glyoxylate and acetyl-CoA to form malate with the release of CoA. Malate is then converted to oxaloacetate by malate dehydrogenase to continue the cycle (yielding NADH) and succinate can be consumed by succinate dehydrogenase to produce fumarate and $FADPH_2$[145].

The glyoxylate cycle allows the conversion of two units of Acetyl-CoA into succinate avoiding the CO_2 releasing steps of the TCA cycle (minimizing carbon losses) and enables the use of C_2 units for biomass production in a more efficient way[145]. The introduction of these enzymes in *Synechocystis*, could

reconnect the disrupted TCA cycle and increase the flux towards fumarate production, potentially aligning production with higher fitness during the night that the one observed for the Δ*fumC* Δ*zwf* strain.
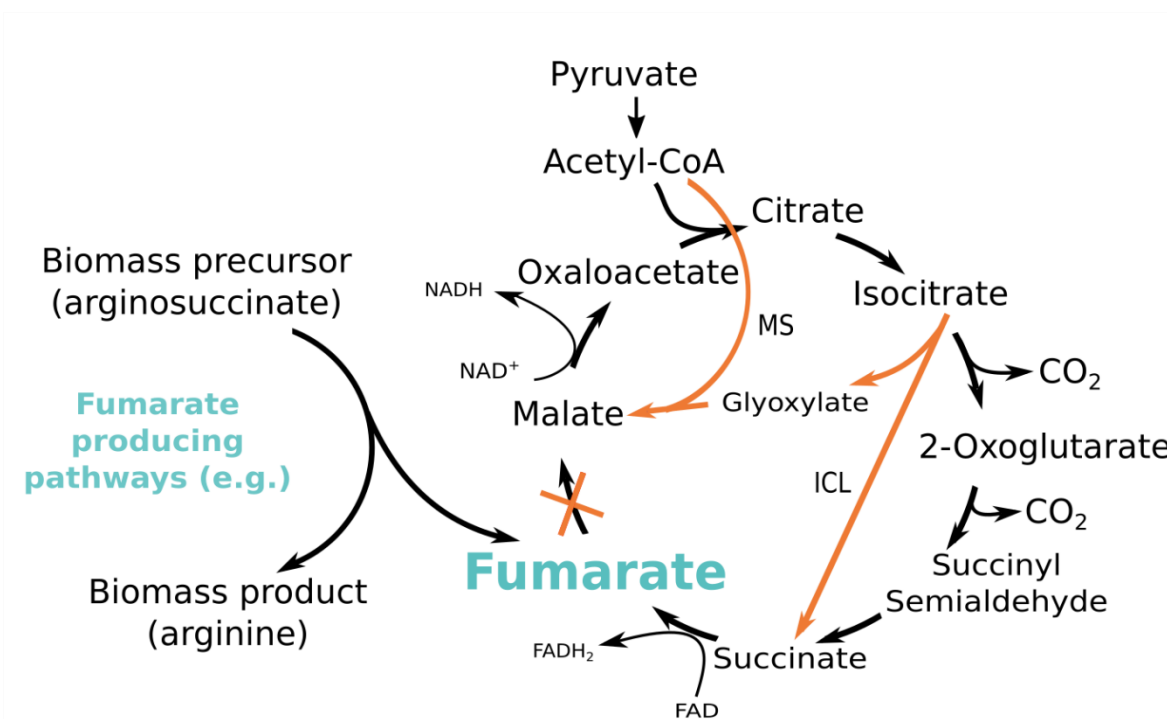

Figure 3.1 Scheme showing the glyoxylate (orange) and TCA (black) cycles. The glyoxylate shunt is comprised of two enzymes: malate synthase (MS) and isocitrate lyase (ICL). The orange X indicates the knockout of the *fumC* gene coding for the enzyme fumarate, leading to the accumulation of fumarate.

Predictions using a genome-scale metabolic model anticipates that the timing of the activation of the glyoxylate shunt enzymes is critical for the stability of the cells. The enzymes cannot be active during the day (only during the night), as it will draw carbon from more preferred pathways, and it will affect the growth rate. An option to regulate their expression is the use of an inducible promoter, but the delays between transcription and translation due to the mRNA or protein stability may affect the timing effect needed in this case. This creates a challenge to find a promoter with the desired expression pattern to activate the shunt enzymes only during the night. Circadian clocks in *Synechocystis*, which are synchronized to solar time, regulate the response to day and night cycles[146]. The existence of rhythms controlling gene expression and various cell processes, increases the chances that there is already a promoter in the genome of *Synechocystis*, able to express the glyoxylate shunt enzymes with the right expression levels at the right time.

In this study, we constructed an integrated library in *Synechocystis*, to find a suitable promoter to express the glyoxylate shunt enzymes with the optimal timing and level of expression required to increase the night-time production of fumarate. The promoter library was grown under continuous cultivation under day and night regime, to allow the natural selection of the promoter we were looking for. Our results showed that although there was expression of the glyoxylate shunt enzymes

during the night, we were not able to find a promoter in the genome of *Synechocystis* that could express the enzymes in the right timing and strength necessary to create a stable strain. Nonetheless, these results show the potential of this method to identify genomic elements with specific expression pattern under certain conditions.

Results and Discussion

Construction of promoter library

The promoter library was constructed by partially digesting the genome of *Synechocystis* in fragments ranging from 200 bp to 1 Kb. Those fragments were inserted on a suicide plasmid that integrates in the *slr0168* locus of *Synechocystis* by homologous recombination. This plasmid contained an insertion site for the DNA fragments, followed by a ribosome binding site, the malate synthase (MS) gene *aceB*, the isocitrate lyase (ICL) gene *aceA* (both from *Chlorogloeopsis fritschii*)[145], a terminator and a kanamycin resistance gene (Figure 3.2).
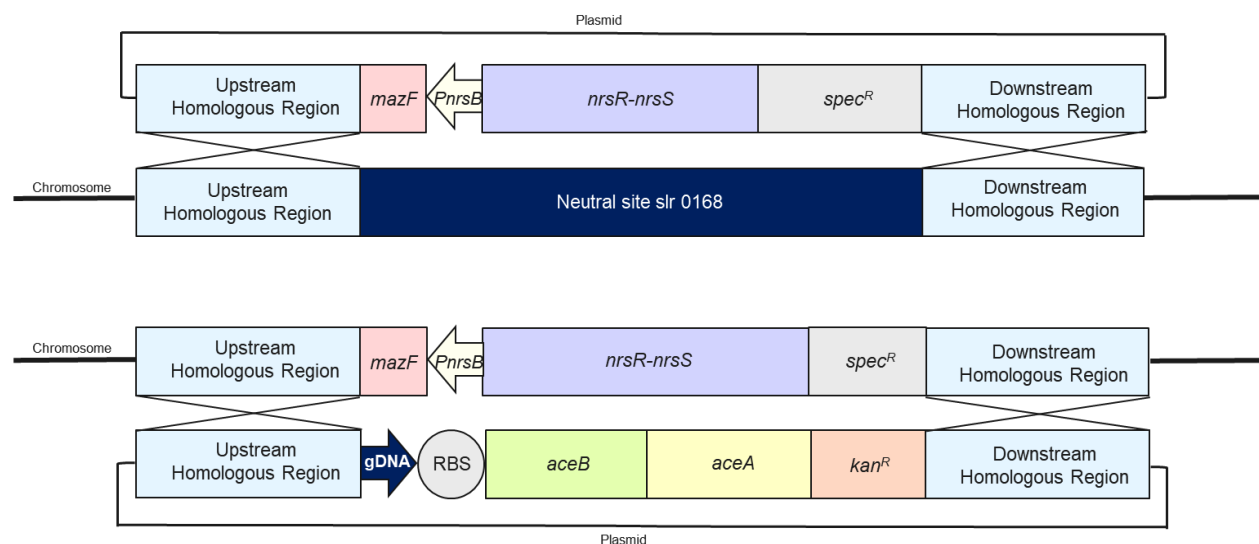

Figure 3.2 Schematic representation of the cassettes used to create the library. (Up) Vector used to insert the *mazF* cassette replacing the native gene *slr0168* in the chromosome of *Synechocystis*. (Down) Vector used to insert the promoter library, malate synthase (MS, *aceB*) and Isocytrate lyase (ICL, *aceA*), in the chromosome, replacing the *mazF* cassette.

We adopted a counter-selection strategy to ensure the full segregation of the library into the *Synechocystis* Δ*fumC* Δ*zwf* strain[118]. This method uses the nickel inducible expression of *mazF*, a protein synthesis inhibitor, as a selection marker only allowing the survival of the cells in which the promoter library was integrated in all the copies of the chromosomes. To implement this strategy, we first constructed a *Synechocystis* Δ*fumC* Δ*zwf* strain in which the *mazF* cassette was fully segregated (through spectinomycin selection) replacing the *slr0168* locus in all the copies of the chromosome. Posteriorly, the *mazF* cassette was replaced by the promoter library cassette (through kanamycin and nickel selection).

The screening of the library was performed by natural selection. *Synechocystis* Δ*fumC* Δ*zwf* carrying the promoter library (LIB-Δ*fumC* Δ*zwf*) was inoculated in a bioreactor under long term cultivation (100 days), mimicking natural day and night conditions. It was expected that if a promoter could activate the glyoxylate shunt enzymes (MS, ICL) during the night having the right expression level, that would confer a fitness advantage to the cell. The disrupted TCA cycle would be reconnected, and it would increase the flux towards fumarate production, by feeding into reactions that produce electron carriers in a more efficient way, avoiding the inherent loss of carbon that occurs in the TCA cycle. This would create a stable mutant and the presence of the shunt enzymes would be selected and maintained over time.

Expression of the shunt enzymes

Previous work showed that *Synechocystis* strain Δ*fum*C Δ*zwf* was able to produce fumarate during the day and during the night[129]. The expression of the glyoxylate shunt enzymes (i.e. MS and ICL) during the night in that strain, should increase both fumarate production and fitness as a result of a higher growth rate (or at least energy production). These two parameters were used to evaluate the expression and timing activation of the foreign glyoxylate shunt enzymes in *Synechocystis*. In addition to the promoter library constructed in the Δ*fum*C Δ*zwf* strain (LIB-Δ*fumC* Δ*zwf*), a promoter library was also constructed in the Δ*fumC* strain (LIB-Δ*fumC*) to be used as negative control for the fumarate night production and to evaluate the genetic stability. In the LIB-Δ*fumC* strain, the gene *zwf* is not deleted, therefore during the night the preferable route to generate energy is the pentose phosphate pathway instead of the TCA cycle. In these conditions, the expression of the glyoxylate shunt enzymes in the LIB-Δ*fumC* strain would be detrimental for the cells, because the glyoxylate cycle cannot be used during the night and a daytime expression of the enzymes would affect the growth rate. As a result, over time, the presence of the glyoxylate shunt enzymes would be selected against via transcriptional or translational mechanisms, such us inactive promoters, and/or mutations in the glyoxylate shunt enzymes genes.

When comparing the fumarate production between the strains LIB-Δ*fumC* Δ*zwf* and LIB-Δ*fumC*, it was expected a similar daytime production for both, but a higher night-time production for the LIB-Δ*fumC* Δ*zwf* strain (Figure 3.3). The results confirmed our prediction, showing a higher night-time fumarate production in the strain LIB-Δ*fumC* Δ*zwf*. Similarly, the fumarate yield along the continuous cultivation time (100 days) showed an increment of fumarate production in the strain LIB-Δ*fumC* Δ*zwf* visible after day 40. This result suggests that malate synthase and isocitrate lyase were expressed during the night, being able to reconnect the TCA cycle and increase the production of fumarate. As the library is composed by a heterogeneous population, where each clone should contain different fragments of the genome, the glyoxylate shunt enzymes are expressed under the control of different promoters with different levels and timings. This heterogeneity may explain the small increase in the fumarate production.
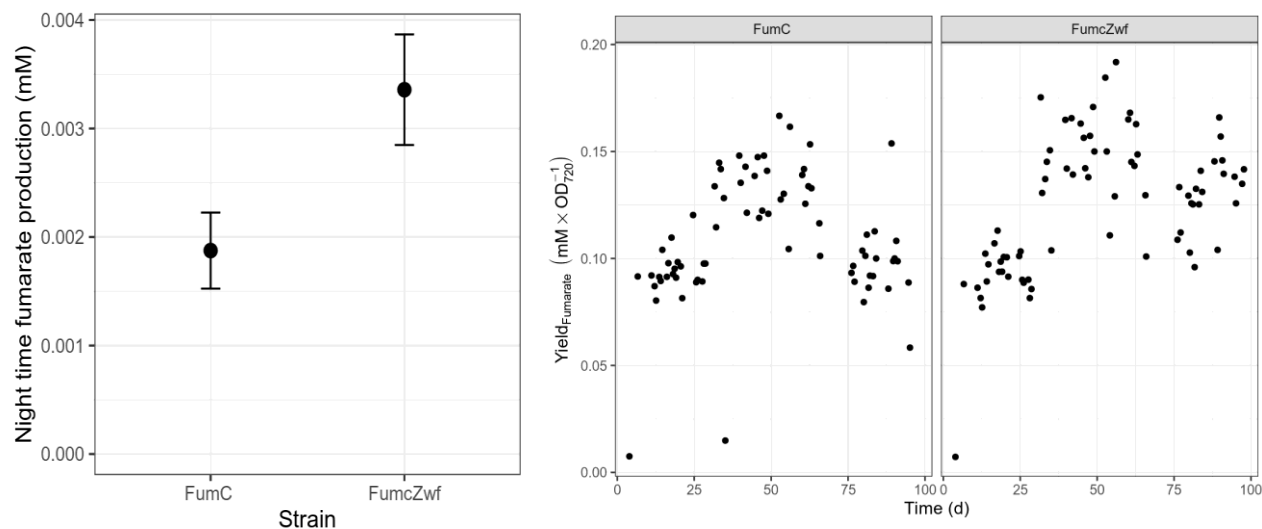
Figure 3.3 Comparison of the fumarate production of the strains LIB-Δ*fumC* Δ*zwf* and LIB-Δ*fumC*. (Left) Overall night- time fumarate production (mM) during continuous cultivation in both strains. Dots indicate mean ± S.E. (Right) Yield of fumarate (mM/OD$_{730}$) over time of continuous cultivation (100 days) in both strains.

It was expected that over time, according to the principles of natural selection, the clone carrying a promoter whose expression pattern of the shunt enzymes could confer a fitness advantage to the cell, would take over the population, prevailing over the clones with lower fitness. Analysis of the growth rate of the LIB-Δ*fumC* strain (negative control) showed no significative change over time of continuous cultivation (100 days). Contrary to what was expected, the LIB-Δ*fumC* Δ*zwf* did not show either a change in the growth rate that could indicate the presence of clones with higher fitness prevailing in the population (Figure 3.4). These results suggested, that even though there were cells in the population able to express the glyoxylate shunt enzymes during the night, increasing the nigh-time production of fumarate, the time and level of expression was not the ideal to confer an additional fitness advantage and instead could be affecting the genetic stability of the cells.
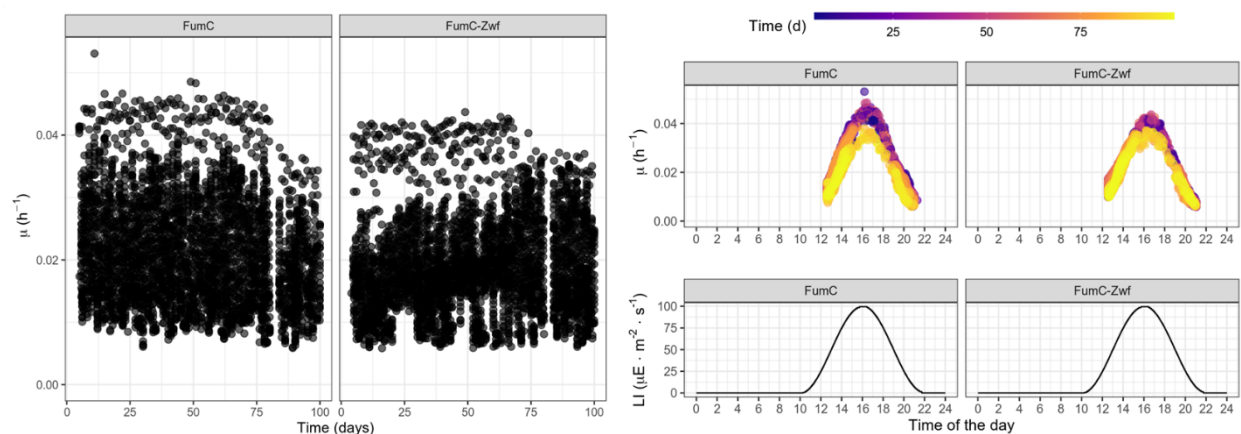

Figure 3.4 Comparison of the growth rate of the strains LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf*. (Left) Growth rate over time (100 days) in both strains. (Right) Variation of the growth rate at different hours of the day over time (100 days) compared with the light intensity at different hours of the day in both strains.

Characterization of the heterogeneity of the population over time

To have a better insight of how the heterogeneity of the population was changing over time, single colonies were isolated, and the size of the DNA fragments inserted was analyzed by PCR. This analysis was performed for clones of the LIB-Δ*fumC* Δ*zwf* and LIB-Δ*fumC* strains. The size characterization was done amplifying the region between the promoter and the beginning of the malate synthase gene (*aceB*), providing also information about the presence of this gene in the cells (Figure 3.5).
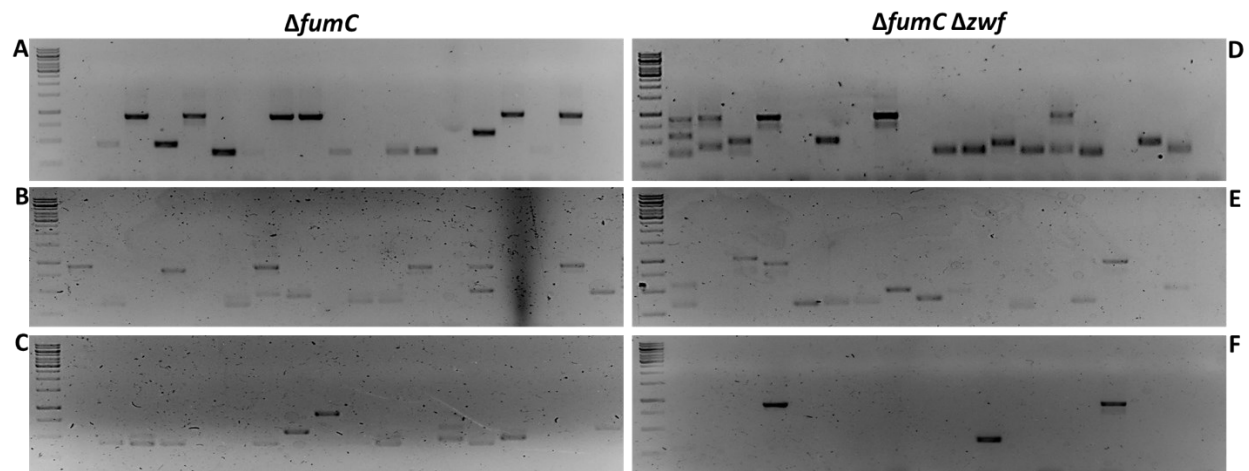

Figure 3.5 Comparison of the size variation over time of the promoter regions of LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf*. (Left) Size variation of LIB-Δ*fumC* after (A) 60 days, (B) 120 days, (C) 180 days of continuous cultivation. (Right) Size variation of LIB-Δ*fumC* Δ*zwf* after (D) 60 days, (E) 120 days, (F) 180 days of continuous cultivation.

The results showed that over time under continuous cultivation, the heterogeneity of both populations was decreasing, but also the number of colonies showing an amplicon was decreasing until it was no longer possible to amplify any band (samples after 210 of continuous cultivation). PCR amplification on the clones were the amplification was not possible, using primers binding to different regions than the malate synthase or isocitrate lyase genes, showed that the amplicon was smaller than expected, confirming that a fragment of that region was missing. These results suggested a deletion in the malate synthase gene, due to negative selection over time.

Genetic characterization of the promoter regions and mutations

The DNA fragments inserted and the integrity of the glyoxylate shunt enzymes were analyzed by sequencing of single clones of the strains LIB-Δ*fumC* Δ*zwf* and LIB-Δ*fumC*. For this analysis, samples of the clones in which the malate synthase was presumably intact (according to the PCR size and the binding of the malate synthase primer) and samples of the clones in which the amplification was not possible, taken on different time points during continuous cultivation, were used.

The analysis of the fragments of gDNA (potential promoters) inserted upstream the ribosome binding site of the LIB-Δ*fumC* and LIB-*fumC* Δ*zwf* clones, showed five unique sequences in the LIB-Δ*fumC*, and six unique sequences in LIB-*fumC* Δ*zwf*, of which three were common for both strains. All these fragments were identified as part of a coding region, instead of an UTR region. As the gDNA fragments

digested had compatible ends, some fragments were composed by pieces of two or three different genes. Some clones were not carrying any promoter region (Table 3.1).

Table 3.1 Characteristics of DNA fragments inserted in the LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf* strains

LIB-Δ*fumC*		
Days of cultivation	**Sequence**	**Annotation**
120 150	GATCAGGTTGCACTGGGTAACAAACCGTTATGACGCAACAGCGGAGCAGTTTGCGGTTGTCGCCCCCGGAATTT TTCAAAGACCACTGCGGGCGCTAAGCTTCCCCCCAAGGCCAACACCGTTTCTCGAAAACGTTTACCCGTAGTTTG CACCGCAGCTTCATTGTCTAAGCCCACCTCCTCAAAGGCGGCAAAGGCATCGGCACTGAGTACTTCCGCCCATTT ATAGCTGTAGTAACCGGCGGCATAGCCCCCGGCAAAGATGTGGCCAAAGGAGCAGAGGAAAGCATTTTCCGGC AGGGGAGGCAAAATGGTGGTGATTTTCGCCAGGCGATCGGGACTGAAACAAGTAAGGTTCAAACGCCGCTAC GCAGTCTGGTCCCTTTCCTTCTACTAATCCCGGCGATCCCGCCACGGATTTTCGCATTCAGGAATGGATGGAGGA ACGGGGAGAAGATGTGCGCAAAAAAATGGCTCGGCTAATTGCTGATCCTGTTGGGTTGTGGTGCTCTGTTACT GGTCTTTAAAGCCATGTTCTTTGGCGGTGTATACGACACCTGGGCCCCCGGTGGTGGGGACGTACGGGT	part of *prlC* and part of *psbC*
120	GATCGCCGTTGGGGTGAGCCAACATCAGAGCCAGATCCAACCGTGTGTGTGTTGCCGGAGTCCCTCCCCGCCGT TACCCCCGTTACTGAAACCCCCTATATAAATTCCATAGACAACATGAGTATTAATTCTGACATTACCAACCA	part of *ccmM*
120	GATCGCCTTCTCGTACAATGCTTTCAAACTCCGGTAGCGGTCTGAATAGTTCTG	part of *slr0607*
120	GATCAGGAATTCCTCAGCTTACCCAGGGAAGTGACTACCACGGTTATGGTTACCCATCAACGCTATTTTCCCGTG GTGGACAAAGATGGAAGGCTATTGCCCCATTTCATCACCATTGCCAACGGTGATCGCCCGTTGGCCATAGTCTG CCTTGAGTTCATGGTTTTCCGTTTCCACAAAACTAGTGGGAAAAATGCCGTAGGTGGGAAAGCCCTTACTTTGG AGGGTGAGGCAAATTTCCAAAAAGTTTTGGACAATTTCCGTTTCATTTTGCAGTAACAAGAA	part of *glyS* and part of *lim17*
150	GATCAACATTTTGTGCGGGCCAGTTTGGACCATCAACCCCTGGGTAAACTACCTACCGAGGCGGTGGAAACAGT ATTGACAGTGGGACAGGACAAACCAGAAATTTTGGTCTTAACCGCCAGCCGTAGTCTAATCATTCCCTCTTTCCA ACGGCTAGAAGAAGCCCTGGTGTACACCTGTGTCCTTACCAAGGC	part of *slr1107*

LIB-Δ*fumC* Δ*zwf*		
Days of cultivation	**Sequence**	**Annotation**
120	GATCAACATTTTGTGCGGGCCAGTTTGGACCATCAACCCCTGGGTAAACTACCTACCGAGGCGGTGGAAACAGT ATTGACAATGGGACAGGCACGAACCAGAAATTTTGGTCTTAACCGCCAGCCGTAGTCTAATCATTCCCTCTTTCC AACGGCTAGAACAAGCCCTGGATGTACACCTGTGTCCTTACCAAGGC	part of *slr1107*
120	GATCGCCATCCCCTCCCTAGCCCAGCCC	part of *slr1557*
120 150 180	GATCAGGTTGCACTGGGTAACAAACCGTTATGACGCAACAGCGGAGCAGTTTGCGGTTGTCGCCCCCGGAATTT TTCAAAGACCACTGCGGGCGCTAAGCTTCCCCCCAAGGCCAACACCGTTTCTCGAAAACGTTTACCCGTAGTTTG CACCGCAGCTTCATTGTCTAAGCCCACCTCCTCAAAGGCGGCAAAGGCATCGGCACTGAGTACTTCCGCCCATTT ATAGCTGTAGTAACCGGCGGCATAGCCCCCGGCAAAGATGTGGCCAAAGGAGCAGAGGAAAGCATTTTCCGGC AGGGGAGGCAAAATGGTGGTGATTTTCGCCAGGCGATCGGGACTGAAACAAGTAAGGTTCAAACGCCGCTAC GCAGTCTGGTCCCTTTCCTTCTACTAATCCCGGCGATCCCGCCACGGATTTTCGCATTCAGGAATGGATGGAGGA ACGGGGAGAAGATGTGCGCAAAAAAATGGCTCGGCTAATTGCTGATCCTGTTGGGTTGTGGTGCTCTGTTACT GGTCTTTAAAGCCATGTTCTTTGGCGGTGTATACGACACCTGGGCCCCCGGTGGTGGGGACGTACGGGT	part of *prlC* part of *psbC*
150	GATCGCCTTCTCGTACAATGCTTTCAAACTCCGGTAGCGGTCTGAATAGTTCTG	part of *slr0607*
120 180	GATCGCCCTGAAATTAAATTAGGATTTATTCCTATTCAAAATATTATTGAGGGTAGAAAAACAAAAGAAAATAAT GCGACTACTAATTTATGTTTAAATTTAAAATTTATAAGTGATTTTTCACACGGATTAGACGAGATAGCCAAAGCT TTACATGCTTATATTAAGCAGTCTTCAAAAAAATATAAAACTACGGACAGATCCACACCACCACAAATCTCGGCAT CATATTTAAAGGAAACTGTCACCTGGGTGGGTTCACCATAGGTTCTGTTGGCGATCGCCGTTGGATAATCTCAG GGATTTGGAGGATGCGGTTAGCTATGGTTGGTTACTGACCACCACGCAGATCCCTCCGTTCAGCAGTATCTAGA ACAACTAGAGCCGGATTCCTTCACCACTGACAATTCTTTTGCCTCCCCTGCCCTCCTTAGCACCGCAACGGAATCG GAAACTCCCATGGTACATAGTTCCGCCGCCCTTCCCGATCGCCCTTTGACCTCCACCGTTCCCTCACGACGGGGA CGCAGTCCAAGACGTTCCCGAGACGATGTTTTCCCCAGCGCCGACAATTCCAGTGGTTTGGCCGTCACCACCCTA TCTCCGGC	part of s*lr6016* part of *sll0188* and part of *sll0169*
210	GATCCTGCCGGGGCTGGTTGAACTTGGCCGGTAGGGCTTGGGGGGAACGGGCTTGCCAATCAGGAACTGGTTC GTTATTAACTTGGCGGGCAACGGTCAGGGCACTTCGAGGTACCATAGGAGTCCT	part of *nlpD*

The three promoters present in both strains correspond to: a sequence composed by fragments of two genes, the *prlc* gene (oligopeptidase A) and *psbC* gene (photosystem II CP43 protein); a sequence from the *slr1107* gene (hypothetical protein) and a sequence from the *slr0607* (hypothetical protein). The DNA fragment composed by parts of *prlc* and *psbC* genes was the most frequent fragment, found in different clones and on different time points of continuous cultivation. The gene product of *prlC* gene (oligopeptidase A) is involved in the hydrolysis of oligopeptides and metal ion binding, and *psbC* gene codes for photosystem II CP43 protein, a transmembrane protein involved in light harvesting.

The promoters only present in the clones of LIB-Δ*fumC* correspond to a fragment of the *ccmM* gene (carbon dioxide concentrating mechanism protein CcmM) and a sequence composed by fragments of two genes, the *glyS* gene (glycyl-tRNA synthetase beta chain) and the lim17 gene (LIM17 protein, with glycopeptide alpha-N-acetylgalactosaminidase activity). The promoters only present in the clones of LIB-Δ*fumC* Δ*zwf* library correspond to a fragment of the *slr1557* gene (coding for a hypothetical protein), a fragment of the *nlpD* gene (lipoprotein NlpD) and a sequence composed by fragments of the *slr6016*, *sll0188* and *sll0169* genes, all coding for hypothetical proteins.

It was expected that promoters of genes related with the circadian rhythm were the best candidates to activate the expression of the glyoxylate shunt enzymes during the night, but it is possible that promoters of genes involved in photosynthesis, with differential expression during light and darkness periods, might activate the glyoxylate shunt enzymes in a timely way. The results also suggest that some regions of annotated genes or genes that are not yet characterized could act as regulatory elements of other open reading frames (ORF) or coding sequences nearby.

The expression of the glyoxylate shunt in the strain LIB-Δ*fumC* Δ*zwf*, according to our simulations, would only be maintained over time if it happens solely during the night. Due to the inability of the LIB-Δ*fumC* strain to use the TCA cycle during the night, the nocturnal expression of the glyoxylate enzymes would not confer any fitness advantage to the cells. Therefore, it was expected that the diurnal expression of the enzymes would affect the growth and would cause a negative selection pressure in both strains.

The analysis of the sequences of the genes coding for the malate synthase and isocitrate lyase in the LIB-Δ*fumC* and the LIB-Δ*fumC* Δ*zwf* strains, showed that even in the case when the region seemed to be intact (according to the size of the PCR amplification), there was a mutation (GCA = alanine changed to GGC = glycine) and a deletion of 57 bp before the start of the malate synthase. More importantly, the start codon of malate synthase was missing, resulting in no translation of the enzyme (Table 3.2). This was the most frequent change found in both strains.

The clones that showed smaller size in the PCR analysis, had a complete deletion of the RBS and malate synthase sequence, and a partial deletion of the isocitrate lyase sequence (different lengths on different clones) in comparison with the expected size (Figure 3.6).

Table 3.2 Characteristics of the mutations found in the shunt genes of LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf*

LIB-Δ*fumC*			
Days of cultivation*	**Found sequence**	**Mutation/deletion before MS**	**Deletion MS-ICL**
120 (4) 150 (2)	RBS-UTR of MS = 44 bp MS-ICL = 3395 bp	GGC = glycine by GCA = alanine Region before MS = 57 bp	ATG of MS
120	RBS-UTR of MS = 0 bp ICL = 462 bp		MS and most of ICL
120	RBS-UTR of MS = 0 bp ICL = 231 bp		MS and most of ICL
120	RBS-UTR of MS = 0 bp ICL = 80 bp		MS and most of ICL
LIB-Δ*fumC* Δ*zwf*			
Days of cultivation*	**Found sequence**	**Mutation/deletion before MS**	**Deletion MS-ICL**
120 (4) 150 (3) 180 (3)	RBS-UTR of MS = 44 bp MS-ICL = 3395 bp	GGC = glycine by GCA = alanine Region before MS = 57 bp	ATG of MS
120	RBS-UTR of MS = 0 bp ICL = 1112 bp		MS and part of ICL
120 (2) 210 (2)	RBS-UTR of MS = 0 bp ICL = 568 bp		MS and part of ICL

* The samples were taken on different time points of continuous cultivation: 120, 150, 180, 210 days (repetitions between brackets). The table shows the sequence found in comparison with the original sequence (RBS-UTR of MS = 101 bp; MS-ICL = 3398 bp) and the deletion/mutation to which it corresponds. MS: malate synthase gene *aceB*. ICL: isocitrate lyase gene *aceA.*

Figure 3.6 Schematic representation of the deletions (grey area) found in LIB-Δ*fumC* Δ*zwf* (A, B, C, D) and LIB-Δ*fumC* (A, B, D). (A) Intact cassette. (B) Deletion of the start codon (atg) of the MS gene (*aceB*). (C) Deletion of the RBS, the MS gene (*aceB*) and part of the ICL gene (*aceA*). (D) Deletion of the promoter, the RBS, the MS gene(*aceB*) and part of the ICL gene (*aceA*).

An increase in the night-time production of fumarate in the LIB-Δ*fumC* Δ*zwf* strain, is an indicator of the expression of the glyoxylate shunt enzymes during the night. However, these results show that on long term, there was not a possible way to express the glyoxylate shunt enzyme in the specific

timing required. The untimely activation, affected the fitness of the cell, resulting in a negative selection of the expression of the glyoxylate shunt enzymes, by transcriptional (deletion of the genes) and translational regulation (deletion of the start codon).

The mutation of one or few nucleotides, or the loss of a fragment of variable size in engineered strains, leading to the loss of gene expression or the protein functionality is described as genetic instability[128]. Genetic instability has been observed in different cyanobacterial production strains. Suppressor mutations such as point mutations, insertion and deletions, were often found in the coding region of the introduced gene(s) coding for a heterologous pathway, leading to the impairment of the functionality of the enzyme(s) and the partial or total loss of product formation[131]. This is an important issue in large industrial cultures and long cultivations. Longer periods of cultivation increase the chances of selecting spontaneous mutations that cause a decrease in the synthesis of the product. Those mutations reduce the burden of product formation and consequently, increase the cell fitness, as it was observed in this case[147].

Even though we could not find a promoter triggering the expression of the glyoxylate shunt enzymes with the required timing to reconnect the TCA cycle and increase the night production of fumarate, this work provided an approach to find native elements such as promoters able to activate gene expression under specific conditions.

Conclusion

In this study we have constructed a fully segregated integrative genomic library in the producing strain *Synechocystis* Δ*fumC* Δ*zwf* (double knockout for the *fumC* and *zwf* genes)*.* We have used long cultivation to screen via natural selection for random genomic DNA fragments in *Synechocystis*, that could act as promoters capable of activating the expression of the malate synthase and isocitrate lyase enzymes exclusively during the night. This, in an attempt to reconnect the disrupted TCA cycle and aiming to increase the fumarate production during the night.

The night production of fumarate was monitored over time as well as the sequence of the genomic fragments inserted and the integrity of the sequence coding for the malate synthase and the isocitrate lyase. Despite the increase of the night-time fumarate production, indicating the expression of the glyoxylate enzymes during that period, it was not possible to obtain the desired timing and/or level of activation necessary for a stable strain. Over time the cells showed different deletions affecting the transcription and/or translation of the glyoxylate shunt enzymes.

These results suggest that the timing/strength of the expression of the shunt enzymes is not adequate and eventually reduces the fitness of the cells. That could be because *Synechocystis* does not have promoters that display the desired transcription profile in timing and intensity of expression. In this case, it would be worthwhile to test a library of exogenous promoters. Another reason could be that the lifetime of the shunt proteins used, is not compatible with the desired expression profile, i.e. the enzymes could possibly be too stable due to slow degradation. In that case it would be worthwhile to test different glyoxylate shunt enzymes from other organisms or increase the degradation rate of these enzymes.

Overall, this study validates our approach to find native genomic elements needed for genetic engineering applications, such us promoters acting under specific conditions, using the construction of a genomic DNA library on the chromosome of the model cyanobacteria *Synechocystis*. This was confirmed by the expression of foreign glyoxylate shunt enzymes in *Synechocystis*, although more studies are required to achieve the desired timing expression needed to generate a genetically stable strain.

Materials and Methods

General cultivation conditions. *Escherichia coli* strain DH5α was used as the host for recombinant plasmid construction, amplification and cloning. It was grown at 37 °C in Luria-Bertani Broth (LB) in an incubator with a shaking speed of 200 rpm or on LB agar. The concentrations of antibiotics used were 50 µg mL^{-1} for kanamycin and spectinomycin.

Synechocystis sp. PCC 6803 Δ*fumC* Δ*zwf* (double knockout for the *fumC* and *zwf* genes) and Δ*fumC* (knockout for *fumC* gene)[129] were cultivated at 30 °C in a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific) under constant moderate red light illumination (~ 30 µE · m^{-2} · s^{-1}), in liquid BG-11 medium supplemented with 10 mM TES-KOH (pH = 8), or on BG-11 agar plates supplemented with 10 mM TES-KOH and 0.3% (w/v) sodium thiosulfate. For the construction of the mutants, kanamycin, spectinomycin and/or nickel sulphate were added to the medium with a final concentration of 50 µg mL^{-1}, 20 µg mL^{-1} and 15 µM, respectively. The concentration of biomass in the cultures was measured by optical density at 730 nm (OD$_{730}$) in a spectrophotometer (Lightwave II, Biochrom).

Plasmid and strain construction. All plasmids, strains and primers are listed in Table 3.3. The plasmid pFLXN-MAF contains the upstream and downstream homologous regions (~ 1 kb each) of the *slr0168* gene, and the MAZF (Spec^R) selection cassette in between. The plasmid pFLXN-RBS-YFP contains the upstream and downstream homologous regions of the *slr0168* gene and the RBS-YFP cassette in between. This cassette is composed by the ribosome binding site from the P*cpcBA* promoter (excluding the promoter region), eYFP, terminator BB0014 and Kan^R. The glyoxylate shunt genes of *Chlorogloeopsis fritschii* PCC 9212 were obtained from the Bryant Lab in a plasmid called pAQ1Ex_cpc_MS-ICL[145]. To construct the plasmid pFLXN-RBS-MS-ICL, the plasmid pFLXN-RBS-YFP was digested with NdeI and BamHI to replace the *eyfp* gene by the *aceB* and *aceA* genes (ML-ICL cassette) taken from pAQ1Ex_cpc_MS-ICL. The plasmid pFLXN-LIB-MS-ICL was constructed introducing the genomic DNA fragments of the library in a BglII site located before the RBS. All the fragments amplified in this study were confirmed by Sanger sequencing (Macrogen Europe, The Netherlands).

Synechocystis mutants MAZF-Δ*fumC* and MAZF-Δ*fumC* Δ*zwf*, were created via natural transformation of the pFLXN-MAF plasmid to the *Synechocystis* strains Δ*fumC* and Δ*fumC* Δ*zwf*, respectively. Full segregation was verified by PCR using the genomic DNA of each mutant as the template for 35 cycles. To obtain a fully segregated genomic DNA promoter library integrated in the chromosomes of *Synechocystis*, we used a liquid transformation to introduce the promoter library to MAZF-Δ*fumC* and MAZF-Δ*fumC* Δ*zwf*, generating the LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf* strains. First, fresh MAZF-Δ*fumC* and MAZF-Δ*fumC* Δ*zwf* cells were collected from 20 mL liquid culture (OD730 ≈ 1). After washing them twice with fresh BG11 medium through centrifugation (3900 rpm, 10 min), cells were

concentrated to a total volume of 200 µL. The cells were mixed with pFLXN-LIB-MS-ICL plasmids to a final DNA concentration of 30 µg mL^{-1}, and then incubated for 5 hours at 30 °C with moderate light intensity (white light, 50 µmol photons $\cdot$ m$^{-2} \cdot$ s^{-1}). After that, the mixture was inoculated in 20 mL of BG-11 (without antibiotic) and incubated in a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific) at 30 °C under constant moderate red-light illumination (~30 µmol photons $\cdot$ m^{-2} $\cdot$ s^{-1}). After 24 hours of further incubation, kanamycin at a final concentration of 50 µg mL^{-1} was added to the cultures. After another 4 days of incubation, nickel sulphate at a final concentration of 15 µM was added to the cultures. When the cultures reached OD$_{730}$ ≈ 1-1.5, they were inoculated in a bioreactor for continuous cultivation.

Genomic DNA promoter library construction in *E. coli*. Genomic DNA of *Synechocystis* was extracted from 1 mL of a culture of OD$_{730}$ about 1. The pellet of the cells was suspended in 200 µL TE buffer and mixed with 200 µL of phenol/chloroform/isoamylalcohol (25:24:1). After the addition of 100 mg of glass beads (0.1 mm diameter), the suspension was vortexed for 5 min and then centrifuged at 14.000 rpm for 5 min. The supernatant was collected and mixed with 200 µL of chlorophorm/isoamylalcohol (24:1). After that, the mixture was centrifuged at 12.000 rpm for 5 min. The supernatant was transferred and precipitated with 1/5 volume of 5 M NaCl and 2 volumes of absolute ethanol and kept at -20 °C for at least 30 min. After centrifugation at 12.000 rpm for 10 min, the DNA pellet was washed with 500 µL 70% (V/V) ethanol and dried at room temperature. Then, it was dissolved in 20 µL MQ water. The RNA in the DNA sample was removed incubating it with RNAse (final concentration 100 µg mL^{-1}) at 37 ˚C for 30 min. The genomic DNA was then precipitated with cold ethanol, dried and re-dissolved in MQ water.

Later, 10 µg of the genomic DNA was partially digested with 5 µL Sau3AI (New England) for 5 hours at 37 °C, generating DNA fragments with a size from 200 bp to 1000 bp. These fragments were purified using MSB spin PCRapace kit (Invitek molecular). Simultaneously, 4 µg of the vector pFLXN-RBS-MS-ICL were digested with 2.5 µL of BglII (Thermo Scientific) at 37 °C for 4 hours. After 1 hour of digestion, alkaline phosphatase was added to the mix for 3 hours at 37 °C to remove the phosphate groups at each end of the vector and prevent vector self-ligation. The vector was purified by gel purification using Gen JET gel extraction kit (Thermo Scientific). The genomic DNA fragments and the vector were ligated using T4 ligase (Thermo Scientific) and the ligation was transformed into DH5 alpha competent *E. coli*. A large-scale transformation was done in *E. coli*, plating the resulting colonies in square petri dishes, until reaching an approximate number of 200.000 colonies. All of them were collected in LB and incubated at 37 ˚C overnight. The plasmid DNA was extracted by maxiprep (Quiagen) and it was later transformed into *Synechocystis*.

Bioreactor cultivation. The LIB-Δ*fumC* and LIB-Δ*fumC* Δ*zwf* strains, were inoculated in a photobioreactor (FMT 151, Photon System Instruments) under turbidostat regime, in which cultures were diluted by 5% (v/v) when OD$_{720}$ (as measured in the bench-top spectrophotometer) was above 1. The bioreactor was set with a light regime of 12 hours of light and 12 hours of darkness. During the light period, a sinusoidal function controlled the light intensity with a maximum intensity of 100 µE $\cdot$ m$^{-2} \cdot$ s^{-1}. Light was provided by a red LED panel (636 nm).

BG-11 supplemented with 10 mM NaHCO$_3$ was used for *Synechocystis* cultivation at 30 °C. Cultures were sparged at a rate of 150 mL · min^{-1} with a mix of N$_2$ and CO$_2$. The actual ratio of N$_2$ and CO$_2$ was set by a gas mixer (GMS 150, PSI) controlled by the bioreactor software with a custom script in order to maintained culture about pH 8. Cultures of the library strains (approximately OD$_{730}$ = 1) were inoculated in the bioreactor to a final volume of 960 mL. Samples were taking on mornings and evenings, recording ID and collecting supernatant for HPLC analysis.

Fumarate quantification. To determine the extracellular fumarate production, samples of 500 µL aliquot were taking every 12 hours on mornings and evenings, recording OD and collecting supernatant for HPLC analysis. Cells were removed through centrifugation for 5 min at 15.000 rpm. The supernatant was filtered before analysis (Sartorius Stedin Biotech, minisart SRP 4, 0.22 µm) and the fumarate concentration was measured by HPLC-UV/VIS (LC-20AT, Prominence, Shimadzu), using ion exclusion Rezex ROA-Organic Acid column (250 x 4.6 mm; Phenomenex) and UV detector (SPD-20A, Prominence, Shimadzu) at 210 nm wavelength. A volume of 10 µL the of samples was injected through an autosampler (SIL-20AC, Prominence, Shimadzu), using 5 mM H$_2$SO$_4$ as eluent at a flow rate of 0.15 mL min^{-1} and a column temperature of 45 °C. To determinate the night-time production of fumarate, the value obtained in the night sample, was subtracted from the value obtained in the next morning sample.

Growth rate measurements. Growth rates were determined for each dilution event of the turbidostat data in each photobioreactor. The log transformed equation for exponential growth was fitted to the recorded OD$_{720}$ per time (in hours), using linear regression, where the slope of the linear model corresponds to the growth rate. The yield on fumarate was calculated by diving the measured concentration of fumarate by the externally measured OD$_{730}$.

Author contributions PCB, WD and FBS designed the experiments; PCB and YE performed the experiments; PCB wrote the manuscript; FBS reviewed it.

Table 3.3 Plasmids, strains and primers used in this study

Plasmids, strains and primers*	Relevant characteristics	Reference
pAQ1Ex_cpc_MS-ICL	Vector containing the MS-ICL cassette	145
pFLXN-MAZF	pFLXN derivate, Amp^{R}Sp^R, containing *slr0168* gene upstream homologous region, MAZF selection cassette, and downstream homologous region	This study
pFLXN-RBS-YFP	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, RBS-YFP cassette, and downstream homologous region	This study
pFLXN-RBS-MS-ICL	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, RBS-MS-IC cassette, and downstream homologous region	This study
pFLXN-LIB-MS-ICL	pFLXN derivate, Amp^{R}Kan^R, containing *slr0168* gene upstream homologous region, promoter library-MS-ICL cassette, and downstream homologous region	This study
E. coli DH5α	Cloning host	MCLAB
Δ*fumC*	*Synechocystis* sp. PCC 6803 *fumC* gene knockout mutant	129
Δ*fumC* Δ*zwf*	*Synechocystis* sp. PCC 6803 *fumC* and *zwf* genes double knockout mutant	129
MAZF-Δ*fumC*	*Synechocystis* sp. PCC 6803 Δ*fumC* with MAZF cassette integrated at *slr0168* locus, Sp^R	This study
MAZF-Δ*fumC* Δ*zwf*	*Synechocystis* sp. PCC 6803 Δ*fumC* Δ*zwf* with MAZF cassette integrated at *slr0168* locus, Sp^R	This study
LIB-Δ*fumC*	*Synechocystis* sp. PCC 6803 Δ*fumC* with promoter library-MS-ICL cassette integrated at *slr0168* locus, Kan^R	This study
LIB-Δ*fumC* Δ*zwf*	*Synechocystis* sp. PCC 6803 Δ*fumC* Δ*zwf* with promoter library-MS-ICL cassette integrated at *slr0168* locus, Kan^R	This study
H2 0168 down Rev	cactctgcactgtgtctgtgc	This study
MS colony Rev	actctagcgcctgtggagtc	This study
BglII rev	atgacgaccttcgatatggc	This study
MS 1 forw	caatgccctaaactctggggc	This study
MS 2 forw	catgatggcacctgggttg	This study
ICL 1 forw	catctgcccacgttcgca	This study
Kan up reverse seq	tcccgttgaatatggctc	This study

*primer sequences are given from 5'→3'

Construction of a family of biosensors to detect fumarate applicable across different clades of microorganisms

Patricia Caicedo Burbano, Age Tjalma, Wei Du, Filipe Branco dos Santos

Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences,
University of Amsterdam, Netherlands

Abstract: Bacteria have evolved to sense and respond to changes in the extracellular environment using two-component systems (TCS). These systems typically consist of a sensor and a response regulator. The sensor histidine kinase (SHK) is located in the membrane and gets activated upon sensing particular signals (e.g. a specific compound). The response regulator gets activated by an active SHK and is able to initiate a biological response after binding to DNA at specific loci. In this study we constructed variants of a fumarate biosensor, based on the DcuSZ/OmpR chimeric TCS developed by Ganesh *et al*. in *E. coli*. Our detection system was made fusing two native *E. coli* SHKs, the fumarate sensor DcuS and the osmolarity sensor EnvZ (DcuSZ). The two-component system was coupled to the expression of a fluorescent protein (eYFP) in response to extracellular fumarate. We have created nine different DcuSZ, which differ in the length of the linker region between the DcuS and EnvZ functional domains. Three, DcuSZ 1, 6 and 9, were found to be functional biosensors of fumarate in *E. coli* with a combined signal dynamic range of 0.1-20 mM. The DcuSZ 6 variant was singled out to be the best candidate to be potentially used in *Synechocystis*.

Keywords: DcuSZ two component system, fumarate biosensor, *E. coli*, *Synechocystis*

Introduction

Cyanobacteria can be engineered to produce carbon compounds relevant for industry. One example of it is fumarate, an intermediate in the TCA cycle, with diverse applications in food industry[132], in plastic production[133] and as a medicine for psoriasis[134]. The use of cyanobacterial cell-factories to produce fumarate requires the constant monitoring of the concentration of this product. Measurements must be accurate and specific for fumarate, and the measuring process should ideally be quick and easy. Currently, the most common methods used to measure the concentration of fumarate in liquid samples, are based on high-performance liquid chromatography (HPLC) or specific enzymatic assays. Both of them present the drawback of being costly and/or labor intensive. Hence, there is the need for an alternative fumarate sensing system to perform high throughput screening of fumarate producing organisms.

Biological tools, such as biosensors, are promising candidates to be used as detection systems with a wide range of detection[134,148]. Biosensors are analytical devices that integrate a biological recognition element from a microorganism (which typically responds to the target compound), with a physical transducer (which converts the biological response to a detectable signal). This will then be used to generate a measurable variable proportional to the concentration of the target compound[149]. Since the first biosensor was developed in 1962[150], they have been intensively studied and implemented in various applications such as environmental monitoring and public health[151,152]. The intrinsic ability of bacteria to recognize fumarate, can be used to design and construct a fumarate biosensor.

Bacteria have evolved to sense and respond to extracellular environment changes using two-component systems (TCS). These systems are highly specific in response to particular compounds and they are very precise in their measurement. This mechanism is ubiquitous among bacteria and it plays essential roles in signaling events such as adaptation, pathogenesis and cell-cell communication[153]. A two-component system consists of a sensor histidine kinase (SHK) and a response regulator (RR). The sensor histidine kinase is located in the membrane (HK is a transmembrane protein). It gets activated at the time of sensing particular signals such as a specific compound, which leads to the activation of the kinase activity and autophosphorylation of a conserved histidine residue. The response regulator gets activated by an active sensor histidine kinase, being able to bind to DNA at specific loci, to turn on (or off) gene transcription, initiating a biological response[154].

Ganesh and co-workers[134] constructed a fumarate biosensor for *E. coli* using two native *E. coli* sensor histidine kinases, DcuS and EnvZ. DcuS belongs to the C_4-dicarboxylate sensing two-component system apparatus, and it is sensitive and selective towards fumarate[155]. The Env/OmpR system, regulates the expression of OmpC and OmpF (outer membrane porines) in response to changes in the osmolarity[156]. Ganesh and co-workers, created a chimeric histidine kinase, integrating the periplasmic sensor domain, transmembrane domains, and cytoplasmic HAMP domain of DcuS (fumarate sensor) with the cytoplasmatic histidine kinase domain of EnvZ (osmolarity sensor) (Figure 4.1).
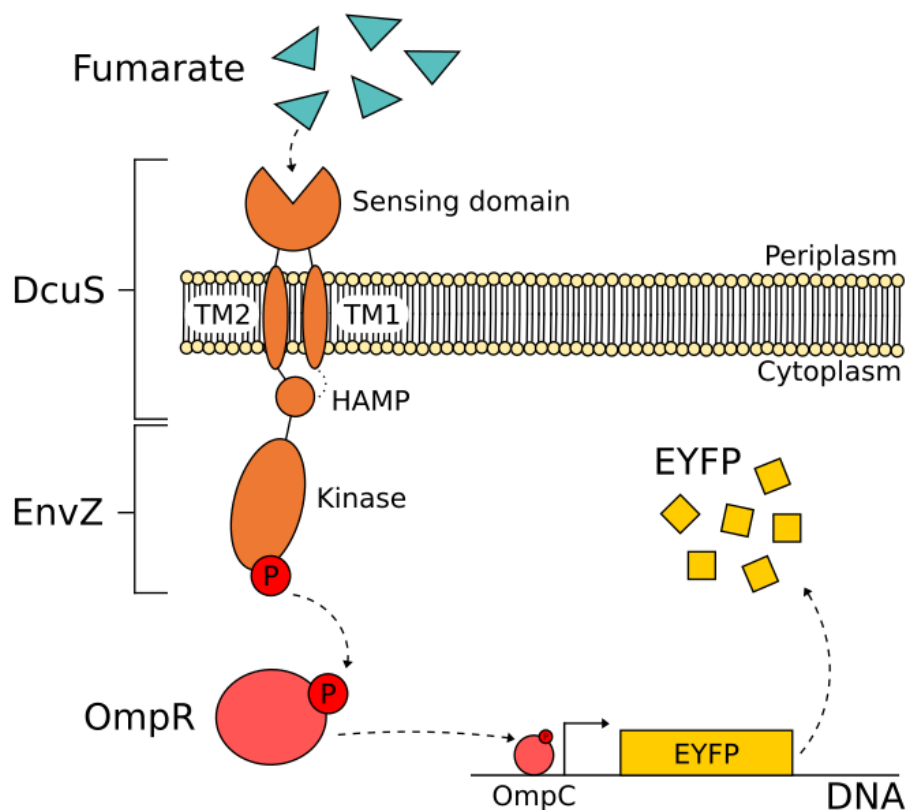
Figure 4.1 Scheme of the biosensor design. The chimeric DcuSZ/OmpR two-component system was created by fusing the DcuS fumarate sensing, transmembrane and HAMP domains to the EnvZ kinase. The system is coupled to the expression of a fluorescent protein (eYFP) under the control of the *ompC* promoter to report the presence of extracellular fumarate.

Different versions of this chimeric histidine kinase DcuS/EnvZ (from now on referred to as DcuSZ) can be created, using different lengths of the linker region between the functional domains of DcuS and EnvZ. The linker region is encoded by the *dcuS* gene and forms an α-helix. A complete turn in an α-helix corresponds approximately to 3.6 amino acids. Therefore, a variation in the linker length in intervals of 5 amino acids, can shorten or extend the α-helix by approximately 1.4 turns per interval changed. This can be used to titrate the different orientations of the EnvZ kinase relative to the membrane, through the membrane bound DcuS domains.

The linker length and kinase orientation should play a central role in the functionality of the system. The creation of *dcuSZ* genes with different lengths could lead to the finding of a variant with optimal orientation and functionality. By changing more than one turn per interval it is possible to scan a considerable range of distances between the DcuS and EnvZ functional domains. These changes may influence the kinetics of binding and activation of the two-component system engineered, possibly resulting in variants with different detection limits and different signal dynamic ranges for the detection of fumarate. In this study, the functionality of 9 different variants of the *dcuSZ* gene to screen for fumarate production were tested in *E. coli*. The variant DcuSZ 6, while showing the best results in *E. coli*, was chosen for potential use in *Synechocystis* sp. PCC 6803.

Results and Discussion

Construction of DcuS/EnvZ (DcuSZ) variants

A fumarate biosensor was constructed using specific parts of the DcuS/EnvZ two component systems. In this system the periplasmic receptor DcuS senses extracellular fumarate, which phosphorylates the histidine kinase domain EnvZ and subsequently phosphorylates the response regulator OmpR. That leads to the activation of the *ompC* promoter, resulting in the expression of the *eyfp* gene. In this work, *dcuSZ* was expressed under the control of the inducible promoter p*LtetO-1*[157] and *ompR* was expressed under the control of the constitutive *Bba J23104* promoter (http://parts.igem.org/Part:BBa_J23104[47]).

Nine *dcuSZ* gene variants were constructed, maintaining the first 1020 base pairs of *dcuS* and the 690 base pairs of e*nvZ*, while changing the length of the linker between them, resulting in different orientation of the EnvZ domain. Modifying the original length of the linker (60 bp), extending or shortening it in intervals of 15 base pairs (*i.e.* 5 amino acids), can change the orientation of the linker α-helix, allowing the characterization of the role of the orientation in the functionality of the biosensor (Figure 4.2).
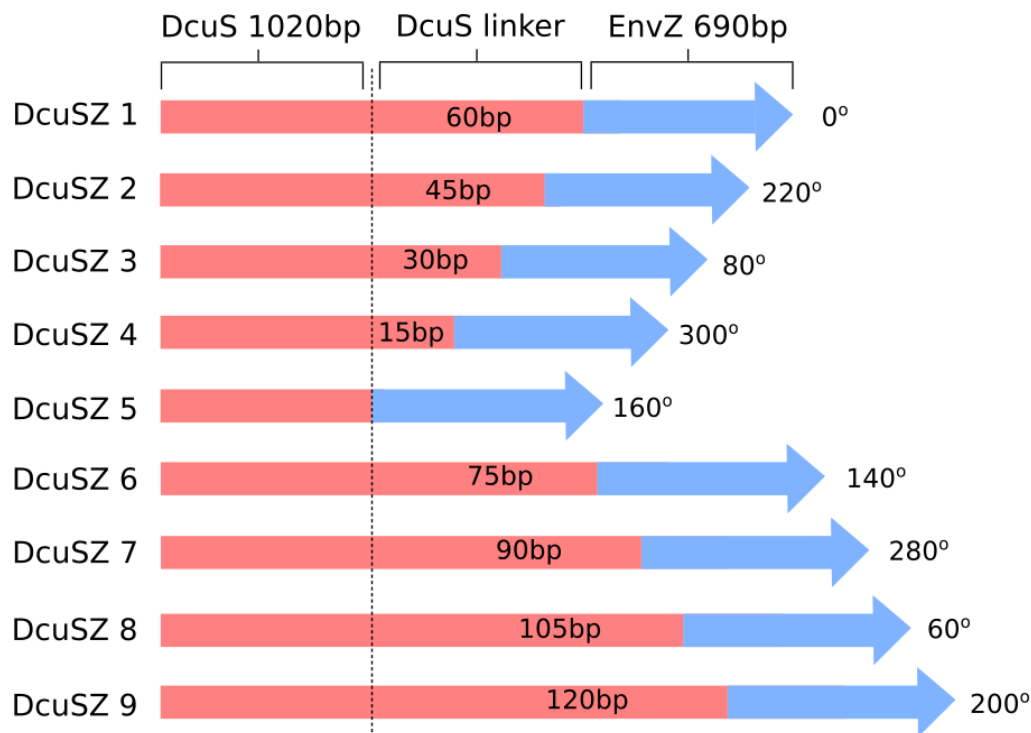

Figure 4.2 Scheme of the 9 DcuSZ variants. The first 1020 base pairs of DcuS and the 690 base pairs of EnvZ remain the same. Linkers are shortened or extended at the union between DcuS and EnvZ. Orientations of the EnvZ kinase domains are displayed relative to the orientation in DcuSZ 1, as it was constructed by Ganesh *et al.*[134]

Screening the functionality of the DucSZ variants in *E. coli*

The functionality of the resulting *E. coli* strains 1-9, containing corresponding DcuSZ 1-9 signaling systems, was characterized within a detection range of 0 to 10 mM of fumarate. The expression of eYFP was measured after 16 and 22 hours of the signaling system induction, by adding anhydrotetracycline (aTc) upon addition of fumarate. The results obtained with this screening method suggested that the response in fluorescence/OD$_{600}$ is robust over time. And more importantly, it indicated that strains 1, 6 and to some extent 9 behave as suitable biosensors for fumarate displaying different sensitivity (Figure 4.3).
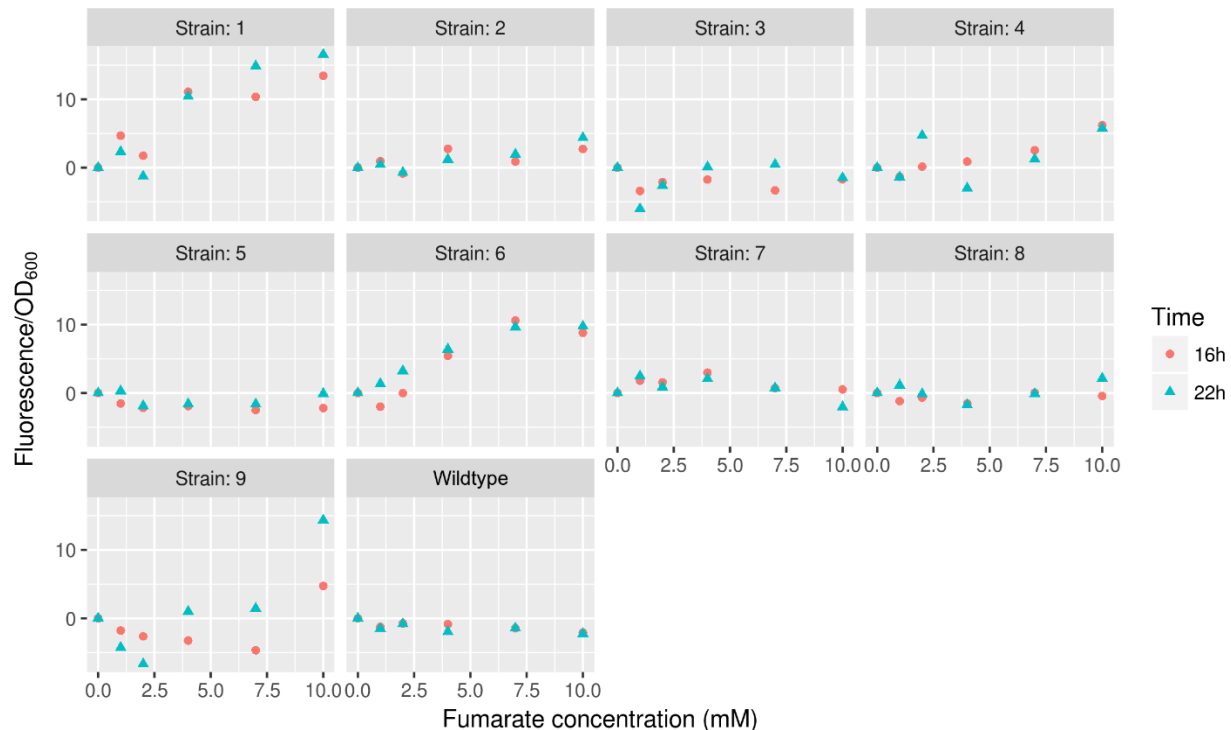

Figure 4.3 Screening of the functionality of constructed strains. The y-axis shows Fluorescence/OD$_{600}$. The x-axis shows fumarate concentration ranging from 0 mM to 10 mM. Each strain is displayed in a separate plot, showing two measured time points, 16 h (o) and 22 h (Δ).

Screening of DcuSZ variants 1, 6, 9

Further characterization was performed on the strains 1, 6 and 9, using a detection range of 0 to 20 mM of fumarate. The detection limits and saturation levels of each sensing strain can be determined by fitting a logistic curve to the data. Examples of the logistic nature of similar sensing systems have been shown before by Pollard *et al.*[158] It was expected that by varying the length of the linker and with that, modifying the orientation of the kinase, the effectiveness of signal transduction in the chimeric DcuSZ protein would change. Even though the expectation was that only the range of signal detection would differ as a consequence, the results showed that the dynamics of the fluorescence output vary extensively as well (Figure 4.4).
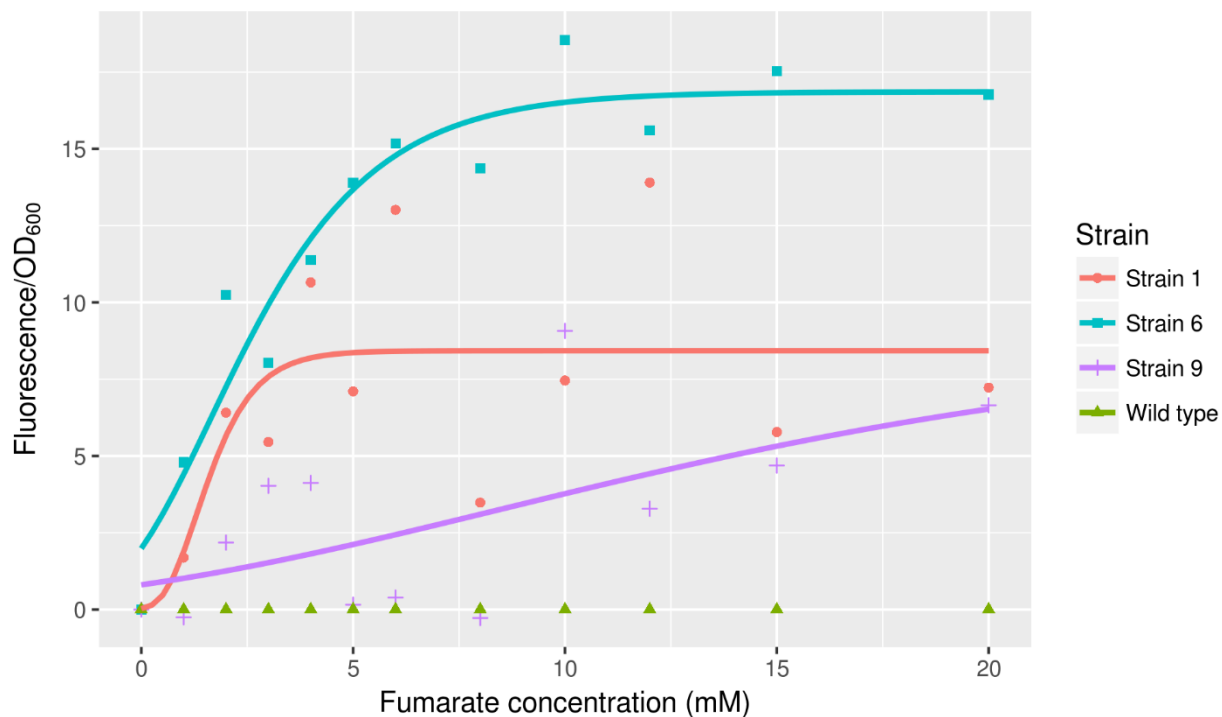
Figure 4.4 Screening of the functionality of the strains 1, 6, 9. The y-axis shows Fluorescence/OD_{600}. The x-axis shows fumarate concentration with a range of 0 mM to 20 mM. The fitted curve is a Gompertz function $y = ae^{(-be^{(-cx)})}$.

Strains 1, 6 and 9 have a relative kinase orientation of 0°, 140° and 200° respectively and are functional biosensors. DcuSZ 1 and 6 are similar in length, but DcuSZ 9 is 15 and 20 amino acids longer than DcuSZ 1 and 6, respectively. The results suggest different detection limits in each strain. The strain 1 can be used to quantify fumarate concentrations up to 3 mM, the strain 6 up to 6 mM and strain 9 up to 20 mM. Of the three strains characterized, strain 9 can measure up to the highest fumarate concentration, however the sensitivity is lower in comparison with strains 1 and 6. Strain 1 shows a high spread in the fluorescence/OD_{600}, which results in more uncertainty about the level of saturation of this biosensor. Strain 6 has the best logistic fit and its curve is the most trustworthy. This variant was therefore selected as the best candidate to be integrated into *Synechocystis* sp. PCC 6803.

Applicability of the fumarate biosensors

This family of biosensors can be used to screen for fumarate production in *Synechocystis* cell factories in different ways. The screening of a big amount of producing cells, for example a library, can be performed by mixing the *E. coli* biosensor with *Synechocystis* cell factories in an emulsion, as previously described[159]. In this emulsion, it is not necessary that *E. coli* cells are able to grow, only that they are able to make fluorescent protein *de novo*. The availability of a preferred carbon compound such as glucose, will most likely prevent the uptake of the fumarate produced[160].

The screening of a smaller range of producing cells, can be performed in a standard 96-well plate, using the supernatant of the fumarate producing *Synechocystis* cells to resuspend the same amount

of biosensor *E. coli* cells for each culture. This was tested using the biosensor strain DcuSZ 6, and the *Synechocystis* strain Δ*fumc*, a knockout of the gene coding for the enzyme fumarase, leading to an accumulation of fumarate[129]. The results were compared with the measurement using HPLC. After resuspension in the supernatant of the Δ*fumc* strain, the biosensor produced a significant signal for 22-60 hours (Figure 4.5). The mean fluorescence/OD_{600} measured over this time period, was approximately 3 expressed in arbitrary units (AU). However, this signal was too low to determine the fumarate concentration, as the lowest calibration measurement with strain DcuSZ 6, was made at 1 mM fumarate, which correspond to a fluorescence/OD_{600} of approximately 5 AU.
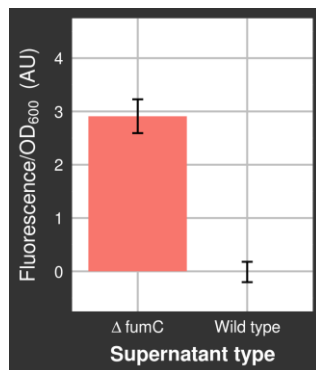

Figure 4.5 Mean fluorescence/OD_{600} signal of 22-60 hours measurements of the supernatant of the Δ*fumc* strain in comparison with the supernatant of the wild type strain. The error bars represent a 95% confidence interval.

The fumarate concentration in the Δ*fumc* supernatant measured by HPLC was 0.087 mM, which confirms that the concentration was too low to be quantified by the biosensor strain DcuSZ 6, employing the current calibration. This result nonetheless suggests that lower concentrations can also be quantified with more thorough calibration, as the sensor still produces a significant and stable signal.

Integration of the variant DcuSZ 6 into Synechocystis

The engineered system can also be integrated in *Synechocystis* directly. This has potentially several advantages. For instance, it prevents the need to cultivate *Synechocystis* and *E. coli* together, which can be difficult; *Synechocystis* does not consume fumarate under aerobic conditions as *E. coli* may do; and it would allow the development of cell factories that can detect their own production. This could be applied to create more complex regulatory circuits, where cells are able to regulate their own fumarate production.

The system was introduced in *Synechocystis,* using two plasmids. The p*LtetO-1-dcuSZ6* and P*ompC-eyfp* cassettes were expressed in a replicative plasmid while *Bba J23104-ompR* cassette was integrated into the chromosomal neutral site NS1, replacing the *psbA1* gene. The full segregation of the *ompR* cassette in all the chromosomes, and the presence of the *dcuSZ6-eyfp* cassette were confirmed by PCR (Figure 4.6).
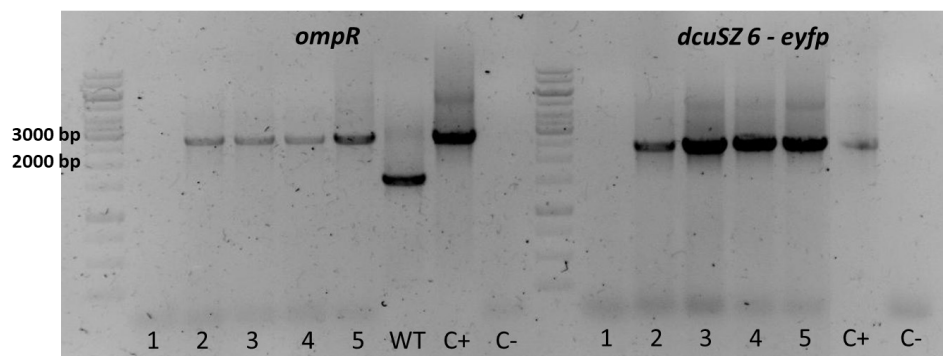
Figure 4.6 Agarose gel showing PCR fragments indicating the presence of the *ompR* and *dcuSZ6-eyfp* cassettes in four clones (2-5) of *Synechocystis*.

This biosensor strain can be potentially used in the future to measure fumarate concentrations without the need of resorting to co-cultures or to the cultivation of *E. coli* on supernatants, facilitating the quantification process during biotechnological production. Although this has not been implemented here, all the elements of the system are in place to use the biosensor in *Synechocystis*.

Conclusion

In this study, a family of biosensors to screen for fumarate production in *Synechocystis* cell factories was developed. We have created and tested 9 variants of DcuSZ fumarate biosensor in *E. coli*. The effect of the length of the linker between the DcuS and EnvZ domains and the kinase orientation in the functionality of the biosensors was evaluated. The results showed a considerable difference in fluorescence response as a result of DcuSZ linker length and kinase orientation. However, we could not identify a general trend in functionality.

Three variants of the biosensor family, DcuSZ 1, 6 and 9, were capable of quantifying fumarate concentration from (at least) 1 up to 20 mM showing diverse dynamic ranges in *E. coli*. Using the strain DcuSZ 6, it was possible to qualitatively detect extracellular fumarate produced by the *Synechocystis* strain Δ*fumc*. The results suggest that this biosensor could be used to quantify fumarate concentrations below 0.1 mM, after performing a more thorough calibration. The variant DcuSZ 6 was integrated into *Synechocystis*, but further analysis is necessary to test the functionality and sensitivity of this fumarate detection system in *Synechocystis* cell factories.

Even though previous work showed that the DcuSZ/OmpR system with the linker length and kinase orientation of variant 1 is highly specific to fumarate[134], the specificity to fumarate of the other variants should be tested by comparing our results to the response to other C4-dicarboxylates (malate, aspartate and succinate).

Materials and Methods

General cultivation conditions. *Escherichia coli* strain DH5α was used as a host to perform molecular cloning. It was grown in liquid Luria-Bertani Broth (LB) at 37 ℃ in a shaking incubator at 200 rpm, or on LB plates containing 1.5% (w/v) agar. The concentrations of antibiotics used in the medium for propagation purposes, were 100 µg mL^{-1} for ampicillin, 50 µg mL^{-1} for kanamycin and 35 µg mL^{-1} for chloramphenicol.

Synechocystis sp. PCC 6803 (glucose tolerant strain, obtained from D. Bhaya, Stanford University, USA) and the strain Δ*fumC* (knockout of *fumC* gene)[129] were cultivated at 30 °C, in liquid BG-11 medium supplemented with 10 mM TES-KOH (pH = 8) under constant moderate red light illumination (~ 30 µE · m^{-2} · s^{-1}) using a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific), or on BG-11 agar plates supplemented with 10 mM TES-KOH and 0.3% (w/v) sodium thiosulfate. For the construction of the strain Syn-DcuZ6, kanamycin in a concentration of 50 µg mL^{-1} and chloramphenicol in a concentration of 20 µg mL^{-1} were added to the medium. The biomass concentration in the cultures was measured by optical density at 730 nm (OD$_{730}$) in a spectrophotometer (Lightwave II, Biochrom).

Plasmid and strain construction. All plasmids, strains and primers are listed in Tables 4.1 and 4.2. Plasmids pCI148 and pCI146 (obtained from C. Immethun, Washington University)[59] were used as a starting point for the construction of the biosensor. The genome of *E. coli* was used as a template to amplify the complete sequence of the *dcuS* gene, utilizing the primers 1-4. These primers generated a point mutation at position 858 (T → C) to remove an Acc65I restriction site without changing the original amino acid (glycine, position 286). Then, using the primers 5-14 the first part of *dcuS* gene was amplified, adding an Acc65I restriction site to the start of the gene. This resulted in *dcuS* fragments with 9 different lengths. The longest consist of 1140 base pairs and the shortest fragment consists of 1020 base pairs, with intervals of 15 bp in between. The sequences of the linker region are described in Table 4.3.

The plasmid pCI148 was used as a template to amplify the *E. coli envZ* gene. This plasmid contains the last 690 bp of the original *envZ* gene. Using the primers 15-24, the partial *envZ* gene was amplified from pCI148, up to and including the MunI restriction site. The *dcuS* and *envZ* fragments were combined by fusion PCR, resulting in 9 different *dcuSZ* genes, whose sequence was confirmed by Sanger sequencing at Macrogen Europe (The Netherlands). The pCI148 plasmid (replicative with an RSF1010 replicon) was digested with Acc65I and MunI restriction enzymes, to introduce the *dcuSZ* genes, replacing the original *cph8* gene. The *dcuSZ* containing plasmids together with the unedited pCI146 plasmid containing the OmpR cassette (Kan^R), were transformed into *E. coli* leading to the fumarate sensing strains *E. coli* DcuSZ 1-9. To construct the strain Syn-DcuZ6, *Synechocystis* sp. PCC 6803 was transformed with the plasmids pDcuZ6 and pCI145. The plasmid pCI145 contains the *ompR* gene, expressed constitutively under the control of the Bba J23104 promoter, and homologous regions to be integrated in the neutral site NS1, of the chromosome, replacing the *psbA1* gene. The OmpR cassette (Kan^R) from the pCI145 plasmid was integrated in the genome of *Synechocystis* by natural transformation, as described previously by Du *et al.*[125] The RSF1010-based pDcuZ6 plasmid

was introduced in *Synechocystis* by conjugation, performed by tri-parental mating using *E. coli* J53 (pRP4) as the helper strain, essentially as described by Savakis *et al.*[161]

Table 4.1 Plasmids and strains used in this study

Plasmids and strains	Relevant characteristics	Reference
pCI145	Integrative (NS1) vector containing f1/pBR322 ori, Bba J23104-*ompR*, Kan^R	59
pCI146	Integrative (NSP1) vector containing M13/pBR322 ori, Bba J23104-*ompR*, Kan^R	59
pCI148	RSF1010 derivate containing PLtetO-1-*cph8*, P*ompC-eyfp,*Cm^R	59
pDcuSZ 1	p148 derivate containing PLtetO-1-*dcuSZ1*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 2	p148 derivate containing PLtetO-1-*dcuSZ2*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 3	p148 derivate containing PLtetO-1-*dcuSZ3*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 4	p148 derivate containing PLtetO-1-*dcuSZ4*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 5	p148 derivate containing PLtetO-1-*dcuSZ5*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 6	p148 derivate containing PLtetO-1-*dcuSZ6*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 7	p148 derivate containing PLtetO-1-*dcuSZ7*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 8	p148 derivate containing PLtetO-1-*dcuSZ8*, P*ompC-eyfp,*Cm^R	This study
pDcuSZ 9	p148 derivate containing PLtetO-1-*dcuSZ9*, P*ompC-eyfp,*Cm^R	This study
E. coli DH5α	Cloning host	MCLAB
E. coli DcuSZ 1	*E. coli* strain containing DcuSZ 1 signaling system	This study
E. coli DcuSZ 2	*E. coli* strain containing DcuSZ 2 signaling system	This study
E. coli DcuSZ 3	*E. coli* strain containing DcuSZ 3 signaling system	This study
E. coli DcuSZ 4	*E. coli* strain containing DcuSZ 4 signaling system	This study
E. coli DcuSZ 5	*E. coli* strain containing DcuSZ 5 signaling system	This study
E. coli DcuSZ 6	*E. coli* strain containing DcuSZ 6 signaling system	This study
E. coli DcuSZ 7	*E. coli* strain containing DcuSZ 7 signaling system	This study
E. coli DcuSZ 8	*E. coli* strain containing DcuSZ 8 signaling system	This study
E. coli DcuSZ 9	*E. coli* strain containing DcuSZ 9 signaling system	This study
Synechocystis sp. PCC 6803	*Synechocystis* sp. PCC 6803 glucose tolerant 6803 wild type	D. Bhaya
Syn-DcuZ6	*Synechocystis* sp. PCC 6803 containing DcuSZ 6 signaling system	This study

Table 4.2 Primers used in this study

#	Primer*	Sequence	Additional Information
1	Fmut1_dcuS	atgagacattcattgcccta	Amplify Mut_1_*dcuS*
2	Rmut1_dcuS	ttcgtcgcggcgcggggt**g**ccgtcgcgtaacacttcc	Amplify Mut_1_*dcuS*
3	Fmut2_dcuS	ggaagtgttacgcgacgg**c**accccgcgccgcgacgaa	Amplify Mut_2_*dcuS*
4	Rmut2_dcuS	tcatctgttcgacctctccc	Amplify Mut_2_*dcuS*
5	For_dcuS	cacggtacccatgagacattcattgcccta	Amplify all variants of *dcuS*
6	Rev1_dcuS	cttaacaccagccgccatatgtccgagaatcacatgcaattt	Amplify *dcuS 1*
7	Rev2_dcuS	cttaacaccagccgccatatgcaatttattcataaattcgtg	Amplify *dcuS 2*
8	Rev3_dcuS	cttaacaccagccgccatatgttcgtgggatcgttcac	Amplify *dcuS 3*
9	Rev4_dcuS	cttaacaccagccgccatatgacgaagtgcgtcagcata	Amplify *dcuS 4*
10	Rev5_dcuS	cttaacaccagccgccatatgatagttgaccagaccgtc	Amplify *dcuS 5*
11	Rev6_dcuS	cttaacaccagccgccatatgcttcagatgcaataatccgag	Amplify *dcuS 6*
12	Rev7_dcuS	cttaacaccagccgccatatgcaactgcttataactcttcagatg	Amplify *dcuS 7*
13	Rev8_dcuS	cttaacaccagccgccatatggagaatgtaatcttccaactgct	Amplify *dcuS 8*
14	Rev9_dcuS	cttaacaccagccgccatatggttattggctgttttgagaatgt	Amplify *dcuS 9*
15	For1_envZ	aaaattgcatgtgattctcggacatatggcggctggtgttaag	Amplify *envZ 1*
16	For2_envZ	cacgaatttatgaataaattgcatatggcggctggtgtta	Amplify *envZ 2*
17	For3_envZ	gtgaacgatcccacgaacatatggcggctggtgttaag	Amplify *envZ 3*
18	For4_envZ	tatgctgacgcacttcgtcatatggcggctggtgttaag	Amplify *envZ 4*
19	For5_envZ	gacggtctggtcaactatcatatggcggctggtgttaag	Amplify *envZ 5*
20	For6_envZ	ctcggattattgcatctgaagcatatggcggctggtgttaag	Amplify *envZ 6*
21	For7_envZ	catctgaagagttataagcagttgcatatggcggctggtgttaag	Amplify *envZ 7*
22	For8_envZ	agcagttggaagattacattctccatatggcggctggtgttaag	Amplify *envZ 8*
23	For9_envZ	acattctcaaaacagccaataaccatatggcggctggtgttaag	Amplify *envZ 9*
24	Rev_envZ	acgctgcacaattgccag	Amplify all variants of *envZ*
25	145 Psb up	ggctatgtcccgcttaaactc	Check insertion of OmpR
26	145 Psb down	gcgagagtgcccttgattaac	Check insertion of OmpR
27	envZ forw	gatacgctgcacaattgccag	Check insertion of DcuZ6
28	EYFP rev	gtcgtgcagcttcatgtggt	Check insertion of DcuZ6

Table 4.3 Sequences of the linker region

#	Sequence result DcuS linker region
1	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattgcatgtgattctcgga
2	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattg
3	gctgacgcacttcgtgaacgatcccacgaa
4	gctgacgcacttcgt
5	-
6	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattgcatgtgattctcggattattgcatctgaag
7	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattgcatgtgattctcggattattgcatctgaagagttataagcagttg
8	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattgcatgtgattctcggattattgcatctgaagagttataagcagttggaagattacattctc
9	gctgacgcacttcgtgaacgatcccacgaatttatgaataaattgcatgtgattctcggattattgcatctgaagagttataagcagttggaagattacattctcaaaacagccaataac

Testing conditions. First, a screening experiment was performed for the *E. coli* strains containing the 9 different DcuSZ variants, to determine the functionality of the biosensors by measuring the expression of fluorescence in response to fumarate. Then, a more detailed characterization experiment was performed for the strains 1, 6 and 9. *E. coli* DH5α was used for both experiments, growing it at 37 °C in liquid mineral medium M9 + arginine (DSMZ 450). This cultivation was done in an incubator with a shaking speed of 200 rpm, in shake flasks (screening) or in glass tubes (characterization). Kanamycin and chloramphenicol were added to the medium in a concentration of 25 μg mL^{-1} and 17.5 μg mL^{-1} respectively. Ampicillin was not added since it affects the transcription of the *ompC* gene[162], and our system is located behind the *ompC* promoter.

Before the start of the experiment, pre-cultures were grown for approximately 24 hours in M9 medium. The cultures were diluted to an OD$_{600}$ of 0.5 at the start of the experiment, adding anhydrotetracycline (aTc) to a final concentration of 100 ng mL^{-1} and fumarate in a concentration ranging from 0 to 10 mM for the screening experiment and from 0 to 20 mM for the characterization experiment. The osmolarity was kept equal by adding the appropriate concentration of NaCl to the fumarate solutions.

The fluorescence of eYFP was measured in a black 96-well plate with a BMG FluoStar Optima plate reader, using an excitation wavelength of 470 nm and an emission wavelength of 510 nm (gain was set to 1000). For the OD$_{600}$ measurement, an ISOGEN Lightwave II spectrophotometer was used. Following the addition of aTc and fumarate, measurements were made after 0, 16, and 22 hours of cultivation for the screening experiment and after 0 and 22 hours of cultivation for the characterization experiment.

To test the functionality of the *E. coli* biosensor strain DcuS6 to detect fumarate produced by *Synechocystis*, the biosensor strain was inoculated and cultured overnight in liquid mineral medium M9 + arginine (DSMZ 450), adding 25 μg mL^{-1} of kanamycin, 17.5 μg mL^{-1} of chloramphenicol and 100 ng mL^{-1} of aTc, in an incubator with a shaking speed of 200 rpm at 37 °C. After 20 hours of culture, two aliquots of 2 mL were taken and centrifuged at 5000 rpm for 10 min, discarding the supernatant. Next, 2 mL of both a Δ*fumc* and a wild type *Synechocystis* culture were centrifuged at 14.500 rpm for 5 min, and 1 mL of each culture´s supernatant was used to resuspend each biosensor cell pellet. Each resuspension was divided over 6 wells (6 replicates), in a black 96 well plate and then placed in a plate reader (FluoStar Optima) at 37 °C for 60 hours. Fluorescence measurements were performed using excitation wavelength of 470 nm and an emission wavelength of 510 nm (gain was set to 1000). Fluorescence and OD$_{600}$ were measured every 15 min during the entire incubation period.

HPLC measurements. A sample of 1 mL of the Δ*fumc* and wild type *Synechocystis* cultures was centrifuged at 15.000 rpm for 10 min, to take 500 μL of supernatant and filter it (Sartorius Stedin Biotech, minisart SRP 4, 0.22 μm) for sample preparation. The concentration of fumarate was measured using HPLC-UV/VIS (LC-20AT, Prominence, Shimadzu), with ion exclusion Rezex ROA-Organic Acid column (250 x 4.6 mm, Phenomenex) and UV detector (SPD-20A, Prominence, Shimadzu) at a wavelength of 210 nm. A volume of 50 μL of the HPLC samples were injected through an autosampler (SIL-20AC, Prominence, Shimadzu), using as effluent 5 mM H$_2$SO$_4$ at a flow rate of 0.15 mL min^{-1} and 45 °C for the column temperature. The retention time of fumarate was determined

between 18.16 and 18.36 min for the system used, and fumarate samples were normalized using a correction factor of 10 mM divided by the measured TES concentration.

Data analysis. For the screening and characterization experiments in *E. coli*, the raw fluorescence data was corrected for position bias of the wells by subtracting the measurement at 0 hour of each well from all the following measurements in that well. The fluorescence was then normalized by dividing it by the corresponding untransformed OD_{600} measurement. To analyze the dynamics of each strain's response and make a better comparison between them, the background fluorescence/OD_{600} was removed by subtracting the 0 mM fumarate fluorescence/OD_{600} value of each strain at each time point from the fluorescence/OD_{600} value of all fumarate concentrations of that strain at that time point.

For the characterization experiment, the wild type fluorescence/OD_{600} was subtracted from fluorescence/OD_{600} of all strains. For the curve fitting of the data, a Gompertz function was used. This is a Hill equation-like logistic equation following the structure: y = ae ^(-be^(-cx)) where *a* is the asymptote and *b* and *c* are positive numbers.

To detect the fumarate concentration in *Synechocystis* using the *E. coli* biosensor strain DcuS6, position bias was corrected for both OD_{600} and fluorescence by subtracting the mean of the first three measurements in each well of all measurements in that well. The mean OD_{600} of the first measurement in 6 replicates was later added. The fluorescence was normalized by dividing it by the OD_{600}. The background fluorescence/OD_{600} was corrected by subtracting the mean of the 6 replicates of the wild type fluorescence/OD_{600} per time point of all separate fluorescence/OD_{600} values at that time point. The average of the fluorescence/OD_{600} over 22-60 hours for every replicate was calculated and the mean was plotted with an error bar showing the 95% confidence interval.

Author contributions PCB, WD and FBS designed the experiments; PCB and AT performed the experiments; PCB and AT wrote the manuscript; FBS reviewed it.

Is gene expression in *Synechocystis* sp. PCC 6803 influenced by genomic location?

Patricia Caicedo Burbano and Filipe Branco dos Santos

Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Netherlands

Abstract: An open question related to the expression of heterologous genes is to what extent can it be affected by its location within the foreigner chromosome. Several factors may affect the efficacy of the process. To name but a few, (i) the surrounding sequences to the gene of interest may have uncharacterized regulatory roles; (ii) the three-dimensional structure of the chromosome can affect the efficacy with which that particular region is expressed; and (iii) the closeness to the origin of replication can vary its copy number and hence its expression level depending on the growth rate. Previous studies in *Synechocystis* sp. PCC 6803 have identified at least six loci that could be disrupted without affecting the viability of the cell or cause any noticeable phenotype, *i.e.* neutral sites. These are situated at different locations around the chromosome of *Synechocystis* and can be used to study how this specifically affects expression of a given sequence. In this study, we have introduced the same cassette in six different loci spread over the chromosome and we have characterized its expression. This cassette consists of an eYFP protein emitting yellow fluorescence under the control of the P*cpc* promoter. The results obtained did not show big differences, in terms of fluorescence levels, irrespective of the locus in which the cassette was cloned. This supports the idea that in this relatively slow-growing organism, and for at least these six tested neutral sites, expression levels are not very much affected by chromosomal location *per se*.

Keywords: gene expression, promoters, neutral sites, *Synechocystis*

Introduction

Photosynthetic cyanobacteria have been studied as host organisms for the production of desired compounds from CO_2 and water, using solar radiation as energy[163]. Although they are considered promising cell factories, there are still various challenges that restrict their capacity to a sufficient degree for commercial applications. To overcome these limitations, synthetic biology approaches using validated genetic control elements have been evaluated in cyanobacterial host[27].

One of the challenges when engineering cyanobacteria is to have precise control over the expression of the introduced genes in a predictable manner. It is known that gene expression is regulated by the combined effect of the physiological state of the cell (especially the activity of the machinery responsible for gene expression) and DNA-binding transcription factors, amongst other specific regulators[164]. However, it remains an open question how much the expression of heterologous genes can be affected by its location within the foreigner chromosome.

There are several factors that may affect the gene expression. To name but a few: (i) the surrounding sequences to the gene of interest may have uncharacterized regulatory roles; (ii) the three-dimensional structure of the chromosome may affect the efficacy with which that particular region is expressed; and (iii) the closeness to the origin of replication can cause variations in the gene copy number and hence in the gene expression level depending on the growth rate.

Previous studies have shown that the expression level of a transcription and translation initiation unit may vary depending on the genetic context. This occurs as a result of different secondary structures formed by different mRNAs within and around the 5´untranslated region (5´UTR) that can hide the ribosome binding site (RBS)[165,166]. The 3´untranslated region (3´UTR) is also involved in gene regulation. For instance, 3´UTRs can be recognized and cleaved by ribonucleases to initiate the decay of mRNA. Or they are rich reservoirs of regulatory RNAs able to regulate the expression of target genes. Additionally, they can interact with regulatory sRNA protecting or promoting 3´UTR degradation, or they can also interact with the 5´UTR region affecting translational initiation and mRNA stability[167].

There are multiple reports of the differences in the expression of a gene depending on its chromosomal location (chromosomal position effect) in prokaryotic organisms such as *Escherichia coli*, *Salmonella typhimurium*, *Lactococcus lactis* and *Bacillus subtilis*[168,169,170,171]. This variability can be caused by the level of DNA compactization and supercoiling[172], and the variation of the promoter strength. However, the major factor of variability in gene expression in bacteria seems to be the distance to the origin of replication. A gene located closer to the origin of replication is replicated before a gene located near the terminus, creating an increase in gene dosage (especially at high rates of growth when multiple rounds of chromosome replication are initiated)[173].

Even though most of the chromosomal-position effects for heterologous genes are predominantly due to gene dosage, a study in *E. coli* shows that the variation of gene expression can be due only to the location[168]. In some polyploid species of cyanobacteria such as *Synechocystis*, the copy number of plasmids and genomes may vary, affecting the gene dosage. However, it is not clear if the location *per se* affects the expression of heterologous genes in these organisms[42].

The integration of heterologous genes in the model cyanobacteria *Synechocystis*, has been usually done in "neutral sites", *i.e.* genomic locations where a modification does not result in a noticeable phenotypic change, nor does it affect cellular viability. Some of these neutral sites, located in the chromosome and in the plasmids of *Synechocystis* have been identified, characterized and their neutrality has been validated in recent studies[174,175].

Ng *et al*., compared the expression of a reported gene inserted in two putative neutral sites located in the chromosome and in one putative neutral site located in an endogenous plasmid of *Synechocystis*. Their results showed that the expression of eYFP was nearly identical for the chromosomal neutral sites, while the expression in the endogenous plasmid was 8 to 14 times higher. This difference increased when the cells went from the exponential to the stationary phase[174]. These results are congruent with the gene dosage effect caused by the plasmid copy number and are in concordance with previous observations[176].

The episomal expression of target genes could be useful for production applications due to the higher expression that they exhibit in comparison with the chromosomal expression. However, its use presents some challenges such as instability, the loss of the plasmid in absence of selective pressure and the variation of the copy number due to nutritional conditions and growth phase. This might compromise the quantification and predictability of the system behavior[175]. The chromosomal integration confers more stability and reduces the need of permanent selective pressure, which is desired for long term cultivations and bioproduction applications. Yet, more analyses are needed to determine if the locus of insertion in the chromosome might affect the expression of a heterologous gene in *Synechocystis*.

As it has been demonstrated in other organisms, the surrounding genomic region and the location in the chromosome of a particular gene may affect its expression. If that is the case for *Synechocystis*, the expression of a reporter gene will vary when inserted at different places of the chromosome. In this study, the impact of genomic location on gene expression was evaluated by measuring the expression of the fluorescent protein eYFP under the control of the strong constitutive promoter P*cpc*, when the reporter cassette is placed in six different neutral sites spread out over the chromosome of *Synechocystis*.

Results and Discussion

Construction of plasmids

To evaluate the effect of the chromosome location in the expression of a heterologous gene in *Synechocystis* sp. PCC 6803, the validated neutral sites *slr1396, slr0271, slr0397, slr0573, sll1476*[175] and *slr0168*[177] were chosen. The gene *slr0168* codes for a protein with unknown function, but it has been shown that its disruption does not affect the viability of the cells[177,178]. The other neutral sites, code for unknow or hypothetical proteins and are not in the vicinity of genes with assigned putative function[175]. The six neutral sites are distributed across the chromosome of *Synechocystis* (Figure 5.1).
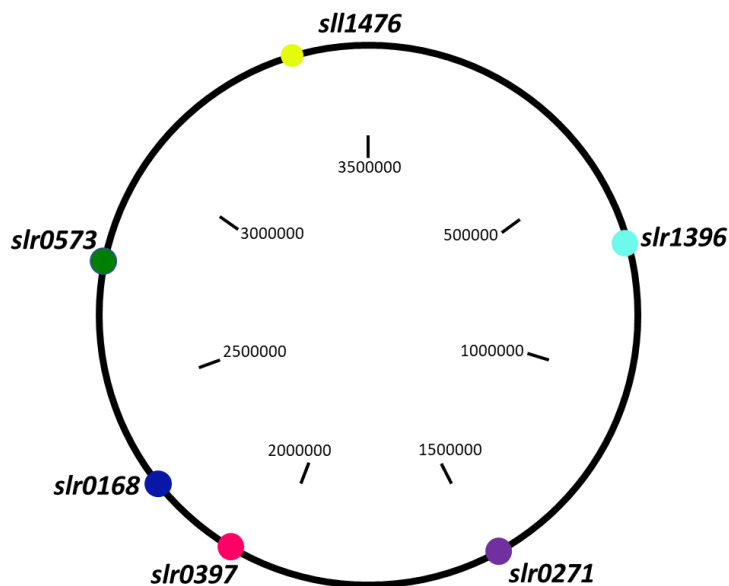
Neutral site	Locus	Bases deleted	Neutral site length (bp)
slr1396	707437-708054	707572-707857	618
slr0271	1524568-1525086	1524568-1525033	519
slr0397	2146801- 2147376	2147211-2147217	576
slr0168	2300515-2302446	2300756-2300761	1932
slr0573	2816517-2816960	2816587-2816919	444
sll1476	3398409-3398717	3398447-3398673	309

Figure 5.1 Scheme and description of the location of the neutral sites in the chromosome of *Synechocystis*.
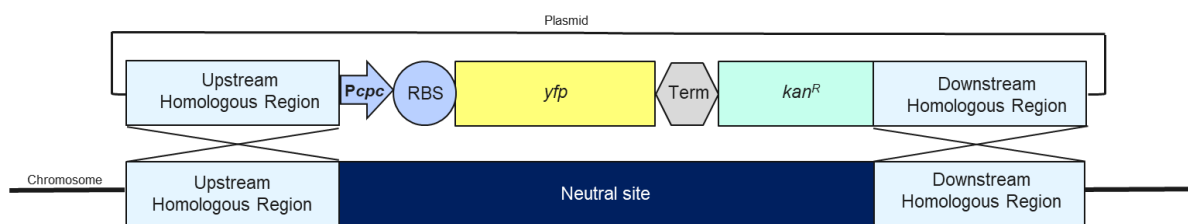

Figure 5.2 Schematic representation of the plasmid used to insert the reporter cassette into the different neutral sites. Double homologous recombination was used to integrate the cassette containing P*cpc* promoter, *eyfp*, and *kan*^R marker intro each neutral site.

Six suicide (integrative) plasmids were constructed to insert the reporter gene into different loci in the chromosome. Each plasmid had homologous regions for each neutral site, to be able to recombine in the chromosome. Between the upstream and downstream homologous regions, a reporter cassette, composed by the P*cpc* promoter, a ribosome binding site, the gene encoding for the fluorescence protein eYFP and a terminator, was inserted (Figure 5.2). In addition, a gene encoding for kanamycin resistance protein was included in this vector for propagation purpose.

Plasmids containing the same cassette without the promoter region were also constructed to be used as negative control for the constitutive expression of eYFP.

EYFP expression on different locations of the chromosome

To determine the effect of the chromosomal location in the expression of eYFP in *Synechocystis*, the fluorescence of the strains with the expression cassette integrated on different neutral sites, was measured when the cells were in exponential phase ($OD_{730} \approx 1$). As a negative control, the strains having a promoterless reporter cassette were included. The analysis of fluorescence of the different *Synechocystis* strains, showed a similar pattern in the expression of eYFP, being *slr0573* the place on the chromosome where the expression of eYFP was somewhat higher (Figure 5.3). A one-way ANOVA analysis was performed, showing that the differences cannot be attributed to chance. These results are in agreement with the results of Pinto and co-workers, who did not find a considerable difference in the expression of the reporter cassette P*psbA2-gfp,* when it was introduced in five of the six neutral sites used in this study.
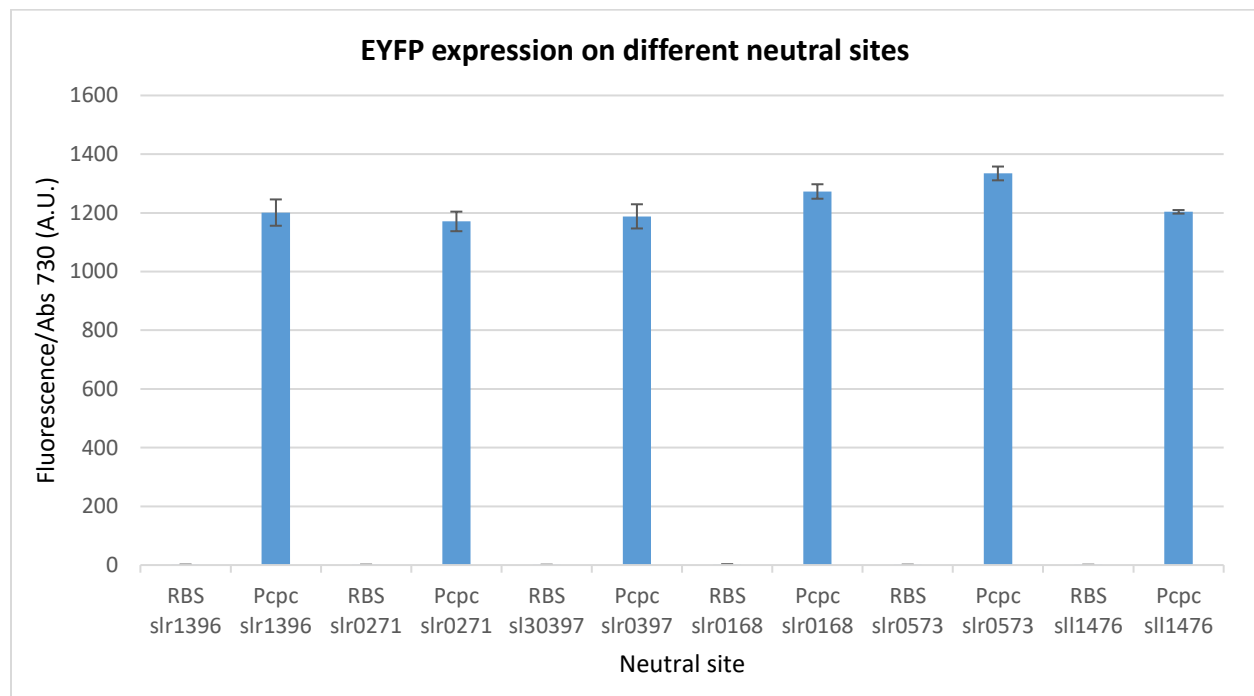

Figure 5.3 Diagram of the eYFP expression under the control of P*cpc*, when the cassette is integrated on different neutral sites of the chromosome of *Synechocystis*. As a negative control, a cassette without a promoter was used (RBS). The fluorescence was normalized to OD_{730}. Bars indicate mean ± S.D. (n = 3).

Several studies have been done to determine the effect of genome location on gene expression using model organisms such as *Escherichia coli* and *Bacillus subtills*. Block *et al.*, found that in in rapidly growing cultures of *E. coli*, the expression levels for gene locations close to the origin of replication were approximately 9-fold higher than the expression on locations close to the replication terminus[173]. Similarly, Murray and Koh reported that the copy number of a gene close to origin of

replication in *B. subtilis* when growing exponentially, can be ≈ 5 times higher than the copy number of a gene close the terminus[179], indicating that gene dosage has a major effect on gene expression in fast growing cells.

When *B. subtills* is cultivated in minimal medium, where the cells grow considerably slower, it reduces the difference in the expression of the reporter gene[179]. In a similar way, Sauer and co-workers showed that when the reporter gene is induced only in the stationary phase (DNA replication is reduced when exponential growth ceases), the difference in gene expression between different location was almost completely abolished[171]. These results show the role that gene dosage, as a consequence of multiple replication forks, plays in gene expression in fast growing cells.

Unlike *E. coli* and *B. subtills*, which have doubling times of 20 and 27 min at 37 °C[179,180], the doubling time of wild-type *Synechocystis* under phototrophic conditions, is around 12 hours[181], although it can be as short as 4.3 h when the growth conditions are optimized[182]. The slow growth of *Synechocystis* may decrease the need of multiple replication forks during the chromosome replication, causing a reduction of the gene dosage effect due to the distance to the replication origin.

Additional studies have shown that other factors such as the chromosomal three-dimensional structure and the genetic context may affect the expression of a heterologous gene. This has been observed by the variations in the expression levels when the same transcriptional and translational units were used to express different genes in *Synechocystis*[56]. Nevertheless, we found that at least for these six tested neutral sites, expression levels are not very much affected by chromosomal location *per se*.

Conclusion

In this study we have integrated a reporter cassette on six different neutral sites in the chromosome of *Synechocystis* to test the effect of location in the expression of heterologous genes. Previous work has shown the regulatory role of the surrounding genomic regions and the influence that other factors such as the three-dimensional chromosome structure and the distance to the origin of replication have in the expression of a gene. However, we did not observe a big variation in the expression of eYFP when its gene was introduced on different chromosomal location.

These results are in concordance with previous reports in *Synechocystis*[174,175]. This suggests that the differential gene expression caused by gene dosage effect in fast growing organisms, might not occur in slow growing organisms (such as *Synechocystis)*, at least not for chromosomal expression in neutral sites. The fact that neutral sites do not encode for proteins essential to the viability of the cells but encode for hypothetical or unknown proteins, might be reducing the effect that the surrounding regions could have in the expression of a heterologous gene. This shows the importance of choosing an adequate insertion site.

The performance of biological parts (biobricks) such as promoters, ribosome binding sites and terminators should ideally be independent of the location in the genome where they are introduced. This functional independence from their surrounding genomic region is essential for them to act as

transferable elements. The characterization of the factors and conditions affecting this independence is necessary for the effective implementation of synthetic biology strategies, namely in *Synechocystis.*

Methods

Strains and general cultivation conditions. *Escherichia coli* strains DH5α was used as the host for recombinant plasmid construction and for plasmid amplification and cloning. It was grown at 37 °C in Luria-Bertani Broth (LB) in a shaking incubator at 200 rpm, or on LB agar. The concentration of antibiotic used was 50 μg mL^{-1} for kanamycin.

Synechocystis sp. PCC 6803 (glucose tolerant, obtained from D. Bhaya) was grown at 30 °C in liquid BG-11 medium supplemented with 10 mM TES-KOH (pH = 8), in a shaking incubator at 120 rpm (Innova 44, New Brunswick Scientific) under constant moderate red light illumination (~ 30 μE · m^{-2} · s^{-1}), or on BG-11 agar plates supplemented with 10 mM TES-KOH (pH = 8) and 0.3% (w/v) sodium thiosulfate. For the construction of the mutants, kanamycin was added to the medium to a final concentration of 50 μg mL^{-1}. Biomass concentration in the cultures was measured by optical density at 730 nm (OD$_{730}$) in a spectrophotometer (Lightwave II, Biochrom).

Plasmid and strain construction. All plasmids, strains and primers are listed in Tables 5.1 and 5.2. To construct the integrative plasmids, the upstream and downstream homologous regions of each neutral site (~ 1 kb each) were amplified from the genomic DNA of *Synechocystis* using Herculase II polymerase (Agilent). Primers were designed to introduce restriction sites in the 5´ of the upstream region, the 3´ of the downstream region (detailed in the Table 5.2), and XbaI in between the two regions. The fragments were fused together using Pfu DNA Polymerase (Thermo Scientific). After gel extraction and purification (Thermo Scientific), the fused fragments were inserted in the vector pFL-XN (both digested with the corresponding pair of enzymes) and ligated using T4 ligase (Thermo Scientific). This resulted in the new plasmid pFLXN (for each neutral site), serving as the background plasmid for the following constructs.

The cassette RBS-YFP, and CPC-YFP were obtained from the plasmid pHKH-cpcBA-His10 by PCR amplification. To generate the CPC-YFP cassette, the region containing the *cpcBA* promoter (including the RBS), e*yfp*, terminator BB0014 and *kan*^{R} was amplified. To generate the RBS-YFP cassette, only the region containing the ribosome binding site from the *cpcBA* promoter (excluding the promoter region), e*yfp*, terminator BB0014 and *kan*^{R} was amplified.

The RBS-YFP and CPC-YFP cassettes were cloned in an intermediate plasmid and digested with XbaI and NheI to generate compatible ends. Then, they were inserted into the pFLXN digested with XbaI, resulting in the plasmids pFLXN-CPC-YFP and pFLXN-RBS-YFP for each neutral site. All the fragments amplified in this study were confirmed by Sanger sequencing at Macrogen Europe (The Netherlands).

Synechocystis strains Syn-RBS-YFP, and Syn-CPC-YFP were created by natural transformation of *Synechocystis* wild type with the corresponding pFLXN-CPC-YFP and pFLXN-RBS-YFP plasmids for each neutral site, as previously described by Du *et al.*[125] Full segregation of each mutant was verified by PCR using the genomic DNA as the template for 35 cycles.

Fluorescence analysis. Precultures of the Syn-CPC-YFP and Syn-RBS-YFP strains were inoculated on 48 well plates with BG-11 medium, 10 mM TES and 50 mM $NaHCO_3$. The cells were reinoculated to OD_{730} = 0.25 and after three days ($OD_{730} \approx 1$), 200 µL were used to measure the total fluorescence with plate reader (FLUOstar optima, BMG labtech) using a 470 nm filter for excitation and a 510 nm for emission. The measurement was done on a 96 well black plate (Greiner-F bottom) using three biological replicates and three technical replicates.

Data analysis. To avoid position bias, three biological replicates, and three technical replicates, were distributed on different positions of the plate. The raw OD_{730} and fluorescence data were corrected by subtracting the blank. Next, the fluorescence was normalized by dividing it by the corresponding OD_{730} measurement. Technical replicates were averaged, and the biological replicates were used to calculate the standard deviation.

Author contributions PCB and FBS designed the experiments; PCB performed the experiments; PCB wrote the manuscript; FBS reviewed it.

Table 5.1 Plasmids used in this study

Plasmids	Relevant characteristics	Reference
pHKH-cpcBA-His10-YFP	Vector containing the CPC-YFP cassette	119
pFL-XN	Biobrick "T" vector (XbaI and NheI) used for functional block assembling, Ap^R, Cm^R	126
pFLXN-RBS-YFP-1396	pFL-XN derivate, Km^R, containing *slr1396* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-RBS-YFP-0271	pFL-XN derivate, Km^R, containing *slr0271* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-RBS-YFP-0397	pFL-XN derivate, Km^R, containing *slr0397* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-RBS-YFP-0168	pFL-XN derivate, Km^R, containing *slr0168* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-RBS-YFP-0573	pFL-XN derivate, Km^R, containing *slr0573* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-RBS-YFP-1476	pFL-XN derivate, Km^R, containing *sll1476* gene upstream homologous region, RBS-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-1396	pFL-XN derivate, Km^R, containing *slr1396* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-0271	pFL-XN derivate, Km^R, containing *slr0271* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-0397	pFL-XN derivate, Km^R, containing *slr0397* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-0168	pFL-XN derivate, Km^R, containing *slr0168* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-0573	pFL-XN derivate, Km^R, containing *slr0573* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study
pFLXN-CPC-YFP-1476	pFL-XN derivate, Km^R, containing *sll1476* gene upstream homologous region, CPC-YFP cassette and downstream homologous region	This study

Table 5.2 Strains and primers used in this study

Strains and primers	Relevant characteristics	Reference
Synechocystis sp.	*Synechocystis* sp. PCC 6803 glucose tolerant wild type	D. Bhaya
Syn-RBS-YFP-1396	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *slr1396*, Km^R	This study
Syn-RBS-YFP-0271	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *slr0271*, Km^R	This study
Syn-RBS-YFP-0397	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *slr0397*, Km^R	This study
Syn-RBS-YFP-0168	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *slr0168*, Km^R	This study
Syn-RBS-YFP-0573	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *slr0573*, Km^R	This study
Syn-RBS-YFP-1476	*Synechocystis* sp. PCC 6803 with RBS-YFP cassette inserted in *sll1476*, Km^R	This study
Syn-CPC-YFP-1396	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *slr1396*, Km^R	This study
Syn-CPC-YFP-0271	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *slr0271*, Km^R	This study
Syn-CPC-YFP-0397	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *slr0397*, Km^R	This study
Syn-CPC-YFP-0168	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *slr0168*, Km^R	This study
Syn-CPC-YFP-0573	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *slr0573*, Km^R	This study
Syn-CPC-YFP-1476	*Synechocystis* sp. PCC 6803 with CPC-YFP cassette inserted in *sll1476*, Km^R	This study
H1 1396 SacI forw	GCAGAGCTCCCCTTATCGGAACCGATGGA	This study
H1 1396 XbaI rev	AGCCTATTTTGCGACCCATCTCTAGAGCAGCGCAATTGATTCCAGT	This study
H2 1396 XbaI forw	ACTGGAATCAATTGCGCTGCTCTAGAGATGGGTCGCAAAATAGGCT	This study
H2 1396 SalI rev	ATTGTCGACGTGAACGGCATCTTCAGGAC	This study
H1 0271 KpnI forw	GCAGGTACCGACGGAATGTTGGCAGGTAG	This study
H1 0271 XbaI rev	CAGCGAATGTGAATTTGCGGTCTAGAGACGAACAGTTTAGCCTCCT	This study
H2 0271 XbaI forw	AGGAGGCTAAACTGTTCGTCTCTAGACCGCAAATTCACATTCGCTG	This study
H2 0271 SalI rev	ATTGTCGACGAGTCTGATTCTGTGGGGCT	This study
H1 0397 KpnI forw	GACGGTACCCTTTGGCCGCTAGCAC	This study
H1 0397 XbaI rev	CTTGTTTCTCTCCATAGGCTTCTAGAGCTGAGGTAAATTTTCGGGC	This study
H2 0397 XbaI forw	GCCCGAAAATTTACCTCAGCTCTAGAAGCCTATGGAGAGAAACAAG	This study
H2 0397 SalI rev	ATTGTCGACCTTTGCTAAATCCTATCAGACTC	This study
H1 0168 KpnI forw	GCAGGTACCAAGTGGGGCACATTGAACGC	This study
H1 0168 XbaI_rev	GTGGACAAATCCCCCAGGTTTCTAGAGGCCACATTGTTGTCAAAGG	This study
H2 0168 XbaI forw	CCTTTGACAACAATGTGGCCTCTAGAAACCTGGGGGATTTGTCCAC	This study
H2 0168 HindIII rev	CGGAAGCTTAGGAGACTTTGGTGGGCTG	This study
H1 0573 KpnI forw	GCAGGTACCGAGAAGGGTGCAAAGGGAAT	This study
H1 0573 XbaI rev	AATCCAAAGCATGGCCAACCTCTAGAAGCTTGGCGAAACTAGGACTA	This study
H2 0573 XbaI forw	TAGTCCTAGTTTCGCCAAGCTCTAGAGGTTGGCCATGCTTTGGATT	This study
H2 0573 SalI rev	ATTGTCGACCACGGTAAACCAAAGTCACC	This study
H1 1476 KpnI_forw	GCAGGTACCGAACTAGCCAAATCCCTGGG	This study
H1 1476 XbaI rev	GGAAGTGTTGTTGCTGTCCATCTAGACGGAGTTCTGCTTTAGGTCA	This study
H2 1476 XbaI forw	TGACCTAAAGCAGAACTCCGTCTAGATGGACAGCAACAACACTTCC	This study
H2 1476 SalI rev	ATTGTCGACGCATTTCGCCAGAGGACC	This study
BglII cpc-RBS forw	AGATCTCCAACTCATAAAGTCAAGTAGGAG	This study
Pcpc forw	CGATCCCGACTTCGTTATAA	This study
KanTn rev	CGCTGAGGTCTGCCTC	This study

General Discussion

Patricia Caicedo Burbano and Filipe Branco dos Santos

Molecular Microbial Physiology Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Netherlands

Cyanobacteria as host for bioproduction

Cyanobacteria are an ancient clade of organisms within the phylogenetic tree of life. They have been key players in biochemical cycles including the ones involving oxygen, carbon and nitrogen. They also have a major role in the carbon cycle via direct assimilation of carbon dioxide or photosynthetic mediated biocalcification[183]. Many efforts have been done to understand the molecular mechanisms and metabolism of cyanobacteria aiming to employ that knowledge in various applications. Especially, in exploring the potential of cyanobacteria to be engineered as a microbial chassis for chemical production and atmospheric CO_2 assimilation[184].

Engineered cyanobacteria can directly use atmospheric CO_2 as a carbon source and sunlight as energy source, to produce a wide variety of useful compounds (biochemicals and biofuels). The latter can be done without competing for arable land or plant biomass (that could be used as food supply) to obtain sugar-based carbon sources, and thereby contributing to the reduction of greenhouse gases[35]. This is usually accomplished by introducing heterologous gene(s) or pathways, for the production of a target compound, using the host cell´s native metabolite(s) as a substrate for the production of the heterologous pathway. It is also common the introduction of gene knockouts, to channel more metabolic flux through the production pathway[32]. In addition, cyanobacteria show relatively fast photoautotrophic growth in comparison with plants, and they are accessible for genetic modification, serving as an attractive microbial host for "direct conversion". This approach has been extensively explored in the last years for the production of a variety of commodity compounds[6] (Table 6.1).

Despite all the progress achieved in the design and construction of cyanobacterial cell factories, it is still challenging to develop factories that meet the economic requirements on a commercial scale[185,186]. One of the reasons is that the current cyanobacterial cell factories are not yet capable of providing the necessary product quantities at a sufficient rate and yield. This can be attributed to a lack of understanding of their metabolism and regulation, and to a lack of adequate tools to engineer the cells in a predictable and controlled manner. New approaches, such as synthetic biology, have been used to tackle this problem. The number of tools designed to control gene expression and edit the genome in cyanobacteria is rapidly expanding, as well as the number of genetic parts and regulators characterized in cyanobacteria. The combination of these elements enables the construction of advanced genetic devices that allow the responsive and complex regulation required to develop a robust host for industrial production. The development of new tools will expand the usefulness of cyanobacteria as chassis in the biotechnological industry[30].

Table 6.1 List of selected compounds produced by cyanobacteria

Compound	Production*	Cyanobacteria	Reference
Acetate	0.2 mM/g DW	*Synechocystis* sp. PCC 6803	125
Acetone	0.036 g/L	*Synechocystis* sp. PCC 6803	187
Acetone	0.0225 g/L	*Synechococcus elongatus* PCC 7942	188
Alka(e)nes	0.026 g/L	*Synechocystis* sp. PCC 6803	189
Alkanes	1200 µg/g DCW	*Anabaena* sp. PCC 7120	190
Amorpha-4,11-diene	0.0198 g/L	*Synechococcus elongatus* PCC 7942	191
Astaxanthin	29.6 mg/g DCW	*Synechocystis* sp. PCC 6803	192
α-Bisabolene	0.0222 g/L	*Synechocystis* sp. PCC 6803	193
α-Bisabolene	0.0006 g/L	*Synechococcus* sp. PCC 7002	194
2,3-Butanediol	12.6 g/L	*Synechococcus elongatus* PCC 7942	195
2,3-Butanediol	2.38 g/L	*Synechococcus elongatus* PCC 7942	196
2,3-Butanediol	1.6 g/L	*Synechococcus* sp. PCC 7002	116
2,3-Butanediol	0.72 g/L	*Synechocystis* sp. PCC 6803	161
2,3-Butanediol	0.496 g/L	*Synechococcus elongatus* PCC 7942	197
2,3-Butanediol	0.028 g/L	*Synechocystis* sp. PCC 6803	161
1-Butanol	0.404 g/L	*Synechococcus elongatus* PCC 7942	198
1-Butanol	0.030 g/L	*Synechococcus elongatus* PCC 7942	199
1-Butanol	0.013 g/L	*Synechococcus elongatus* PCC 7942	200
1-Butanol	0.0065 g/L	*Synechococcus elongatus* PCC 7942	199
Caffeic acid	0.0072 g/L	*Synechocystis* sp. PCC 6803	201
β-Carotene	0.68 µg/mL/OD_{730}	*Synechocystis* sp. PCC 6803	202
β-Caryophyllene	0.00005 g/L	*Synechocystis* sp. PCC 6803	203
p-Coumaric acid	0.0826 g/L	*Synechocystis* sp. PCC 6803	204
Dhurrin	0.005 g/L	*Synechocystis* sp. PCC 6803	205
Dhurrin	0.0042 g/L	*Synechocystis* sp. PCC 6803	206
Dihydroxyacetone	0.0786 g/L	*Synechococcus elongatus* PCC 7942	207
Erythritol	0.256 g/L	*Synechocystis* sp. PCC 6803	119
Ethanol	5.50 g/L	*Synechocystis* sp. PCC 6803	8
Ethanol	3.856 g/L	*Synechococcus elongatus* PCC 7942	208
Ethanol	2.3 g/L	*Synechocystis* sp. PCC 6803	209
Ethanol	2.2 g/L	*Synechococcus* sp. PCC 7002	210
Ethanol	0.819 g/L	*Synechocystis* sp. PCC 6803	161
Ethanol	0.596 g/L	*Synechocystis* sp. PCC 6803	211
Ethylene	0.021 g/L	*Synechococcus elongatus* PCC 7942	51
Ethylene	0.011 g/L	*Synechocystis* sp. PCC 6803	212
Ethylene	0.0010 g/L	*Synechocystis* sp. PCC 6803	213

Table 6.1 Continued

Compound	Production*	Cyanobacteria	Reference
α-Farnesene	0.013 g/L	*Synechococcus elongatus* PCC 7942	214
α-Farnesene	0.0046 g/L	*Synechococcus elongatus* PCC 7942	215
α-Farnesene	0.000310 g/L	*Anabaena* sp. PCC 7120	216
Fatty acids	0.197 g/L	*Synechocystis* sp. PCC 6803	217
Fatty acids	0.130 g/L	*Synechococcus* sp. PCC 7002	218
Fatty acids	0.040 g/L	*Synechococcus elongatus* PCC 7942	219
Fatty acids	0.032 g/L	*Synechococcus elongatus* PCC 7942	220
Fatty alcohols	10.3 mg/g DCW	*Synechocystis* sp. PCC 6803	221
Fatty alcohols	2.87 mg/g DCW	*Synechocystis* sp. PCC 6803	222
Fatty alcohols	0.000200 g/L	*Synechocystis* sp. PCC 6803	57
Fumarate	0.5 mM/g DW	*Synechocystis* sp. PCC 6803	129
Geranyllinalool	360 µg/g DCW	*Synechocystis* sp. PCC 6803	223
Glucose and fructose	0.045 g/L	*Synechococcus elongatus* PCC 7942	224
Glucosylglycerol	1.64 g/L	*Synechocystis* sp. PCC 6803	225
L-Glutamate	2.3 g/L	*Synechocystis* sp. PCC 6803	226
Glycerol	3.5 g/L	*Synechococcus* sp. PCC 7002	227
Glycerol	1.32 g/L	*Synechocystis* sp. PCC 6803	228
Glycerol	1.17 g/L	*Synechococcus elongatus* PCC 7942	207
Glycerol	1.16 g/L	*Synechococcus elongatus* PCC 7942	229
Glycogen	3.5 g/L	*Synechococcus* sp. PCC 7002	227
Guanidine	0.5865 g/L	*Synechocystis* sp. PCC 6803	230
Heparosan	0.0000028 g/L	*Synechococcus elongatus* PCC 7942	231
Heptadecane	4 µg/g DCW	*Synechococcus sp. NKBG15041c*	232
(*R*)-3-Hydroxybutyrate	1.84 g/L	*Synechocystis* sp. PCC 6803	233
3-Hydroxybutyrate	1.2 g/L	*Synechococcus elongatus* PCC 7942	234
3-Hydroxybutyrate	0.533 g/L	*Synechocystis* sp. PCC 6803	235
3-Hydroxybutyrate	0.232 g/L	*Synechocystis* sp. PCC 6803	127
p-Hydroxyphenylacetaldoxime	0.066 g/L	*Synechocystis* sp. PCC 6803	205
3-Hydroxypropionic acid	0.837 g/L	*Synechocystis* sp. PCC 6804	236
3-Hydroxypropionic acid	0.659 g/L	*Synechococcus elongatus* PCC 7942	237
3-Hydroxypropionic acid	0.0317 g/L	*Synechococcus elongatus* PCC 7942	207
Isobutanol	0.608 g/L	*Synechocystis* sp. PCC 6802	115
Isobutanol	0.550 g/L	*Synechococcus elongatus* PCC 7942	238
Isobutanol	0.450 g/L	*Synechococcus elongatus* PCC 7942	11
Isobutanol	0.298 g/L	*Synechocystis* sp. PCC 6803	239
Isobutyraldehyde	1.10 g/L	*Synechococcus elongatus* PCC 7942	11
Isoprene	1.26 g/L	*Synechococcus elongatus* PCC 7942	240
Isoprene	0.0003 g/L	*Synechocystis* sp. PCC 6803	241
Isoprene	5.4 mg/g DCW	*Synechocystis* sp. PCC 6803	242

Table 6.1 Continued

Compound	Production*	Cyanobacteria	Reference
Isoprene	2.8 mg/g DCW	*Synechocystis* sp. PCC 6803	243
Isoprene	1.8 mg/g DCW	*Synechocystis* sp. PCC 6803	244
Isopropanol	0.146 g/L	*Synechococcus elongatus* PCC 7942	245
Isopropanol	0.0265 g/L	*Synechococcus elongatus* PCC 7942	10
Itaconic acid	0.0145 g/L	*Synechocystis* sp. PCC 6803	246
D-Lactic acid	2.17 g/L	*Synechocystis* sp. PCC 6803	247
D-Lactic acid	1.06 g/L	*Synechocystis* sp. PCC 6803	248
D-Lactic acid	0.829 g/L	*Synechococcus elongatus* PCC 7942	249
D-Lactic acid	0.055 g/L	*Synechococcus elongatus* PCC 7942	224
L-Lactic acid	1.84 g/L	*Synechocystis* sp. PCC 6803	250
L-Lactic acid	0.836 g/L	*Synechocystis* sp. PCC 6803	251
L-Lactic acid	0.469 g/L	*Synechocystis* sp. PCC 6803	251
L-Lactic acid	0.015 g/L	*Synechocystis* sp. PCC 6803	252
Limonene	0.0067 g/L	*Synechocystis* sp. PCC 6803	253
Limonene	0.005 g/L	*Synechococcus elongatus* PCC 7942	254
Limonene	0.004 g/L	*Synechococcus* sp. PCC 7002	194
Limonene	0.001 g/L	*Synechocystis* sp. PCC 6803	255
Limonene	0.000172 g/L	*Anabaena* sp. PCC 7120	256
(*S*)-Linalool	0.0116 g/L	*Synechocystis* sp. PCC 6803	226
Lysine	0.4 g/L	*Synechococcus* sp. PCC 7002	257
D-Mannitol	1.10 g/L	*Synechococcus* sp. PCC 7002	258
D-Mannitol	11 µM/OD$_{730}$	*Synechocystis* sp. PCC 6803	259
13R-Manoyloxide	0.002 g/L	*Synechocystis* sp. PCC 6803	206
13R-Manoyloxide	0.45 mg/g DCW	*Synechocystis* sp. PCC 6803	260
2-Methyl-1-butanol	0.2 g/L	*Synechococcus elongatus* PCC 7942	261
3-Methyl-1-butanol	0.070 g/L	*Synechococcus elongatus* PCC 7942	238
β-Phellandrene	0.000090 g/L	*Synechocystis* sp. PCC 6803	262
β-Phellandrene	0.000050 g/L	*Synechocystis* sp. PCC 6803	263
β-Phellandrene	10 mg/g DCW	*Synechocystis* sp. PCC 6803	264
β-Phellandrene	5.95 mg/g DCW	*Synechocystis* sp. PCC 6803	265
β-Phellandrene	3.2 mg/g DCW	*Synechocystis* sp. PCC 6803	266
Pigments	0.0045 g/L	*Synechocystis* sp. PCC 6803	202
Poly-3-hydroxybutyrate	0.05 g/L	*Synechococcus* sp. PCC 7002	267
1,2-Propanediol	0.150 g/L	*Synechococcus elongatus* PCC 7942	268
1,3-Propanediol	1.22 g/L	*Synechococcus elongatus* PCC 7942	269
1,3-Propanediol	0.288 g/L	*Synechococcus elongatus* PCC 7942	229
Squalene	0.0051 g/L	*Synechocystis* sp. PCC 6803	270
Squalene	11.98 mg/L/OD$_{730}$	*Synechococcus elongatus* PCC 7942	271

Table 6.1 Continued

Compound	Production*	Cyanobacteria	Reference
Squalene	4.98 mg/L/OD$_{730}$	*Synechococcus elongatus* PCC 7942	191
Squalene	0.67 mg/L/OD$_{750}$	*Synechocystis* sp. PCC 6803	272
Succinate	0.430 g/L	*Synechococcus elongatus* PCC 7942	273
Sucrose	8.7 g/L	*Synechococcus elongatus* UTEX 2973	274
Sucrose	2.69 g/L	*Synechococcus elongatus* PCC 7942	275
Sucrose	0.138 g/L	*Synechocystis* sp. PCC 6803	276
Tocopherols	0.250 mg/g DCW	*Synechocystis* sp. PCC 6803	277
Trehalose	5.7 g/L	*Synechococcus elongatus* PCC 7942	278
(+)-Valencene	0.0096 g/L	*Synechocystis* sp. PCC 6803	226
Zeaxanthin	0.95 µg/mL/OD$_{730}$	*Synechocystis* sp. PCC 6803	202

* Titers reported were measured after different days of cultivation. Different strategies were used to express the same compound. Direct comparison is not possible.

Synthetic biology toolbox in cyanobacteria

In recent years, remarkable efforts have been made to adapt molecular techniques from other organisms, characterize new genetic elements and develop new tools. This was done with the goal of improving the cyanobacterial toolbox and facilitating the cellular reprogramming to achieve higher production yields. Some of these regulatory parts and tools include promoters, ribosome binding sites, terminators, reporters, modular vectors, expression plasmids, markerless selection systems, small RNA regulatory tools, riboswitches, CRISPR/Cas systems, kinetic and genome-scale stoichiometric models[19,27].

The characterization and engineering of promoters for cyanobacteria, has been an important area of research aiming to provide more options when a promoter that meets specific requirements is needed[30]. In addition to the well-known and commonly used native constitutive promoters such as P*cpcB*, P*sbA2* and P*rnpB*[56,126,279], efforts are being made to identity other strong promoters, as it has been reviewed by Sun *et al.*[19] Some examples are the native promoters P$_{A2520}$ and P$_{A2579}$ from *Synechococcus* sp. PCC 7002, able to drive strong expression of heterologous genes[280], and nine native promoters characterized for their expression in *Synechocystis* sp. PCC 6803[61].

One strategy that has been proven to be efficient, is the optimization of existing promoters as it was observed for the truncated promoters P$_R$-P$_S$[281], P*cpc560*[282] and P*psbA2S*[68]. Other strategies that have been useful for identification or optimization of new promoters are the error prone strategy[24] and studies based on RNA deep sequencing data[280], which have been valuable resources for the identification of promoters with a wide range of activity strengths.

In **Chapter 2**, the potential of a method to construct fully segregated genomic libraries in polyploid organisms was tested to mine native promoters in *Synechocystis*. As a conceptual framework, 72 clones isolated from the promoter library were characterized by the expression of the yellow fluorescence protein (eYFP) leading to the identification of sequences with a broad range of promoter activities. The same method was used in **Chapter 3** to find a promoter able to express the foreign glyoxylate shunt enzymes in a timely manner when introduced into *Synechocystis*. These results show the utility of constructing fully segregated genomic libraries to find and to characterize native genetic

elements like promoters, especially when looking for a promoter that meet a specific requirement or that can be inducible under specific conditions.

Inducible promoters are especially valuable when the gene(s) or pathway(s) introduced in the host organism exert heavy metabolic burden or when the end products or the intermediates are toxic for the cells[19]. The inducible promoters available for cyanobacteria have been extensively reviewed by Berla *et al.*[18] and Sun *et al.*[19] A large number of them are induced by varying metal concentrations, such as iron, copper, zinc and nickel. These promoters exhibit a high dynamic range of induction but are not appropriated for large scale applications. This, due to the toxicity and cost of the inducers, whereby the use of systems inducible by sugars, macronutrients or environmental signals would be preferable for being non-toxic and economical for industrial production[22].

Promoters repressed by LacI and activated by isopropy-β-D-thiogalactoside (IPTG) like P*trc*[11] and promoters from *E. coli* repressed by TetR and induced by anhydrotetracycline (aTc) like PL03[283], have been demonstrated to be functional in cyanobacteria. Given that for cyanobacteria, *L*-arabinose is an ideal heterologous feedstock, the arabinose inducible promoter P_{BAD} has been tested in *Synechocystis* sp. PCC 6803[59] and optimized in *Synechococcus elongatus* PCC 7942[284]. Similarly, the rhamnose inducible promoter P_{rhaBAD} from *E. coli* was introduced and tested in *Synechocystis* sp. PCC 6803[60], and the vanillate induced promoter P_{van} from *Corynebacterium glutamicum* has been tested in *Synechococcus elongatus* PCC 7942[285]. Some of these promoters have been also applied to faster growing strains such as *Synechococcus* sp. PCC 7002, and *Synechococcus elongatus* UTEX 2973, although for these promising producing strains the number of inducible promoters is still limited[27].

Due to the characteristics of the cyanobacterial oxygenic photosynthesis, the production of heterologous oxygen sensitive products/enzymes could be affected. As an alternative to avoid that, two "cyanobacteria-specific" promoters, anaerobic and dark inducible, have been constructed. The oxygen responsive system carrying the promoter PO_2 was developed in *Synechocystis*, using the FNA (fumarate and nitrate reduction) system from *E. coli*. This system senses the transitions between aerobic and anaerobic conditions, allowing the induced expression of genes under dark or anaerobic conditions[286]. Another dark inducible promoter system, carrying the P*ompC* promoter, was constructed using the native dark sensing protein Cph1 from *Synechocystis* and the histidine and osmotic response system in *E. coli*, allowing induced expression of genes under dark conditions[59].

Although most of the parts developed for cyanobacteria regulate gene expression at a transcriptional level, other parts can be used to regulate gene expression at a post-transcriptional and translational levels. Ribosome binding sites (RBS) are key elements regulating gene expression in a translational level. The rate of protein production from a mRNA transcript, is determined by the strength of the RBS to recruit ribosomes for translation. The translational efficiency can be affected by the sequence and position of a particular RBS[27]. Significant efforts have been made recently to develop RBS libraries for cyanobacteria. RBS sequences from the registry of standard biological parts as well as native RBS elements have been characterized in *Synechocystis*[32,61,68]. These efforts are being combined with the development of synthetic RBSs based on *in silico* modelling tools for *Synechocystis* sp. PCC 6803[9,69,287], *Synechococcus* sp. PCC 7002[24] and *Synechococcus elongatus* PCC 7942[20,287].

In addition to the commonly used regulatory parts, important developments have been made implementing new technologies for synthetic gene control in cyanobacteria. One of these technologies are the clustered regularly short palindromic repeats (CRISPR)-mediated modification systems, using CRISPR associated protein 9 (Cas9) and CRISPR associated endonuclease from

Prevotella and *Fancisella* 1 (Cpf1/Cas12a). They have been successfully applied to edit the genome of different cyanobacteria species[87,88]. A similar technology, the CRISPR-interference systems, have been tested in limited studies but have shown the ability to repress multiple genes simultaneously[29,86,90].

Challenges to exploit the full potential of cyanobacteria as cell factories

Despite the recent efforts and advances to expand the cyanobacterial toolbox, the current tools available for genetic engineering of cyanobacteria are not yet good enough to exploit the full potential that these organisms have. There are still several challenges limiting the synthetic biology potential of cyanobacterial cell factories that need to be addressed, such as the limited modularity of the parts to be used independently as biobricks, the need for tunable regulation and predictability of the outcome, the effect of the genome context in the gene expression of an heterologous gene, and the challenges associated with gene dosage and polyploidy to create stable engineered strains.

The design of modular, standardized parts that can be combined with other genetic elements to obtain a predictable outcome, has permitted the development of complex heterotrophic genetic circuits[288]. Ideally, modular parts would be characterized in a way that allows an accurate prediction of their performance in the context of an interconnected system[289]. Yet, in reality, it is still difficult to characterize them in a way that can be fully transferable to other organisms[31]. Natural biological systems and their components are not decoupled. They perform their function in the cell, in the midst of an immense amount of interactions with other biomolecules[290]. Cross talk or unknown interactions between the natural components and other parts of the host cell such as metabolites, protein and genes, may cause a deficient performance or the failure of the system[165].

Many of the modular parts and synthetic biology tools developed for heterotrophs organisms, do not perform as robustly when they are transferred to cyanobacterial strains[56]. It has been observed that many *E. coli* promoters do not show similar characteristics when they are used in cyanobacteria[32]. This may be due to species-dependent distinctions in key features such as promoter recognition by endogenous transcription machinery or RBS sequences, causing unpredictable gene expression[33,42,56].

The efforts to improve and expand the toolbox of characterized genetic elements are crucial to enable the design of more refined circuits in cyanobacteria. However, the ability to fully predict the expression output from a particular element is still elusive. The context-specific characteristics of a construct can affect the performance of a determined combination of promoter-RBS, which reduces the modularity of promoters and RBSs. It has been observed that secondary structures can be formed between a heterologous gene. This can hamper the transcription and/or translation processes, resulting in expression variability when using an identical promoter-RBS cassette to express different genes. Because of that, during the design phase it is often difficult to predict the probable expression level that a construct will have[27].

Another challenge is that a tunable regulation of gene expression is required to optimize product yields and titers[30]. Unlike the well-defined library of promoters characterized in other model organisms, the number of constitutive and inducible promoters well characterized in cyanobacteria is rather small, and often a substantial variation is observed in the level of expression achieved for specific heterologous genes. The development of inducible genetic systems in cyanobacteria have been compromised by some limiting factors such as the toxicity of the inducers, leakiness when the inducer is not present and inducer photolability[27].

Besides the biological parts and tools to engineering an organism, a reasonably good understanding of the biological system is needed to thrive in synthetic biology applications. In the case of cyanobacteria, its metabolism remains only partially elucidated. This creates limitations for modelling as it hinders the predictability of their simulations[22]. Even though more metabolic routes and reactions are being discovered, key regulation mechanisms continue being unknown[206,291]. Recent work in metabolic modelling has provided crucial tools for studying metabolism[292,293,294,295], but more studies are necessary for further understanding of the cyanobacterial metabolism that can improve synthetic biology and metabolic engineering applications.

One of the main challenges for cyanobacterial engineering is the multiple copy of their chromosome per cell that cyanobacteria often possess[296]. Due to this, the generation of stable transformant strains results in a time consuming and labor-intensive process requiring extensive segregation to achieve homogeneity in all the copies of the chromosome. Even though homologous recombination is very efficient in cyanobacteria, it is difficult to guarantee the edition of multiple genes on the same chromosome copy. Because of that, it is challenging and sometimes impossible to use complex synthetic biology methods that have been successfully applied in other organisms such us *E. coli* and yeast[22].

There is still a need to characterize more tools in the context they are going to be used and find more genetic elements that can be tailored to the particular needs of the design. It is also necessary to develop new tools and methods that can help to circumvent the challenges limiting the engineering of cyanobacteria, speed up the process and allow high throughput screening systems.

Engineering cyanobacteria

Even though engineering cyanobacteria is laborious due to the complexity of their metabolism, the progress in synthetic biology and systems biology methods are increasing the possibilities to rationally design them[30]. Cyanobacteria are easier to engineer in comparison with other photosynthetic organisms, such as plants and algae because genetic manipulation is relatively simple[297]. For some of the genetically accessible strains, only natural transformation is required to introduce foreign DNA[18].

The introduction of foreign DNA in cyanobacteria can be done by integrating it into the genome or by using self-replicating foreign plasmids. This can be accomplished through natural transformation, using mechanisms of endogenous DNA uptake; through conjugation, using a helper *E. coli* strain; or through electroporation[298]. In other microbial hosts used for bioproduction like *E. coli* and *S. cerevisiae*, the preferred method to express heterologous genes is through self-replicating plasmids. On the contrary, in cyanobacteria the expression of heterologous genes is mainly done by gene integration into the chromosomes or into the endogenous plasmids[35].

Typically, the introduction of heterologous genes in chromosomes or endogenous plasmids is performed using locus-targeting vectors, which are integrated by homologous recombination into the loci of interest or into neutral sites. The latter are non-transcribing sequences or intergenic sites that can be modified without affecting the cell physiology[35]. Most of the self-replicative plasmids used in cyanobacteria are RSF1010 derivatives, but other chimeric shuttle vectors have been created by fusing endogenous cyanobacterial plasmids with plasmids of *E. coli* origin[56,61,117].

Episomal expression usually shows higher expression levels in comparison with chromosomal expression, because plasmids have a higher copy number[33,56,174]. This can be beneficial for bioproduction to achieve higher titers of the desired compound[115]. However, chromosomal expression grants genetic stability, which is desired for long-term cultivations and bioproduction applications. In addition, it can avoid the possibility of plasmid loss, and the uncertainty in the assessment of the performance of a determined construct due the variability of the plasmid copy number[34,35].

The cyanobacteria species used as chassis, usually have more than one copy of their chromosome in each cell, which requires full genome segregation in order to obtain a homogenous genotype when a modification is made in the chromosome or in one of the native plasmids. That represents a challenge to introduce or delete a gene and screen for mutations[44,45]. The integration of heterologous genes is usually facilitated by the use of selection markers, such as antibiotic resistance cassettes, which are necessary to select transformants. Nevertheless, the number of antibiotic markers is limited, and the accumulation of markers from integrated cassettes may inhibit cellular metabolism and be deleterious to cell viability. This, because the addition of antibiotics during cultivation is required to achieve segregation and guarantee the maintenance of the genes integrated[35,195].

Due to the above, methods for markerless integration or techniques for curing the marker are preferred, especially for up scaling applications and mass cultivation in bioreactors, or for multiloci editing (multiple gene knockins and knockouts). Markerless edition is generally done by the integration of a cassette containing a resistance marker and a killer gene, followed by the removal of that cassette and counterselection of the cells in which it has been eliminated[35]. Some of the markerless systems that have been used in cyanobacteria are the *sacB* system from *Bacillus subtilis* and the *mazF* system. The *sacB* system encodes a levansucrase that uses glucose to generate a lethal product acting as killer gene for negative selection[100,299]. The *mazF* system, employs a nickel inducible promoter for the expression of the endoribonuclease *mazF*, an inhibitor of protein synthesis acting as a counter selection marker[118].

Other methods that have been used for the construction of markerless mutants in cyanobacteria are the rps12-mediated gene replacement[299], the acrylate-*acsA* counter selection system[300], the FLP/FRT recombination system[301] and the Cre/Lox recombination system[302]. One major impediment for the application of these methods is the long time required to achieve one markerless editing (at least four weeks). The variable chromosome copy number increases the editing time, because full segregation of the introduced gene along with the selection marker in all the chromosomes is necessary to achieve a homogenous population. Also, to avoid that it can be diluted through homologous recombination by the wild type sequences of the chromosomes that were not modified[35].

Libraries in cyanobacteria

In addition of the rational design approach, which is the construction of engineered systems guide by rational design parameters, the random approach, by means of the construction of DNA libraries, has proven to be a very useful tool. The random approach can be used to study complex genetic phenotypes, discern tolerance mechanism, generate and identify mutants with desired characteristics and identify and characterized genomic parts[34]. Native genomic parts with a particular function can be identified by creating a library of genomic DNA fragments that can be characterized under different conditions with the help or a reporter gene. Sequences of the genome that possess

regulatory activity such as promoters, and that are induced under specific conditions, can be identified by screening the library using the conditions of interest.

In cyanobacteria, these libraries can be constructed in a replicative plasmid or can be integrated into the chromosome of the host. It is relatively easier to construct a library on a replicative plasmid, since the ploidy of cyanobacteria (the presence of multiple copies of the chromosome in each cell) makes the integration of the library in each chromosome a time consuming and difficult process. However, despite segregation is not necessary when using a replicative plasmid, its copy number might be even more variable than the copy number of the chromosomes[34]. This may cause uncertainty to characterize the properties of the genomic fragments (e.g. the strength of a promoter).

Native genetic elements such as promoters, are important elements for synthetic biology applications. The identification of new native promoters with different strengths and inducible properties would be of great utility to expand the cyanobacterial toolbox. The recent work focused on the development and characterization of promoter libraries in cyanobacteria, has been done mainly generating and characterizing variants of existing promoters such as λPL, Ptrc, P*psbA*, P*cpcB* and P*rbcl*. These libraries were constructed in replicative plasmids by conjugation[48,61,283] or they were integrated in the host chromosome by homologous recombination[20,24,112].

The libraries constructed in the chromosome would have the advantage of being more stable and have a relatively constant copy number in comparison with the libraries constructed in a replicative plasmid. However, the polyploidy of cyanobacteria represents a major challenge. It is necessary to track the insertion of the library, chromosome by chromosome and it creates a limitation for the construction of large genomic libraries.

In **Chapter 2** of this thesis, a method is presented to construct fully segregated genomic libraries in polyploid organisms using a counter-selection approach[303]. This method, which combines positive (through the use of antibiotic) and negative selection (through the expression of a toxic gene), enables only the cells where the library is fully integrated to be able to survive. As a proof of concept of the functionality of the method, 72 native sequences of *Synechocystis* with promoter activity were characterized, showing a wide range of strength. This method can be applied to other polyploid organism to identify genes or genetic elements for fundamental research or biotechnological applications.

Genomic libraries can also be used to find genetic elements when specific characteristics are required. The success of engineering cyanobacterial systems relies on the capacity to create stable strains and achieve the fine-tuning necessary to express the genes of interest under the specific conditions needed. The metabolism can be modified adding or removing parts of a metabolic pathway, but the expression levels and the timing of expression may be critical for metabolic regulation. In some cases, the genetic elements required for this fine-tuned regulation are not available. It is possible that regulatory elements that can perform in a desired manner already exist in the genome of the host organism or in a closely related one.

In **Chapter 3**, the method developed to construct genomic libraries in polyploid organisms was applied to find a promoter able to express the heterologous glyoxylate shunt enzymes in a timely manner. This with the aim of increasing the night-time production of fumarate in the strain *Synechocystis* Δ*fumC* Δ*zwf* (double knockout for the *fumC* and *zwf* genes). In an attempt to reconnect the TCA cycle that was disrupted by the knockout of the *fumC* gene, the malate synthase and

isocitrate lyase enzymes from *Chlorogloeopsis fritschii* where introduced. These enzymes should be expressed exclusively during the night, because day expression affects the growth rate. To find a suitable promoter that could trigger the expression of the enzymes with the desired pattern, a genomic library was constructed and screened under day and light regime.

Even though an increment in the night-time production of fumarate was observed, indicating the expression of the glyoxylate shunt enzymes during that period, it was not possible to obtain the timely activation needed to create a stable strain. Over time, the cells suffered different deletions, leading to an impaired transcription and/or translation of the enzymes. This suggests that the strength or the timing of expression of the enzymes reduced the fitness of the cells, causing genetic instability. Despite the impossibility to find a promoter with the desired characteristics in *Synechocystis*, these results show the applicability of the method developed to find new genetic elements. Especially, when a specific performance is needed and the tools available are not suitable for that task.

Genetic instability

The modification of an organism, or its genetic information, might alter the stability of the genome, or the functionality of the organism due to the loss of fitness. It has been observed that the redirection of the metabolism of cyanobacteria can result in an unexpected imbalance of the homeostasis of the cell. That leads to the occurrence of mutations that impair the functionality of the enzymes and reduce the product formation[128].

Obtaining high titers in producing strains depends mostly on sustained microbial productivity during rather long cultivation periods. This requires higher number of cell divisions, increasing the risk of acquiring genetic lesions. The individuals with higher fitness will eventually be selected over the rest, dominating the population. Because of that, populations can acquire mutants that hamper the productivity, to relieve to some degree the product formation burden. The cells with a lower burden will generally grow faster, outcompeting the slower growing producing cells. As a result, over time the population will be dominated by non-producing, yet faster growing, cells. This represents a major challenge in the design and development of (cyanobacterial) cell factories.

Genetic instability of production strains has been reported in different cyanobacterial species, such as *Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942 and *Synechococcus* sp. PCC 7002, when those strains were engineered for the production of many different compounds such as ethylene, lactic acid, isopropanol, mannitol and ethanol[128,147]. In revertant cells, suppressor mutations were often found in the coding region of the heterologous gene(s) introduced. These mutations ranged from point mutations, to small and large deletions and insertions. Many of them, alter an active site residue or lead to a truncated protein, caused by frame shift or by the introduction of stop codons[120]. This results in severe impairment of the functionality of the enzyme(s), reducing or completely suppressing product formation.

The expression of a heterologous production pathway may affect various cellular processes. Some of these processes are: the levels of transcription and/or translation due to competition for RNA polymerases or ribosomes[304], the availability of substrates[305], the capacity of cellular transport due to potential limitations in membrane occupancy[306] and the ability to maintain metabolites in homeostatic levels, which may be toxic for the cell if they are accumulated at high levels[307].

Given that the phenotypic instability of the producing strains appears from the occurrence of genetic mutations, it is of great importance to understand how these changes in the genome occur[308]. Although not much research about this topic has been done directly in cyanobacteria, the mechanism of generation of mutations has been studied in model organisms like *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas aeruginosa*, which share the basic replication machinery.

Currently, there are six known common sources of genetic variation in prokaryotes:

I. DNA replication errors: the error rate per base pair is about 10^{-10} per generation in *E. coli*, nearly one mutation in the genome in the course of 2000 cell divisions[308]. However, due to the very high number of cells present in a population, around half million cells will carry a mutation in only 1 mL of culture (OD_{600} = 1), even after one division[309]. If the change in a single base results in an amino acid substitution, this is sufficient to cause the possible inactivation of the encoded protein[310].

II. Homologous recombination errors: this generally happens between two homologous regions (neighboring long direct repeats), where DNA sequences between repeats can be either duplicated or deleted, depending on how the homologous recombination occurs. The repeated use of biological parts such as promoters and terminators, can increase the frequency of occurrence of these errors, affecting the functionality of the expressed constructs[311].

III. DNA replication slippage on simple sequence repeats (SSR) errors: DNA polymerase may temporally stall during DNA replication and then realign the nascent strand to the template[312,313]. It is very probable that misalignments happen at sites with SSRs, causing duplication or deletions of the sequences between them.

IV. Mobile genetic elements errors: these mobile elements (e.g. transposons, lysogenic phages) are DNA sequences that can integrate into a new site of their genome through transposition[314]. This can result in gene insertional inactivation, duplications, deletions and inversions[315].

V. Plasmid loss: the variability of plasmid copy number affects the target gene expression level, when it is expressed on a self-replicative plasmid, which could vary the productivity. In addition, plasmid loss is a key factor contributing to limited expression of heterologous genes[316], especially in the case of high copy number plasmids, due to the large metabolic burden.

VI. Stress responsive error prone DNA replication/repair: during growth-limited stress conditions such as hypoxia, starvation or antibiotics[317], errors could be introduced through error-prone DNA replication/repair. These stress conditions are very common under long term cultivation, and/or the accumulation of toxic bioproducts. The generation and selection of fitter mutants could affect the target product formation.

Instability can also occur during DNA methylation at specific sequences[318]. Methylation generates mismatches for base pairing, converting methylated bases in hotspots for mutations[319]. Basically, the instability of producing strains, is mainly caused by genome instability determined by Darwinian selection[320]. The rate at which these mutations manifest as strain instability, is determined by many factors such as the mode of cultivation, how often the revertant (non-producing) cells appear in the population and how revertant cells behave in comparison with the producing strain.

Some strategies have been proposed to achieve stable producing strains over a continued period of time:

I. Enhancing microbial genome stability: mutations in the heterologous pathway(s) impairing productivity could be theoretically prevented by reducing the host mutation rate, through the elimination of mechanism that can lead to mutations. Some of the options are the removal of *recA*, which encodes a key component of the homologous recombination machinery that prevents gene inactivation by homologous recombination[321]; deletion of mobile elements from the genome[322]; and deletion of the polymerases mediating error-prone DNA replication caused by stress conditions[323].

In addition to minimize the mutation rates in the host, another approach would be to avoid the presence of sequences prone to mutation in the heterologous synthetic construct. However, reducing the number of mutations to a neglectable level is difficult due to the average size of a microbial population. In the case of *Synechocystis* in an industrial set up, it is hard to fully prevent the emergence of revertant cells, therefore, the scope of this strategy is very limited.

II. Coupling growth and product formation: heterologous production pathways could be designed in such a way that product formation becomes beneficial or mandatory for the growth of the cells. Product formation can be coupled to catabolism or anabolism processes. Products that are generated as (by)products of anabolism, generally can be redirected into intermediary metabolism. However, if the enzymes involved in this assimilation are knocked out, a stoichiometric coupling between product formation and anabolism is possible.

Computational tools have been used to find candidate compounds in which the production is stoichiometrically coupled to cell growth. One of them is the algorithm FRUITS (Find Reactions Usable In Tapping Side Products). This algorithm analyzes existing genome scale metabolic models to identify side products of anabolism, whose production can be coupled by deletion of their re-utilization/assimilation pathway(s), leading to production being strictly coupled to biomass synthesis[125].

III. Reducing the burden of product formation: producing cells would be less easily outcompeted by the revertant cells if the size of the burden could be manipulated. This is possible by expressing the genes encoding a product-forming pathway under inducible control. The genes can be induced only after growth has ceased. This avoids direct competition for resources between product formation and growth, reducing the risk that non-producing mutants overgrown the producing cells. Inducible promoters and genetic regulation systems based on transcription factors can be used for this purpose[324].

Signal transduction systems based on transcription factors, usually contain a sensory input module, various control devices and an output module (able to modulate transcription). Their use is preferable because they allow flux adjustments in a dynamic mode, to manage the trade-offs between product formation and growth[325,326]. Flexible inducible regulation systems can be useful to cope with the metabolic burden imposed to the microorganisms, potentially improving strain stability. Still, the limited availability of sensors responsible to metabolites, defy their wide applications and therefore, new approaches to design or identify new sensor are needed[327,328,329,330].

The development of new sensors not only contributes to the construction of inducible regulation systems that can increase product formation and improve strain stability[331,332,333,334,335,336], they can also be a useful tool to construct cell factories that can detect and regulate their own production. In cyanobacteria, a few sensors have been constructed responding to physical, chemical or metabolic state signals such as light[59,337], oxygen[286], nitrate, and L-arabinose levels[59].

Chapter 4 describes the construction of nine variants of a fumarate biosensor based on the chimeric two-component system DcuSZ/OmpR developed by Ganesh *et al*. in *E. coli*[134]. This detection system was made fusing the fumarate sensor DcuS and the osmolarity sensor EnvZ (DcuSZ) and coupling the expression of eYFP in response to extracellular fumarate. By changing the length of the linker between the DcuS and EnvZ functional domains, it was tested if the variants could have different signal dynamic ranges and different detection limits for the detection of fumarate. Three of the nine variants, DcuSZ 1, 6, 9, behaved as functional biosensors, showing a combined signal dynamic range of 0.1-20 mM fumarate in *E. coli*, being DcuSZ 6 the variant with higher potential to be used in *Synechocystis*.

These sensors along with other biological parts can be combined to create complex devices able to control gene expression at different levels (e.g. transcription and translation). These devices could be used to fine-tune the expression of genes of interest and achieve the precise control necessary in synthetic biology[30]. Multipart devices have been designed and constructed for *Anabaena* sp. PCC 7120[338], *Synechococcus* sp. PCC 7002[63], *Synechococcus elongatus* PCC 7942[285] and *Synechocystis* sp. PCC 6803[59,286].

Role of chromosome location in genetic functionality

The expression of heterologous genes is necessary to implement synthetic metabolic pathways for production in cyanobacteria. This is usually achieved by integrating the heterologous genes into the chromosomes or into one of the endogenous plasmids through homologous recombination. The integration sites commonly used are called "neutral sites". They are loci that are typically intergenic sites or non-transcribing sequences that can be disrupted without affecting the cell physiology, viability, or causing any detectable phenotype[35,175].

Various neutral sites have been identified and utilized for the expression of heterologous genes in cyanobacteria, including five sites regularly used in *Synechococcus*, and three commonly used in *Synechocystis*, *slr0168, slr2030-31* and *psbA1*[35]. However, it is not well known yet how the gene context and the location of insertion of a heterologous gene affect its expression in cyanobacteria, and if it is possible to achieve the independence required for the biological parts to be modular. The selection of the integration site is relevant in terms of stability and functionality of the insert and it is an important consideration to prevent affecting cell viability. Pinto and colleagues carried out a systematic study to identify and characterize neutral sites in the chromosome of *Synechocystis.* Their work showed that from all the ORFs (open reading frame) initially considered, putatively encoding hypothetical or unknown proteins, many of them were not suitable due to the possibility of non-coding regions possessing regulatory function through *cis* acting sites[175].

Besides the regulatory role of the surrounding regions to the insertion site, other factors can affect the expression of heterologous genes. Secondary structures can be formed between the surrounding regions, altering transcription. The proximity of the inserted gene to the origin of replication, might cause differences in gene dosage as it has been observed in other prokaryotes[168,169,170,171]. The variation in gene dosage has been observed in fast growing organisms, because at high growth rates multiple rounds of chromosome replication are initiated. Because of this, genes located closer to the origin of replication will be replicated before the genes located close to the terminus. More information is needed to determine if a slow growth organism like *Synechocystis* presents gene dosage variations related with the proximity of the origin of replication.

In **Chapter 5**, an expression cassette consisting of eYFP protein under the control of the strong promoter P*cpc*, was introduced in six different loci located on different regions of the chromosome of *Synechocystis*. This with the goal of evaluating if the gene expression was affected by the proximity to the origin of replication or by the surrounding regions next to the inserted sequence. The results did not show big differences in the expression of eYFP when inserted in different loci of the chromosome. This suggests that slow growing organisms like *Synechocystis* might not have gene dosage variations associated to multiple rounds of chromosomal replication, observed in fast growing organisms. At least for these six *loci*, expression levels are not very much affected by chromosomal location *per se*, confirming the usefulness of these *loci* as neutral sites for the insertion of heterologous genes.

Concluding remarks

The characteristics that cyanobacteria possess make them a suitable chassis for biotechnological applications. It has been shown that cyanobacteria have the capacity to produce compounds of high value. However, the tools designed for heterotrophic organisms like *Escherichia coli* and *Saccharomyces cerevisiae* are not always transferable to cyanobacteria. Even though specific tools have been designed, more refined tools are still necessary to achieve the fine-tuning required to engineer cyanobacteria. This has caused a delay in their development as an industrial host. The design and development of new genetic tools would significantly contribute to exploit the full potential of cyanobacteria as cell factories for bioproduction, to make it an economically viable approach for up scaling processes[27].

Many recent studies have been focused on the characterization of biological parts, including expression vectors, promoters, ribosome binding sites and terminators. However, it is still necessary the identification of new parts and the development of methodologies that facilitate the engineering of cyanobacteria. The construction of genomic libraries could speed up the identification and characterization of biological parts. Furthermore, the construction of genetic libraries (knockout, expression and genomic libraries) could improve the understanding of the cyanobacterial metabolism. Other techniques that could improve and expand the efficiency to modify the cyanobacterial genome are the markerless modification systems, CRISPR/Cas based genome editing and high throughput methods, which would allow multiple genome modifications and the assembly of complex metabolic pathways, necessary for advanced engineering projects[27,34,88,339].

The work presented in this thesis, contributes with new tools to expand the synthetic biology toolbox in *Synechocystis*, with the potential to be applied to other cyanobacteria species. A strategy to develop fully segregated libraries in polyploid organisms was described and validated by using it to find new native promoters in *Synechocystis*. This method was also used to find a promoter able to express heterologous enzymes in *Synechocystis* in a timely manner, since a promoter with the desired expression pattern has not yet been described. This showed the potential of this method to mine genetic elements performing under specific conditions. A detection system to detect fumarate was also described, and the effect that the chromosome insertion site has on the expression of heterologous genes in *Synechocystis* was studied through an expression cassette inserted on different loci of the chromosome.

Despite the progress made in the last years, there are still some challenges that need to be tackled in order to construct efficient and stable cyanobacterial cell factories. Further improvements in the genetic and metabolic engineering of cyanobacteria for compounds production are required to transform these processes into economically viable alternatives for industrial bioproduction. In addition to increasing product yields, more efforts are necessary to develop and optimize cyanobacterial strains adaptable to large scale cultivation systems. Further investigation towards a better understanding of the cyanobacterial metabolism, and the development of more sophisticated techniques and resources should significantly increase our knowledge and capacity to use cyanobacteria as a sustainable biotechnological platform.

Author contributions PCB, WD, FBS and KH wrote the original manuscript that was published; PCB used it as a base to write this chapter; FBS reviewed it.

Summary

One of the main challenges faced by human societies, lies in the balancing of human activities with those that occur around it naturally on Earth. The increasing accumulation of atmospheric CO_2 for nearly 200 years is posing serious environmental threats such as global warming and ocean acidification that have caused alarming consequences. One promising biotechnological solution to counter this is the usage of cyanobacteria as green cell factories. These hosts could act as chassis in which additional (enzymatic) components can be added, resulting in the direct conversion of CO_2 into products of interest, fuelled by (sun) light, and having O_2 as the only by-product.

Synechocystis sp. PCC 6803, a model microorganism of the phylum cyanobacteria, has been widely studied as a sustainable alternative for biotechnological production of different compounds. Extensive molecular manipulation is required to construct cell factories, but the tools available for *Synechocystis* and other species of cyanobacteria are limited in comparison with the tools developed for chemoheterotrophic model organisms such as *Escherichia coli* and S*accharomyces cerevisiae.* The current tools used for cyanobacteria have been to a great extend inherited from the toolbox developed from these organisms. These tools are not always adequate for cyanobacteria, and consequently perform less optimally.

In order to exploit the full potential of cyanobacterial cell factories and use them for industrial biotechnological applications, it is necessary to expand the synthetic biology toolbox of cyanobacteria. This thesis describes the development of synthetic biology techniques that allow controlled expression of target pathways in cyanobacterial cell factories and that can speed up the engineering of these strains.

Chapter 1: Introduction, describes the background and main questions that this research work attempts to address. This chapter provides an overview of how synthetic biology has been applied to improve biotechnological applications and how it has been used to engineer cyanobacteria. We describe the current tools and standardized parts available for cyanobacteria and the challenges of working with these organisms due to their particular biological characteristics. We discuss how the development of new synthetic biology tools can facilitate the engineering process and tackle some of the challenges described.

Chapter 2: **Construction of fully segregated genomic libraries in polyploid organisms such as *Synechocystis* sp. PCC 6803**, describes the development of a novel method to construct genomic libraries in organisms that, like *Synechocystis*, have multiple copies of their chromosome in each cell. This method assures the full segregation of the inserted DNA in all its copies, without the need to track the insertion chromosome by chromosome. As a proof of concept, the method was applied to construct a promoter library for the characterization of *Synechocystis* native sequences showing promoter activity that could be used to regulate gene expression. The library was screened by the expression of eYFP in the cells containing the inserted DNA fragments. The fluorescence was measured using FACS, and a fraction of the cells emitting fluorescence signal were characterized, giving as a result 72 unique sequences with promoter activity, showing a broad range of protein (eYFP) expression.

Chapter 3: Construction of a promoter library in *Synechocystis* sp. PCC 6803 to increase the night production of fumarate, describes how the method for creating fully segregated libraries in polyploid organisms was used to find a promoter able to express the enzymes of a heterologous glyoxylate

shunt in a timely manner. This, to attempt the overproduction of fumarate at night. *Synechocystis* can be engineered to produce fumarate during the day by removing the *fumC* gene. Night-time production is additionally possible by deleting the z*wf* gene, which diverges flux from the pentose phosphate pathway to the TCA cycle, leading to an increase in the production of fumarate also during the night. The incorporation of the glyoxylate shunt enzymes from *Chlorogloeopsis fritschii* can increase the night-time production of fumarate even further without a negative effect on fitness provided that they are expressed exclusively during the night. The library was characterized by natural selection under continuous cultivation using a day/night regime. Over time, the cells selected against the expression of the shunt enzymes, due to the genetic instability caused by its un-timing expression. This suggests that *Synechocystis* does not have promoters that display the desired transcription profile in timing and intensity of expression, or the lifetime of the shunt enzymes used are not compatible with the desired expression profile.

Chapter 4: **Construction of a family of biosensors to detect fumarate applicable across different clades of microorganisms**, describes a novel approach to construct biosensors that allows the detection and selection of individuals displaying better production characteristics in a fumarate production population. We describe the construction of variants of a fumarate biosensor, based on the DcuSZ/OmpR chimeric two-component system developed in *E. coli*. This system was constructed by fusing two native *E. coli* sensor histidine kinases, sensor DcuS and osmolarity sensor EnvZ (DcuSZ), and coupling the system to the expression of eYFP in response to extracellular fumarate. Nine variants of DcuSZ were constructed in *E. coli* by changing the linker length region between the DcuS and EnvZ functional domains. Three of these variants, DcuSZ 1, 6, 9 behaved as functional biosensors, able to detect fumarate concentrations ranging from 0.1 to 20 mM. The variant DcuSZ 6 was integrated in *Synechocystis*, being the best candidate to have functional biosensor activity in this organism.

Chapter 5: **Is gene expression in *Synechocystis* sp. PCC 6803 influenced by genomic location?**, analyzes if the genomic location of the insertion site of a heterologous gene, may lead to differential gene expression. One of the problems for the expression of heterologous genes is that expression may be affected by factors such as the regulatory role of the surrounding nucleotides to the gene of interest, the chromosomal three-dimensional structure, or the proximity to the origin of replication. Some loci of the chromosome of *Synechocystis* sp. PCC 6803 have been identified and validated as neutral sites, loci that can be disrupted without affecting the viability or physiology of the cell. In this chapter, we compared the expression of a reporter cassette, carrying the eYFP protein under the control of the P*cpc* promoter, when it was inserted on six different locations of the chromosome (neutral sites). The goal was to determine if the site of integration affects the strength of the promoter. The results did not show a significant difference in the expression of the eYFP protein regarding its insertion site, supporting the idea that in organisms with slow growth like *Synechocystis*, and for at least these six loci, the levels of expression are not very much affected by chromosomal location *per se*.

Chapter 6: General discussion, outlines the main findings of this thesis and discusses the challenges for the application of synthetic biology for the synthesis of products in cyanobacterial cell factories. We discuss the potential applications of the tools developed in this thesis and suggest some strategies that could be used to create more stable strains. We review the recent advances in cyanobacterial engineering and identify areas where further research is necessary to achieve the full potential of cyanobacterial cell factories on an industrial scale.

Samenvatting

Een van de grootste uitdagingen voor de mensheid is het vinden van een balans tussen de activeiten van de mens en de natuurlijke processen van onze planeet. De toenemende ophoping van atmosferische koolstofdioxide (CO_2) over bijna 200 jaar vormt serieuze bedreigingen voor het milieu, zoals de opwarming van de aarde en de verzuuring van de oceanen. Een veelbelovende biotechnologische oplossing om dit tegen te gaan is het gebruik van cyanobacteriën (blauwalg) als groene fabrieken. De blauwalg kan gebruikt worden als basis, waar we (enzymatische) componenten aan toe kunnen voegen, wat leidt tot de directe omzetting van CO_2 in nuttige eindproducten met (zon)licht als brandstof en zuurstof als enige bijproduct.

Synechocystis sp. PCC 6803, een blauwalg en modelorganisme, is veel bestudeerd als duurzaam alternatief voor de biotechnologische productie van allerlei chemische stoffen. Uitgebreide moleculaire modificaties zijn nodig om de fabriekjes op te bouwen, maar de gereedschappen die beschikbaar zijn voor de manipulatie van blauwalgen zijn beperkt in vergelijking met die van chemoheterotrofe modelorganismen zoals *Escherichia coli* en *Saccharomyces cerevisiae*. De gereedschappen die nu gebruikt worden voor blauwalgen zijn voor een groot deel overgenomen van deze modelorganismen, maar zijn niet altijd even geschikt.

Om het uiterste uit deze blauwalgfabrieken te kunnen halen en om ze toe te passen voor industriële biotechnologisce doeleinden, is het nodig om de synthetisch-biologische-gereedschapskist voor blauwalgen uit te breiden. In dit proefschrift hebben wij synthetische-biologie technieken ontwikkeld waardoor wij in staat zijn om de activiteit van specifieke metabole routes te reguleren in blauwalgen die de ontwikkeling van deze cellijnen zal versnellen.

Hoofdstuk 1: De introductie beschrijft de achtergrond van de hoofdvragen die in dit onderzoek worden behandeld. Dit hoofdstuk biedt een overzicht van hoe synthetische biologie is toegepast om biotechnologische toepassingen te verbeteren en hoe het is gebruikt om blauwalgen the manipuleren. We beschrijven de huidige gereedschappen en de gestandaardiseerde onderdelen die beschikbaar zijn voor blauwalgen en de uitdagingen van het werken met deze organismen en hun bijzondere biologische eigenschappen. We bespreken hoe de ontwikkeling van nieuwe synthetische-biologie gereedschappen het ontwikkelingsproces kan faciliteren en bespreken we enkele uitdagingen.

Hoofstuk 2: De ontwikkeling van volledig gesegregeerde genetische bibliotheken in polyploïde organismen zoals *Synechocystis* sp. PCC 6803, beschrijft een nieuwe methode voor het genereren van genomische bibliotheken in organismen, zoals *Synechocystis*, die meerdere kopieën van hun chromosomen hebben in elke cel. Deze methode verzekerd een volledige opname van de ingevoegde DNA in alle kopieën, zonder de noodzaak om de invoeging per chromosoom te volgen. Deze methode is toegepast als 'proof-of-concept' voor het maken van een promotorbibliotheek om promotor-activiteit vertonende DNA-fragmenten uit *Synechocystis* te karakteriseren, zodat ze gebruikt kunnen worden om gen uitting te reguleren. De bibliotheek werd doorzocht op uitting van eYFP in cellen die de toegevoegde DNA-fragmenten bevatten. De fluorscentie van eYFP werd met de FACS gemeten en

verdere bestudering van een deel van de fluorescerende cellen, met een breed fluorescentiebereik van eYFP, resulteerde in 72 unieke sequenties met promotoractiveit.

Hoofdstuk 3: De bouw van een promotorbibliotheek in *Synechocystis* sp. PCC 6803 om nachtelijke fumaraatproductie te verhogen, beschrijft hoe de in hoofdstuk 2 ontwikkelde methode is gebruikt om een promotor te vinden die in staat is om tijdig de enzymen van een heterologe glyoxylaatshunt tot uiting te brengen, om zo de overproductie van fumaraat 's nachts op gang te brengen. Fumaraatproductie gedurende de dag kan op gang gebracht worden in *Synechocystis* door het verwijderen van het *fumC* gen. Nachtelijke productie is vervolgens mogelijk door het verwijderen van het *zwf* gen, dat activiteit omleid van de pentosefosfaatcascade naar de citroenzuurcyclus. Het incorporeren van de glyoxylaatshuntenzymen van *Chlorogloeopsis fritschii* kan de nachtelijke fumaraatproductie verder verhogen zonder negatieve bijeffecten op de fitness, mits de enzymen enkel 's nachts tot uitdrukking komen. De bibliotheek werd gekarakteriseerd door middel van natuurlijke selectie onder continue cultivering met een dag/nacht ritme. Op den duur trad er selectie op tegen de expressie van de shuntenzymen vanwege de expressie op ongewenste tijden en de genetische instabiliteit dat hierdoor werd veroorzaakt. Dit suggereert dat *Synechocystis* niet beschikt over promotoren met het geschikte transcriptie profiel in termen van timing en intensiteit, óf de levensduur van de gebruikte shuntenzymen is niet verenigbaar met het gewenste expressieprofiel.

Hoofdstuk 4: Het maken van een familie van biosensoren om fumaraat waar te nemen die toepasbaar zijn op verschillende claden van microorganismen, beschrijft een vernieuwende aanpak om biosensoren te genereren die de detectie en selectie van individuen met betere productiekarakteristieken toestaat in een fumaraatproducerende populatie. We beschrijven het genereren van fumaraatbiosensorvarianten, gebaseerd op het DcuSZ/OmpR chimere tweecomponentensysteem, ontwikkeld in *E. coli*. Dit systeem werd samengesteld door de fusie van twee *E. coli*-eigen histidine kinases, sensor DcuS en osmolariteitssensor EnvZ (DcuSZ), en werd gekoppeld aan de expressie van eYFP in reactie op extracelulaire fumaraat. Negen DcuSZ varianten werden in *E. coli* gegenereerd door de 'linker'-regio lengte tussen de functionele domeinen van DcuS en EnvZ te veranderen. Drie van deze varianten, DcuSZ 1, 3 en 9, bleken functionele biosensoren, in staat om fumaraatconcentraties van 0,1 tot 20 mM waar te nemen. Variant DcuSZ 6 was de beste candidaat met functionele biosensoractiviteit en werd in *Synechocystis* geïntegreerd.

Hoofdstuk 5: Wordt genexpressie in *Synechocystis* sp. PCC 6803 door genetische lokalisatie beïnvloed?, analyseert of de insertiepositie van een heteroloog gen invloed heeft op de genexpressie. Een van de problemen voor de expressive van heterologe genen is dat de expressie beïnvloed kan worden door factoren zoals de regulerende rol van omliggende nucleotiden, de driedimensionale structuur van de chromosomen of de nabijheid van de 'origin of replication'. Sommige *loci* van *Synechocystis* sp. PCC 6803 zijn geïdentificeerd en gevalideerd als 'neutral sites', plekken die verstoord kunnen worden zonder levensvatbaarheid of fysiologie van de cel aan te tasten. In dit hoofdstuk vergeleken we de expressie van een 'reporter' gencassette, met eYFP en de P*cpc* promotor, geïntegreerd in zes verschillende 'neutral sites' van het chromosoom. Ons doel was om te bepalen of de locatie van integratie de sterkte van de promotor beïnvloedde. De resultaten toonden geen significante verschillen in expressie van het eYFP eiwit. Dit ondersteunde het idee dat in langzaam

groeiende organismen zoals *Synechocystis*, en tenminste voor de zes geteste locaties, het expressieniveau niet per se door genetische lokalisatie wordt beïnvloed.

Hoofdstuk 6: De algemene discussie zet de voornaamste bevindingen van dit proefschrift uiteen en bespreekt de uitdagingen voor het toepassen van synthetische biologie voor de vorming van producten in blauwalg celfabrieken. We bespreken de mogelijke toepassing van de in dit proefschrift ontwikkelde gereedschappen en stellen strategieën voor die behulpzaam kunnen zijn in het tot stand brengen van stabielere cellijnen. Tenslotte evalueren we de recente vooruitgang in de manipulatie van blauwalgen en identificeren we gebieden die verder onderzoek behoeven om de volledige potentiaal van blauwalgfabrieken te benutten op een industriële schaal.

References

(1) Jungblut, A. D., Hawes, I., Mountfort, D., Hitzfeld, B., Dietrich, D. R., Burns, B. P., and Neilan, B. A. (2005) Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ. Microbiol. 7*, 519–529.

(2) Potts, M., and Whitton, B. A. (2012) Introduction to the Cyanobacteria, in *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (Whitton, B. A., Ed.) 1st ed., pp 1–13. Springer Netherlands.

(3) Blankenship, R. E., and Hartman, H. (1998) The origin and evolution of oxygenic photosynthesis. *Trends Biochem. Sci. 23*, 94–97.

(4) Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science (80-.). 281*, 237–240.

(5) Stal, L. J., and Moezelaar, R. (1997) Fermentation in cyanobacteria. *FEMS Microbiol. Rev. 21*, 179–211.

(6) Angermayr, S. A., Gorchs Rovira, A., and Hellingwerf, K. J. (2015) Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol. 33*, 352–361.

(7) Zhou, J., Zhu, T., Cai, Z., and Li, Y. (2016) From cyanochemicals to cyanofactories: A review and perspective. *Microb. Cell Fact. 15*, 1–9.

(8) Gao, Z., Zhao, H., Li, Z., Tan, X., and Lu, X. (2012) Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy Environ. Sci. 5*, 9857–9865.

(9) Xiong, W., Morgan, J. A., Ungerer, J., Wang, B., Maness, P. C., and Yu, J. (2015) The plasticity of cyanobacterial metabolism supports direct CO_2 conversion to ethylene. *Nat. Plants 1*, 1–6.

(10) Kusakabe, T., Tatsuke, T., Tsuruno, K., Hirokawa, Y., Atsumi, S., Liao, J. C., and Hanai, T. (2013) Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metab. Eng. 20*, 101–108.

(11) Atsumi, S., Higashide, W., and Liao, J. C. (2009) Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol. 27*, 1177–1180.

(12) Oliver, J. W. K., and Atsumi, S. (2014) Metabolic design for cyanobacterial chemical synthesis. *Photosynth. Res. 120*, 249–261.

(13) Dittmann, E., Gugger, M., Sivonen, K., and Fewer, D. P. (2015) Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends Microbiol. 23*, 642–652.

(14) Gao, X., Sun, T., Pei, G., Chen, L., and Zhang, W. (2016) Cyanobacterial chassis engineering for enhancing production of biofuels and chemicals. *Appl. Microbiol. Biotechnol. 100*, 3401–3413.

(15) Oliver, N. J., Rabinovitch-Deere, C. A., Carroll, A. L., Nozzi, N. E., Case, A. E., and Atsumi, S. (2016) Cyanobacterial metabolic engineering for biofuel and chemical production. *Curr. Opin. Chem. Biol. 35*, 43–50.

(16) Luan, G., and Lu, X. (2018) Tailoring cyanobacterial cell factory for improved industrial properties. *Biotechnol. Adv. 36*, 430–442.

(17) Knoot, C. J., Ungerer, J., Wangikar, P. P., and Pakrasi, H. B. (2018) Cyanobacteria: Promising biocatalysts for sustainable chemical production. *J. Biol. Chem. 293*, 5044–5052.

(18) Berla, B. M., Saha, R., Immethun, C. M., Maranas, C. D., Moon, T. S., and Pakrasi, H. B. (2013) Synthetic biology of cyanobacteria: Unique challenges and opportunities. *Front. Microbiol. 4*, 1–14.

(19) Sun, T., Li, S., Song, X., Diao, J., Chen, L., and Zhang, W. (2018) Toolboxes for cyanobacteria: Recent advances

and future direction. *Biotechnol. Adv. 36*, 1293–1307.

(20) Wang, B., Eckert, C., Maness, P. C., and Yu, J. (2018) A Genetic Toolbox for Modulating the Expression of Heterologous Genes in the Cyanobacterium Synechocystis sp. PCC 6803. *ACS Synth. Biol. 7*, 276–286.

(21) Sengupta, A., Sunder, A. V., Sohoni, S. V., and Wangikar, P. P. (2019) Fine-Tuning Native Promoters of Synechococcus elongatus PCC 7942 to Develop a Synthetic Toolbox for Heterologous Protein Expression. *ACS Synth. Biol. 8*, 1219–1223.

(22) Vavitsas, K., Crozet, P., Vinde, M. H., Davies, F., Lemaire, S. D., and Vickers, C. E. (2019) The synthetic biology toolkit for photosynthetic microorganisms. *Plant Physiol. 181*, 14–27.

(23) Higo, A., Isu, A., Fukaya, Y., and Hisabori, T. (2018) Spatio-Temporal Gene Induction Systems in the Heterocyst-Forming Multicellular Cyanobacterium Anabaena sp. PCC 7120. *Plant Cell Physiol. 59*, 82–89.

(24) Markley, A. L., Begemann, M. B., Clarke, R. E., Gordon, G. C., and Pfleger, B. F. (2015) Synthetic Biology Toolbox for Controlling Gene Expression in the Cyanobacterium Synechococcus sp. strain PCC 7002. *ACS Synth. Biol. 4*, 595–603.

(25) Li, S., Sun, T., Xu, C., Chen, L., and Zhang, W. (2018) Development and optimization of genetic toolboxes for a fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. *Metab. Eng. 48*, 163–174.

(26) Fujisawa, T., Narikawa, R., Maeda, S. I., Watanabe, S., Kanesaki, Y., Kobayashi, K., Nomata, J., Hanaoka, M., Watanabe, M., Ehira, S., Suzuki, E., Awai, K., and Nakamura, Y. (2017) CyanoBase: A large-scale update on its 20th anniversary. *Nucleic Acids Res. 45*, D551–D554.

(27) Santos-Merino, M., Singh, A. K., and Ducat, D. C. (2019) New applications of synthetic biology tools for cyanobacterial metabolic engineering. *Front. Bioeng. Biotechnol. 7*, 1–24.

(28) Cheng, A., and Lu, T. K. (2012) Synthetic Biology: An Emerging Engineering Discipline. *Annu. Rev. Biomed. Eng. 14*, 155–178.

(29) Gordon, G. C., Korosh, T. C., Cameron, J. C., Markley, A. L., Begemann, M. B., and Pfleger, B. F. (2016) CRISPR interference as a titratable, trans-acting regulatory tool for metabolic engineering in the cyanobacterium Synechococcus sp. strain PCC 7002. *Metab. Eng. 38*, 170–179.

(30) Immethun, C. M., and Moon, T. S. (2018) Synthetic gene regulation in cyanobacteria, in *Synthetic Biology of Cyanobacteria. Advances in Experimental Medicine and Biology*, pp 317–355.

(31) Sengupta, A., Pakrasi, H. B., and Wangikar, P. P. (2018) Recent advances in synthetic biology of cyanobacteria. *Appl. Microbiol. Biotechnol. 102*, 5457–5471.

(32) Heidorn, T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., and Lindblad, P. (2011) Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions, in *Methods in Enzymology*, pp 539–579.

(33) Wang, B., Wang, J., and Meldrum, D. R. (2012) Application of synthetic biology in cyanobacteria and algae. *Front. Microbiol. 3*, 1–15.

(34) Ramey, C. J., Barón-Sola, Á., Aucoin, H. R., and Boyle, N. R. (2015) Genome Engineering in Cyanobacteria: Where We Are and Where We Need to Go. *ACS Synth. Biol. 4*, 1186–1196.

(35) Xia, P. F., Ling, H., Foo, J. L., and Chang, M. W. (2019) Synthetic Biology Toolkits for Metabolic Engineering of Cyanobacteria. *Biotechnol. J. 14*, 1–11.

(36) Gale, G. A. R., Osorio, A. A. S., Mills, L. A., Wang, B., Lea-Smith, D. J., and McCormick, A. J. (2019) Emerging species and genome editing tools: Future prospects in cyanobacterial synthetic biology. *Microorganisms 7*, 1–36.

(37) Zavřel, T., Sinetova, M. A., Búzová, D., Literáková, P., and Červený, J. (2015) Characterization of a model cyanobacterium Synechocystis sp: PCC 6803 autotrophic growth in a flat-panel photobioreactor. *Eng. Life Sci. 15*,

122–132.

(38) Bernstein, H. C., Konopka, A., Melnicki, M. R., Hill, E. A., Kucek, L. A., Zhang, S., Shen, G., Bryant, D. A., and Beliaev, A. S. (2014) Effect of mono- And dichromatic light quality on growth rates and photosynthetic performance of Synechococcus sp. PCC 7002. *Front. Microbiol. 5*, 1–9.

(39) Bernstein, H., McClure, R., Hill, E. A., Markillie, L., Chrisler, W., Romine, M., Mcdermott, J., Posewitz, M. C., Bryant, D. A., Allan, K., Fredrickson, J., and Beliaev, A. S. (2016) Unlocking the Constraints of Cyanobacterial Productivity: Acclimations Enabling Ultrafast Growth. *MBio 7*, 1–10.

(40) Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., Koppenaal, D. W., Brand, J. J., and Pakrasi, H. B. (2015) Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. *Sci. Rep. 5*, 1–10.

(41) Loza-Correa, M., Gomez-Valero, L., and Buchrieser, C. (2010) Circadian clock proteins in prokaryotes: Hidden rhythms? *Front. Microbiol. 1*, 1–11.

(42) Camsund, D., and Lindblad, P. (2014) Engineered transcriptional systems for cyanobacterial biotechnology. *Front. Bioeng. Biotechnol. 2*, 1–9.

(43) Krishnakumar, S., Gaudana, S. B., Digmurti, M. G., Viswanathan, G. A., Chetty, M., and Wangikar, P. P. (2015) Influence of mixotrophic growth on rhythmic oscillations in expression of metabolic pathways in diazotrophic cyanobacterium Cyanothece sp. ATCC 51142. *Bioresour. Technol. 188*, 145–152.

(44) Griese, M., Lange, C., and Soppa, J. (2011) Ploidy in cyanobacteria. *FEMS Microbiol. Lett. 323*, 124–131.

(45) Zerulla, K., Ludt, K., and Soppa, J. (2016) The ploidy level of synechocystis sp. PCC 6803 is highly variable and is influenced by growth phase and by chemical and physical external parameters. *Microbiol. (United Kingdom) 162*, 730–739.

(46) Saecker, R. M., Record, M. T., and Dehaseth, P. L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - Promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol. 412*, 754–771.

(47) iGEM. (2020) Registry of Standard Biological Parts. *http://parts.igem.org/Catalog*.

(48) Camsund, D., Heidorn, T., and Lindblad, P. (2014) Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *J. Biol. Eng. 8*, 1–23.

(49) Muramatsu, M., and Hihara, Y. (2006) Characterization of high-light-responsive promoters of the psaAB genes in Synechocystis sp. PCC 6803. *Plant Cell Physiol. 47*, 878–890.

(50) Muramatsu, M., and Hihara, Y. (2007) Coordinated high-light response of genes encoding subunits of photosystem I is achieved by AT-rich upstream sequences in the cyanobacterium Synechocystis sp. strain PCC 6803. *J. Bacteriol. 189*, 2750–2758.

(51) Sakai, M., Ogawa, T., Matsuoka, M., and Fukuda, H. (1997) Photosynthetic conversion of carbon dioxide to ethylene by the recombinant cyanobacterium, Synechococcus sp. PCC 7942, which harbors a gene for the ethylene-forming enzyme of Pseudomonas syringae. *J. Ferment. Bioeng. 84*, 434–443.

(52) Iwaki, T., Haranoh, K., Inoue, N., Kojima, K., Satoh, R., Nishino, T., Wada, S., Ihara, H., Tsuyama, S., Kobayashi, H., and Wadano, A. (2006) Expression of foreign type I ribulose-1,5-bisphosphate carboxylase/ oxygenase (EC 4.1.1.39) stimulates photosynthesis in cyanobacterium Synechococcus PCC7942 cells. *Photosynth. Res. 88*, 287–297.

(53) Imashimizu, M., Fujiwara, S., Tanigawa, R., Tanaka, K., Hirokawa, T., Nakajima, Y., and Higo, J. (2003) Thymine at -5 Is Crucial for cpc Promoter Activity of Synechocystis sp. Strain PCC 6714. *J. Bacteriol. 185*, 6477–6480.

(54) Deng, M. De, and Coleman, J. R. (1999) Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ.*

Microbiol. 65, 523–528.

(55) Liu, X., Fallon, S., Sheng, J., and Curtiss, R. (2011) CO2-limitation-inducible Green Recovery of fatty acids from cyanobacterial biomass. *Proc. Natl. Acad. Sci. U. S. A. 108*, 6905–6908.

(56) Huang, H. H., Camsund, D., Lindblad, P., and Heidorn, T. (2010) Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res. 38*, 2577–2593.

(57) Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., and Lu, X. (2011) Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab. Eng. 13*, 169–176.

(58) Liu, X., and Curtiss, R. (2009) Nickel-inducible lysis system in Synechocystis sp. PCC 6803. *Proc. Natl. Acad. Sci. U. S. A. 106*, 21550–21554.

(59) Immethun, C. M., DeLorenzo, D. M., Focht, C. M., Gupta, D., Johnson, C. B., and Moon, T. S. (2017) Physical, chemical, and metabolic state sensors expand the synthetic biology toolbox for Synechocystis sp. PCC 6803. *Biotechnol. Bioeng. 114*, 1561–1569.

(60) Kelly, C. L., Taylor, G. M., Hitchcock, A., Torres-Méndez, A., and Heap, J. T. (2018) A Rhamnose-Inducible System for Precise and Temporal Control of Gene Expression in Cyanobacteria. *ACS Synth. Biol. 7*, 1056–1066.

(61) Liu, D., and Pakrasi, H. B. (2018) Exploring native genetic elements as plug-in tools for synthetic biology in the cyanobacterium Synechocystis sp. PCC 6803. *Microb. Cell Fact. 17*, 1–8.

(62) Ferreira, E. A., Pacheco, C. C., Pinto, F., Pereira, J., Lamosa, P., Oliveira, P., Kirov, B., Jaramillo, A., and Tamagnini, P. (2018) Expanding the toolbox for Synechocystis sp. PCC 6803: validation of replicative vectors and characterization of a novel set of promoters. *Synth. Biol. 3*, 1–15.

(63) Zess, E. K., Begemann, M. B., and Pfleger, B. F. (2016) Construction of new synthetic biology tools for the control of gene expression in the cyanobacterium Synechococcus sp. strain PCC 7002. *Biotechnol. Bioeng. 113*, 424–432.

(64) Wu, J., Zhao, F., Wang, S., Deng, G., Wang, J., Bai, J., Lu, J., Qu, J., and Bao, Q. (2007) cTFbase: A database for comparative genomics of transcription factors in cyanobacteria. *BMC Genomics 8*, 1–7.

(65) Thiriet-Rupert, S., Carrier, G., Chénais, B., Trottier, C., Bougaran, G., Cadoret, J. P., Schoefs, B., and Saint-Jean, B. (2016) Transcription factors in microalgae: Genome-wide prediction and comparative analysis. *BMC Genomics 17*, 1–16.

(66) Kelly, C. L., Taylor, G. M., Šatkutė, A., Dekker, L., and Heap, J. T. (2019) Transcriptional terminators allow leak-free chromosomal integration of genetic constructs in cyanobacteria. *Microorganisms 7*, 1–9.

(67) Chen, Y. J., Liu, P., Nielsen, A. A. K., Brophy, J. A. N., Clancy, K., Peterson, T., and Voigt, C. A. (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods 10*, 659–664.

(68) Englund, E., Liang, F., and Lindberg, P. (2016) Evaluation of promoters and ribosome binding sites for biotechnological applications in the unicellular cyanobacterium Synechocystis sp. PCC 6803. *Sci. Rep. 6*, 1–12.

(69) Thiel, K., Mulaku, E., Dandapani, H., Nagy, C., Aro, E. M., and Kallio, P. (2018) Translation efficiency of heterologous proteins is significantly affected by the genetic context of RBS sequences in engineered cyanobacterium Synechocystis sp. PCC 6803. *Microb. Cell Fact.*

(70) Storz, G., Vogel, J., and Wassarman, K. M. (2011) Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol. Cell. 43*, 880–891.

(71) Kopf, M., and Hess, W. R. (2015) Regulatory RNAs in photosynthetic cyanobacteria. *FEMS Microbiol. Rev. 39*, 301–315.

132–143.

(90) Yao, L., Cengic, I., Anfelt, J., and Hudson, E. P. (2016) Multiple Gene Repression in Cyanobacteria Using CRISPRi. *ACS Synth. Biol. 5*, 207–212.

(91) Yao, L., Shabestary, K., Björk, S. M., Asplund-Samuelsson, J., Joensson, H. N., Jahn, M., and Hudson, E. P. (2020) Pooled CRISPRi screening of the cyanobacterium Synechocystis sp PCC 6803 for enhanced industrial phenotypes. *Nat. Commun. 11*, 1–13.

(92) Higo, A., Isu, A., Fukaya, Y., Ehira, S., and Hisabori, T. (2018) Application of CRISPR Interference for Metabolic Engineering of the Heterocyst-Forming Multicellular Cyanobacterium Anabaena sp. PCC 7120. *Plant Cell Physiol. 59*, 119–127.

(93) Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res. 41*, 7429–7437.

(94) Huang, C. H., Shen, C. R., Li, H., Sung, L. Y., Wu, M. Y., and Hu, Y. C. (2016) CRISPR interference (CRISPRi) for gene regulation and succinate production in cyanobacterium S. elongatus PCC 7942. *Microb. Cell Fact. 15*, 1–11.

(95) Kimura, A., Hamada, T., Morita, E. H., and Hayashi, H. (2002) A high temperature-sensitive mutant of Synechococcus sp. PCC 7002 with modifications in the endogenous plasmid, pAQ1. *Plant Cell Physiol. 43*, 217–223.

(96) Voshol, G. P., Meyer, V., and van den Hondel, C. A. M. J. J. (2015) GTP-binding protein Era: A novel gene target for biofuel production. *BMC Biotechnol. 15*, 1–13.

(97) Suzuki, J. Y., and Bauer, C. E. (1995) A prokaryotic origin for light-dependent chlorophyll biosynthesis of plants. *Proc. Natl. Acad. Sci. U. S. A. 92*, 3749–3753.

(98) Zhu, H., Ren, X., Wang, J., Song, Z., Shi, M., Qiao, J., Tian, X., Liu, J., Chen, L., and Zhang, W. (2013) Integrated OMICS guided engineering of biofuel butanol-tolerance in photosynthetic Synechocystis sp. PCC 6803. *Biotechnol. Biofuels 6*, 1–14.

(99) Eaton-Rye, J. J. (2004) The Construction of Gene Knockouts in the Cyanobacterium Synechocystis sp. PCC 6803., in *Photosynthesis Research Protocols. Methods In Molecular Biology, vol 274.* (Carpentier, R., Ed.), pp 309–324. Humana Press.

(100) Viola, S., Rühle, T., and Leister, D. (2014) A single vector-based strategy for marker-less gene replacement in Synechocystis sp. PCC 6803. *Microb. Cell Fact. 13*, 1–12.

(101) Watabe, K., Mimuro, M., and Tsuchiya, T. (2014) Development of a high-frequency in vivo transposon mutagenesis system for synechocystis sp. PCC 6803 and synechococcus elongatus PCC 7942. *Plant Cell Physiol. 55*, 2017–2026.

(102) Nozzi, N. E., and Atsumi, S. (2015) Genome Engineering of the 2,3-Butanediol Biosynthetic Pathway for Tight Regulation in Cyanobacteria. *ACS Synth. Biol. 4*, 1197–1204.

(103) Xiong, W., Cano, M., Wang, B., Douchi, D., and Yu, J. (2017) The plasticity of cyanobacterial carbon metabolism. *Curr. Opin. Chem. Biol. 41*, 12–19.

(104) Zhang, F., and Keasling, J. (2011) Biosensors and their applications in microbial metabolic engineering. *Trends Microbiol. 19*, 323–329.

(105) Mehrotra, P. (2016) Biosensors and their applications - A review. *J. Oral Biol. Craniofacial Res. 6*, 153–159.

(106) Nierman, W. C., and Feldblyum, T. V. (2001) Genomic Library, in *Encyclopedia of Genetics* (Brenner, S., and Miller, J. H. B. T.-E. of G., Eds.), pp 865–872. Academic Press, New York.

(107) Nicolaou, S. A., Gaida, S. M., and Papoutsakis, E. T. (2012) Exploring the combinatorial genomic space in

Escherichia coli for ethanol tolerance. *Biotechnol. J. 7*, 1337–1345.

(108) Reyes, L. H., Almario, M. P., and Kao, K. C. (2011) Genomic library screens for genes involved in n-butanol tolerance in Escherichia coli. *PLoS One 6*, 1–9.

(109) Aikens, J., and Turner, R. J. (2009) Transgenic photosynthetic microorganisms and photobioreactor.

(110) Peabody, G. L., Winkler, J., and Kao, K. C. (2014) Tools for developing tolerance to toxic chemicals in microbial systems and perspectives on moving the field forward and into the industrial setting. *Curr. Opin. Chem. Eng. 6*, 9–17.

(111) Jahn, M., Vorpahl, C., Hübschmann, T., Harms, H., and Müller, S. (2016) Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microb. Cell Fact. 15*, 1–12.

(112) Sengupta, A., Madhu, S., and Wangikar, P. P. (2020) A Library of Tunable, Portable, and Inducer-Free Promoters Derived from Cyanobacteria. *ACS Synth. Biol. 9*, 1790–1801.

(113) Pope, M. A., Hodge, J. A., and Nixon, P. J. (2020) An Improved Natural Transformation Protocol for the Cyanobacterium Synechocystis sp. PCC 6803. *Front. Plant Sci. 11*, 1–9.

(114) Ng, W. O., Zentella, R., Wang, Y., Taylor, J. S. A., and Pakrasi, H. B. (2000) phrA, the major photoreactivating factor in the cyanobacterium Synechocystis sp. strain PCC 6803 codes for a cyclobutane-pyrimidine-dimer- specific DNA photolyase. *Arch. Microbiol. 173*, 412–417.

(115) Miao, R., Liu, X., Englund, E., Lindberg, P., and Lindblad, P. (2017) Isobutanol production in Synechocystis PCC 6803 using heterologous and endogenous alcohol dehydrogenases. *Metab. Eng. Commun. 5*, 45–53.

(116) Nozzi, N. E., Case, A. E., Carroll, A. L., and Atsumi, S. (2017) Systematic Approaches to Efficiently Produce 2,3-Butanediol in a Marine Cyanobacterium. *ACS Synth. Biol. 6*, 2136–2144.

(117) Chen, Y., Taton, A., Go, M., London, R. E., Pieper, L. M., Golden, S. S., and Golden, J. W. (2016) Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium Synechococcus elongatus PCC 7942. *Microbiol. (United Kingdom) 162*, 2029–2041.

(118) Cheah, Y. E., Albers, S. C., and Peebles, C. A. M. (2013) A novel counter-selection method for markerless genetic modification in Synechocystis sp. PCC 6803. *Biotechnol. Prog. 29*, 23–30.

(119) van der Woude, A. D., Perez Gallego, R., Vreugdenhil, A., Puthan Veetil, V., Chroumpi, T., and Hellingwerf, K. J. (2016) Genetic engineering of Synechocystis PCC6803 for the photoautotrophic production of the sweetener erythritol. *Microb. Cell Fact. 15*, 1–12.

(120) Du, W., Angermayr, S. A., Jongbloets, J. A., Molenaar, D., Bachmann, H., Hellingwerf, K. J., and Branco Dos Santos, F. (2017) Nonhierarchical Flux Regulation Exposes the Fitness Burden Associated with Lactate Production in Synechocystis sp. PCC6803. *ACS Synth. Biol. 6*, 395–401.

(121) Kopf, M., Klähn, S., Scholz, I., Matthiessen, J. K. F., Hess, W. R., and Voß, B. (2014) Comparative Analysis of the Primary Transcriptome of Synechocystis sp. PCC 6803. *DNA Res. 21*, 527–539.

(122) Zabarovsky, E. R. (2006, September) Genomic DNA Libraries, Construction and Applications. *Rev. Cell Biol. Mol. Med.*

(123) Friehs, K. (2004) Plasmid Copy Number and Plasmid Stability BT - New Trends and Developments in Biochemical Engineering, in *Advances in Biochemical Engineering* (Scheper, T., Ed.), pp 47–82. Springer Berlin Heidelberg, Berlin, Heidelberg.

(124) Paulsson, J., and Ehrenberg, M. (2001) Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys. 34*, 1–59.

(125) Du, W., Jongbloets, J. A., van Boxtel, C., Pineda Hernández, H., Lips, D., Oliver, B. G., Hellingwerf, K. J., and

Branco Dos Santos, F. (2018) Alignment of microbial fitness with engineered product formation: obligatory coupling between acetate production and photoautotrophic growth. *Biotechnol. Biofuels 11*, 1–13.

(126) Zhu, T., Xie, X., Li, Z., Tan, X., and Lu, X. (2015) Enhancing photosynthetic production of ethylene in genetically engineered Synechocystis sp. PCC 6803. *Green Chem. 17*, 421–434.

(127) Carpine, R., Du, W., Olivieri, G., Pollio, A., Hellingwerf, K. J., Marzocchella, A., and Branco dos Santos, F. (2017) Genetic engineering of Synechocystis sp. PCC6803 for poly-β-hydroxybutyrate overproduction. *Algal Res. 25*, 117–127.

(128) Jones, P. R. (2014) Genetic instability in cyanobacteria - An elephant in the room? *Front. Bioeng. Biotechnol. 2*, 1–5.

(129) Du, W., Jongbloets, J. A., Guillaume, M., Van De Putte, B., Battaglino, B., Hellingwerf, K. J., and Branco Dos Santos, F. (2019) Exploiting Day- And Night-Time Metabolism of Synechocystis sp. PCC 6803 for Fitness-Coupled Fumarate Production around the Clock. *ACS Synth. Biol. 8*, 2263–2269.

(130) Klamt, S., and Mahadevan, R. (2015) On the feasibility of growth-coupled product synthesis in microbial strains. *Metab. Eng. 30*, 166–178.

(131) Du, W., Caicedo Burbano, P., Hellingwerf, K. J., and Branco Dos Santos, F. (2018) Challenges in the Application of Synthetic Biology Toward Synthesis of Commodity Products by Cyanobacteria via "Direct Conversion," in *Synthetic Biology of Cyanobacteria. Advances in Experimental Medicine and Biology, vol 1080.* (Zhang W., S. X., Ed.), pp 3–26. Springer, Singapore.

(132) Goldberg, I., and Rokem, J. S. (2014) Fumaric Acid Biosynthesis and Accumulation, in *Bioprocessing of Renewable Resources to Commodity Bioproducts* (Bisaria, V. S., and Kondo, A., Eds.), pp 409–434.

(133) Chen, X., Dong, X., Wang, Y., Zhao, Z., and Liu, L. (2015) Mitochondrial engineering of the TCA cycle for fumarate production. *Metab. Eng. 31*, 62–73.

(134) Ganesh, I., Ravikumar, S., Lee, S. H., Park, S. J., and Hong, S. H. (2013) Engineered fumarate sensing Escherichia coli based on novel chimeric two-component system. *J. Biotechnol. 168*, 560–566.

(135) Shinde, S., Zhang, X., Singapuri, S. P., Kalra, I., Liu, X., Morgan-Kiss, R. M., and Wang, X. (2020) Glycogen metabolism supports photosynthesis start through the oxidative pentose phosphate pathway in cyanobacteria. *Plant Physiol. 182*, 507–517.

(136) Wan, N., DeLorenzo, D. M., He, L., You, L., Immethun, C. M., Wang, G., Baidoo, E. E. K., Hollinshead, W., Keasling, J. D., Moon, T. S., and Tang, Y. J. (2017) Cyanobacterial carbon metabolism: Fluxome plasticity and oxygen dependence. *Biotechnol. Bioeng. 114*, 1593–1602.

(137) Kreth, J., Lengeler, J. W., and Jahreis, K. (2013) Characterization of Pyruvate Uptake in Escherichia coli K-12. *PLoS One 8*, 1–7.

(138) Ihlenfeldt, M. J. A., and Gibson, J. (1977) Acetate uptake by the unicellular cyanobacteria Synechococcus and Aphanocapsa. *Arch. Microbiol. 113*, 231–241.

(139) Zhang, S., and Bryant, D. A. (2014) Learning new tricks from an old cycle: the TCA cycle in cyanobacteria, algae and plants. *Perspect. Phycol. 1*, 73–86.

(140) Tang, K. H., Tang, Y. J., and Blankenship, R. E. (2011) Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front. Microbiol. 2*, 1–23.

(141) Eley, J. H. (1988) Glyoxylate cycle enzyme activities in the cyanobacterim Anacystis nidulans. *ournal Phycol. 24*, 586–588.

(142) Bandyopadhyay, A., Elvitigala, T., Welsh, E., Stöckel, J., Liberton, M., Min, H., Sherman, L. A., and Pakrasi, H. B.

(2011) Novel Metabolic Attributes of the Genus Cyanothece, Comprising a Group of Unicellular Nitrogen-Fixing Cyanobacteria. *MBio 2*, 1–10.

(143) Kornberg, H. L., and Krebs, H. A. (1957) Synthesis of Cell Constituents From C2-Units By A Modified Tricarboxylic Acid Cycle. *Nature 179*, 988–991.

(144) Kornberg, H. (2000) Krebs and his trinity of cycles. *Nat. Rev. Mol. Cell Biol. 1*, 225–228.

(145) Zhang, S., and Bryant, D. A. (2015) Biochemical validation of the glyoxylate cycle in the cyanobacterium Chlorogloeopsis fritschii strain PCC 9212. *J. Biol. Chem. 290*, 14019–14030.

(146) Tu, B. P., and McKnight, S. L. (2006) Metabolic cycles as an underlying basis of biological oscillations. *Nat. Rev. Mol. Cell Biol. 7*, 696–701.

(147) Cassier-Chauvat, C., Veaudor, T., and Chauvat, F. (2016) Comparative genomics of DNA recombination and repair in cyanobacteria: Biotechnological implications. *Front. Microbiol. 7*, 1–13.

(148) Wan Jusoh, W. N. A., and Wong, L. S. (2014) Exploring the Potential of Whole Cell Biosensor : A Review in Environmental Applications. *Int. J. Chem. Environ. Biol. Sci. 2*, 52–56.

(149) Su, L., Jia, W., Hou, C., and Lei, Y. (2011) Microbial biosensors: A review. *Biosens. Bioelectron. 26*, 1788–1799.

(150) Clark, L. C., and Lyons, C. (1962) Electrode Systems for Continuous Monitoring in Cardiovascular Surgery. *Ann. N. Y. Acad. Sci. 102*, 29–45.

(151) Altintas, Z. (2018) Biosensors and Nanotechnology: Applications in Health Care Diagnostics (Altintas, Z., Ed.). Wiley.

(152) Halilovic, A., Merdan, E., Kovacevic, Z., and Pokvic, L. G. (2019) Review of Biosensors for Environmental Field Monitoring. *2019 8th Mediterr. Conf. Embed. Comput. MECO 2019 - Proc.* 1–8.

(153) Hazelbauer, G. L., and Lai, W. C. (2010) Bacterial chemoreceptors: providing enhanced features to two-component signaling. *Curr. Opin. Microbiol. 13*, 124–132.

(154) Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-Component Signal Transduction. *Annu. Rev. Biochem. 69*, 183–215.

(155) Davies, S. J., Golby, P., Omrani, D., Broad, S. A., Harrington, V. L., Guest, J. R., Kelly, D. J., and Andrews, S. C. (1999) Inactivation and regulation of the aerobic C4-dicarboxylate transport (dctA) gene of Escherichia coli. *J. Bacteriol. 181*, 5624–5635.

(156) Forst, S. A., and Roberts, D. L. (1994) Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *Res. Microbiol. 145*, 363–373.

(157) Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., and Voigt, C. A. (2012) Genetic programs constructed from layered logic gates in single cells. *Nature 491*, 249–253.

(158) Bi, S., Pollard, A. M., Yang, Y., Jin, F., and Sourjik, V. (2016) Engineering Hybrid Chemotaxis Receptors in Bacteria. *ACS Synth. Biol. 5*, 989–1001.

(159) Bachmann, H., Fischlechner, M., Rabbers, I., Barfa, N., Dos Santos, F. B., Molenaar, D., and Teusink, B. (2013) Availability of public goods shapes the evolution of competing metabolic strategies. *Proc. Natl. Acad. Sci. U. S. A. 110*, 14302–14307.

(160) Görke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nat. Rev. Microbiol. 6*, 613–624.

(161) Savakis, P. E., Angermayr, S. A., and Hellingwerf, K. J. (2013) Synthesis of 2,3-butanediol by Synechocystis sp. PCC6803 via heterologous expression of a catabolic pathway from lactic acid- and enterobacteria. *Metab. Eng. 20*,

121–130.

(162) Batchelor, E., Silhavy, T. J., and Goulian, M. (2004) Continuous control in bacterial regulatory circuits. *J. Bacteriol. 186*, 7618–7625.

(163) Lau, N. S., Matsui, M., and Abdullah, A. A. A. (2015) Cyanobacteria: Photoautotrophic Microbial Factories for the Sustainable Synthesis of Industrial Products. *Biomed Res. Int. 2015*, 1–9.

(164) Berthoumieux, S., De Jong, H., Baptist, G., Pinel, C., Ranquet, C., Ropers, D., and Geiselmann, J. (2013) Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. *Mol. Syst. Biol. 9*, 1–11.

(165) Cardinale, S., and Arkin, A. P. (2012) Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol. J. 7*, 856–866.

(166) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods 10*, 354–360.

(167) Ren, G. X., Guo, X. P., and Sun, Y. C. (2017) Regulatory 3' untranslated regions of bacterial mRNAs. *Front. Microbiol. 8*, 1–6.

(168) Bryant, J. A., Sellars, L. E., Busby, S. J. W., and Lee, D. J. (2014) Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic Acids Res. 42*, 11383–11392.

(169) Schmid, M. B., and Roth, J. R. (1987) Gene location affects expression level in salmonella typhimurium. *J. Bacteriol. 169*, 2872–2875.

(170) Thompson, A., and Gasson, M. J. (2001) Location Effects of a Reporter Gene on Expression Levels and on Native Protein Synthesis in Lactococcus lactis and Saccharomyces cerevisiae. *Appl. Environ. Microbiol. 67*, 3434–3439.

(171) Sauer, C., Syvertsson, S., Bohorquez, L. C., Cruz, R., Harwood, C. R., Van Rij, T., and Hamoen, L. W. (2016) Effect of Genome Position on Heterologous Gene Expression in Bacillus subtilis: An Unbiased Analysis. *ACS Synth. Biol. 5*, 942–947.

(172) Peter, B. J., Arsuaga, J., Breier, A. M., Khodursky, A. B., Brown, P. O., and Cozzarelli, N. R. (2004) Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli. *Genome Biol. 5*, 1–16.

(173) Block, D. H. S., Hussein, R., Liang, L. W., and Lim, H. N. (2012) Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Res. 40*, 8979–8992.

(174) Ng, A. H., Berla, B. M., and Pakrasi, H. B. (2015) Fine-tuning of photoautotrophic protein production by combining promoters and neutral sites in the cyanobacterium Synechocystis sp. strain PCC 6803. *Appl. Environ. Microbiol. 81*, 6857–6863.

(175) Pinto, F., Pacheco, C. C., Oliveira, P., Montagud, A., Landels, A., Couto, N., Wright, P. C., Urchueguía, J. F., and Tamagnini, P. (2015) Improving a Synechocystis-based photoautotrophic chassis through systematic genome mapping and validation of neutral sites. *DNA Res. 22*, 425–437.

(176) Berla, B. M., and Pakrasi, H. B. (2012) Upregulation of Plasmid Genes during Stationary Phase in Synechocystis sp. Strain PCC 6803, a Cyanobacterium. *Appl. Environ. Microbiol. 78*, 5448–5451.

(177) Kunert, A., Hagemann, M., and Erdmann, N. (2000) Construction of promoter probe vectors for Synechocystis sp. PCC 6803 using the light-emitting reporter systems Gfp and LuxAB. *J. Microbiol. Methods 41*, 185–194.

(178) Williams, J. G. K. (1988) Construction of Specific Mutations in Photosystem II Photosynthetic Reaction Center by Genetic Engineering Methods in Synechocystis 6803. *Methods Enzymol. 167*, 766–778.

(179) Murray, H., and Koh, A. (2014) Multiple Regulatory Systems Coordinate DNA Replication with Cell Growth in

Bacillus subtilis. *PLoS Genet. 10*, 1–15.

(180) Gibson, B., Wilson, D. J., Feil, E., and Eyre-Walker, A. (2018) The distribution of bacterial doubling times in the wild. *Proc. R. Soc. B Biol. Sci. 285*, 1–9.

(181) Vermass, W. F. J., Rutherford, A. W., and Hansson, O. (1988) Site-directed mutagenesis in photosystem II of the cyanobacterium Synechocystis sp. PCC 6803: Donor D is a tyrosine residue in the D2 protein. *Proc. Natl. Acad. Sci. 85*, 8477–8481.

(182) van Alphen, P., Abedini Najafabadi, H., Branco dos Santos, F., and Hellingwerf, K. J. (2018) Increasing the Photoautotrophic Growth Rate of Synechocystis sp. PCC 6803 by Identifying the Limitations of Its Cultivation. *Biotechnol. J. 13*, 1–8.

(183) Couradeau, E., Benzerara, K., Gérard, E., Moreira, D., Bernard, S., Brown, G. E., and López-García, P. (2012) An early-branching microbialite cyanobacterium forms intracellular carbonates. *Science (80-.). 336*, 459–462.

(184) Carroll, A. L., Case, A. E., Zhang, A., and Atsumi, S. (2018) Metabolic engineering tools in model cyanobacteria. *Metab. Eng. 50*, 47–56.

(185) Wijffels, R. H., Kruse, O., and Hellingwerf, K. J. (2013) Potential of industrial biotechnology with cyanobacteria and eukaryotic microalgae. *Curr. Opin. Biotechnol. 24*, 405–413.

(186) Nielsen, J., and Keasling, J. D. (2016) Engineering Cellular Metabolism. *Cell 164*, 1185–1197.

(187) Zhou, J., Zhang, H., Zhang, Y., Li, Y., and Ma, Y. (2012) Designing and creating a modularized synthetic pathway in cyanobacterium Synechocystis enables production of acetone from carbon dioxide. *Metab. Eng. 14*, 394–400.

(188) Chwa, J. W., Kim, W. J., Sim, S. J., Um, Y., and Woo, H. M. (2016) Engineering of a modular and synthetic phosphoketolase pathway for photosynthetic production of acetone from CO2 in Synechococcus elongatus PCC 7942 under light and aerobic condition. *Plant Biotechnol. J. 14*, 1768–1776.

(189) Wang, W., Liu, X., and Lu, X. (2013) Engineering cyanobacteria to improve photosynthetic production of alka(e)nes. *Biotechnol. Biofuels 6*, 1–9.

(190) Kageyama, H., Waditee-Sirisattha, R., Sirisattha, S., Tanaka, Y., Mahakhant, A., and Takabe, T. (2015) Improved Alkane Production in Nitrogen-Fixing and Halotolerant Cyanobacteria via Abiotic Stresses and Genetic Manipulation of Alkane Synthetic Genes. *Curr. Microbiol. 71*, 115–120.

(191) Choi, S. Y., Lee, H. J., Choi, J., Kim, J., Sim, S. J., Um, Y., Kim, Y., Lee, T. S., Keasling, J. D., and Woo, H. M. (2016) Photosynthetic conversion of CO2 to farnesyl diphosphate-derived phytochemicals (amorpha-4,11-diene and squalene) by engineered cyanobacteria. *Biotechnol. Biofuels 9*, 1–12.

(192) Diao, J., Song, X., Zhang, L., Cui, J., Chen, L., and Zhang, W. (2020) Tailoring cyanobacteria as a new platform for highly efficient synthesis of astaxanthin. *Metab. Eng. 61*, 275–287.

(193) Sebesta, J., and Peebles, C. A. (2020) Improving heterologous protein expression in Synechocystis sp. PCC 6803 for alpha-bisabolene production. *Metab. Eng. Commun. 10*, 1–9.

(194) Davies, F. K., Work, V. H., Beliaev, A. S., and Posewitz, M. C. (2014) Engineering limonene and bisabolene production in wild type and a glycogen-deficient mutant of Synechococcus sp. PCC 7002. *Front. Bioeng. Biotechnol. 2*, 1–11.

(195) Kanno, M., Carroll, A. L., and Atsumi, S. (2017) Global metabolic rewiring for improved CO2 fixation and chemical production in cyanobacteria. *Nat. Commun. 8*, 1–11.

(196) Oliver, J. W. K., Machado, I. M. P., Yoneda, H., and Atsumi, S. (2013) Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. U. S. A. 110*, 1249–1254.

(197) Oliver, J. W. K., Machado, I. M. P., Yoneda, H., and Atsumi, S. (2014) Combinatorial optimization of

cyanobacterial 2,3-butanediol production. *Metab. Eng. 22*, 76–82.

(198) Lan, E. I., Ro, S. Y., and Liao, J. C. (2013) Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy Environ. Sci. 6*, 2672–2681.

(199) Lan, E. I., and Liao, J. C. (2012) ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A. 109*, 6018–6023.

(200) Lan, E. I., and Liao, J. C. (2011) Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab. Eng. 13*, 353–363.

(201) Xue, Y., Zhang, Y., Grace, S., and He, Q. (2014) Functional expression of an Arabidopsis p450 enzyme, p-coumarate-3-hydroxylase, in the cyanobacterium Synechocystis PCC 6803 for the biosynthesis of caffeic acid. *J. Appl. Phycol. 26*, 219–226.

(202) Lagarde, D., Beuf, L., and Vermaas, W. (2000) Increased Production of Zeaxanthin and Other Pigments by Application of Genetic Engineering Techniques to Synechocystis sp. Strain PCC 6803. *Appl. Environ. Microbiol. 66*, 64–72.

(203) Reinsvold, R. E., Jinkerson, R. E., Radakovits, R., Posewitz, M. C., and Basu, C. (2011) The production of the sesquiterpene β-caryophyllene in a transgenic strain of the cyanobacterium Synechocystis. *J. Plant Physiol. 168*, 848–852.

(204) Xue, Y., Zhang, Y., Cheng, D., Daddy, S., and He, Q. (2014) Genetically engineering Synechocystis sp. Pasteur Culture Collection 6803 for the sustainable production of the plant secondary metabolite p-coumaric acid. *Proc. Natl. Acad. Sci. U. S. A. 111*, 9449–9454.

(205) Wlodarczyk, A., Gnanasekaran, T., Nielsen, A. Z., Zulu, N. N., Mellor, S. B., Luckner, M., Thøfner, J. F. B., Olsen, C. E., Mottawie, M. S., Burow, M., Pribil, M., Feussner, I., Møller, B. L., and Jensen, P. E. (2016) Metabolic engineering of light-driven cytochrome P450 dependent pathways into Synechocystis sp. PCC 6803. *Metab. Eng. 33*, 1–11.

(206) Vavitsas, K., Rue, E. Ø., Stefánsdóttir, L. K., Gnanasekaran, T., Blennow, A., Crocoll, C., Gudmundsson, S., and Jensen, P. E. (2017) Responses of Synechocystis sp. PCC 6803 to heterologous biosynthetic pathways. *Microb. Cell Fact. 16*, 1–11.

(207) Wang, Y., Tao, F., Ni, J., Li, C., and Xu, P. (2015) Production of C3 platform chemicals from CO2 by genetically engineered cyanobacteria. *Green Chem. 17*, 3100–3110.

(208) Velmurugan, R., and Incharoensakdi, A. (2020) Heterologous Expression of Ethanol Synthesis Pathway in Glycogen Deficient Synechococcus elongatus PCC 7942 Resulted in Enhanced Production of Ethanol and Exopolysaccharides. *Front. Plant Sci. 11*, 1–12.

(209) Luan, G., Qi, Y., Wang, M., Li, Z., Duan, Y., Tan, X., and Lu, X. (2015) Combinatory strategy for characterizing and understanding the ethanol synthesis pathway in cyanobacteria cell factories. *Biotechnol. Biofuels 8*, 1–12.

(210) Wang, M., Luan, G., and Lu, X. (2020) Engineering ethanol production in a marine cyanobacterium Synechococcus sp. PCC7002 through simultaneously removing glycogen synthesis genes and introducing ethanolgenic cassettes. *J. Biotechnol. 317*, 1–4.

(211) Dexter, J., and Fu, P. (2009) Metabolic engineering of cyanobacteria for ethanol production. *Energy Environ. Sci. 2*, 857–864.

(212) Lee, T. C., Xiong, W., Paddock, T., Carrieri, D., Chang, I. F., Chiu, H. F., Ungerer, J., Hank Juo, S. H., Maness, P. C., and Yu, J. (2015) Engineered xylose utilization enhances bio-products productivity in the cyanobacterium Synechocystis sp. PCC 6803. *Metab. Eng. 30*, 179–189.

(213) Ungerer, J., Tao, L., Davis, M., Ghirardi, M., Maness, P. C., and Yu, J. (2012) Sustained photosynthetic conversion of CO2 to ethylene in recombinant cyanobacterium Synechocystis 6803. *Energy Environ. Sci. 5*, 8998–9006.

(214) Pattharaprachayakul, N., Lee, H. J., Incharoensakdi, A., and Woo, H. M. (2019) Evolutionary Engineering of Cyanobacteria to Enhance the Production of α-Farnesene from CO2. *J. Agric. Food Chem. 67*, 13658–13664.

(215) Lee, H. J., Lee, J., Lee, S. M., Um, Y., Kim, Y., Sim, S. J., Choi, J. Il, and Woo, H. M. (2017) Direct Conversion of CO2 to α-Farnesene Using Metabolically Engineered Synechococcus elongatus PCC 7942. *J. Agric. Food Chem. 65*, 10424–10428.

(216) Halfmann, C., Gu, L., Gibbons, W., and Zhou, R. (2014) Genetically engineering cyanobacteria to convert CO2, water, and light into the long-chain hydrocarbon farnesene. *Appl. Microbiol. Biotechnol. 98*, 9869–9877.

(217) Liu, X., Sheng, J., and Curtiss, R. (2011) Fatty acid production in genetically modified cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A. 108*, 6899–6904.

(218) Ruffing, A. M. (2014) Improved free fatty acid production in cyanobacteria with Synechococcus sp. PCC 7002 as host. *Front. Bioeng. Biotechnol. 2*, 1–10.

(219) Ruffing, A. M. (2013) Borrowing genes from Chlamydomonas reinhardtii for free fatty acid production in engineered cyanobacteria. *J. Appl. Phycol. 25*, 1495–1507.

(220) Ruffing, A. M., and Jones, H. D. T. (2012) Physiological Effects of Free Fatty Acid Production in Genetically Engineered Synechococcus elongatus PCC 7942. *Biotechnol. Bioeng. 109*, 2190–2199.

(221) Kaczmarzyk, D., Cengic, I., Yao, L., and Hudson, E. P. (2018) Diversion of the long-chain acyl-ACP pool in Synechocystis to fatty alcohols through CRISPRi repression of the essential phosphate acyltransferase PlsX. *Metab. Eng. 45*, 59–66.

(222) Yao, L., Qi, F., Tan, X., and Lu, X. (2014) Improved production of fatty alcohols in cyanobacteria by metabolic engineering. *Biotechnol. Biofuels 7*, 1–9.

(223) Formighieri, C., and Melis, A. (2017) Heterologous synthesis of geranyllinalool, a diterpenol plant product, in the cyanobacterium Synechocystis. *Appl. Microbiol. Biotechnol. 101*, 2791–2800.

(224) Niederholtmeyer, H., Wolfstädter, B. T., Savage, D. F., Silver, P. A., and Way, J. C. (2010) Engineering cyanobacteria to synthesize and export hydrophilic products. *Appl. Environ. Microbiol. 76*, 3462–3466.

(225) Tan, X., Du, W., and Lu, X. (2015) Photosynthetic and extracellular production of glucosylglycerol by genetically engineered and gel-encapsulated cyanobacteria. *Appl. Microbiol. Biotechnol. 99*, 2147–2154.

(226) Matsudaira, A., Hoshino, Y., Uesaka, K., Takatani, N., Omata, T., and Usuda, Y. (2020) Production of glutamate and stereospecific flavors, (S)-linalool and (+)-valencene, by Synechocystis sp. PCC6803. *J. Biosci. Bioeng. 130*, 464–470.

(227) Aikawa, S., Nishida, A., Ho, S. H., Chang, J. S., Hasunuma, T., and Kondo, A. (2014) Glycogen production for biofuels by the euryhaline cyanobacteria Synechococcus sp. strain PCC 7002 from an oceanic environment. *Biotechnol. Biofuels 7*, 1–8.

(228) Savakis, P., Tan, X., Du, W., Branco Dos Santos, F., Lu, X., and Hellingwerf, K. J. (2015) Photosynthetic production of glycerol by a recombinant cyanobacterium. *J. Biotechnol. 195*, 46–51.

(229) Hirokawa, Y., Maki, Y., Tatsuke, T., and Hanai, T. (2016) Cyanobacterial production of 1,3-propanediol directly from carbon dioxide using a synthetic metabolic pathway. *Metab. Eng. 34*, 97–103.

(230) Wang, B., Dong, T., Myrlie, A., Gu, L., Zhu, H., Xiong, W., Maness, P., Zhou, R., and Yu, J. (2019) Photosynthetic production of the nitrogen-rich compound guanidine. *Green Chem. 21*, 2928–2937.

(231) Sarnaik, A., Abernathy, M. H., Han, X., Ouyang, Y., Xia, K., Chen, Y., Cress, B., Zhang, F., Lali, A., Pandit, R., Linhardt, R. J., Tang, Y. J., and Koffas, M. A. G. (2019) Metabolic engineering of cyanobacteria for photoautotrophic production of heparosan, a pharmaceutical precursor of heparin. *Algal Res. 37*, 57–63.

(232) Yoshino, T., Liang, Y., Arai, D., Maeda, Y., Honda, T., Muto, M., Kakunaka, N., and Tanaka, T. (2015) Alkane production by the marine cyanobacterium Synechococcus sp. NKBG15041c possessing the α-olefin biosynthesis pathway. *Appl. Microbiol. Biotechnol. 99*, 1521–1529.

(233) Wang, B., Xiong, W., Yu, J., Maness, P. C., and Meldrum, D. R. (2018) Unlocking the photobiological conversion of CO2 to (R)-3-hydroxybutyrate in cyanobacteria. *Green Chem. 20*, 3772–3782.

(234) Ku, J. T., and Lan, E. I. (2018) A balanced ATP driving force module for enhancing photosynthetic biosynthesis of 3-hydroxybutyrate from CO2. *Metab. Eng. 46*, 35–42.

(235) Wang, B., Pugh, S., Nielsen, D. R., Zhang, W., and Meldrum, D. R. (2013) Engineering cyanobacteria for photosynthetic production of 3-hydroxybutyrate directly from CO2. *Metab. Eng. 16*, 68–77.

(236) Wang, Y., Sun, T., Gao, X., Shi, M., Wu, L., Chen, L., and Zhang, W. (2016) Biosynthesis of platform chemical 3-hydroxypropionic acid (3-HP) directly from CO2 in cyanobacterium Synechocystis sp. PCC 6803. *Metab. Eng. 34*, 60–70.

(237) Lan, E. I., Chuang, D. S., Shen, C. R., Lee, A. M., Ro, S. Y., and Liao, J. C. (2015) Metabolic engineering of cyanobacteria for photosynthetic 3-hydroxypropionic acid production from CO2 using Synechococcus elongatus PCC 7942. *Metab. Eng. 31*, 163–170.

(238) Li, X., Shen, C. R., and Liao, J. C. (2014) Isobutanol production as an alternative metabolic sink to rescue the growth deficiency of the glycogen mutant of Synechococcus elongatus PCC 7942. *Photosynth. Res. 120*, 301–310.

(239) Varman, A. M., Xiao, Y., Pakrasi, H. B., and Tang, Y. J. (2013) Metabolic engineering of Synechocystis sp. Strain PCC 6803 for isobutanol production. *Appl. Environ. Microbiol. 79*, 908–914.

(240) Gao, X., Gao, F., Liu, D., Zhang, H., Nie, X., and Yang, C. (2016) Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO2. *Energy Environ. Sci. 9*, 1400–1411.

(241) Bentley, F. K., Zurbriggen, A., and Melis, A. (2014) Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Mol. Plant 7*, 71–86.

(242) Chaves, J. E., Rueda-Romero, P., Kirst, H., and Melis, A. (2017) Engineering Isoprene Synthase Expression and Activity in Cyanobacteria. *ACS Synth. Biol. 6*, 2281–2292.

(243) Englund, E., Shabestary, K., Hudson, E. P., and Lindberg, P. (2018) Systematic overexpression study to find target enzymes enhancing production of terpenes in Synechocystis PCC 6803, using isoprene as a model compound. *Metab. Eng. 49*, 164–177.

(244) Chaves, J. E., Romero, P. R., Kirst, H., and Melis, A. (2016) Role of isopentenyl-diphosphate isomerase in heterologous cyanobacterial (Synechocystis) isoprene production. *Photosynth. Res. 130*, 517–527.

(245) Hirokawa, Y., Suzuki, I., and Hanai, T. (2015) Optimization of isopropanol production by engineered cyanobacteria with a synthetic metabolic pathway. *J. Biosci. Bioeng. 119*, 585–590.

(246) Chin, T., Sano, M., Takahashi, T., Ohara, H., and Aso, Y. (2015) Photosynthetic production of itaconic acid in Synechocystis sp. PCC6803. *J. Biotechnol. 195*, 43–45.

(247) Varman, A. M., Yu, Y., You, L., and Tang, Y. J. (2013) Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microb. Cell Fact. 12*, 1–8.

(248) Zhou, J., Zhang, H., Meng, H., Zhang, Y., and Li, Y. (2014) Production of optically pure D-lactate from CO2 by blocking the PHB and acetate pathways and expressing d-lactate dehydrogenase in cyanobacterium Synechocystis sp. PCC 6803. *Process Biochem. 49*, 2071–2077.

(249) Li, C., Tao, F., Ni, J., Wang, Y., Yao, F., and Xu, P. (2015) Enhancing the light-driven production of D-lactate by engineering cyanobacterium using a combinational strategy. *Sci. Rep. 5*, 1–11.

(250) Angermayr, S. A., and Hellingwerf, K. J. (2013) On the use of metabolic control analysis in the optimization of cyanobacterial biosolar cell factories. *J. Phys. Chem. B 117*, 11169–11175.

(251) Angermayr, S. A., Van Der Woude, A. D., Correddu, D., Vreugdenhil, A., Verrone, V., and Hellingwerf, K. J. (2014) Exploring metabolic engineering design principles for the photosynthetic production of lactic acid by Synechocystis sp. PCC6803. *Biotechnol. Biofuels 7*, 1–15.

(252) Joseph, A., Aikawa, H., Sasaki, K., Tsuge, Y., Matsuda, F., Tanaka, T., and Kondo, A. (2013) Utilization of lactic acid bacterial genes in synechocystis sp. PCC 6803 in the production of lactic acid. *Biosci. Biotechnol. Biochem. 77*, 966–970.

(253) Lin, P. C., Saha, R., Zhang, F., and Pakrasi, H. B. (2017) Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium Synechocysti s sp. PCC. *Sci. Rep. 7*, 1–10.

(254) Wang, X., Liu, W., Xin, C., Zheng, Y., Cheng, Y., Sun, S., Li, R., Zhu, X. G., Dai, S. Y., Rentzepis, P. M., and Yuan, J. S. (2016) Enhanced limonene production in cyanobacteria reveals photosynthesis limitations. *Proc. Natl. Acad. Sci. U. S. A. 113*, 14225–14230.

(255) Kiyota, H., Okuda, Y., Ito, M., Hirai, M. Y., and Ikeuchi, M. (2014) Engineering of cyanobacteria for the photosynthetic production of limonene from CO2. *J. Biotechnol. 185*, 1–7.

(256) Halfmann, C., Gu, L., and Zhou, R. (2014) Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO2 and H2O. *Green Chem. 16*, 3175–3185.

(257) Korosh, T. C., Markley, A. L., Clark, R. L., McGinley, L. L., McMahon, K. D., and Pfleger, B. F. (2017) Engineering photosynthetic production of L-lysine. *Metab. Eng. 44*, 273–283.

(258) Jacobsen, J. H., and Frigaard, N. U. (2014) Engineering of photosynthetic mannitol biosynthesis from CO2 in a cyanobacterium. *Metab. Eng. 21*, 60–70.

(259) Wu, W., Du, W., Gallego, R. P., Hellingwerf, K. J., Van Der Woude, A. D., and Dos Santos, F. B. (2020) Using osmotic stress to stabilize mannitol production in Synechocystis sp. PCC6803. *Biotechnol. Biofuels 13*, 1–12.

(260) Englund, E., Andersen-Ranberg, J., Miao, R., Hamberger, B., and Lindberg, P. (2015) Metabolic Engineering of Synechocystis sp. PCC 6803 for Production of the Plant Diterpenoid Manoyl Oxide. *ACS Synth. Biol. 4*, 1270–1278.

(261) Shen, C. R., and Liao, J. C. (2012) Photosynthetic production of 2-methyl-1-butanol from CO2 in cyanobacterium Synechococcus elongatus PCC7942 and characterization of the native acetohydroxyacid synthase. *Energy Environ. Sci. 5*, 9574–9583.

(262) Formighieri, C., and Melis, A. (2014) Regulation of β-phellandrene synthase gene expression, recombinant protein accumulation, and monoterpene hydrocarbons production in Synechocystis transformants. *Planta 240*, 309–324.

(263) Bentley, F. K., García-Cerdán, J. G., Chen, H. C., and Melis, A. (2013) Paradigm of Monoterpene (β-phellandrene) Hydrocarbons Production via Photosynthesis in Cyanobacteria. *Bioenergy Res. 6*, 917–929.

(264) Formighieri, C., and Melis, A. (2016) Sustainable heterologous production of terpene hydrocarbons in cyanobacteria. *Photosynth. Res. 130*, 123–135.

(265) Betterle, N., and Melis, A. (2018) Heterologous Leader Sequences in Fusion Constructs Enhance Expression of Geranyl Diphosphate Synthase and Yield of β-Phellandrene Production in Cyanobacteria (Synechocystis). *ACS Synth. Biol. 7*, 912–921.

(266) Formighieri, C., and Melis, A. (2015) A phycocyanin·phellandrene synthase fusion enhances recombinant protein expression and β-phellandrene (monoterpene) hydrocarbons production in Synechocystis (cyanobacteria). *Metab. Eng. 32*, 116–124.

(267) Zhang, S., Liu, Y., and Bryant, D. A. (2015) Metabolic engineering of Synechococcus sp. PCC 7002 to produce poly-3-hydroxybutyrate and poly-3-hydroxybutyrate-co-4-hydroxybutyrate. *Metab. Eng. 32*, 174–183.

(268) Li, H., and Liao, J. C. (2013) Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO2 to 1,2-propanediol. *Microb. Cell Fact. 12*, 1–9.

(269) Hirokawa, Y., Maki, Y., and Hanai, T. (2017) Improvement of 1,3-propanediol production using an engineered cyanobacterium, Synechococcus elongatus by optimization of the gene expression level of a synthetic metabolic pathway and production conditions. *Metab. Eng. 39*, 192–199.

(270) Pattanaik, B., Englund, E., Nolte, N., and Lindberg, P. (2020) Introduction of a green algal squalene synthase enhances squalene accumulation in a strain of Synechocystis sp. PCC 6803. *Metab. Eng. Commun. 10*, 1–7.

(271) Choi, S. Y., Wang, J. Y., Kwak, H. S., Lee, S. M., Um, Y., Kim, Y., Sim, S. J., Choi, J. Il, and Woo, H. M. (2017) Improvement of Squalene Production from CO2 in Synechococcus elongatus PCC 7942 by Metabolic Engineering and Scalable Production in a Photobioreactor. *ACS Synth. Biol. 6*, 1289–1295.

(272) Englund, E., Pattanaik, B., Ubhayasekera, S. J. K., Stensjö, K., Bergquist, J., and Lindberg, P. (2014) Production of squalene in Synechocystis sp. PCC 6803. *PLoS One 9*, 1–8.

(273) Lan, E. I., and Wei, C. T. (2016) Metabolic engineering of cyanobacteria for the photosynthetic production of succinate. *Metab. Eng. 38*, 483–493.

(274) Song, K., Tan, X., Liang, Y., and Lu, X. (2016) The potential of Synechococcus elongatus UTEX 2973 for sugar feedstock production. *Appl. Microbiol. Biotechnol. 100*, 7865–7875.

(275) Ducat, D. C., Avelar-Rivas, J. A., Way, J. C., and Silvera, P. A. (2012) Rerouting carbon flux to enhance photosynthetic productivity. *Appl. Environ. Microbiol. 78*, 2660–2668.

(276) Du, W., Liang, F., Duan, Y., Tan, X., and Lu, X. (2013) Exploring the photosynthetic production capacity of sucrose by cyanobacteria. *Metab. Eng. 19*, 17–25.

(277) Qi, Q., Hao, M., Ng, W. O., Slater, S. C., Baszis, S. R., Weiss, J. D., and Valentin, H. E. (2005) Application of the Synechococcus nirA promoter to establish an inducible expression system for engineering the Synechocystis tocopherol pathway. *Appl. Environ. Microbiol. 71*, 5678–5684.

(278) Qiao, Y., Wang, W., and Lu, X. (2020) Engineering cyanobacteria as cell factories for direct trehalose production from CO2. *Metab. Eng. 62*, 161–171.

(279) Lindberg, P., Park, S., and Melis, A. (2010) Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. *Metab. Eng. 12*, 70–79.

(280) Ruffing, A. M., Jensen, T. J., and Strickland, L. M. (2016) Genetic tools for advancement of Synechococcus sp. PCC 7002 as a cyanobacterial chassis. *Microb. Cell Fact. 15*, 1–14.

(281) Chungjatupornchai, W., and Fa-aroonsawat, S. (2014) The rrnA promoter as a tool for the improved expression of heterologous genes in cyanobacteria. *Microbiol. Res. 169*, 361–368.

(282) Zhou, J., Zhang, H., Meng, H., Zhu, Y., Bao, G., Zhang, Y., Li, Y., and Ma, Y. (2014) Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. *Sci. Rep. 4*, 1–6.

(283) Huang, H. H., and Lindblad, P. (2013) Wide-dynamic-range promoters engineered for cyanobacteria. *J. Biol. Eng. 7*, 1–11.

(284) Cao, Y. Q., Li, Q., Xia, P. F., Wei, L. J., Guo, N., Li, J. W., and Wang, S. G. (2017) AraBAD based toolkit for gene expression and metabolic robustness improvement in synechococcus elongatus. *Sci. Rep. 7*, 1–10.

(285) Taton, A., Ma, A. T., Ota, M., Golden, S. S., and Golden, J. W. (2017) NOT Gate Genetic Circuits to Control Gene Expression in Cyanobacteria. *ACS Synth. Biol. 6*, 2175–2182.

(286) Immethun, C. M., Ng, K. M., Delorenzo, D. M., Waldron-Feinstein, B., Lee, Y. C., and Moon, T. S. (2016) Oxygen-responsive genetic circuits constructed in Synechocystis sp. PCC 6803. *Biotechnol. Bioeng. 113*, 433–442.

(287) Taton, A., Unglaub, F., Wright, N. E., Zeng, W. Y., Paz-Yepes, J., Brahamsha, B., Palenik, B., Peterson, T. C., Haerizadeh, F., Golden, S. S., and Golden, J. W. (2014) Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Res. 42*, 1–16.

(288) Popp, P. F., Dotzler, M., Radeck, J., Bartels, J., and Mascher, T. (2017) The Bacillus BioBrick Box 2.0: Expanding the genetic toolbox for the standardized work with Bacillus subtilis. *Sci. Rep. 7*, 1–13.

(289) Pasotti, L., Politi, N., Zucca, S., Cusella de Angelis, M. G., and Magni, P. (2012) Bottom-up engineering of biological systems through standard bricks: A modularity study on basic parts and devices. *PLoS One 7*, 1–10.

(290) Young, E., and Alper, H. (2010) Synthetic biology: Tools to design, build, and optimize cellular processes. *J. Biomed. Biotechnol. 2010*, 1–12.

(291) Stensjö, K., Vavitsas, K., and Tyystjärvi, T. (2018) Harnessing transcription for bioproduction in cyanobacteria. *Physiol. Plant. 162*, 148–155.

(292) Nogales, J., Gudmundsson, S., Knight, E. M., Palsson, B. O., and Thiele, I. (2012) Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proc. Natl. Acad. Sci. U. S. A. 109*, 2678–2683.

(293) Knoop, H., Gründel, M., Zilliges, Y., Lehmann, R., Hoffmann, S., Lockau, W., and Steuer, R. (2013) Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of Synechocystis sp. PCC 6803. *PLoS Comput. Biol. 9*, 1–15.

(294) Broddrick, J. T., Rubin, B. E., Welkie, D. G., Du, N., Mih, N., Diamond, S., Lee, J. J., Golden, S. S., and Palsson, B. O. (2016) Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proc. Natl. Acad. Sci. U. S. A. 113*, E8344–E8353.

(295) Abernathy, M. H., Yu, J., Ma, F., Liberton, M., Ungerer, J., Hollinshead, W. D., Gopalakrishnan, S., He, L., Maranas, C. D., Pakrasi, H. B., Allen, D. K., and Tang, Y. J. (2017) Deciphering cyanobacterial phenotypes for fast photoautotrophic growth via isotopically nonstationary metabolic flux analysis. *Biotechnol. Biofuels 10*, 1–13.

(296) Watanabe, S., Ohbayashi, R., Kaneski, Y., Saito, N., Chibazakura, T., Soga, T., and Yoshikawa, H. (2015) Intensive DNA replication and metabolism during the lag phase in Cyanobacteria. *PLoS One 10*, 1–16.

(297) Formighieri, C. (2015) Solar-to-Fuel Conversion in Algae and Cyanobacteria 1st ed. Springer International Publishing.

(298) Vioque, A. (2007) Transformation of cyanobacteria, in *Transgenic Microalgae as Green Cell Factories. Advances in Experimental Medicine and Biology, vol 616.* (León R., Galván A., F. E., Ed.), pp 12–22. Springer, New York, NY.

(299) Clerico, E. M., Ditty, J. L., and Golden, S. S. (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria, in *Circadian Rhythms. Methods in Molecular Biology, vol 362* (Rosato, E., Ed.), pp 155–171. Humana Press.

(300) Begemann, M. B., Zess, E. K., Walters, E. M., Schmitt, E. F., Markley, A. L., and Pfleger, B. F. (2013) An Organic Acid Based Counter Selection System for Cyanobacteria. *PLoS One 8*, 1–12.

(301) Tan, X., Liang, F., Cai, K., and Lu, X. (2013) Application of the FLP/FRT recombination system in cyanobacteria for construction of markerless mutants. *Appl. Microbiol. Biotechnol. 97*, 6373–6382.

(302) Zhang, Y., Pu, H., Wang, Q., Cheng, S., Zhao, W., Zhang, Y., and Zhao, J. (2007) PII is important in regulation of nitrogen metabolism but not required for heterocyst formation in the cyanobacterium Anabaena sp. PCC 7120. *J. Biol. Chem. 282*, 33641–33648.

(303) Caicedo-Burbano, P., Smit, T., Pineda Hernández, H., Du, W., and Branco Dos Santos, F. (2020) Construction of

Fully Segregated Genomic Libraries in Polyploid Organisms Such as Synechocystis sp. PCC 6803. *ACS Synth. Biol. 9*, 2632–2638.

(304) de Jong, H., Geiselmann, J., and Ropers, D. (2017) Resource Reallocation in Bacteria by Reengineering the Gene Expression Machinery. *Trends Microbiol. 25*, 480–493.

(305) Vickers, C. E., Bongers, M., Liu, Q., Delatte, T., and Bouwmeester, H. (2014) Metabolic engineering of volatile isoprenoids in plants and microbes. *Plant, Cell Environ. 37*, 1753–1775.

(306) Zhuang, K., Vemuri, G. N., and Mahadevan, R. (2011) Economics of membrane occupancy and respiro-fermentation. *Mol. Syst. Biol. 7*, 1–9.

(307) Nicolaou, S. A., Gaida, S. M., and Papoutsakis, E. T. (2010) A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. *Metab. Eng. 12*, 307–331.

(308) Renda, B. A., Hammerling, M. J., and Barrick, J. E. (2014) Engineering reduced evolutionary potential for synthetic biology. *Mol. Biosyst. 10*, 1668–1678.

(309) Beneyton, T., Coldren, F., Baret, J. C., Griffiths, A. D., and Taly, V. (2014) CotA laccase: High-throughput manipulation and analysis of recombinant enzyme libraries expressed in E. coli using droplet-based microfluidics. *Analyst 139*, 3314–3323.

(310) Guo, H. H., Choe, J., and Loeb, L. A. (2004) Protein tolerance to random amino acid change. *Proc. Natl. Acad. Sci. U. S. A. 101*, 9205–9210.

(311) Watt, V. M., Ingles, C. J., Urdea, M. S., and Rutter, W. J. (1985) Homology requirements for recombination in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A. 82*, 4768–4722.

(312) Moxon, R., Bayliss, C., and Hood, D. (2006) Bacterial contingency loci: The role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet. 40*, 307–333.

(313) Mirkin, E. V., and Mirkin, S. M. (2007) Replication Fork Stalling at Natural Impediments. *Microbiol. Mol. Biol. Rev. 71*, 13–35.

(314) Kazazian, H. H. (2004) Mobile Elements: Drivers of Genome Evolution. *Science (80-.). 303*, 1626–1632.

(315) Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) The complete genome sequence of Escherichia coli K-12. *Science (80-.). 277*, 1453–1462.

(316) Grabherr, R., Nilsson, E., Striedner, G., and Bayer, K. (2002) Stabilizing plasmid copy number to improve recombinant protein production. *Biotechnol. Bioeng. 77*, 142–147.

(317) Galhardo, R. S., Hastings, P. J., and Rosenberg, S. M. (2007) Mutation as a Stress Response and the Regulation of Evolvability. *Crit. Rev. Biochem. Mol. Biol. 42*, 399–435.

(318) Holliday, R., and Grigg, G. W. (1993) DNA methylation and mutation. *Mutat. Res. - Fundam. Mol. Mech. Mutagen. 285*, 61–67.

(319) Lee, H., Popodi, E., Tang, H., and Foster, P. L. (2012) Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *Proc. Natl. Acad. Sci. U. S. A. 109*, E2774–E2783.

(320) Darmon, E., and Leach, D. R. F. (2014) Bacterial Genome Instability. *Microbiol. Mol. Biol. Rev. 78*, 1–39.

(321) Vapnek, D., Alton, N. K., Bassett, C. L., and Kushner, S. R. (1976) Amplification in Escherichia coli of enzymes involved in genetic recombination: Construction of hybrid ColE1 plasmids carrying the structural gene for exonuclease I. *Proc. Natl. Acad. Sci. U. S. A.*

(322) Pósfai, G., Plunkett, G., Fehér, T., Frisch, D., Keil, G. M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S. S., Arruda, M. de, Burland, V., Harcum W., S., and Blattner, F. R. (2006) Emergent Properties of Reduced-Genome Escherichia coli. *Science (80-.). 312*, 1044–1046.

(323) Csörgo, B., Fehér, T., Tímár, E., Blattner, F. R., and Pósfai, G. (2012) Low-mutation-rate, reduced-genome Escherichia coli: An improved host for faithful maintenance of engineered genetic constructs. *Microb. Cell Fact. 11*, 1–13.

(324) Shen-Orr, S. S., Milo, R., Mangan, S., and Alon, U. (2002) Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet. 31*, 64–68.

(325) Bulter, T., Lee, S. G., Wong, W. W. C., Fung, E., Connor, M. R., and Liao, J. C. (2004) Design of artificial cell-cell communication using gene and metabolic networks. *Proc. Natl. Acad. Sci. U. S. A. 101*, 2299–2304.

(326) Brockman, I. M., and Prather, K. L. J. (2015) Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnol J. 10*, 1360–1369.

(327) Liu, D., Evans, T., and Zhang, F. (2015) Applications and advances of metabolite biosensors for metabolic engineering. *Metab. Eng. 31*, 35–43.

(328) Feng, J., Jester, B. W., Tinberg, C. E., Mandell, D. J., Antunes, M. S., Chari, R., Morey, K. J., Rios, X., Medford, J. I., Church, G. M., Fields, S., and Baker, D. (2015) A general strategy to construct small molecule biosensors in eukaryotes. *Elife 4*, 1–23.

(329) Younger, A. K. D., Dalvie, N. C., Rottinghaus, A. G., and Leonard, J. N. (2017) Engineering Modular Biosensors to Confer Metabolite-Responsive Regulation of Transcription. *ACS Synth. Biol. 6*, 311–325.

(330) Xiong, D., Lu, S., Wu, J., Liang, C., Wang, W., Wang, W., Jin, J. M., and Tang, S. Y. (2017) Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metab. Eng. 40*, 115–123.

(331) Tsao, C. Y., Hooshangi, S., Wu, H. C., Valdes, J. J., and Bentley, W. E. (2010) Autonomous induction of recombinant proteins by minimally rewiring native quorum sensing regulon of E. coli. *Metab. Eng. 12*, 291–297.

(332) Zhang, F., Carothers, J. M., and Keasling, J. D. (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol. 30*, 354–359.

(333) Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., Petzold, C. J., Mukhopadhyay, A., Lee, T. S., Adams, P. D., and Keasling, J. D. (2013) Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol. 31*, 1039–1046.

(334) Skjoedt, M. L., Snoek, T., Kildegaard, K. R., Arsovska, D., Eichenberger, M., Goedecke, T. J., Rajkumar, A. S., Zhang, J., Kristensen, M., Lehka, B. J., Siedler, S., Borodina, I., Jensen, M. K., and Keasling, J. D. (2016) Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol. 12*, 951–958.

(335) Bothfeld, W., Kapov, G., and Tyo, K. E. J. (2017) A Glucose-Sensing Toggle Switch for Autonomous, High Productivity Genetic Control. *ACS Synth. Biol. 6*, 1296–1304.

(336) Peng, B., Plan, M. R., Carpenter, A., Nielsen, L. K., and Vickers, C. E. (2017) Coupling gene regulatory patterns to bioprocess conditions to optimize synthetic metabolic modules for improved sesquiterpene production in yeast. *Biotechnol. Biofuels 10*, 1–16.

(337) Hirose, Y., Shimada, T., Narikawa, R., Katayama, M., and Ikeuchi, M. (2008) Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein. *Proc. Natl. Acad. Sci. U. S. A. 105*, 9528–9533.

(338) Higo, A., Isu, A., Fukaya, Y., and Hisabori, T. (2017) Designing synthetic flexible gene regulation networks using RNA devices in cyanobacteria. *ACS Synth. Biol. 6*, 55–61.

(339) Behler, J., Vijay, D., Hess, W. R., and Akhtar, M. K. (2018) CRISPR-Based Technologies for Metabolic Engineering in Cyanobacteria. *Trends Biotechnol. 36*, 996–1010.