UNIVERSITY OF CALIFORNIA

Los Angeles

Towards improving carbon fixation in plants: Cyanobacteria as a model organism

A dissertation submitted in partial satisfaction of the

requirements for the Degree of Philosophy

in Chemical Engineering

by

Fabienne Duchoud

2016

ABSTRACT OF THE DISSERTATION

Towards improving carbon fixation in plants: Cyanobacteria as a model organism

by

Fabienne Duchoud

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2016

Professor James C. Liao, Chair

With the expanding global population and the associated growing demand for food and energy, researchers have been looking into ways to improve plants productivity, and more particularly the way they convert carbon dioxide to biomass during photosynthesis. Targeting carbon capture efficiency seems a good opportunity to raise crop productivity both for food and biomass-based biofuels. Most approaches have focused on optimizing the CBB cycle and more particularly the

key carboxylating enzyme, Rubisco, because of its rather low catalytic rate and poor substrate specificity. Here we propose an alternative approach by engineering an orthogonal artificial carbon fixation pathway into photosynthetic organism. We choose cyanobacteria, a model system for phototrophic eukaryotes, while also implementing the new pathway into the model plant organism, *Arabidopsis thaliana*. The new CO_2 fixation cycle would not rely on Rubisco and would fix carbon with only 30% of the energy requirement compared to the native CBB cycle. Over the course of this study we encountered a number of logistical problems that thwarted the implementation of the orthogonal carbon fixation cycle. First, the Rubisco mutant strain that could have served as a platform for efficient screening of our pathway was too sick to be used as such. Secondly, our pathway relied on a pyruvate ferredoxin:oxidoreductase for one of its carboxylation step, which is known to be oxygen sensitive. Efforts to evolve an oxygen tolerant mutant of this enzyme were unsuccessful. Lastly, a possible kinetic trap that could prevent the cycle from running was identified. The accumulation of setbacks forced us to modify our initial pathway cycle design. The final pathway, termed rGS2, does not use pyruvate ferredoxin:oxidoreductase for carbon fixation and avoids the kinetic trap. It does lose any energy efficiency gain over the CBB cycle but could nevertheless still increase carbon fixation efficiency as it does not solely depend on Rubisco. Implementation of the rGS2 in cyanobacteria increased the intracellular acetyl-CoA pool and led to secretion of α- ketoisocaproate. The engineered strain was genetically unstable, which was probably the result of toxic intermediates being built up from the rGS2 pathway. In the future, we hope to optimize the rGS2 pathway by testing new homologues enzymes or alternative steps to improve the rGS2 related toxicity.

In plants, the original synthetic carbon fixation pathway, the rGS cycle, was engineered in both wild type *Arabidopsis* and in two mutant strains defective in their endogenous CBB cycle. In the

three different backgrounds, mutants with enhance biomass and/or height were identified, suggesting that the rGS pathway may be functional and could be working as an optional pathway for CO_2 fixation. But like in cyanobacteria, genetic instability was observed, and genotyping of the T3 mutant plants could not confirm the presence of all heterologous rGS genes, possibly the results of DNA recombination between homologous regions on the T-DNA causing some of the fragment to loop out.

Numerous obstacles were encountered throughout the course of this work, but the results and lessons learned could later be applied to build more robust orthogonal carbon fixation cycles, and improve production of downstream pathways.

.

The dissertation of Fabienne Duchoud is approved.

Katrina Mae Dipple

Tatiana Segura

Philip A. Romero

James C. Liao, Committee Chair

University of California, Los Angeles

2016

To my loving family

Table of Content

Acknowledgements

I would like to thank my PI, Dr. James C. Liao as well as the DOE ARPA-E for the financial support (PETRO grant Award DE-AR0000201), without whom this project would not have been possible.

A large part of this project was accomplished as a group and could not have been done without all the people that took part in it. From the cyanobacteria part team I would like to thank Dr. Luisa Gronenberg, Hannah Park, Hans Sebastian and Xiaoqian Li. From the plants team I would like to thank Reem Elteriefi, Dr. Molly Leung, Dr. Samuel Mainguet and Dr. Avinash Srivastava which performed all the plant plasmids construction and/or plants transformation and early screening.

I also would like to thank current and past members of the lab for their support, more particularly Paul Lin for all his pertinent advices and countless late motorcycle rides, and specialy Paul Opgenorth for always being there for me and for patiently editing the long chapters I sent him while writing this thesis.

I would lastly like to thank my family. Despite being physically far they were always there for me, to support me and give me the strength and mental support to continue and never give up even in the hardest times.

VITA

Education

2006: B.S. in Chemical and Biochemical Engineering, Swiss Federal Institute of Technology, Lausanne.

2008: M.S. in in Chemical and Biochemical Engineering, Swiss Federal Institute of Technology, Lausanne.

Publication

Duchoud Fabienne, Derrick S.W. Chuang, James C. Liao, “Cyanobacteria as a Host Organism”, *Industrial Biotechnology: Microorganisms*, Ed. Christoph Wittmann, Ed. James Liao. Wiley, 2016.

1. Introduction

The increasing world population and concomitant rise in food and energy demand have raised attention and efforts towards increasing the efficiency and productivity of photosynthesis in crop use for food or energy production. Photosynthesis is the process used by plants, algae and certain bacteria like cyanobacteria, to convert solar energy into chemical energy by converting CO_2 and water into sugars. Current photosynthesis represents the upper limit of the theoretical yield for plant-based conversion of solar to chemical energy[1]. Improving photosynthesis would not only contribute to better food security in the future, but would also help in the development of low-cost production of biomass-based biofuel[2] which is gaining increasing attention as a source of renewable energy. To date, however, increasing the photosynthetic efficiency has only played a minor role in improving the yield potential of crops, though several potential target enzymes have been identified[3,4]. Photosynthetic can be broken up into two parts commonly referred to as the light and dark reactions. The light reactions involve the electron transport chain which serves to harvest light and converts it into chemical energy in the form of ATP. The dark reactions of photosynthesis serve to run the CBB cycle and turn ATP and CO_2 into sugars which are the main carbon and energy storage molecules of the cell. The light reactions of photosynthesis are catalyzed by complex group of mainly membrane proteins on which basic research is still being performed to more fully understand. The dark reactions however are well known, extensively studied soluble proteins which can potentially be improved or swapped out to increase the overall photosynthetic efficiency of an organism. The most important candidate is the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), which catalyzes the first major step of the Calvin Benson Bassham (CBB) cycle (Fig. 1-1A), pathway used by all photosynthetic

organisms for carbon fixation[5]. Despite its central role, Rubisco is thought to be the primary rate limiting step of the CBB cycle because of its general slow catalytic turnover and substrate promiscuity reacting with both CO_2 and O_2. Each Rubisco molecule only fix 3-10 CO_2 per second, while typical enzymes of central metabolism can carry out thousands of reactions in the same time[6]. In addition, Rubisco can react with O_2, in a process called photorespiration, which reduces the overall carbon fixation efficiency and wastes energy in the form of ATP. Previous work to improve CO_2 fixation have shown that improving specificity via site-directed mutagenesis comes with a lost in the catalytic rate and vise-versa. Furthermore, different forms of the enzymes are found in nature which differ in their kinetic properties, such as their CO_2/O_2 specificity, substrate affinity and catalytic efficiency[7], also indicating some evolution related to the surroundings of the organism they belong to. These results led to the popular opinion that Rubisco has already been naturally optimized for the best compromise between substrate specificity and maximum rate of catalytic turnover[8,9] and would be a bottleneck enzyme for the CBb cycle. Thus, further improvement of carbon fixation needs to concentrate on other targets. Some studies focusing on overexpression of CBB cycle enzymes showed some encouraging results, but others were inconclusive[10]. This suggests that improving the existing CBB cycle and more generally carbon fixation, without fundamentally changing the pathway network structure is unlikely to be successful. Beside the CBB cycle, nature has evolved five alternative autotrophic pathways found in some prokaryotes, including the reductive tricarboxylic acid (TCA) cycle, the reductive acetyl-CoA (Woods-Ljungdahl) pathway, the 3-hydroxypropionate/glyoxylate cycle, and two related cycles: the 3-hydroxypropionate /4-hydroxybutyrate cycle and the dicarboxylate/4-hydroxybutyrate cycle. All these alternative pathways use energy (ATP) more efficiently (Table 1-1) than the CBB cycle, but some are not

well characterized. The reductive TCA cycle (Fig. 1-1C) enzymes have been fully identified and characterized, and possess very high energy efficiency. The reductive TCA cycle is found in diverse autotrophic bacteria and archea, all anaerobes or microaerobic hyperthermophiles[11]. The stoichiometry of the rTCA cycle allows for the fixation of two CO_2 molecules which produce one acetyl-CoA which can be further carboxylated to pyruvate, a key metabolite in central metabolism. In addition, an artificial energy efficient carbon fixation cycle could be built upon a synthetic pathway already developed in our laboratory, called the reverse glyoxylate shunt (rGS)[12]. The rGS allows for conversion of malate and succinate to acetyl-CoA without carbon loss (Fig. 1-1B). Extension of this pathway by addition of five enzymatic steps would form a complete carbon fixation cycle, called the reverse glyoxylate shunt cycle (rGS cycle). Because of our previous work, we favored the rTCA cycle and the artificial rGS cycle, as potential carbon fixation pathways to increase the total carbon fixation efficiency in plants. The rTCA and rGS cycles require only 2 and 3 ATP molecules respectively to produce one molecule of pyruvate from CO_2, as compared to 7 ATP molecules required for the CBB cycle. In addition, these carbon fixation cycles do not utilize Rubisco for CO_2 fixation, and thus do not have the photorespiration associated carbon and energy loss. The reductive TCA cycle is slightly more efficient than the rGS, but involves two oxygen sensitive ferredoxin-oxidoreductase enzymes that have to be compatible with the reduce ferredoxin generated by the host photosystems. The other option, rGS, fixes CO_2 through only one oxygen sensitive ferredoxin-oxidoreductase, the other carboxylation step being either a pyruvate carboxylate from bacteria or the native phosphoenolpyruvate carboxylase in plants. Instead of producing sugars like the CBB cycle, both pathways produce pyruvate, a precursor for a wide variety of liquid fuels.

The overall purpose of this project is to improve carbon fixation efficiency in photosynthetic organisms through metabolic engineering. This will be accomplished through implementation an alternative carbon fixation pathway and evaluation of its compatibility with the light reactions of photosynthesis which are already present. The new pathway has the potential to improve carbon fixation efficiency and channels the carbon flux to pyruvate, a direct precursor of biofuels. The pathway will first be tested in model photosynthetic organisms and, if successful, could later be incorporated into plants, to improve the productivity of both food and fuel crops.

Chapter 2 is an introduction to cyanobacteria as a simple biological model organism for plants and carbon fixation, as well as for CO_2 conversion to different industrially relevant bioproducts.

Chapter 3 will discuss the construction of a mutant cyanobacteria strain with impaired Rubisco activity that could serve as a platform for the implementation and analysis of a new carbon fixation pathway. Amichay et al. had previously demonstrated a *Synechocystis* sp.PCC 6803 mutant with inactivated *rbcL* the functional replacement of the Rubisco operon with a *Rhodospirullum rubrum* gene encoding a homodimer Rubisco. We constructed a similar mutant from wild-type *Synechocystis* that shows impaired growth under atmospheric CO_2 concentration. Although the mutant had the desire phenotype for identification of potential complementary carbon fixation pathways, it could not be used as a screening platform because of the inability to further transform and engineer the alternative carbon fixation pathways into it.

Chapter 4 describes the attempts at the implementation of a new carbon fixation pathway, the reverse glyoxylate shunt (rGS) cycle, in cyanobacteria as well as the challenges encountered. The full rGS requires two enzymes believed to be oxygen sensitive, fumarate reductase and pyruvate ferredoxin oxydoreductase, which is problematic for implementation into an organism

like cyanobacteria that generates oxygen as a side product of photosynthesis. Fumarate reductase is generally only expressed anaerobically[13], and pyruvate ferredoxin oxydoreductase is known to be highly sensitive to oxygen[14]. Moreover computational modeling of the rGS cycle allowed us to identify a possible kinetic trap that may prevent it from fully functioning. These led us to slightly modify the design of the cycle to a linear pathway that could theoretically allow for more efficient acetyl-CoA production from CO_2. In this new linear pathway acetyl-CoA could be later channeled into industrially relevant compounds such as isoprenoids, fatty acids or long chain alcohols. This new approach did not give rise to the expecting results but allowed us to deduct that part of the rGS pathway may have been active.

Chapter 5 will discuss a new variation of the rGS termed the reverse glyoxylate shunt 2 (rGS2). This design maintains the three first steps of the rGS pathway to produce glyoxylate and acetyl-CoA, but uses a new route to recycle glyoxylate, inspired by the glycerate pathway. It allows to avoid both the kinetic trap seen in the rGS cycle as well as the two oxygen sensitive enzymes from rGS linear pathway. We will describe the improvement made on the three first steps of the pathway, followed by how the new recycling part allowed us to produce α-ketoisocaproate from CO_2, the final intermediate metabolite in leucine synthesis and a potential therapeutic in patients with chronic kidney disease as part of keto-acid supplementation.

Chapter 6 describes the implementation of the reverse glyoxylate shunt (rGS) into the plant model *Arabidopsis thaliana*. We identified mutant lines transformed with the rGS genes that showed enhanced biomass and/or height compare to their parent lines, wild-type *Arabidopsis* or mutant strains defective in their endogenous CBB cycle. The integrated genes were however unstable and we could not confirm the expected genotype.

1.1 Tables and Figures

Pathway	CO_2	HCO_3^-	Reducing equivalents	ATP equivalents
CBB	3	0	5	7
reductive acetyl-CoA	3	0	5	1
3HP/glyoxylate	0	3	5	7
3HP/4HB	1	2	5	6
decarboxylate/4HB	2	1	5	5
rTCA	3	0	5	2
rGS*	2	1	5	3

Table 1-1: Autotrophic carbon fixation pathways. CBB, Calvin Benson Bassham cycle; 3HP, 3-hydroxypropionate, 4HB, 4-hydroxybutyrate; rTCA, reductive tricarboxylic acid cycle; rGS, reverse glyoxylate shunt cycle. * The rGS cycle is the only synthetic carbon fixation pathway presented in the table.

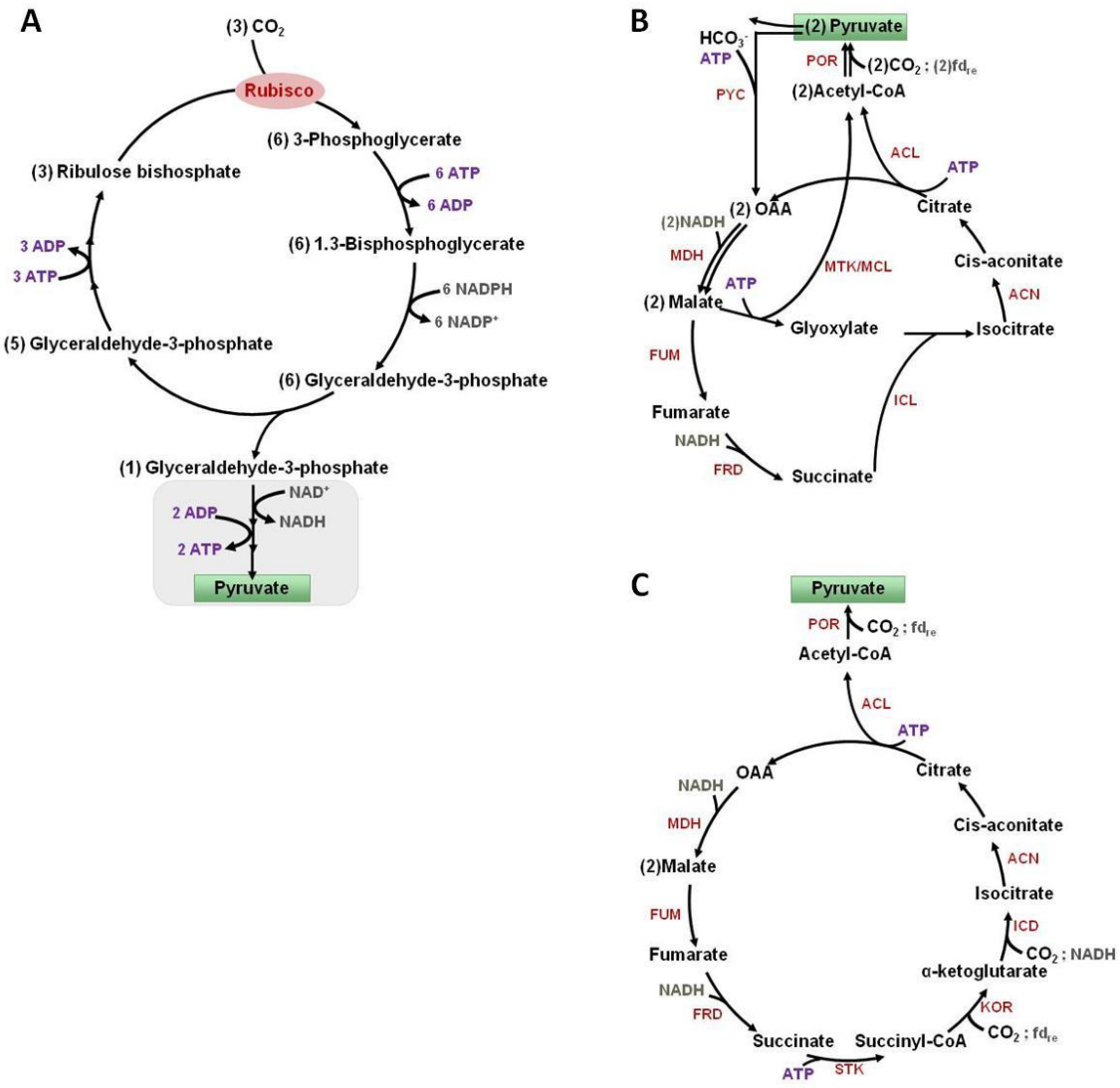


Figure 1-1. Schematic representation of the Calvin Benson Bassham (CBB) cycle (A), the reverse Glyoxylate Shunt (rGS) cycle (B) and of the reductive tricarboxylic acid (rTCA) cycle (C) for production of one molecule of pyruvate, highlighted in green. The CBB cycle direct product is glyceraldehydes-3-phosphate. Further conversion to pyrvuvate is represented in the grey box. All cycles use energy (ATP) and reducing power (NADH, NADH equivalent and reduced ferredoxin) from photosynthesis. Rubisco, ribulose 1,5-bisphosphate

carboxylase/oxygenase; PYC, pyruvate carboxylase; MDH, malate dehydrogenase; ACN, aconitase; FUM, fumarase; FRD, fumarate reductase; MTK, malate thiokinase; MCL, malyl-CoA lyase; ICL, isocitrate lyase; ACL, acetyl-CoA lyase; POR, pyruvate ferredoxine oxidoreductase, STK, succinate thiokinase; KOR, α-ketoglutarate oxidoreductase; ICD, isocitrate dehydrogenase, fd_{re}, reduced ferredoxin.

1.2 References

1. Zhu, X.-G., Long, S. P. & Ort, D. R. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr. Opin. Biotechnol.* **19,** 153–159 (2008).
2. Himmel, M. E. *et al.* Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **315,** 804–807 (2007).
3. Zarzycki, J., Axen, S. D., Kinney, J. N. & Kerfeld, C. A. Cyanobacterial-based approaches to improving photosynthesis in plants. *J. Exp. Bot.* **64,** 787–798 (2013).
4. Evans, J. R. & Caemmerer, S. von. Enhancing Photosynthesis. *Plant Physiol.* **155,** 19–19 (2011).
5. H M Miziorko & Lorimer, and G. H. Ribulose-1,5-Bisphosphate Carboxylase-Oxygenase. *Annu. Rev. Biochem.* **52,** 507–535 (1983).
6. Ellis, R. J. Biochemistry: Tackling unintelligent design. *Nature* **463,** 164–165 (2010).
7. Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. & Scott, S. S. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* **59,** 1515–1524 (2008).
8. Tcherkez, G. G., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci.* **103,** 7246–7251 (2006).
9. Ninomiya, N., Ashida, H. & Yokota, A. in *Photosynthesis. Energy from the Sun* 867–870 (Springer, 2008).

10. Ruan, C.-J., Shao, H.-B. & Silva, J. A. T. da. A critical review on the improvement of photosynthetic carbon assimilation in C3 plants using genetic engineering. *Crit. Rev. Biotechnol.* **32,** 1–21 (2012).

11. Hügler, M., Wirsen, C. O., Fuchs, G., Taylor, C. D. & Sievert, S. M. Evidence for Autotrophic CO2 Fixation via the Reductive Tricarboxylic Acid Cycle by Members of the ε Subdivision of Proteobacteria. *J. Bacteriol.* **187,** 3020–3027 (2005).

12. Mainguet, S. E., Gronenberg, L. S., Wong, S. S. & Liao, J. C. A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab. Eng.* **19,** 116–127 (2013).

13. Jones, H. M. & Gunsalus, R. P. Transcription of the Escherichia coli fumarate reductase genes (frdABCD) and their coordinate regulation by oxygen, nitrate, and fumarate. *J. Bacteriol.* **164,** 1100–1109 (1985).

14. Horner, D. S., Hirt, R. P. & Embley, T. M. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Mol. Biol. Evol.* **16,** 1280–1291 (1999).

2. Cyanobacteria as a Host Organism

Disclaimer: This chapter was accepted in February 2016 to be published with the same title in the Wiley book series *Industrial Biotechnology: Microorganisms*.

2.1 Abstract

Cyanobacteria are a diverse group of photosynthetic organisms, the only prokaryotes capable of performing plant-like photosynthesis. Their minimal growth requirements, established genetic manipulation tools and carbon neutral production process make them promising candidates for sustainable production of bulk products such as biofuels and chemicals. Efforts still need to be made in the development of improved strains with more efficient carbon fixation, faster conversion of CO_2 to products and increased yield. The recent efforts and progresses in biotechnology are facilitating the path towards realizing the potential of cyanobacteria as an industrial platform for the production of a broad range of products.

2.2 Introduction

2.2.1 Introduction and relevance: cyanobacteria as a host organism

Cyanobacteria, a group of prokaryotes previously known as blue-green algae, belong to the most ancient group of photosynthetic organisms and are the only prokaryotes capable of carrying out plant like oxygenic photosynthesis. In fact, plants chloroplasts are considered to be originated from cyanobacteria through endosymbiosis. With a theoretical maximum photosynthetic

efficiency in the range of 8-10% [1], compare to 4.6-6% for terrestrial plants [2], they convert solar energy, CO_2 and water into chemicals with only minimal nutrients requirement. Cyanobacteria exhibit a large morphological diversity and are comprised of unicellular, filamentous, colonial and planktonic types. They can thrive in a broad range of habitats, from aquatic ecosystems, including fresh and marine water, to terrestrial ecosystems. As a result of this large biodiversity, cyanobacteria produce a wide array of molecules, which have gained great importance in the biotechnological and chemical industries, with compounds that may act as pesticide, antibacterial, antiviral, and antifungal [3]. Moreover because cyanobacteria are susceptible to genetic transformation, recent investment has been made for their potential use as a platform for production of biofuels and chemicals. In addition to having superior photosynthesis capabilities to plants, cyanobacteria cultivation does not require agricultural land, minimizing competition with food crops. Extensive research has been carried out mainly in the *Synechococcus* and *Synechocystis* genera, with some lesser effort accomplished in *Anabena* and *Cyanothece*.

This chapter first gives a general introduction to cyanobacteria physiology as well as their genetic tools, which are important background for utilizing these organisms for industrial applications. We then review works on the development of strains as potential cell factories for the production of commercially relevant compounds, with a focus on primary metabolites.

2.3 General description of cyanobacteria

2.3.1 A diverse bacterial group and its metabolisms

Cyanobacteria form a phylum of bacteria that can perform plant like photosynthesis. The cyanobacteria phylum includes some 2000 species from 150 genera and 5 orders, which vary greatly in shapes and sizes. All cyanobacteria are unicellular, though many grow in aggregates of

cells, in more structured colonies, or in filaments. They represent the most genetically diverse group of microorganisms and survive in a large range of habitats across the globe but are most common in freshwater, marine, and terrestrial environments. They can be found in extreme ecological niches such as deserts, hot springs and hypersaline waters. Besides their wild diversity in terms of morphology and habitats, cyanobacteria also differ in terms of physiology and metabolic capacities [4]. They have the ability, throughout different species, to synthesize a broad range of structurally and functionally diverse natural products [3].

Most cyanobacteria carry out oxygenic photosynthesis which allows them to use minimal nutrients for growth, namely carbon dioxide, water and some salts as their carbon and nitrogen sources, while deriving energy from light. Some of them can also utilize organic compounds for heterotrophy. Many cyanobacteria fix atmospheric nitrogen under what seems to be aerobic conditions [5]. This is surprising as nitrogenase, the enzymatic complex responsible for nitrogen fixation, is very oxygen sensitive. Cyanobacteria deal with that problem by separating, either spatially or temporarily, the two processes of oxygenic photosynthesis and nitrogen fixation. In filamentous strains the nitrogenase is confined to heterocysts, differentiated thick-walled cells that lack an oxygenic photosystem and are the site of N_2 fixation. Other species, unicellular as well as filamentous, induce nitrogenase activity only in the dark periods of light-dark growth cycles, when photosynthesis is shut down.

Even though oxygenic photosynthesis is the principal mode of growth of these organisms, some species can also be found in anoxic environments [6]. These cyanobacteria are able to switch to sulfide dependent anoxygenic photosynthesis where inorganic sulfur compounds act as electron donors [7,8]. Furthermore, other species show a distinct ability to assimilate sugars or other organic

compounds in the light (photoheterotrophic growth) [9,10]. It is also known that certain cyanobacteria can survive for prolonged periods of time in complete darkness. They have indeed the ability to accumulate and store a range of different compounds that can be utilize later when nutrients are low, as it is the case during a dark phase. Among them is glycogen that serves as a carbon source and can be respired to generate energy, polyphosphate to provide a source of phosphorus, cyanophycin to supply nitrogen, and poly-β-hydroxybutyrate, even though its function is uncertain [11]. Under dark anoxic condition, some species can perform fermentation for maintenance requirements [12].

2.3.2 Nitrogen fixation

Cyanobacteria are one of the few organisms, all prokaryotes, capable of fixing atmospheric nitrogen, thus decreasing their dependence on reduced nitrogen sources [13]. They play a major role in the global nitrogen cycle. Many strains of cyanobacteria, termed diazotrophs, are able to fix N_2. Because they simultaneously produce oxygen through photosynthesis and can fix nitrogen, they possess particularly well developed mechanisms to protect the extremely oxygen sensitive nitrogen-fixing enzyme, nitrogenase. The two main strategies employed are: spatial separation and temporal separation.

Spatial separation is mainly used by filamentous heterocysts-forming cyanobacteria [14]. When the cells are deprived of dissolved inorganic nitrogen, some of them differentiate into heterocysts in a semiregularly spaced pattern along the filament. These thick-walled cells are the site of N_2 fixation. They have decreased permeability to O_2 and high rate of respiration to protect nitrogenase from oxygen inactivation. They also only contain photosystem I for energy

production in the form of ATP, but lack photosystem II to avoid oxygen production [15]. Heterocysts rely on vegetative cells for carbon sources as they cannot fix CO_2.

While all heterocystous cyanobacteria can fix nitrogen through nitrogenase, only some non-heterocystous cyanobacteria can (see [16] for a review of these strains). In addition, a gradation of tolerance to O_2 exists among those strains capable of N_2 fixation. Most do it only under micro-oxic or anoxic conditions, yet a few can fix nitrogen aerobically [16]. The main strategy employed to prevent inactivation of nitrogenase by O_2 is to temporally separate photosynthesis and its associated oxygen production from nitrogen fixation. Typically, nitrogen is fixed in the dark while photosynthetic O_2 production occurs in the light [17–20]. During the transition from light to dark phase, cellular respiration plays a critical role, consuming the intracellular oxygen generated by photosynthesis to reach low enough O_2 concentration levels to allow for nitrogenase activity. Besides serving as a metabolic strategy for oxygen consumption, respiration also sustains the process of nitrogen fixation by generating the reducing equivalents and ATP needed by this energy-intensive process [21]. In many cases, the pattern of nitrogenase activity and respiration indeed overlap in cultures grown under alternating light and dark periods [22]. In the dark, respiration relies on the glycogen stored during the photosynthetic light period as a carbon source [23], while nitrogen is accumulated as cyanophycin that can later be used under light conditions [24].

Only a few cyanobacteria do not follow these two main nitrogen fixation strategies. Members of the genus *Trichodesmium* show nitrogenase activity peaking in the light [25], although they do not technically produce heterocyst. A lot of ongoing studies have been conducted to elucidate how *Trichodesmium* fix both CO_2 and N_2 concomitantly without any obvious spatial separation. It seems that part of the solution is achieved by restricting nitrogenase localization to a subset of

cells along the trichomes where photosynthesis is downregulated [26]. Temporal separation may also play a role in metabolic partition of CO_2 and N_2 fixation in *Trichodesmium* as carbon fixation rate is higher in the morning and nitrogen fixation peaks in the afternoon [27].

2.3.3 Circadian clock

Circadian rhythms are endogenous oscillations of about 24 hours a cycle observed in many physiological activities in a wide range of organisms [28]. These rhythms are controlled by a circadian clock assumed to be a fundamental function of most living organisms to cope with the predictable environment changes that come with the daily day/night cycle [29]. Circadian time keeping was initially assumed to be restricted to eukaryotes, until 1986 when Huang and coworkers demonstrated bacterial circadian rhythms in the cyanobacteria *Synechococcus* sp. RF-1 for nitrogen fixation first [30,31] and later amino acid uptake [32]. These rhythms persisted for some time in absence of external cues, were reset by light signals and were temperature compensated, satisfying the three characteristics of a true circadian oscillator [33]. Since then, circadian rhythms have been identified in a number of cyanobacterial strains, with *Synechococcus elongatus* PCC7942 emerging as the main model system to further study the circadian clock mostly because of its genetic tractability and available genetic tools. To date, cyanobacteria remain the simplest organisms and only prokaryotes known to have a true circadian clock.

In *S. elongatus* the *kai* gene cluster, comprised of *kaiA, kaiB* and *kaiC,* is responsible for the generation of circadian oscillations [34]. It has been shown that almost all genes in this organism are rhythmically controlled by these 3 genes [35]. Most of them can be divided into two groups with either high expression late in the day like genes involved in nitrogen fixation or high expression in the dark phase just before the next light phase comes around such as genes for

glycogen synthesis. Additionally, the circadian clock also regulates cell division timing [36,37] as well as chromosome compaction [38,39]. A full review on cyanobacterial circadian clock has recently been published [40].

2.3.4 Light/dark regulation

The physiology of cyanobacteria as a photosynthetic organism is greatly impacted by the availability of light in its environment. Cyanobacteria are one of the few groups capable of carrying out both photosynthesis and respiration in the same cellular compartment. Light plays a role in the regulation of both pathways. Under high light intensity, photosynthesis prevails, and its electron flow capacity is greater than that of the respiratory chain. Under low light or in complete dark, the tendency is reversed [41].

Another response to changes in light quality and quantity occurs at the level of the light-harvesting apparatus of cyanobacteria. The light harvesting complexes surround the reaction center of a photosystem, enhancing light absorption due to their high pigment density [42]. To adapt to the different lights in their natural environment the antenna of the light harvesting apparatus can adjust the pigment composition, pigment organization and size [43,44]. These strategies tend to be long-term acclimations (hours) to light conditions, while on the short term (seconds to minutes), modifications in pigment-pigment interactions or rapid chemical or structural variation of pigments help to survive when the variations of light intensity becomes too high [45].

Finally, physiological adaptations between light and dark metabolism involve a diversity of mechanisms for the activation/inactivation of enzymes controlled by the redox state of the cell [46]. Disulphide/dithiol exchange catalysed by thioredoxins is a classic example of redox regulation

already well-established in higher plants [47]. In cyanobacteria, despite a presumable common origin, the functions of thioredoxins were poorly understood until recently. As in plant chloroplasts, cyanobacterial thioredoxins receive reducing equivalent from the photosynthetic electron transport chain via the ferredoxin dependent thioredoxin reductase [48]. Screening for potential thioredoxin target proteins has led to the *in vitro* identification of 77 candidates thioredoxin target proteins in *Synechocystis sp. PCC 6803* [49]. Only a small portion, approximately a third, has homologues in other photosynthetic organisms. These common thioredoxin targets are known in plants to control carbon fixation and storage, nitrogen and sulfur metabolism, stress response and protein translation and folding. The other proteins identified as possible receptor of redox signals might have evolved separately and many still need to be investigated to evaluate their significance *in vivo* as true targets of thioredoxin regulation.

2.4 Genetic tools

2.4.1 Transformation

Introduction of foreign DNA into cyanobacteria has been demonstrated in laboratory for several strains, and is now a common practice [50]. A few unicellular cyanobacteria are naturally competent for transformation, and can uptake foreign DNA from their environment in the form of plasmid or linear DNA [51]. Amongst naturally competent strains are the model freshwater cyanobacteria *Synechococcus elongatus* PCC 7942 and *Synechocystis* PCC 6803, as well as the marine *Synechococcus* PCC 7002 and the thermophile *Thermosynechococcus elongatus* BP-1 [52–56].

The DNA is incorporated in the cyanobacteria genome through homologous recombination when the foreign DNA is flanked by adjacent fragments from the recipient's genome. Double crossover is favored over single crossover during natural transformation, probably due to a fragmentation process of the DNA during incorporation [56]. This robust mechanism for integration of foreign DNA has allowed the generation of mutants with gene disruption or heterologous gene overexpression when the incoming DNA is incorporated into a neutral site.

Natural transformation has not been demonstrated with any filamentous cyanobacteria. However, conjugation with *E. coli* carrying the mobilization gene (*mob*) has been widely used to transfer plasmid DNA into these strains, particularly in cyanobacteria from the genus *Nostoc* and *Anabaena* [57–60]. Conjugation has also been successfully used with unicellular strains like *Synechococcus* PCC 7942, PCC6301 and *Synechocystis* PCC 6803 when transfer of an intact self replicating plasmid is preferred, but also with plasmid containing DNA that can be integrated into the genome [61,62]. In brief, this technique is based on the ability to mobilize DNA from one bacterium usually *E. coli* carrying a conjugal plasmid, to another bacterium [50]. The DNA to be transferred into the cyanobacterium is cloned on a cargo plasmid that carries a *bom* site (*oriT* region). The same plasmid, or a helper plasmid provided *in trans*, carries the *mob* gene encoding a specific nickase, that recognizes the *bom* site and cuts the plasmid to be transferred. Additionally, because restriction of incoming DNA in the recipient cell may occur, the DNA is protected by methylases cloned in a helper plasmid. As mentioned above, unlike natural transformation, conjugation results in a circular plasmid in the recipient strain. It can be either a self-replicating plasmid or can be integrated in the genome if it contains homologous flanking DNA fragment from the receiver genomic DNA. Electroporation can also be used to transform a number of cyanobacterial strains. It is generally done using plasmids developed for conjugation,

but the procedure has a higher probability to be mutagenic [63].

Most cyanobacteria strains are multiploid, carring an estimated 8-72 genome copies per cell [64], with *Synechocystis* containing up to 218 copies during exponential growth phase [65]. However, after transformation, the newly introduced DNA fragment is inserted in only one chromosome. Thus, numerous growth cycles with selection pressure are likely required to obtain complete segregation. Moreover, the choice of integration loci plays a significant role in the likeliness of achieving full segregation as disruption of an essential gene would make it impossible [66].

2.4.2 Promoters

Gene expression in cyanobacteria, whether on a self-replicating plasmid or integrated in the genome, needs to be driven by a specific promoter that can be recognized by the host. Only a limited number of promoters have been evaluated for this purpose in cyanobacteria [50]. Most known promoters are native to the host organism from genes required for photosynthesis and expressed in relatively high levels, such as P_{rbc}, P_{psbA2} and P_{cpc}, or metal-ion induced promoters like the P_{nrsB} promoter linked to nickel homeostasis in *Synechocystis* [67]. Even though some native promoters are well characterized with clear environmental triggers for regulation, their utilization is limited by potential cross-talk with the regulatory network of the organism. Several foreign promoters, orthogonal to cyanobacteria, have also been characterized for gene expression. Strong *E. coli* promoters P_{trc} and P_{tac} have been frequently used in cyanobacteria, with or without a regulated expression of LacI. But because the RNA polymerase (RNAP) holo enzymes in cyanobacteria are quite different from most bacteria, common promoters from *E.coli* might perform differently in cyanobacteria. For example, while P_{trc} is not inducible in *Synechocystis* PCC 6803, it can be regulated by LacI in *Synechococcus* 7942 [68]. Other IPTG inducible promoter

have been found to be more suited for use in *Synechocystis*, like PA1lacO-1 that exhibits an eight-fold induction ratio [69]. Recently, a library of TetR-regulated promoters with varying dynamic ranges was designed in *Synechocystis*. The best performing promoter induced a 290-fold induction ratio under red light [70]. A "super-strong" promoter, P_{cpc560}, that contains two predicted promoters and depends on 14 predicted transcription factor binding sites (TFBS) crucial for its promoter strength, has been characterized [67]. Proteins expressed under its regulation where produced to up to 15% of the total soluble protein content. In *E. coli*, strong promoters usually contain only one or two TFBS. This example illustrates how important it is to understand the difference in cellular machinery between cyanobacteria and other bacteria like *E. coli*, to design a list of cyanobacteria promoters. The structural difference of the RNAP mentioned above, as well as the different sets of sigma factors present in cyanobacteria [71], further support the need to build a basic understanding of these organism-specific differences to engineer and identify promoters.

2.4.3 Terminators

The importance of transcriptional termination in regulating natural genetic systems is well known [72], yet when it comes to characterize elements that control transcription and translation in bacteria, most research focuses on the initiation part [73,74]. Similarly, in cyanobacteria, very little work has been done to characterize terminators. It is known nevertheless, that the introduction of a terminator downstream of an overepressed gene is important for the stabilization of the transcript [75] and, in the case of chromosomal insertion, will prevent effect on the adjacent genes. In some strains, for example *Synechococcus* PCC7942, it has been reported that out of the two signals for transcription termination identified in bacteria [76], intrinsic Rho-independent

terminators play a larger role than Rho-dependent terminators [77]. In accordance with this finding, it was previously shown that most bacteria possess a homolog of the *E. coli* Rho protein, but surprisingly the cyanobacteria *S. elongatus* and *Synechocystis* PCC 6803 do not [76]. Only a few terminators, native or foreign, have been successfully used in cyanobacteria. Some examples are the bacteriophage T7 terminator, the strong terminator of the gene *rrnB* of *E. coli*, and the cyanobacteria RuBisCO terminator.

2.4.4 Ribosome binding sites

The ribosome binding site (RBS) affects the frequency with which the translation of downstream target genes begins. The Shine-Dalgarno (SD) sequence of the RBS is usually more or less conserved in bacteria such as *E. coli* where 57% of the genes contain this sequence. In *Synechocystis* PCC 6803 it was found that only 26% of the genes possess a conserved core SD sequence [78]. Recently, an RBS complementary to the anti-SD of *Synechocystis* PCC 6803 was tested and was found to be 2 to 4 fold more efficient than other RBS with core SD sequence not exactly complementary to the anti-SD [50]. The spacing between the core SD and the start codon also plays an important role in the effectiveness of an RBS. In the past few years, three independent translation rate calculators have been developed. The RBS Calculator [79], developed in 2009 and further updated in 2011, the RBS Designer [80] released in 2010 and finally the UTR Designer in 2013 [81]. Each calculator demonstrates comparable accurate prediction when compared to experimental data in *E. coli*. Unfortunately, because these three calculators were developed for *E. coli*, and because of the lack of conserved RBS sequence in cyanobacteria as mentioned above, they might not show the same accuracy if used for these organisms. It would be useful if similar models could be developed specifically for RBS prediction in cyanobacteria.

2.5 Improving photosynthetic efficiency

2.5.1 Improving light harvesting

In cyanobacteria, light is captured through light harvesting antenna, supramolecular complexes consisting of proteins and photosynthetic pigments [82]. The absorbed light can then be converted to chemical energy by the process known as photosynthesis. Although solar to biomass energy conversion of cyanobacteria is 2-3 time higher than for crop plants [83], photosynthesis efficiency is still theoretically limited to 8-10% conversion, attainable under low-light conditions [1]. Under ambient sunlight (high-light) conditions, the energy conversion efficiency drops substantially lower to 1-2% for the best species [84]. In fact, under high light, cyanobacteria cannot absorb all incoming sunlight, as the photons absorption rate by the light harvesting antennae is much faster than the rate at which photosynthesis can utilize them. This results in wasteful dissipation of excess light mostly via non-photochemical quenching (NPQ), which dissipates excess energy as heat [45]. Moreover, the shading effect of cells on the surface layer of high-density cultures causes limit light penetration for the cells in the lower layers [1]. To address this issue, it has been proposed to use strains with smaller antenna sizes. This approach can reduce both the problem of saturation at the surface and shading deeper in the culture. Even grown in high density, these mutants should still allow for good light penetration into the inner layers of the photobioreactor, thus resulting in more uniform illumination and increased whole-culture productivity. This concept, first demonstrated in green microalgae [85–87], is known as the Truncated Light-harvesting Antenna (TLA) concept [1,88,89]. Mutants expressing genetically truncated light-harvesting chlorophyll antenna size (*tla*) show enhanced solar energy conversion efficiency and productivity. RNA interference (RNAi) technology to down-regulate the whole light-harvesting

complex gene family also validated the concept independently [90]. Various strategies have been employed in cyanobacteria to create or isolate this kind of mutants. Disruption of genes encoding phycobilisome subunits first failed to produce mutants with increased biomass productivity [91]. More recently, a phycocyanin-deletion mutant of *Synechocystis sp. PCC 6803* with a highly truncated phycobilisome antenna size allowed for a 57% improvement in the productivity as compared to the wild type when grown under high cell density and saturating light [84]. Although theoretical improvements in photosynthetic energy conversion efficiency of TLA cells could be as high as 3-fold that of the wild type [1], the study demonstrates the possibility of improved productivity in cyanobacteria by reducing the light harvesting antenna size, in opposition to what was previously described [92].

2.5.2 Improving carbon fixation

Cyanobacteria use the Calvin-Benson-Bassham (CBB) cycle to fix CO_2, similar to algae and higher plants. The first step is catalyzed by ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), which converts ribulose-1,5-bisphosphate (RuBP) and CO_2 to two molecules of 3-phosphoglycerate (3PG). One 3PG is recycled to RuBP, while the other goes to central metabolic pathways. Although being the first major step in carbon fixation, RuBisCO is also the primary rate limiting factor of the CBB cycle, because of its slow catalytic turnover and low specificity. Each RubisCO molecule only fix 3-10 CO_2 per second, while some enzymes can carry out thousands of reactions in the same time [93]. In addition, RubisCO can take O_2 as a substrate in a process called photorespiration, which reduces the overall carbon fixation efficiency. Studies have shown that improving specificity via site-directed mutagenesis comes with a lost in the catalytic rate. Thus, RuBisCO is thought to have been optimized for the best

compromise between substrate specificity and maximum rate of catalytic turnover [94,95]. Moreover, cyanobacteria and some microalgae have evolved an extremely effective CO_2 concentrating mechanism (CCM) composed of at least two elements: a C_i transport system and RuBisCO-rich polyhedral carboxysomes. This mechanism allows active uptake of inorganic carbon from the external medium in order to elevate the CO_2 concentration around the active site of RuBisCO [96]. Rubisco can in turn operate at higher efficiency lowering the need for high specificity [97]. In cyanobacteria, Rubisco is encapsulated with carbonic anhydrase in unique micro-compartments, the carboxysomes. Carbonic anhydrase catalyzes the conversion of HCO_3^- to CO_2, the latter staying locked in the carboxysomes in the vicinity of RuBisCO [98]. In *Synechococcus 7942* it has been shown that cells with more carboxysomes have approximately a 50% higher carbon fixation rate than cells with fewer [99]. Expression of a heterologous RuBisCO from *Allochromatium vinosum*, a purple-sulfur bacterium, also resulted in about a 50% increase in CO_2 assimilation [100]. More recently, overexpression of *Synechococcus* 6301 RuBisCO in *Synechococcus 7942* led to a twofold higher isobutyraldehyde production, possibly due to a more efficient carbon fixation [101].

In addition, cyanobacteria possess a second carbon fixation mechanism through phosphoenolpyruvate carboxylase (PPC) that is responsible for close to 25% of CO_2 fixation [102]. PPC fixes HCO_3^- rather than CO_2 and combines it with phosphoenolpyruvate to form oxaloacetate and inorganic phosphate. An *in silico* modeling study has proposed to couple PPC with the core C4 plant carbon fixation cycle to increase the overall carbon fixation rate compare to the CBB cycle [103].

2.6 Direct conversion of CO_2 into biofuels and chemicals

By capturing atmospheric CO_2 as the carbon source, cyanobacterial biofuel production is an intriguing solution to relieve the excess amount of greenhouse gases released by the fossil fuel combustion. Ethanol is the most common biofuel production in various microorganisms. The first ethanol production in cyanobacteria is accomplished by introducing a pyruvate decarboxylase, which directly decarboxylates pyruvate following by a primary alcohol dehydrogenase [104]. This heterologous pathway enables to direct carbon flux from pyruvate to acetaldehyde and then ethanol. On the other hand, longer chain alcohols such as 1-butanol [105], isobutanol [101] and 2-methyl-1-butanol [106] have been produced from CO_2 using *S. elongatus* 7942. In addition, several other chemicals have been produced through heterologous pathway expression in cyanobacteria [107,108] (Fig. 2-1 and Table 2-1). In order to efficiently utilize cyanobacteria for production of biofuels and chemicals, several important strategies can be employed such as altering cofactor preference [109,110], developing a synthetic driving force [111], managing the toxicity of the desired product [112], and reducing byproduct formation [106]. These strategies improve cyanobacteria for specific and high titer production.

2.6.1 Fuels and chemicals from Acetyl-CoA

There are many compounds that can be produced from acetyl-coenzyme A (CoA), including 1-butanol, acetone, 2-propanol, 3-hydroxypropinate and 3-hydroxybutyrate. Ethanol production has been discussed above and has been attempted in industry. Photosynthetic production of 1-butanol in oxygenic cyanobacteria has been difficult. 1-Butanol is naturally produced by *Clostridia* via the CoA-dependent pathway as part of the acetone-butanol-ethanol (ABE) fermentation. Implementation of this pathway in more user-friendly hosts was successful after

rounds of improvements to build-up both NADH and acetyl-CoA driving forces, enabling production of up to 15g/L 1-butanol anaerobically in *E. coli* [123]. In contrast, engineering a similar pathway in the oxygenic cyanobacterium *S. elongatus* PCC 7942 was even more difficult. Expression of the same modified 1-butanol pathway enzymes in this oxygenic organism failed to produce any 1-butanol due to the oxygen sensitivity of pathway enzymes [124]. The product was detected only under anoxic conditions or with photosystem II inhibited. The problem was solved by using an oxygen-tolerant butyraldehyde dehydrogenase (PduP) [105], and by introducing a malonyl-CoA based synthetic pathway [125] as shown in Fig.2-2. This new approach inspired by fatty acid biosynthesis utilizes ATP as a driving force, thus enabling the organism to compensate for the thermodynamically unfavorable condensation of two acetyl-CoA molecules. The combined strategy enhanced 1-butanol production to 404mg/L [105].

Other fuels and chemicals produced from acetyl-CoA in cyanobacteria include acetone, 2-propanol, 3-hydroxypropinate and 3-hydroxybutyrate. One of the difficulties in 2-propanol production is the step catalyzed by acetate/3-ketoacid CoA transferase. This enzyme needs acetate as a substrate in order to catalyze the reaction toward acetoacetate biosynthesis. However, under autotrophic condition, cyanobacteria do not produce much acetate, and the acetate production pathway remains unclear. Providing external acetate cannot increase 2-propanol titer, potentially due to regulation of acetate uptake. 3-Hydroxypropionate (3HP) is an important compound as it can be converted to several industrially important chemicals such as acrylic acid, acrylamide, and 1,3-propanediol. Furthermore, biodegradable and biocompatible poly-3HP can be made by using 3HP as the monomer. Three pathways have been proposed for 3HP production by microbial platforms (Fig. 2-3). The glycerol dependent pathway [126,127] needs a coenzyme B12 dependent glycerol dehydratase. However, providing external coenzyme B12

raises additional cost. Moreover, B12-independent glycerol dehydratase is an oxygen sensitivity enzyme which is not an ideal candidate for application in oxygenic cyanobacteria [105]. Thus, direct transfer of glycerol dependent pathway into cyanobacteria for 3HP production would become difficult. On the other hand, 3HP can be produced from malonyl-CoA, a precursor for fatty acid synthesis which is common to microorganisms. By introducing different malonyl-CoA reductases and malonate semialdehyde reductases, 3HP productions achieve to titers between 124mg/L to 659mg/L in *S. elongatus* PCC 7942. Another pathway is through β-alanine, which is able to redirect carbon fluxes from phosphoenolpyruvate to 3HP. By coupling the β-alanine dependent pathway with the malonyl-CoA dependent pathway, 3HP production by utilizing both pathways outperforms only the individual pathways. However, this enhancement is potentially limited by 3HP toxicity to cyanobacteria. It is also reported that knocking out acetyl-CoA synthase is able to enhance tolerance to 3HP and other organic acids in *Synechococcus* sp. PCC 7002 [128].

2.6.2 Fuels and chemicals from keto acids

Keto acids can be decarboxylated to aldehydes and then reduced to the corresponding alcohols. The most common example is the conversion of pyruvate to ethanol, which has been demonstrated in multiple organisms, including cyanobacteria. The keto acid intermediates in several amino acid biosynthesis pathways can also be converted to their corresponding branched chain higher alcohols by introducing a keto acid decarboxylase and a primary alcohol dehydrogenase [129]. The first keto acid pathway dependent biofuel production in cyanobacteria is through valine biosynthesis to produce isobutyraldehyde and isobutanol [112] (Fig.2-4). Isobutyraldehyde is an important precursor for industrial applications and isobutanol is a

promising alternative energy as it has higher energy density and lower hygroscopicity than ethanol. In order to enhance carbon flux from pyruvate to valine biosynthesis, an *alsS* gene from *B. subtilis*, *ilvC* and *ilvD* genes from *E. coli*, and *kivd* from *L. lactis* have been overexpressed. In this strain, isobutyraldehyde titer achieved 723 mg/L. Since aldehyde is a toxic compound to cells and Rubisco is considered as one of the limiting steps in photosynthetic organisms, further improvement is accomplished by *in-situ* product removal and overexpression of *rbcLS* from *S. elongatus* PCC6301. The combined strategy successfully enhanced the isobutyraldehyde titer up to 1.1 g/L. Isobutanol can be produced by introducing an alcohol dehydrogenase in the isobutyraldehyde production strains. Among the three alcohol dehydrogenases used in the test such as *yqhD* from *E.coli*, *adh2* from *S. cerevisiae* and *adhA* from *L. lactis*, *yqhD* is a NADPH-utilizing enzyme and indicates the highest isobutanol titer which is 450 mg/L. This result suggests the importance of the cofactor choice in cyanobacterial biofuel production.

2-Methyl-butanol (2MB) is synthesized via the isoleucine pathway (Fig. 2-4), which is similar to the valine biosynthesis pathway. One of the challenges for 2MB production is that it passes over the same cassette of genes as for isobutanol production. A recent study [106] indicated that by enhancing the citramalate pathway to direct the carbon flux from pyruvate to α-ketobutyrate, it led to 2MB production to about 200mg/L with minimum side-products such as 1-propanol and isobutanol. Kinetic analysis of native AHAS from *S. elongatus* PCC7942 demonstrated a preferred selectivity toward α-ketobutyrate.

Cyanobacteria are known to synthesize glycogen as a storage compound. Knocking out glucose-1-phosphate adenylyltransferase (*glgC*), the initial step for glycogen synthesis, has a negative effect on growth and glycogen synthesis under high light conditions ($150 \mu Es^{-1}m^{-2}$). This effect

is attributed to the lack of an electron sink. However, the isobutanol pathway can serve as a metabolic sink for replacing glycogen synthesis to rescue growth in the Δ*glgC* mutant [130]. This phenotype may be exploited as a selection platform for enhancing product formation.

2.6.3 Chemicals from TCA cycle intermediates

The TCA cycle oxidizes acetyl-CoA derived from glycogen degradation in the dark to produce NADH and $FADH_2$ for ATP generation in the respiratory chain. Moreover, TCA cycle intermediates are important precursors for amino acid biosynthesis. In cyanobacteria, it had been generally considered that the TCA cycle was not complete, lacking the 2-oxoglutarate dehydrogenase complex (2-OGDH), until a recent study showed that their unique TCA cycle is closed by 2-oxoglutarate decarboxylase (2-OGDC) and succinic semialdehyde dehydrogenase (SSADH) [131].

It is generally accepted that cyanobacterial TCA cycle is not very active under phototrophic conditions. A recent study further proved that it operates in a bifurcated way [116]. However, the bifurcated pattern switches to a dominantly cyclic pattern by introducing an ethylene forming enzyme from *Pseudomonas*. Ethylene can be sustainably and efficiently produced from 2-oxoglutarate while also activating the TCA cycle metabolism. The enhanced flux via the remodeled TCA is 37% of the total fixed carbon which is higher than the 13% found in wild type. This result indicates that the metabolic malleability of cyanobacterial TCA cycle depends on what it produces.

2.6.4 Hydrogen

In recent years, hydrogen produced by living organisms has gained increased attention. Hydrogen is particularly attractive because its utilization is carbon-free, generating only water

and no greenhouse gases. A number of photosynthetic organisms naturally produce small quantities of hydrogen. In cyanobacteria several enzymes are involved in hydrogen metabolism: nitrogenase which produces hydrogen as a side product during the reduction of nitrogen to ammonia, a hydrogenase that can catalyze the reversible reaction $2H^+ + 2e^- \rightarrow H_2$, and an uptake hydrogenase with the function of recapturing the hydrogen produced by nitrogenase [132]. Several species of cyanobacteria able to make H_2 have been described and are reviewed elsewhere [133]. Studies have been focusing on selecting strains with specific hydrogen metabolism, improving conditions in bioreactors cultivation and increasing H_2 production by genetic engineering [134]. This last point focuses on reinforcing the electron flow towards hydrogenases. In one study, hydrogen production in *S. elongatus* sp. 7942 was approximately doubled by strengthening the electron flow through overexpression of *Clostridium acetobutylicum* ferredoxin [135]. Effort has also been made to eliminate competitive pathways in term of reducing agents consumption. In *Synechococcus* 7002, inactivation of the gene coding for lactate dehydrogenase, the main fermentative reductant sink, resulted in a fivefold higher H_2 production than wild type under dark anoxic conditions [136]. Another major problem to overcome for the development of more efficient biohydrogen production is the high oxygen sensitivity of hydrogenases [137]. In current processes, H_2 production is temporally separated from O_2 generating photosynthesis by first growing the cells normally and then transferring them in anaerobic conditions coupled with PSII inactivation [135]. Hydrogen is then produced through fermentation using internal carbohydrates. Engineering of current hydrogenases or discovery of new ones with higher O_2 tolerance remains one of the biggest challenges for a sustained continuous hydrogen production in cyanobacteria.

2.7 Conclusions

Cyanobacteria possess several advantageous features, making them promising host organisms for various sustainable biotechnological applications. They are quite easy to genetically maniputate, have only simple growth requirements and can fix carbon dioxide to produce industrially relevant molecules using sunlight as energy. Moreover, several downstream products such as sucrose [138], 2,3-butanediol [115] and ethylene [116] have been shown to act as carbon sinks able to pull the upstream carbon flux toward product formation, enhancing the photosynthetic efficiency. These results highlight the potential of cyanobacterial platforms to make progress towards carbon-neutral biofuel and chemical production.

However, although it is promising, there are still a number of challenges that need to overcome to see cyanobacteria become an industrial microorganism. The productivity of metabolites of interest is still low and needs to be increased through both metabolic engineering to create improved strains and development of more efficient bioreactors. With advances in both fields, cyanobacteria may become a valuable addition to the current industrial microbial platforms.

2.8 Tables and Figures

Compound	Organism	Titers	Reference
acetone	*Synechocystis* sp. PCC6803	36mg/L	Zhou et al.(2012)[113]
2-propanol	*S.elongatus* sp. PCC7942	26.5mg/L	Kusakabe et al.(2013)[114]
2,3-butanediol	*S.elongatus* sp. PCC7942	2.4g/L	Oliver et al.(2013)[115]
1-butanol	*S.elongatus* sp. PCC7942	404mg/L	Lan et al.(2013)[105]
ethylene	*Synechocystis* sp. PCC6803	718±19μl $l^{-1}h^{-1}$ perA_{730nm}	Xiong et al.(2015)[116]
1,2-propanediol	*S.elongatus* sp. PCC7942	150mg/L	Li and Liao(2013)[117]
3-hydroxybutyrate	*Synechocystis* sp. PCC6803	533.4 mg/L	Wang et al.(2013)[118]
3-hydroxypropionate	*S.elongatus* sp. PCC7942	665 mg/L	Lan et al.(2015)[119]
	Synechocystis sp. PCC6803	837 mg/L	Wang et al.(2015)[120]
isobutanol	*S.elongatus* sp. PCC7942	450mg/L	Atsumi et al.(2009)[112]
isobutyraldehyde	*S.elongatus* sp. PCC7942	1.1 g/L	Atsumi et al.(2009)[112]
2-methyl-1-butanol	*S.elongatus* sp. PCC7942	200mg/L	Shen and Liao (2012)[106]
fatty acid	*Synechocystis* sp. PCC6803	197 mg/L	Liu et al.(2011)[121]
ethanol	*Synechocystis* sp. PCC6803	5.5 g/L*	Gao et al.(2012)[122]

*****by photobioreactor

Table 2-1.Titers for biofuel and chemical productions from cyanobacteria

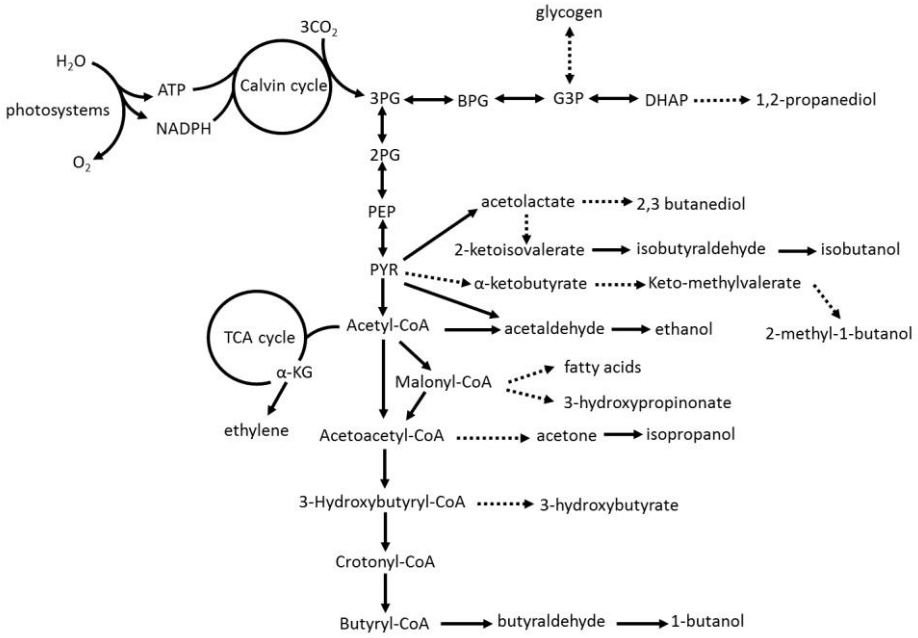


Figure 2-1. Overview of pathway for biofuels and chemicals produced from cyanobacteria

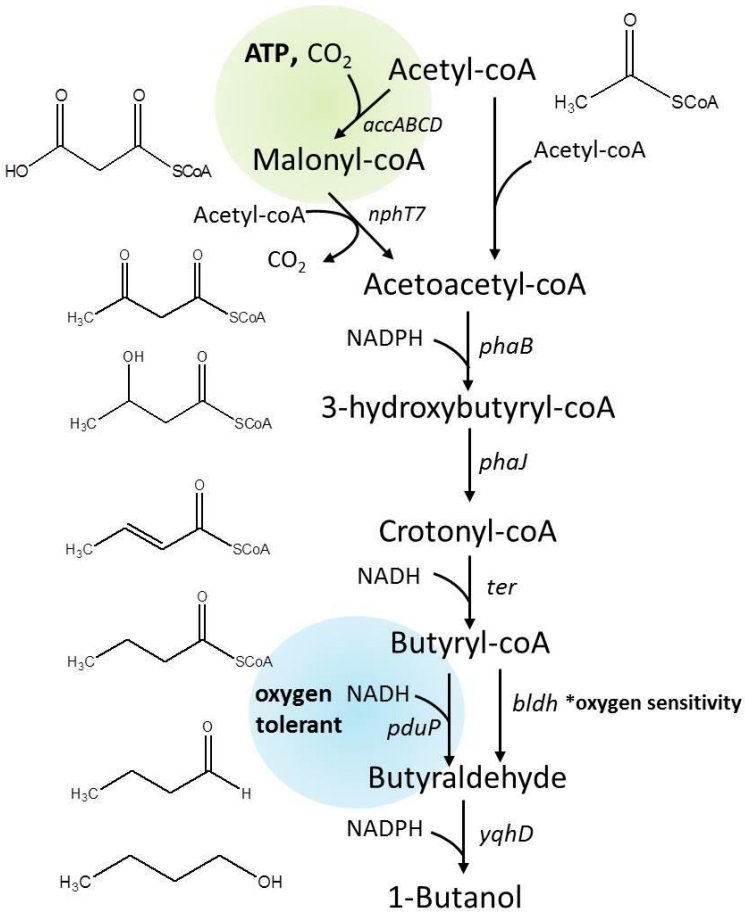


Figure 2-2. 1-Butanol production via the CoA-dependent pathway implemented in *S. elongatus* PCC7942. Gene symbols are: *accABCD*, acetyl-coA carboxylase; *nphT7*, acetoacetyl-coA synthase; *phaB*, acetoacetyl-coA reductase; *phaJ*, R-specific enoyl-coA hydratase; *ter*, trans enoyl-coA reductase; *pduP*, CoA-acylating propionaldehyde dehydrogenase; *bldh*, butyraldehyde dehydrogenase; *yqhD*, NADPH-dependent alcohol dehydrogenase.

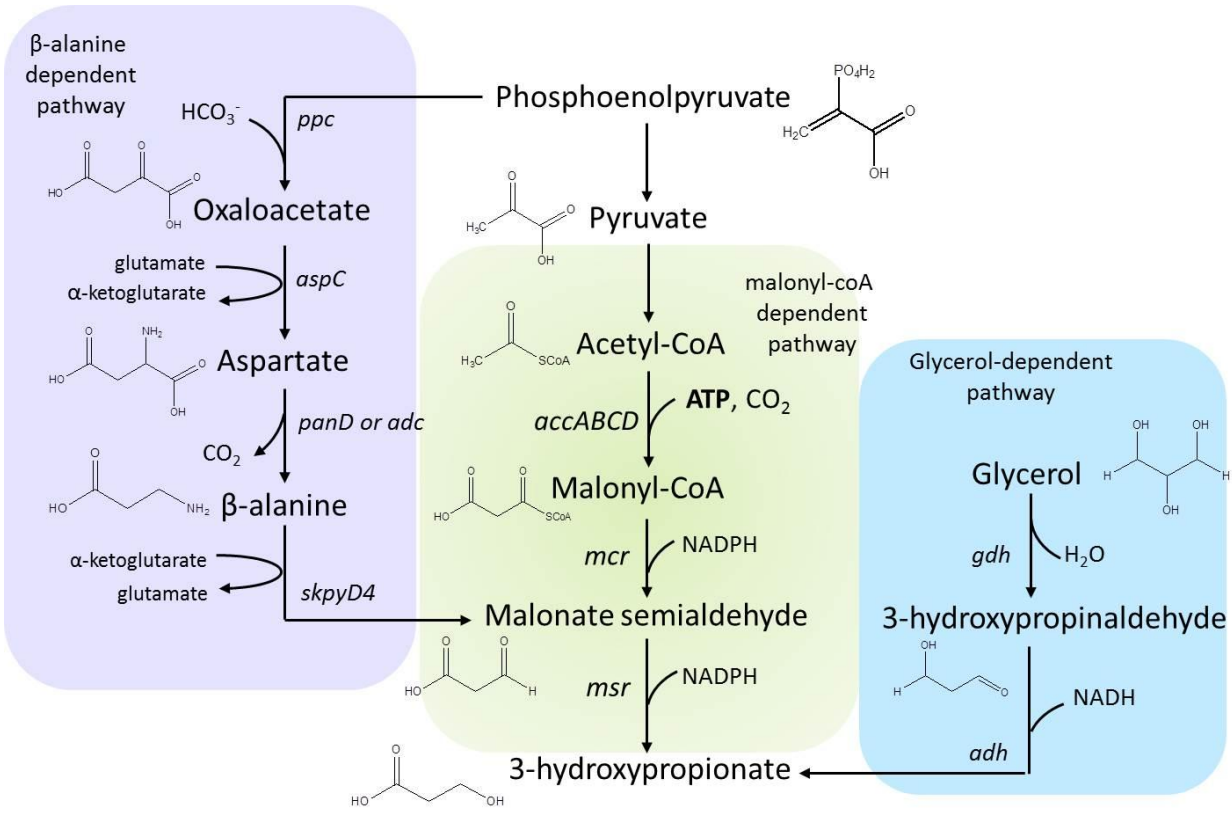


Figure 2-3. 3-Hydroxypropionate (3HP) synthesis via glycerol-dependent, malonyl-CoA-dependent, and β-alanine-dependent pathways. Gene symbols are: *ppc*, phosphoenolpyruvate carboxylase; *aspC*, aspartate transaminase; *panD* or *adc*: aspartate decarboxylase; *skpyD4*, β-alanine aminotransferase; *mcr*, malonyl-coA reductase; *msr*, malonate semialdehyde reductase; *gdh*, glycerol dehydratase; *adh*, alcohol dehydrogenase.

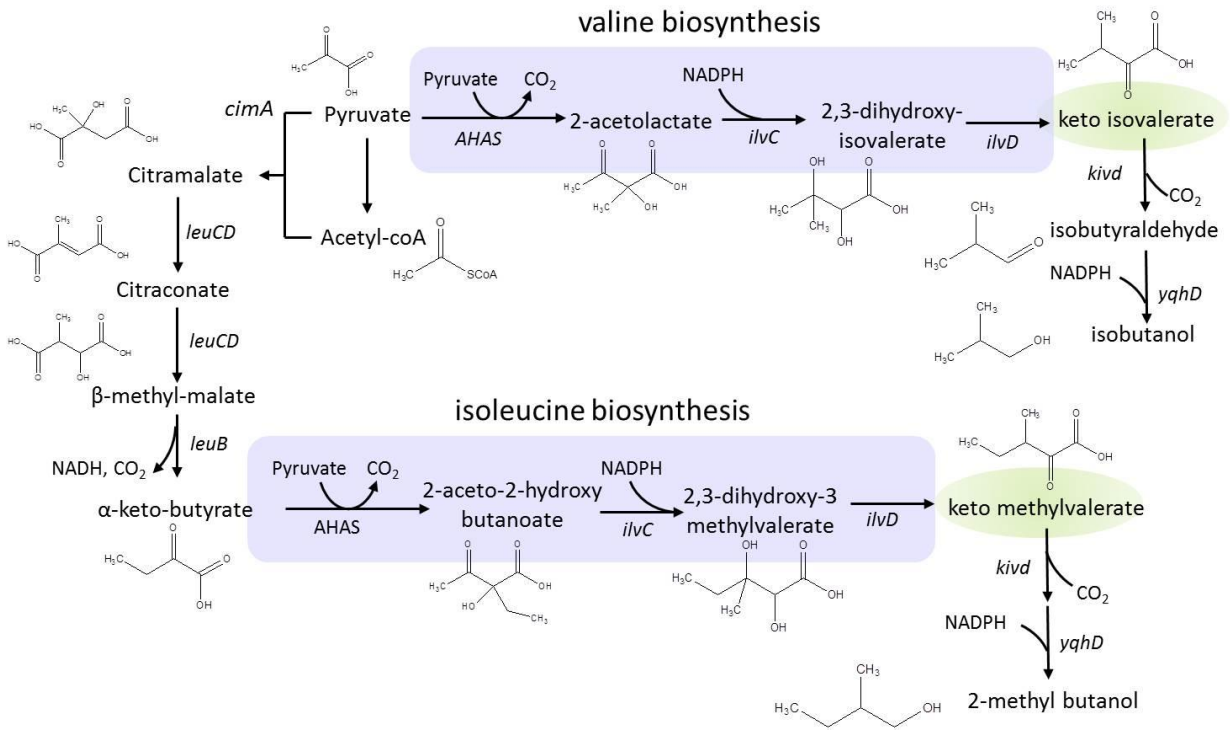


Figure 2-4: Isobutanol and 2-methyl butanol (2MB) production via keto acid pathway in cyanobacteria. Gene/protein symbols are: *cimA*, citramalate synthase; *leuCD*, isopropylmalate isomerase; *leuB*, 3-isopropylmalate dehydrogenase; AHAS, acetohydroxyacid synthase; *ilvC*, acetohydroxy acid isomeroreductase; *ilvD*, dihydroxy acid dehydratase; *kivd*, ketoisovalerate decarboxylase; *yqhD*, alcohol dehydrogenase

2.9 References

1. Melis, A. Solar energy conversion efficiencies in photosynthesis: minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.* **177,** 272–280 (2009).

2. Zhu, X.-G., Long, S. P. & Ort, D. R. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr. Opin. Biotechnol.* **19,** 153–159 (2008).

3. Rastogi, R. P. & Sinha, R. P. Biotechnological and industrial significance of cyanobacterial secondary metabolites. *Biotechnol. Adv.* **27,** 521–539 (2009).

4. Stal, L. J. The metabolic versatility of the mat-building cyanobacteria Microcoleus chthonoplastes and Oscillatoria limosa and its ecological significance. *Algol. Stud. Für Hydrobiol. Suppl. Vol.* **64,** 453–467 (1991).

5. Herrero, A., Muro-Pastor, A. M. & Flores, E. Nitrogen control in cyanobacteria. *J. Bacteriol.* **183,** 411–425 (2001).

6. Castenholz, R. W. Ecology of blue-green algae in hot springs. *Bot. Monogr.* (1973).

7. Cohen, Y., Jørgensen, B. B., Revsbech, N. P. & Poplawski, R. Adaptation to hydrogen sulfide of oxygenic and anoxygenic photosynthesis among cyanobacteria. *Appl. Environ. Microbiol.* **51,** 398–407 (1986).

8. Peschek, G. A., Löffelhardt, W. & Schmetterer, G. *The phototrophic prokaryotes*. (Springer Science & Business Media, 2012).

9. Fay, P. Heterotrophy and nitrogen fixation in Chlorogloea fritschii. *J. Gen. Microbiol.* **39,** 11–20 (1965).

10. Rippka, R. Photoheterotrophy and chemoheterotrophy among unicellular blue-green algae. *Arch. Für Mikrobiol.* **87,** 93–98 (1972).

11. Kromkamp, J. Formation and functional significance of storage products in cyanobacteria. *N. Z. J. Mar. Freshw. Res.* **21,** 457–465 (1987).

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

13. Cyanobacterial nitrogen fixation (CYANOFIX). (1999).

14. Maldener, I. & Muro-Pastor, A. M. Cyanobacterial heterocysts. *eLS* (2010).

15. Kumar, K., Mella-Herrera, R. A. & Golden, J. W. Cyanobacterial heterocysts. *Cold Spring Harb. Perspect. Biol.* **2,** a000315 (2010).

16. Bergman, B., Gallon, J. R., Rai, A. N. & Stal, L. J. N2 fixation by non-heterocystous cyanobacteria. *FEMS Microbiol. Rev.* **19,** 139–185 (1997).

17. Gallon, J. R., LaRue, T. A. & Kurz, W. G. W. Photosynthesis and nitrogenase activity in the blue-green alga Gloeocapsa. *Can. J. Microbiol.* **20,** 1633–1637 (1974).

18. Mitsui, A. *et al.* Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature* **323,** 720–722 (1986).

19. Misra, H. S. & Tuli, R. Differential expression of photosynthesis and nitrogen fixation genes in the cyanobacterium Plectonema boryanum. *Plant Physiol.* **122,** 731–736 (2000).

20. Compaoré, J. & Stal, L. J. Oxygen and the light–dark cycle of nitrogenase activity in two unicellular cyanobacteria. *Environ. Microbiol.* **12,** 54–62 (2010).

21. Bandyopadhyay, A., Elvitigala, T., Liberton, M. & Pakrasi, H. B. Variations in the rhythms of respiration and nitrogen fixation in members of the unicellular diazotrophic cyanobacterial genus Cyanothece. *Plant Physiol.* **161,** 1334–1346 (2013).

22. Colon-Lopez, M. S., Sherman, D. M. & Sherman, L. A. Transcriptional and translational regulation of nitrogenase in light-dark-and continuous-light-grown cultures of the unicellular cyanobacterium Cyanothece sp. strain ATCC 51142. *J. Bacteriol.* **179,** 4319–4327 (1997).

23. Schneegurt, M. A., Sherman, D. M., Nayar, S. & Sherman, L. A. Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium Cyanothece sp. strain ATCC 51142. *J. Bacteriol.* **176,** 1586–1597 (1994).

24. Li, H., Sherman, D. M., Bao, S. & Sherman, L. A. Pattern of cyanophycin accumulation in nitrogen-fixing and non-nitrogen-fixing cyanobacteria. *Arch. Microbiol.* **176,** 9–18 (2001).

25. Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B. & Carpenter, E. J. Trichodesmium, a globally significant marine cyanobacterium. *Science* **276,** 1221–1229 (1997).

26. Bergman, B., Sandh, G., Lin, S., Larsson, J. & Carpenter, E. J. Trichodesmium–a widespread marine cyanobacterium with unusual nitrogen fixation properties. *FEMS Microbiol. Rev.* **37,** 286–302 (2013).

27. Finzi-Hart, J. A. *et al.* Fixation and fate of C and N in the cyanobacterium Trichodesmium using nanometer-scale secondary ion mass spectrometry. *Proc. Natl. Acad. Sci.* **106,** 6345–6350 (2009).

28. Edmunds, L. N. Chronobiology at the cellular and molecular levels: models and mechanisms for circadian timekeeping. *Am. J. Anat.* **168,** 389–431 (1983).

29. Kippert, F. Endocytobiotic coordination, intracellular calcium signaling, and the origin of endogenous rhythms. *Ann. N. Y. Acad. Sci.* **503,** 476–495 (1987).

30. Grobbelaar, N., Huang, T. C., Lin, H. Y. & Chow, T. J. Dinitrogen-fixing endogenous rhythm in Synechococcus RF-1. *FEMS Microbiol. Lett.* **37,** 173–177 (1986).

31. Huang, T.-C., Tu, J., Chow, T.-J. & Chen, T.-H. Circadian rhythm of the prokaryote Synechococcus sp. RF-1. *Plant Physiol.* **92,** 531–533 (1990).

32. Chen, T.-H., Chen, T.-L., Hung, L.-M. & Huang, T.-C. Circadian rhythm in amino acid uptake by Synechococcus RF-1. *Plant Physiol.* **97,** 55–59 (1991).

33. Dong, G., Kim, Y.-I. & Golden, S. S. Simplicity and complexity in the cyanobacterial circadian clock mechanism. *Curr. Opin. Genet. Dev.* **20,** 619–625 (2010).

34. Ishiura, M. *et al.* Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. *Science* **281,** 1519–1523 (1998).

35. Ito, H. *et al.* Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus. *Proc. Natl. Acad. Sci.* **106,** 14168–14173 (2009).

36. Dong, G. *et al.* Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in Synechococcus elongatus. *Cell* **140,** 529–539 (2010).

37. Mori, T., Binder, B. & Johnson, C. H. Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. *Proc. Natl. Acad. Sci.* **93,** 10183–10188 (1996).

38. Smith, R. M. & Williams, S. B. Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium Synechococcus elongatus. *Proc. Natl. Acad. Sci.* **103,** 8564–8569 (2006).

39. Woelfle, M. A., Xu, Y., Qin, X. & Johnson, C. H. Circadian rhythms of superhelical status of DNA in cyanobacteria. *Proc. Natl. Acad. Sci.* **104,** 18819–18824 (2007).

40. Cohen, S. E. & Golden, S. S. Circadian Rhythms in Cyanobacteria. *Microbiol. Mol. Biol. Rev.* **79,** 373–385 (2015).

41. Vermaas, W. F. in *eLS* (John Wiley & Sons, Ltd, 2001).

42. Croce, R. & van Amerongen, H. Natural strategies for photosynthetic light harvesting. *Nat. Chem. Biol.* **10,** 492–501 (2014).

43. Green, B. & Parson, W. W. *Light-harvesting antennas in photosynthesis*. **13,** (Springer Science & Business Media, 2003).

44. Gutu, A. & Kehoe, D. M. Emerging perspectives on the mechanisms, regulation, and distribution of light color acclimation in cyanobacteria. *Mol. Plant* **5,** 1–13 (2012).

45. Niyogi, K. K. & Truong, T. B. Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.* **16,** 307–314 (2013).

46. Ansong, C. *et al.* Characterization of protein redox dynamics induced during light-to-dark transitions and nutrient limitation in cyanobacteria. *Front. Microbiol.* **5,** (2014).

47. Buchanan, B. B. Regulation of CO 2 assimilation in oxygenic photosynthesis: the ferredoxin/thioredoxin system: perspective on its discovery, present status, and future development. *Arch. Biochem. Biophys.* **288,** 1–9 (1991).

48. Florencio, F. J., Pérez-Pérez, M. E., López-Maury, L., Mata-Cabana, A. & Lindahl, M. The diversity and complexity of the cyanobacterial thioredoxin systems. *Photosynth. Res.* **89,** 157–171 (2006).

49. Lindahl, M. & Kieselbach, T. Disulphide proteomes and interactions with thioredoxin on the track towards understanding redox regulation in chloroplasts and cyanobacteria. *J. Proteomics* **72,** 416–438 (2009).

50. Heidorn, T. *et al.* Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods Enzymol.* **497,** 539–579 (2011).

51. Kufryk, G. I., Sachet, M., Schmetterer, G. & Vermaas, W. F. J. Transformation of the cyanobacterium Synechocystis sp. PCC 6803 as a tool for genetic mapping: optimization of efficiency. *FEMS Microbiol. Lett.* **206,** 215–219 (2002).

52. Grigorieva, G. & Shestakov, S. Transformation in the cyanobacterium Synechocystis sp. 6803. *FEMS Microbiol Lett* **13,** 367–370 (1982).

53. Shestakov, S. V. & Khyen, N. T. Evidence for genetic transformation in blue-green alga Anacystis nidulans. *Mol. Gen. Genet. MGG* **107,** 372–375 (1970).

54. Stevens, S. E. & Porter, R. D. Transformation in Agmenellum quadruplicatum. *Proc. Natl. Acad. Sci.* **77,** 6052–6056 (1980).

55. Onai, K., Morishita, M., Kaneko, T., Tabata, S. & Ishiura, M. Natural transformation of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1: a simple and efficient method for gene transfer. *Mol. Genet. Genomics* **271,** 50–59 (2004).

56. Iwai, M., Katoh, H., Katayama, M. & Ikeuchi, M. Improved Genetic Transformation of the Thermophilic Cyanobacterium, Thermosynechococcus elongatus BP-1. *Plant Cell Physiol.* **45,** 171–175 (2004).

57. Wolk, C. P., Vonshak, A., Kehoe, P. & Elhai, J. Construction of shuttle vectors capable of conjugative transfer from Escherichia coli to nitrogen-fixing filamentous cyanobacteria. *Proc. Natl. Acad. Sci.* **81,** 1561–1565 (1984).

58. Cohen, M. F., Wallis, J. G., Campbell, E. L. & Meeks, J. C. Transposon mutagenesis of Nostoc sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* **140,** 3233–3240 (1994).

59. Flores, E. & Wolk, C. P. Identification of facultatively heterotrophic, N2-fixing cyanobacteria able to receive plasmid vectors from Escherichia coli by conjugation. *J. Bacteriol.* **162,** 1339–1341 (1985).

60. Murry, M. A. & Wolk, C. P. Identification and initial utilization of a portion of the smaller plasmid ofAnabaena variabilis ATCC 29413 capable of replication inAnabaena sp. strain M-131. *Mol. Gen. Genet. MGG* **227,** 113–119 (1991).

61. Tsinoremas, N. F., Kutach, A. K., Strayer, C. A. & Golden, S. S. Efficient gene transfer in Synechococcus sp. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. *J. Bacteriol.* **176,** 6764–6768 (1994).

62. Marraccini, P., Bulteau, S., Cassier-Chauvat, C., Mermet-Bouvier, P. & Chauvat, F. A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera Synechococcus and Synechocystis. *Plant Mol. Biol.* **23,** 905–909 (1993).

63. Vioque, A. in *Transgenic microalgae as green cell factories* 12–22 (Springer, 2007).

64. Herdman, M., Janvier, M., Rippka, R. & Stanier, R. Y. Genome Size of Cyanobacteria. *Microbiology* **111,** 73–85 (1979).

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

66. Vermaas, W. Molecular genetics of the cyanobacteriumSynechocystis sp. PCC 6803: Principles and possible biotechnology applications. *J. Appl. Phycol.* **8,** 263–273 (1996).

67. Zhou, J. *et al.* Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. *Sci. Rep.* **4,** (2014).

68. Camsund, D. & Lindblad, P. Engineered transcriptional systems for cyanobacterial biotechnology. *Front. Bioeng. Biotechnol.* **2,** (2014).

69. Guerrero, F., Carbonell, V., Cossu, M., Correddu, D. & Jones, P. R. Ethylene synthesis and regulated expression of recombinant protein in Synechocystis sp. PCC 6803. (2012).

70. Huang, H.-H. & Lindblad, P. Wide-dynamic-range promoters engineered for cyanobacteria. *J Biol Eng* **7,** (2013).

71. Imamura, S. & Asayama, M. Sigma factors for cyanobacterial transcription. *Gene Regul. Syst. Biol.* **3,** 65 (2009).

72. Peters, J. M., Vangeloff, A. D. & Landick, R. Bacterial Transcription Terminators: The RNA 3′-End Chronicles. *J. Mol. Biol.* **412,** 793–813 (2011).

73. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10,** 354–360 (2013).

74. Cambray, G. *et al.* Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res.* **41,** 5139–5148 (2013).

75. Sharma, N. K., Rai, A. K. & Stal, L. J. *Cyanobacteria: an economic perspective*. (John Wiley & Sons, 2013).

76. De Hoon, M. J., Makita, Y., Nakai, K. & Miyano, S. Prediction of transcriptional terminators in Bacillus subtilis and related species. *PLoS Comput Biol* **1,** e25 (2005).

77. Vijayan, V., Jain, I. H. & O'Shea, E. K. A high resolution map of a cyanobacterial transcriptome. *Genome Biol.* **12,** R47 (2011).

78. Ma, J., Campbell, A. & Karlin, S. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J. Bacteriol.* **184,** 5733–5745 (2002).

79. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27,** 946–950 (2009).

80. Na, D. & Lee, D. RBSDesigner: software for designing synthetic ribosome binding sites that yields a desired level of protein expression. *Bioinformatics* **26,** 2633–2634 (2010).

81. Seo, S. W. *et al.* Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metab. Eng.* **15,** 67–74 (2013).

82. Ting, C. S., Rocap, G., King, J. & Chisholm, S. W. Cyanobacterial photosynthesis in the oceans: the origins and significance of divergent light-harvesting strategies. *Trends Microbiol.* **10,** 134–142 (2002).

83. Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *science* **332,** 805–809 (2011).

84. Kirst, H., Formighieri, C. & Melis, A. Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. *Biochim. Biophys. Acta BBA-Bioenerg.* **1837,** 1653–1664 (2014).

85. Kirst, H., García-Cerdán, J. G., Zurbriggen, A. & Melis, A. Assembly of the light-harvesting chlorophyll antenna in the green alga Chlamydomonas reinhardtii requires expression of the TLA2-CpFTSY gene. *Plant Physiol.* **158,** 930–945 (2012).

86. Kirst, H., Garcia-Cerdan, J. G., Zurbriggen, A., Ruehle, T. & Melis, A. Truncated photosystem chlorophyll antenna size in the green microalga Chlamydomonas reinhardtii upon deletion of the TLA3-CpSRP43 gene. *Plant Physiol.* **160,** 2251–2260 (2012).

87. Polle, J. E., Kanakagiri, S.-D. & Melis, A. tla1, a DNA insertional transformant of the green alga Chlamydomonas reinhardtii with a truncated light-harvesting chlorophyll antenna size. *Planta* **217,** 49–59 (2003).

88. Melis, A. Photosynthesis-to-fuels: from sunlight to hydrogen, isoprene, and botryococcene production. *Energy Environ. Sci.* **5,** 5531–5539 (2012).

89. Kirst, H. & Melis, A. The chloroplast signal recognition particle (CpSRP) pathway as a tool to minimize chlorophyll antenna size and maximize photosynthetic productivity. *Biotechnol. Adv.* **32,** 66–72 (2014).

90. Mussgnug, J. H. *et al.* Engineering photosynthetic light capture: impacts on improved solar energy to biomass conversion. *Plant Biotechnol. J.* **5,** 802–814 (2007).

91. Page, L. E., Liberton, M. & Pakrasi, H. B. Reduction of photoautotrophic productivity in the cyanobacterium Synechocystis sp. strain PCC 6803 by phycobilisome antenna truncation. *Appl. Environ. Microbiol.* **78,** 6349–6351 (2012).

92. Lea-Smith, D. J. *et al.* Phycobilisome-deficient strains of Synechocystis sp. PCC 6803 have reduced size and require carbon-limiting conditions to exhibit enhanced productivity. *Plant Physiol.* **165,** 705–714 (2014).

93. Ellis, R. J. Biochemistry: Tackling unintelligent design. *Nature* **463,** 164–165 (2010).

94. Tcherkez, G. G., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci.* **103,** 7246–7251 (2006).

95. Ninomiya, N., Ashida, H. & Yokota, A. in *Photosynthesis. Energy from the Sun* 867–870 (Springer, 2008).

96. Price, G. D., Badger, M. R., Woodger, F. J. & Long, B. M. Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot.* **59,** 1441–1461 (2008).

97. Price, G. D. Inorganic carbon transporters of the cyanobacterial CO2 concentrating mechanism. *Photosynth. Res.* **109,** 47–57 (2011).

98. Espie, G. S. & Kimber, M. S. Carboxysomes: cyanobacterial RubisCO comes in small packages. *Photosynth. Res.* **109,** 7–20 (2011).

99. Savage, D. F., Afonso, B., Chen, A. H. & Silver, P. A. Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science* **327,** 1258–1261 (2010).

100. Iwaki, T. *et al.* 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 (2006).

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

102. Yang, C., Hua, Q. & Shimizu, K. Metabolic flux analysis in Synechocystis using isotope distribution from 13 C-labeled glucose. *Metab. Eng.* **4,** 202–216 (2002).

103. Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. Design and analysis of synthetic carbon fixation pathways. *Proc. Natl. Acad. Sci.* **107,** 8889–8894 (2010).

104. Deng, M. D. & Coleman, J. R. Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ. Microbiol.* **65,** 523–528 (1999).

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

106. Shen, C. R. & Liao, J. C. 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 (2012).

107. Nozzi, N. E., Oliver, J. W. K. & Atsumi, S. Cyanobacteria as a Platform for Biofuel Production. *Front. Bioeng. Biotechnol.* **1,** (2013).

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

109. Varman, A. M., Yu, Y., You, L. & Tang, Y. J. Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microb. Cell Factories* **12,** 117 (2013).

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

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

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

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

114. Kusakabe, T. *et al.* Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metab. Eng.* **20,** 101–108 (2013).

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

116. Xiong, W. *et al.* The plasticity of cyanobacterial metabolism supports direct CO2 conversion to ethylene. *Nat. Plants* **1,** 15053 (2015).

117. Li, H. & Liao, J. C. Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO2 to 1,2-propanediol. *Microb. Cell Factories* **12,** 4 (2013).

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

119. Lan, E. I. *et al.* Metabolic engineering of cyanobacteria for photosynthetic 3-hydroxypropionic acid production from CO2 using Synechococcus elongatus PCC 7942. *Metab. Eng.* **31,** 163–170 (2015).

120. Wang, Y. *et al.* Biosynthesis of platform chemical 3-hydroxypropionic acid (3-HP) directly from CO2 in cyanobacterium Synechocystis sp. PCC 6803. *Metab. Eng.* (2015). doi:10.1016/j.ymben.2015.10.008

121. Liu, X., Sheng, J. & Iii, R. C. Fatty acid production in genetically modified cyanobacteria. *Proc. Natl. Acad. Sci.* **108,** 6899–6904 (2011).

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

123. Shen, C. R. *et al.* Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Appl. Environ. Microbiol.* **77,** 2905–2915 (2011).

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

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

126. Tokuyama, K. *et al.* Increased 3-hydroxypropionic acid production from glycerol, by modification of central metabolism in Escherichia coli. *Microb. Cell Factories* **13,** 64 (2014).

127. Chu, H. S. *et al.* Metabolic engineering of 3-hydroxypropionic acid biosynthesis in Escherichia coli. *Biotechnol. Bioeng.* **112,** 356–364 (2015).

128. Begemann, M. B. *et al.* An Organic Acid Based Counter Selection System for Cyanobacteria. *PLoS ONE* **8,** e76594 (2013).

129. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451,** 86–89 (2008).

130. Li, X., Shen, C. R. & Liao, J. C. 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 (2014).

131. Zhang, S. & Bryant, D. A. The tricarboxylic acid cycle in cyanobacteria. *Science* **334,** 1551–1553 (2011).

132. Tamagnini, P. *et al.* Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiol. Mol. Biol. Rev.* **66,** 1–20 (2002).

133. Dutta, D., De, D., Chaudhuri, S. & Bhattacharya, S. K. Hydrogen production by Cyanobacteria. *Microb. Cell Factories* **4,** 36 (2005).

134. Lindblad, P. Cyanobacterial H2 metabolism: knowledge and potential/strategies for a photobiotechnological production of H2. *Biotecnol. Apl.* **16,** 141–144 (1999).

135. Ducat, D. C., Sachdeva, G. & Silver, P. A. Rewiring hydrogenase-dependent redox circuits in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 3941–3946 (2011).

136. McNeely, K., Xu, Y., Bennette, N., Bryant, D. A. & Dismukes, G. C. Redirecting Reductant Flux into Hydrogen Production via Metabolic Engineering of Fermentative Carbon Metabolism in a Cyanobacterium. *Appl. Environ. Microbiol.* **76,** 5032–5038 (2010).

137. Stripp, S. T. *et al.* How oxygen attacks [FeFe] hydrogenases from photosynthetic organisms. *Proc. Natl. Acad. Sci.* **106,** 17331–17336 (2009).

138. Ducat, D. C., Avelar-Rivas, J. A., Way, J. C. & Silver, P. A. Rerouting Carbon Flux To Enhance Photosynthetic Productivity. *Appl. Environ. Microbiol.* **78,** 2660–2668 (2012).

139. Stal, L. J. Physiological ecology of cyanobacteria in microbial mats and other communities. *New Phytol.* **131,** 1–32 (1995).

3. Construction of Calvin cycle mutants suitable to assess new carbon fixation pathways

3.1 Introduction

Photosynthetic organisms have evolved over millions of years to be optimized for growth and carbon capture through the Calvin Benson Bassham (CBB) cycle. All photosynthetic organisms fix atmospheric carbon dioxide through the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco)[1]. Rubisco represents the first major step in carbon fixation, but is often pointed out as inefficient because of its slow catalytic rate and poor specificity, reacting with both CO_2 and O_2 molecules. Natural variations in Rubisco kinetic properties exist among diverse spieces[2]. The different forms of Rubisco found in nature can be divided in four different groups based on their differences in primary sequence (Table 3-1), although form IV, also referred as the Rubisco-like protein (RLP), does not catalyse the distinctive carboxylation and oxygenation reactions[3]. RLPs have been shown to play a role in sulfur metabolism, some of them functioning as tautomerases/enolases in a methionine salvage pathway[4,5]. The four different holoenzyme forms of Rubisco usually show unique structures, but they all contain at least one large subunit (L) dimer which represents the catalytic unit containing two active sites. Form I, the most abundant, is found in eukaryotes, like plants and green algae, and bacteria. Four dimmers of L form the catalytic core of form I with eight small subunits on the top and bottom of the L8 octomeric core[6]. Form I is the only form that contains small subunits. High structural similarities have been shown throughout the many form I Rubisco from different organisms[7], but a further division exists based on the sequence of their large subunit[8]. Form I Rubisco is categorized into two main subgroups, a “green-like” and a “red-like” subgroup. Additionally, the “green-like” group is further subdivided into subclasse IA, regrouping various proteobacterial

and marine cyanobacterial large subunits, and subgroup IB, containing plant, green algae and most of cyanobacterial large subunits. The "red-like" subgroup is comprised of subclasse IC, which comprises various proteobacterial large subunits, and subclasse ID, which includes large subunits from chromophytic and rhodophytic algae. The remaining three other types of Rubisco are composed solely of large subunits. Form II is formed by two to eight dimers of L depending of its origin. The first Rubisco structure to ever be crystallized, from the bacterium *Rhodospirillum rubrum,* belongs to the type II and is only a simple dimer[9,10]. Form III is present only in some archaea[11] mostly anaerobic extremophiles and consists of homodimers or pentamers of dimers[12]. Form III Rubiscos usually have a strong affinity for O_2[13], and thus are highly oxygen sensitive[14].

Form III Rubiscos are believed to fulfill a catabolic step in the 5- phospho-D-ribose-1-pyrophosphate (RuPP) pathway, metabolizing RuBP produced from RuPP[15]. Form IV (RLP) appears to have an L2 structure[16,17]. Despite diverse primary structures, Rubisco form I and II, and maybe form III, have a conserved catalytic process. They do nevertheless differ in their kinetic properties, such as their CO_2/O_2 specificity substrate affinity and catalytic efficiency[8], indicating some evolution related to the environment of the organism they belong to. Rubisco is in fact thought to have been optimized depending of its surroundings for the best compromise between substrate specificity and maximum rate of catalytic turnover[18,19]. Therefore, Rubiscos with higher carboxylating rate but low CO_2 affinity are found in organisms that thrive in high carbon dioxide atmosphere, or that possess CO_2-concentrating mechanisms (CCM) that rise CO_2 concentration in the vicinity of Rubisco. In the past, the natural catalytic diversity of Rubisco has been taking advantage of to construct a cyanobacteria mutant in which the native Rubisco (Form I) was replaced by the *Rhodospirillum rubrum rbcL* gene[20,21] which codes for a a simple

homodimer Form II Rubisco with an inferior CO_2 specificity[22,23]. The mutant characterized by the inability to grown in air and a general extreme sensitivity to the CO_2/O_2 ratio also lacked carboxysomes, the micro-compartiments that encapsulate Rubisco with carbonic anhydrase facilitating carbon fixation. This strategy allowed to create a carbon fixation deficient *Synechocystis* sp. PCC 6803 mutant, which had not been possible via total Rubisco knockout. Construction of a Rubisco null mutant in *Synechocystis* has indeed been shown to be impossible[21]. The carboxylase function seems essential in this organism despite its capability of heterotrophic growth on glucose, probably because of the role that plays the CBB cycle in maintaining the cell redox balance. The described Rubisco inactivatation upon functional replacement with a Rubisco gene of a different photosynthetic organism offers an alternative approach.

The aim of this study was to prepare several mutant strains based on the replacement strategy, targeting Rubisco in the unicellular heterotrophe cyanobacteria *Synechocystis* sp. PCC 6803 strains, which is naturally competent for DNA transformation and thus quiet easy to manipulate genetically. The purpose was to create different background strains suitable to serve as platforms for the identification of new carbon fixation cycles. The idea behind this is that it would be difficult to see an improvement in carbon fixation on top of the already well established and naturally optimized CBB cycle. By constructing mutants impaired in their ability to fix CO_2 and with decreased growth rate, we will have the opportunity to detect any degree of rescue related to a new carbon fixation pathway.

Here we substitute the native *Synechocystis* Rubisco (large and small subunit) with genes from five different organisms (*Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, *Methanosarcina activorans*, *Methanocaldococcus jannaschii*, and *Rhodospirillum rubrum*). These different

candidates were chosen as they possess diverse forms of Rubisco, different from the cyanobacteria IB type. A fully segregated mutant was obtained only in *Synechocystis* with a transformant expressing *R. rubrum* Rubisco like previously reported[20].

3.2 Materials and Methods

3.2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich or Fisher Scientifics unless otherwise specified. KOD and KOD xtreme DNA polymerases were purchased from EMD Millipore and used for gene amplification from genomic or plasmid DNA. T5-Exonuclease (Epicenter), Taq DNA ligase (New England Biolabs) and *Phusion* DNA polymerase (Fisher) were purchased individually and used to make the assembly master mix (AMM) used for cloning as described by Gibson[24].

3.2.2 Bacterial strains and DNA manipulations

All chromosomal manipulations were carried out by homologous recombination of plasmid DNA into a glucose tolerant wild-type (WT) *Synechocystis* sp. PCC 6803 genome (kind donation from Pr. Vermaas, Arizona State University). All plasmids were constructed using the isothermal DNA assembly method[24]. *Escherichia coli* XL-1 Blue (Stratagene) was used to construct, propagate and store all plasmids in this study.

3.2.3 Culture medium and growth conditions

Synechocystis 6803 wild-type and mutant strains were grown at 30 °C in BG-11 medium or solid BG-11 medium as described by Eaton-Rye[25] under continuous illumination (35-40 μmol photons m^{-2} s^{-1} unless otherwise specified) in ambient air unless otherwise specified. BG-11 medium

was supplemented with gentamicin for GenR strains. Cell growth was monitored by measuring the optical density at 730 nm (OD_{730}) with a Beckman Coulter DU800 spectrophotometer. Mixotrophic cultures of *Synechocystis* 6803 were started in BG-11 medium supplemented with 5mM glucose as an organic carbon source. When growth under high CO_2 concentration was tested, cell cultures were placed in a BD GasPak EZ Anaerobe Gas-generating Pouch System with an Indicator (BD Biosciences). As described by the manufacturer, the Anaerobe Gas Generating Pouch System produces an atmosphere containing approximately 10% carbon dioxide and less than 1% oxygen.

E. coli strain XL-1 Blue was grown at 37 °C in LB broth or on solid LB agar 1.5% (wt/vol). Gentamicim (50 μg/ml) was added to the LB medium when required for the propagation of plasmids in *E. coli*.

3.2.4 Plasmids construction

The vector pLG101 was constructed using Gibson isothermal assembly[24] between four DNA fragments: colE1 origin, a Gentamicin resistance cassette, 1Kb DNA fragment directly upstream of the *rbcL* gene in Synechocystis 6803 and 1Kb fragment directly downstream of the *rbcS* gene. All primers used to amplify the four fragments were designed with approximately 25 bp of overlapping sequence. pLG101 is a vector designed to integrate any foreign DNA into the genome of synechocystis 6803 by replacing both the *rbcL* and *rbcS* genes. pLG101 was then opened by PCR just downstream of the *rbcL* upstream homologous region. Five individual Rubisco genes from genomic DNA of *Rodospirullum rubrum*, *Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, *Methanosarcina aceticorans,* and *Methanocaldococcus jannaschii* were PCR amplified by respective primers and cloned into pLG101 at the end of the upstream

homologous region under the control of a cbb operon promoter, creating the vectors pLG106, pFD29, pFD30, pFD31 and pFD32 respectively.

The recombination vector pFD21 for integration into the psbA2 site was also constructed using the isothermal Gibson assembly method between four fragments: psbA2 upstream-flanking region, a spectinamycin cassette, psbA2 downstream-flanking region, and the replication origin p15A. All primers used to amplify the four fragments were designed with approximately 25 bp of overlapping sequence. Integration in the open-reading frame of psbA2 would eventually allow to use the strong, light inducible psbA2 promoter to drive expression of any gene that would be cloned in frame with the 5' regulatory region while deleting the native psbA2 gene encoding the D1 protein of photosystem II. Because a compensatory expression of psbA3 exists when psbA2 is deleted[26] no phenotypic change has been reported from the use of the site.

3.2.5 Transformation of *Synechocystis* PCC. 6803

Transformation of *Synechocystis* was performed using a double homologous recombination system, and the heterologous Rubisco genes were integrated along with a Gentamycin resistance cassette into the target site of *Synechocystis* genomic DNA, producing strains M106, M29, M30, M31 and M32. Specifically, 5mL of *Synechocystis* 6803 from a mid-log phase culture (OD_{730nm}=0.5-0.8) was centrifuge at 4000 rcf for 10 min at room temperature. The pellet was suspended in 200μl of fresh BG-11 medium. Plasmid DNA (10 μl) with a concentration of 100-300μg/mL was then added to the *Synechocystis* cell suspension. The mixture was incubated under normal light conditions (35-40μE m^{-2} s^{-1}) for 6 hours, with a flicking at 3 hours. The culture was then diluted with 2mL of BG-11 with glucose (5mM) and shacked at 120rpm for 24 hours under normal light conditions. The culture was then centrifuged at 4000 rcf for 10 min at

room temperature and the supernatant was discarded except for 200µl in which the cells were resuspended. The culture was spread onto a BG-11 agar plate supplemented with 2.5µg/ml Gentamicin for selection of successful recombination. Colonies were detected after 8-10 days and screened by colony PCR using KOD Xtreme for successful integration of the insert into the genomic DNA. Transformants were repeatedly streaked on BG-11 plates (without glucose) with increasing amount of antibiotic (5µg/ml, 7.5µg/ml and 10µg/ml) and under anaerobic conditions with 10% CO_2, until segregation was confirmed by colony PCR.

3.3 Results and Discussion

3.3.1 Functional replacement of the WT Rubisco Gene in *Synechocystis*

To more easily enable screening of new synthetic carbon fixation pathways in cyanobacteria, the native carboxylating enzyme must be removed from the host cell. Full knockouts of Rubisco in *Synechocystis* have never been obtained and the carboxylase function is essential. Cyanobacteria is polyploid containing multiple copies of each chromosome[27,28], and attempts to insertionally inactivate the rbc genes always lead to meroploids containing both residual wt *rbc* sequence as well as the inactivated ones[21]. Thus, we decided to employ an alternate method used successfully in the past, which consist of the functional replacement of the *rbc* operon with an heterologous Rubisco gene. To do so we constructed plasmids in which a 1kb fragment directly upstream of *rbc*L, containing the regulatory region of the cyanobacterial *rbc* operon, was fused individually to the start codon of 5 heterologous rubisco genes, followed by a gentamicin cassette and the 3' region (1kb) directly downstream of *rbc*S (Fig. 3-1A). These constructs allowed the insertion of the heterologous rubisco and the Gen^R gene into the *rbc* locus via site specific recombination between the homologous regions flanking the *rbc*L and *rbc*S. Transformation of WT

synechocystis with each plasmid (pLG106, pFD29, pFD30, pFD31, pFD32) separately gave gentamicin resistant colonies after about 8-10 days, resulting in mutants M106, M29, M30, M31 and M32 (Fig. 3-1B). The transformants were continuously restreaked on gentamicin plates approximately every 2 weeks to allow for gene segregation. To help with segregation, the agar plates were placed in an anaerobic atmosphere with 10% CO_2. A fully segregated mutant was obtained only with transformants M106 after 6 months of consecutive rounds of restreaking. The genomic analysis of M106 is shown in figure 3-2. No wt band could be detected after colony PCR of the transformants (Fig. 3-2A), and insertion via double crossover was confirmed with three different sets of primers (Fig. 3-2B). This mutant is the very similar to the one constructed by Amichay et al.[20](Syn6803Δrbc), except for the fact that ours was constructed directly from *Synechocystis* wt while Amichay and its collegues transformed a mutant with an already impaired carboxylase previously built by Pierce et al[21]. Their first attempt to replace directly *Synechocystis* Rubisco with the *R. rubrum* Rubisco yields no colonies.

3.3.2 Growth Characterization of mutant M106

Determination of growth rates of *Synechocystis* mutant M106 revealed that it was unable to grow in air (0.035% CO_2/21% O_2) in the absence or presence of glucose (Fig. 3-3A). This absence of heterotrophic growth in presence of glucose might be associated with photoinhibition caused by an excess of reducing equivalents derived from photosystem II and glucose. To circumvent this problem and bypass CO_2 fixation as a means to maintain redox balance[29] we tried to grow the mutant in the presence of DMSO, an electron acceptor, but without success (Fig. 3-3B). It has also been demonstrated that this type of cyanobacteria mutants are deficient in carboxysomes. Electron microscopy of a similar strain revealed the absence of the typical polyhedral electron-

dense structures found in WT[21]. Carboxysomes are polyhedral microcompartment composed of multiple shell proteins in which rubisco and carbonic anhydrase reside[30]. The co-localization of rubisco and carbonic anhydrase inside the microcompartments is believed to allow for efficient sequestration and utilization of bicarbonate as carbonic anhydrase converts it into carbon dioxide the proper substrate of rubisco. Co-localization of these enzymes also increases rubisco efficiency by elevating local carbon dioxide substrate levels around it. Furthermore, it has been speculated that the shell proteins possess specific permeability properties that sequester carbon dioxide in the vicinity of rubisco further increasing its substrate availability. The aforementioned functions of carboxysomes and their probable functional absence in the mutant M106 could contribute to explain the O_2 sensitivity of the mutant as well as its requirement for high CO_2 concentration atmosphere to grow.

When using 10% CO_2 in anaerobic atmosphere (10% CO_2/<0.1% O_2) and lower light intensity (15μE m^{-2} s^{-1}) M106 mutant growth rate was identical to that measured with the wt under those same conditions (Fig. 3-3C), but still far from the maximal growth rate of wt cells when grown in air (0.035% CO_2/21% O_2).

3.3.3 Transformation of mutant M106

Mutant M106 genotype was suitable to serve as a platform for the identification of new carbon fixation cycles. If transformation with genes coding for a new carbon fixation pathway turned out to be possible, this selection platform would allow to easily detect any degree of growth rescue related to a new pathway.

Incubation of the mutant M106 in presence of a plasmid vector bearing a spectinomycin resistance cassette surrounded by the psbA2 site flanking regions (pFD21) did not yield any

transformants. Incubation of the cells in high CO_2 environment prior to their transfer in air like described by Amichay et al[20] did not give any transformant either, despite our flanking sequences being in the range discussed in their study. This difficulty to transform our rubisco mutant compare to what have been described before might be explained by: (1) Mutant M106, despite a growth rate similar to wt under high CO_2 anaerobic conditions, showed a yellowish phenotype, sign of unhealthy cells in the early stage of photobleaching, even though the cells were grown under low light conditions (15μE m^{-2} s^{-1}). Mutant M106 cells might have been too damaged to survive transformation; (2) the transformation procedure established previously on Syn6803Δrbc was used to restore the native *rbc* locus, procuring some clear advantages to the cells. In our case, M106 was transformed with a plasmid for integration of a spectinomycin resistance marker in a new locus. The possible related stress added on the cells that now needed to express two different antibiotic resistance genes might also explain the absence of transformants. Further evaluation was performed by going through the transformation procedure on M106 cells without addition of any plasmid DNA. No cells survived the procedure, indicating a too strong phenotype, where the cells are too sick to be transformed, unless the new acquired DNA provides them a clear advantage.

3.4 Conclusion

The purpose of the present study was to create a suitable host that can be used for the introduction and evaluation of new carbon fixation pathways. Here we report on the construction of a Rubisco mutant (M106) in which the entire chromosomal region cloning for Rubisco, and thus containing both *rbc*L and *rbc*S genes, had been replaced by a type II Rubisco from *R. rubrum*. We were able to show full replacement, full segregation and the impaired growth

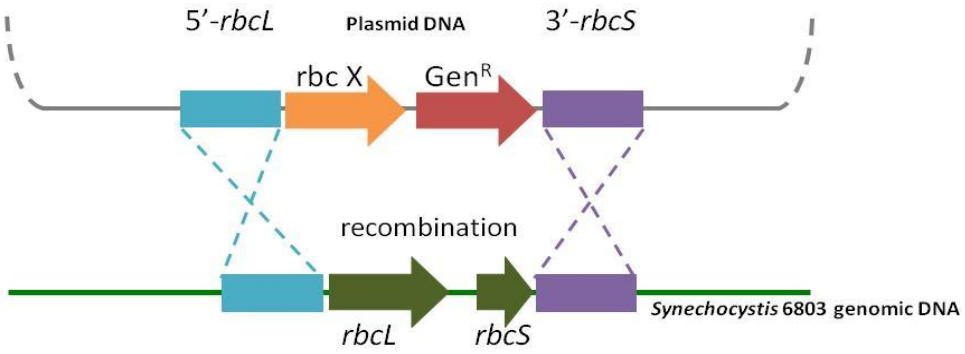
phenotype desired, all of which were required traits for our selection platform. Unfortunately we were unable to transform the mutant with an additional plasmid that would have allowed the insertion via homologous recombination of the genes coding for the new carbon fixation pathway. This final mutant is the same as the one constructed by Amichay et al.[20], but ours was constructed directly from *Synechocystis* wild-type. Amichay and its colleagues transformed a mutant with an already impaired carboxylase previously built by Pierce et al.[21] because their first attempt to replace directly *Synechocystis* Rubisco with the *R. rubrum* Rubisco yields no colonies.

3.5 Tables and Figures

Rubisco Form	Subunit composition	Rubisco activity	Enzymatic function	Distribution
IA	L_8S_8	+	CBB cycle	Cyanobacteria, proteobacteria
IB	L_8S_8	+	CBB cycle	Cyanobacteria, green algae, plants
IC	L_8S_8	+	CBB cycle	Proteobacteria
ID	L_8S_8	+	CBB cycle	Non-green algae
II	L_2 and L_n	+	CBB cycle	Proteobacteria, archaea
III	L_2 and $(L_2)_5$	+	RuPP cycle	Archaea
IV (RLP)	L_2?	-	Methionine salvage pathway	Bacteria, archaea

Table 3-1. Rubisco enzymes forms, their properties and principal distribution. CBB: Calvin Benson Bassham, RuPP: 5-phospho-D-ribose-1-pyrophosphate, RLP: Rubisco-like protein.

A



B

Strain	Gene expressed	Rubisco Form	Organism
M106	*rbcL*	Form II	*Rhodospirillum rubrum*
M29	*rbpL*	Form II	*Rhodobacter sphaeroides*
M30	*cbbM*	Form II	*Rhodobacter capsulatus*
M31	*rbcL*	Form III	*Methanococcus activvorans*
M32	*rbcL*	Form III	*Methanocaldococcus jannaschii*

Figure 3-1. (A) Schematic representation of recombination to integrate a heterologous Rubisco gene in place of the native *Synechocystis rbc* operon. Different CO_2 fixation genes (rbc X) were used to replace *rbc*L and *rbc*S. (B) List of strains with different carboxylase genes overexpressed used in this study.

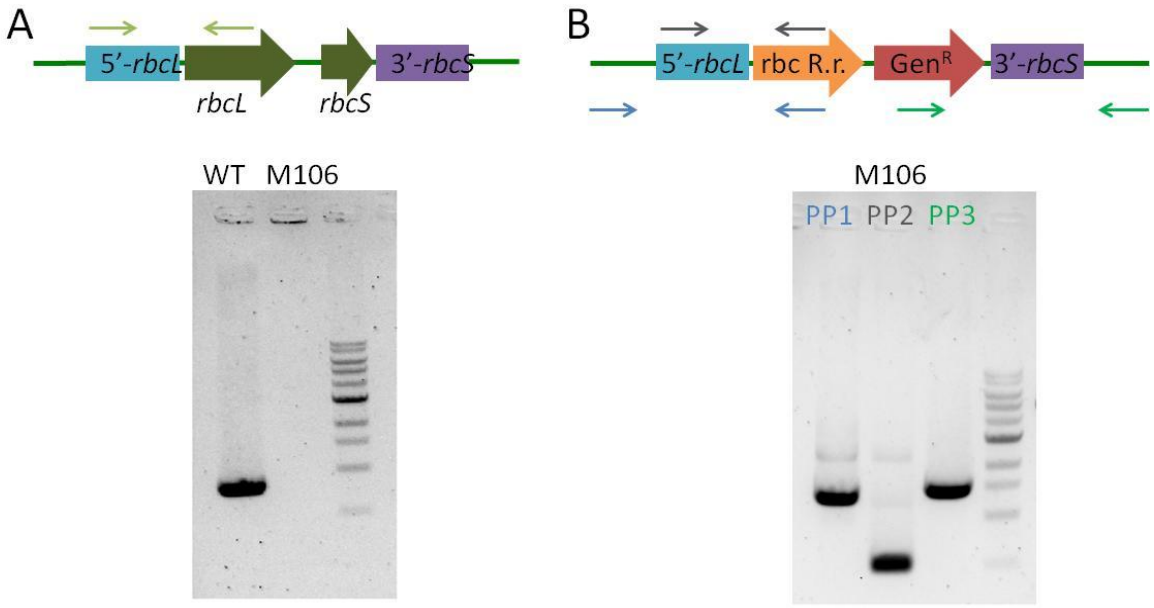


Figure 3-2. Colony PCR analysis of the genome in Mutant M106. Four different sets of primers, specified above each gel picture and illustrated as arrows were used. One set was used on both WT *Synechocystis* and M106 (A) to show the absence of the rbc operon in the M106 mutant compare to WT *Synechocystis*. Three sets of primers were used on the mutant M106 (B) to specifically show the presence of the *R. rubrum* Rubisco and its insertion by double crossover.

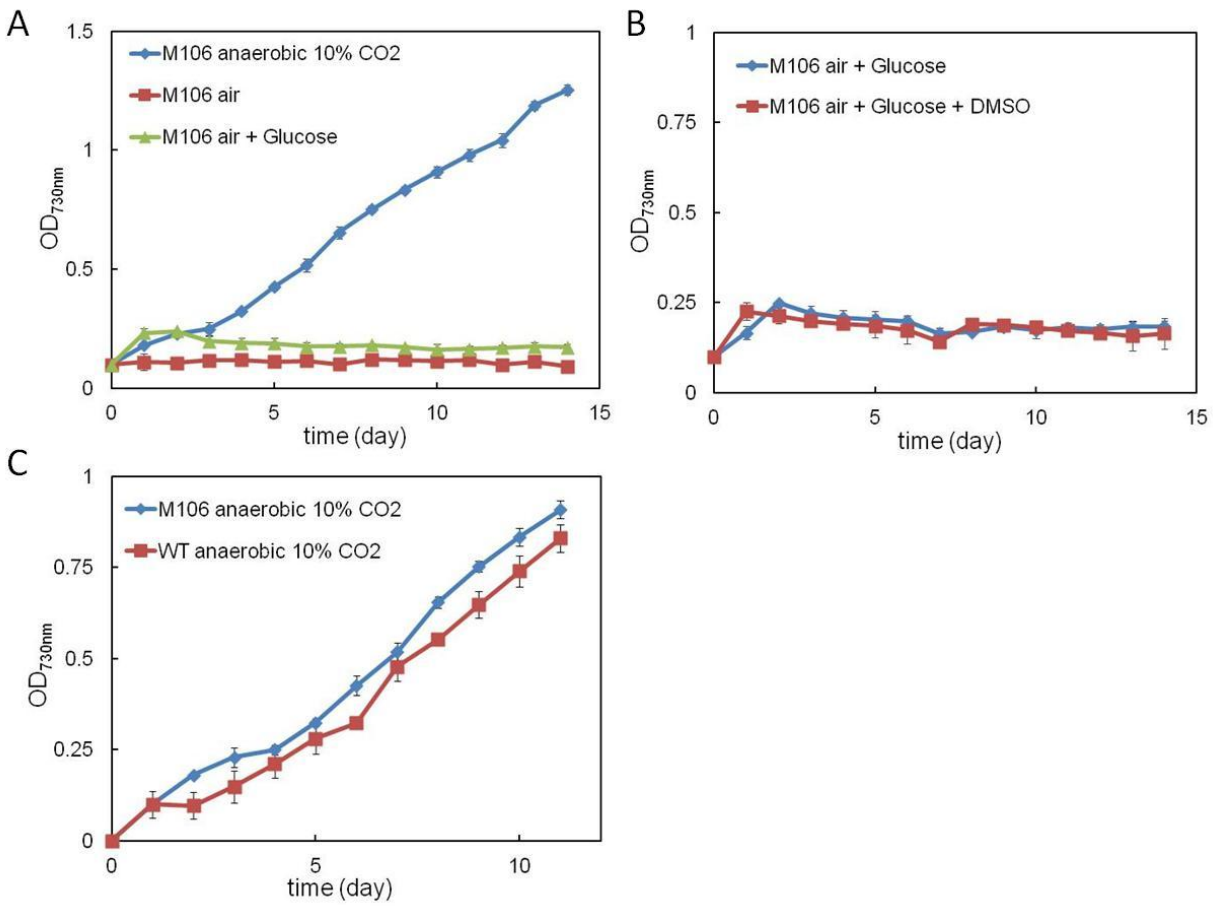


Figure 3-3. (A) Growth curve of mutant M106 under three different conditions in BG-11: anaerobic atmosphere with 10% CO_2, aerobic air (0.035% CO_2/21% O_2) and aerobic air supplemented with 5mM glucose. (B) Growth curve of mutant M106 in BG-11 supplemented with 5mM glucose with and without DMSO (5μM). (C) Comparison of wt *Synechocystis* 6803 and mutant M106 growth under the same conditions; BG-11, low light (15μE m^{-2} s^{-1}), anaerobic atmosphere with 10% CO_2.

3.6 References

1. H M Miziorko & Lorimer, and G. H. Ribulose-1,5-Bisphosphate Carboxylase-Oxygenase. *Annu. Rev. Biochem.* **52,** 507–535 (1983).

2. Jordan, D. B. & Ogren, W. L. Species variation in kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **227,** 425–433 (1983).

3. Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. & Scott, S. S. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* **59,** 1515–1524 (2008).

4. Hanson, T. E. & Tabita, F. R. A ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from Chlorobium tepidum that is involved with sulfur metabolism and the response to oxidative stress. *Proc. Natl. Acad. Sci.* **98,** 4397–4402 (2001).

5. Singh, J. & Tabita, F. R. Roles of RubisCO and the RubisCO-Like Protein in 5-Methylthioadenosine Metabolism in the Nonsulfur Purple Bacterium Rhodospirillum rubrum. *J. Bacteriol.* **192,** 1324–1331 (2010).

6. G Schneider, Y Lindqvist & Branden, and C. I. Rubisco: Structure and Mechanism. *Annu. Rev. Biophys. Biomol. Struct.* **21,** 119–143 (1992).

7. Andersson, I. & Taylor, T. C. Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **414,** 130–140 (2003).

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

9. Andersson, I. *et al.* Crystal structure of the active site of ribulose-bisphosphate carboxylase. *Nature* **337,** 229–234 (1989).

10. Schneider, G., Lindqvist, Y. & Lundqvist, T. Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from Rhodospirillum rubrum at 1.7 Å resolution. *J. Mol. Biol.* **211,** 989–1008 (1990).

11. Watson, G. M. F., Yu, J.-P. & Tabita, F. R. Unusual Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase of Anoxic Archaea. *J. Bacteriol.* **181,** 1569–1575 (1999).

12. Ezaki, S., Maeda, N., Kishimoto, T., Atomi, H. & Imanaka, T. Presence of a Structurally Novel Type Ribulose-bisphosphate Carboxylase/Oxygenase in the Hyperthermophilic Archaeon,Pyrococcus kodakaraensis KOD1. *J. Biol. Chem.* **274,** 5078–5082 (1999).

13. Kreel, N. E. & Tabita, F. R. Substitutions at Methionine 295 of Archaeoglobus fulgidus Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Affect Oxygen Binding and CO2/O2 Specificity. *J. Biol. Chem.* **282,** 1341–1351 (2007).

14. Finn, M. W. & Tabita, F. R. Synthesis of Catalytically Active Form III Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase in Archaea. *J. Bacteriol.* **185,** 3049–3059 (2003).

15. Finn, M. W. & Tabita, F. R. Modified Pathway To Synthesize Ribulose 1,5-Bisphosphate in Methanogenic Archaea. *J. Bacteriol.* **186,** 6360–6366 (2004).

16. Li, H., Sawaya, M. R., Tabita, F. R. & Eisenberg, D. Crystal Structure of a RuBisCO-like Protein from the Green Sulfur Bacterium Chlorobium tepidum. *Structure* **13,** 779–789 (2005).

17. Tabita, F. R. *et al.* Function, Structure, and Evolution of the RubisCO-Like Proteins and Their RubisCO Homologs. *Microbiol. Mol. Biol. Rev.* **71,** 576–599 (2007).

18. Ninomiya, N., Ashida, H. & Yokota, A. in *Photosynthesis. Energy from the Sun* 867–870 (Springer, 2008).

19. Tcherkez, G. G., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci.* **103,** 7246–7251 (2006).

20. Amichay, D., Levitz, R. & Gurevitz, M. Construction of a Synechocystis PCC6803 mutant suitable for the study of variant hexadecameric ribulose bisphosphate carboxylase/oxygenase enzymes. *Plant Mol. Biol.* **23,** 465–476 (1993).

21. Pierce, J., Carlson, T. J. & Williams, J. G. A cyanobacterial mutant requiring the expression of ribulose bisphosphate carboxylase from a photosynthetic anaerobe. *Proc. Natl. Acad. Sci. U. S. A.* **86,** 5753–5757 (1989).

22. Jordan, D. B. & Ogren, W. L. Species variation in kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **227,** 425–433 (1983).

23. Whitney, S. M., Houtz, R. L. & Alonso, H. Advancing Our Understanding and Capacity to Engineer Nature's CO2-Sequestering Enzyme, Rubisco. *Plant Physiol.* **155,** 27–35 (2011).

24. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

25. Eaton-Rye, J. in *Photosynthesis Research Protocols* (ed. Carpentier, R.) 309–324 (Humana Press, 2004).

26. Mohamed, A. & Jansson, C. Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium Synechocystis 6803. *Plant Mol. Biol.* **13,** 693–700 (1989).

27. Herdman, M., Janvier, M., Rippka, R. & Stanier, R. Y. Genome Size of Cyanobacteria. *Microbiology* **111,** 73–85 (1979).

28. Vermaas, W. Molecular genetics of the cyanobacteriumSynechocystis sp. PCC 6803: Principles and possible biotechnology applications. *J. Appl. Phycol.* **8,** 263–273 (1996).

29. Richardson, D. J. *et al.* The role of auxiliary oxidants in maintaining redox balance during phototrophic growth of Rhodobacter capsulatus on propionate or butyrate. *Arch. Microbiol.* **150,** 131–137 (1988).

30. Cannon, G. C. *et al.* Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. *Appl. Environ. Microbiol.* **67,** 5351–5361 (2001).

4. Implementation of a heterologous carbon fixation pathway in cyanobacteria

4.1 Introduction

Photosynthesis is the process used by plants, algae and certain bacteria like cyanobacteria to convert solar energy, CO_2 and water into chemical energy in the form of sugars. Much of the limitation of CO_2 assimilation in photosynthetic organisms has been attributed to the catalytic properties of the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), the carbon fixing enzyme of the Calvin-Benson-Bassham (CBB) cycle for carbon fixation. Not only is Rubisco an inefficient enzyme because of its low turnover number, but it also catalyses two competing reactions: carboxylation and oxygenation. Previous work to improve CO_2 fixation mainly focused on Rubisco and how to improve the enzyme's undesirable characteristics[1]. However, results have shown that improving specificity often comes with a lost in the catalytic activity, suggesting that Rubisco may have naturally been optimized for the best compromise between substrate specificity and maximum turnover rate[2,3]. Other approaches focusing on overexpression of different enzymes of the CBB cycle or enzymes linked to carbon transportation showed some encouraging outcomes while others were inconclusive[4,5]. Taken as a whole, these results suggest that improving the existing CBB cycle without fundamentally changing the carbon fixation network is unlikely to be successful.

In the quest to enhance photosynthesis, cyanobacteria have been used for decades as model organisms because of their evolutionary relationship to chloroplasts, rapid life cycle compared to plants, and ease of genetic manipulation[6,7]. Thus, molecular understanding of cyanobacteria carbon fixation and its possible improvement provide tools that could be further exploited for plant engineering.

The particular aim of this study is to develop and implement a new pathway for more efficient CO_2 fixation in cyanobacteria. Implementation in cyanobacteria will allow for an analysis of the pathway efficiency in a photosynthetic organism. We chose two different cyanobacteria model organisms, *Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803, as well as two pathways for CO_2 fixation, the reductive tricarboxylic acid (TCA) cycle and a synthetic pathway modelled on a reverse version of the glyoxylate shunt. The reductive TCA cycle (Fig. 4-3B) is found in diverse autotrophic anaerobic and microaerophilic bacteria and archea. It allows for the fixation of two molecules of CO_2 while forming one molecule of acetyl-CoA, which is further carboxylated to pyruvate, from which all other central metabolites can be produced. The process requires only two ATP, compared to seven for the formation of pyruvate though the CBB cycle, making the reductive TCA cycle 3.5 time more energy efficient. On the other hand the reductive TCA involves two unusual enzymes for fixing CO_2, pyruvate:ferredoxin oxidoreductase (POR) and α-ketoglutarate ferredoxin oxidoreductase (KOR), both of are oxygen sensitive and need to be compatible with the reduce ferredoxin generated by the host photosystems. An alternative to this cycle is to use pyruvate carboxylase and the reversal of the glyoxylate shunt, which has previously been shown reversible[8], to complete the cycle (Fig. 4-3A). The resulting pathway, termed reverse glyoxylate shunt (rGS) cycle uses one more ATP than the reductive TCA cycle, but fixes CO_2 through only one ferredoxin-oxidoreductase (POR), taking advantage for the other carboxylation step of the highly active pyruvate carboxylase from bacteria. Both the rTCA and the rGS cycle do not utilize Rubisco for CO_2 fixation, and thus do not have the carbon and energy loss associated with photorespiration and the CBB cycle. Here we sought to implement these two CO_2 fixation cycles in cyanobacteria. After a first screening of enzymes for each step of the passways we favored the rGS over the rTCA cycle, this last one involves more oxygen

sensitive enzymes. Furthermore, implementation of the full pathway was carried out in *Synechococcus elongatus* PCC 7942 only, as transformation efficiency in *Synechocystis* sp. PCC 6803 with the required larger plasmids was unsuccessful. Because no oxygen tolerant pyruvate:ferredoxin oxydoreductase was identified, we report the integration of a linear pathway that fixes carbon through pyruvate carboxylase, on top of the CBB cycle, to theoretically produce acetyl-CoA or acetate if an acetyl-CoA ligase is co-expressed. Experimentally, however, only fumarate was detectable in small concentration, leading to the conclusion that only the first half of the pathway was active.

4.2 Materials and Methods

4.2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich or Fisher Scientifics unless otherwise specified. KOD and KOD xtreme DNA polymerases were purchased from EMD Millipore and used for gene amplification from genomic or plasmid DNA. T5-Exonuclease (Epicenter), Taq DNA ligase (New England Biolabs) and *Phusion* DNA polymerase (Fisher) were purchased individually and used to make the assembly master mix (AMM) used for cloning as described by Gibson[9].

4.2.2 Bacterial strains and DNA manipulations

Two different strains of cyanobacteria were used in this study: 1- A glucose tolerant wild-type (WT) *Synechocystis* sp. PCC 6803, hereafter referred as *Synechocystis* (obtained from Pr. W. Vermaas, Arizona State University); 2-WT *Synechococcus elongatus* PCC 7942, hereafter referred as *Synechococcus* (obtained from Pr. Susan S. Golden, University of California, San

Diego). All chromosomal manipulations were carried out by homologous recombination of plasmid DNA into *S. elongatus* genome at neutral site I (NSI)[10] , II (NSII)[11] and III (NSIII)[12]. All plasmids were constructed using the isothermal DNA assembly method[9]. *Escherichia coli* XL-1 Blue (Stratagene) was used to construct, propagate and store all plasmids in this study, as well as for protein expression and enzyme assays.

For POR selection and evolution, the *E. coli* strains JCL301 and JCL302 constructed in house were used. Both strains are blocked in their ability to convert into acetyl-CoA or acetate. JCL301 is based on the recombinant strain YYC202[13] but with point mutation disruptions of genes replaced by complete chromosomal deletions. JCL301 thus has four knochouts, of genes coding for pyruvate dehydrogenase (*aceEF*), pyruvate formate lyase (*pflB*) and pyruvate oxidase (*poxB*). In the strain *E. coli* JCL302, the gene coding for succynyl-CoA synthase (*SucC*) has been deleted on top of JCL301 genetic bachground as it has been identified as a possible loophole during longer evolution experiments.

4.2.3 Culture medium and growth conditions

All *S. elongatus* PCC 7942 strains were grown at 30 °C under continuous illumination supplied by four Lumichrome F30W-1XX 6500K 98CRI light tubes (100 μmol photons m^{-2} s^{-1} unless otherwise specified) in solution of BG-11 (Sigma, Cyanobacteria BG-11 Freshwater Solution 50x) or on modified BG-11 agar (1.5% w/v) plates[14] . Strains were cultured in 250 mL screw cap flasks or 5ml glass tubes. 50 mM $NaHCO_3$ was added to the medium. Cell growth was monitored by measuring OD_{730} with Beckman Coulter DU800 spectrophotometer. *Synechocystis* 6803 wild-type weas grown at 30 °C in BG-11 medium or solid BG-11 medium as described by Eaton-Rye[15] under continuous illumination (35-40 μmol photons m^{-2} s^{-1} unless otherwise specified) in

a under ambient air unless otherwise specified. Cell growth was monitored by measuring the optical density at 730 nm (OD_{730}) with a Beckman Coulter DU800 spectrophotometer. BG-11 media were supplemented with antibiotics (20 μg⁄mL spectinomycin and 10 μg⁄mL kanamycin, 5 μg⁄mL gentamycin, 5 μg⁄mL chloramphenicol) when appropriate. Mixotrophic cultures of *Synechocystis* 6803 were started in BG-11 medium supplemented with 5mM glucose as an organic carbon source.

For production analysis, *Synechococcus* cells in exponential phase were diluted to an OD_{730} of 0.4 in BG-11 medium supplemented with 50 mM $NaHCO_3$, 0.5 mM IPTG and the necessary antibiotics.

E. coli strain XL-1 Blue was grown at 37 °C in LB broth or on solid LB agar 1.5% (wt/vol). Spectinomycin (100 μg⁄mL), kanamycin (50 μg/ml), gentamycin (15 μg/ml), and chloramphenicol (20 μg/ml), were added to the LB medium when required for the propagation of plasmids in *E. coli*.

For high-throughput growth of *Synechococcus* in 96-well plate 150 μl of cyanobacteria liquid culture of OD_{730} equals to 0.1 was aliquoted in each well of the 96-well plate. Also the outer wells were not used for actual samples (they receive more light than the rest), they were filled with WT *Synechococcus.* This insured that each sample well had a similar surrounded environment. The 96-well plate was placed on an elevated rack on a ground shaker, so as to get equal airflow on all sides of the plate. The plate was covered with a transparent plastic lid and parafilmed to minimize evaporation. A shaking speed of 400rpm was used and light intensity of 30.5 μE. All strains were induced 24 hours after the start of the experiment and OD was monitored twice a day using a 96 well plate reader Bio-TEK powerwave XS.

4.2.4 Plasmids Construction

For single enzymatic step evaluation at least 4 homologs of each non-native cyanobacteria genes required for the rGS and/or rTCA cycles were amplified from the genome of different organisms (see Table 4-1). For integration into *Synechococcus* the genes were cloned into a NSI vector backbone[16] using the Gibson method of DNA assembly, under the control of an IPTG-inducible promoter PLlacO1 and followed by a spectinomycin resistance cassette, $Spec^R$, as the selective marker. For *Synechocystis* the genes were cloned into the recombination vector pFD21 for integration into the psbA2 site (described in chapter 1), also using the isothermal Gibson assembly method. Integration in the open-reading frame of psbA2 allow to use the strong, light inducible psbA2 promoter to drive expression of any gene that would be cloned in frame with the 5' regulatory region while deleting the native psbA2 gene encoding the D1 protein of photosystem II. Because a compensatory expression of psbA3 exists when psbA2 is deleted[17] no phenotypic change has been reported from the use of the site.

Plasmids containing multiple genes were also assembled using the Gibson DNA assembly method. For the initial rounds of assembly the fragments were all amplified by PCR. However the larger plasmids were built sequentially by adding a unique restriction site (ascI) after the first gene. This allowed the vector to be opened and assembled with the second gene in such a way as to destroy the restriction site at the front and regenerate it at the back of the newly added gene (Fig. 4-1). This method allowed for iterative assembly of the large vectors, when PCR of such large fragment became difficult. RGS genes were cloned, as listed in table 2, on 3 separate plasmids each one targeting NSI, NSII or NSII. When *aceB* from *E. coli*, or *acd* from *P. furiosus* were added, they were cloned in NSIII.

4.2.5 Transformation of *S. elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803

S. elongatus is naturally competent and was transformed as followed: 2ml of cells at mid-log phase ($OD_{730}$of 0.4–0.6) were spun down at room temperature and resuspended in 200µl of fresh BG-11. Cells were then incubated with 1-3 µg of plasmid DNA overnight in the dark, before being spread on BG-11 plates with appropriate antibiotics for selection of successful recombination. Colonies grown on BG-11 agar plates were analyzed by PCR using gene-specific primers to verify integration of inserted genes into the recombinant strain. In all cases, eight individual colonies were analyzed by PCR for verification. One positive colony was then propagated and tested for downstream experiments.

Synechocystis is also naturally competent and was transformed as followed: 5mL of *Synechocystis* 6803 from a mid-log phase culture (OD_{730nm}=0.5-0.8) was centrifuge at 4000 rcf for 10 min at room temperature. The pellet was suspended in 200µl of fresh BG-11 medium. Plasmid DNA (10 µl) with a concentration of 100-300µg/mL was then added to the *Synechocystis* cell suspension. The mixture was incubated under normal light conditions (35-40µE m^{-2} s^{-1}) for 6 hours, with a flicking at 3 hours. The culture was then diluted with 2mL of BG-11 with glucose (5mM) and shacked at 120rpm for 24 hours under normal light conditions. The culture was then centrifuged at 4000 rcf for 10 min at room temperature and the supernatant was discarded except for 200µl in which the cells were resuspended. The culture was spread onto a BG-11 agar plate supplemented with the right antibiotic for selection of successful recombination. Colonies were detected after 8-10 days and screened by colony PCR using KOD Xtreme for successful integration of the insert into the genomic DNA. When an electroporation step was added for transformation, a pulser from Bio-rad Laboratories like previously described[18].

4.2.6 Enzyme Assays

Enzyme assays were conducted by using Bio-Tek PowerWave XS microplate spectrophotometer or a Beckman Coulter DU800 spectrophotometer when anaerobic cuvettes were used. Cell-free extracts were prepared as followed: Cyanobacteria liquid cultures (25ml) were grown until OD_{730nm} reaches 0.4-0.6, and were then induced with 1 mM IPTG. The induced cultures were further grown for 2 days to allow protein expression. Cells were collected by centrifugation at 5,250 × g for 15 min (4°C). The pellet was then resuspended in 1 mL of the appropriate assay buffer and mix with 500µl of 0.1 mm glass beads (Biospec). The sample was then homogenated using a bead beater (biospec). The soluble protein fraction was collected after centrifugation. Purified protein was collected after His-spin column purification following the manufacturer protocol (Zymo Research).Total protein measurements were made using the Pierce Coomassie Plus Assay Reagent (ThermoFisher Scientific). When appropriate the protein samples were run on SDS/ PAGE using precast gels from Biorad following normal protocol.

Aconitate activity (ACN) was monitored by the increase of absorbance at 240 nm which correspond to the conversion of isocitrate to aconitate[19]. The reaction mixture contained 100mM potassium phosphate buffer pH 7.5 and crude extract. The reaction was initiated with the addition of 8 µl of isocitrate, 250mM, for a final reaction volume of 200 µl. Aconitase activity is expressed as U per mg of total soluble protein at 30 °C, where U = conversion of 1 µmol of substrate/min.

Malate dehydrogenase (MDH) activity was measured in the direction of the reduction of oxaloacetate to malate. The concomitant oxidation of NADH was monitored by a decrease in absorbance at 340 nm[20]. A 200 µl reaction mixture contained 100mM potassium phosphate

buffer pH 7.5, 0.375 mM NADH, and total soluble protein from bacteria crude extract. The reaction was started with the addition of oxaloacetate (0.6 mM final concentration). Mdh activity is expressed as U of NADH oxidized per mg of soluble protein at 30 °C (U = μmol/min).

Fumarase (FUM) activity was measured according to a modified protocol based on the Sigma-aldrich one[21]. The reaction mixture was composed of 100mM potassium phosphate buffer pH 7.5 and 50mM malate. The bacteria crude extract was added to start the reaction for a final reaction volume of 200 μl. Absorbance was measured at 240 nm at 30 °C. The activity is expressed as μmol of fumarate formed per min per mg of soluble protein (U/mg of total soluble protein).

Quinol:fumarate reductase (FRD) assay was determined by a methyl viologen-linked assay. Since this assay requires reduced methyl viologen, which can be oxidized by exposure to air, this assay was performed anaerobically in a sealed cuvette. Working in an anaerobic chamber, reagents were added in 1-ml quartz cuvettes with stoppers (reagents were kept anaerobic during the course of the assay). The reaction mixture contained 75 mM sodium phosphate buffer (pH 6.8), 0.2 mM methyl viologen, and cell extract. Freshly made 20 mM sodium dithionite was then injected into each cuvette with a syringe until the absorbance at 585 nm reached 0.8 to 0.9, which represented half-reduced methyl viologen. An anaerobic solution of sodium fumarate (final concentration, 5 mM) was added, and the methyl viologen oxidation was recorded at 585 nm. Fumarate reductase activity was expressed in Units (U) of methyl viologen oxidized per min per mg of soluble protein.

NADH: fumarate reductase (FRD) activity was assayed by monitoring the oxidation of NADH at 340 nm. The assay was carried out at 30 °C in 100 mM sodium phosphate pH 7.5, with 0.375 mM NADH, and bacteria soluble crude extract. To start the reaction fumarate pH 7.5 was added

for a final concentration of 50 mM in 200 µl. Fumarate reducatse activity is expressed as U of NADH oxidized per mg of soluble protein at 30 °C (U = µmol/min).

Succinate thiokinase (STK) assay was determined discontinuously in a reaction mixture containing 50 mM Tris buffer pH 7.5, 10 mM $MgCl_2$, 0.08 mM ATP, 0.05 mM CoA and the cells soluble crude extract. The reaction was started with the addition of succinate for a final concentration of 50mM. Samples were taken every 2 minutes and reacted with excess 5.5'-dithiobis(2-nitrobenzoate) (DTNB). DTNB reacts with the –SH group on the unreacted free CoA to form 5-thio-2-nitrobenzoic acid (TNB). Absorbance of TNB was measured at 412 nm. Succinate thiokinase activity is expressed as U per mg of total soluble protein at 30 °C, where U = conversion of 1 µmol of substrate/min.

Pyruvate carboxylase (PYC) activity was determined by coupling the reaction to malate dehydrogenase (MDH) from porcine heart (sigma), which oxidizes NADH. The assay mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 5 mM ATP, 0.6 mM Acetyl-CoA, 15 mM $NaHCO_3$, 0.2 mM NADH, 5 mM pyruvate, 0.5 units/ml MDH and protein extract or purified protein. The reaction was started at 30 °C with the addition of pyruvate, and decrease in absorbance was followed at 340 nm. The activity is expressed as µmol of malate formed per min per mg of soluble protein (U/mg of total soluble protein), which is equivalent to the amount of NADH oxidized.

ATP-citrate lyase (ACL) catalyzes the cleavage of CoA-activated citrate into acetyl-CoA and oxaloacetate (OAA). The activity of this enzyme was measured by coupling the production of OAA to the activity of malate dehydrogenase (MDH) which oxidizes NADH and can therefore be monitored spectrophotometrically. The reaction mixture contained 250 mM Tris-HCL (pH

8.5), 1 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 20 mM citrate, 10mM ATP, 0.2mM NADH, 0.2 mM CoA, 0.5 units/ml MDH and soluble cell lysate. The reaction was started at 30 °C with the addition of citrate, and decrease in absorbance at 340 nm was followed. The activity is expressed as U per mg of total soluble protein at 30 °C, where U = conversion of 1 μmol of substrate/min.

Isocitrate lyase (ICL) activity in the isocitrate-forming direction was coupled with that of commercially available isocitrate ($NADP^+$) dehydrogenase from *Bacillus subtilis*[22]. Glyoxylate- and succinate-dependent reduction of $NADP^+$ was measured by monitoring the formation of NADPH at 340nm. The assay mixture (200μl) was composed of 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 100 mM NaCl, 1 mM dithiothreitol (DTT), 5 mM $NADP^+$, 10 mM succinate, 10 mM glyoxylate,0.3U of ICD and soluble cell lysate. The reduction of $NADP^+$ to NADPH was followed at 340nm, and the activity of ICL expressed as U per mg of total soluble protein at 30 °C, where U = conversion of 1 μmol of substrate/min.

Malate thiokinase (MTK) and malyl-CoA lyase (MCL) activities were tested simultaneously. The assay used works on the following principle: MTK performs the ATP-dependent condensation of malate and CoA into malyl-CoA. In turn, MCL cleaves malyl-CoA into acetyl-CoA and glyoxylate, the latter reacting with phenylhydazine to form glyoxylate-phenylhydrazone. Formation of glyoxylate-phenylhydrazone is recorded at 324 nm. The assay mixture contained 50 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 2 mM phenylhydrazine, 5 mM malate, 1 mM ATP, 0.5 mM CoA, and soluble protein extract. The reaction was started at 30 °C with the addition of malate, and increase in absorbance was followed at 324 nm. The activity is expressed as μmol of glyoxylate formed per min per mg of soluble protein (U/mg of total soluble protein).

Isocitrate dehydrogenase (ICD) the reversible convertion of α-ketoglutarate to isocitrate utilizing either NADP(H) or NAD(H). The assay was performed in the oxidative direction, the decarboxylation of isocitrate, which releases CO_2 and produces NAD(P)H. The reaction mixture (200µl) was composed of 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 100 mM NaCl, 1 mM dithiothreitol (DTT), 0.5mM $NADP^+$, 10mM isocitrate and soluble cell lysate. The reaction was started at 30 °C with the addition of isocitrate and increase NADPH absorbance at 340nm was followed. The activity is expressed as µmol of product formed per min per mg of soluble protein (U/mg of total soluble protein).

Alpha-ketoglutarate oxidoreductase (KOR) catalyzes the reductive carboxylation of succinyl-CoA to α-ketoglutarate. Its activity was tested using an aerobic *in vitro* assay adapted from Kerscher et al[23]. This assay measures KOR activity in the oxidative direction, leading to the production of reduced ferredoxin. The ferredoxin in turn reduces cytochrome C, which can then absorb light at 550nm. The reaction mixture contained 50 mM Tris-HCl (pH 7.5), 10mM α-ketoglutarate, 0.9 µg/ml cytochrome C, 0.05 mM CoA, 1 mM EDTA, and soluble cell lysate. The increase in absorbance at 550nm was recorded at 30 °C after addition of the substrate α-ketoglutarate. The enzymatic activity was expressed as U per mg of total soluble protein at 30 °C, where U = conversion of 1 µmol of substrate/min.

Pyruvate ferredoxin oxidoreductase (POR) activity was measured using an aerobic *in vitro* assay similar to the KOR one. This assay measures POR activity in the oxidative direction, leading to the production of reduced ferredoxin that reduces cytochrome C, which absorbs light at 550nm. The reaction mixture contained 50 mM Tris-HCl (pH 7.5), 10mM pyruvate, 0.9 µg/ml cytochrome C, 0.05 mM CoA, 1 mM EDTA, and soluble cell lysate. The reaction was started

with the addition of pyruvate and the increase in absorbance at 550nm was recorded at 30 °C. The activity is expressed as μmol of product formed per min per mg of soluble protein (U/mg of total soluble protein). Anaerobic activity was determined by a methyl viologen-linked assay, similar to the one used to determine quinol:fumarate reductase activity. The same sealed cuvettes were used in an anaerobic chamber following a protocol previously described[24].

4.2.7 POR evaluation and evolution in E. coli JCL301/302 plateform

Evaluation of possible oxygen tolerant POR was accomplished in JCL301. Plasmid cloning for *Streptomyces coelicolor* 2-oxoacid oxidoreductase (OOR) under the control of a pTac promoter was transformed in JCL301 and followed by selection on antibiotic plates. JCL301-S.c.OOR was first grown in SOB medium overnight. Cells were spun down and washed 3 times in 1% glucose M9 to remove residual SOB. Finally they were resuspended in 3ml M9 with 1% glucose, 25mM Trimethylamine N-oxide (TMAO) as electron acceptor to the desired starting OD_{600} (0.05-0.2) and purged in an atmosphere of 5% hydrogen, 10% carbon dioxide balanced in nitrogen. Different amount of air were then injected in the hermetically capped tubes after removing the same volume of gas to keep the pressure the same.

Evolution of POR was accomplished using the new platform JCL302 under aerobic condition in minimal medium (M9) supplemented with 1% glucose, 25mM TMAO and 0.2 mM acetate which allows, in 24 hours, to reach a maximum OD_{600} of 0.3. The liquid cultures were passed down every day to a starting OD_{600} of 0.05. The same evolution procedure was continued until some culture reached in 24 hours a significant higher OD_{600}. If this result could be repeated the POR plasmid was purified from the strain and retransformation in JCL302 to check if the growth

under the same culture conditions could be repeated. If this result could be repeated, the evolution was carried on without any acetate added in the medium.

4.2.8 Quantification of the products (read-outs)

The alcohol compounds produced were quantified by an Agilent model 6850 gas chromatograph (GC) equipped with a flame ionization detector and and DB-FFAP capillary column (30 m, 0.32 mm i.d., 0.25 film thickness) from Agilent Technologies. Culture samples (1 mL) were centrifuged for 5 min at 15,000 rpm, and the supernatant was retrieved and mixed with 0.1% v⁄v 2-methyl-pentanol (100 µL) as internal standard. Helium gas was used as the carrier gas with an inlet pressure of 9.52 psi. The injector and detector were maintained at 225 °C. A 1 µL sample was injected. The GC oven temperature was initially maintained at 85 °C for 3 minutes before a temperature ramp of 45 °C/ min raised it to 235 °C where it was held for 1 min before completion of the analysis. Column flow rate was 1.7 m*l*⁄ min. Alcohols in the sample were identified and quantified by comparing to 0.001% v⁄v standard. The amount of alcohol in the samples was calculated based on the ratio of its integrated area compare to the 0.001% standard normalized by the internal standard integrated area.

Other secreted metabolites were quantified by a high-performance liquid chromatography. Culture samples (1 mL) were centrifuged for 5 min at 15,000×*g*. The supernatant was analyzed by Agilent 1200 HPLC equipped with a BioRad HPX87 column using a photodiode array detector at 210 nm absorbance. The mobile phase used was 30 mM H_2SO_4 at a constant flow rate of 0.4 mL/min. The column was maintained at 35 °C and the injection volume was 20 µL.

4.3 Results and Discussion

4.3.1 Evaluation of *Synechococcus* and *Synechocystis* native genes participating in the rGS/rTCA cycles

Cyanobacteria possess an unusual TCA cycle that lacks the enzyme α-ketoglutarate dehydrogenase. The cycle is instead completed by the action of two alternative enzymes[25]. Nevertheless, a few of the enzymatic steps required for the rGS and/or rTCA cycle to function have corresponding native enzyme in *Synechococcus* and *Synechocystis* (Fig. 4-2, A and C*)*. For those native steps, enzyme activities were tested in cell lysate to determine whether or not they would be sufficient as part of the alternative carbon fixation cycle. Overall the enzyme activity of 5 native genes was tested in *Synechococcus* and *Synechocystis*: fumarase, fumarate reductase, aconitase, isocitrate dehydrogenase, malate dehydrogenase, and succinyl-CoA synthetase. The enzymes tested showed low or no activity. Only 3 enzymes in *Synechocystis,* malate dehydrogenase, aconitase and fumarase had detectable activity using our enzyme assays (Fig. 4-2B), respectively 4.38×10^{-4} U per mg of total cellular protein, 1.43×10^{-4} U per mg of total cellular protein, and 5.36×10^{-4} U per mg of total cellular protein. Although these numbers can seem small, it is important to notice that activities are expressed per total amount of soluble protein, and in cyanobacteria proteins make up a large fraction of the biomass with for example high amounts of antenna pigments containing the phycobiliproteins that account for up to 60% of the soluble protein content[26]. These data show that out of the natives genes overlapping with the ones of our two alternative carbon fixation cycles, all will need to be overexpressed in Synechococcus, while in Synechocystis malate dehydrogenase, aconitase and fumarase might

have enough activity natively. Only fumarate reductase, isocitrate dehydrogenase and succinyl-CoA synthetase would have to be overexpressed.

4.3.2 Gene screening for enzymatic step of the rGS/rTCA cycles

Implementation of the rGS and rTCA cycles in cyanobacteria requires the overexpression of a series of enzymatics steps, a few of them being non-native to cyanobacteria. We decided to screen individually for activity all non-native rGS and rTCA genes (pyc, mtk, mcl, icl, kor, acl) as well as genes representing important carbon fixation steps (por, kor, icd, pyc) (Fig. 4-3). Each gene was cloned individually into the NSI vector backbone for integration into the *Synechococcus* genome, driven by an IPTG inducible PLlacO1 promoter, and into the psbA2 vector backbone for *Synechocystis*. All genes cloned were successfully sequenced and transformed into *Synechococcus* and *Synechocystis*. We encountered some difficulties transforming *Synechocystis*, and decided to focus on the *Synechococcus* strains for single gene evaluation. Their integration into the genome was confirmed by colony PCR. The different varieties of each gene were then tested in cyanobacteria soluble crude extract using *in vitro* enzyme assays to determine which were most active.

Activity for all the non-carboxylating enzymes assayed was found. Malate thiokinase (MTK) and malyl-CoA lyase (MCL) were tested together for the convertion of malate to glyoxylate and acetyl-CoA so that the activity of both enzymes was determined at once. In this two-step conversion malate is first converted to malyl-CoA, which is then split into acetyl-CoA and glyoxylate. We tested four combinations of Mtk and Mcl enzymes (*Methylococcus capsulatus* (Mtk), *Rhodobacter sphaeroides* (Mcl); *Ruegeria pomeroyi* (Mtk), *Rhodobacter sphaeroides* (Mcl); *Methylobacillus flagellates* (Mtk), *Rhodobacter sphaeroides* (Mcl); *Methylococcus*

capsulatus (Mtk), *Methylobacterium extorquens (Mcl)*). We found that three of our four combinations of Mtk and Mcl genes were active (Fig. 4-4). The two most active combinations contain the *M. capsulatus* Mtk and vary only in their Mcl. They were kept as candidates for the full rGS pathway integration. Isocitrate lyase (ICL) activity was mesured in the isocitrate-forming direction by a coupled assay with a commercially available isocitrate ($NADP^+$) dehydrogenase. We tested the activity of enzymes from 5 sources: *Ralstonia eutropha, Escherichia coli, Rhodobacter capsulatus, Saccharomyces cerevisiae, Idiomarina loihiensis*, all expressed in *Synechococcus*. We found that all 5 enzymes were active in whole cell lysate (Fig. 4-5), with the gene from *R. eutropha* being the most active. We decided to keep this gene for the full rGS constructs. ATP-citrate lyase (ACL) activity was measured by an assay coupling the production of oxaloacetate to the activity of malate dehydrogenase which oxidizes NADH and can therefore be easily monitored colorimetrically. We tested the activity of four different ACL enzymes from *Ralstonia eutropha*, *Chlorobium tepidum, Aspergillus terreus* and *Homosapiens sapiens* and found that two of them were active (Fig. 4-6). *C. tepidum* ACL showed stronger activity than that of *H. sapiens*. We have decided to test both active Acl in our final strains to keep a wider range of activity available as the optimum one is not known.

The rGs and rTCA cycle possess together 4 carboxylating enzymes: isocitrate dehydrogenase (ICD), α-ketoglutarate ferredoxin oxidoreductase (KOR), pyruvate ferredoxin oxidoreductase (POR) that fixe CO_2, and pyruvate carboxylase (PYC) that fixes bicarbonate[27]. Pyruvate carboxylase activity was measured by a coupled assay with malate dehydrogenase similar to the ACL assay. Enzyme activity was first assayed using whole cell lysate expressing five different PYC from *Bacillus subtilis* (B. s.), *Rhodobacter sphaeroides* (R. s.), *Staphylococcus aureus* (S. a.), *Saccharomyces cerevisiae* (S. c.) and *Listeria monocytogenes* (L. m.), however this approach

yielded only very weak signal (data not shown). Therefore the two proteins that showed promising sign of activity in the first round, the PYC from *S. cerevisiae* and *B. subtilis* were His-tagged and purified by Ni-affinity chromatography (Fig. 4-7A) before measuring activity. Both genes were found to be active (Fig. 4-7B), with specific activity of 0.085 U/mg protein for the *B. subtilis* pyc and 0.037 U/mg protein for the *S. cerevisiae* one. The most active one from *B. subtilis* was chosen for the all pathway construct. In the rTCA cycle only, the second carbon fixation step is catalyzed by an α-ketoglutarate ferredoxin oxidoreductase (KOR) which catalyzes the reductive carboxylation of succinyl-CoA to α-ketoglutarate. The activity of seven version of KOR was tested using an aerobic *in vitro* assay which measures KOR activity in the oxidative direction. Genes from *Hydrogenobacter thermophiles, Methanococcus maripaludis, Moorella thermoacetica, Thermotoga maritime, Sulfolobus acetocaldarius, Sulfolobus takadei* and *Halobacterium salinarum* were all expressed in *Synechococcus* and tested from soluble cell lysate. Unfortunately no activity was detected for any of these genes under our aerobic assay conditions. Most KOR genes are known to be O_2 sensitive and our preparation of whole cell lysate (aerobically) may have damaged the KOR enzymes. The KOR from *H. salinarum*[23] and from the Sulfolobus organisms are thought to be O_2 tolerant, these may be active *in vivo* in cyanobacteria. Therefore these three genes were kept as candidates to be evaluated *in vivo* in the context of the full rTCA cycle. Isocitrate dehydrogenase (ICD) catalyzes the third CO_2 fixation step (after PYC and KOR) for the rTCA cycle. This enzyme converts α-ketoglutarate to isocitrate utilizing either NADPH or NADH as a source of reducing power. We tested the activity of five different ICD genes expressed in *Synechococcus.* The genes chosen were from the organisms *Pseudomonas syringae, Saccharomyces cerevisiae, Vibrio fischeri, Ralstonia eutropha* and *Hydrogenobacter thermophiles*. In order to simplify the assay the enzymatic reaction was

evaluated in the oxidative direction, which releases CO_2 and produces NAD(P)H. We found activity of 0.325 U per mg of total cellular protein and 0.283 U per mg of total cellular protein for the genes from *V. fischeri* and *P. syringae* respectively. Since *Synechococcus* has a native ICD gene we also saw some activity in the wild type strain (0.996 U per mg of total cellular protein) (Fig. 4-8). The last carboxylating enzyme is pyruvate ferredoxin oxidoreductase (POR) which catalyzes the reductive carboxylation of acetyl-CoA to pyruvate in a reaction very similar to the reductive carboxylation of performed by KOR. Five version of POR from *Escherichia coli, Helicobacter pylori, Moorella thermoacetica, Desulfovibrio africanus* and *Synechococcus elongatus* were tested using an aerobic *in vitro* assay which measures POR activity in the oxidative direction. Many POR assays described in the literature use O_2 sensitive indicators such as methyl viologen and must therefore be performed anaerobically[28]. The cytochrome C assay used in this study can be performed under normal atmosphere. However, when we performed this assay on lysed *Synechococcus* cultures expressing one of the five POR genes tested we could not detect any activity. Most POR genes are known to be O_2 sensitive, even though some, such as the *Synechococcus* native gene, *nifJ*, function in O_2 generating organisms. We believe that in our preparation of whole cell lysate (aerobically) the POR enzymes might have been damaged. Therefore several of the POR genes tested here will be evaluated again in the context of the full cycle during *in vivo*. But because the carboxylating steps presented the most difficulties we decided at this point to drop the rTCA cycle and focus on the rGS which possesses only two carboxylation steps (PYC and POR) compare to four for rTCA (PYC, KOR, ICD and POR).

4.3.3 Construction of a library of plasmids for integration of multigene rGS cassettes

Integration of the full rGS cycle in cyanobacteria requires the delivery and integration in their genomic DNA of 8 to 11 heterologous genes. To avoid integrating one very large DNA fragment, we decided to split the genes into three reasonably sized portions that will be targeted to three different neutral sites in *Synechococcus* and/or *Synechocystis* genome. These sites have been described in the literature and most have been tested previously in our laboratory. We chose to split up the genes in such a way that only one of the plasmids differs between the rGS overexpressing strains. One plasmid was reserved only for POR as the activity of this step had to be tested *in vivo* and we wanted to be able to easily swap out PORs in our final strains. If POR is alone on one plasmid/in one site it can be manipulated more easily. For each site several versions of each integration plasmid were cloned (Table 4-2). Three additional native genes have to be overexpressed in *Synechococcus*, and might not be necessary in *Synechocystis* according to our native genes activity evaluation. These genes, cloning for malate dehydrogenase (MDH), fumarase (FUM) and aconitase (ACN), were all combined on the plasmid carrying the pyruvate carboxylase (PYC) and fumarate reductase (FRD) genes. We made one version of this plasmid containing the three additional genes and one without since the native level of gene expression, while undetectable with our assays, might still be sufficient for activity of the rGS. To increase the chance of success several variations of genes will be integrated into the final strains. Table 4-2 shows the genes will be varied in the final constructs: ACL, MCL, PYC and POR. We also made plasmids lacking one essential component of the rGS (MTK/MCL) to be used for making negative control strains.

4.3.4 Integration of the rGS cycle in cyanobacteria and growth evaluation of the strains

Transformation of the large multigene DNA fragments in the two cyanobacteria strains was successful only in *Synechococcus*. Trouble-shooting to increase the transformation efficiency in *Synechocystis* was unsuccessful. The completed plasmids are large in size, up to 14500bp for *Synechocystis* (~11500bp fragment integrating into the chromosome, and ~1000bp homologous regions on each side). To the best of our knowledge these are the largest plasmids that been reported to be transformed into cyanobacteria. Our transformation technique has worked well for smaller plasmid sizes, and we have confirmed transformations of a plasmid with an insert size of 7500bp into the WT *Synechocystis* strain. The problem might be due to inefficiencies of getting the large DNA fragments through the physical barrier of the cell membrane. We have attempted some alternative transformation approaches using electroporation that could help increase the porosity of the cell membrane but were unsuccessful. Using the *Synechococcus* transformation method on *Synechocystis* did not work either.

In *Synechococcus* transformation of the large multigene DNA fragments was successful. In total 45 full rGS strains carrying different variations of the genes were engineered (Table 4-3). The genotype of all these strains was verified by PCR. A first round of testing based on growth rate to identify strains with improved properties was performed with cells cultured in glass test tubes. Because of the large number of samples, these tests had to be performed in batches, which was not optimal as it may have contributed to some variability in the results. The growth curves were measured under regular (high) light conditions. The optimum light condition and CO_2 concentration to favor rGS mediated growth improvements was unknown. We chose a high light / high CO_2 condition for our initial test based on the rationale that excess reducing power and light can damage cyanobacteria, and an additional carbon fixation cycle may provide a sink for

the excess reducing power, allowing the engineered strains to accumulate less damage and grow better. These preliminary sets of testing (data not shown), although not performed on every strain, gave us some important information. For the rGS strains, the control strains (containing ACL, ICL, PYC, FRD and POR, but missing MTK, MCL) grew more poorly than many of the full strains. Also some strains expressing the *R. sphaeroides* Mcl died; however, the strains expressing the *M. extorquens* Mcl gene were able to grow at a rate similar to wild type. It remained unclear why the control strains grew so poorly, two scenarios could account for this observation, in the ideal case where the pathway would be fully functional: 1) Expressing ACL, ICL, PYC, FRD and POR poses a very high metabolic burden on the cell, due to protein expression or enzyme activity, which causes growth defects. Expression of the full rGS pathway rescues this growth defect by either increasing carbon fixation rate or utilizing toxic intermediates. 2) The NSI plasmid of the negative control strains somehow is more toxic to the cell than that of the full (*M.e. mcl*) rGS strains. The full rGS strain genes pose less of a burden on the cell, however this lack of toxicity is unrelated to the activity of the rGS.

Even though using test tubes as a method for evaluation of growth gave us some preliminary ideas on a few strains behavior, it introduced a lot of variability and was not particularly high throughput. To date no effective method for high through-put growth of cyanobacteria has been reported. To establish such a method that reliably allows for comparison of growth rate of the rGS strains, we tested photosynthetically growth in 96-well plates. Some issues made this task challenging and were addressed as follow a) To ensure that all samples receive equal amount of light, the side-wells in a clear 96-well plate were not use for actual samples, b) parafilmed hard plastic lids were used to avoid evaporation, c) condensation was preventing by elevating the plate on a rack allowing airflow on all sides, and thus equalizing the temperature above and

below the 96-well plate, d) adequate agitation was necessary to prevent flocculation. The small size of the wells demands a good agitation in the limit of not causing the samples to spill. The growth method developed and used is described under materials and methods. A consistent growth up to OD_{730}= 1 was observed, after which the cells started to flocculate. This growth up to this critical point was enough to record reliable rates of exponential growth. The reproducibility of this method was tested by growing WT *Synecococcus* in multiple wells and on multiple plates, and comparing the rates. The growth was very reproducible across all wells of a plate (except the outermost wells, which were not included) (Fig. 4-9). This method could further be used to test growth rates of the many different rGS strains constructed. Each of them was grown at least twice on each plate and on at least two plates under the same conditions to get accurate, statistically significant measures of growth rate. The growth rate of each tested strain and their corresponding control strain was determined (Fig. 4-10). A number of control strains showed growth rates significantly higher than when they were first tested in glass tubes. The most probable explanation is that these strains evolved to grow better and are now able to grow almost as well as the full rGS strains. This would have happened during a long decontamination process after a fungal contamination of a number of strains. They had to be cultivated for multiple rounds, both in liquid cultures and on plates, to get rid of the fungus. Regardless of the control strains behavior, the full strains do not show an increase growth phenotype compare to WT (Fig. 4-10). Therefore, even though this method allows for high through-put growth test of cyanobacteria, it is not appropriate in this case. If the full rGS pathway is working, it is still possible that the extra carbon fixed does not go into biomass but is consumed in another ways. Before testing this hypothesis we decided to make sure that all enzyme of the cycle are active in

the new constructs. To do so, series of enzyme assays for each step were performed (see next paragraph).

4.3.5 Evaluation of enzymatic activities in the first generation of full rGS strains

Evaluation of the growth rate of a large number of cyanobacterial strains with versions of the rGS cycle did not show significant growth improvement over wild type. To determine which steps could be improved, assessment of the enzyme activities of the rGS strains was needed. Each enzyme had been tested for activity individually but only prior to construction of the multigene plasmids and the full rGS strains. A few representative versions of the final strains expressing different version of the multigene plasmids (Fig. 4-11) were picked to test the activity of each rGS enzyme and see which, if any, were limiting. We found that activity of some key enzymes was very low or undetectable in cyanobacterial cell lysate. Results are summarized in table 4-4. Icl was only active in the negative control strain transformed with pLG125 but not in the full rGS strain tested with pLG128 integrated. This could be attributed to some mutations occuring only in certain strains during integration of the construct into the cyanobacterial genome. PYC showed no activity in all the strains tested. We have had similar problems detecting pyc activity in our first individual screening since activity levels are always low, and ended up His-tagging this enzyme. FRD had no detectable activity, although it might be a result of the enzyme assay used. The frd gene we used has an activity that depends on membrane bound electron carriers and was therefore difficult to assay *in vitro*. MDH and FUM were found to be active. Acn showed some activity but lower than the one measured in WT. Mtk/mcl had activity in the full strain tested and not in the negative control ones like expected. As a second mean of evaluation we also measured the rGS genes activity in E. coli XL1Blue lysate of cells

expressing the rGS plasmids. Gene expression is generally higher *E. coli* lysate. Like in cyanobacteria some of the enzymes were much less active than expected from previous screening experiments. This set of experiments allowed us to notice a correlation between enzyme activity of heterologously expressed genes from the same constructs in *E.coli* and cyanobacteria. In general, rGS enzymes active in *Synechococcus* were also in *E.coli* and unactive ones in *Synechococcus* were not working in *E.coli* either. From here forward, we tested enzymes activity in *E.coli* first as an evaluation and screening platform. Overexpressed protein levels are often higher in *E.coli* lysate and measuring enzyme activity is easier as when using cyanobacteria soluble crude extract.

4.3.6 Construction of the second generation of full rGS strains

Evaluation of enzymes activity in our first generation of full rGS strains allowed us to identify problematic steps that showed low or now activity. New versions of the NSI and NSII rGS plasmids were constructed with new variants of the genes in question (acl, frd, pyc and acn). The acl from *H. sapiens* was replaced by the one from *C.tepidum* which already showed higher activity during the first round of screening. A new soluble frd from *Trypanosoma brucei* that uses NADH as a reducing agent was cloned to replace the *E. coli* frd which is membrane bound and uses quinone pool as electron carrier. The change was thought to help with the cofactor balance of the pathway and assessment of the new frd by enzyme assay was easier. Synechococcus *acn* gene was swapped for the *E. coli* equivalent as it was found more active in other ongoing projects in our laboratory (data not shown). Lastely, PYC had undetectable activity in our final strains as well as when the rGS plasmids were expressed in *E.coli*. Similar problems had been faced in the past during the individual gene screening since activity levels of

this protein were already low. The new rGS constructs were designed with a HIS-tag added to the pyc-coding gene. This allowed for purification and concentration of the protein before evaluation of its activity. A schematic representation of the new rGS constructs is shown in figure 4-12 and the enzymatic activity of each step was confirmed in *E.coli* lysate.

Additionally our lab developed a computer model of the rGS which allowed identifying a potential rate imbalance. According to the simulation the irreversible step catalyzed by mtk/mcl could lead to imbalances that would prevent the cycle from functioning[29]. At this branchpoint malate is the substrate for two different reactions, the production of glyoxylate and the production of succinate. These two compounds then recombined to form isocitrate, and thus must be produced in equal amounts. Fumarase and fumarate reductase are reversible enzymes, so if too much succinate is being made its excess pool can drive the reaction in the reverse direction. However, mtk/mcl is not reversible. Anticipating the possibility of this rate imbalance, we decided to build an additional variation of the cycle that involves a malate synthase enzyme that performs the reverse reaction of mtk/mcl (Fig. 4-13). This, in effect, makes the mtk/mcl step reversible allowing for a better balancing of the reaction rates of the cycle. The malate synthase from *E. coli* (*aceB*), was choosen and cloned in the neutral site III integration plasmid with the por genes.

4.3.7 Bioprospecting for a POR active under aerobic conditions

One major challenge of implementing the rGS cycle was to find a pyruvate:ferredoxin oxidoreductase (POR) that is actively expressed in a photosynthetic organism (when grown in light conditions) and that is active in the presence of oxygen. Five version of POR from *Escherichia coli, Helicobacter pylori, Moorella thermoacetica, Desulfovibrio africanus* and

Synechococcus elongatus were already tested using an aerobic *in vitro* assay and no activity was detected tested (see section 4.3.2 Gene screening for enzymatic step of the rGS/rTCA cycles). Some native PORs from cyanobacteria have been shown to be expressed under aerobic conditions[30]. We decided to test again the *Synechococcus* POR as well as the *Synechocystis* one both expressed in neutral site III (NSIII). Their activity could successfully be assayed under anaerobic conditions but was not observable under aerobic conditions (data not shown). The next promising candidate genes tested were from O_2 tolerant organisms (*Hydrogenobacter thermophilus* and *Halobacterium salinarum*). The first is an aerobic hyperthermophilic autotrophs, while *Halobacterium salinarum*[23] is an extremely halophilic obligate aerobic archaeon. Our lab had previously confirmed activity in the decarboxylating direction of the *Halobacterium salinarum* POR under aerobic conditions when natively expressed. The downside is that this protein is adapted to very high intracellular salt concentrations, and works optimally in solution with 3M KCl. When the enzyme assay was performed in a buffer not containing KCl or other salts, 39% of its maximal activity was retained. Several attempts of expressing the protein in *E. coli* were unsuccessful at producing an active recombinant enzyme. One possible problem is the very high content of acidic residues of the native protein, which is a common protein adaptation aimed at increasing its solubility in highly saline environments[31]. Cyanobacteria and *E. coli* are distant organisms and the unsuccessful expression of a protein in one does not exclude success in the other. The three new POR candidates were cloned and transformed in *Synechococcus* to test their expression and activity in the cyanobacterial strain using the O_2 tolerant assay. The cells were prepared as described in Materials and Methods with some adjustments: cells were harvested and lysed under anaerobic conditions, and the assay was also set up and initiated anaerobically. Only after some activity was detected the reaction cuvette

was opened up to air and the reaction was allowed to proceed for at least another hour. No decrease in activity was noticeable. This seemed to indicate that *H. thermophilus* and *H. salinarum* PORs are actively expressed in *S. elongatus* under photosynthetic conditions, and are able to function in the presence of O_2. Attempts to repeat this experiment failed at detecting any POR activity even under anaerobic conditions. Low protein expression or unstable expression might have been the cause. Beside the aerobic osmotolerant *Halobacterium*, a few aerobic bacteria have been reported to possess a POR or POR-like protein that may be less O_2-sensitive[30]. The oxoacid oxidoreductase (OOR) from *Streptomyces coelicolor*, a soil-dwelling obligate aerobe bacterium, is one of them. As no POR for *S. coelicolor* is annotated, we blasted his genome against the closest annotated PFOR like protein (from Frankia sp. EUlK1) to identify the coding sequence to be cloned. Testing of this enzyme was performed using a selection platform in *E.coli* (JCL302 or JCL301) as described in Materials and Methods. Using such a platform allows to test for the true activity of the enzyme *in vivo*. We suspect that the candidate PORs might be somewhat tolerant to O_2 but to a level probably lower than the atmospheric oxygen concentration. Preparation of cell lysate for enzyme assay *in vitro* could damage the proteins and affect their activity. *E. coli* JCL301 and 302 are entirely blocked in their ability to convert pyruvate into acetyl-CoA or acetate, resulting in acetate auxotrophy during growth on glucose in minimal medium. Without addition of acetate JCL301 and 302 can be used as selection platforms for a foreign acetyl-CoA synthesis pathway from pyruvate, like POR. JCL301 was used to test *S. coelicolor* OOR. The strain transformed with a plasmid coding *S. coelicolor* OOR was grown in M9 1% glucose, 25mM TMANO under low O_2 concentration. To do so, the cultures were first placed in an atmosphere of 5% hydrogen, 10% carbon dioxide balanced in nitrogen before being hermetically capped. Different amount of air were then

injected in the tubes. No growth was observed for the control (JCL301) for all conditions tested like expected. S. *coelicolor* OOR could rescue JCL301 and cultures with a higher air volume injected showed better growth (Fig. 4-14 A). These results would have indicated the possibility of a POR-like protein capable to perform the decarboxylation reaction of pyruvate to acetyl-CoA in the presence of oxygen. Unfortunately, after further evaluation of the culture conditions we discovered that the rescued growth was due to the antibiotic solution used in the experiment (Fig. 4-14B). The OOR was cloned in a plasmid backbone containing a chloramphenicol cassette for selection. Chloramphenicol is poorly soluble in water and is commonly desolved in pure ethanol. When the experiment was tested with a water soluble chloramphenicol, no growth rescue of JCL301 could be observed. The ethanol in the culture could have act as a precursor for acetyl-CoA, thus rescuing the growth of the acetate auxotrophe. Further attempt to evolve S. coelicolor OOR in JCL302 under aerobic condition like described in Materials and Methods was unsuccessful. After 16 rounds of evolution, one strain reached higher OD_{600} (1.5) after 24 hours under the tested conditions (aerobic, M9, 1% glucose, no acetate). Purification of the OOR plasmid from this strain and retransformation in JCL302 did not show any growth under the same culture conditions. A mutation in the strain genome not directly related to our platform is probably the cause of the growth first observed.

4.3.8 Evaluation of the Linear rGS (LrGS) through acetate production

Because no active POR could be identified under our conditions, we decided to test the remaining part of the rGS pathway through production of acetate. RGS without the POR step from acetyl-CoA to pyruvate is not a carbon fixation cycle anymore, but still allows to fix on extra carbon while producing acetyl-CoA. Acetyl-CoA is a precursor of acetate through the

activity of an acetyl-CoA ligase ADP-forming (ACD), and acetate can serve as a read-out for the functionality of the upstream pathway (Fig. 4-15). The NSI and NSII plasmids (pXL55 and pFD210) integrated in the acetate producing *Synechococcus elongatus* strain (termed LrGS+acd) were tested for acetate production and compared to WT *Synechococcus* expressing only *acd* without the LrGS genes, as well as to the LrGS control strain lacking mtk/mcl (Fig. 4-16). The genes encoded by these two plasmids are: in NSI, *H. sapiens acl, R. eutropha icl, M. capsulatus mtkCD* and *M. estorquen mcl, and in NSII, E. coli frd, B. subtilis pyc, E. coli acn, E. coli mdh* and *S. elongatus fumC*. In NSIII, *acd* from *Pyrococcus furiosus* is expressed, as well as *aceB*. Cells were cultured in shake flasks, fed bicarbonate daily and induced with 1mM IPTG when OD_{730} reached about 0.5. Samples of 1 ml were taken at day 4 after induction (based on preliminary data). HPLC analysis was then conducted on the samples as described in Materials and Methods. The LrGS+acd strain produced less acetate (1.182mM/OD_{730}) that its control lacking mtk/mcl (1.411mM/OD_{730}), and less than WT+acd (1.453mM/OD_{730}). These results show that acetate production is independent of the full LrGS. No acetate was produced by LrGS without *acd* like expected. We did notice some small fumarate production in all LrGS strains (full and control) compare to WT. Moreover the control strain did show more fumarate production (0.057mM/OD_{730}), compare to LrGS+acd (0.007mM/OD_{730}), or LrGS alone (0.011mM/OD_{730}). This could indicate that the left branch of the rGS, from pyruvate to oxaloacetate, malate and then fumarate, had some kind of activity. The highest fumarate concentration in the control strain could be explained by the absence of mtk/mcl. In LrGs, mtk/mcl drained some of the malate away to acetyl-CoA and glyoxylate, leaving less of its pool available to form fumarate. This would also mean that mtk/mcl is indeed active in the LrGs strain.

4.4 Conclusion

In this study we aimed at developing and optimizing a rewired CO_2 fixation pathway in cyanobacteria, with the purpose, if successful, to later implement it in plants. Cyanobacteria have been used for decades as a biologically simplistic model system for phototrophic eukaryotes. They also have the advantage to grow faster than plant, making them more suited for a first stage evaluation of our alternative CO_2 fixation pathway. Our two candidate pathways were the reductive tricarboxylic acid (rTCA) cycle found in some microorganisms, and a novel synthetic reverse glyoxylate shunt (rGS) cycle, which has been partially demonstrated in *E. coli* in our laboratory[8]. Both cycles could theoretically improve carbon fixation efficiency and channel the carbon flux to pyruvate, a direct precursor of biofuels like branched chain higher alcohols. They fixe CO_2 with only 30% of the ATP requirement compared to the native Calvin cycle and do not utilize Rubisco for carbon fixation, thus avoiding the photorespiration associated carbon loss.

After a first screening of at least four variations of each gene participating in the rTCA and/or rGS cycle, we decided to drop the rTCA pathway and focus on the rGS. This last one possesses only two carboxylation steps (PYC and POR) compare to four for rTCA (PYC, KOR, ICD and POR). Our *in vitro* assay screening, performed in *Synechococcus* lysate expressing the heterologous genes, showed that the carboxylation steps were the more challenging. Only PYC showed some detectable activity when his-purified, and all ICD genes tested could only be evaluated in the decarboxylation direction. No activity could was detected for POR and KOR in our aerobic conditions. Moreover, the rTCA cycle has recently been predicted thermodynamically unfavorable, at least in *E. coli*[32]. It is expected to be infeasible precisely because of the sequential reaction of KOR and ICD. Both enzymes are already moderately unfavorable on their own, and the sequential operation of the two produces an energetic barrier

that cannot be overcome even when consider the most extreme substrates concentrations within the physiological range in *E. coli*. The authors still pointed out that the reductive TCA cycle operates in other organisms where the cytosolic conditions are significantly different, like a lower pH (pH< 7), small CoA concentration ([CoA] << 1mM) and/or very high CO_2 concentration due to carbon concentrating mechanisms. *Synechococcus* physiological pH is generally comprised between pH 7.0 to 9.0[33] and its carbon concentration mechanism raises CO_2 concentration mainly in the carboxysomes around Rubisco, not in the cytosol. There are relatively few reports available on the intracellular pool of free CoA, but it seems that based on the available informations, the rTCA cycle would be unfavorable in *Synechococcus* like in *E. coli*.

The integration of rGS requires expression in cyanobacteria of 8 to 11 heterologous genes. We decided to split them up on three plasmids for recombination in three different neutral sites to minimize the size of the DNA fragment integrated. This still represents to our knowledge the largest plasmids reported to be integrated into cyanobacteria. Transformation was only successful in the strain *Synechococcus elongates* PCC 7942, but failed in the second strain we tested, *Synechocystis* sp. PCC 6803.

The next big challenge we faced was to find an active O_2-resistant POR, the last carboxylating enzyme of the cycle. POR proteins are known to be highly sensitive to oxygen[34], and cyanobacteria generate O_2 as a side product of photosynthesis. We screened 9 PORs mostly from organisms that tolerate some level of oxygen, but were unsuccessful. A platform in *E. coli* was further developed to evolve these PORs. Despite our effort no oxygen resistant enzyme was found, which led us to take a new direction with our project. Without the last step, the rGS cycle become a linear pathway, termed Linear reverse Glyoxylate Shunt (LrGS), that still allows to fix

carbon through PYC on top of the CBB cycle to produce acetyl-CoA at a lower energy cost than through the CBB cycle alone. An acetyl-CoA ligase (ADP-forming) was expressed in *Synechococcus* on top of the LrGS genes to produce acetate from acetyl-CoA. Production of acetate could then be used as a read-out to evaluate the activity of the pathway. No increase in acetate production was detected in LrGS compare to the control strains. However, a small concentration of fumarate, undetectable in WT, was present in all LrGS strains. This could indicate a small carbon flux through part of the rGS, from pyruvate to fumarate. The fumarate detected was also significantly higher, although still in small concentration, in the control strain without MTK/MCL. This trend could be explained by active MTK/MCL enzymes in the full strain, which, absent in the control, do not drain malate away to acetyl-CoA and glyoxylate, leaving its pool available to form more fumarate. The LrGS may thus be partially active. The fact that it shares enzymatic steps with the native TCA cycle of cyanobacteria, and thus interacts directly with the central metabolic pathways may explain its low and partial activity. Additionally the identification of a possible kinetic trap at the MTK/MCL and FUM junction might also prevent the cycle from running. An additional enzyme has been added to prevent the potential rate imbalance, but the exact rate equilibrium at that junction would be difficult to assess *in vivo*. Further improvement of the LrGs will require evaluating one by one the step of the pathway to identify any rate limiting step. Moreover, an alternative pathway to recycle glyoxylate might be necessary if the flux past the MTK/MCL point cannot be detected.

4.5 Tables and Figures

Enzymatic step	Gene	Organism of origin
Isocitrate lyase (Icl)	*iclA*	*Ralstonia eutropha*
	aceA	*Escherichia coli*
	aceA	*Rhodobacter capsulatus*
	Icl1	*Saccharomyces cerevisiae*
	aceA	*Idiomarina loihiensis*
Pyruvate carboxylase (Pyc)	*pyc*	*Bacillus subtilis*
	pyc	*Rhodobacter sphaeroides*
	pyc	*Staphylococcus aureus*
	pyc	*Saccharomyces cerevisiae*
	pyc	*Listeria monocytogenes*
ATP-citrate lyase (Acl)	*acl*	*Ralstonia eutropha*
	aclAB	*Chlorobium tepidum*
	acl	*Aspergillus terreus*
	acl	*Homosapiens sapiens*
Malate thiokinase (Mtk) and Malyl-CoA lyase (Mcl)	*sucCD/ mclA*	*Methylococcus capsulatus (Mtk)/ Rhodobacter sphaeroides (Mcl)*
	mtk/ mclA	*Ruegeria pomeroyi (Mtk)/ Rhodobacter sphaeroides (Mcl)*
	mtk/ mclA	*Methylobacillus flagellates (Mtk)/ Rhodobacter sphaeroides (Mcl))*
	sucCD/ mcl	*Methylococcus capsulatus (Mtk)/ Methylobacterium extorquens (Mcl)*
Pyruvate – ferredoxin oxidoreductase (Pfor)	*por*	*Escherichia coli*
	por	*Helicobacter pylori*
	por	*Moorella thermoacetica*
	por	*Desulfovibrio africanus*
	nifJ	*Synechococcus elongatus*
Isocitrate dehydrogenase (Icd)	*icd*	*Pseudomonas syringae*
	icd	*Sacheromyces cerevisiae*
	icd	*Vibrio fischeri*
	icd	*Ralstonia eutropha*
α-ketoglutarate - ferredoxin oxidoreductase (Kor)	*kor*	*Hydrogenobacter thermophiles*
	kor	*Methanococcus maripaludis*
	kor	*Moorella thermoacetica*
	kor	*Thermotoga maritime*
	kor	*Sulfolobus acetocaldarius*
	kor	*Sulfolobus takadei*
	kor	*Halobacterium salinarum*

Table 4-1: The non-native genes of the rGS were cloned from a variety of organisms (column 3) into a NSI vector and transformed into *S. elongates* genome. Successful integration into neutral site 1 was confirmed by colony PCR.

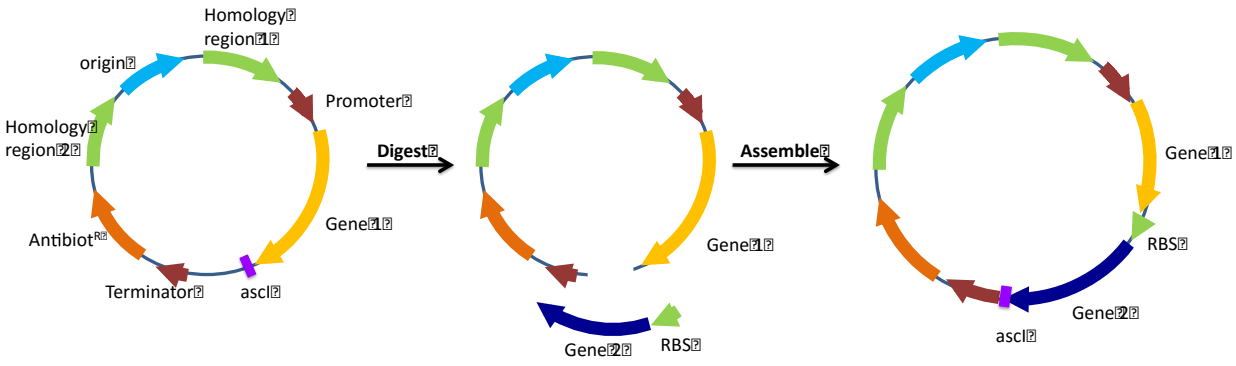


Figure 4-1. Assembly of large vectors by iterative gene insertion (restriction digestion of vector by ascI and then assembly with a PCR amplified insert). Each successive round of assembly regenerates an ascI site after the inserted gene while destroying the one in front.

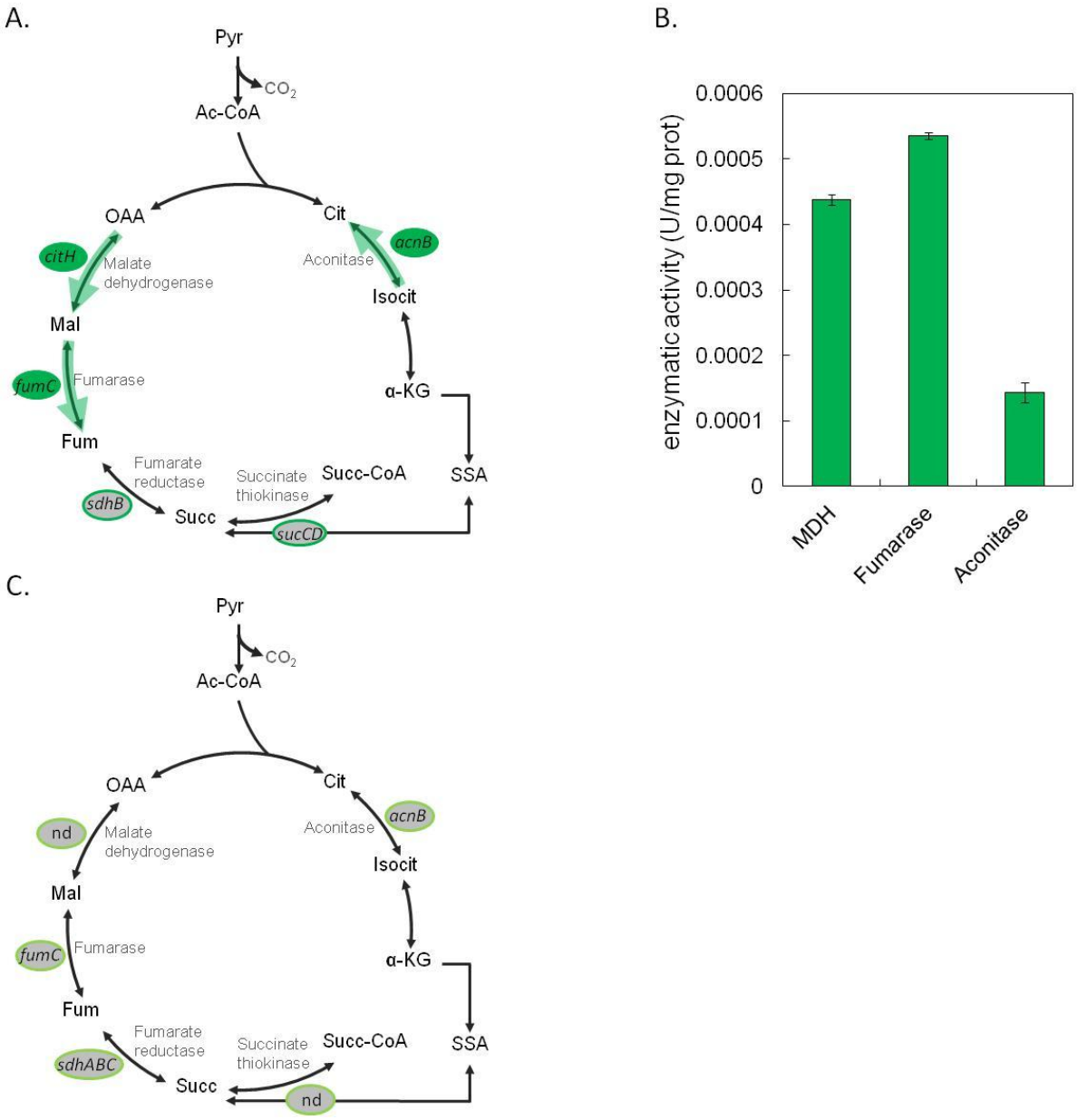


Figure 4-2: (A) Schematic representation of the tricarboxylic acid (TCA) cycle in *Synechocystis* sp. PCC 6803. For clarity cofactors have been omitted. Activity of some native genes (circled in ovals) was tested in lysates of wild type as described in Material and Methods. Enzymes that were active to a sufficient level are highlighted in green; those that were only minimally active, or had no activity are in grey. For the active enzymes, a green arrow indicates the direction in

which the activity was tested. (B) Enzymatic activity of the 3 enzymes found significantly active in *Synechocystis.* MDH (malate dehydrogenase) activity was found to be 4.38×10^{-4} U per mg of total cellular protein, fumarase activity 5.36×10^{-4} U per mg of total cellular protein and aconitase activity 1.43×10^{-4} U per mg of total cellular protein. Error represents the standard deviation between n = 3 cultures. (C) Schematic summary of the tricarboxylic acid (TCA) cycle in *Synechococcus elongatus* PCC 7942. For clarity cofactors have been omitted. Activity of some native genes (circled in ovals) was tested in lysates of wild type as described in Material and Methods. No enzymatic activity could be detected in our conditions for all the enzymes tested.

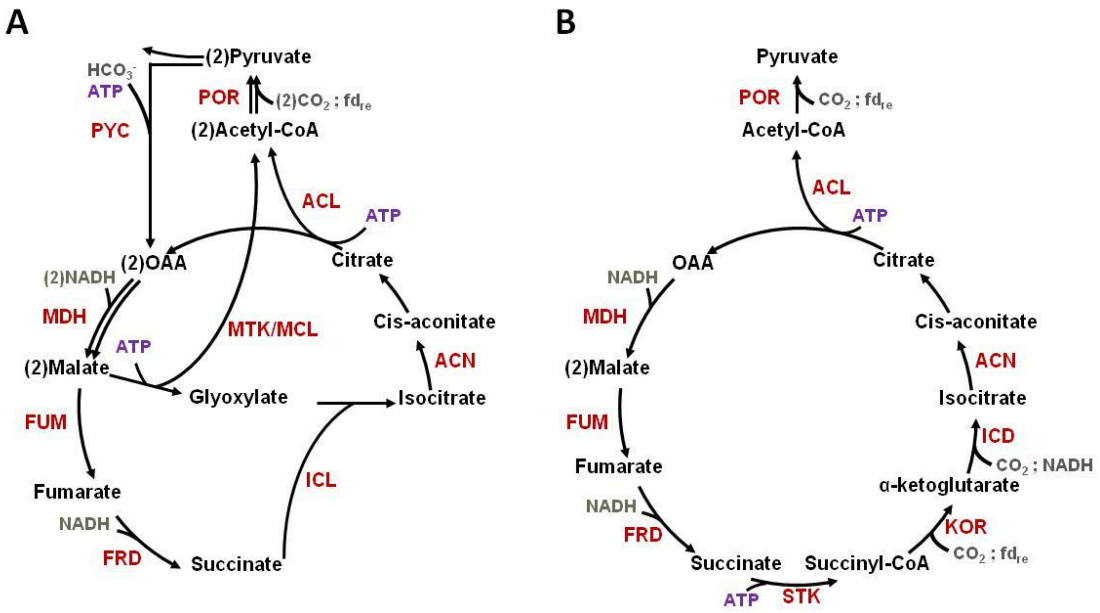


Figure 4-3. Schematic representation of the reverse Glyoxylate Shunt (rGS) cycle (A) and of the reductive tricarboxylic acid (rTCA) cycle (B). Both pathways produce one molecule of pyruvate using energy (ATP) and reducing power (NADH, NADH equivalent and reduced ferredoxin) from photosynthesis. PYC, pyruvate carboxylase; MDH, malate dehydrogenase; ACN, aconitase; FUM, fumarase; FRD, fumarate reductase; MTK, malate thiokinase; MCL, malyl-CoA lyase; ICL, isocitrate lyase; ACL, acetyl-CoA lyase; POR, pyruvate ferredoxine oxidoreductase, STK, succinate thiokinase; KOR, α-ketoglutarate oxidoreductase; ICD, isocitrate dehydrogenase, fd_{re}, reduced ferredoxin.

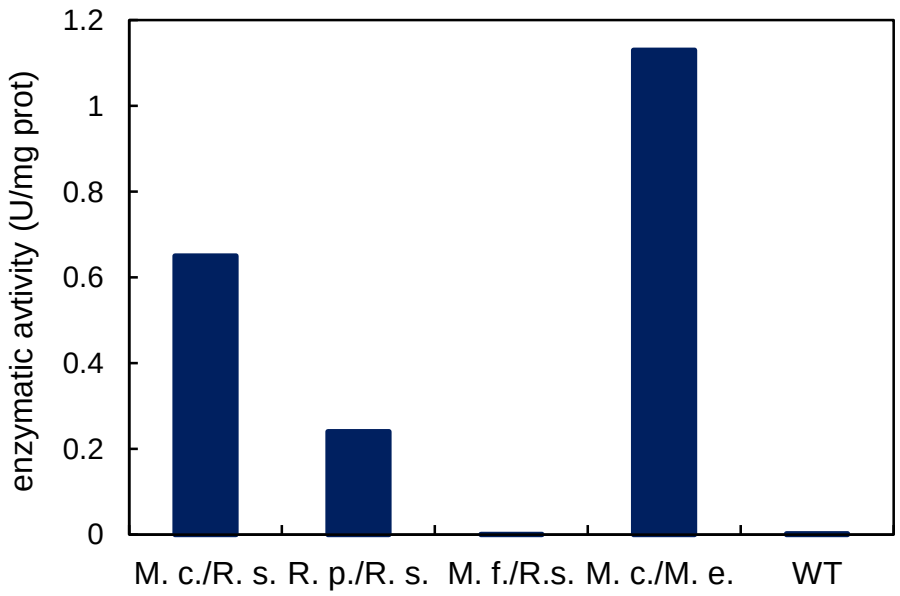


Figure 4-4. Comparison of the activity level of the expressed malate thiokinase (MTK) and malyl-CoA lyase (MCL) combinations in *Synechococcus*. The combination *Methylobacillus flagellates* (*M.f.*), *Rhodobacter sphaeroides* (*R.s.*) was found to have no activity. Activity was found to be 0.65 U per mg of total cellular protein for the combination *Methylococcus capsulatus* (*M. c.*), *Rhodobacter sphaeroides* (*R.s.*), 0.24 U per mg of total cellular protein for *Ruegeria pomeroyi* (*R. p.*) et *Rhodobacter sphaeroides* (*R. s.*), and 1.13 U per mg of total cellular protein for the combination *Methylococcus capsulatus* (*M. c.*), *Methylobacterium extorquens (M. e.*). The two constructs with *M. capsulatus* Mtk showed higher activity.

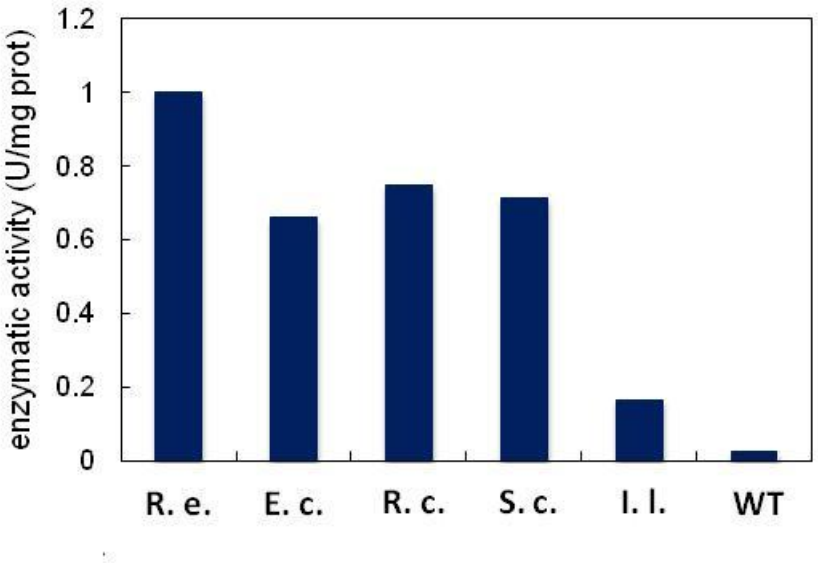


Figure 4-5. Activity of whole cell lysate of *Synechococcus* expressing isocitrate lyase genes from different organisms: *Ralstonia eutroph* (*R. e.*), *Escherichia coli* (*E. c.*), *Rhodobacter capsulatus* (*R. c.*), *Saccharomyces cerevisiae* (*S. c.*), *Idiomarina loihiensis* (*I. l.*). All genes tested showed activity. The highest activity of 0.99 U per mg of total cellular protein was found for the *R. eutropha* ICL.

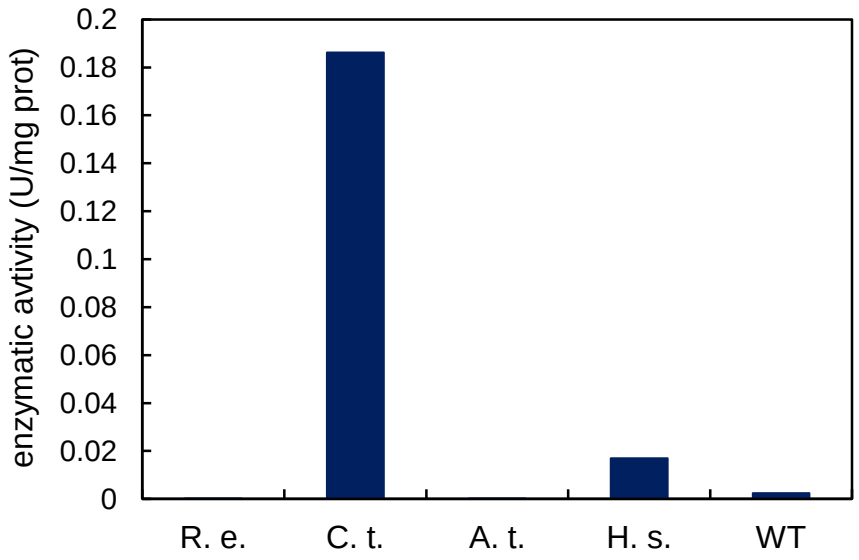
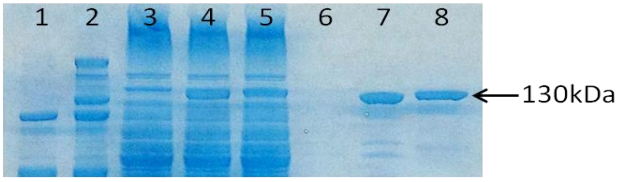


Figure 4-6. Activity of the ATP-citrate lyase (ACL) enzymes that were expressed and assayed in *Synechococcus. Chlorobium tepidum* and *Homo sapiens* were both active, but at different levels. *C. tepidum* ACL showed stronger activity (0.186 U per mg of total cellular protein) than that of *H. sapiens* (0.017 U per mg of total cellular protein).

A.



B.

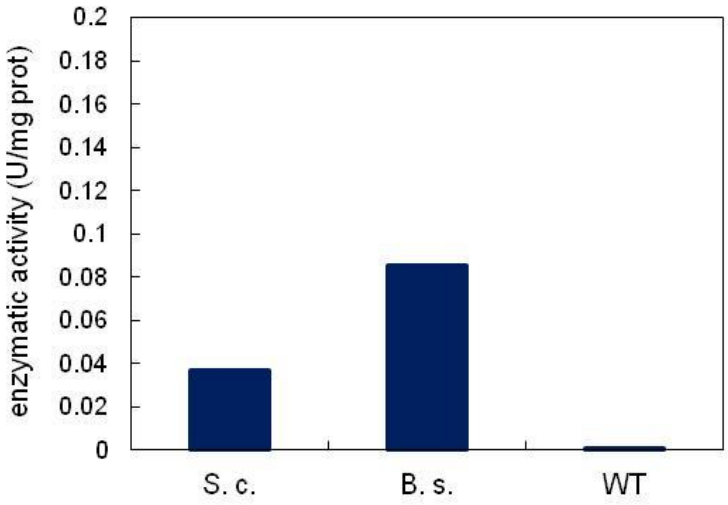


Figure 4-7. Pyruvate carboxylase (PYC) activity assay. His-tagged proteins were purified by Ni-affinity chromatography. (A) Gel shows whole cell lysate and purified proteins. Columns indicate: 1-Low molecular weight ladder; 2-High molecular weight ladder; 3-WT cell lysate; 4-Cell lysate of *Synechococcus* expressing *S. c.* PYC; 5-Cell lysate of *Synechococcus* expressing *B. s.* PYC; 6-WT elute after protein purification; 7-Purified *S. c.* PYC; 8-Purified *B. s.* PYC. (B) Bar graph shows activity of purified PYC from *Saccharomyces cerevisiae* (S. c.) 0.037 U/mg protein and *Bacillus subtilis* (B. s.) 0.085 U/mg protein.

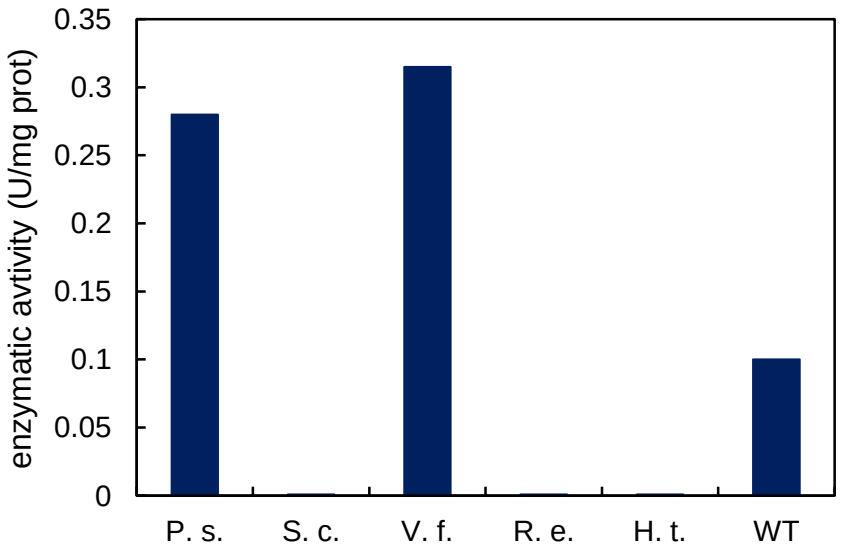


Figure 4-8. Activity of isocitrate dehydrogenase (ICD) in whole cell lysate of *Synechococcus* expressing the genes indicated: *Pseudomonas syringae* (*P. s.*), *Saccharomyces cerevisiae* (*S. c.*), *Vibrio fischeri* (*V. f.*), *Ralstonia eutropha* (*R. e.*) and *Hydrogenobacter thermophiles* (*H. t.*), as well as the native activity in WT *Synechococcus* lysate (WT). Two heterologous genes showed activity, 0.283 U per mg of total cellular protein for *Pseudomonas syringae* ICD and 0.325 U per mg of total cellular protein for *Vibrio fischeri* ICD. *Synechococcus* has a native ICD gene which explains the activity found in the WT strain lysate (0.996 U per mg of total cellular protein).

Synechocystis sp. PCC6803		
Integration site	**rGS enzymes**	**Resistance**
PsbA2	ACL (x2), ICL, MTK, MCL (x2)	Spec
Negative control	ACL, ICL	Spec
slr0168	PYC (x2), FRD	Kan
ssl0410	POR (x7)	Cam
Synechococcus elongatus pcc7942		
Integration site	**rGS enzymes**	**Resistance**
NSI	ACL, ICL, MTK, MCL (x2)	Spec
Negative control	ACL, ICL	Spec
NSII	PYC (x2), FRD (MDH, FUM, ACN)	Kan
NSIII	POR (x7)	Cam and Amp

Table 4-2. List of enzymes expressed in the constructs used to integrate rGS into three neutral sites of *Synechocystis* or *Synechococcus.* Negative control strains were also constructed by omitting one essential component of the rGS (MTK/MCL) in the psbA2 and NSI plasmids. Genes coding for some enzymes were varied and the number of variations is indicated in brackets. The origin of each gene coding for the rGS enzymes were: ACL: *H. sapiens* and *C. tepidum* (*Synechocystis* only); ICL: *R. eutropha*; MTK: *M. capsulatus*; MCL: *M. extorquens* and *R. sphaeroides*; PYC: *B. subtilis* and *S. cerevisiae*; FRD: *E. coli*; MDH: *E. coli*; ACN: *S.elongatus*; FUM: *S.elongatus*; POR: *H. pylori*, *M. thermoacetica*, *D. africanus*, *T. denticola*, *T. maritime*, *H. salinarum* and *S. acetocaldarius.*

Strain	Genes in NSI site	NSII	NSIII
180	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S.7942*)
181	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*S.7942*)
182	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S.7942*)
186	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S.6803*)
187	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*S.6803*)
188	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S.6803*)
192	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*H. pylori*)
193	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*H. pylori*)
194	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*H. pylori*)
198	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*M. thermoacetica*)
199	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*M. thermoacetica*)
200	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*M. thermoacetica*)
204	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*D. africanus*)
205	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*D. africanus*)
206	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*D. africanus*)
210	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*T. denticola*)
211	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*T. denticola*)
212	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*T. denticola*)
216	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*T. maritima*)
217	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*T. maritima*)
218	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*T. maritima*)
231	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*H. salinarum)*
232	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*H. salinarum)*
233	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*H. salinarum)*
237	Acl (*H. sapiens*); Icl (*R. eutropha*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S. acetocaldarius)*
238	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*)	POR (*S. acetocaldarius)*
239	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (*E. coli*); Pyc (*S. cerevisiae*)	POR (*S. acetocaldarius)*
243	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*S.elongatus*)
244	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) /* Mcl *(R. sphaeroides*)	Frd (E. coli); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*S.elongatus*)

247	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (S.7942); Mdh (*E. coli*); FumC (S. *elongatus*)	POR (*S.6803*)
248	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*S.6803*)
251	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*H. pylori*)
252	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*H. pylori*)
255	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*M. thermoacetica*)
256	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*M. thermoacetica*)
259	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*D. africanus*)
260	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*D. africanus*)
263	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*T. denticola*)
264	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*T. denticola*)
267	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*T. maritima*)
268	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*T. maritima*)
271	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*H. salinarum)*
272	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*H. salinarum)*
275	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (M. extorquens*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*S. acetocaldarius)*
276	Acl (*H. sapiens*); Icl (*R. eutropha*); MtkCD (*M. capsulatus) / Mcl (R. sphaeroides*)	Frd (*E. coli*); Pyc (*B. subtilis*); Acn (*S.7942*); Mdh (*E. coli*); FumC (*S. elongatus*)	POR (*S. acetocaldarius)*

Table 4-3. All rGS strains that have been constructed are listed by strain number. Columns 2-4 detail which version of each gene is found in neutral site I, II or III (NSI, NSII, NSIII). Strains

highlighted in white are control strains missing key genes of the rGS cycle (*mtkCD* and *mcl*). Strains highlighted in dark grey contain three additional genes in NSII (*acn*, *mdh* and *fumC*).

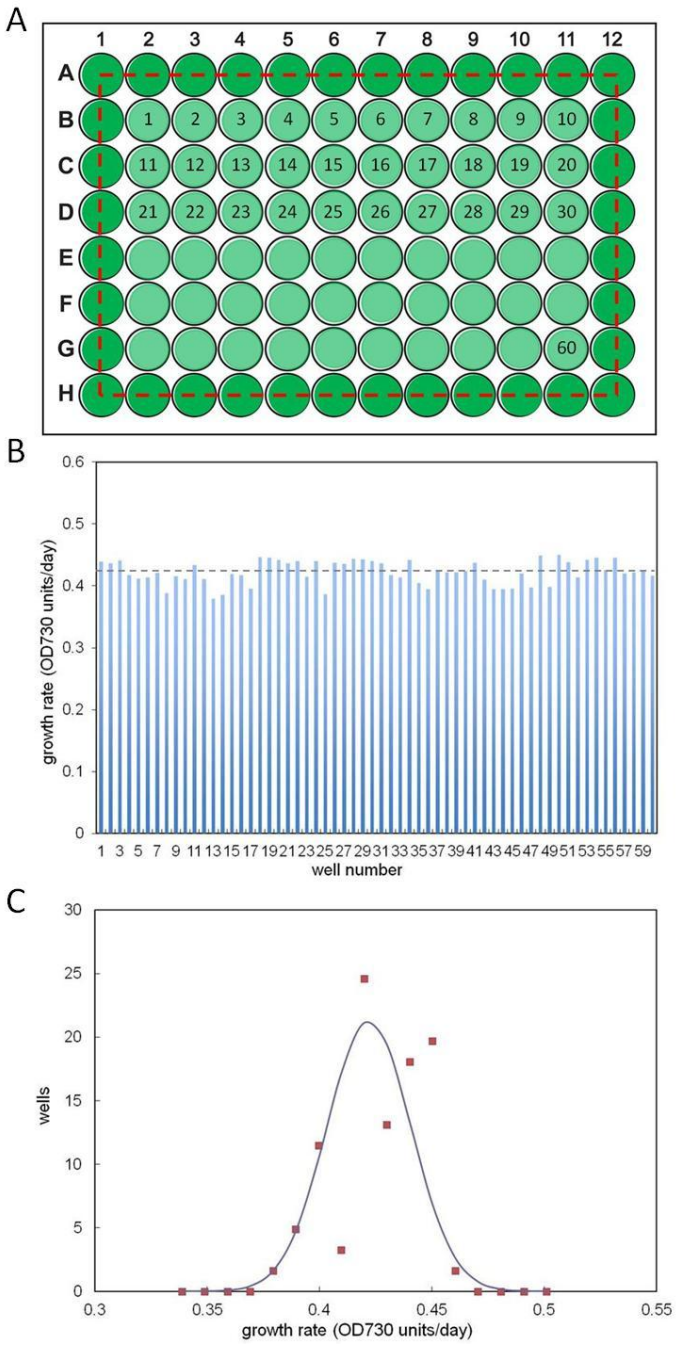


Figure 4-9: (A) Illustration of a 96-well plate used to measure growth rates of various rGS strains and WT *Synechococcus*. The outer wells contain WT *Synechococcus* and were not taking into account for growth analysis. Only the 60 inside wells were used for growth analysis of rGS strains and each strain was cultivated in at least 2 different locations per plate and on 2 different

plates. (B) Reproducibility of growth rate among WT *Synechococcus* samples grown in 60 wells of a 96-well plate (outside wells excluded). Growth rate during exponential growth is shown for 60 samples. The mean growth rate is 0.42 OD_{730} units/day and the standard deviation is 0.019. (c) The distribution of growth rates is plotted as number of wells at each growth rate (OD_{730} units/day). The red squares represent experimental values and the theoretical normal distribution is shown in blue.

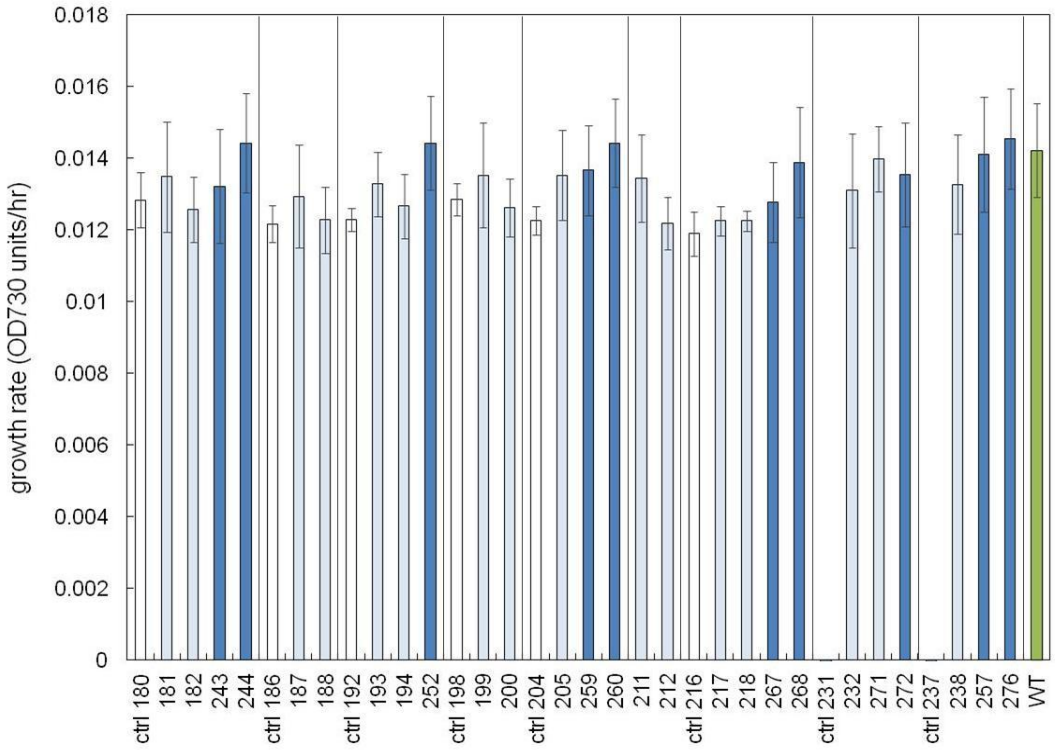


Figure 4-10. Growth rate of various rGS strains, negative control strains and WT *S. elongatus* (green) cultured in 96-well plates. Strains expressing the rGS without overexpression of the 3 native genes are shown in light blue, while strains with overespression of ACN, MDH and FUM are in dark blue. Control strains lacking MTK/MCL are labeled "ctrl" and shown in white. The genes expressed in each strain are detailed in table 4-3. Briefly, the all contained *H. sapiens* Acl, *R. eutropha* Icl, *E. coli* Frd and *M. capsulatus* Mtk (full strains only). Each set of strains separated by vertical lines, varies in their POR gene. Inside one set, when 2 full strains exist for one control, they vary in term of Mcl (from *M. extorquens* or *R sphaeroides*) and in term of Pyc (*B. subtilis* or *S. cerevisiae*).Note that the first 6 control strains 180, 186, 192, 198, 204 and 216 were cultured for a long time before use in this experiment, while control strains 231 and 237 (no growth) were recently constructed. The older control strains used to grow poorly in preliminary data and might have evolved to grow better. Error bars represent SD between $n \geq 4$ cultures.

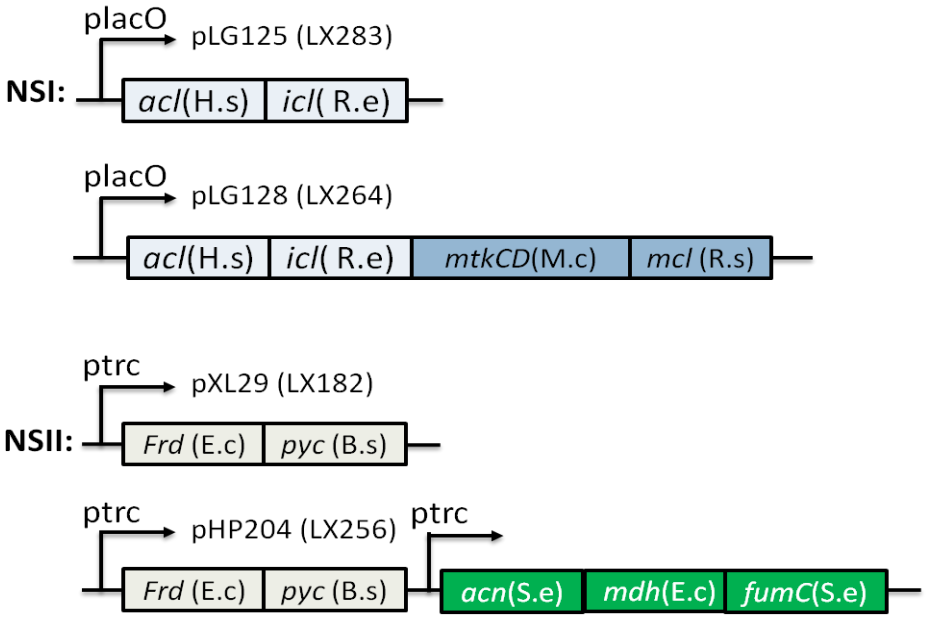


Figure 4-11: Gene region of plasmid constructs used for NSI and NSII integration sites. Plasmid pLG125 was used to construct negative control strains (like LX283) lacking the mtk and mcl genes required for rGS to function. Plasmid pLG128 contains the full set of genes for NSI. Plasmids pXL29 and pHP204 were used for integrated in NSII of the remaining rGS genes, with pHP204 containing three extra genes (*acn*, *mdh*, *fum*).

Gene	Activity in *Synechococcus*	Activity in *E. coli*
acl H.s.	✗	✗
icl R.e.	✗	✓ (50)
frd E.c.	Difficult to assay	Difficult to assay
pyc B.s.	✗	✗
acn S.e.	✗	✗
mtk M.c. / mcl R.s.	✓ (0.58)	✓ (2.5)
mdh E.c.	✓ (97)	-
fum S.e.	✓ (56)	-

Table 4-5: Summary of rGS enzymes activity measurements in *Synechococcus* lysate of final strains and in *E. coli* lysate from strain expressing the constructs pLG125, pLG128 or pHP204. Activities are expressed in U per mg of total soluble protein, where U = conversion of 1 µmol of substrate/min. *Icl* from *R.eutroph*a was not active in the rGS full strain of *Synechococcus* but showed an activity of 13.8 U/mg of total soluble protein in the negative control (LX283 expressing pLG125). This difference could be attributed to mutations in the gene or RBS that would have occurred during integration in *Synechococcus* in some strains. *Acn* from *Synechococcus* (*S.e.*) had measureable activity, between 1.5-2.0 U/mg of total soluble protein in the strains tested, but this value was lower that the one measured in the WT control (3.4 U/mg of total soluble protein) and was thus considered null. MDH and FUM activity were not tested in *E.coli* lysate as their activity was already pretty high in *Synechococcus*.

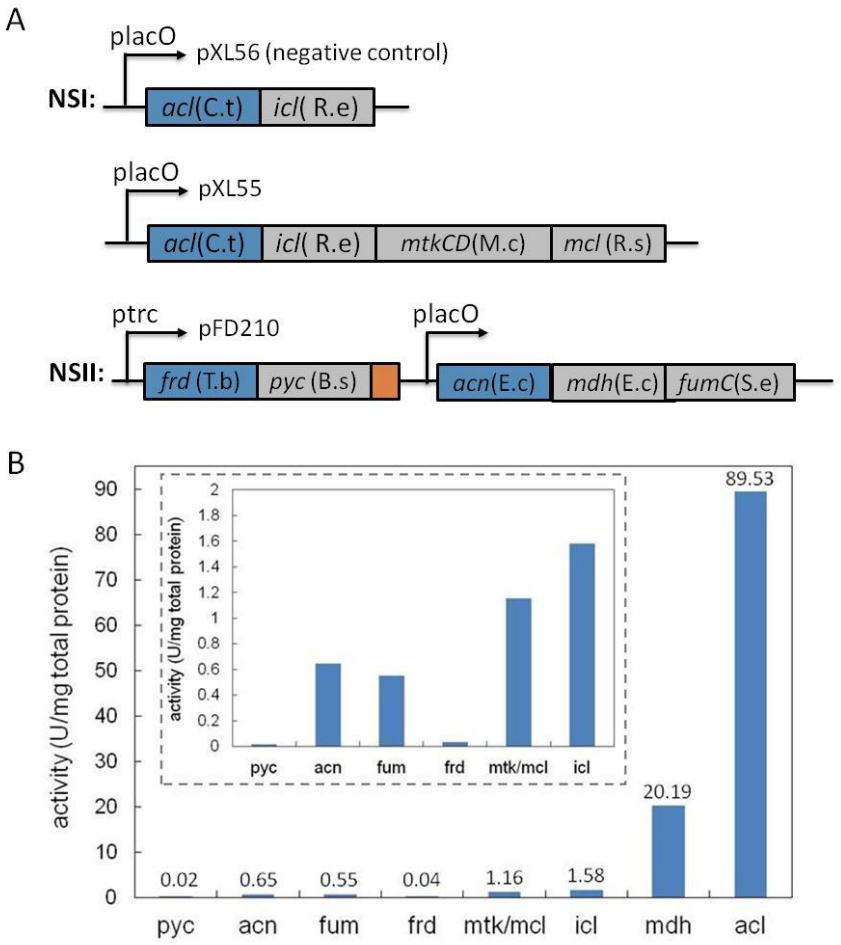


Figure 4-12. (A) New plasmid constructs for NSI and NSI integration sites, pXL55 and pFD210 respectively. pXL56 is the negative control plasmid. The genes in grey were unchanged from the first generation of constructs. The genes in blue were newly cloned to fix the low/no activity measured with the previous constructs. The orange region after pyc indicates a HIS-tag. (B) Activity of each rGS enzyme tested in E. coli XL1Blue lysate transformed with pXl55 or pFD210. The cells were induced with IPTG at OD600 = 0.4, and collected by centrifugation 3 hours later before being lysed by sonication for enzyme assay. The dash-lined graph represents a zoom of the main graph as two enzymes, mdh and acl, had much stronger activity than the others.

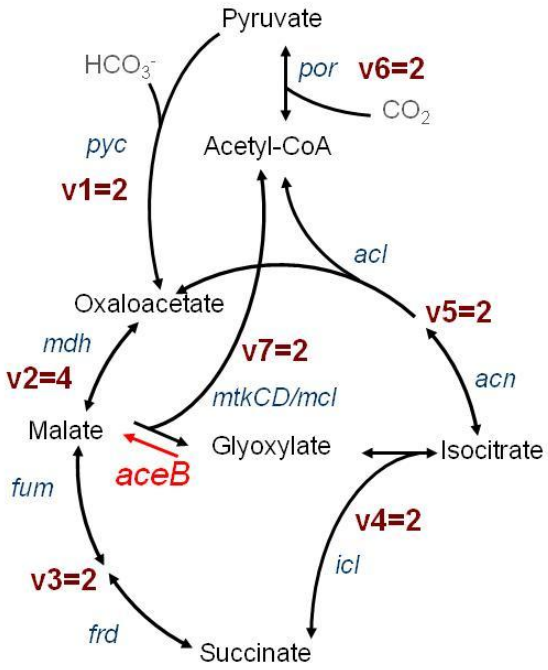


Figure 4-13. The rGS cycle and its fluxes at equilibrium. Without the malate synthase (*aceB*) step (red arrow) a possible rate-imbalance exists as rates v7 andv3 must be finely balanced. Glyoxylate and succinate are both formed from malate and are needed in stochiometric amount to further form isocitrate. If the malate synthase step is inserted both v7 and v3 are reversible and the rates of the branchpoint reactions do not need to be finely tuned, making the cycle more robust.

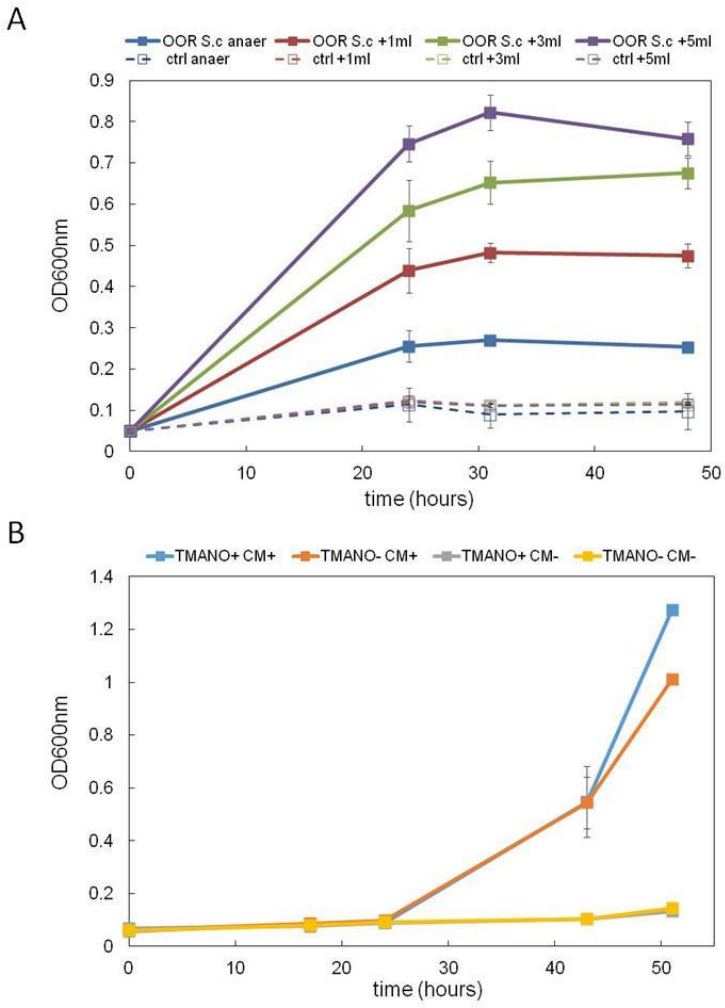
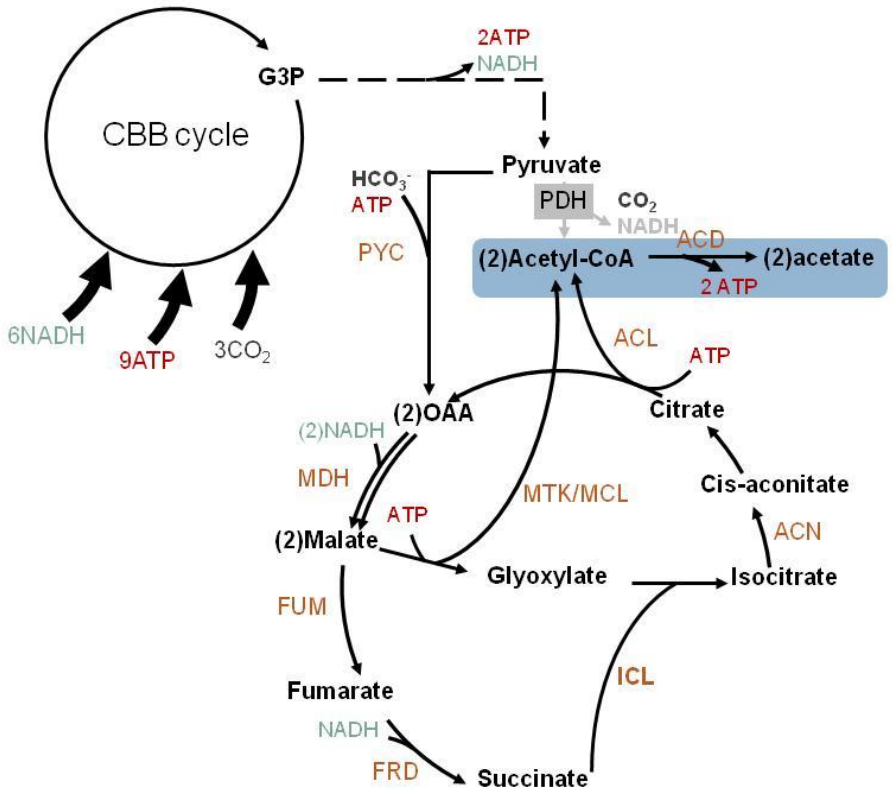


Figure 4-14. (A) Growth rescue of JCL301 by *S. coelicolor* OOR in M9 medium supplemented with 1% glucose, 25mM TNAMO. Chloramphenicol (20 μg/ml) was added to the culture expressing *S. coelicolor* OOR on plasmid pFD84.The cells were cultured in 5ml hermetically capped tubes in an atmosphere of 5% hydrogen, 10% carbon dioxide balanced in nitrogen with 0mL (blue), 1mL (red), 3mL (green) or 5mL (purple) replaced with air.The dash lines represent the growth of JCL301 alone in the same conditions. (B) Growth of JCL301 in M9 medium supplemented with 1% glucose, with or without 25mM TNAMO and with or without chloramphenicol under aerobic conditions. Growth was observed only in the cultures containing

chloramphenicol (20 μg/ml). Chloramphenicol is desolved in 100% ethanol to make the stock solution (1000x). Ethanol can be metabolized back to acetyl-CoA which would rescue JCL301 growth. The growth rescue observed in (A) was actually due to the antibiotic solution in the media (chloramphenicol desolved in ethanol).

A



B

	CBB cycle + linear rGS	CBB cycle + PDH
CO_2 fixation enzyme(s)	Rubisco and PYC	Rubisco
CO_2 to acetyl-CoA reaction	$4CO_2$+8NAD(P)H+10ATP→ 2acetyl-COA	$4CO_2$+8NAD(P)H+14ATP→ 2acetyl-COA
CO_2 to acetate reaction	$4CO_2$+8NAD(P)H+8ATP →2acetate	$4CO_2$+8NAD(P)H+12ATP →2acetate

Figure 4-15. (A) Linear rGS pathway with acetate read out. (B) Comparison between linear rGS and pyruvate dehydrogenase (PDH) for production of acetyl-CoA and acetate from CO_2 in cyanobacteria. The linear rGS has great advantage in acetyl-CoA production from CO_2 in term of ATP saving. Acetyl-CoA can further be converted into acetate by ADP-forming acetyl-CoA synthetase (ACD).

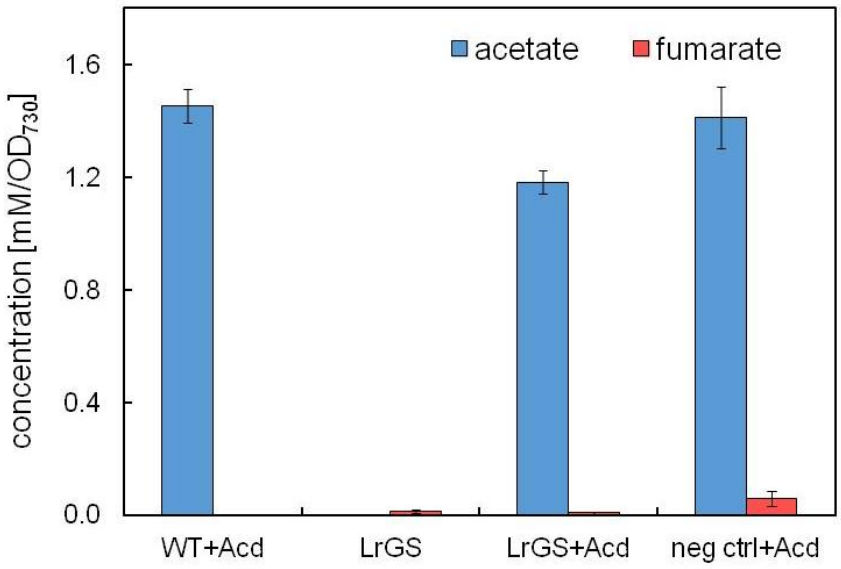


Figure 4-16. Production of acetate and fumarate by recombinant *Synechococcus* strains. LrGS strain is overexpressing genes from plasmid pXL55 and pFD210. Neg ctrl represents the negative control LrGS strain lacking mtk/mcl with genes from plasmid pXL56 and pFD210 overepressed. All stain termed "+Acd" have the *Pyrococcus furiosus* acetate-CoA ligase ADP-forming (*acd*) overexpressed in NSIII (as well as the *aceB* gene). Production is normalized by OD_{730}. Cells were cultured (50ml) in shake flasks, fed 50mM bicarbonate daily and induced with 1mM IPTG when OD_{730} reached about 0.5. Samples of 1 ml were taken at day 4 after induction and HPLC analysis was then conducted on the samples as listed under Materials and Methods.

4.6 References

1. Spreitzer, R. J. & Salvucci, M. E. Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.* **53,** 449–475 (2002).

2. Tcherkez, G. G., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci.* **103,** 7246–7251 (2006).

3. Ninomiya, N., Ashida, H. & Yokota, A. in *Photosynthesis. Energy from the Sun* 867–870 (Springer, 2008).

4. Ruan, C.-J., Shao, H.-B. & Silva, J. A. T. da. A critical review on the improvement of photosynthetic carbon assimilation in C3 plants using genetic engineering. *Crit. Rev. Biotechnol.* **32,** 1–21 (2012).

5. Zarzycki, J., Axen, S. D., Kinney, J. N. & Kerfeld, C. A. Cyanobacterial-based approaches to improving photosynthesis in plants. *J. Exp. Bot.* **64,** 787–798 (2013).

6. Reyes-Prieto, A., Weber, A. P. M. & Bhattacharya, D. The origin and establishment of the plastid in algae and plants. *Annu. Rev. Genet.* **41,** 147–168 (2007).

7. Martin, W. *et al.* Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12246–12251 (2002).

8. Mainguet, S. E., Gronenberg, L. S., Wong, S. S. & Liao, J. C. A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab. Eng.* **19,** 116–127 (2013).

9. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

10. Bustos, S. A. & Golden, S. S. Light-regulated expression of the psbD gene family in Synecbococcus sp. strain PCC 7942: evidence for the role of duplicated psbD genes in cyanobacteria. *Mol. Gen. Genet. MGG* **232,** 221–230 (1992).

11. Andersson, C. R. *et al.* Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.* **305,** 527–542 (2000).

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

13. Zelić, B. *et al.* Fed-Batch Process for Pyruvate Production by Recombinant Escherichia coli YYC202 Strain. *Eng. Life Sci.* **3,** 299–305 (2003).

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

15. Eaton-Rye, J. in *Photosynthesis Research Protocols* (ed. Carpentier, R.) 309–324 (Humana Press, 2004).

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

17. Mohamed, A. & Jansson, C. Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium Synechocystis 6803. *Plant Mol. Biol.* **13,** 693–700 (1989).

18. Zang, X., Liu, B., Liu, S., Arunakumara, K. K. I. U. & Zhang, X. Optimum conditions for transformation of Synechocystis sp. PCC 6803. *J. Microbiol. Seoul Korea* **45,** 241–245 (2007).

19. Soeda, E., Torikata, T. & Akune, S. Sarcosomal aconitase of silkworm. *J. Biochem. (Tokyo)* **73,** 529–536 (1973).

20. Malate Dehydrogenase - Assay. Available at: http://www.worthington-biochem.com/mdh/assay.html. (Accessed: 10th March 2016)

21. Enzymatic Assay of Fumarase | Sigma-Aldrich. Available at: http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-fumarase.html. (Accessed: 10th March 2016)

22. Singh, S. K., Miller, S. P., Dean, A., Banaszak, L. J. & LaPorte, D. C. Bacillus subtilis isocitrate dehydrogenase. A substrate analogue for Escherichia coli isocitrate dehydrogenase kinase/phosphatase. *J. Biol. Chem.* **277,** 7567–7573 (2002).

23. Kerscher, L. & Oesterhelt, D. Ferredoxin is the coenzyme of alpha-ketoacid oxidoreductases in Halobacterium halobium. *FEBS Lett.* **83,** 197–201 (1977).

24. Noth, J., Krawietz, D., Hemschemeier, A. & Happe, T. Pyruvate:Ferredoxin Oxidoreductase Is Coupled to Light-independent Hydrogen Production in Chlamydomonas reinhardtii. *J. Biol. Chem.* **288,** 4368–4377 (2013).

25. Zhang, S. & Bryant, D. A. The Tricarboxylic Acid Cycle in Cyanobacteria. *Science* **334,** 1551–1553 (2011).

26. Colyer, C. L., Kinkade, C. S., Viskari, P. J. & Landers, J. P. Analysis of cyanobacterial pigments and proteins by electrophoretic and chromatographic methods. *Anal. Bioanal. Chem.* **382,** 559–569 (2005).

27. Jitrapakdee, S. *et al.* Structure, Mechanism and Regulation of Pyruvate Carboxylase. *Biochem. J.* **413,** 369–387 (2008).

28. Narikawa, S. & Nakamura, M. Differentiation of obligate anaerobes by assay of pyruvate:Ferredoxin oxidoreductase activity. *Eur. J. Clin. Microbiol.* **6,** 74–75 (1987).

29. Lee, Y., Lafontaine Rivera, J. G. & Liao, J. C. Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. *Metab. Eng.* **25,** 63–71 (2014).

30. Schmitz, O., Gurke, J. & Bothe, H. Molecular evidence for the aerobic expression of nifJ, encoding pyruvate:ferredoxin oxidoreductase, in cyanobacteria. *FEMS Microbiol. Lett.* **195,** 97–102 (2001).

31. Marg, B.-L., Schweimer, K., Sticht, H. & Oesterhelt, D. A two-alpha-helix extra domain mediates the halophilic character of a plant-type ferredoxin from halophilic archaea. *Biochemistry (Mosc.)* **44,** 29–39 (2005).

32. Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim. Biophys. Acta BBA - Bioenerg.* **1827,** 1039–1047 (2013).

33. Billini, M., Stamatakis, K. & Sophianopoulou, V. Two Members of a Network of Putative Na+/H+ Antiporters Are Involved in Salt and pH Tolerance of the Freshwater Cyanobacterium Synechococcus elongatus. *J. Bacteriol.* **190,** 6318–6329 (2008).

34. Horner, D. S., Hirt, R. P. & Embley, T. M. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Mol. Biol. Evol.* **16,** 1280–1291 (1999).

5. Alterations of the reverse glyoxylate shunt cycle for carbon fixation

5.1 Introduction

Photosynthesis, the basis of plants, algae and certain bacteria growth, describes the process of capturing sunlight energy and converting it into biochemical energy. Carbon dioxide (CO_2) is fixed through the Calcin-Benson-Bassham (CBB) cycle using the sun energy to create higher energy molecules such as sugars. Improving photosynthesis would help not only the world crop, and thus food production, but also could contribute to more efficient production of advanced biomass-based biofuels. Biofuels represents a renewable energy source and promising alternative to petroleum-based fuels. However, their current production is limited by the ineffective processes used to convert plant matter to fuel and by the low efficiency of photosynthesis, mostly attributed to the inefficiency of the CBB cycle main enzyme, ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco).

Our group proposed to develop and implement a new carbon fixation pathway in plants, with the goal of improving carbon assimilation efficiency and later biomass-based biofuels production. The new CO_2 fixation cycle, termed reverse Glyoxylate Shunt (rGS) (Fig. 5-1A) fixes CO_2 with only 30% of the ATP requirement compared to the native Calvin cycle and does not utilize Rubisco for carbon fixation, thus will not have the photorespiration associated carbon loss. This new synthetic pathway also channels the carbon flux to pyruvate, a direct precursor of biofuels, in place of sugars.

To test the compatibility of the rGS cycle with the photosynthetic machinery we previously decided to implement it first in cyanobacteria, a model system for phototrophic eukaryotes. As described in chapter 4, the original design had to be altered to form a linear pathway. The

original rGS required a pyruvate ferredoxin: oxidoreductase enzyme (POR) to catalyze the conversion of acetyl-CoA to pyruvate. Despite our effort, no oxygen tolerant POR active under our conditions could be identified. Without the last step, the rGS cycle become a linear pathway, termed the Linear reverse Glyoxylate Shunt (LrGS), that still allows to fix carbon more efficiently through both the CBB cycle and PYC to produce acetyl-CoA (Fig. 5-1B). An acetyl-CoA ligase (ACD) was added that catalyzed the conversion of acetyl-CoA to acetate, which allowed for a better read-out to evaluate the activity of the pathway. In *Synechococcus elongatus*, overexpression of the LrGS with ACD did not increase the acetate production compare to the control strains. However, the presence of a small concentration of fumarate only in the LrGs strains seemed to indicate the presence of a small carbon flux from possibly pyruvate to fumarate. Additionally, a control LrGS strain missing MTK/MCL produced more fumarate, which could be explained by active MTK/MCL enzymes in the full strain: absent in the control, they do not drain malate away to acetyl-CoA and glyoxylate, leaving it available as a direct substrate for fumarate production. These results from chapter 4 could indicate a slight partial activity of LrGS in *Synechococcus*, until the branched point formed by MTK/MCL and FUM around malate. This junction had been identified as a possible kinetic trap because of the MTK/MCL irreversibility that may prevent the cycle from running. A malate synthase enzyme that catalyzes the reverse reaction of MTK/MCL was added to make that step reversible and attempt to avoid the rate imbalance. The exact rate equilibrium at that junction would be difficult to assess *in vivo*, and even with now reversible reactions consuming malate, the balance around that junction may still be problematic for the whole LrGS to function.

In this chapter we first focused on improving the enzymatic activity of the first three steps of the pathway, termed rGS0.5 (Fig. 5-1), to produce glyoxylate and acetyl-CoA, this later being the

product of the LrGS. We then proposed a new way of recycling glyoxylate, inspired by the glycerate pathway. The glycerate pathway is used by certain bacteria like *E. coli* to convert glyoxylate into phosphoenolpyruvate (PEP) and supply carbon skeletons for biosynthesis[1]. Some cyanobacteria like *Synechocystis* sp. PCC 6803, natively use the glycerate pathway as one of three routes to metabolize 2-phosphoglycolate (2PG)[2], the toxic product of photorespiration[3,4]. Two molecules of 2PG are converted to CO_2 and 3-phosphoglycerate (3PGA) that can reenter the CBB cycle.

Our new carbon fixation design, termed the reversed Glyoxylate Shunt 2 (rGS2), form a new carbon fixation cycle that produce acetyl-CoA (Fig. 5-1C). Compare to the rGS design (chapter 4), this new cycle has the advantage of not containing the kinetic node around malate where MTK/MCL and FUM have to compete for the same substrate. However, rGS2 requires equivalent amount of ATP compared to the CBB cycle, but could nevertheless still increase carbon fixation efficiency as it does not solely rely on Rubisco. Implementation of this new cycle in *Synechococcus* led to increase intracellular acetyl-CoA pool and further secretion α-ketoisocaproate. α- Ketoisocaproate (KIC; 2-keto-4-metyl-pentanoate) is the final intermediate metabolite in the leucine synthesis pathways and its corresponding branched-chain ketoacid. In microorganisms and plants, KIC and leucine are synthesized from the addition of acetyl-CoA to the direct precursor of valine, α-ketoisovalerate (KIV). KIV is itself synthesized via several steps starting from two molecules of pyruvate (Fig. 5-2). KIC could be used in treatment of patients with chronic kidney disease and serving as a nitrogen-free substitute to provide their leucine requirement[5–7]. This is of great interest since a low-nitrogen diet reduces hyperammonemia and dietary protein intolerance usually linked with some hepatic disorders and renal failure[8]. KIC is also used as a nutritional supplement because of its possible anti-catabolic

properties which may reduce muscle degradation after high-intensity exercise such as resistance training[9]. Because of the multiple application of KIC, significant efforts have been made towards its production, mainly by chemical synthesis using different methods[10]. But requirements for expensive catalyst and specific infrastructure result in high production costs. Microbial synthesize and enzymatic transformation represent alternative processes for the production of KIC. Moreover, an organism producing KIC could be further engineered to yield some derived products of interest like 3-methyl-1-butanol, a potential future biofuel[11,12]. Production of KIC from glucose and acetate in engineered *Corynebacterium glutamicum* yielded 9.2 ± 0.4 g/L[13], while a whole-cell bioconversion method from leucine using engineered *E. coli* led to maximal KIC titer of 69.1 g/L[14].

In the present study we demonstrate the implementation of the rGS2 cycle in cyanobacteria, which lead to intracellular acetyl-CoA accumulation and further KIC secretion, but also an impaired cells growth possibly because of the toxicity of one rGS2 intermediate. The rate of KIC production in the recombinant culture decreased in subcultures, probably from competition with faster growing mutants that evolve against rGS2 activity and its related toxicity.

5.2 Materials and Methods

5.2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich or Fisher Scientifics unless otherwise specified. KOD and KOD xtreme DNA polymerases were purchased from EMD Millipore and used for gene amplification from genomic or plasmid DNA. T5-Exonuclease (Epicenter), Taq DNA ligase (New England Biolabs) and *Phusion* DNA polymerase (Fisher) were purchased

individually and used to make the assembly master mix (AMM) used for cloning as described by Gibson[15].

5.2.2 Bacterial strains, DNA manipulations and plasmids construction

The wild-type (WT) strain of cyanobacteria *Synechococcus elongatus* PCC 7942, hereafter referred as *Synechococcus* (obtained from Pr. Susan S. Golden) was used as the host for rGS2 implementation. All chromosomal manipulations were carried out by homologous recombination of plasmid DNA into *S. elongatus* genome at neutral site I (NSI)[16,17] and II (NSII)[18]. *Escherichia coli* XL-1 Blue (Stratagene) was used to construct, propagate and store all plasmids in this study. *E. coli* BL21 (DE3) was used for protein expression and purification. Strains and plasmids used in this study are listed in Table 1.

For rGS2 integration in *Synechococcus* genes were cloned under the control of an IPTG-inducible promoter (PLlacO1 or Ptrc) into either a NSI vector fragment[17] (Sp^R) an NSII vector fragment[18] (Km^R) using the Gibson method of DNA assembly[15]. Both inserts (target genes) and vector backbones were amplified by PCR using KOD Xtreme DNA polymerase followed by a purification step using a PCR purification Kit (Zymo Research). Cloned were screened for the correct inserts by colony PCR and DNA sequencing (Genewiz).

5.2.3 Culture medium and growth conditions

All *Synechococcus* strains were grown at 30 °C under continuous illumination supplied by four Lumichrome F30W-1XX 6500K 98CRI light tubes (100 μmol photons m^{-2} s^{-1} unless otherwise specified) in solution of BG-11 (Sigma, Cyanobacteria BG-11 Freshwater Solution 50x) or on modified BG-11 agar (1.5% w/v) plates[19] . Strains were cultured in 250 mL screw cap flasks or 5ml glass tubes. 50 mM $NaHCO_3$ was added to the medium. Cell growth was monitored by

measuring OD_{730} with Beckman Coulter DU800 spectrophotometer. BG-11 media was supplemented with antibiotics (20 μg⁄mL spectinomycin (Sp) and 10 μg⁄mL kanamycin (Km)) when appropriate.

E. coli strain XL-1 Blue was grown at 37 °C in LB broth or on solid LB agar 1.5% (wt/vol). Spectinomycin (100 μg⁄mL) or kanamycin (50 μg/ml) were added to the LB medium when required for the propagation of plasmids in *E. coli* or for protein expression. To induce protein expression, strains were induced with 0.5mM IPTG (unless otherwise specified) when OD_{730} for S*ynechococcus* and OD_{600} for *E. coli* reached 0.5.

5.2.4 Transformation of *Synechococcus elongatus* PCC 7942

Synechococcus is naturally competent and was transformed as followed: 2ml of cells at mid-log phase ($OD_{730}$of 0.4–0.6) were spun down at room temperature and resuspended in 200μl of fresh BG-11. Cells were then incubated with 1-3 μg of plasmid DNA overnight in the dark, before being spread on BG-11 plates with appropriate antibiotics for selection of successful recombination. Colonies grown on BG-11 agar plates were analyzed by PCR using gene-specific primers to verify integration of inserted genes into the recombinant strain. In all cases, eight individual colonies were analyzed by PCR for verification. One positive colony was then propagated and tested for downstream experiments.

5.2.5 Enzyme Assays

Enzyme assays were conducted by using Bio-Tek PowerWave XS microplate spectrophotometer. Cell-free extracts were prepared as followed: Cyanobacteria liquid cultures (25ml) were grown until OD_{730nm} reaches 0.4-0.6, and were then induced with 1 mM IPTG. The induced cultures were further grown for 2 days to allow protein expression. Cells were collected

by centrifugation at 5,250 × g for 15 min (4°C). The pellet was then resuspended in 1 mL of the appropriate assay buffer and mix with 500µl of 0.1 mm glass beads (Biospec). The sample was then homogenated using a bead beater (biospec). The soluble protein fraction was collected after centrifugation. Purified protein was collected after His-spin column purification (Zymo Research).Total protein measurements were made using the Pierce Coomassie Plus Assay Reagent (ThermoFisher Scientific).

Malate dehydrogenase activity was measured in the direction of the reduction of oxaloacetate to malate. The concomitant oxidation of NADH was monitored by a decrease in absorbance at 340 nm[20]. A 200 µl reaction mixture contained 100mM potassium phosphate buffer pH 7.5, 0.375 mM NADH, and total soluble protein from bacteria crude extract. The reaction was started with the addition of oxaloacetate (0.6 mM final concentration). Enzymatic activity is expressed as U of NADH oxidized per mg of soluble protein at 30 °C (U = µmol/min).

Pyruvate carboxylase activity was determined by coupling the reaction to malate dehydrogenase from porcine heart (sigma), which oxidizes NADH. The assay mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 5 mM ATP, 0.6 mM Acetyl-CoA, 15 mM $NaHCO_3$, 0.2 mM NADH, 5 mM pyruvate, 0.5 Units/ml malate dehydrogenase and protein extract or purified protein. The reaction was started at 30 °C with the addition of pyruvate, and decrease in absorbance was followed at 340 nm. The activity is expressed as µmol of malate formed per min per mg of soluble protein (U/mg of total soluble protein), which is equivalent to the amount of NADH oxidized.

Phosphoenolpyruvate carboxylase was measured using a modified protocol based on the pyruvate carboxylase assay where 5 mM pyruvate was replaced by 4 mM phosphoenolpyruvate.

Malate thiokinase and malyl-CoA lyase activities were tested simultaneously. The assay used works on the following principle: Malate thiokinase performs the ATP-dependent condensation of malate and CoA into malyl-CoA. In turn, malyl-CoA lyase cleaves malyl-CoA into acetyl-CoA and glyoxylate, the latter reacting with phenylhydazine to form glyoxylate-phenylhydrazone. Formation of glyoxylate-phenylhydrazone is recorded at 324 nm. The assay mixture contained 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 2 mM phenylhydrazine, 5 mM malate, 1 mM ATP, 0.5 mM CoA, and soluble protein extract. The reaction was started at 30 °C with the addition of malate, and increase in absorbance was followed at 324 nm. The activity is expressed as μmol of glyoxylate formed per min per mg of soluble protein (U/mg of total soluble protein).

Glyoxylate carboligase and tatronate semialdehyde reductase were assayed simultaneously in a coupled assay. Enzymes activity was followed spectrophotometrically by measuring the decrease in absorbance at 340nm resulting from the utilization of NADH. The reaction mixture included the following: 100 mM Tris-HCl (pH 7.5), 5mM $MgCl_2$, 0.2mM NADH; 0.5mM thiamine diphosphate; 5mM glyoxylate and an appropriate amount of cell lysate. All reactions were carried out at 30 °C. A unit of activity is defined as that amount of enzyme catalyzing the formation of one μmol of NADH per minute in the reaction mixture. Specific activity is defined in terms of activity units per mg of protein (U/mg of protein).

Glycerate 2-kinase catalyzed the ATP-dependent phosphorylation of glycerate to 2-phosphoglycerate and ADP. Its activity was determined at 30°C by coupling 2-phosphoglycerate formation to the oxidation of NADH via enolase, phosphoenolpyruvate carboxylase and malate dehydrogenase. The reaction mixture contained 100 mM Tris-HCl (pH 7.5), 5 mM D-glycerate, 5 mM $MgCl_2$, 0.25 mM NADH, 2.5 mM ATP, 1 Units/ml enolase, 1.2 phosphoenolpyruvate

carboxylase, 0.5 Units/ml malate dehydrogenase and cell lysate. The activity is expressed as μmol of NADH formed per min per mg of soluble protein (U/mg of total soluble protein).

5.2.6 Alpha-ketoisocaproate production

A loopful of *Synechococcus* was used to inoculate fresh 5 mL BG-11 with 50 mM $NaHCO_3$. The culture was grown for 3 days before being diluted to $OD_{730\,nm}$ of 0.4 to 0.6 in glass tubes containing 5ml BG11 supplemented with the appropriate antibiotics, 0.25mM thiamine diphosphate (TPP; cofactor of glyoxylate carboligase) and 0.05mM pantothenate a precursor in coenzyme-A (CoA) synthesize. Cultures were induced with 0.5 mM IPTG as final concentration and purged in an atmosphere of 5% hydrogen, 10% carbon dioxide balanced in nitrogen. 60 μl of growing culture was sampled everyday for cell density measurements and 400 μl was sampled every other day for alpha-ketoisocaproate (KIC) production measurements. Everyday 200 μl of fresh BG-11 with 1 M $NaHCO_3$, appropriate antibiotics, and IPTG, TPP and pantothenate were added back to the culture. Method for KIC quantification is listed below.

5.2.7 Quantification analysis of alpha-ketoisocaproate and acetyl-CoA

Alpha-ketoisocaproate was quantified by a high-performance liquid chromatography (HPLC). Culture samples (400 μl) were centrifuged for 5 min at 15,000×*g*. The supernatant was analyzed by Agilent 1200 HPLC equipped with a BioRad HPX87 column using a photodiode array detector at 210 nm absorbance. The mobile phase used was 30 mM H_2SO_4 at a constant flow rate of 0.6 mL/min. The column was maintained at 30 °C and the injection volume was 20 μL.

Intracellular acetyl-CoA was analyzed by ion-pair liquid chromatography-triple quadrupole mass spectrometry, IP-LC/QqQ-MS, like previously described[21] using a Shimadzu Nexera UHPLC system coupled with LCMS 8030 plus (Shimadzu Co., Japan).

5.3 Results

5.3.1 Improving rGS0.5 enzymatic steps

The rGS0.5 is a three steps pathway that produces acetyl-CoA and glyoxylate and starts with the formation of oxaloacetate (OAA) through a carboxylation steps (Fig. 5-1). Four enzymes are required for this pathway: 1) pyruvate carboxylase (PYC), or phosphoenolpyruvate (PEP) carboxylase to form OAA; 2) malate dehydrogenase (MDH) to catalyze the formation of malate from OAA; 3) malate thiokinase (MTK) and malyl-CoA lyase (MCL) that catalyze together the convertion of malate to glyoxylate and acetyl-CoA. Based on previous results obtained in Chapter 4, step 1 and 3 enzymatic activities needed to be improve as the ones measured were much lower than MDH activity, up to 1000 folds difference between PYC and MDH when activity were measured in *E. coli* expressing the cyanobacterial rGS plasmids.

The synthesis of oxaloacetate through carboxylation can be accomplished from either pyruvate using a pyruvate carboxylase (PYC), or from phosphoenolpyruvate (PEP) using a PEP carboxylase (PPC). In Chapter 4 we screened 5 different pyc genes for activity in *Synechococcus*, but even the most active one had only detectable activity if HIS-purified, that could not be detected in cyanobacteria lysate (Fig. 5-4A). PYC is a biotin-dependant enzyme that catalyses the carboxylation of pyruvate to oxaloacetate using Mg^{2+} and ATP as cofactors. It is found in many prokaryotes, yeast and animals that catalyses the carboxylation of pyruvate to oxaloacetate, an important anaplerotic reaction to replenish the Krebs cycle of intermediates that were withdrawn for anabolic purposes[22]. However, only some prokaryotes use PYC as the exclusive anaplerotic enzyme. Many bacteria and plants avoid the need for PYC by utilizing PPC, converting directly PEP to oxaloacetate. In comparaison to PYC, PPC does not require

biotin or ATP. We decided to replace PYC by PPC in our design, and cloned and purified two enzymes: *Escherichia coli* PPC and *Corynebacterium glutamicum* PPC. PPC enzymes activity is known to be regulated by a few effectors. In particular, *E. coli* PPC is inhibited by aspartate and malate, and activated by acetyl-CoA and fructose-1,6-bisphosphate[23]. Inhibition by aspartate had also been discribed for *C. glutamicum* PPC, while no effect from acetyl-Coa was reported[24]. Because both malate and acetyl-CoA are part of the rGS2 as intermediate and product respectively, we first decided to assess their effect on both purified PPC enzymes expressed in *E.coli* (Fig. 5-3). In our conditions, both enzymes showed their maximum activity when 0.6mM acetyl-CoA was supplied (0.83U/mg for *E.coli* PPC and 1.25U/mg for *C. glutamicum* PPC). Above that concentration no significant improvement in activity was observed (data not shown). Without acetyl-CoA *E.coli* PPC activity dropped to almost zero meanwhile the *C. glutamicum* one kept about 30% of its maximum activity (Fig. 5-3A). The effect of malate was also stronger on *E. coli* PPC compare to the *C. glutamicum* one. At 1mM malate *E.coli* PPC activity was about 70% of its maximum measured, while *C. glutamicum* activity was almost unchanged (Fig. 5-3B). At 5mM the activity dropped to 16% for the *E.coli* PPC compare to 49% for the *C. glutmicum* one. The activity of both enzymes was then measured when expressed in S*ynechococcus* as part of our carbon fixation pathway (Fig. 5-4A). Both PPC enzymes showed more activity than *E. coli* PYC when measured in cyanobacterial lysate. *C. glutmicum* PPC was approximately twice as active as *E. coli* one.

The last direct step in the production of acetyl-CoA through the rGS2 pathway before glyoxylate recycling is the two steps conversion of malate to acetyl-CoA and glyoxylate catalyzed by a malate thiokinase (MTK) and a malyl-CoA lyase (MCL). We previously screened 4 different MTK/MCL combinations for activity in *Synechococcus* (Chapter 4). The *Methylococcus*

capsulatus (*M. c.*) MTK/*Methylobacterium extorquens (M. e.*) MCL pair showed the highest activity followed by the *Methylococcus capsulatus* (*M. c.*) MTK/*Rhodobacter sphaeroides* (*R.s.*) MCL pair. The first combination was later dropped after focusing on a few more promising strains of the first rGS cycle (Chapter 4). The little difference in growth of these strains might however not have been due to the MCL activity.

The native *M. capsulatus* MTK was originally used by Mainguet et al[25] for engineering of rGS in *E.coli*. *M. capsulatus* MTK enzyme is encoded by the SucCD-2 operon and is annotated as succinyl-CoA synthetase, but showed the highest activity out of the 15 putative MTK they tested. In their study, as well as in our previous data (chapter 4), the native SucCD-2 operon was cloned under the control of a promoter of choice. Here we tested the influence on MTK activity of a well defined *E. coli* Ribosome Binding Site (RBS), AGGAGATATACC[26], introduced in between *sucC* and s*ucD*. Activity in Synechococcus lysate coupled with *R. sphaeroides* MCL was evaluated (Fig. 5-4B). An 80 fold rise in activity was observed with the new RBS. We then reevaluated the expression of the combination *Methylococcus capsulatus* (*M. c.*) MTK/*Methylobacterium extorquens (M. e.*) MCL as part of the new pathway design. This further increased the combined MTK/MCL activity of about 40% (Fig. 5-4B).

Having improved the first and last step of rGS0.5, we reevaluated enzymatic activities of all the three steps in *Synechococcus* lysate of cells expressing in NSI *C. glutamicum* PPC, *E. coli* PPC, *M.capsulatus* MTK (*sucC* and *sucD* with new RBS) and *M. extorquens* MCL (plasmid pFD258, Table 5-1). Figure 4C shows each activity measured as described in Materials and Methods. MDH still showed higher activity, but only one order of magnitude apart compare to the three orders of magnitude previously measured.

5.3.2 Implementation of the full rGS2 pathway and KIC production

With appropriate enzymes identified for acetyl-CoA production through rGS0.5, we proceeded to establish the glyoxylate recycling part of the cycle. We cloned the non-native genes *gcl* from R. *eutropha*, *glxR* and *garK* from *E. coli* under the control of a pLacO1 promoter into the NSII vector backbone that contains kanamycin resistance marker to give pXL145 (Table 1). All genes were first tested in *E.coli* for their individual and combined activity (data not shown). The last step of the recycling pathway that converts 2-phospho-D-glycerate to PEP and water (Fig. 5-1C) was not overexpressed as *Synechococcus* possess a native enolase gene (*eno*). We integrated these genes by homologous recombination at NSII into the genome of *Synechococcus* strains rGS0.5-ppcEc and rGS0.5-ppcCg (expressing *mdh*, *SucCD* and *ppc* from either *E. coli* or *C. glutamicum* at Neutral Site I (NSI) under the control of a pLacO1 promoter) and selected for successful transformant. Enzymatic activity of successful transformants was measured in *Synechococcus* lysate as described in Materials and Methods (Fig. 5-4D).

Addition of glyoxylate recycling on top of rGS0.5 form the new cycle termed rGS2. This new pathway was implemented in *Synechococcus* using both *E.coli ppc* (rGS2-ppcEc) and *C. glutamicum ppc* (rGS2-ppcCg). The version expressing the *C. glutamicum ppc* showed strong growth retardation compared to the *E.coli* version and only started growing about 6 days post induction (Fig. 5-5). Compared to WT, rGS2 with *E.coli ppc* already presented a slower growth rate. Moreover both rGS2 strains showed a yellowish color, indication of a metabolic stress. Interestingly, the two corresponding rGS0.5 strains had growth similar to WT and no apparent change in color. This result points out to a possible toxicity of one of the intermediate of the glyoxylate recycling part of the pathway that would be absent in the rGS0.5 strains. Moreover,

C. glutamicum ppc being more active *in vitro* might equal to a higher flux through rGS0.5, leading to a larger glyoxylate pool to be recycled. If an intermediate of the recycling pathway is indeed toxic, this would explain the difference in growth observed between the two versions of the rGS2.

We further investigated rGS2-ppcEc strain as the one expressing *C. glutamicum ppc* could barely grow. As shown in Figure 6B, strain rGS2 had a larger acetyl-CoA pool at day 4 after induction compared to WT and strain rGS0.5. Analysis by HPLC of the strain supernatant showed production of 13.7 mg/L of α-Ketoisocaproate (Fig. 5-6C), but the rate of KIC production in the recombinant culture decreased as a result of competition with faster growing KIC-non-forming mutants (Fig. 5-6A,C). This result supports our hypothesis of a toxic intermediate formed through the rGS2 pathway, thus giving a disadvantage to cells expressing a fully active cycle. The cells from strain rGS2-ppcEc with a restored growth phenotype and no KIC production had a 70% decreased in Mtk/Mcl activity compared to the initial strain (data not shown).

5.4 Discussion and Conclusion

Expression of a new synthetic pathway, termed rGS2, in *Synechococcus elongatus* PCC 7942 increase the intracellular acetyl-CoA pool and enables photosynthetic production of KIC, the last intermediate in leucine biosynthesis, and a possible treatment for patients with chronic kidney disease also used as a nutritional supplement because of its potential anti-catabolic properties. Strains expressing the rGS2 exhibit a yellowish phenotype and growth deficiencies, which was markedly stronger in the line expressing the *ppc* gene from *C.glutamicum* compared to the one expressing the *E.coli ppc* (rGS2-ppcEc). *C.glutamicum ppc* had demonstrated higher activity when assayed *in vitro* and showed lower inhibition by malate, an intermediate of the

RGS2 cycle (Fig. 5-3). In addition, it is interesting to note that strains expressing only the first half of the pathway, rGS0.5, did not exhibit any significant growth deficiencies, independently of which *ppc* gene was expressed (Fig. 5-5). Thus, the growth retardation observed in rGS2 strains is likely a result of the combined activities of the heterologous enzymes responsible for the recycling of glyoxylate, the second half of the pathway. Reversion of RGS2-ppcEc subcultures to growth rate similar to WT was observed and seemed to be associated with the decrease and further cessation of KIC production. Data showed an inverse correlation between growth and KIC production in the substrains (Fig. 5-6A and C). We hypothesized that this phenomenon was due to a strong selection against the RGS2 cycle, probably related to the toxicity of one of the rGS2 intermediate metabolite. Tartronate semialdehyde seemed like a candidate as toxicity of aldehydes in cyanobacteria has been described previously[17,27]. If the enzyme directly consuming tartronate semialdehyde (GlxR) is slower than the one producing it (Gcl), tartronate semialdehyde might accumulate and reach toxic intracellular levels. Both enzymes were assayed in a coupled reaction, making the identification of the limiting step difficult, but in any case, the activity measured was the lowest of all the heterologously expressed rGS2 enzymes (Fig. 5-4). One substrain that stopped to produce KIC showed a 70% decrease in Mtk/Mcl activity, the couple of enzyme responsible for the formation of glyoxylate, the direct substrate of tartronate semialdehyde in rGS2. This might be due to a selection for secondary mutations that result in a decrease in tartronate semialdehyde production, in that case through a decrease in Mtk/Mcl activity. Genetic instability has previously been observed in transgenic strains of cyanobacteria[28,29], but despite appearing like a valid issue, not many studies have been conducted so far[30].

The highest KIC productivity obtained in *Synechococcus* in this study, 13.7 mg/L, was in an early culture of rGS2-ppcEc, 8 days post-induction, when growth had reached a plateau. However, KIC secretion has been previously reported in a *Synechococcus* sp. PCC 7002 mutant deficient in glycogen synthesis, and reached a maximum of 250μM (32mg/L) under nitrogen starvation[31]. KIC was overproduced at the cost of a complete growth inhibition along other central metabolites presumably due to an unspecific carbon repartitioning. In our work, increase of the intracellular acetyl-CoA pool in rGS2 lines could explain an increase in KIC, as acetyl-CoA participate directly in its synthesis from α-ketoisovalerate (KIV), the last substrate in valine biosynthesis (Fig. 5-2). However, it may be premature to consider cyanobacteria a potential host for production of acetyl-CoA and related down products, at least until improved strategies for better genetic stability are developed. Addressing the potential toxicity of the rGS2 intermediate, tartronate semialdehyde, could provide a step in the right direction. One possibility would be to recycle glyoxylate through a different pathway that does not involve any aldehyde as intermediate, or at least that does not allow for its accumulation. One such pathway involves the formation of hydroxypyruvate from glyoxylate using a serine-glyoxylate aminotransferase. A second step catalyzes by hydroxypyruvate reductase can then converts hydroxypyruvate to glycerate falling back in the rGS2 cycle. If this approach is successful cyanobacteria may be reconsidered for engineered production of acetyl-CoA and its derivatives like KIC from carbon dioxide using both the rGS2 and the CBB cycle.

5.5 Tables and Figures

Plasmid	Genotypes	Reference
pCDFDuet	SpecR; CDF ori; PT7::MCS	Novagen
pCDF-ppc	SpecR; CDF ori; PT7::ppc_{EC} (his-tagged)	This work
pCDF-ppcC	SpecR; CDF ori; PT7::ppc_{CG} (his-tagged)	This work
pXL55	SpecR; NSI targeting; PLlacO1::$aclB_{CT}$.icl_{RE}.$sucCD_{MC}$.mcl_{RS}	This work (Chapter 4)
pFD210	KanR; NSII targeting; PTrc::frd_{TB}.pyc_{BS}-PLlacO1::acn_{EC}.mdh_{EC}.$fumC_{SE}$	This work (Chapter 4)
pFD250	SpecR; NSI targeting; PLlacO1:: $sucC_{MC}$.$sucD_{MC}$.mcl_{RS}-PLlacO1::ppc_{EC}.mdh_{EC}	This work
pFD258	SpecR; NSI targeting; PLlacO1:: $sucC_{MS}$.$sucD_{MC}$.mcl_{ME}-PLlacO1::ppc_{CG}.mdh_{EC}	This work
pXL145	KanR; NSII targeting; PLlacO1:: $garK_{EC}$.$glxR_{EC}$.gcl_{RE}	This work

SpecR: Spectinomycin resistant; KanR: Kanamycin resistant. In plasmid genotypes, subscripts indicate the source of the gene: EC, *Escherichia coli*; CG, *Corynebacterium glutamicum*; CT, *Chlorobium tepidum*; RE, *Ralstonia eutropha*; MC, *Methylococcus capsulatus*; RS, *Rhodobacter sphaeroides*; TB, *Trypanosoma brucei*; BS, *Bacillus subtilis*; SE, *Synechococcus elongatus*; ME, *Methylobacterium extorquens*.

Table 5-1. Plasmids list.

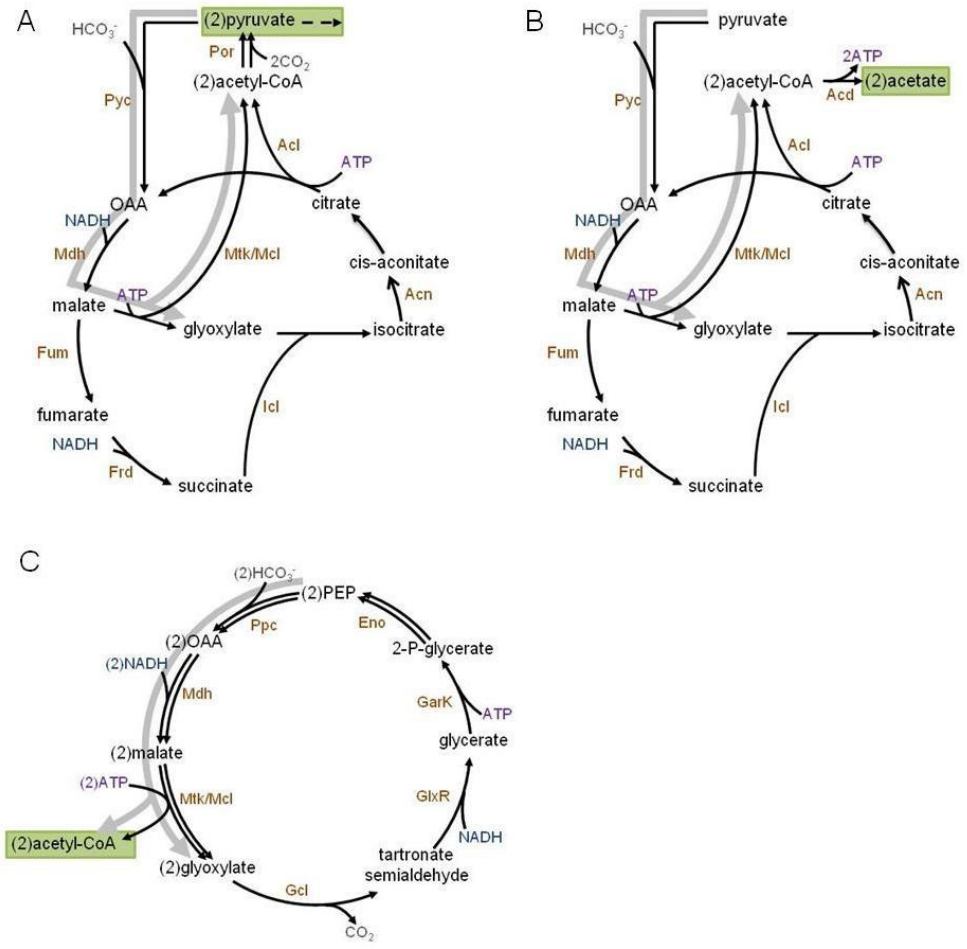


Figure 5-1. Schematic representation of (A) the reverse Glyoxylate Shunt (rGS) cycle; (B) the Linear reverse Glyoxylate Shunt (LrGS) and (C) the reverse Glyoxylate Shunt 2 (rGS2). Products of each pathway are highlighted in green. Grey arrows show the reverse Glyoxylate Shunt 0.5 (rGS0.5) in each panel, which produces acetyl-CoA and glyoxylate. Pyc, pyruvate carboxylase; Mdh, malate dehydrogenase; Acn, aconitase; Fum, fumarase; Frd, fumarate reductase; Mtk, malate thiokinase; Mcl, malyl-CoA lyase; Icl, isocitrate lyase; Acl, acetyl-CoA lyase; Por, pyruvate ferredoxine oxidoreductase; Acd, acetate-CoA ligase ADP-forming, Ppc, phosphoenolpyruvate carboxylase; Gcl, glycerate carboligase; GlxR, tertronate semialdehyde reductase; GarK, glycerate 2-kinase; Eno, enolase.

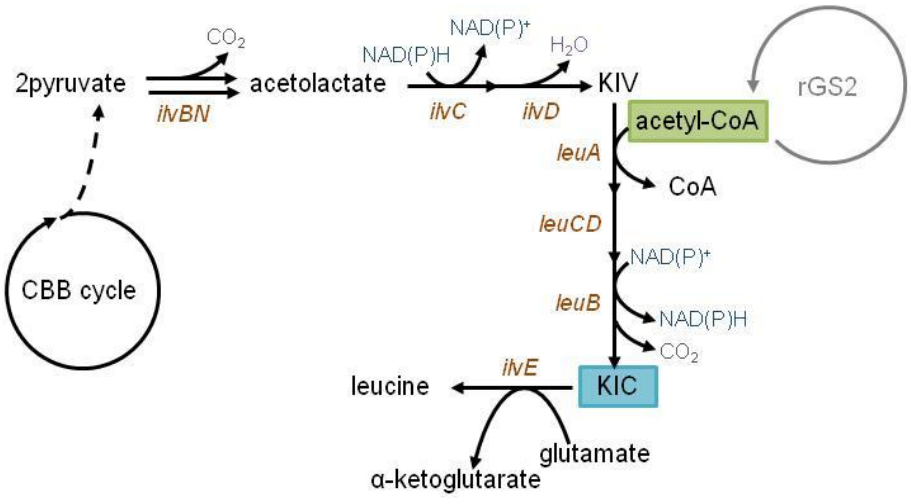


Figure 5-2. Schematic representation of α-ketoisocaproate (KIC) biosynthesis *in Synechococcus elongatus* PCC. 7942. KIV, α-ketoisovalerate.

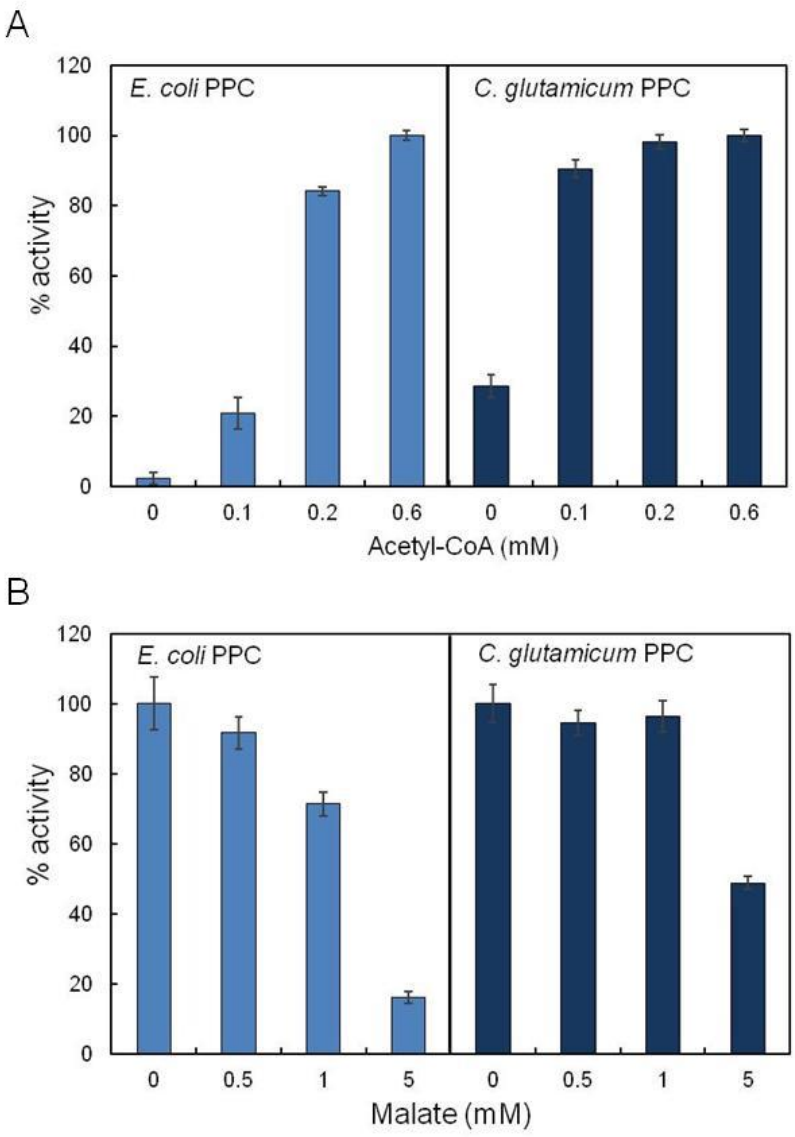


Figure 5-3. Effect of (A) acetyl-CoA and (B) malate on *E.coli* and *C.glutamicum* phosphoenolpyruvate carboxylase (PEP). Reactions were performed as described in Materials and Methods. Reactions were performed in triplicate. Activity is reported as percentage of activity with 0.6mM acetyl-CoA (A) and no malate (B).

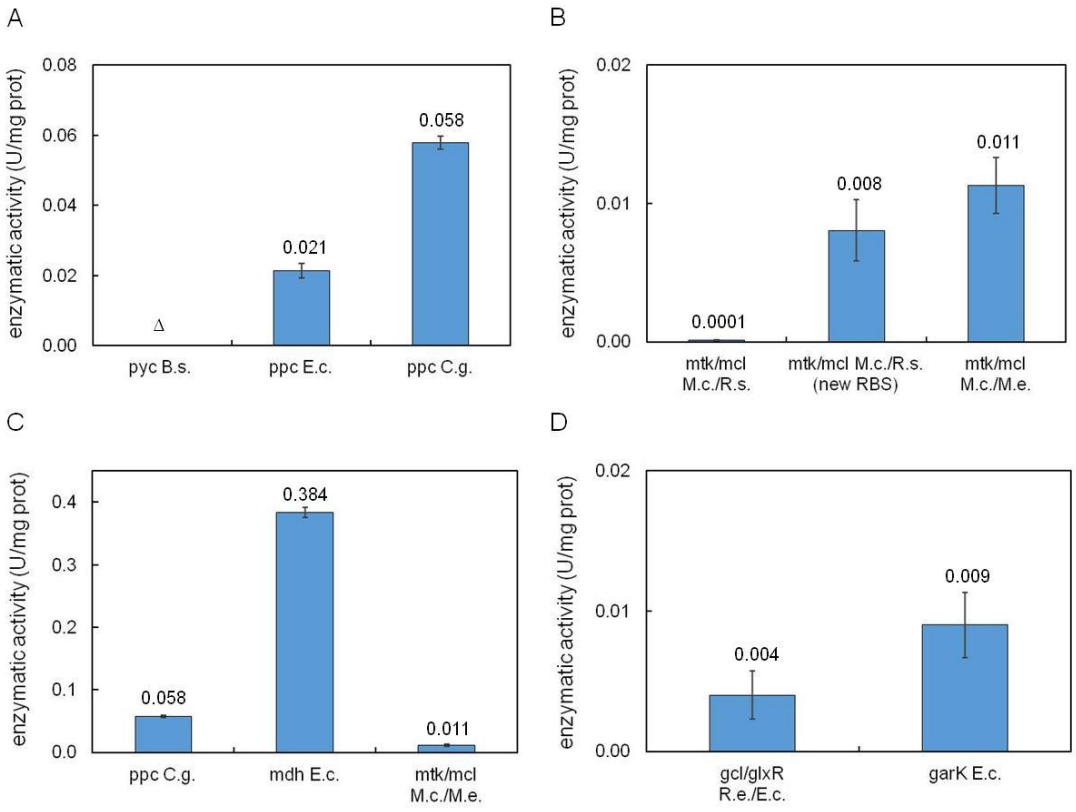


Figure 5-4. Activity of different enzymes expressed as part of the rGS, rGS0.5 or rGS2 in whole cell lysate of *Synechococcus*. Pyc, pyruvate carboxylase; ppc, phosphoenolpyruvate carboxylase; mtk/mcl, malate thiokinase and malyl-CoA lyase; mdh, malate dehydrogenase; gcl/glxR, glyoxylate carboligase and tatronate semialdehyde reductase; garK, glycerate 2-kinase; B.s., *Bacillus subtilis*; E.c., *Escherichia coli*; C.g., *Corynebacterium glutamicum*; M.c., *Methylococcus* capsulatus; R.s., *Rhodobacter sphaeroides*; M.e.; *Methylobacterium extorquens*; R.e., *Ralstonia eutropha*. Δ activity of *pyc* from *Bacillus subtilis* could not be detected in whole cell lysate. (new RBS) indicates the replacement of the native ribosomal binding site by a well define *E. coli* RBS in between *sucC* and *sucD*, the two genes cloning for *M. capsulatus* mtk. Error bar shows standard deviation for n=3.

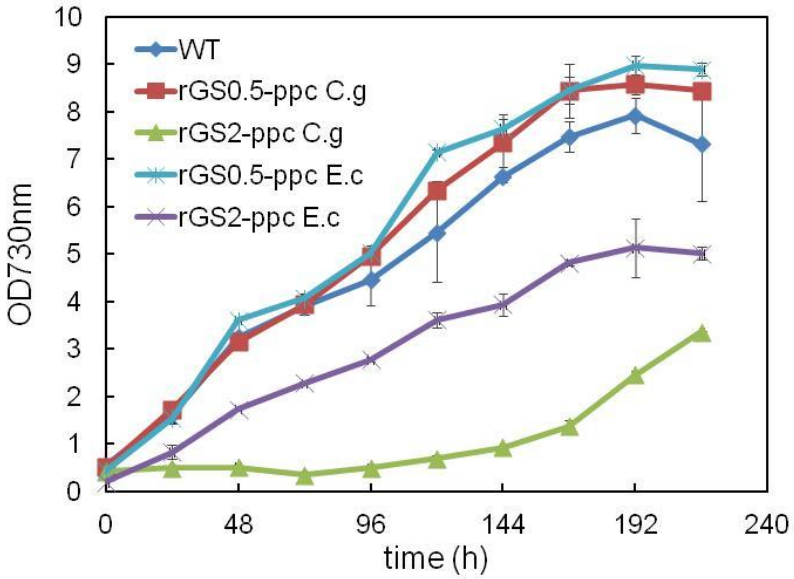


Figure 5-5. Effect of two phosphoenolpyruvate carboxylase (ppc) in rGS2 strains. Cell density as a function of time of rGS2 strains expressing *E. coli ppc* or *C. glutamicum ppc*, and rGS0.5 expressing *C. glutamicum ppc*. Cells were induced at day 0 as described in Materials and Methods.

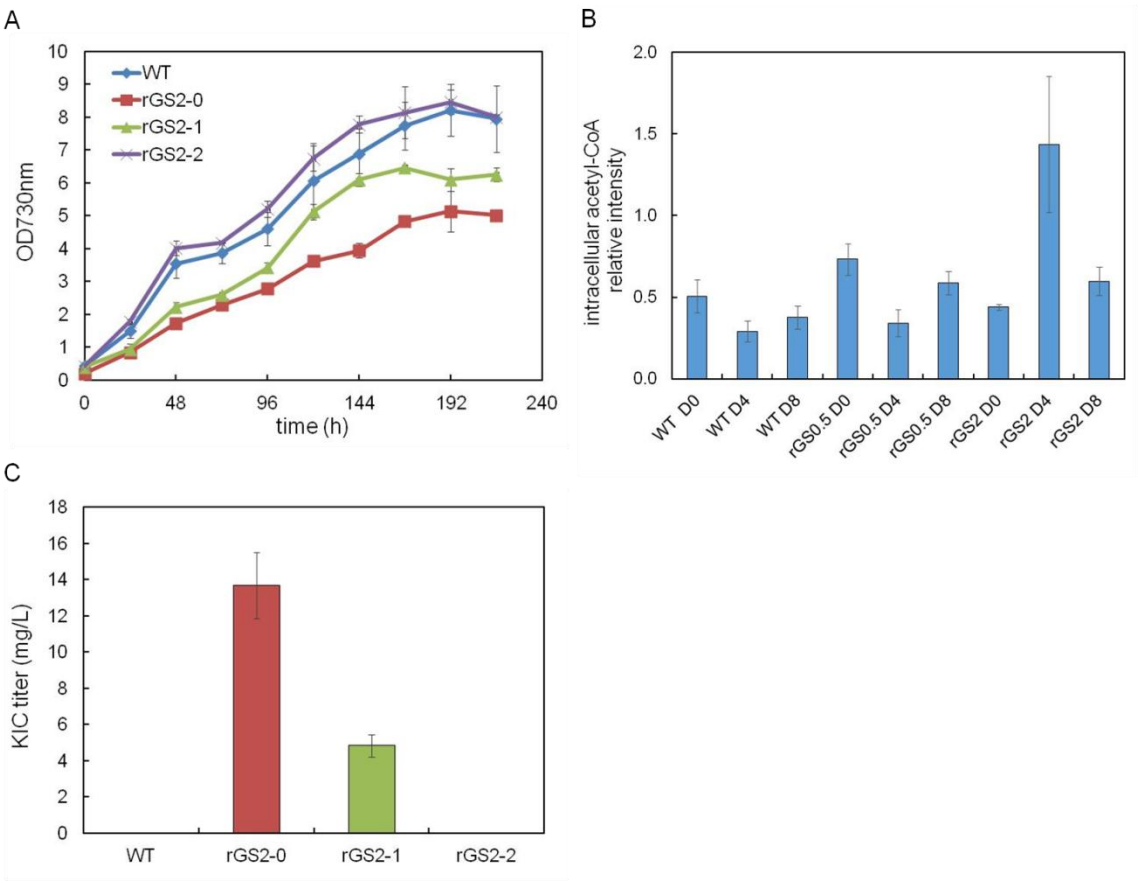


Figure 5-6. (A) Cell density as a function of time of rGS2 strains (expressing *E. coli ppc*) compared to WT. (B) Intracellular CoA relative intensity of acetyl-CoA in WT, rGS0.5 and the corresponding rGS2 strain (rGS2-0, expressing *E. coli ppc*) at day 0, 4 and 8. (C) Maximum alpha-ketoisocaproate (KIC) accumulation of rGS2 strains observed at day 8 after induction. The -0, -1 and -2 show the instability of the rGS2 strain and -1 indicates a variant from colonies streaked out from the original transformed culture (-0), while -2 indicates a variant from colonies streaked out from the culture of rGS2-1. Error bar shows standard deviation for n=3.

5.6 References

1. Kim, B. H. & Gadd, G. M. *Bacterial Physiology and Metabolism*. (Cambridge University Press, 2008).

2. Eisenhut, M. *et al.* The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 17199–17204 (2008).

3. Husic DW, Husic DH & Tolbert NE. The oxidative photosynthetic carbon cycle or C2 cycle. *Crit Plant Sci* 45–100 (1987).

4. Norman, E. G. & Colman, B. Purification and Characterization of Phosphoglycolate Phosphatase from the Cyanobacterium Coccochloris peniocystis1. *Plant Physiol.* **95,** 693–698 (1991).

5. Aparicio, M. *et al.* Keto acid therapy in predialysis chronic kidney disease patients: final consensus. *J. Ren. Nutr. Off. J. Counc. Ren. Nutr. Natl. Kidney Found.* **22,** S22–24 (2012).

6. Holecek, M. Three targets of branched-chain amino acid supplementation in the treatment of liver disease. *Nutr. Burbank Los Angel. Cty. Calif* **26,** 482–490 (2010).

7. Chang, J. H. *et al.* Influence of ketoanalogs supplementation on the progression in chronic kidney disease patients who had training on low-protein diet. *Nephrol. Carlton Vic* **14,** 750–757 (2009).

8. Teschner, M. & Heidland, A. in *Therapy of Renal Diseases and Related Disorders* (eds. Suki, W. N. & Massry, S. G.) 675–695 (Springer US, 1991).

9. van Someren, K. A., Edwards, A. J. & Howatson, G. Supplementation with beta-hydroxy-beta-methylbutyrate (HMB) and alpha-ketoisocaproic acid (KIC) reduces signs and symptoms of exercise-induced muscle damage in man. *Int. J. Sport Nutr. Exerc. Metab.* **15,** 413–424 (2005).

10. Cooper, A. J. L., Ginos, J. Z. & Meister, A. Synthesis and properties of the .alpha.-keto acids. *Chem. Rev.* **83,** 321–358 (1983).

11. Connor, M. R. & Liao, J. C. Engineering of an Escherichia coli strain for the production of 3-methyl-1-butanol. *Appl. Environ. Microbiol.* **74,** 5769–5775 (2008).

12. Connor, M. R., Cann, A. F. & Liao, J. C. 3-Methyl-1-butanol production in Escherichia coli: random mutagenesis and two-phase fermentation. *Appl. Microbiol. Biotechnol.* **86,** 1155–1164 (2010).

13. Bückle-Vallant, V., Krause, F. S., Messerschmidt, S. & Eikmanns, B. J. Metabolic engineering of Corynebacterium glutamicum for 2-ketoisocaproate production. *Appl. Microbiol. Biotechnol.* **98,** 297–311 (2014).

14. Song, Y. *et al.* One-step biosynthesis of α-ketoisocaproate from l-leucine by an Escherichia coli whole-cell biocatalyst expressing an l-amino acid deaminase from Proteus vulgaris. *Sci. Rep.* **5,** (2015).

15. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

16. Bustos, S. A. & Golden, S. S. Light-regulated expression of the psbD gene family in Synecbococcus sp. strain PCC 7942: evidence for the role of duplicated psbD genes in cyanobacteria. *Mol. Gen. Genet. MGG* **232,** 221–230 (1992).

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

18. Andersson, C. R. *et al.* Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.* **305,** 527–542 (2000).

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

20. Malate Dehydrogenase - Assay. Available at: http://www.worthington-biochem.com/mdh/assay.html. (Accessed: 10th March 2016)

21. Noguchi, S. *et al.* Quantitative target analysis and kinetic profiling of acyl-CoAs reveal the rate-limiting step in cyanobacterial 1-butanol production. *Metabolomics* **12,** (2016).

22. Jitrapakdee, S. *et al.* Structure, Mechanism and Regulation of Pyruvate Carboxylase. *Biochem. J.* **413,** 369–387 (2008).

23. Morikawa, M., Izui, K., Taguchi, M. & Katsuki, H. Regulation of Escherichia coli Phosphoenolpyruvate Carboxylase by Multiple Effectors In Vivo: I. Estimation of the Activities in the Cells Grown on Various Compounds. *J. Biochem. (Tokyo)* **87,** 441–449 (1980).

24. O'Regan, M. *et al.* Cloning and nucleotide sequence of the phosphoenolpyruvate carboxylase-coding gene of Corynebacterium glutamicum ATCC13032. *Gene* **77,** 237–251 (1989).

25. Mainguet, S. E., Gronenberg, L. S., Wong, S. S. & Liao, J. C. A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab. Eng.* **19,** 116–127 (2013).

26. Yan, Y., Lee, C.-C. & Liao, J. C. Enantioselective synthesis of pure (R,R)-2,3-butanediol in Escherichia coli with stereospecific secondary alcohol dehydrogenases. *Org. Biomol. Chem.* **7,** 3914–3917 (2009).

27. Yao, L., Qi, F., Tan, X. & Lu, X. Improved production of fatty alcohols in cyanobacteria by metabolic engineering. *Biotechnol. Biofuels* **7,** 94 (2014).

28. Takahama, K., Matsuoka, M., Nagahama, K. & Ogawa, T. Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the psbAI locus. *J. Biosci. Bioeng.* **95,** 302–305 (2003).

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

30. Jones, P. R. Genetic Instability in Cyanobacteria – An Elephant in the Room? *Front. Bioeng. Biotechnol.* **2,** (2014).

31. Davies, F. K., Work, V. H., Beliaev, A. S. & Posewitz, M. C. Engineering Limonene and Bisabolene Production in Wild Type and a Glycogen-Deficient Mutant of Synechococcus sp. PCC 7002. *Front. Bioeng. Biotechnol.* **2,** (2014).

6. Engineering of the reverse glyoxylate shunt cycle in *Arabidopsis thaliana*

6.1 Introduction

As the world fossil fuels limited reserves are being depleted, more and more attention is given to plant-based energy sources. But in recent years, serious doubts have been raised about crop-based fuels feasibility. Plants and plant waste from various sources can be used to produce automotive fuels, but technologies for low-cost production of advanced biomass-based biofuels are limited. Processes to convert plant matter to fuel are ineffective as is photosynthesis when compared to photovoltaic cells. Even the most efficient crops converts only 0.5% of solar energy into usable chemical energy[1]. In a direct comparison, photovoltaic systems can convert energy and store it in chemical bonds at about a 10% efficiency[2]. Added to the fact that most biofuels production involves growing the feedstock, plant-based biofuels became a large consumer of cropland. In a world that has a limited cropland capacity; every acres of land designated for fuel production is one less acre available to grow food crops. This leads to higher food prices or tropical deforestation to make more space, both of which lead to adverse effects on the economy or environment. Improving the yield potential of plants might thus present a better answer for energy crop to become more competitive and attractive. To date, improving photosynthetic efficiency has only played a minor role in improving the yield potential of crops, may it be for food or energy.

In plants, carbon fixation occurs through the Calvin–Benson-Bassham (CBB) cycle, and more specifically through ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), the key enzyme of this cycle. Its low turnover number and poor specificity between CO_2 and O_2 makes him inefficient. Furthermore, the cycle contains many phosphorylated intermediates and thus needs a

significant amount of energy in the form of ATP to power this pathway. Previous studies, in metabolic engineering, have principally focused on overexpression of CBB cycle steps to improve CO_2 fixation[3], but results were mixed and inconclusive. This suggests that improving carbon fixation without fundamentally changing the structure of the native CBB cycle seems unlikely to be successful.

In this study we propose to implement a new artificial carbon fixation cycle, the reverse glyoxylate shunt (rGS) into the plant model *Arabidopsis thaliana* that will be working with the existing CBB cycle to improve efficiency (Fig. 6-1). Part of the cycle has already been demonstrated to be functional when tested in *Escherichia coli* in a previous study[4]. The new CO_2 fixation cycle fixes carbon with only 30% of the ATP requirement compared to the native Calvin cycle and does not utilize Rubisco for carbon fixation, thus will not have the carbon and energy loss associated with photorespiration. The rGS also channels carbon flux directly towards pyruvate, a direct precursor of biofuels, instead of sugars like the CBB cycle does. The new pathway needs only 3 ATP to produce pyruvate, compare to 7 for the CBB cycle.

To demonstrate the functionality of the new pathway and facilitate the identification of successful transgenic lines we decided to implement the rGS into WT *Arabidopsis* and into mutant strains defective in their endogenous CBB cycle. We chose two distinct lines: a sedoheptulose-1-7 bisphosphatase (SBPase) T-DNA insertion line[5] , and a Rubisco (RBCS) double mutant[6]. Both mutants have T-DNA insertion in essential genes of the CBB cycle and have shown strong phenotypes marked by growth retardation.

The transgenic lines that express all the foreign rGS genes are expected to rescue at least partially the impaired mutants' phenotype. We hypothesized that the addition of a functional new carbon fixation cycle should help recover growth in plants with a damaged native CBB cycle.

This strategy should allow for easier identification among the transformant population as the effect of rGS on WT *Arabidosis* is unknown. This project represents the first effort to rewire carbon fixation in plants.

6.2 Materials and Methods

6.2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich or Fisher Scientifics unless otherwise specified. KOD xtreme DNA polymerases were purchased from EMD Millipore and used for gene amplification from genomic or plasmid DNA. Genotyping from gDNA was performed using Mytaq HS Red Mix from Bioline. T5-Exonuclease (Epicenter), Taq DNA ligase (New England Biolabs) and *Phusion* DNA polymerase (Fisher) were purchased individually and used to make the Gibson assembly master mix (AMM) for cloning as described[7].

6.2.2 Plant material and growth conditions

Plants of Arabidopsis thaliana ecotype Columbia were used in all experiments. For the complementation of CBB mutants experiments, the two background lines used were an SBPase T-DNA insertion line described by Liu et al.[5] (obtained by the Arabidopsis Biological Resource Center), and a Rubisco double mutant termed rbcs1A X rbcs3B[6] recreated in our laboratory for this study. T-DNA insertion lines of rbcs1a (GABI_608_F01) and rbcs3b (SALK_117835) were obtained from the Nottingham Arabidopsis Stock Centre and the ABRC respectively. The homozygous mutant lines of RBCS1A and RBCS3B were identified by genotyping and the double mutant of both genes was generated by reciprocal sexual crossing.

All seeds were surface-sterilized with 20% bleach for 15 min, washed with distilled water, and plated on ½ MS medium (0.8% agar) with or without 1% sucrose. Plates were kept at 4°C for 3 days before seeds were germinated and transplanted to soil. Plants are grown in SunGro-Metro-Mix 350 into 3.5-inch-square pots and cultivated in a controlled-environment chamber (Percival Scientific, IA, USA) at 120-140mmol photons m-2 s-1 under conditions of 14 h day and 10 h night at 21°C during day time and 19°C during night time. For root growth experiments, seeds on plates were geminated, grown vertically for 10 days.

6.2.3 Plasmid constructs and plant transformation

For overexpression of the rGS pathway, genes required for each step were split into two binary vectors: p318, a DS-RED containing vector provided by Metabolix, and pMol-E, a binary vector constructed for this study. This plasmid contains two in-planta selection markers (Basta^R and Kanamycin^R), an *Escherichia coli* origin of replication and selection marker (ori6K and ampicillin) and a single restriction site AscI where various inserts will be inserted.

Each gene selected for expression of the rGS in *Arabidopsis* were first cloned independently in pMol-E as part of an expression cassette in open-reading frame with one promoter (CaMV 35S or superpromoter), a transit peptide to the chloroplast (11 variations, each containing a unique restriction site), and followed by a terminator (4 variations). The expression cassettes were then cloned in groups into three subclones: pMolA, pMolB and pMolD. The respective genes for each constructs are presented in Table 6-1. Briefly pMolA cloned for aconitase (ACN), malate dehydrogenase (MDH), fumarase (FUM) and fumarate reducatse (FRD), pMolB for acetyl-CoA lyase (ACL) and pyruvate ferredoxine oxidoreductase (POR), and pMolC for malate thiokinase (MTK), malyl-CoA lyase (MCL), isocitrate lyase (ICL) and pyruvate carboxylase (PYC). In a

second round of cloning subclones B and D were combined into binary vector p318 to obtain pDS-BD31. Plasmid pMolA was renamed pBR6.A.

All plasmids were assembled using the Gibson DNA assembly method[7]. For the initial rounds of assembly the fragments were all amplified by PCR. However the larger plasmids were built sequentially by adding a unique restriction site (ascI) after the first gene. This allowed the vector to be opened and assembled with the second gene in such a way as to destroy the restriction site at the front and regenerate it at the back of the gene. This method allowed for iterative assembly of the larger vectors.

Plasmids pDS-BD31 and pBR6.A were transformed into A*grobacterium tumefaciens* (LBA 4404) and three independent cotransformation experiments (pBR6+pDS31) were performed in WT plants, SBPase suppressor lines and Rubisco suppressor lines. Independent sequential transformation was also performed. Transgenic plants were generated by a floral dip method[8] and first screened on solid plates containing 50 mg L^{-1} kanamycin.

6.2.4 Screening of the rGS transgenic lines, propagation and genotyping

Positive transgenic lines (T1 generation) in all background strains (WT, SBPase and Rubisco suppressor lines) were first identified on selection plates, before being screened for both Basta resistance on ½ MS medium and on soil, and for fluorescence with a Ds-Red filter. Positive strains were then checked for transcripts levels of all the rGS genes transformed. Semi-quantitative RT-PCR and quantitative two-steps RT-PCR were performed. Total RNA was isolated from leaf samples. Complementary DNA was produced using 1 µg total RNA with the Qiagen cDNA synthesis kit (Qiagen, Hilden, Germany) and the cDNA was subsequently used as

a template for qPCR with gene-specific primers. The plant specific EF4A2 (At1g54270) gene was used as a control for constitutive gene expression.

Seeds from T1 plants showing promising growth over parental lines (WT, SBPase and Rubisco suppressor lines) and positive for rGS genes (termed WrGs, SrGs and RrGS respectively) were grown into T2 and T3 generation to perform segregation analysis. PCR on gDNA using Mytaq was used to genotype the T3 plants. Quality of the gDNA and PCR protocol was confirmed through detection of native genes: PEP carboxylase, phosphoribulokinase, acetyl-CoA carboxylase, mevalonate kinase, the mitochondrial NADH dehydrogenase subunit 5 and the chloroplast photosynthetic electron transport A. SBPase and Rubisco disruption were confirmed with T-DNA-specific oligonucleotides and two oligonucleotides flanking the putative T-DNA insertion site. PCR on gDNA was also used for further analysis of the inserted rGS T-DNA fragment. The longer PCR fragments were sequenced by primer walking.

6.2.5 Double blinded experiment

The best lines of WrGS, SrGS and RrGS were picked based on growth parameters like plant biomass, and plant height to perform a double blinded assay and reevaluate the parameters for all strains at once. WT, SBPase and Rubisco suppressor lines were used as controls. Each line was grown initially on ½ MS Media supplemented with 1% sucrose to have uniform seedlings. Four days old homogenous seedlings were transferred to soil (11-15 per line) and identified through a triple coding system so that their true identity remained hidden until the end of the experiment. The plants were grown in soil in growth chambers until full maturity (70 days for WrGS and RrGS lines, and 120 days for SrGS lines) where leaf area and plant height were measured.

Watering of the plants was then stopped to let them uniformly go to senescence. The plants were then harvest and fresh and dry biomass was recorded.

6.3 Results

6.3.1 Characterization platform lines for rGS evaluation

To assess the potential of the rGS pathway in term of improved plant growth, we chose to implant this cycle in a background line with a defective CBB cycle and marked phenotype. If carbon fixation is impaired in plants we would expect to see some kind of growth retardation, which would easily allow for screening of complementation by the rGS pathway. We decided to use two different background strains: a sedoheptulose-1-7 bisphosphatase (SBPase) T-DNA insertion line described by Liu et al[5] , and a Rubisco (RBCS) double mutant termed rbcs1A X rbcs3B recreated in this study based on Izumi et al[6]. Both mutants have T-DNA insertion in essential genes of the CBB cycle and had been described to show growth retardation. *Arabidopsis thaliana* Col-0 T-DNA insertion line (SALK_130939) at the SBPase locus (AT3G55800) was originally obtained by the Arabidopsis Biological Resource Center (ABRC). For convenience, we renamed this SALK_130939 line SBPase. Homozygous mutant lines were isolated by genomic PCR and their phenotype was confirmed. Homozygous seedlings had small primary roots, and almost no lateral roots. Once transferred into soil, homozygous lines showed less chlorophyll, less leaf area, delayed flowering, along with significant growth arrest although heterozygous plants appears normal. Seed setting was also compromised in the SBPase homozygous plants. They could nevertheless survive, flower and set seeds.

The Rubisco double mutant, termed rbcs1AXrbcs3B, was created by sexual crossing between homozygous mutant lines of rbcs1a (GABI_608_F01) and rbcs3b (SALK_117835) as described

by Izumi et al. Each homozygous line has a T-DNA insertion impairing one of the four Rubisco small subunits (RBCS) described in *Arabidopsis*, RBCS1A and RBCS3B respectively. The two T-DNA-inserted lines were obtained from the Nottingham Arabidopsis Stock Centre (rbcs1a) and the ABRC (rbcs3b), and homozygous mutant lines of rbcs3b and rbcs3b were isolated by genotyping. Semi-quantitative RT-PCR and qRT-PCR analysis in rbcs1A, rbcs3B and rbcs1aXrbcs3B mutant lines confirmed that rbcs1A was a null mutant and rbcs3b a knockdown mutants like described previously. The individual mutants showed almost no growth difference when compared to wild-type (WT) at the same age, while the double mutant growth rate was significantly compromised. The time to bolting was also delayed in the double mutant, from 23 days after germination in WT or the single mutants, to 40 days. Marked chlorosis was also noticeable in the leaves, as well as dark green veins indicating altered carbon and nitrogen assimilation and partitioning.

Both the SBPase and Rubisco double mutant showed strong phenotypes linked to an impairment of critical enzymes of the CBB cycle. They both represent a desirable kind of mutant to test a potential rescue by a new carbon fixation pathway like the rGS. We decided to test rGS implementation in both these strains as well as in WT plants.

6.3.2 Generation and genotyping of rGS mutants in 3 backgrounds strains

Genes coding for the full rGS pathway were introduced into both homozygous CBB mutants (SBPase and rbcs1aXrbcs3B, termed RBCS hereafter) as well as in WT plants. This was accomplished through transformation of the two plasmids pDS-BD31 and pBR6.A, either sequentially or simultaneously. 422 transformants were screened for Basta resistance and DS-Red selection marker (Fig. 6-2), and 9-10 lines per parental strains were selected based on

promising growth over parental lines. Selected T1 lines were screened to confirm the presence of all rGS genes in their genome. Reverse-transcribed cDNA from purified RNA was used as template for semi-quantitative RT-PCR. Out of the 11 genes transformed into WT, SBPase and RBCS mutant lines, transcripts of 10 genes were detectable (Fig. 6-3). Aconitase transcripts could not be detected in any of the lines. This might be due to a potential low sensibility of the primers set used for that gene. Even when the plasmid pBR6 (positive control) was used as template, only a faint band could be detected. No detectable transcripts of FRD, FUM, MTK (*sucD* gene) and PYC were present in the WT (negative control) plants like expected, but their presence was confirmed in all the transgenic rGS lines tested. Selected lines were grown in the T2 generation and segregation analysis was performed on these plants before quantifying various growth parameters such as plant biomass, and plant height (results presented below). Based on these parameters we further selected 9 lines in total, 3 transgenic lines in the SBPase mutant background (SrGS), 2 in the RBCS mutant background (RrGS) and 4 in WT (WrGS) that demonstrated improved grown compared to their parental line. These lines were gorwn in T3 generation and their phenotypic properties were further evaluated through a double-blinded experiment (results presented below) and genotyping was performed from purified gDNA. The evaluation of rGS genes transcripts done on T1 plants used primers set detecting only short gene sequence (approx. 200bp). Confirmation of the full rGS genes length integrating in the mutants' genome would give a better indication of the new pathway contribution in the mutants' phenotype. Genotyping of full rGS genes turned out to be difficult in all T3 mutant strains tested, even though native genes were easily detectable. Although some heterologous genes could be detected by PCR occasionally, results were not always reproducible throughout plants from the same line (data not shown). Full genome sequencing was also unsuccessful and thus inconclusive

(Illumina, Manhattan Beach, CA). We decided to analyze the inserted T-DNA fragment of the 9 selected strains using a different approach. Instead of targeting each heterologous gene individually by PCR, we started from one known integrated gene and tried to PCR all possible adjacent sequences. We knew DS-red was integrated in that strain as fluorescence was detectable, and PCR of the full gene using purified gDNA confirmed it. A forward primer was design in the DS-red gene and multiple reverse primers binding on all other genes present on pDS-BD31 were pooled in different reaction. Any PCR product obtained was then sequenced using primer walking starting at the DS-red gene. This new method was successful for one WrGS strain: WrGS-58 (Fig. 6-4). Out of the 6 rGS genes originally in pDS-BD31, only part of the PYC gene was integrated in the plant genome. DS-red, GUS, as well as some of the plasmid backbone sequences were also confirmed. These results indicate a strong instability of the T-DNA in our transgenic lines, which could explain why some genes present in the T1 generation cannot be detected in T3, or only in some individual plants of the same line. The large length of the T-DNA, as well as some repeated sequences like promoters might be at the origin of the DNA instability. We could imagine that recombination may have occurred between homologous regions of the T-DNA looping out some parts of the DNA fragment.

6.3.3 Phenotype of SBPase complementary lines

The segregation analysis performed on selected SrGS lines grown in T2 generation showed segregation as per Mendelian ration (Fig. 6-5). The best rGS transgenic lines that showed partial complementation of the SBPase homozygous mutant background were selected (SrGS-12A, -4D, -3A). The plants exhibited a larger number of leaves and higher plant height (30-34 cm), almost 5 times compare to the SBPase mutant which is only 6-7 cm tall at the same physiological age

(Fig. 6-6A). Root growth was also significantly higher when plants were grown on non-sugar plates (Fig. 6-6B). Growth rate of these complemented lines was still very slow at the early growth phase and they attained maximum plant biomass and height only 120-140-d post germination whereas WT plants reached it much earlier (80-90-d post germination). At 120 days both growth parameters, biomass and height, were quantified (Fig. 6-7). Two lines out of the three tested (srG4D and srG3A) showed significant increase in biomass and plant height compare to the control sbp-/- . The results indicate that CO_2 fixation through the rGS pathway might be active in the two complemented SrGS mutants, showing partial rescue compared to the parental strain deficient in SBPase-dependent carbon assimilation.

6.3.4 Phenotype of Rubisco complementary lines

Analysis on selected RrGS lines (RrGS-53A and -101A) grown in T2 generation showed segregation according to Mendel's laws (Fig. 6-8). All RrGS lines showed significant improvement in plant growth over the rbc-/- mutant, especially in the early growth stage (Fig. 6-9A). Rosette area was also increased even though the number of leaves remained unchanged compare to the control (Fig. 6-9C-D). The two complemented lines tested had increase dry biomass (20-25%) but no significant increase in height compare to their parental line (Fig. 6-10). The highest plant biomass can be explained by the widest rosette area observed in the RrGS lines. Collectively these data suggest that the rGS pathway is functional and could be working as an optional pathway for CO_2 fixation. Carbon fixed through the new pathway could complement the lowest CO_2 assimilation rate observed in the rbc -/- mutant[6], and increase total biomass.

6.3.5 Phenotype of WT *Arabidopsis* supplemented with the rGS cycle

Combined impact of rGS and an uncompromised Calvin Cycle was also evaluated in plants through examination of WrGS strains which are WT *Arabidopsis* with the full set of rGS genes integrated. The most promising line selected were WrGs-19C, -3, -33, -58. Above ground and below ground plant height as well as aerial biomass were measured. All selected WrGS lines screened in our growth conditions showed significant increase in plant growth, both aerial and below ground, compare to WT plants (Fig. 6-11). Likewise, we also noticed a significant increase in plant biomass (20%) in all the lines and in plant height (5-15%) in three out of four strains (Fig. 6-12). These results suggest the presence of an active rGS cycle in the WrGS strain, which by fixing additional carbon compare to plant with the CBB cycle alone (WT), improves plant growth.

6.4 Discussion and Conclusion

Introducing an entirely new carbon fixation cycle such as rGS in plants requires the transformation of 11 non-native genes. The timeline and milestones of the PETRO program under which this project was funded prevented a traditional approach in which each step would be implemented individually and validated sequentially. Therefore, all necessary genes needed to be delivered simultaneously, in a single transformation event. The probability for the resulting transformants to functionally express all the genes concertedly is expected to be low, due to stochastic gene silencing events. A reliable and effective screening method was thus desirable to allow for easy identification of transgenic lines successfully expressing the new CO_2 fixation pathway. We selected, on top of WT *Arabidopsis*, two individual strains described in the literature to serve as screening platforms for a functional rGS: a SBPase mutant and a RBCS double mutant. Both strains have a damaged endogenous CBB cycle resulting in lower carbon

assimilation rate. They present strong phenotypes with impaired growth. We generated 422 rGS transformants in the three parental lines and chose the most promising lines in term of partially rescued growth for further characterization. Out of the selected lines we identified at least two strains per parental line with the rGS overexpressed showing improved phenotype. Two line of the SBPase mutant transform with rGS showed both enhanced plant biomass and height. In the RBCS mutant only biomass was significantly affected. Two strains had higher plant biomass than their parental line and showed wider rosette area. Finally, WT plants transformed with rGS allowed to examine the combined impact of rGS and a functional Calvin Cycle. At least three lines exhibited an increased in plant biomass and height. Collectively these results suggest that the rGS pathway may be functional and could be working as an optional pathway for CO_2 fixation. Unfortunately genotyping of the T3 plants, used for the final phenotype characterization, turned out to be difficult. The present of all rGS genes in the mutant lines could not be confirmed. Full genome sequencing did not give any usable results, and detection by PCR of individual rGS gene was nor reproducible between different generation of the same line, neither throughout plants of the same generation in one strain. One specific rGS mutant line was shown to have only the selection marker gene, some plasmid backbone and half of one of the rGS gene integrated at this site of the genome. Integration of the non-T-DNA part of the binary vector (the vector backbone) has previously been reported and investigated[9]. On the other hand, and to the best of our knowledge, instability of such big fragment of integrated T-DNA has not been explained yet. Usually, silencing of transgenes occurs at the RNA level[10]. It could be possible that recombination between homologous regions, like promoters, on the T-DNA caused some of the fragment to loop out. This could explain why some genes detected in the T1 generation disappeared in the T3, and why genotyping between plants of the same line is not

consistent. Nevertheless, some T3 plants still demonstrate improved phenotype compared to their parental strains. These are encouraging results towards creating plants with an improve carbon fixation ability. To understand the cause behind our observed phenotype, it would be beneficial in the future to study the overexpression of each rGS gene individually and in smaller group. This approach would allow identifying what part of the rGS pathway, if some, contributed to the highest plant biomass observed. If the full pathway needs to be implemented, caution should be taken to avoid homologous sequences on the T-DNA, and all genes might have to be split between more binary vectors.

6.5 Tables and Figures

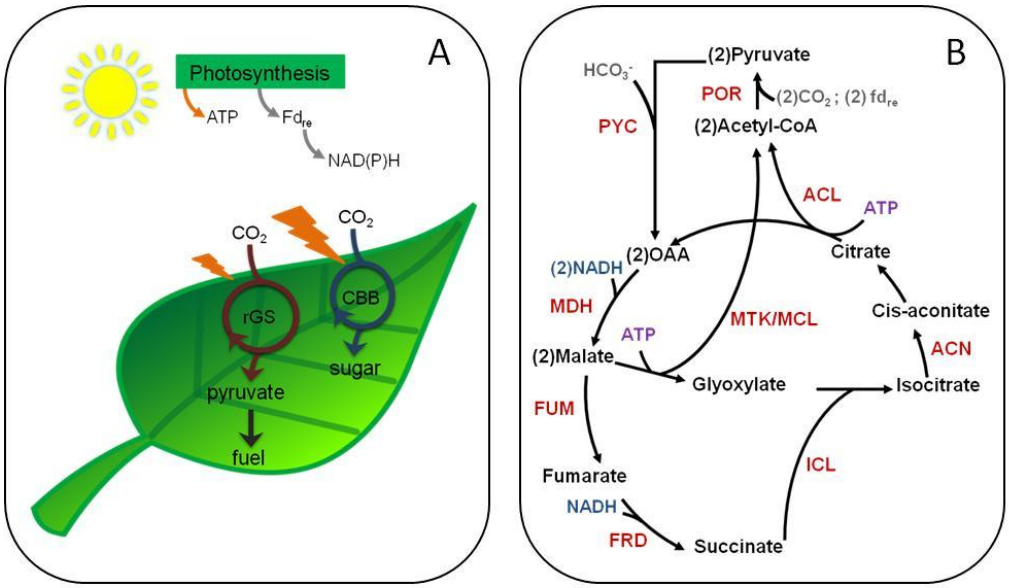


Figure 6-1: (A) Schematized summary of this study: a new carbon fixation cycle, the reverse glyoxylate shunt (rGS), will be clone in *Arabidopsis* chloroplasts. The pathways fix carbon more efficiently than the Calvin–Benson-Bassham (CBB) cycle and produces pyruvate, a direct fuel precursor, instead of sugar. Both the CBB cycle and the rGS cycle use energy and reducing power from photosynthesis, but the rGS requires less ATP when comparing production of one molecule of pyruvate. Fd_{re}, reduced ferredoxin. (B) The rGS cycle and its enzymes: PYC, pyruvate carboxylase; MDH, malate dehydrogenase; ACN, aconitase; FUM, fumarase; FRD, fumarate reductase; MTK, malate thiokinase; MCL, malyl-CoA lyase; ICL, isocitrate lyase; ACL, acetyl-CoA lyase; POR, pyruvate ferredoxine oxidoreductase.

pMolA					
promoter	Transit peptide/restriction site	enzyme	*gene*	origin	terminator
35S	TP09-At opt / BlnI	ACN	*aco1*	*Arabidopsis* ACO1	OCS
35S	TP01-At opt / BglII	MDH	*mdh* (NADP)	*C. reinhardtii*	ADH1
AmasPmas	TP02-At opt / SwaI	FUM	*fumC*	*Synechocystis* sp.PCC 6803	HSP
35S	TP03-At opt / BstENI	FRD	*frdS*	*S. cerevisiae*	OCS
pMolB					
promoter	Transit peptide/restriction site	enzyme	*gene*	origin	terminator
AmasPmas	TP16 / AgeI	ACL	*acl*	*H. sapiens*	UBQ5
35S	TP16 / SanDI	POR	*nifJ*	*Synechocystis* sp.PCC 6803	ADH1
pMolC					
promoter	Transit peptide/restriction site	enzyme	*gene*	origin	terminator
35S	TP04-At opt / StuI	MTK	*sucC*	*M. capsulatus*	ADH1
35S	TP04-At opt / StuI	MTK	*sucD*	*M. capsulatus*	ADH1
AmasPmas	TP04-At opt / XbaI	MCL	*mcl*	*M. extorquens*	HSP
35S	TP07-At opt / HpaI	ICL	*iclA*	*R. eutropha*	OCS
AmasPmas	TP16 / XmaI	PYC	*pyc*	*L. lactis*	UBQ5

Table 6-1: Summary of the subclones used to construct the rGS binary vectors. The table indicates the genes contained in the different subclones pMolA, pMolB, pMolC, as well as the origin of the genes, their respective promoter, transit peptide and terminator. Two different promoters were used in front of transit peptide: the original Cauliflower Mosaic Virus 35S Promoter (35S) and the mannopine promoter with its synthase transcriptional activating element (AmasPmas). ACN, aconitase; MDH, malate dehydrogenase; FUM, fumarase; FRD, fumarate reductase; ACL, acetyl-CoA lyase; POR, pyruvate ferredoxine oxidoreductase, MTK, malate thiokinase; MCL, malyl-CoA lyase; ICL, isocitrate lyase; PYC, pyruvate carboxylase.

A

B

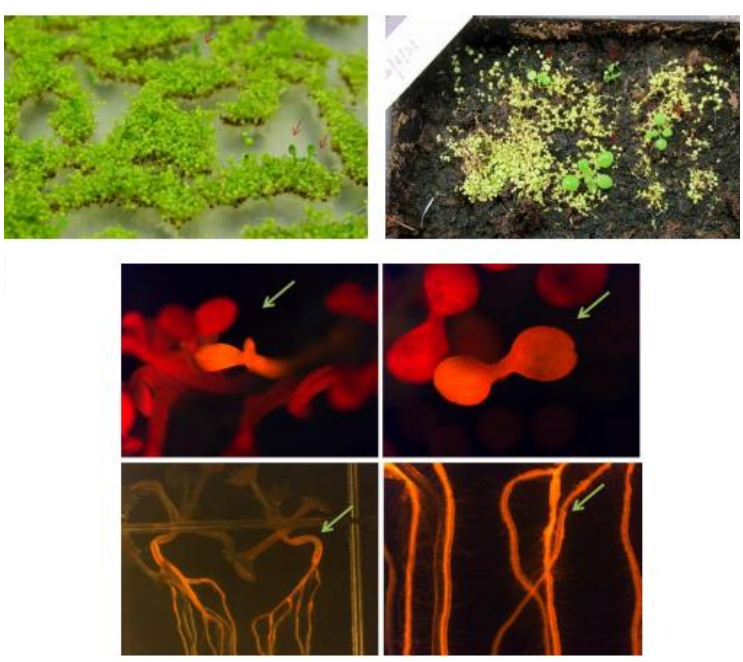


Figure 6-2: *Arabidopsis* WT and CBB mutants co-transformed with pBR6 and pDS-BD31 showing basta resistant transgenic lines on ½ MS medium and on soil (A) and fluorescent light with DS-Red filter (B).

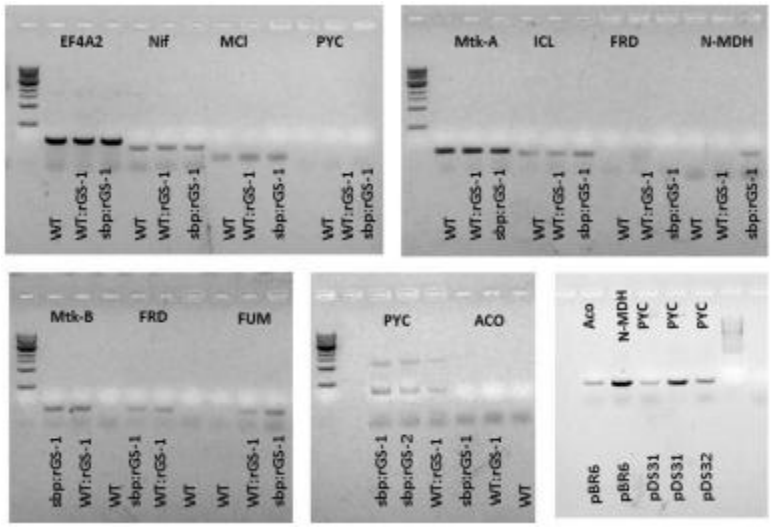


Figure 6-3. Semi-quantitative RT-PCR for determining rGS transcripts in T1seedlings of WT and SBPase suppressor lines transformed with the rGS enzymes (WT:rGS and sbp:rGS). MTK (*sucD* gene), FRD, PYC and FUM transcripts were only detected in sbp:rGS and WT:rGS lines but not in WT, confirming that rGS transgenic lines have detectable transcripts of various genes encrypted by plasmids pBR6 and pDS-BD31. The plant specific EF4A2 (At1g54270) gene was used as a control for constitutive gene expression.

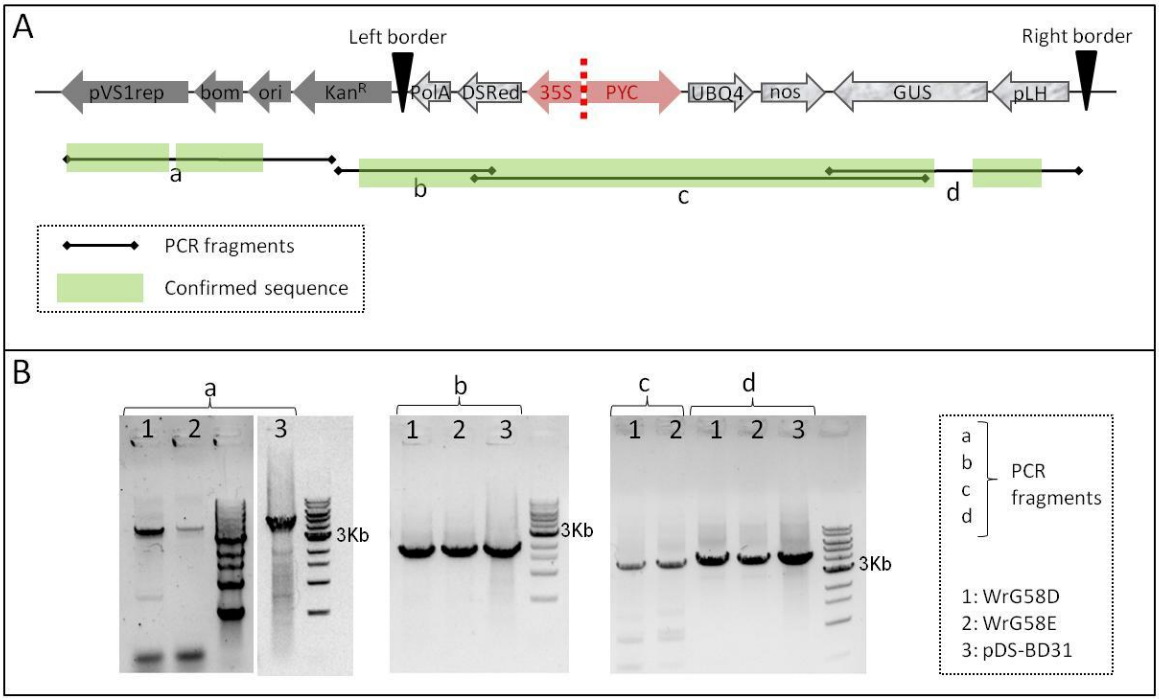


Figure 6-4: (A) Schematization of the fragment of DNA that was confirmed to be inserted in the genome of T3 plants from a line of WT *Arabidopsis* transformed with the full rGS genes: the line WrGS58 (plant WrGS58D and WrGS58E). Four PCR fragments could be amplified (a, b, c, d) and the segments highlighted in green were confirmed by sequencing. Part of the binary vector backbone (genes represented in dark grey) was integrated in the gDNA of both plants. From the T-DNA, a stretch of approximately 20Kb was missing between the truncated 35S promoter of DSRed and about half of the PYC gene. The genes coding for ACL, POR, MTK, MCL and ICL were all missing. (B) PCR results for the 4 fragments mentioned in panel A. The plasmid pDS-BD31 was used as a control. Fragment c could not be amplified on the plasmid as it would be much larger: five more rGS genes are cloned between the 35S promoter of DSRed and the PYC gene. pVS1 rep = region for stability in Agrobacterium; bom = site for transmobilization of vector from *E. coli* to *Agrobacterium*; ori = origin of replication in *E. coli*; Kan^R = kanamycin resistance gene; polA = poly(A) tail; DSRed = reporter gene (red fluorescent protein); 35S =

cauliflower mosaic virus 35S promoter; PYC = pyruvate carboxylase; UBQ4 = ubiquitin terminator ; nos = 3’ end of nopaline synthase gene; Gus = β-glucuronidase gene; pLH = promoter of the oleate 12-hydroxylase (LFAH12) gene.

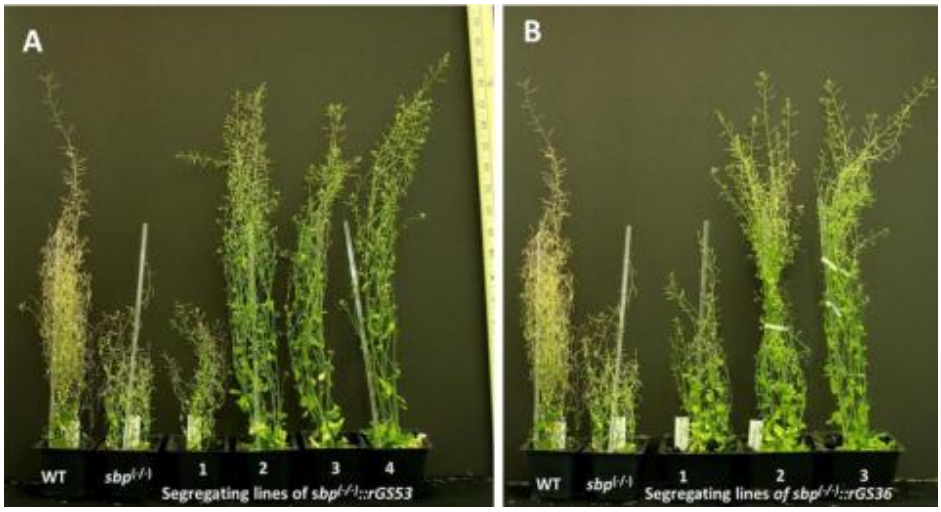


Figure 6-5: Segregation analysis of two independent SrGS lines (sbp$^{(-/-)}$::rGS53 and sbp$^{(-/-)}$::rGS36) segregating in T2 generation as per Mendelian ratio. Plant 1 of both segregating lines shows sbp(-/-0 phenotype and plant 2-4(A), or 2-3 (B) are complemented lines. All complemented lines grow higher that the control (sbp-/-) and reach WT height, although only 120-140 days post germination whereas WT plants reach this stage much earlier (80-90 days post germination).

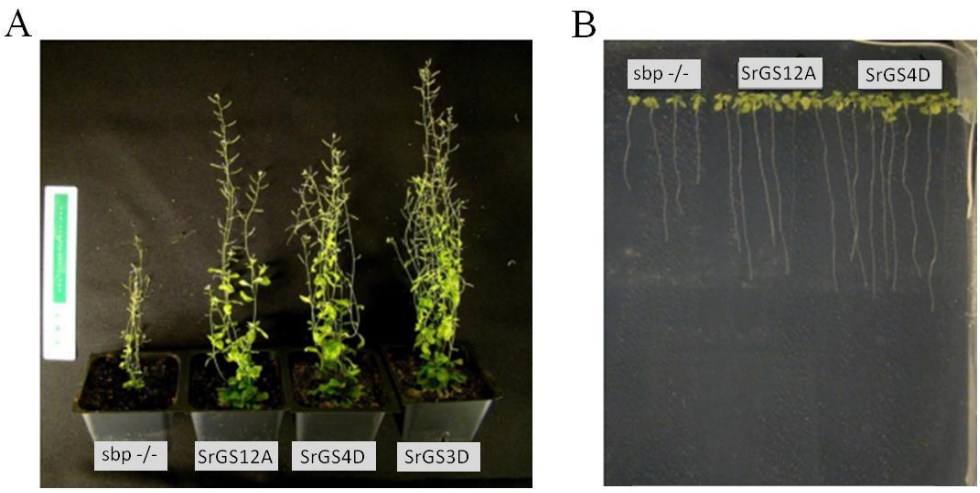


Figure 6-6: (A) Complementation of the SBPase mutant (sbp-/-) by the rGS pathway in three separate SrGS lines (SrGS12A, -4D and -3A). The SrGS lines show a rescued growth phenotype compare to their parental strain. (B) Comparative root growth study of the SBPase mutant and SrGS lines on no-sugar (1/2 MS) plates. Roots of both SrGS12A and SrGS4D grow longer than the mutant ones.

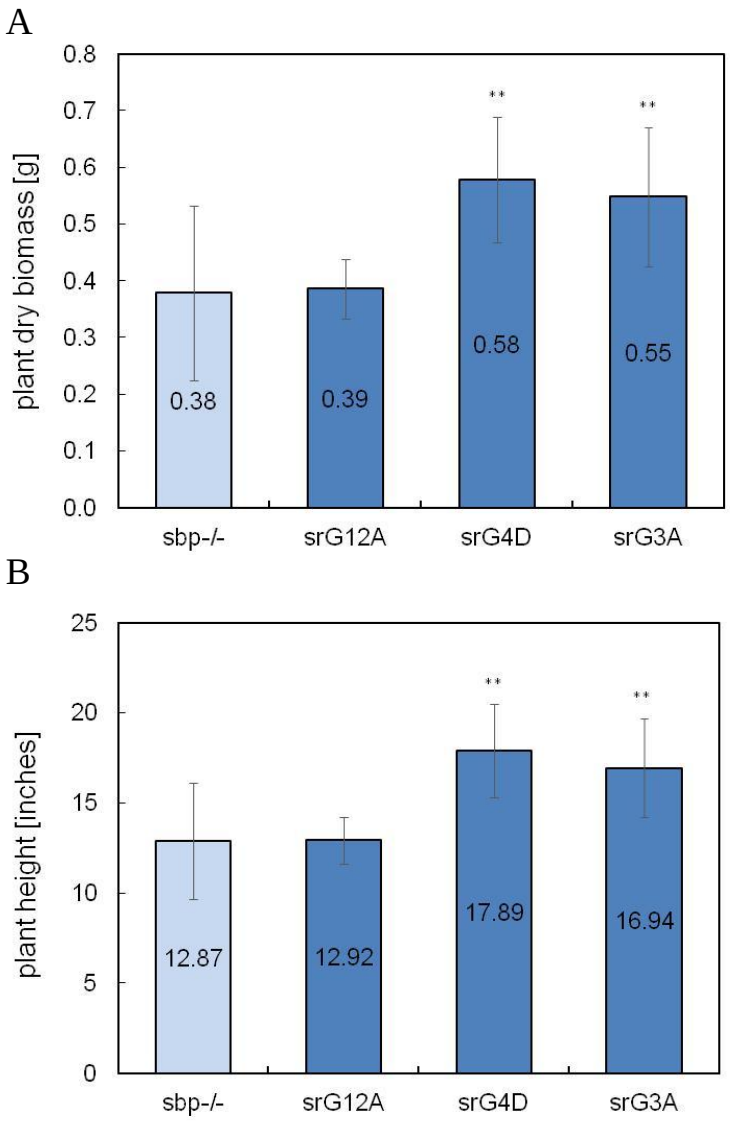


Figure 6-7. Comparative growth analysis of SrGS and its parental strain, the SBPase suppressor line (sbp-/-). Two lines (srG4D and srG3A) showed significant increase in biomass (A) and plant height (B) compare to the control sbp-/-. Plant height was measured at 120 days before harvesting the plants for biomass measurement. * $p<0.05$ and ** $p<0.01$ according to Student's t-test ($n \geq 12$).

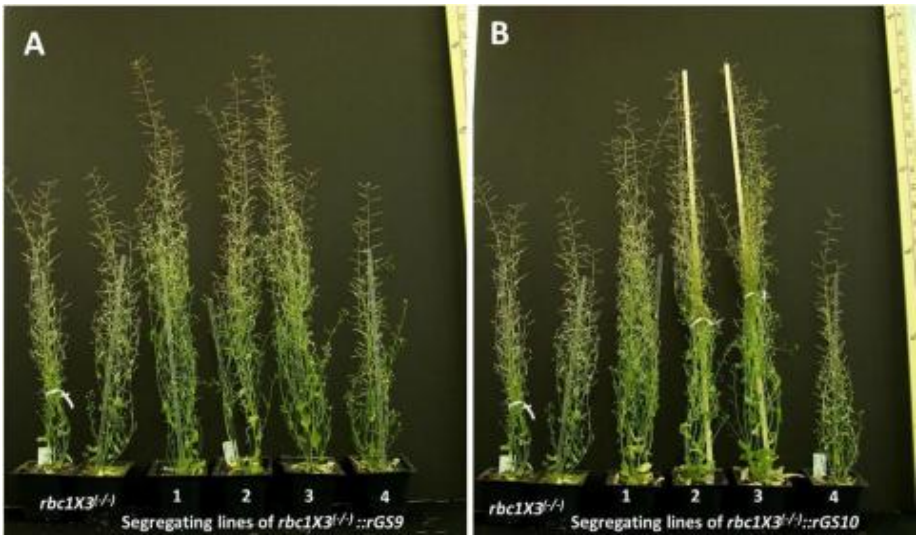


Figure 6-8. Segregation analysis of two independent RrGS lines (rbc1X3$^{(-/-)}$::rGS9 and rbc1X3$^{(-/-)}$::rGS10). Plant 4 of the segregating lines are rbcs-/- and plant 1-3 are complemented lines in both panels (A and B). Plant 4 shows similar height to the two far left plants representing rbc-/- mutants of the same age. RrGS lines show significant improvement in plant biomass and plant height compare to their parental line.

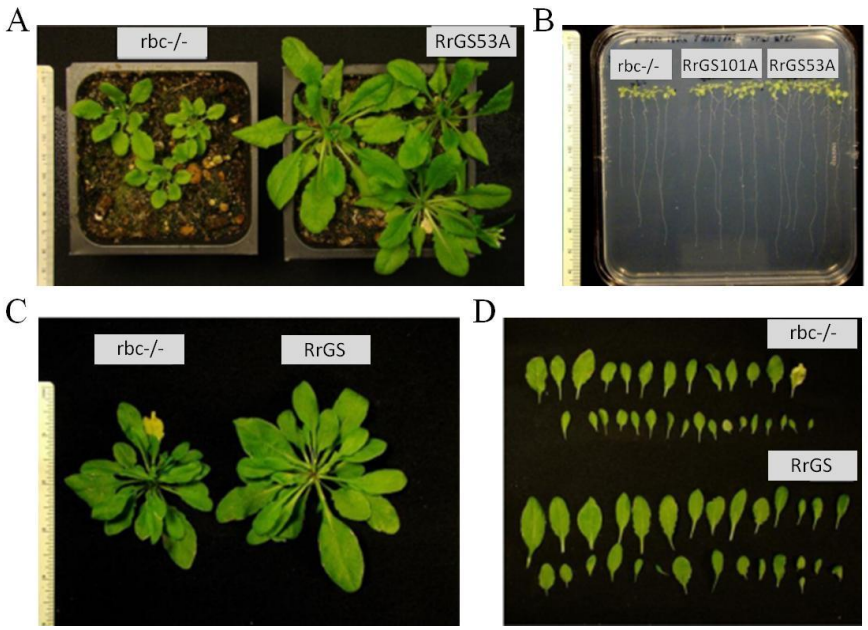


Figure 6-9. Comparative early seedling growth of the RrGS lines on soil (A) and ½ MS Media without sucrose (B). All seedlings on plate were initially germinated and grown for 4 days on 1% sucrose for homogenous plants growth and later the seedlings were transferred to no-sugar plates. RrGS plants show wider total rosetta area (C) but total number of leaves remain unchanged (D).

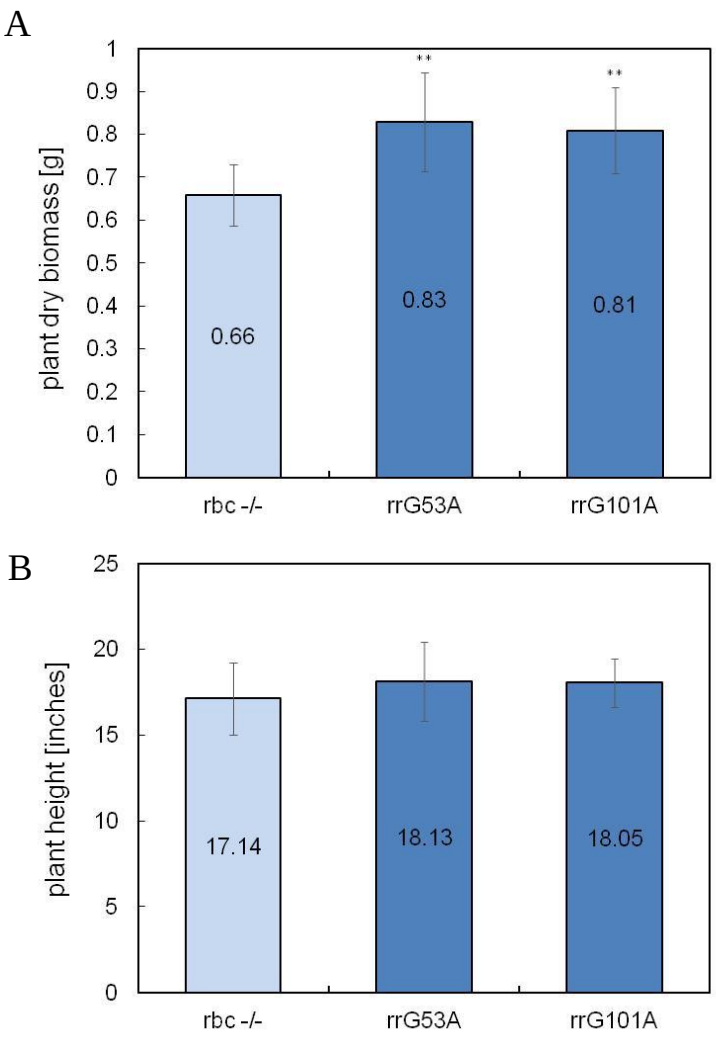


Figure 6-10. Comparative growth analysis of RrGS and its parental strain, the Rubisco suppressor line (rbc-/-). Plants showed significant increase in biomass (A) but not in plant height (B) compared to rbc-/-. Plant height was measured at 70 days before harvesting the plants for biomass measurement. * $p<0.05$ and ** $p<0.01$ according to Student's t-test (n=11-12).

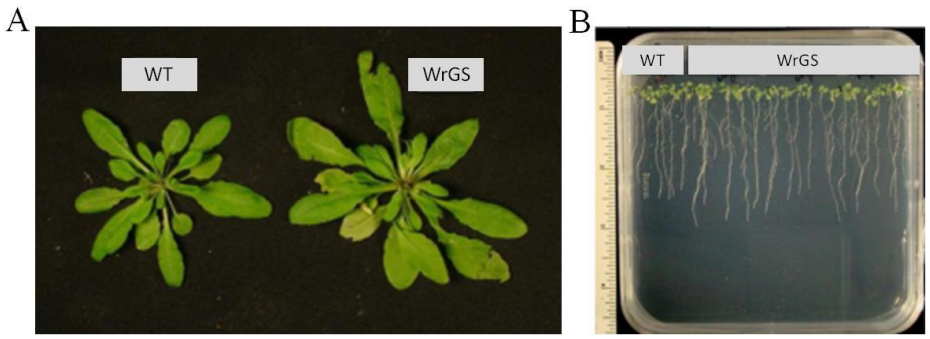


Figure 6-11. Comparative growth analysis of wrGS and WT plants shows a larger total rosette area in WrGS (A) and longer roots (B).

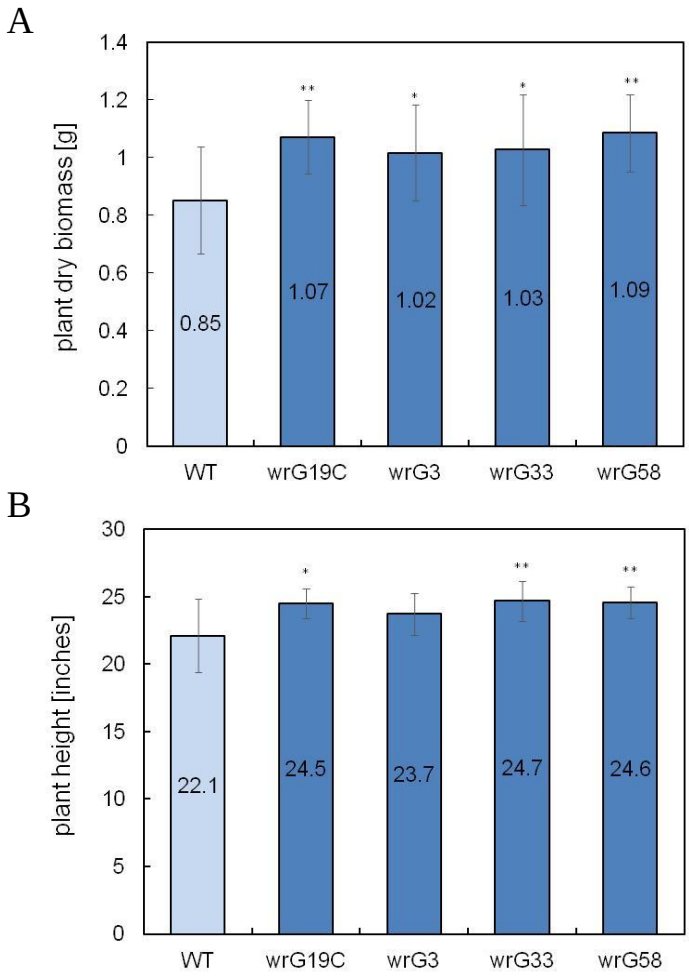


Figure 6-12. Comparative growth analysis of WrGS and WT plants showed significant increase in Biomass (A) and plant height (B) of all WrGs lines compared to WT. Plant height was measured at 70 days before harvesting the plants for biomass measurement. * $p<0.05$ and ** $p<0.01$ according to Student's t-test ($n \geq 11$).

6.6 References

1. Brown, L. R. in *Plan B 4.0: Mobilizing to Save Civilization* (W.W. Norton & Company, 2009).

2. Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *science* **332,** 805–809 (2011).

3. Ruan, C.-J., Shao, H.-B. & Silva, J. A. T. da. A critical review on the improvement of photosynthetic carbon assimilation in C3 plants using genetic engineering. *Crit. Rev. Biotechnol.* **32,** 1–21 (2012).

4. Mainguet, S. E., Gronenberg, L. S., Wong, S. S. & Liao, J. C. A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab. Eng.* **19,** 116–127 (2013).

5. Liu, X.-L., Yu, H.-D., Guan, Y., Li, J.-K. & Guo, F.-Q. Carbonylation and Loss-of-Function Analyses of SBPase Reveal Its Metabolic Interface Role in Oxidative Stress, Carbon Assimilation, and Multiple Aspects of Growth and Development in Arabidopsis. *Mol. Plant* **5,** 1082–1099 (2012).

6. Izumi, M., Tsunoda, H., Suzuki, Y., Makino, A. & Ishida, H. RBCS1A and RBCS3B, two major members within the Arabidopsis RBCS multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. *J. Exp. Bot.* **63,** 2159–2170 (2012).

7. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

8. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J. Cell Mol. Biol.* **16,** 735–743 (1998).

9. Kononov, M. E., Bassuner, B. & Gelvin, S. B. Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J. Cell Mol. Biol.* **11,** 945–957 (1997).

10. Schubert, D. *et al.* Silencing in Arabidopsis T-DNA Transformants: The Predominant Role of a Gene-Specific RNA Sensing Mechanism versus Position Effects. *Plant Cell* **16,** 2561–2572 (2004).