Design and Improvement of the Biosynthesis of 2,3-Butanediol from CO_2 by Metabolic Engineering of Cyanobacterium *Synechococcus elongatus* PCC7942

By

JOHN WILLIAM KIDDER OLIVER

B.S. (University of California, Irvine) 2006

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Shota Atsumi (Chair)

_____

Xi Chen

_____

Michael Toney

Committee in Charge

2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Dedicated to Mum, Dad, Jess, Jen, and Mags;

in loving memory of Jack Osborn

John William Kidder Oliver
December 2014
Chemistry

Design and Improvement of the Biosynthesis of 2,3-Butanediol from CO_2 by Metabolic Engineering of Cyanobacterium *Synechococcus elongatus* PCC7942

Abstract

This dissertation describes metabolic engineering of cyanobacterium *Synechococcus elongatus* PCC7942 as a photosynthetic host for the conversion of CO_2 into 2,3-butanediol. Current advances in pathway design, genetic tool development, and yield improvement are described (Chapter 1). A pathway for the synthesis of 2,3-butanediol is designed based on collective concepts of pathway strength, robustness, and irreversibility, and extensively tested through the generation of mutants (Chapter 2). This pathway is then optimized through modulation of translation by combinatorial mixing of ribosome binding sites (Chapter 3). Finally, photosynthetic productivity is investigated through expression of an exogenous pathway targeting every step between fixation and product (Chapter 4). All materials and methods are given separately for easy reference (Chapter 5).

Acknowledgements

I am fortunate to have a great number of people to acknowledge for their help, support, and collaboration during my degree. I would like to thank first my advisor Shota Atsumi, for his teaching, discussions, constant support, and enthusiasm. I also want to say thank you to Jacquelyn Gervay-Hague who gave me training and guidance, and helped me believe in a move from chemistry to biotechnology.

Thank you to Iara Machado, who taught me everything I needed to know with infinite patience and humor, and accepted a deluge of questions and disbelieving argument on my half, with grace, erudition, and excitement. I was stubborn, and I received many gifts that I cannot pay back, but that I will endeavor to pay forwards.

Full credit for my graduation goes to my family who have supported me from the other side of the world through evolving and innovative methods, who deserve to have me come home and work on the farm, and who will have to settle for long distance again. This dissertation is dedicated to their love and support, and in loving memory of my grandfather John Osborn, who has been and will always be, one of my greatest role models.

A big thank you to Nick, Lexi, Kathy, Ben, Max, and Layne who created a world of support and friendship that otherwise would not of existed and made me a part of it.

I would not be where I am without Lisa Anderson, who gave me the ability to walk through mountains among giants. Why are we in the mountains? Because graduate school can be just like carrying baskets into the heart of the forest to pick wild raspberries. That is the secret. Thank you for your endless support and intelligent partnership.

Thank you to Hisanari Yoneda, for being a tireless team mate, and voice of reason in California, and a mentor and guide for two months in Fuji. To Masahiro Kanno, for a shared excitement in photosynthesis, and for helping me survive in Tokyo and Fuji. I want to thank the members of the Atsumi lab for an open and thrilling environment to be a scientist in. Thank you to Gabe Rodriguez and Jordan McEwen for brain pretzels and trying anything. To Nicole and Anna for problem solving discussions, dedication to the MarPro project, and trying anything. To Christy, Yohei, Shuchi, and Edna for making lab a great place. Every step of the way I was given positive constructive feedback; we celebrated the successes, brushed off the failures, debated the answers, built the equipment, and made the improbable possible.

I would like to thank Mike Toney for frank illuminating discussions, for providing guidance in modern academia, and for serving on my committee. Thank you to Xi Chen for her mentorship and for serving on my committee.

Thank you to Judy Kjelstrom for making the designated emphasis in biotechnology possible, and for creating a community of supportive and inspired students.

Special thanks to Blaine Hudson and Mike Sisto from the machine shop and facilities, who worked with me on many building projects.

I want to especially acknowledge every teacher and tutor that put effort into my education. The high of epiphany is lifelong.

Thank you

Table of Contents

List of figures

List of Tables

Copyright

Chapters 1,2, and 3 are reproduced in part from published works. The citations are as follows:

Chapter 1: Oliver, J. W. & Atsumi, S. Metabolic design for cyanobacterial chemical synthesis. *Photosynth Res* 120, 249-261, doi:10.1007/s11120-014-9997-4 (2014).

Chapter 2: Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* 110, 1249-1254, doi:10.1073/pnas.1213024110 (2013).

Chapter 3: Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Combinatorial optimization of cyanobacterial 2,3-butanediol production. *Metab. Eng.* 22C, 76-82, doi:10.1016/j.ymben.2014.01.001 (2014).

Chapter 1 Introduction

This chapter closely resembles published work. Citation: Oliver, J. W. & Atsumi, S. Metabolic design for cyanobacterial chemical synthesis. Photosynth Res 120, 249-261, doi:10.1007/s11120-014-9997-4 (2014).

1.1 Overview

Forecasts of energy shortage, increasing CO_2 levels, and surging financial support for renewable processes is driving the development of photosynthetic systems for chemical production. To meet this goal, researchers have been engineering cyanobacteria for proficient production of chemicals directly from CO_2. Cyanobacteria are a diverse group of prokaryotic Gram-negative autotrophs, that collectively are responsible for capture of large amounts of CO_2 from the open ocean, and for fixing biologically available nitrogen in the soil.[1] It is likely that cyanobacterial cells are the evolutionary precursor of chloroplasts, and their simple photosynthetic structure allows for highly efficient growth, with doubling times as fast as 4 hours.[2,3] Cyanobacteria can be found in a diverse range of biological niches which provide strains with unique capabilities including growth in high salinity, at a wide range of temperatures, and under high concentrations of $_{CO2.}$[4,5] Despite these competitive advantages, cyanobacteria naturally produce chemicals of value in only trace amounts, and instead have been cultured in labs primarily for the study of photosynthesis and circadian rhythm.[6-8] Accelerated development of metabolic engineering tools in fermentative organisms such as *Escherichia coli* has enabled adaption of these tools and pathways to alternative organisms, leading to the emergence of cyanobacteria as a photosynthetic host for chemical synthesis in a direct process.[9] Many cyanobacteria will naturally take up foreign DNA and recombine it into their genome without modification or pretreatment of cells.[10] This allows for the expression of

heterologous genes from cyanobacterial hosts. The relative simplicity of cyanobacteria as compared to eukaryotic algae has lead to faster development of engineering tools for cyanobacteria.[11]

Proof of concept pathways have been constructed for the production of more than 17 feedstock chemicals; however, differences in host environment and regulation of metabolism between cyanobacteria and model fermentative organisms such as *E. coli*, in which the majority of genetic characterization has been undertaken, have resulted in limited titers from these pathways in photosynthetic hosts (Table 1-2). Improving pathway and strain design to overcome these obstacles in cyanobacteria is an ongoing challenge for metabolic engineers. This chapter will cover the design and effect of heterologous pathways in cyanobacteria for the production of carbon based chemicals. Many factors will impact the development of cyanobacterial technologies as they move from the lab to industry. For challenges in the implementation of cyanobacterial chemical production, readers are referred to recent reviews.[12,13]

1.2 Concepts of Strain Design

Chemical production from cyanobacteria is dependent on two goals: efficient and renewable harvest of light and CO_2, and conversion of harvested light and CO_2 into chemical products in high productivity. Achieving the first goal requires one to choose a host with a robust natural photosynthetic system, as well as to design synthetic pathways that optimize photosynthetic stoichiometry and minimize loss of fixed carbon. Achieving the second goal requires metabolic control of carbon partitioning between cell growth and chemical production.

As metabolic design in cyanobacteria has progressed, key concepts have emerged for designing pathways with emphasis on increased carbon partitioning to synthetic products (**Figure 1-1**). A strong pathway requires a *driving force* (i.e.: high enzyme activity and irreversible reactions), *stable enzymes* (active in the photosynthetic environment), *low toxicity* of accumulating products (or efficient removal), and an *abundant source* of carbon substrate. The effect of these concepts in varied metabolic contexts is apparent in the production section of this chapter.

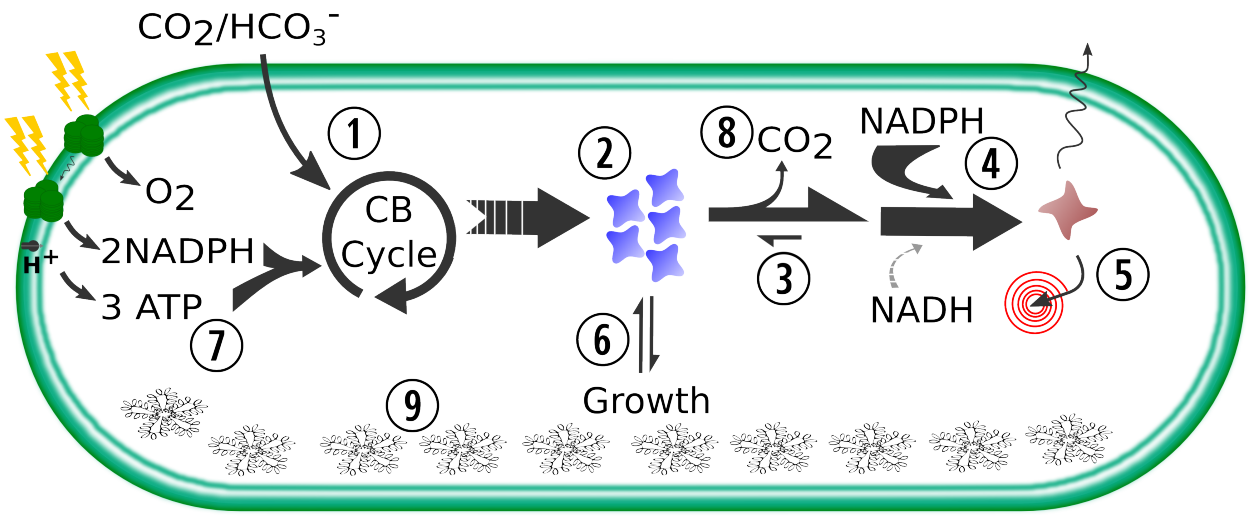


Figure 1-1 Design in Metabolic Engineering for Photosynthetic Organisms. Key concepts of design illustrated are: 1, Efficient carbon fixation; 2, Abundant metabolite pool; 3, Driving force; 4, Abundant cofactor; 5, Low toxicity; 6, Growth increases metabolite pool; 7, Stoichiometry with linear electron transport ; 8, Low carbon loss; 9, Complete segregation.

1.2.1 Driving force

Engineering photosynthetic microbes for chemical production requires two systems to function in tandem in the cell. One for growth and replication, and the other for efficient conversion of carbon to chemical products. Drawing carbon flux away from natural processes during growth requires heterologous pathways to outcompete the activity of native enzymes. This can be achieved by utilizing enzymatic steps that have low reversibility such as decarboxylation[14-16], or by coupling steps to an abundant energetic

cofactor such as ATP[17]. It has become apparent that cofactor abundance significantly impacts product titer in cyanobacteria.[14,16-18] The reducing cofactor NADPH is a direct product of photosynthesis, and is ten times more abundant than NADH in *Synechocystis* 6803.[19] Changing enzymes or expressing a transhydrogenase to utilize the more abundant cofactor increased the titer of ethanol, 1-butanol, 1,2-propanediol, and lactate pathways. Additionally, high activity of enzymes is dependent on robustness of the enzyme in the host cell environment. In many instances, enzymes that have high activity in *E. coli* are significantly less active when transferred into the photosynthetic environment of cyanobacteria, and screening for highly active enzymes in the host cell or under a range of conditions is necessary to identify useful candidate enzymes.[14,20]

1.2.2 Low toxicity

There is strong evidence for a common sense rule that toxicity of the product results in a limited titer. This can be seen by comparing titers and production rates of acetoin and 2,3-butanediol[14], acetaldehyde and ethanol[16], and acetol and 1,2-propanediol[18]. In each of these cases, a low titer of the toxic intermediate (acetoin, acetaldehyde, or acetol) was greatly increased upon addition of a final step to convert the intermediate into a less toxic product (2,3-butanediol, ethanol, or 1,2-proanediol). Efforts are also being made to identify and engineer photosynthetic organisms with increased tolerance to key toxic chemicals [21-25]. In particular, transcriptomics has identified genes regulated in response to stress from free fatty acid (FFA) production, and selective engineering of expression levels lead to an increase in FFA production relative to control strains [26]. Small unsaturated molecules appear to be particularly innocuous to cyanobacteria, and represent a potential avenue for high titer production [27].

1.2.3 Abundant carbon pool

A cyanobacterial cell competes for carbon partitioning at every step from fixation to endpoints in metabolism. If a heterologous pathway is initiated with a strong driving force it can partition carbon at key points, however, care must be taken to fulfill one of the two criteria for an abundant carbon pool. Firstly, a heterologous pathway will ideally withdraw carbon as close as possible to fixation to maximize the pool of available carbon and minimize the number of competing pathways. Secondly, as carbon is drawn away, the steps between fixation and the start of the heterologous pathway that remain catalyzed by native enzymes will ideally become upregulated. These criteria are supported by the fact that the strong production rates (and thus carbon flux) have come from pathways utilizing the central metabolite pyruvate as a starting point [14,28]. Pyruvate is only three steps from carbon fixation and is essential for amino acid synthesis and growth. Depleting the pyruvate pool through an artificial pathway creates a selection force to compensate for that depletion by increasing flux to pyruvate from carbon fixation. This would lead to a metabolic state in which the cell also provides increased starting material for the heterologous pathway. In another example, depleting the acyl-acyl carrier protein (acyl-ACP) pool limits growth, by slowing the synthesis of fatty acids for cell membranes, and leads to as much as a 100-fold increase in the synthesis of acyl-ACP per cell, allowing for increased fatty acid production.[29] The strength of natural regulation is perhaps most apparent in the endogenous pathway for sucrose. Natural carbon partitioning measured over plants, algae, and cyanobacteria shows that 5% and 10% of flux goes to terpenoid synthesis and fatty acid synthesis respectively, while the majority of carbon is directed to the synthesis of sugars [30]. In response to salt shock, cells naturally

produce sucrose as an osmoprotectant. Causing sucrose to be artificially secreted forces the cell to continually upregulate production in order to survive salt stress, leading to high titers and production rates.[31]

1.2.4 Cofactor stoichiometry and carbon balances

Photosynthesis generates ATP and NADPH in a limited ratio and consumption of these equivalents in non-stoichiometric amounts may influence the efficiency of chemical production. To calculate the stoichiometry of energy molecules involved in a biosynthetic pathway, "Gold standard" equations can be used to equate cofactors and fixed carbon.[32] It is generally agreed that absorption of four photons through linear electron flow provides one ATP and one NADPH when the Q-cycle is not active, and up to 1.5 ATP per NADPH can be produced if the cycle is active.[33] Additional ATP can be generated through the consumption of NADPH in cyclic electron flow (CEF) around photosystem I (PSI), however this requires additional photons and is thought to be active primarily during stress responses.[34] The energetic cost of carbon fixation in heterologous pathways can be calculated using as a starting point the requirement of 5 NADPH and 8 ATP per molecule of 3-phospho-D-glycerate, which includes the cost of regenerating Ribulose-1,5-bisphosphate (RuBP) substrate after carbon fixation.[32] Additional cofactor consumption in the pathway can then be included by balancing equations. This approach was used to analyze fatty acid and isoprene production[32], and has been expanded here to present ATP:NADPH consumption ratios for heterologous pathways discussed in this chapter as well as key endogenous intermediates (Figure 1-2). Most pathways fall within the dynamic ratio of 1:1 to 1:1.5 that is enabled by switching on and off the Q cycle, indicating that chemical production would not be hindered by imbalance in cofactors

during linear electron transport. Isopropanol, 3-hydroxybutyrate, and sucrose, all consume slightly more than one ATP per NADPH. The one outlier is acetone, which consumes acetyl-CoA without requiring additional reducing equivalents. This may indicate the requirement of increased CEF or other response mechanisms in this pathway. In the case of sucrose, a slight imbalance does not appear to inhibit production rates.[31]

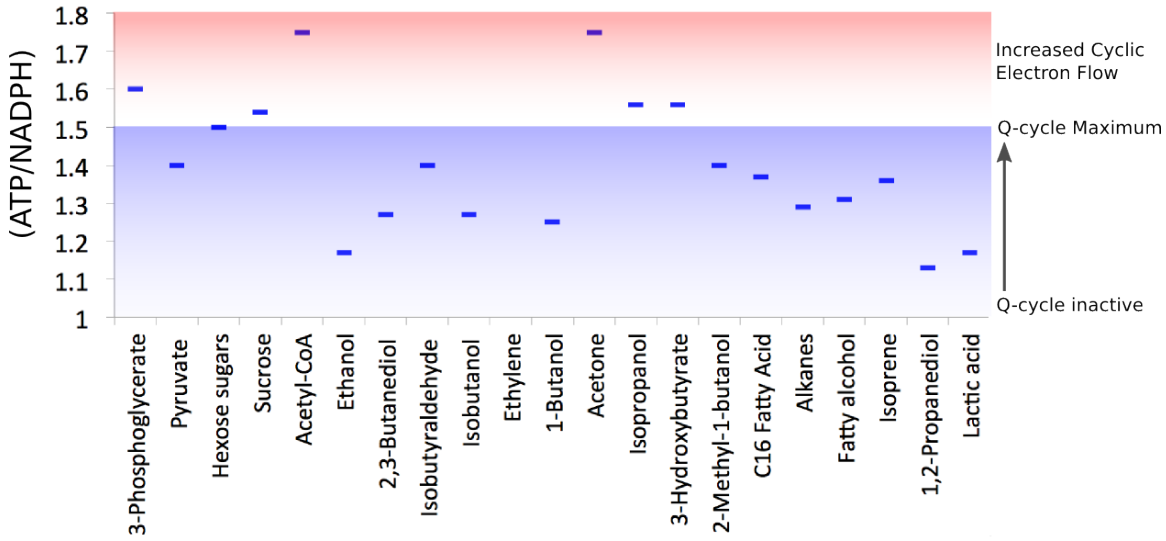


Figure 1-2 The ratio of ATP to NADPH in key heterologous and endogenous pathways. Photosynthesis can provide a stoichiometry of 1:1 to 1:1.5 for ATP:NADPH during linear electron flow.

Separate from energy and cofactor ratios, chemical production is fundamentally limited by carbon conservation. In pathways that release decarboxylated CO_2, a greater number of CO_2 molecules need to be fixed in the Calvin-Benson (CB) cycle to generate each carbon atom in the product molecule. This limits the rate at which product can be formed to a percentage of the rate of carbon fixation. A measure of carbon conservation can be calculated as the number of carbon atoms in the product molecule divided by the total number of CO_2 fixed by the pathway in question (Table 1-1). Perfect conservation of carbon is unusual in heterologous pathways and is achieved only by production of 1,2-propanediol and lactate. The greatest loss of fixed CO_2 occurs in the pathway for

ethylene, which includes several decarboxylation processes, resulting in only 33% of the carbon fixation rate remaining fixed in the product. It is worth noting that hexoses represent stoichiometric capture of carbon and cofactors, as might be expected from high carbon partitioning in the cell.[30]

Table 1-1 Carbon conservation in endogenous and heterologous pathways

	Carbon fixed in pathway	**Carbon stored in product**	**CO_2 Stored /Fixed**
3-Phosphoglycerate	*3*	*3*	*100%*
Pyruvate	*3*	*3*	*100%*
Hexose sugars	*6*	*6*	*100%*
Sucrose	*12*	*12*	*100%*
Acetyl-CoA	*3*	*2*	*67%*
Ethanol	3	2	67%
2,3-butanediol	6	4	67%
Isobutyraldehyde	6	4	67%
Isobutanol	6	4	67%
Ethylene*	6	2	33%
1-Butanol	7	4	57%
Acetone	6	3	50%
Isopropanol	6	3	50%
3-hydroxybutyrate	6	4	67%
2-methyl-1-butanol	9	5	56%
C16 Fatty Acid*	24	16	67%
Alkanes*	24	15	63%
Fatty alcohol *	24	16	67%
Isoprene	6	5	83%
1,2-propanediol	3	3	100%
Lactic Acid	3	3	100%

*HCO_3^- fixation during formation of malonyl-CoA, and during the ethylene pathway, is not included in this total as these steps do not require reducing cofactors. Including these fixation steps would lower the reported ratio of the respective pathway. Italics indicate endogenous pathways.

Table 1-2 Chemical synthesis of listed compounds from CO_2 in cyanobacteria

Host	Chemical	Rate (mg/L/h/OD)	Titer (mg/L)	Reference
7942	Ethanol	NA	230	[35]
6803	Ethanol	10	550	[28]
6803	Ethanol	NA	5500	[16]
7942	2,3-butanediol	4.5	2400	[14]
6803	2,3-butanediol	NA	420	[36]
7942	Isobutyraldehyde	3.1	1100	[15]
7942	Isobutanol	NA	450	[15]
6803	Isobutanol	NA	300	[19]
7942	Ethylene	0.38	NA	[37]
7942	Ethylene	0.53	NA	[38]
6803	Ethylene	0.19	NA	[39]
6803	Ethylene	0.45	NA	[27]
7942	1-butanol	NA	13	[40]
7942	1-butanol	NA	300	[17]
7942	1-butanol	NA	400	[20]
6803	Acetone	NA	36	[41]
7942	isopropanol	NA	27	[42]
6803	3-hydroxybutyrate	NA	530	[43]
7002	polyhydroxyalkanoate	NA	52% CDW	[44]
7942	2-methyl-1-butanol	0.56	200	[45]
6803	Free fatty acids	NA	200	[29]
6803	Alkanes	NA	26	[46]
6803	Alkanes	NA	2.3	[47]
6803	Fatty alcohol	NA	1.7	[48]
6803	Isoprene	0.01	NA	[49]
6803	beta-caryophyllene	2.8×10^{-4}	NA	[50]
7942	1,2-propanediol	NA	150	[18]
7942	Sucrose	9.5	NA	[31]
7942	Hexose sugars	NA	45	[51]
7942	Lactic acid	2.3	55	[51]
6803	Lactic acid	NA	290	[52]
6803	Lactic acid	NA	15	[53]

NA: not available; CDW: cell dry weight; all numbers have been rounded to 2 significant figures.

1.3 Cyanobacterial engineering tools

Metabolic engineering, which enables the design and installation of genetic expression systems to create biosynthetic pathways in cells, has made significant and elegant progress in model fermentative organisms such as *E. coli.*[54] The genetic tractability of

cyanobacteria makes them attractive targets for metabolic engineering as renewable technology. However, many of the genetic elements that control the central dogma of translation and transcription in cyanobacteria are not fully characterized and appear to behave differently than those in *E. coli*. *Synechococcus elongatus sp.* PCC 7942 (*S. elongatus* 7942) has a number of unique sigma factors involved in regulation in response to light and circadian rhythm[55] and numerous studies have shown disparate transcriptional expression between *E. coli* and cyanobacteria from common promoters.[56,57] Studies investigating the use of promoters for controlled protein expression in cyanobacteria are limited. To date, most cyanobacterial engineering has been based on the use of *E. coli* LacI-repressed promoters for inducible expression. In *Synechocystis sp.* PCC 6803 (*Synechocystis* 6803), these promoters function poorly[58,59], and lack of repression by LacI continues to be a problem in *S. elongatus* 7942 and *Synechocystis* 6803.[14,39,58] In *S. elongatus* 7942, expression from the LacI repressed P_{trc} promoter has been observed to attenuate over time when cells are cultured under continuous light after a dark pulse.[60] Recently a TetR repressed promoter, which is known to be inactive in *Synechocystis* 6803, was successfully modified to enable protein expression in this host, achieving a 290-fold increase in induction, with strong repression.[58] The strong constitutive promoter P_{rbc} from the RuBisCO operon in *S. elongatus* 7942 was also recently engineered to increase expression and remove light regulation.[44] A system for post-transcriptional repression by riboswitch motifs was successfully transferred to *S. elongatus* 7942.[61] Riboswitches allow induction by addition of a molecule capable of interfering with binding patterns in specially coded untranslated regions in mRNA. The use of self-binding mRNA motifs gives the significant advantage

of not requiring high expression of repressor proteins. A distinct system using riboregulators has also been tested in *Synechocystis* 6803.[62] This system relies on expression of short RNA sequences to interfere with self-binding RNA, instead of an externally added molecule, which may have broad application as pathway designs become more complex. The use of ribosome binding sites (RBSs) to control protein expression has maintained a low profile in cyanobacteria, with RBS either mimicking synthetic RBS used in *E. coli*, or being cloned (purposefully or unknowingly) along with native gene or promoter sequences. The sequence of the untranslated region preceding a gene, including the RBS, can have a significant effect on gene expression in cyanobacteria.[63] This effect can be used to optimize protein balance in a production pathway by varying the RBS preceding each gene in a single operon.[63]

Cyanobacteria present a unique engineering challenge through the existence of multiple chromosomes per cell. Extra steps required for segregation of chromosomes can be reduced in *Synechococcus sp*. PCC 7002 (*Synechococcus* 7002) by recombining over a native gene that causes sensitivity to organic acids such as acrylate, 3-hydroxypropionate, and propionate.[64] Resistance is installed by removal of the gene (*acsA*), such that unmodified genomes can be selected against, aiding fast segregation. A similar technique is used in antibiotics free recombination, although this requires full segregation of a mutant strain to be verified before recombination.[29,65]

1.4 Chemical production in cyanobacteria

Metabolic pathways described in this chapter are summarized in Figure 1-3. Chemical production rates and titers are listed in Table 1-2. Although they are extremely relevant to the understanding of strong pathway design and carbon partitioning, relevant references

for work in *E. coli* have been omitted due to space constraints, all but a few of the synthetic pathways installed in cyanobacteria have previously been constructed in *E. coli* for heterotrophic production from glucose. Readers are directed to recent reviews[9,66], and references within publications cited here for more information on chemical production from fermentative systems. Some values for titer and productivity are estimated from graphical data and all values have been rounded to two significant digits.

1.4.1 Decarboxylation driven synthesis

Pathways that express heterologous proteins at decarboxylation steps have resulted in relatively high production rates and titers, despite the cost of lower carbon conservation ratios (Table 1-1 and Table 1-2). The shortest of these pathways utilizes pyruvate decarboxylase (Pdc), which removes CO_2 from pyruvate to generate acetaldehyde. Acetaldehyde can then be reduced to ethanol by an alcohol dehydrogenase (Adh). Expression of the *adh* and *pdc* genes of *Zymomonas mobilis* from a shuttle vector in *S. elongatus* 7942 led to ethanol synthesis in titers up to 230 mg/L.[35] This work was repeated in *Synechocystis* 6803 by expression of the same genes from a locus on the genome achieving a maximum of about 550 mg/L when grown in a bioreactor.[28] Switching dependency of the Adh cofactor from NADH to NADPH by utilizing an endogenous gene increased ethanol production by ten-fold to 5,500 mg/L.[16] With an NADPH-dependent Adh in the construct, increasing expression of the pathway led to increased titer, where previously increased protein had no effect or led to toxicity. Decarboxylative condensation of two pyruvate molecules forms a new carbon-carbon bond resulting in the four carbon amino acid intermediate 2-acetolactate. Decarboxylation at this step forms acetoin, a compound in butter flavoring, which can be

reduced to give the chemical feedstock and fuel, 2,3-butanediol. This series of reactions has been assembled by introducing acetolactate synthase endcoded by *alsS* from *Bacillus subtilis*, acetolactate decarboxylase encoded by *alsD* from *Aeromonas hydrophila*, and a secondary alcohol dehydrogenase encoded by *sadh* from *Clostridium beijerincki*, into the genome of *S. elongatus* 7942 under the control of the P_{LlacO1} promoter. This construct enabled production of 2,400 mg/L 2,3-butanediol.[14] Cells grew to an optical density about half that of a control strain without the pathway, which is a similar phenotype to other highly productive strains for ethanol[16] and sucrose[31], and is assumed to be the effect of carbon being drawn away from central metabolism. By comparing cell growth and production per day, as much as 60% of biomass by weight was calculated to be redirected to products during peak rates.[14] An analogous pathway was transformed into the genome of *Synechocystis* 6803, consisting of an acetolactate decarboxylase from *Brevibacillus brevis*, and an acetoin reductase from *Enterobacter sp.*, enabling production of 430 mg/L of 2,3-butanediol.[36]

Decarboxylation of 2-ketoisovalerate (KIV), the ketoacid precursor for valine synthesis, produces isobutyraldehyde, a volatile chemical feedstock that can be efficiently harvested by gas stripping.[15] The addition of a single ketoisovalerate decarboxylase, encoded by *kivd* from *Lactococcus lactis,* onto the genome of *S. elongatus* 7942 under the control of the P_{trc} promoter, resulted in low titers of isobutyraldehyde. Coexpression of heterologous genes to generate KIV, including decarboxylative condensation of pyruvate, produced a significant increase in titer. This was further boosted by 1.4-fold to 1,100 mg/L by simultaneous overexpression of *rbcLS* encoding RuBisCO. Including an NADPH-dependent aldehyde dehydrogenase encoded by *yqhD* from *E. coli,* enabled

conversion of isobutyraldehyde to isobutanol, a drop-in fuel, at 450 mg/L[15]. Recently, this process was partially recreated in *Synechocystis* 6803 by integration of *kivd* and *adhA* (NADH-dependent) from *L. lactis* into the genome under control of a P_{tac} promoter.[19] Increasing IPTG concentration increased mRNA of the two heterologous genes but resulted in reduced isobutanol production, indicating a bottleneck is present elsewhere in the carbon flux. *Synechocystis* 6803 is naturally mixotrophic, and is thus able to consume glucose as a carbon and energy source in addition to photosynthetic carbon fixation. When the isobutanol strain was grown in the presence of ^{13}C fully labeled glucose and 50 mM of unlabeled bicarbonate, isobutanol production was increased by 26%. However, labeled carbon made up only 5-10% of amino acids, compared to 70%-90% in wild type, and only 12% of the isobutanol produced was labeled with carbons from glucose. Production dropped sharply when it was grown with glucose only. Together these data indicate imbalance or interference of the isobutanol pathway with normal glucose metabolism. With an *in situ* oleyl alcohol solvent trap and periodic renewal of bicarbonate, productions of 300 mg/L isobutanol from glucose and bicarbonate (mixotrophic growth) and 240 mg/L isobutanol from bicarbonate alone (autotrophic growth) were achieved.[19] Cultures grown on bicarbonate alone without the solvent trap produced 110 mg/L of isobutanol, which is similar in scale to that of *kivd-yqhD* mutants in *S. elongatus* 7942 that lack any other modifications (~50 mg/L).[15]

Production of ethylene is distinct in decarboxylative synthesis as it hinges on a single heterologous step that occurs many steps from fixation. The ethylene forming reaction utilizes 2-oxoglutarate, which performs an essential role in the cell as a carrier for the transfer of amine groups to newly forming amino acids. In *S. elongatus* 7942, an

ethylene forming enzyme (encoded by *efe* from *Pseudomonas syringae*) was inserted downstream of the *psbAI* promoter.[38] Ethylene was produced at a rate of ~0.3 mg/L/h (0.53 mg/L/h/OD). However, production was quickly attenuated, and sequencing revealed an inactivating insertion mutation in the *efe* gene that could be observed in multiple isolates at an identical location in the genetic code. The sequence of *efe* was modified to avoid mutational hotspots and correct transformants of *Synechocystis* 6803 were constructed.[27,39] Strains containing the optimized gene were stable against mutations over multiple generations. Expression from a *lac* series promoter, on a broad-host-range self-replicating plasmid, enabled production of up to ~0.24 mg/L/h of ethylene (~0.19 mg/L/h/OD).[39] When integrated into the genome under control of the pea plant strong *psbA* promoter, a similar production level of 0.35 mg/L/h was observed.[27] Insertion of a second copy of the *efe* gene at the *psbA2* locus doubled production, and had no observable negative effect on metabolism. Concentration of cells to an optical density of 15, illumination with extreme high light (600 μmol photons/m^2/s), and daily resuspension into fresh 5x medium enabled a production rate of 7.12 mg/L/h.[27] It should be noted that even under these optimized conditions, specific productivity can be calculated as ~0.45 mg/L/h/OD, which is close to the productivity reported previously from low OD cultures [38,39].

1.4.2 Acetyl-CoA as a building block for synthesis

Acetyl-CoA is synthesized in the cell by decarboxylation of pyruvate. This two-carbon molecule plays a central role in fatty acid biosynthesis, and provides a dynamic pool of carbon substrate for the synthesis of linear carbon chains. Pathways developed to utilize acetyl-CoA in fermentative hosts have required significant modification to match the

photosynthetic environment in cyanobacteria. An illustrative example of the challenges in acetyl-CoA pathway transfer is that of 1-butanol. A five gene pathway for 1-butanol production was installed into the genome of *S. elongatus* 7942 in two separate operons controlled by P_{trc} and P_{LlacO1} promoters [40]. Production was not observed, and titers reached 13 mg/L only after anoxic fermentation in which cells were placed in the dark and purged of oxygen. Modification of the pathway to utilize endogenous ATP driven conversion of acetyl-CoA to malonyl-CoA, and inclusion of decarboxylation in the first step of the heterologous pathway provided enough driving force to allow for accumulation of trace amounts directly from light and CO_2 without anoxic conditions[17] Replacing three NADH dependent enzymes with NADPH dependent substitutes further raised the titer to 30 mg/L.[17] Finally, replacement of an oxygen sensitive enzyme in the pathway with an oxygen tolerant enzyme increased production by 20-fold to 400 mg/L.[20] The final 1-butanol pathway (Figure 1-3) still includes two NADH dependent steps; however, this requirement is perfectly balanced by generation of two NADH molecules during synthesis of acetyl-CoA, which is used as substrate in the pathway.

Production of the widely used solvent and chemical feedstock acetone from acetyl-CoA faced similar challenges in cyanobacteria. The native *phaA* gene encoding a 3-ketothiolase in *Synechocystis* 6803 was overexpressed in this host to generate acetoacetyl-CoA from acetyl-CoA as the first step, differing from the ATP approach used in 1-butanol.[41] Two more heterologous genes are then required to convert acetoacetyl-CoA to acetone. Similarly to the production scheme for 1-butanol described above, incubation of cells under anoxic conditions, after exponential growth in the light, was required to produce measureable levels of acetone. Increasing expression levels of

pathway enzymes and knocking out competing pathways for acetate and polyhydroxybutyrate (PHB) led to a titer of 36 mg/L under anoxic conditions.[41]

In *S. elongatus* 7942, a pathway analogous to that for acetone was expanded by the addition of a secondary alcohol dehydrogenase to reduce acetone to isopropanol, a 3-carbon solvent and fuel alcohol.[42] This pathway used a distinct acetyl-coA transferase in the second step, which consumes acetate and reforms the starting substrate acetyl-CoA. Production during growth in continuous light was achieved (22 mg/L) only after the addition of acetate, suggesting that endogenously produced acetate or acetyl-CoA was limiting in these conditions. Transferring cells to dark anaerobic conditions improved production of isopropanol to 27 mg/L.[42]

Two other chemical feedstocks, 3-hydroxybutyrate (3HB) and the corresponding polymer polyhydroxyalkanoate (PHA), have been synthesized in cyanobacteria from acetyl-CoA. As with 1-butanol, acetone, and isopropanol, 3HB and PHA pathways start by generating the common intermediate acetoacetyl-CoA in the first step. *Synechocystis* 6803 was utilized to convert acetyl-CoA to 3HB by the addition of three genes (*phaA, phaB* (*Ralstonia eutropha*) and *tesB* (*E. coli*)), enabling a titer of 530 mg/L from CO_2 in light oxic conditions after limitation of phosphorous.[43] In another study, hydroxyacyl-CoA intermediates of fatty acid degradation were polymerized into PHA chains in *Synechococcus* 7002 by the addition of a similar three gene heterologous pathway (*phaA, phaB, phaC* (*Wautersia eutropha*)) using a shuttle vector and a modified P_{rbc} promoter. Cells produced PHA at concentrations up to 52% of dry cell weight.[44,65]

Production of branched carbon chains can be achieved by addition of acetyl-CoA to ketoacids, a process that is catalyzed naturally in *S. elongatus* 7942 by LeuABCD as

part of the leucine biosynthesis pathway. Carbon partitioning to 2-methyl-1-butanol was achieved in *S. elongatus* 7942 by the addition of an engineered citramalate synthase gene to direct addition of acetyl-CoA to pyruvate, removing carbon flux from other ketoacid intermediates. This pathway is unique in that heterologous steps are used to direct flow at the beginning, and again at the end of the pathway, while intermediate steps are catalyzed by endogenous pathways. Culturing in light oxic conditions resulted in a yield of 200 mg/L for 2-methyl-1-butanol.[45]

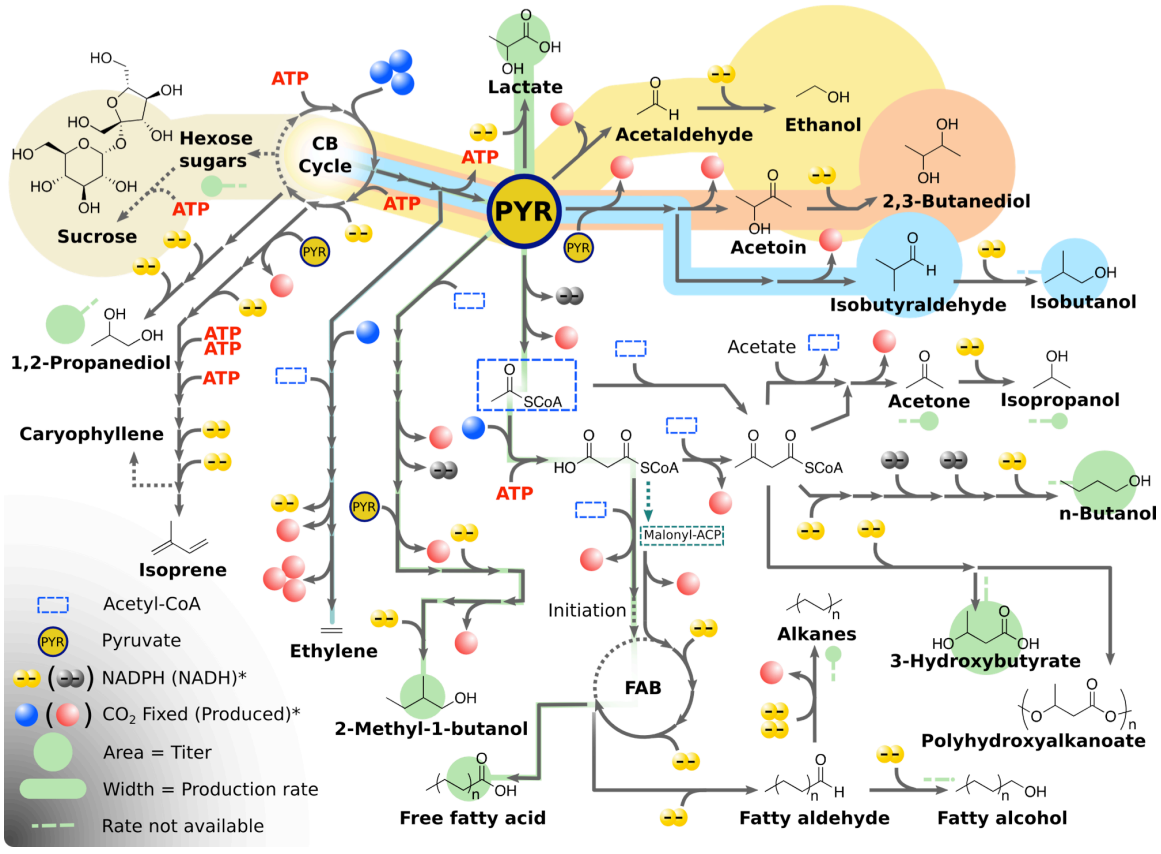


Figure 1-3 Metabolic pathways for chemical production in cyanobacteria. Width of colored backgrounds indicate relative production rates. Areas of circles indicate relative final titers. Dashed colored lines indicate no production rate is available.

1.4.3 Longer carbon chains

The chemical characteristics of acyl-chains make them strong candidates for use as fuels, and the potential for enzymatic control over chain length and saturation has lead to a boom of research in engineering around fatty acid biosynthesis (FAB). Acyl-chains of varied length and saturation are made naturally by cyanobacteria for the generation and maintenance of membranes. The pathway to generate acyl-chains starts with a CO_2 fixation step, combining bicarbonate with acetyl-CoA to form malonyl-CoA. The CoA moiety is then exchanged with ACP, which leads to recognition and incorporation into the FAB pathway. Malonyl-ACP is the central building block of fatty acid chain elongation and undergoes substitution reactions with the thioester bond of acyl-ACP intermediates, lengthening the chain by two carbons each cycle.

To facilitate production of free fatty acid (FFA), the biosynthesis pathway can be short circuited by expression of a thioesterase to hydrolyze the thioester linkage of acyl-ACPs, releasing ACP and a FFA, which then either enters the beta-oxidation pathway for fatty acid degradation or is excreted by the cell. In *Synechocystis* 6803, a thioesterase encoded by *tesA* from *E. coli* was used to replace the acyl-ACP synthetase gene (*slr1609*), resulting in an increase of secreted fatty acids from 1.8 mg/L (wild type) to 84 mg/L.[29] Through six generations of modifications including increasing copies of acetyl-CoA carboxylase, additional thioesterases, and mutations to weaken cell walls and limit competing pathways, titer was increased to 200 mg/L.[29] A subsequent study observed a significant decrease in FFA production if *slr1609* (encoding an acyl-ACP synthetase) is left intact, and identified selective secretion of predominantly unsaturated FFA when the cell membrane is not weakened.[67] Using a promoter induced by limited CO_2 to express a

suite of lipolytic genes or a single thermostable lipase in dense cultures enabled recovery of an additional 27 mg/L and 44 mg/L of FFA respectively from cell membranes after processing.[68,69]

Alkanes and fatty alcohols are desirable as direct fuel products that do not require further chemical modification. Alkanes are naturally produced in some cyanobacteria from intermediates of fatty acid synthesis by an acyl-ACP reductase (AAR) paired with an aldehyde-deformylating oxygenase (ADO).[70,71] Expression of these genes in *Synechocystis* 6803 from either one or two loci on the genome enabled alkane titers of 2.3 mg/L and 26 mg/L respectively.[47,48] Fatty alcohols can be produced in *Synechocystis 6803* at titers of 1.7 mg/L by increased expression of the native gene cluster *accBCDA*, encoding four subunits of an acetyl-CoA carboxylase.[48] In these studies, culture conditions likely affected titers, but uniformly low production is likely due to low activity of AAR and ADO.[47] Due to an endogenous fatty aldehyde reductase, increased expression of the native AAR gene (*orf1594*) in *S. elongatus* 7942 resulted in the production of fatty alcohols. This led to an increase in total esterifiable acyl chains produced from *S. elongatus* 7942, from ~1.7 mg/L to 4.2 mg/L, and a fatty acid secretion rate of 0.63 mg/L/h/OD.[46] Kaiser et al also demonstrated the production of triacylglycerols (TAG) in *S. elongatus* 7942 through the expression of the diacylglycerol acyl transferase gene from *Acinetobacter sp.* ADP1. TAG production was verified by thin layer chromatography and by visibile presence of lipid bodies using microscopy. However, TAG toxicity prevented strong production.[46] Long chain alkenes are naturally produced in small amounts in *Synechococcus 7002* and this has been linked to the activity

of a modular-synthase encoding gene.[72] Increasing expression of this gene by replacing the promoter increased titers 5-fold to 3.5 mg/L/OD.[72]

1.4.4 Terpenoids

The biosynthesis of terpenoids presents a challenge to metabolic engineers due to the distance of the carbon pool from fixation. In prokaryotes, including cyanobacteria, the metabolic precursors for terpenoid synthesis, dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP), are produced in 7 steps from glyceraldehyde-3-phosphate through the methylerythritol pathway. Expression of a codon optimized isoprene synthase (encoded by *ispS* (*Pueraria montana*)) from the *psbA2* locus on the genome of *Synechocystis* 6803 enabled conversion of DMAPP to isoprene at detectable levels.[73] Measurement of transformants in a sealed bioreactor, with periodic dilution of cells into fresh medium, showed a linear accumulation rate of isoprene (0.01 mg/L/h/OD).[49] Expression of the beta-caryophyllene synthase gene (*QHS1* (*Artemisia annua*)) from the *psbA2* locus in *Synechocystis* 6803, enabled synthesis of beta-caryophyllene from the steroid precursor farnesyl diphosphate, which is produced naturally in two steps from DMAPP and IPP. Production was measured by extraction of beta-caryophyllene from cells to be 2.2×10^{-5} mg/gDCW/h.[50] The low titers seen in these studies indicate the limitations on production imposed by carbon partitioning in the methylerythritol pathway.[30]

1.4.5 Non-decarboxylative synthesis by heterologous pathways

Only two heterologous pathways expressed in cyanobacteria achieve complete conservation of fixed CO_2 in their target molecule by avoiding decarboxylation (Table 1-1). Synthesis of 1,2-propanediol was achieved in *S. elongatus* 7942 by targeting the CB

cycle intermediate dihydroxy acetone phosphate (4 steps from CO_2 fixation).[18] Expression of methylglyoxal synthase, encoded by *mgsA* from *E. coli*, allowed for conversion of dihydroxy acetone phosphate to the diacarbonyl intermediate methylglyoxal, which is subsequently reduced by heterologous primary and secondary alcohol dehydrogenases YqhD (*E.coli*) and GldA (*E.coli*) respectively to produce 1,2-propanediol. Switching the NADH dependent alcohol dehydrogenase to an NADPH dependent homolog increased production from 22 mg/L to 150 mg/L.[18]

The shortest heterologous pathway engineered in cyanobacteria is the non-decarboxylative reduction of pyruvate to produce lactic acid. Despite the apparent simplicity of this synthesis, increasing productivity has been challenging. Addition of a lactate transporter and lactate dehydrogenase (encoded by *lldP and ldhA* respectively (*E. coli*)) to *S. elongatus 7942* resulted in accumulation of lactate in the supernatant to 22 mg/L.[51] In another study, addition of lactate dehydrogenase encoded by *ldh* from *B. subtilis* to the genome of *Synechocystis* 6803 led to accumulation of 63 mg/L lactate in the supernatant.[52] This was further enhanced 5-fold by coexpression of a transhydrogenase encoded by *sth* from *Pseudomonas aeruginosa* to increase NADH availability, reaching a lactate titer of 290 mg/L. Another analysis of *Synechocystis* 6803 screened for heterologous lactate dehydrogenase activity (*ldhB* and *ldhX* (*L. lactis*); *ldhL* and *ldh* (*Lactobacillus plantarum*); *ldhL* (*Lactobacillus rhamnosus*)) and confirmed similar titers (15 mg/L) as those seen without a transhydrogenase.[53]

1.4.6 Endogenous production of sugars

Cyanobacterial production of sugars achieves stoichiometric carbon conservation and maximum cofactor ratio of 1:1.5 (Table 1-1, Figure 1-2). Endogenous production can be

enhanced by increasing the rate of sugar transport out of the cell, knocking out competing pathways, and stimulating production by changing culturing conditions. Transferring *S. elongatus* 7942 cells to high saline conditions (100-200 mM NaCl) induces synthesis of sucrose as an osmoprotectant. Addition of a sucrose transporter gene (*cscB* (*E. coli*)) to the genome leads to accumulation of sucrose in the supernatant.[31] Knocking out competing pathways for glycogen (Δ*glgC*) and sucrose degradation (Δ*invA*) increased production to a maximum rate of 36 mg/L/h, which corresponds to 80% of biomass.[31] Production of sugars other than sucrose can similarly be improved. Addition of a monosaccharide transporter gene (*glf* (*Z. mobilis*) to the genome of *S. elongatus* 7942 allowed for secretion of glucose and fructose into the media, after induction with salt stress.[51] Fructose and glucose production was also enhanced by addition of a heterologous invertase encoded by *invA* from *Z. mobilis* to reach titers of ~29 mg/L and ~58 mg/L respectively.

1.5 Calculating practical theoretical yields

While chemical production has increased in photosynthetic systems, calculations for percent yields have remained elusive. Theoretical values for photon and carbon stoichiometry can be extrapolated from fundamental laws as described in the stoichiometry section above. However, applying these values to a production experiment requires precise knowledge of how light and CO_2 enters and leaves the system. During production, light is usually provided in excess, and the amount captured is unknown. Similarly, when measuring CO_2 uptake, the equilibrium between CO_2 and HCO_3^-, as well as between dissolved and gaseous phases, makes quantification of CO_2 uptake difficult. However, when light and carbon are provided in excess, it is known that the limits of

photosynthesis are reflected by either photosynthetic electron transport (PET) or RuBisCO turnover rates, and that the maximum rate imposed by the relationship between these limits can be approximated in any system by the amount of oxygen evolved at photosystem II (PSII).[74] Theoretically, fixation of 1 CO_2 equates to the generation of 0.83 O_2, however when measured, the ratio is more closely approximated as 1 to 1.[74] This reflects that energy is also expended for the carbon concentrating mechanisms in cyanobacteria[75], and that a higher ratio of PSI to PSII is likely to lower the absorption rate of stoichiometric photons.[76] Assuming a ratio of 1 carbon fixed per oxygen evolved, the maximal rate of photosynthesis under optimum conditions of high CO_2 and light has been measured in *Synechocystis* 6803 at 220 μmol CO_2/mg Chl/h[77], in *S. elongatus* 7942 at 120 μmol CO_2/mg Chl/h[74], and in *Synechococcus* 7002 at 550 μmol CO_2/mg Chl/h.[78] The top three production rates for chemical synthesis in cyanobacteria (ethanol, sucrose, and 2,3-butanediol) have been achieved in *S. elongatus* 7942. Using the upper range of chlorophyll synthesis in *S. elongatus* of 5.9 μg Chl/ml/OD[76], the maximal CO_2 assimilation rate becomes 710 μmol CO_2/L/h/OD. For ethanol, which requires 3 CO_2 fixed per molecule, the maximal production rate becomes 240 μmol/L/h/OD. Ethanol production has been measured at 220 μmol/L/h/OD which equates to 92% of the photosynthetic maximum calculated here.[16] Sucrose requires fixation of 12 carbons, which gives a maximal production rate of 59 μmol/L/h/OD. Sucrose production was measured at 9.5 mg/L/h/OD which converts to 28 μmol/L/h/OD giving a 47% yield. 2,3-Butanediol requires 6 CO_2 per molecule and has a theoretical maximum production rate of 120 μmol/L/h/OD. The peak production rate was measured as 50 μmol/L/h/OD corresponding to 42%.[14] These calculations indicate carbon partitioning close to that

observed by decreased biomass ~60% for sucrose and 2,3-butanediol. Refining the values used in these calculations specifically for well-characterized cyanobacterial lab strains and during production experiments will greatly aid metabolic engineering efforts to optimize cells for maximum chemical production.

1.6 Conclusion

The engineering of metabolic pathways in cyanobacteria is rapidly progressing with characterization of new tools, design and implementation of new pathways, and elucidation of core concepts for adapting engineering to the photosynthetic environment. As production is fine-tuned, greater amounts of carbon can be redirected away from growth processes. During the design of new pathways, attention to the stoichiometry of carbon conservation and cofactor ratios can provide valuable insight into efficiency and product yield. Biomass measurements and calculations for theoretical maximums indicate that heterologous biosynthetic pathways may already be partitioning significant amounts of fixed carbon into chemical production, however, a standardized and accessible system for quantifying efficiency is still desirable.

References

1 Field, C. B. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **281**, 237-240, doi:10.1126/science.281.5374.237 (1998).

2 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, doi:10.1073/pnas.182432999 (2002).

3 Xu, Y. *et al.* Altered carbohydrate metabolism in glycogen synthase mutants of *Synechococcus sp.* strain PCC 7002: Cell factories for soluble sugars. *Metab. Eng.* **16**, 56-67, doi:10.1016/j.ymben.2012.12.002 (2013).

4 Ono, E. & Cuello, J. L. Carbon Dioxide Mitigation using Thermophilic Cyanobacteria. *Biosys. Eng.* **96**, 129-134, doi:10.1016/j.biosystemseng.2006.09.010 (2007).

5 Vishnivetskaya, T. in *Permafrost Soils* Vol. 16 *Soil Biology* (ed Rosa Margesin) Ch. 6, 73-84 (Springer Berlin Heidelberg, 2009).

6 Vinyard, D. J. *et al.* Natural variants of photosystem II subunit D1 tune photochemical fitness to solar intensity. *J. Biol. Chem.* **288**, 5451-5462, doi:10.1074/jbc.M112.394668 (2013).

7 Paddock, M. L., Boyd, J. S., Adin, D. M. & Golden, S. S. Active output state of the *Synechococcus* Kai circadian oscillator. *Proc. Natl. Acad. Sci. USA* **110**, E3849-3857, doi:10.1073/pnas.1315170110 (2013).

8 Bell-Pedersen, D. *et al.* Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet* **6**, 544-556 (2005).

9 Rabinovitch-Deere, C. A., Oliver, J. W., Rodriguez, G. M. & Atsumi, S. Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* **113**, 4611-4632, doi:10.1021/cr300361t (2013).

10 Vioque, A. in *Transgenic Microalgae as Green Cell Factories* Vol. 616 *Advances in Experimental Medicine and Biology* (eds Rosa León, Aurora Galván, & Emilio Fernández) Ch. 2, 12-22 (Springer New York, 2007).

11 Gimpel, J. A., Specht, E. A., Georgianna, D. R. & Mayfield, S. P. Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr. Opin. Chem. Biol.* **17**, 489-495, doi:10.1016/j.cbpa.2013.03.038 (2013).

12 Berla, B. M. *et al.* Synthetic biology of cyanobacteria: unique challenges and opportunities. *Front. Microbiol.* **4**, 246, doi:10.3389/fmicb.2013.00246 (2013).

13 Nozzi, N. E., Oliver, J. W. K. & Atsumi, S. Cyanobacteria as a Platform for Biofuel Production. *Front. Bioeng. Biotechnol.* **1**, doi:10.3389/fbioe.2013.00007 (2013).

14 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).
15 Atsumi, S., Higashide, W. & Liao, J. C. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* **27**, 1177-1180, doi:10.1038/nbt.1586 (2009).
16 Gao, Z. X., Zhao, H., Li, Z. M., Tan, X. M. & Lu, X. F. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energ. Environ. Sci.* **5**, 9857-9865, doi:Doi 10.1039/C2ee22675h (2012).
17 Lan, E. I. & Liao, J. C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc. Natl. Acad. Sci. USA* **109**, 6018-6023, doi:10.1073/pnas.1200074109 (2012).
18 Li, H. & Liao, J. C. Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO2 to 1,2-propanediol. *Microb. Cell. Fact.* **12**, 4, doi:10.1186/1475-2859-12-4 (2013).
19 Varman, A. M., Xiao, Y., Pakrasi, H. B. & Tang, Y. J. Metabolic engineering of *Synechocystis sp.* strain PCC 6803 for isobutanol production. *Appl. Environ. Microbiol.* **79**, 908-914, doi:10.1128/AEM.02827-12 (2013).
20 Lan, E. I., Ro, S. Y. & Liao, J. C. Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energ. Environ. Sci.* **6**, 2672, doi:10.1039/c3ee41405a (2013).
21 Zhu, H. *et al.* Integrated OMICS guided engineering of biofuel butanol-tolerance in photosynthetic *Synechocystis sp.* PCC 6803. *Biotechnol. Biofuels.* **6**, 106, doi:10.1186/1754-6834-6-106 (2013).
22 Wang, J. *et al.* RNA-seq based identification and mutant validation of gene targets related to ethanol resistance in cyanobacterial *Synechocystis sp.* PCC 6803. *Biotechnol. Biofuels.* **5**, 89, doi:10.1186/1754-6834-5-89 (2012).
23 Ruffing, A. M. & Jones, H. D. Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnol. Bioeng.* **109**, 2190-2199, doi:10.1002/bit.24509 (2012).
24 Qiao, J. *et al.* Quantitative iTRAQ LC-MS/MS proteomics reveals metabolic responses to biofuel ethanol in cyanobacterial *Synechocystis sp.* PCC 6803. *J. Proteome Res.* **11**, 5286-5300, doi:10.1021/pr300504w (2012).
25 Liu, J., Chen, L., Wang, J., Qiao, J. & Zhang, W. Proteomic analysis reveals resistance mechanism against biofuel hexane in *Synechocystis sp.* PCC 6803. *Biotechnol. Biofuels.* **5**, 68, doi:10.1186/1754-6834-5-68 (2012).
26 Ruffing, A. M. RNA-Seq analysis and targeted mutagenesis for improved free fatty acid production in an engineered cyanobacterium. *Biotechnol. Biofuels.* **6**, 113, doi:10.1186/1754-6834-6-113 (2013).
27 Ungerer, J. *et al.* Sustained photosynthetic conversion of CO2 to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energ. Environ. Sci.* **5**, 8998-9006, doi:Doi 10.1039/C2ee22555g (2012).
28 Dexter, J. & Fu, P. C. Metabolic engineering of cyanobacteria for ethanol production. *Energ. Environ. Sci.* **2**, 857-864, doi:Doi 10.1039/B811937f (2009).

29 Liu, X., Sheng, J. & Curtiss, R., 3rd. Fatty acid production in genetically modified cyanobacteria. *Proc. Natl. Acad. Sci. USA* **108**, 6899-6904, doi:10.1073/pnas.1103014108 (2011).
30 Melis, A. Carbon partitioning in photosynthesis. *Curr. Opin. Chem. Biol.* **17**, 453-456, doi:10.1016/j.cbpa.2013.03.010 (2013).
31 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, doi:10.1128/AEM.07901-11 (2012).
32 Vermaas, W. in *Photosynthesis Research for Food, Fuel and the Future Advanced Topics in Science and Technology in China* Ch. 74, 353-357 (Springer Berlin Heidelberg, 2013).
33 Shikanai, T., Munekage, Y. & Kimura, K. Regulation of proton-to-electron stoichiometry in photosynthetic electron transport: physiological function in photoprotection. *J. Plant Res.* **115**, 0003-0010, doi:10.1007/s102650200001 (2002).
34 Takahashi, H., Clowez, S., Wollman, F. A., Vallon, O. & Rappaport, F. Cyclic electron flow is redox-controlled but independent of state transition. *Nat. Commun.* **4**, 1954, doi:10.1038/ncomms2954 (2013).
35 Deng, M. D. & Coleman, J. R. Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ. Microbiol.* **65**, 523-528 (1999).
36 Savakis, P. E., Angermayr, S. A. & Hellingwerf, K. J. Synthesis of 2,3-butanediol by *Synechocystis sp.* PCC6803 via heterologous expression of a catabolic pathway from lactic acid- and enterobacteria. *Metab. Eng.*, doi:10.1016/j.ymben.2013.09.008 (2013).
37 Sakai, M., Ogawa, T., Matsuoka, M. & Fukuda, H. Photosynthetic conversion of carbon dioxide to ethylene by the recombinant cyanobacterium, *Synechococcus sp*. PCC 7942, which harbors a gene for the ethylene-forming enzyme of *Pseudomonas syringae*. *J. Ferment. Bioeng.* **84**, 434-443, doi:Doi 10.1016/S0922-338x(97)82004-1 (1997).
38 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).
39 Guerrero, F., Carbonell, V., Cossu, M., Correddu, D. & Jones, P. R. Ethylene synthesis and regulated expression of recombinant protein in *Synechocystis sp.* PCC 6803. *PLoS One* **7**, e50470, doi:10.1371/journal.pone.0050470 (2012).
40 Lan, E. I. & Liao, J. C. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab. Eng.* **13**, 353-363, doi:10.1016/j.ymben.2011.04.004 (2011).
41 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, doi:10.1016/j.ymben.2012.03.005 (2012).

42 Kusakabe, T. *et al.* Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metab. Eng.* **20C**, 101-108, doi:10.1016/j.ymben.2013.09.007 (2013).
43 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, doi:10.1016/j.ymben.2013.01.001 (2013).
44 Miyasaka, H., Okuhata, H., Tanaka, S., Onizuka, T. & Akiyama, H. *Polyhydroxyalkanoate (PHA) Production from Carbon Dioxide by Recombinant Cyanobacteria*. 197-215 (InTech, 2013).
45 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 and Environmental Science* **5**, 9574-9583, doi:Doi 10.1039/C2ee23148d (2012).
46 Kaiser, B. K. *et al.* Fatty aldehydes in cyanobacteria are a metabolically flexible precursor for a diversity of biofuel products. *PLoS One* **8**, e58307, doi:10.1371/journal.pone.0058307 (2013).
47 Wang, W., Liu, X. & Lu, X. Engineering cyanobacteria to improve photosynthetic production of alka(e)nes. *Biotechnol. Biofuels.* **6**, 69, doi:10.1186/1754-6834-6-69 (2013).
48 Tan, X. *et al.* Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab. Eng.* **13**, 169-176, doi:10.1016/j.ymben.2011.01.001 (2011).
49 Bentley, F. K. & Melis, A. Diffusion-based process for carbon dioxide uptake and isoprene emission in gaseous/aqueous two-phase photobioreactors by photosynthetic microorganisms. *Biotechnol. Bioeng.* **109**, 100-109, doi:10.1002/bit.23298 (2012).
50 Reinsvold, R. E., Jinkerson, R. E., Radakovits, R., Posewitz, M. C. & Basu, C. The production of the sesquiterpene beta-caryophyllene in a transgenic strain of the cyanobacterium *Synechocystis*. *J. Plant Physiol.* **168**, 848-852, doi:10.1016/j.jplph.2010.11.006 (2011).
51 Niederholtmeyer, H., Wolfstadter, 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, doi:10.1128/AEM.00202-10 (2010).
52 Angermayr, S. A., Paszota, M. & Hellingwerf, K. J. Engineering a cyanobacterial cell factory for production of lactic acid. *Appl. Environ. Microbiol.* **78**, 7098-7106, doi:10.1128/AEM.01587-12 (2012).
53 Joseph, A. *et al.* Utilization of lactic acid bacterial genes in *synechocystis sp.* PCC 6803 in the production of lactic acid. *Biosci., Biotechnol., Biochem.* **77**, 966-970, doi:10.1271/bbb.120921 (2013).
54 Lee, J. W. *et al.* Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **8**, 536-546, doi:10.1038/nchembio.970 (2012).
55 Imamura, S. & Asayama, M. Sigma factors for cyanobacterial transcription. *Gene. Regul. Syst. Bio.* **3**, 65-87 (2009).

56 Nair, U., Thomas, C. & Golden, S. S. Functional elements of the strong psbAI promoter of *Synechococcus elongatus* PCC 7942. *J. Bacteriol.* **183**, 1740-1747, doi:10.1128/JB.183.5.1740-1747.2001 (2001).
57 Chungjatupornchai, W. & Fa-aroonsawat, S. The rrnA promoter as a tool for the improved expression of heterologous genes in cyanobacteria. *Microbiol. Res.*, doi:10.1016/j.micres.2013.09.010 (2013).
58 Huang, H. H. & Lindblad, P. Wide-dynamic-range promoters engineered for cyanobacteria. *J. Biol. Eng.* **7**, 10, doi:10.1186/1754-1611-7-10 (2013).
59 Rosgaard, L., de Porcellinis, A. J., Jacobsen, J. H., Frigaard, N. U. & Sakuragi, Y. Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants. *J. Biotechnol.* **162**, 134-147, doi:10.1016/j.jbiotec.2012.05.006 (2012).
60 Kitayama, Y. *et al.* An in vivo dual-reporter system of cyanobacteria using two railroad-worm luciferases with different color emissions. *Plant Cell Physiol.* **45**, 109-113, doi:Doi 10.1093/Pcp/Pch001 (2004).
61 Nakahira, Y., Ogawa, A., Asano, H., Oyama, T. & Tozawa, Y. Theophylline-Dependent Riboswitch as a Novel Genetic Tool for Strict Regulation of Protein Expression in Cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant Cell Physiol.* **54**, 1724-1735, doi:10.1093/pcp/pct115 (2013).
62 Abe, K. *et al.* Design of riboregulators for control of cyanobacterial (*Synechocystis*) protein expression. *Biotechnol. Lett.*, doi:10.1007/s10529-013-1352-x (2013).
63 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Combinatorial optimization of cyanobacterial 2,3-butanediol production. *Metab. Eng.* **22C**, 76-82, doi:10.1016/j.ymben.2014.01.001 (2014).
64 Begemann, M. B. *et al.* An organic Acid based counter selection system for cyanobacteria. *PLoS One* **8**, e76594, doi:10.1371/journal.pone.0076594 (2013).
65 Akiyama, H. *et al.* Antibiotics-free stable polyhydroxyalkanoate (PHA) production from carbon dioxide by recombinant cyanobacteria. *Bioresour. Technol.* **102**, 11039-11042, doi:10.1016/j.biortech.2011.09.058 (2011).
66 Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320-328, doi:10.1038/nature11478 (2012).
67 Hu, P. *et al.* Metabolic phenotyping of the cyanobacterium *Synechocystis 6803* engineered for production of alkanes and free fatty acids. *Applied Energy* **102**, 850-859, doi:10.1016/j.apenergy.2012.08.047 (2013).
68 Liu, X., Fallon, S., Sheng, J. & Curtiss, R., 3rd. CO2-limitation-inducible Green Recovery of fatty acids from cyanobacterial biomass. *Proc. Natl. Acad. Sci. USA* **108**, 6905-6908, doi:10.1073/pnas.1103016108 (2011).
69 Liu, X. & Curtiss, R., 3rd. Thermorecovery of cyanobacterial fatty acids at elevated temperatures. *J. Biotechnol.* **161**, 445-449, doi:10.1016/j.jbiotec.2012.08.013 (2012).
70 Schirmer, A., Rude, M. A., Li, X., Popova, E. & del Cardayre, S. B. Microbial biosynthesis of alkanes. *Science* **329**, 559-562, doi:10.1126/science.1187936 (2010).

71 Pandelia, M. E. *et al.* Substrate-triggered addition of dioxygen to the diferrous cofactor of aldehyde-deformylating oxygenase to form a diferric-peroxide intermediate. *J. Am. Chem. Soc.* **135**, 15801-15812, doi:10.1021/ja405047b (2013).
72 Mendez-Perez, D., Begemann, M. B. & Pfleger, B. F. Modular synthase-encoding gene involved in alpha-olefin biosynthesis in *Synechococcus sp.* strain PCC 7002. *Appl. Environ. Microbiol.* **77**, 4264-4267, doi:10.1128/AEM.00467-11 (2011).
73 Lindberg, P., Park, S. & Melis, A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab. Eng.* **12**, 70-79, doi:10.1016/j.ymben.2009.10.001 (2010).
74 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. *Photosynthesis Res.* **88**, 287-297, doi:10.1007/s11120-006-9048-x (2006).
75 Fridlyand, L., Kaplan, A. & Reinhold, L. Quantitative evaluation of the role of a putative CO2-scavenging entity in the cyanobacterial CO2-concentrating mechanism. *BioSyst.* **37**, 229-238 (1996).
76 Kopecna, J., Komenda, J., Bucinska, L. & Sobotka, R. Long-term acclimation of the cyanobacterium *Synechocystis sp. PCC 6803* to high light is accompanied by an enhanced production of chlorophyll that is preferentially channeled to trimeric photosystem I. *Plant Physiol.* **160**, 2239-2250, doi:10.1104/pp.112.207274 (2012).
77 Marcus, Y., Altman-Gueta, H., Wolff, Y. & Gurevitz, M. Rubisco mutagenesis provides new insight into limitations on photosynthesis and growth in *Synechocystis PCC6803*. *J. Exp. Bot.* **62**, 4173-4182, doi:10.1093/jxb/err116 (2011).
78 Klughammer, B., Sultemeyer, D., Badger, M. R. & Price, G. D. The involvement of NAD(P)H dehydrogenase subunits, NdhD3 and NdhF3, in high-affinity CO2 uptake in *Synechococcus sp*. PCC7002 gives evidence for multiple NDH-1 complexes with specific roles in cyanobacteria. *Mol. Microbiol.* **32**, 1305-1315 (1999).

Chapter 2 Cyanobacterial conversion of carbon dioxide to 2,3-butanediol

This chapter closely resembles published work. Citation: Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. Proc. Natl. Acad. Sci. USA 110, 1249-1254, doi:10.1073/pnas.1213024110 (2013)

2.1 Overview

Conversion of CO_2 for the synthesis of chemicals by photosynthetic organisms is an attractive process for establishing energy independence from fossil reserves. However, synthetic pathway construction in cyanobacteria is still in its infancy compared to model fermentative organisms. Here we systematically developed the 2,3-butanediol (23BD) biosynthetic pathway in *Synechococcus elongatus* PCC7942 as a model system to establish design methods for efficient exogenous chemical production in cyanobacteria. We identified 23BD as a target chemical with low host toxicity, and designed an oxygen-insensitive, cofactor-matched biosynthetic pathway coupled with irreversible enzymatic steps to create a driving force toward the target. Production of 23BD from CO_2 reached 2.38 g/L, which is a significant increase for chemical production from exogenous pathways in cyanobacteria. This work demonstrates that design methods can continue to increase chemical production in cyanobacteria.

Amid rising global energy demands and pressing environmental issues, interest is growing in the production of fuels and chemicals from renewable resources. Petroleum consumption reached 37.1 quadrillion BTU in the United States in 2008, of which a large majority (71%) was liquid fuel in the transportation sector. Petroleum and natural gas account for 99% of the feedstocks for chemicals such as plastics, fertilizers and pharmaceuticals in chemical industry.[1] Considering rapidly increasing world population

and exhaustion of fossil fuels, the development of sustainable processes for energy and carbon capture (ECC), to produce fuels and chemicals, is crucial for the human society.

ECC by cyanobacteria is also directed toward mitigating increasing atmospheric CO_2 concentrations. According to the US Energy Information Administration[2], world energy-related CO_2 emissions in 2006 were 29 billion metric tons, which is an increase of 35% from 1990. Accelerating accumulation of atmospheric CO_2 is not only due to increased emissions from world growth and intensifying carbon use, but also from a possible attenuation in the efficiency of the world's natural carbon sinks.[3] As a result, atmospheric levels of CO_2 have increased by ~25% over the past 150 years and it has become increasingly important to develop new technologies to reduce CO_2 emissions. Many creative solutions have been proposed and argued for carbon capture, each with varied environmental side-effects, and costs.[4] Sequestration by photosynthetic microorganisms in which CO_2 is biologically converted to valuable chemicals is an important addition to the toolbox for overall capture of CO_2.[5-7]

Photosynthetic microorganisms including cyanobacteria are currently being engineered for platforms to convert solar energy to biochemicals renewably.[5-7] They possess many advantages over traditional terrestrial plants with regard to biochemical production. For example, the photosynthetic efficiency of photosynthetic microorganisms is higher than plants, and photosynthetic microorganisms can be cultivated in locations that do not compete with traditional agricultural crops.[8] Cyanobacteria are collectively responsible for almost 50% of global photosynthesis and are found in a wide range of environments.[9] While cyanobacteria have many similar features with algae in this context, many cyanobacterial species feature more simple genetic structures and faster growth

rates.[5] As a result genetic engineering methods for cyanobacteria are more advanced than those for algae.[10-12]

Cyanobacteria have the biochemical machinery required to fix CO_2, but lack the critical components to generate fuels and chemicals efficiently. An increased understanding of cellular systems enables us to construct novel systems using synthetic biology, assembling the components and control systems into new combinations.[6,7,13] This approach can be applied to produce valuable chemicals, that the cyanobacteria host strains do not produce naturally, by constructing new biosynthetic pathways.[6] Recently, cyanobacteria have been engineered to produce various chemicals including isobutyraldehyde[14], isobutanol[14], 1-butanol[15], ethylene[16], isoprene[17], acetone[18], fatty acids[19], and fatty alcohols[20] through exogenous biosynthetic pathways. This approach in cyanobacteria is significantly less developed compared to a model organism such as *Escherichia coli*. However, results in *E. coli* cannot be directly translated into cyanobacteria. For example, an engineered *E. coli* strain containing the 1-butanol pathway produced more than 30 g/L 1-butanol[21], but the cyanobacterial strain with the same pathway produced only trace amounts of 1-butanol.[22] We thus seek to establish an efficient method for biosynthetic pathway construction in cyanobacteria.

We aim to establish in this work that matching the choice of target chemicals and pathways to the chosen photosynthetic production host is beneficial for cyanobacterial chemical production. First, low toxicity toward cells is important. Isobutyraldehyde production in *S. elongatus* has achieved the highest production rate in literature, however, due to high toxicity of isobutyraldehyde this system requires constant isolation of the product to achieve high productivity.[14] Second, since photosynthesis produces oxygen,

activities of oxygen-sensitive enzymes are usually problematic in cyanobacteria, with the result that pathways such as 1-butanol and acetone from anaerobic organisms are inefficient in cyanobacteria.[18,22] Third, chemicals that are secreted from cells without overexpression of transporters or additional genetic modifications to cells are ideal to make the system simple and inexpensive. Lastly, providing irreversible reactions that create a driving force to direct flux toward the target chemicals, is important for construction of efficient pathways.[15,21,23]

In this work we used production of 2,3-butanediol (23BD) (Figure 2-1) in *Synechococcus elongatus* PCC7942 (*S. elongatus*) as a model system and followed a systematic approach in which various enzymes were compared both in *E. coli* and *S. elongatus* to optimize chemical production. Microbial fermentation of 23BD has been under investigation for many years as an alternate route for chemical production.[24,25] 23BD can be converted by dehydrogenation to methyl ethyl ketone (MEK), which is a liquid fuel additive and useful industrial solvent.[26] Furthermore, the catalytic conversion of 23BD to 1,3-butadiene, which is a precursor for a diversity of polymer and copolymer materials, has been well established.[27] 23BD has also been used in the manufacturing of plasticizers, inks, fumigants, and explosives.[28] Concurrently, methods for improving product recovery efficiency, which is an important determinant of the total production cost and a major limitation of biochemical production from microorganisms[29,30], have been well-established.[24,31] This current work will provide a stronger understanding of the potential of cyanobacteria as a platform for production of valuable biochemicals such as 23BD, while providing important fundamental insight to facilitate future progress along such lines.

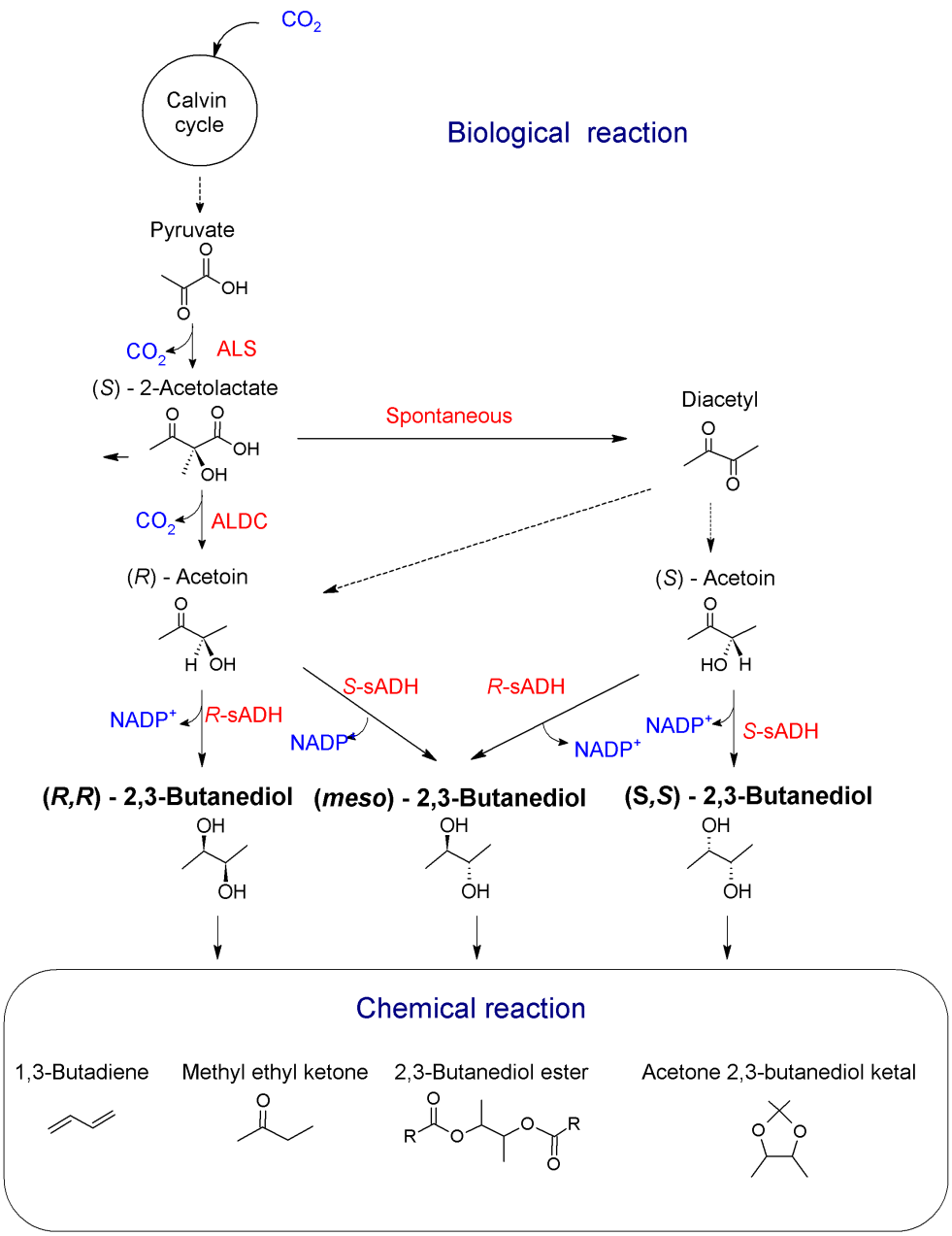


Figure 2-1 The pathway for acetoin and 2,3-butanediol production in *S. elongatus* PCC7942. The acetoin/2,3-butanediol production pathway contains three enzymatic steps from pyruvate. ALS: acetolactate synthase; ALDC: acetolactate decarboxylase; sADH: secondary alcohol dehydrogenase.

2.2 23BD exhibits low toxicity in *S. elongatus*.

To increase the titer and duration of chemical production, low toxicity or constant removal of the product is necessary. Because constant removal and purification of small concentrations during production is not cost-effective on an industrial scale, it was a prerequisite of this study that the chemical target be tolerated at an acceptable volume of greater than 1% (10 g/L) by production strains. To evaluate acetoin and 23BD toxicity, we tested the growth of *S. elongatus* over 72 h in the presence of 23BD or acetoin (Figure 2-2). Growth rates between 0 and 24 h decreased by 18% and 42% in the presence of 0.1 and 0.2 g/L acetoin, respectively (Figure 2-2a). Growth in the presence of 0.2 and 1.0 g/L acetoin stopped after 24 h, while growth in the absence of acetoin reached a plateau at 48 h (Figure 2-2a). These results indicate acute toxicity for this precursor. This is comparable to isobutyraldehyde and isobutanol, which prevent growth of *S. elongatus* at 1 g/L[14]. Conversely growth rates between 0 and 24 h in the presence of 10 and 30 g/L 23BD decreased by only 2.2 and 8.9%, respectively (Figure 2-2b), surpassing our benchmark goal for product tolerance. Growth in the presence of 30 g/L 23BD stopped at 24 h, while that in the presence of 10 g/L 23BD reached a plateau at 48 h similar to that seen in the absence of 23BD (Figure 2-2b). These results indicate that 23BD is a suitable target for high titer and long-term cyanobacterial production, as long as high flux through acetoin can be maintained to prevent accumulation of the toxic intermediate.

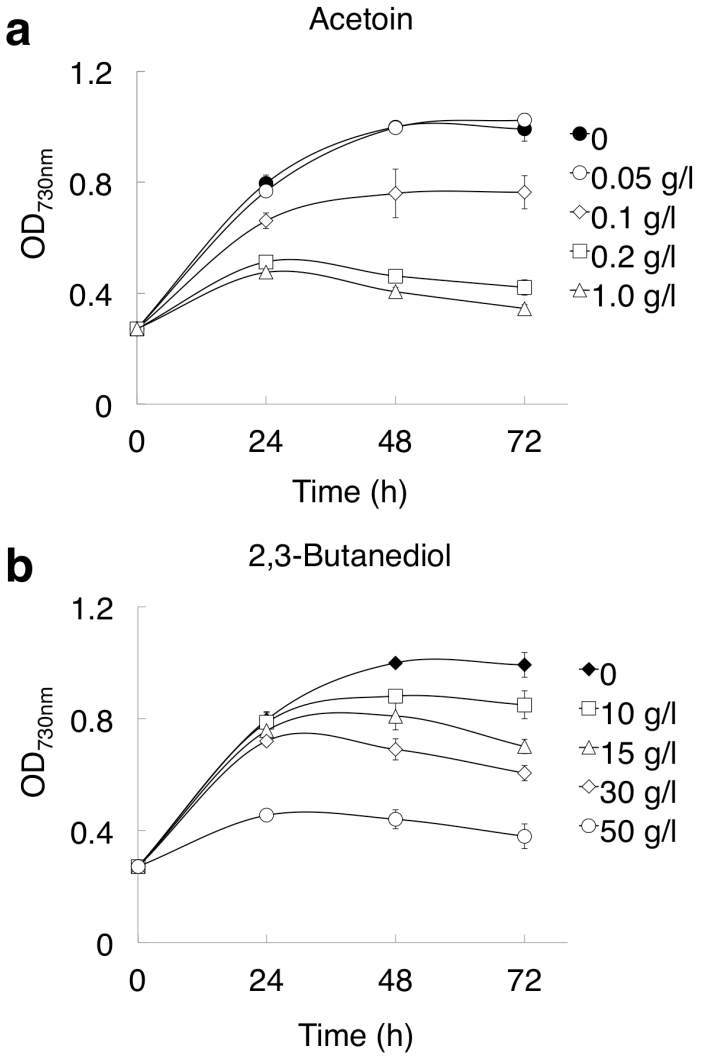


Figure 2-2 Effect of acetoin and 2,3-butanediol on growth. (**a**) Time course for the growth of *S. elongatus* in the absence of acetoin (●) or in the presence of 0.05 g/L (○), 0.1 g/L (◇), 0.2 g/L (□), and 1.0 g/L (△) acetoin. (**b**) Time course for the growth of *S. elongatus* in the absence of 23BD (●) or in the presence of 10 g/L (□), 15 g/L (△), 30 g/L (◇), and 50 g/L (○) 23BD. Cells were incubated in BG-11 containing 50mM $NaHCO_3$ at 30°C. Error bars are standard deviation, n = 3.

2.3 Construction of the acetoin biosynthetic pathway

Acetoin can be produced by the decarboxylation of 2-acetolactate. In this pathway (Figure 2-1) two pyruvate molecules are converted into 2-acetolactate by acetolactate synthase (ALS) encoded by *alsS*. 2-Acetolactate is then decarboxylated to yield acetoin by 2-acetolactate decarboxylase (ALDC) encoded by *alsD*. Pyruvate, the source of carbon for the pathway, is produced naturally through the fixation of three CO_2 molecules

in the Calvin-Benson-Bassham cycle.[32] Conversion of pyruvate to 2-acetolactate occurs naturally during valine/leucine biosynthesis, albeit in low amounts.[33] Previously the *alsS* gene, which encodes ALS from *Bacillus subtilis* (*B. s.*), was overexpressed to increase carbon flux to 2-acetolactate for the production of isobutyraldehyde and was reported to have relatively high activity.[14] To identify strong ALDC candidates, we used the bioinformatics tool, BRaunschweig ENzyme DAtabase (BRENDA)[34] and a comprehensive literature review. We limited our search to O_2 insensitive enzymes, and looked for reports of strong acetoin production. We were further restricted by the need to match pre-sequencing literature reports to chronologically consistent strain names, which now match currently available gene sequences. Based on these criteria, six *alsD* genes were selected (Table 2-1).

Table 2-1 List of acetolactate decarboxylase and secondary alcohol dehydrogenase used in this study

	Strain source	Cofactor	Chirality	K_m (mM)	k_{cat} (s^{-1})	Reference
ALS						
B. subtilis (*B. s.*)				13.6	121	[35]
ALDC						
E. aerogenes (*E. a.*)	ATCC 13048			NA	NA	[36]
E. cloacae (*E. c.*)	ATCC 13047			NA	NA	[36]
B. licheniformis (*B. l.*)	ATCC 14580			NA	NA	[36]
B. subtilis (*B. s.*)*	str 168			NA	NA	[36]
A. hydrophila (*A. h.*)*	ATCC 7966			NA	NA	[36]
G. xylinus (*G. x.*)	NBRC 3288			NA	NA	[36]
ADH						
C. parapsilosis (*C. p.*)*	M203011	NADPH	*S*	1868	7.4	[37] This study
L. pseudomesenteroides (*L. p.*)*	CHCC 2114	NADPH	*S*	5.1#	18.3	[38]
C. beijerinckii (*C. b.*)*	NRRL B593	NADPH	*R*	8.3	8.2	[39]
T. brockii (*T. b.*)*	HTD4	NADPH	*R*	0.23	0.91	[39]

* Genes were synthesized with codon optimized for expression in *S. elongatus*. #, apparent K_m; NA, not available

To test acetoin production, each *alsD* gene was overexpressed with *alsS* (*B. s.*) under the isopropyl-b-D-thiogalactoside (IPTG) inducible promoter P_{LlacO1}[40] in *E. coli* (Figure 2-3a). The cells were cultured in modified M9 medium (see Methods), containing

50 g/L of glucose, at 30˚C for 16 and 40 h, respectively. A control strain expressing only *alsS* (*B. s.*) produced 0.2 g/L acetoin indicating that 2-acetolactate decomposes to acetoin in small amounts, which is consistent with previous observations.[41,42] When *alsD* was coexpressed, more than 20 g/L of acetoin was produced, indicating that autodecarboxylation is not a major contributor to 2-acetolactate conversion (Figure 2-3b). All ALDC except that from *Enterobacter cloacae* (*E. c.*) were active in *E. coli*, and displayed a pattern of activity that was consistent through 16 h and 40 h of production (Figure 2-3b). The strain expressing *alsD* from *Aeromonas hydrophila* (*A. h.*) was the highest producer (21.0 g/L) followed by the strains expressing *alsD* from *Gluconacetobacter xylinus* (*G. x.*)(17.8 g/L), *alsD* from *Bacillus licheniformis* (*B. l.*)(16.7 g/L), and *alsD* from *Enterobacter aerogenes* (*E. a.*)(16.0 g/L)(Figure 2-3b). The strain expressing codon-optimized *alsD* (*B. s.*), which is the natural gene partner to the *alsS* (*B. s.*) used in the production operon, produced the least acetoin (6.6 g/L), which demonstrates that native pathways do not necessarily maintain their integrity when transferred to new hosts; screening of multiple candidates is necessary to find optimal genes for pathway optimization in each new host (Figure 2-3b).

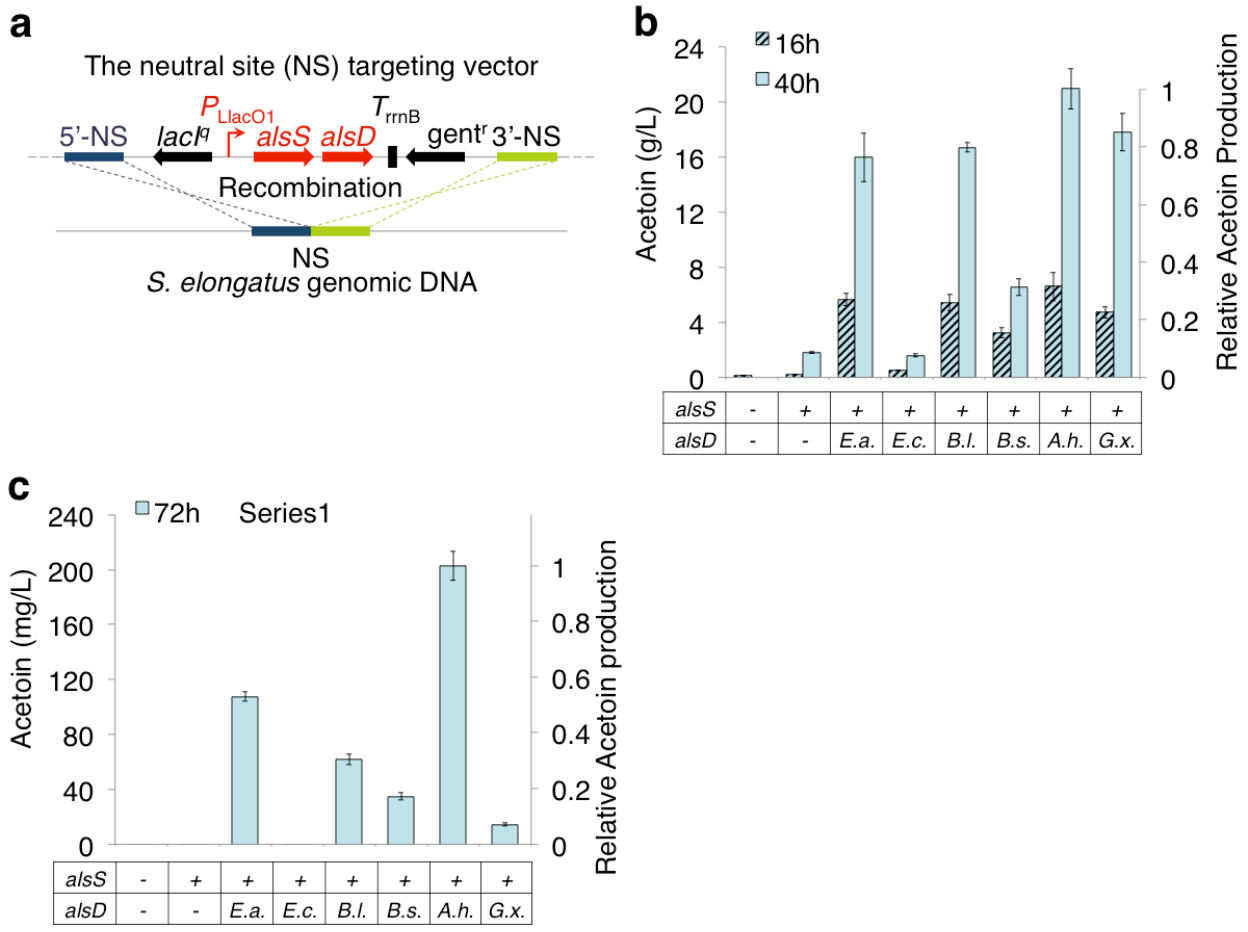


Figure 2-3 Acetoin production in modified strains. (a) Schematic representation of recombination to integrate *alsS* and *alsD* genes into the *S. elongatus* chromosome. (b) Acetoin production in modified *E. coli.* Cells were grown for 16 h (gray) and 40 h (blue). (c) Acetoin production in modified *S. elongatus*. Cells were grown for 72 h. *alsS* indicates inclusion (+) of *alsS* (*B.s.*) or absence (-) of the gene. *alsD* indicates the source organism for the *alsD* gene (Table 1).

We obtained an estimate of activity for ALDC proteins by coexpression with ALS in crude lysate from *E. coli* (Figure 2-4). Our results show a range of activity similar to that of productivity *in vivo*. A significant lead in activity for ALDC (*A. h.*)(388 nmol/min/mg) was seen relative to the moderate activity of ALDC (*E. a.*, *B. l.*, and *G. x.*)(184, 171, and 50 nmol/min/mg respectively), and negligible activity of ALDC (*E. c.*and *B. s.*)(Figure 2-4). Cell extracts of a strain expressing *alsS* in the absence of an *alsD*

gene showed a background rate of 19 nmol/min/mg for spontaneous 2-acetolactate conversion under assay conditions.

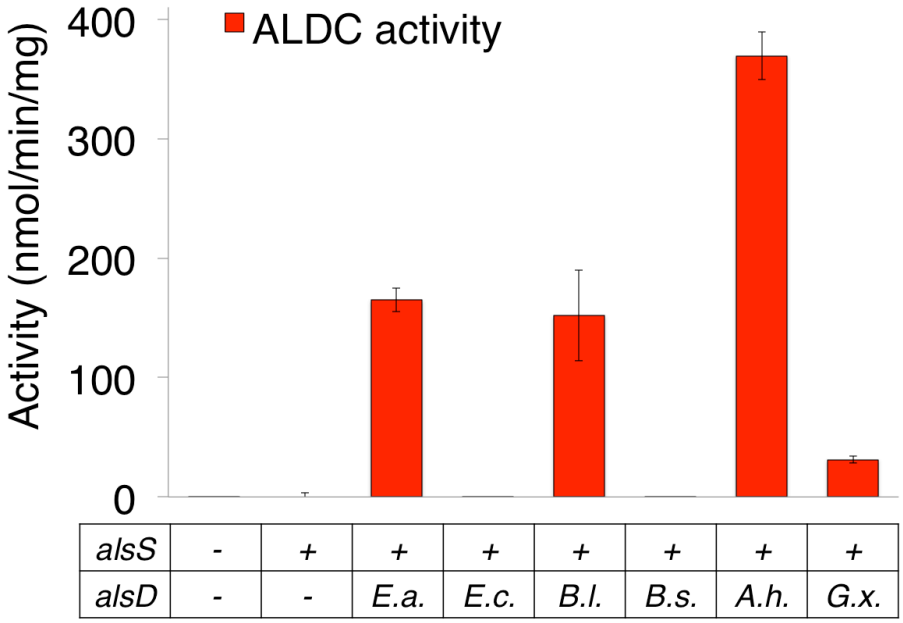


Figure 2-4 Activity of ALDC coexpressed with ALS. Strains containing alsS and alsD were induced for 4 hours and activity was tested in crude lysate. This is an approximation of ALDC activity for internal comparison only. An excess of substrate is assumed due to high ALS activity (300-600 nmol/min/mg) and not quantified. In the case of ALDC (A. h.) substrate may be limiting. Error represents standard deviation between n = 3 cultures.

2.4 Acetoin production in *S. elongatus* from CO_2

Following our screening strategy for pathway optimization, ALDC activity was compared *in vivo* in the photosynthetic cell environment of *S. elongatus,* based on production of acetoin during heterologous *alsS* and *alsD* expression. Each strain was cultured in 125 mL shake flasks with 25 mL BG-11 containing 50 mM $NaHCO_3$ in constant light (55 μE s^{-1} m^{-2}) at 30°C for 72 h (Figure 2-3c). Strains expressing *alsD* from *E. a.*, *B. l.*, *B. s.*, *A. h.*, and *G. x.*, produced 108 mg/L, 62 mg/L, 35 mg/L, 203 mg/L, and

14 mg/L respectively (Figure 2-3c). Control strains, and the strain expressing *alsD* (*E. c.*), did not produce a measureable amount of acetoin in this host. Based on these results, we had two *alsD* genes (from *E. a*. and *A. h.*) capable of moderate and high production of acetoin respectively in *S. elongatus*. To avoid excessive acetoin toxicity we chose *alsD* (*E. a.*) as a starting point for sADH analysis.

2.5 Constructing the 23BD biosynthetic pathway

Acetoin can be reduced by a secondary alcohol dehydrogenase (sADH) to produce 23BD (Figure 2-1). Identification of strong sADH candidates followed the same method used for ALDC, however, in addition to low oxygen sensitivity, two more criteria were added. Firstly, we limited our search to NADPH-dependent sADH as this cofactor is expected to have higher bioavailability during photosynthesis.[22] Secondly, reduction of acetoin by sADH is a diastereoselective reaction, allowing us to choose enzymes to install either an *R* or *S* stereocenter. Two NADPH utilizing sADH with R-installing reaction sites had been characterized previously in *E. coli.*[39] The availability of sADH with S-installing reaction sites and NADPH as a cofactor, however, was limited. In the end we chose four *adh* genes, two with *R*-installing reaction sites, and two with *S*-installing reaction sites (Table 1). Plasmids were constructed harboring *alsS* (*B. s.*), *alsD* (*E. a.*), and each of the four *adh* under P_{LlacO1} (Figure 2-5a)(See Methods).

The resulting *E. coli* strains were cultured in modified M9 medium, containing 50 g/L glucose, at 30˚C for 40 h (Figure 2-5b). The concentration of acetoin remained high for three out of the four strains indicating that sADH activity is limiting in this cell environment. The fourth strain, expressing *adh* (*Clostridium beijerinckii* (*C. b.*)), maintained a relatively low acetoin concentration (less than 6% of total production) and

produced 13.8 g/L total 23BD as a mixture of (*R,R*)-23BD and *meso*-23BD stereoisomers forming 74% and 21% of total production respectively (Figure 2-5b). The strains expressing *adh* (*Thermoanaerobium brockii* (*T. b.*)) and *adh* (*Candida parapsilosis* (*C. p.*)) produced 2.4 g/L and 14.2 g/L 23BD respectively with both isomers formed in roughly equal amounts in each. High stereoselectivity was achieved in the strain expressing *adh* (*Leuconostoc pseudomesenteroides* (*L. p.*)), which produced 9.1 g/L meso-23BD exclusively. Enzyme activities measured from crude cell lysate isolated during production were high for both sADH (*C. p.*) and sADH (*C. b.*), at 265 and 440 nmol min^{-1} mg^{-1} respectively, when excess substrate was used (Figure 2-5b). However, for the strain expressing *adh* (*C. p.*), accumulation of acetoin in the supernatant during production indicates that the enzyme turnover rate at the substrate concentrations present within the cell is comparable or slower than the rate of secretion. Relatively low activity of sADH (*T. b.*) was consistent with accumulation of acetoin in the strain indicating that sADH activity is a bottleneck for production in this *E. coli* strain (Figure 2-5b). The major 23BD product of each of the *adh* expressing strains except the strain containing *adh* (*C. p.*) matched the stereochemistry predicted by previous characterization.[37-39]

2.6 23BD production in *S. elongatus* from CO_2

To screen the differences in 23BD productivity in *S. elongatus* each of the plasmids used for 23BD production in *E. coli* was used for transformation of *S. elongatus* (see Methods). The engineered strains were cultured in 125 mL baffled shake flasks with 25 mL BG-11 containing 50 mM $NaHCO_3$ in constant light (55 μE s^{-1} m^{-2}) at 30°C.

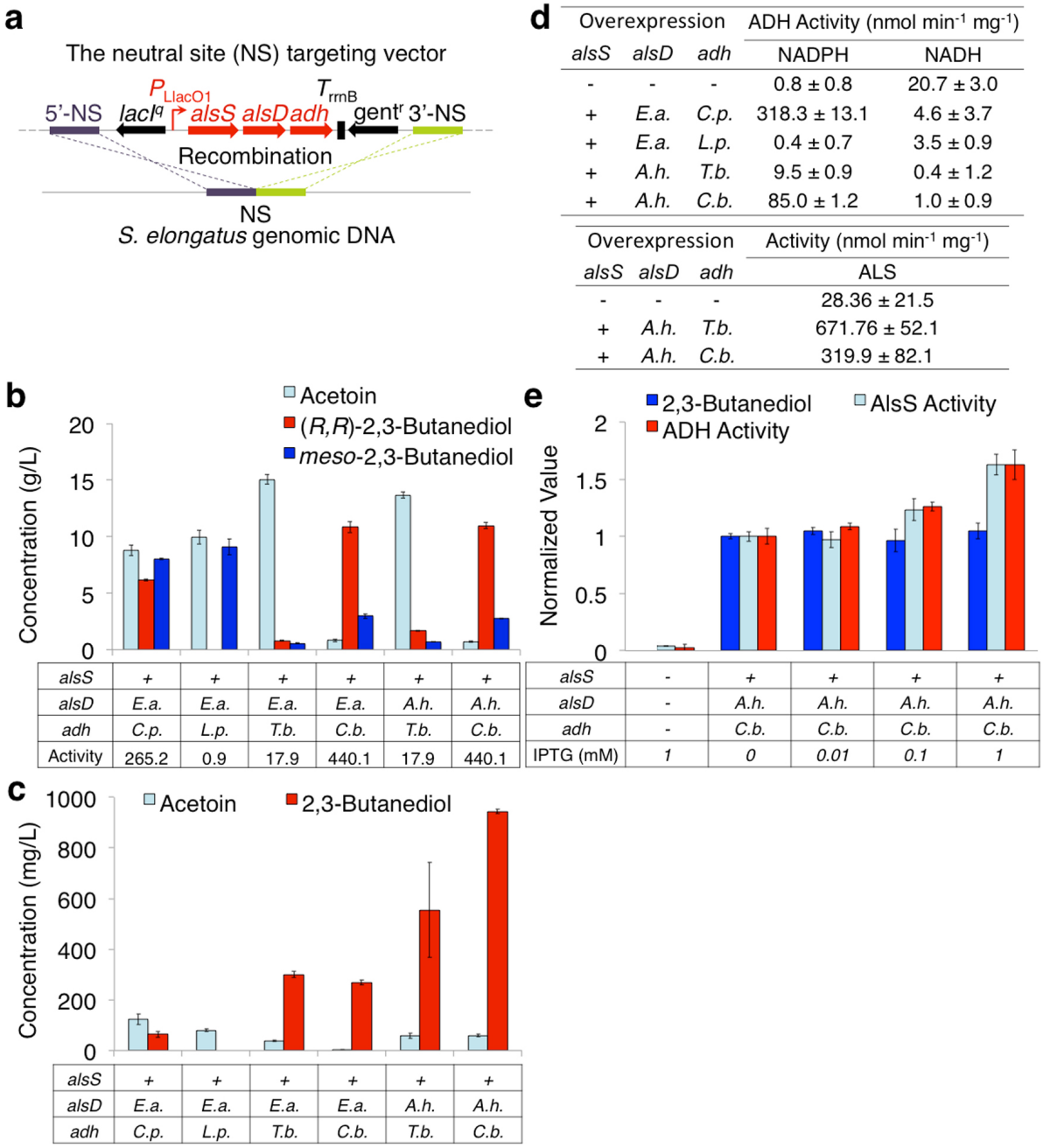


Figure 2-5 2,3-Butanediol production in modified strains. (**a**) Schematic representation of recombination to integrate *alsS*, *alsD*, and *adh* genes into the *S. elongatus* chromosome. (**b**) Production in modified *E. coli.* Cells were grown for 40 hours. Bars represent acetoin (light blue), *meso*-23BD (red), and (*R,R*)-23BD (dark blue). *alsS* indicates inclusion (+) of *alsS* (*B.s.*). *alsD* and *adh* rows indicate the source organism for the gene (Table 1). Activity is that of sADH expressed in *E. coli* and measured in cell extract (nmol NADPH min^{-1} mg^{-1}). (**c**) 23BD production in modified *S. elongatus*. Cells were grown for 72 h. (**d**) Specific activities of ALS and sADH in cell extracts from modified *S. elongatus* strains. (**e**) Effect of IPTG in modified *S. elongatus*. Cells were grown for 72 h after induction with the specified concentration of IPTG. Values are normalized to those from uninduced cultures. Bars represent 23BD production (dark blue), Activity of sADH protein measured with NADPH as cofactor (red), and Activity of ALS protein (light blue). Error indicates s.d. (n=3).

23BD production was detected in three out of four *S. elongatus* strains (Figure 2-5c). Measurement of sADH performance in *S. elongatus* was made by comparison of acetoin and 23BD concentrations after 72 hours of growth, using the less active ALDC (*E. a.*) to lower toxicity in cases when acetoin conversion was low. The strain expressing *adh* (*T. b.*) produced 301 mg L^{-1} (*R,R*)-23BD with trace amounts of *meso*-23BD but also allowed for accumulation of acetoin (Figure 2-5c). The strain expressing *adh* (*C. b.*) produced 270 mg/L (*R,R*)-23BD (major) and undetectable levels of acetoin indicating high flux through the intermediate. The strain expressing *adh* (*C. p.*) produced 65 mg L^{-1} 23BD with *meso*-23BD as the primary product and accumulated toxic levels of acetoin. The remaining *S*-installing *adh* (*L. p.*) was not active in *S. elongatus,* resulting only in accumulation of acetoin. Enzyme activities measured in crude cell lysate isolated during production showed a roughly 9-fold higher activity for sADH (*C. b.*) than for sADH (*T. b.*), 85.0 and 9.5 nmol min^{-1} mg^{-1} respectively, which could explain the accumulation of acetoin in the less active strain (Figure 2-5c&d). Activity for sADH (*C. p.*) was roughly 5-fold higher than sADH (*C. b.*); however, low production and acetoin accumulation was observed, similar to the result in *E. coli*. Enzyme activity could not be detected for the strain expressing *adh* (*L. p.*), indicating that the sADH enzyme is responsible for lack of production (Figure 2-5c&d).

The two sADH with highest production and lowest acetoin accumulation were further tested with the more active ALDC encoded by *alsD* (*A. h.*). Both strains increased production, yielding 568 mg L^{-1} from *adh* (*T. b.*) and 952 mg L^{-1} from *adh* (*C. b.*), the latter of which is 3 fold higher than production with *alsD* (*E. a.*) over 72 hours (Figure

2-5c). Both strains also showed increased acetoin concentrations, although neither reached toxic levels, accumulating 59 mg L^{-1} and 61 mg L^{-1} respectively.

In order to have an inducible expression system, *lacI*^{q}, which encodes the *E. coli* lac repressor, was cloned upstream of P_{LlacO1} (Figure 2-3a). The efficiency of LacI repression in the *S. elongatus* strain containing *alsS* (*B. s*), *alsD* (*A. h.*), and *adh* (*C. b.*) was investigated by testing 23BD production with or without various concentrations of IPTG (Figure 2-5e). Interestingly, 23BD production without IPTG was similar to that with IPTG (Figure 2-5e), suggesting that P_{LlacO1} was leaky in this construct. The promoter and coding region of *lacI*^{q} were verified by Sanger sequencing. This phenomena has been reported with other IPTG-inducible promoters in *S. elongatus* PCC7942 [43,44] and *Synechocystis* sp. PCC6803.[45] However, ALS and sADH activities increased with elevated concentrations of IPTG (Figure 2-5e). ALS and sADH activities both increased 1.2- and 1.6- fold with 0.1 and 1.0 mM IPTG, respectively, over that of cultures without IPTG induction (Figure 2-5e). These results indicate that LacI repressed P_{LlacO1} in *S. elongatus* and IPTG induced expression from P_{LlacO1}. From this we infer that leaky expression of genes under P_{LlacO1} without IPTG is significant, even though LacI repression is functioning.

To estimate the amount of installed enzymes translated in engineered *S. elongatus*, cell extracts of the engineered strain containing *alsS* (*B. s.*), *alsD* (*A. h.*), and *adh* (*C. b.*), and non-engineered wild-type strains were analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)(Figure 2-6). The bands of ALS were detected only in samples from the engineered strain. The amounts of ALS are estimated at about 0.21% without IPTG, and 0.28, 0.37 and 0.34% with 0.01, 0.1, and 1

mM IPTG, respectively, of total protein. The bands of ALDC and sADH were not visible on the gel, corresponding to less than 0.02% of total protein.

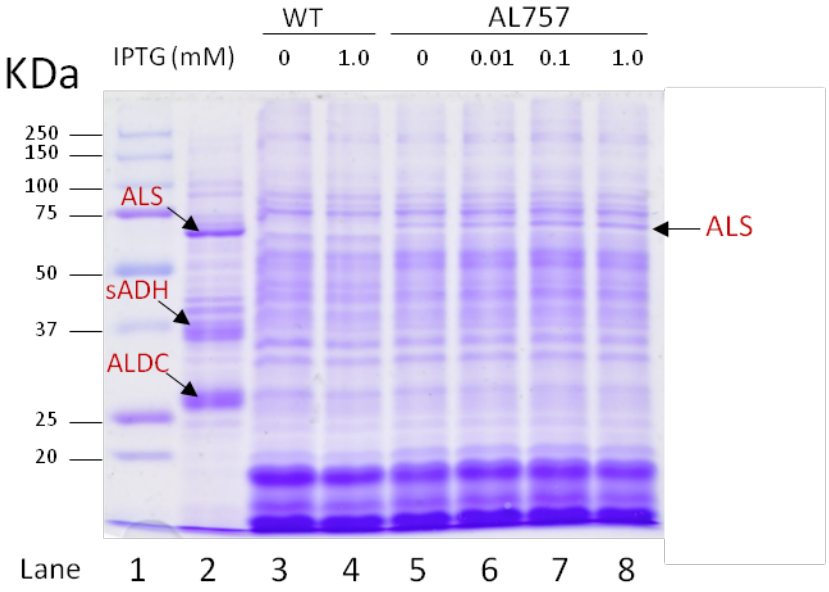


Figure 2-6 SDS-PAGE analysis of the 23BD strain. Cell extracts of the wild-type and AL757 were analyzed with SDS-PAGE. Lane 1, size ladder; Lane 2, the mixture of cell extracts of E. coli overexpressing alsS, alsD (A. h.), and adh (C. b.); Lane 3 and 4, cell extracts of the wild-type without (Lane 3) and with 1.0 mM IPTG (Lane 4); Lane 5-8, cell extracts of AL757 without IPTG (Lane 5), and with 0.01 (Lane 6), 0.1 (Lane 7), and 1.0 mM (Lane 8) IPTG. The calculated molecular weights of AlsS, AlsD, and Adh are 62.1, 37.7, and 29.0 KDa, respectively.

2.7 Long-term production of 23BD in *S. elongatus*

We verified stability of the highest producing strains by maintaining continuous production in 25 ml cultures at 30°C in the presence of constant light (Figure 2-7)(see Methods). The strain containing *adh* (*C. b.*) reached a total yield of 2.38 g L^{-1} (*R,R*)-23BD and a maximum production rate of 9847 mg L^{-1} h^{-1} (3 day average)(Figure 2-7b). Production was sustained for 21 days. The strain containing *adh* (*T. b.*) showed similar results reaching a total yield of 1.97 g L^{-1}, achieving a maximum production rate of 7757 mg L^{-1} h^{-1} (3 day average), and sustaining production for 21 days (Figure 2-7b). After 21 days, production in both cultures dropped off sharply and was not restored when cells

were resuspended in fresh medium, indicating that changes in the culture population, such as spontaneous mutations that restore flux to metabolism, fundamentally impair production over time. Strains containing the 23BD biosynthetic pathway showed reduced growth compared to control strains, as expected mirroring the rate of carbon redirection from central metabolism (Figure 2-7a, b&d). A second control strain containing only *alsS* (*B. s.*) and *alsD* (*E.a.*) produced acetoin up to toxic levels after the stationary phase was reached (Figure 2-8) and showed impaired growth beyond what is attributable to carbon redirection.

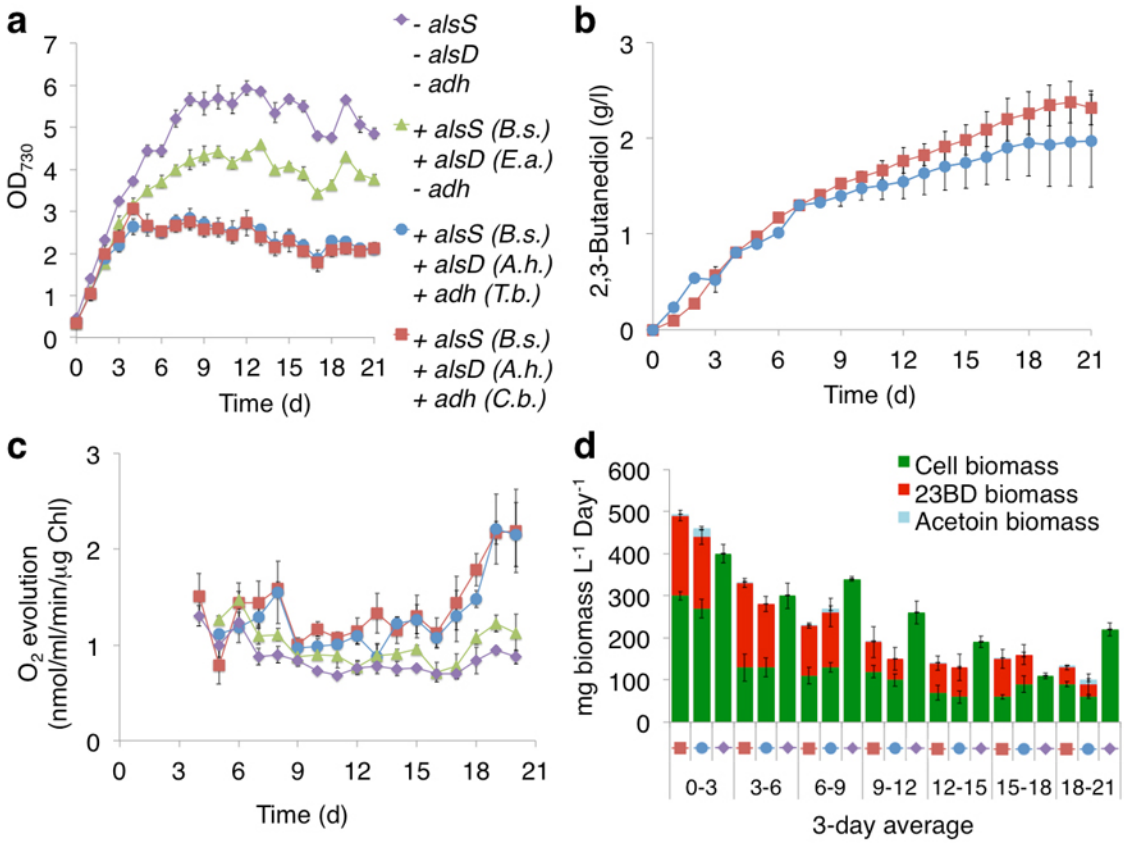


Figure 2-7 Long term 2,3-Butanediol production. (**a-c**) Summary of results for 23BD production in continuous cultures. Red: *S. elongatus* containing *alsS* (*B. s.*), *alsD* (*A. h.*) and *adh* (*C. b.*). Blue: *S. elongatus* containing *alsS* (*B. s.*), *alsD* (*A. h.*) and *adh* (*T. b.*). Green: *S. elongatus* containing *alsS* (*B. s.*) and *alsD* (*E. a.*). Purple: *S. elongatus* without *alsS*, *alsD* or *adh*. (**a**) Time courses for growth. (**b**) Total 23BD production. (**c**) Photosynthetic efficiency. (**d**) Total biomass production per day. Error bars indicate s.d. (n = 3).

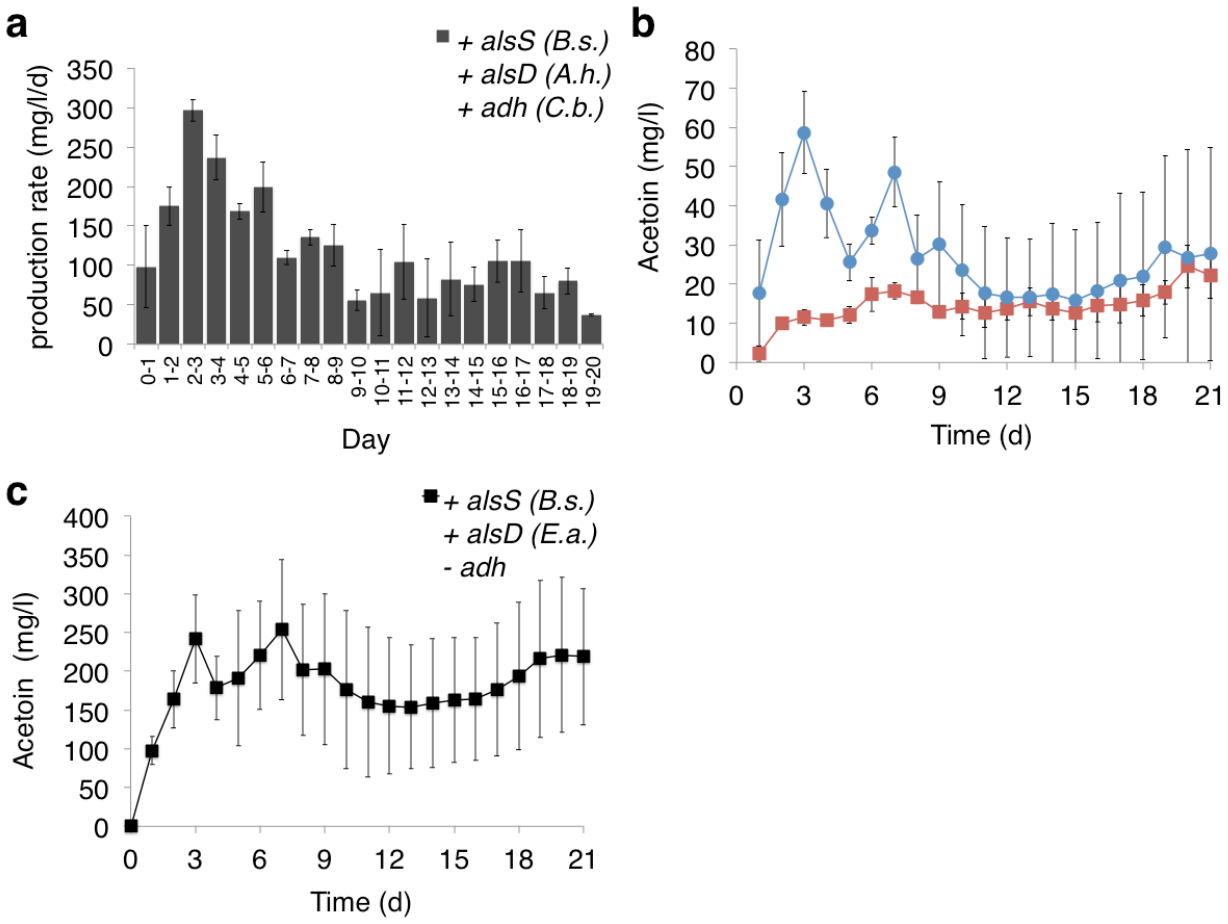


Figure 2-8 23 BD and acetoin production in *S. elongatus*. (a) 23BD daily production by AL757 (with integrated alsS, alsD (A. h.), and adh (C. b.))(Figure 2-7) (b) Acetoin concentration during long-term 23BD production experiments (Figure 2-7). Red squares: AL757 (with integrated alsS (B. s.), alsD (A. h.) and adh (C. b.)). Blue circles: AL756 (with integrated alsS (B. s.), alsD (A. h.) and adh (T. b.)). (c) Long-term acetoin production by AL763 (with integrated alsS and alsD (E. a.)).

2.8 Evaluating the photosynthetic efficiency of production strains

Evolution of O_2 from illuminated cells during continuous production was measured to verify whether the 23BD overproduction pathway could affect the photosynthetic system (Figure 2-7c). Both strains expressing the 23BD biosynthetic pathway displayed a slightly higher rate of O_2 evolution per mg of chlorophyll compared to control strains (Figure 2-7c). This rate increased during late stages of production. Both control strains, each with no production pathway, or only the acetoin production pathway expressed,

displayed similar rates of O_2 evolution (Figure 2-7c). This trend follows the amount of fixed carbon diverted away from central metabolism, indicating that the burden placed on the cell by overproduction could stimulate a positive effect on the cells photosynthetic efficiency (Figure 2-7c&d).

2.9 Discussion

In this study, we describe for the first time, production of 23BD and acetoin directly from CO_2 and light, through engineering of the cyanobacterium, *S. elongatus*. We systematically approached pathway design to match the production pathway to a photosynthetic host. Engineered strains achieved a production rate of 9847 mg l^{-1} h^{-1} and final titer of 2.4 g l^{-1}, with sustained production for 21 days. These values, achieved during continuous production from CO_2 and light, compare favorably with other studies. The rate is 1.6 fold higher than that for isobutyraldehyde (6,230 mg l^{-1} h^{-1}), and significantly higher than other products overproduced from exogenous pathways (Figure 2-9). The percentage of biomass produced as 23BD ranges from 30% to 60% (Figure 2-7d), which compares favorably to the maximum of 80% achieved during endogenous sucrose overproduction.[43]

To construct the 23BD pathway, we theorized that low toxicity would improve culture sustainability and production in *S. elongatus*. The negative effect of toxicity on pathway flux is reinforced by low production of acetoin, which is toxic above 0.1 g/L in *S. elongatus* (Figure 2-2.a.), from the 23BD pathway without coexpression of *adh* (Figure 2-8c). Addition of a strong sADH to the pathway to convert acetoin to 23BD increases total production ten-fold (Figure 2-7b and S4c), even though reduction by sADH is not an

irreversible step, while production of acetoin is. We also proposed that matching genes to their host could improve pathway function. All genes were screened for 23BD and acetoin production in *E. coli* concurrently with cyanobacteria using identical operons. The patterns of production exhibited by the genes were different between hosts (Figure 2-3&Figure 2-5). The productivities of strains expressing *alsD* (*B. l.*) and *alsD* (*G. x.*) in *S. elongatus* were much lower (30% and 7% of top production respectively)(Figure 2-3c), than strains overexpressing the same genes in *E. coli* (80% and 85% of top production respectively)(Figure 2-3b). Conversely sADH (*T. b.*), which displayed severely attenuated production in *E. coli* achieved significant production in *S. elongatus*. Additionally, the enzyme encoded by *adh* (*L. p.*) was entirely inactive in *S. elongatus* despite production of 9.1 g L^{-1} *meso*-23BD in *E. coli*.

Production of 23BD was increased in strains expressing *alsD* (*A. h.*), indicating that the second step of the pathway, catalyzed by ALDC (*E. a.*) was limiting in the strain containing *alsS*, *alsD* (*E. a.*), and *adh* (*T. b.* or *C. b.*) (Figure 2-5c). Gradient analysis of the IPTG expression system using a strain expressing *alsS* (*B. s.*), *alsD* (*A. h.*), and *adh* (*C. b.*) (Figure 2-5e), showed that even though enzyme concentrations in total protein increased, as measured by activity, the amount of 23BD produced did not increase. This implies that the amount of enzyme translated through leaking of the P_{lacO1} was sufficient to achieve the maximum production of 23BD (Figure 2-5e). From these results we assume that the supply of pyruvate from the Calvin-Benson-Bassham (CBB) cycle[46] is limiting in this strain. Thus, further enzyme or pathway optimization would likely not be effective until the substrate pool of pyruvate from carbon fixation can be increased.

Using 23BD production as a model system allowed for the inclusion of stereoselectivity as part of our pathway design (Figure 2-1). Chirality can be costly to install in chemical synthesis; however, biological control offers a much simpler route to chiral products. In all known cases in nature, acetoin is generated from 2-acetolactate containing an *R*-stereocenter[47] resulting in (*S,S*)-23BD not being observed. However, autodecarboxylation of 2-acetolactate, or enolate racemization of acetoin, could possibly form (*S*)-acetoin in the cell and result in (*S,S*)-23BD production in the presence of *S*-installing sADH enzymes. We designed two pathways, one for each possible stereocenter formed during reduction (Figure 2-5 & Table 1). In *S. elongatus* both R-installing strains tested in long-term production consistently produced (*R,R*)-23BD as the major product (Figure 2-7), although a trace amount of *meso*-23BD was observed in production by the strain expressing *adh* (*T. b.*). The strain containing *adh* (*C. p.*) produced mixed isomers in *E. coli* and in *S. elongatus* (Figure 2-5). These results indicate that sADH (*C. p.*) does not show stereoselectivity toward (*R*)-acetoin, although this enzyme catalyzes an *S*-installing reaction toward 2-hydroxyacetophenone[37]. We assume that (*R*)-acetoin is too small to bind to the catalytic site stereospecifically. In this study no production of (*S,S*)-23BD was detected, indicating that degradation products do not contribute significantly to the pathway. With further engineering to remove endogenous *S. elongatus* sADH activity, observable in enzyme assays with NADH as a cofactor (Figure 2-5d), it is likely that the purity of these chiral products can be increased.

During *in vivo* production acetoin is continuously secreted as evidenced by equilibration into culture supernatant (Figure 2-5 & Figure 2-8b). We can infer from this that enzymes with low velocity at intracellular acetoin concentrations or high K_m may not

be effective. Reported K_m for sADH from *C. b.* and *T. b.* are 8.3 mM, and 0.23 mM respectively[39], while reported k_{cat} are 8.2 and 0.91 respectively. The higher K_m of sADH (*C. b.*) may explain why production from this strain surpassed that of the strain expressing *adh* (*T. b.*) only after acetoin was increased by inclusion of *alsD* (*A. h.*). This may also explain similar production from both strains in *S. elongatus*, but differing production in *E. coli* where acetoin is more abundant within the cell due to higher supernatant concentration.

Enzyme activities measured in crude *S. elongatus* cell lysate were roughly four-fold higher for sADH (*C. p.*) than for sADH (*C. b.*) at 318.3 and 85.0 nmol min^{-1} mg^{-1}, respectively (Figure 2-5d), although the strain containing *adh* (*C. p.*) produced 4-fold less 23BD than that of the strain containing *adh* (*C. b.*)(Figure 2-5c). The K_m of sADH (*C. p.*) was much larger than those of other sADHs (Table 1), given that the native substrate of the enzyme has a phenyl ring adjacent to the reaction site and binding for acetoin without the phenyl ring may be inefficient.[37] Accumulation of acetoin in the strain expressing *adh* (*C. p.*) results in up to 120 mg/L during production. Assuming uniform diffusion this corresponds to 1.1 mM concentrations in the cell. In this case the enzyme would have difficulty competing with diffusion of acetoin, yet would display high activity under assay conditions where acetoin concentration is high (25 mM).

During long-term production, evolution of O_2 per mg of chlorophyll increased in production strains relative to a control containing the recombination cassette but no *alsS*, *alsD*, or *adh* genes (Figure 2-7c). This indicates that the stress imposed on metabolism by production elicits an increase in photosynthetic efficiency. Chlorophyll and O_2

production have been seen to increase in roughly equal amounts during similar overproduction of sucrose in engineered *S. elongatus*.[43]

For further improvements of the productivities, it is important to gain understanding of the engineered strains at the systems level using systematic and quantitative analysis methods. To identify modifications to improve the productivities, we will compare the engineered strains with the wild type at the transcriptome, proteome, and metabolome levels. It is also important to increase a number of available parts such as promoter and RBS. These parts for pathway construction in *S. elongatus* are limited compared to model organisms such as *E. coli*. In this study, we used a promoter, P_{LlacO1}, because this promoter has been shown to be functional in *S. elongatus*.[14] However, we showed P_{LlacO1} was leaky and had a narrow expression range in *S. elongatus* (Figure 2-5e). Additionally, it has been known that the composition of the cyanobacterial holopolymerase is quite different from those in most bacteria, thus *E. coli* promoters might perform differently when introduced into *S. elongatus*.[11] To achieve efficient production, more controllable systems for expression of synthetic pathways would be desirable.

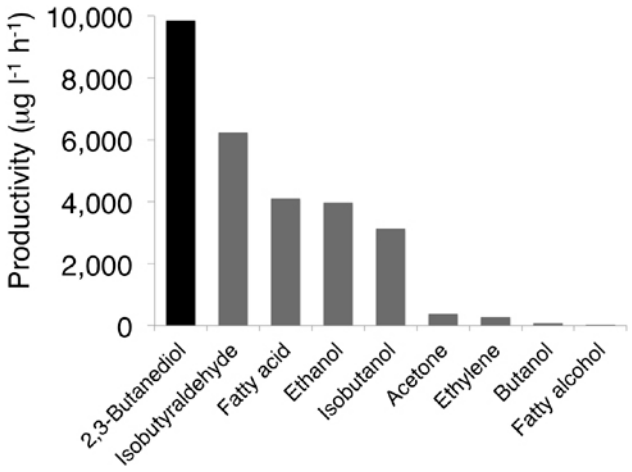


Figure 2-9 Comparison of productivities for chemicals produced from engineered cyanobacteria. Source data is obtained from: 23BD (this work), isobutyraldehyde and isobutanol [14], fatty acid [19], ethanol [48], acetone [18], ethylene [16], butanol [15], fatty alcohol [20]. The detailed calculation is described in Supporting Notes.

2.10 Conclusion

Defining the engineering principles for photosynthetic organisms is an important landmark in the search for sustainable technologies. Biological production of 23BD by heterotrophic microbes has attracted attention for many years because of the existence of natural fermentative producers, and the chemical's potential as a versatile carbon feedstock for plastics, solvents, and fuel.[24,25] The biosynthetic production rate and titer presented here mark the second large increase in cyanobacterial yields in the last 3 years, due to increased understanding in pathway design.

2.11 Specific Methods

2.11.1 Plasmid construction

Plasmids used and constructed are listed in Table 2-2. All plasmids except pAL60 were constructed using sequence and ligase independent cloning (SLIC)[49] in *E. coli* XL1-Blue

(Agilent Technologies, Santa Clara, CA). Primers for constructions and genotype verifications are listed in Table S2.

A neutral site (NS) located between Synpcc7942_0893 (903,564 - 904,283 bp) and Synpcc7942_0894 (904,845 - 905,417 bp) in the *S. elongatus* chromosome was used for insertion of an expression cassette. This region was amplified with primers MC173 and MC176. PCR products were digested with AatII and AvrII and cloned into pZE12-luc[40] cut with the same enzyme, creating pAL60.

The fragment containing P_{LlacO1} and *alsS* (*B.s.*) genes was amplified with primers IM103 and IM11 and *lacI*^{q} was amplified with primers IM39 and AK3 from pSA69.[33] The resulting fragments were inserted into pAL60 by SLIC, creating pAL301.

To clone *alsD* (*E. a.*), we used genomic DNA of *E. aerogenes* ATCC13048 (ATCC) as a PCR template with primers IM16 and IM17. To clone *alsD* (*E. c.*), genomic DNA of *E. cloacae* ATCC 13047 (ATCC) was used as a PCR template with primers IM19 and IM20. To clone *alsD* (*B. l.*), genomic DNA of *B. licheniformis* ATCC 14580 (ATCC) was used as a PCR template with primers IM23 and IM24. To clone *alsD* (*G. x.*), genomic DNA of *G. xylinus* (NBRC 3288) was used as a PCR template with primers IM21 and IM22. *alsD* (*B. s*) and *alsD* (*A. h.*) were chemically synthesized by DNA2.0 Inc. (Menlo Park, CA) to optimize codon usage for *S. elongatus*. Each *alsD* gene was cloned into downstream of *alsS* (*B. s.*) on pAL301 by SLIC, creating pAL302, pAL303, pAL304, pAL305, pAL306 and pAL307 (Table S1). To construct plasmid pAL312, we used plasmid pAL301 as a PCR template and primers IM114 and IM11 to amplify the entire plasmid, without the *alsS* gene. The resulting fragment was assembled by SLIC. All four *adh* genes were chemically synthesized by DNA2.0 Inc. (Menlo Park, CA) to

optimize codon usage for S. *elongatus*. Each *adh* gene was cloned into downstream of *alsD* (*E. a.*) on pAL302 by SLIC, creating pAL308, pAL309, pAL310 and pAL315 (Table S1). The *adh* (*T.b.*) and *adh* (*C. b.*) genes were clone into downstream of *alsD* (*A. h.*) on pAL306 by SLIC, creating pAL299 and pAL300, respectively. To purify sADH (*C. p.*), a 6xHis tag was fused to the C-terminus. The *adh* (*C. p.*) gene was amplified with primers IM150 and IM151 from pAL308. pETDuet-1 (Novagen) was amplified with primers IM152 and IM153. The two fragments were fused by SLIC, creating pAL410.

Table 2-2 Strains and plasmids used in this study

Strains	Genotype	Source
***E. Coli* strain**		
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10* (*Tetr*)]	Agilent Technologies
***S. elongatus* strains**		
PCC7942	wild type	S. Golden
AL723	P_{LlacO1} and Gent^R integrated at NS	This study
AL756	*alsS* (*B. s.*)-*alsD* (*A. h.*)-*adh* (*T. b.*) integrated at NS	This study
AL757	*alsS* (*B. s.*)-*alsD* (*A. h.*)-*adh* (*C. b.*) integrated at NS	This study
AL762	*alsS* (*B. s.*) integrated at NS	This study
AL763	*alsS* (*B. s.*)-*alsD (E. a.)* integrated at NS	This study
AL764	*alsS* (*B. s.*)-*alsD* (*E. c.)* integrated at NS	This study
AL765	*alsS* (*B. s.*)-*alsD* (*B. l.*) integrated at NS	This study
AL766	*alsS* (*B. s.*)-*alsD* (*B. s.*) integrated at NS	This study
AL767	*alsS* (*B. s.*)-*alsD* (*A. h.*) integrated at NS	This study
AL768	*alsS* (*B. s.*)-*alsD* (*G. x.*) integrated at NS	This study
AL769	*alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*C. p.*) integrated at NS	This study
AL770	*alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*L. p.*) integrated at NS	This study
AL771	*alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*T. b.)* integrated at NS	This study
AL772	*alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*C. b.*) integrated at NS	This study
Plasmids		
pAL60	NS targeting vector; ColE1 ori; *P*trc; Amp^R	This study
pSA69	P15A ori; Kan^R, $P_L$lacO$_1$:*alsS* (*B. s.*)-*ilvC* (*E. coli*)-*ilvD* (*E. coli*)	[33]
pAL299	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*A. h.*)-*adh* (*T. b.*); Gent^R	This study
pAL300	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*A. h.*)-*adh* (*C. b.*); Gent^R	This study
pAL301	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*); Gent^R	This study
pAL302	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. a.*); Gent^R	This study
pAL303	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. c.*); Gent^R	This study
pAL304	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*B. l.*); Gent^R	This study

pAL305	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*B. s.*); $Gent^R$	This study
pAL306	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*A. h.*); $Gent^R$	This study
pAL307	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*G. x.*); $Gent^R$	This study
pAL308	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*C. p.*); $Gent^R$	This study
pAL309	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*L. p.*); $Gent^R$	This study
pAL310	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*C. b.*); $Gent^R$	This study
pAL312	As pAL60, but P_{LlacO1}: $Gent^R$	This study
pAL315	As pAL60, but P_{LlacO1}: *alsS* (*B. s.*)-*alsD* (*E. a.*)-*adh* (*T. b.*); $Gent^R$	This study
pAL410	Derivative of pETDuet-1 with *adh* (*C. p.*)	This study

Table 2-3 Oligonucleotides used in this study

Name	Sequence 5' → 3'	Plasmid constructed
MC173	CTATGACGTCGGCGTTTTCTGCTACATGGGCCGTGAG	pAL60
MC176	CTAACCTAGGGGAAGTCCAGCGCAATCAGCGGAGTTG	pAL60
MC181	CTAAATGCATTAAGTTGTTACTAGTGCTTGGATTCTCACC	
IM15	CTAGAGAGCTTTCGTTTTCATGAG	
IM88	CGGGGAACTCATGAAAACGAAAGCTCTCTA	
IM89	GTTGGGTAGCAGACAATGCGGGGGATCTGG	
AK1	GGTTTTGCACCAGGATCCCGCTCGAGTTGACGCGTGCTTA	pAL301
AK3	TCACTGCCCGCTTTCCAGTC	pAL301
IM11	AGATCCTCTAGAGTCGACCTG	pAL299, pAL300, pAL312
IM15	CTAGAGAGCTTTCGTTTTCATGAG	pAL302
IM16	TAGGTCGACGAGGAATCACCATGAATCATGCTTCAG	pAL302
IM17	AGGTCGACTCTAGAGGATCTCTAACTTTCTACTGAACGGA	pAL302
IM19	TAGGTCGACGAGGAATCACCATGAGCGCCCTGCTAA	pAL303
IM20	CAGGTCGACTCTAGAGGATCTTTAGTTTTCGACGGA	pAL303
IM21	TAGGTCGACGAGGAATCACCATGGAAATAGGCTTTA	pAL307
IM22	CAGGTCGACTCTAGAGGATCTTCAGCCGCCCTCGGC	pAL307
IM23	CTAGGTCGACGAGGAATCACCATGAAAAGTGCAAG	pAL304
IM24	CAGGTCGACTCTAGAGGATCTTTACTCGGGATTGCCT	pAL304
IM27	TAGGTCGACGAGGAATCACCATGAAACGTGAGTCG	pAL305
IM28	CAGGTCGACTCTAGAGGATCTCTACTCGGGAGAACC	pAL305
IM29	TAGGTCGACGAGGAATCACCATGGAAACTAATAGC	pAL306
IM30	CAGGTCGACTCTAGAGGATCTCTAACCCTCAGCCGC	pAL306
IM39	CGGGATCCTGGTGCAAAACCTTTCGCGGTA	pAL301
IM44	GTACCTTTCTCCTCTTCTAACTTTCTACTGAACGGATGGC	pAL308, pAL309, pAL310, pAL315
IM45	TAGAAGAGGAGAAAGGTACATGAAAGGTTTTGCCA	pAL310
IM46	CAGGTCGACTCTAGAGGATCTCTACAGGATTACGAC	pAL310
IM47	GTTAGAAGAGGAGAAAGGTACATGAAGGGTTTCGC	pAL315
IM48	CAGGTCGACTCTAGAGGATCTCTATGCCAAAATGAC	pAL315
IM49	TTAGAAGAGGAGAAAGGTACATGGGGGAGATTGAG	pAL308
IM50	CAGGTCGACTCTAGAGGATCTCTAGGGGCATGTGTAA	pAL308
IM51	TTAGAAGAGGAGAAAGGTACATGACAAAGAAAGT	pAL309
IM52	AGGTCGACTCTAGAGGATCTCTAGTGAAACTGCATG	pAL309
IM103	AGTTGACGCGTGCTTATCATAATTGTGAGCGGATAACAAT	pAL301

IM114	GGTCGACTCTAGAGGATCTTGTACCTTTCTCCTCTTTAA	pAL312
IM125	GTACCTTTCTCCTCTTCTAACCCTCAGCCGCACGGATAGC	pAL299, pAL300
IM150	AGAAGGAGATATACCATGGGGGAGATTGAGTCCTA	pAL410
IM151	CTAGTGGTGATGATGGTGATGGGGGCATGTGTAACCGCCATCGATA	pAL410
IM152	ATCACCATCATCACCACTAGGCTGAGCAATAACTAGCATA	pAL410
IM153	CCCATGGTATATCTCCTTCTTAAAGTTAAAC	pAL410

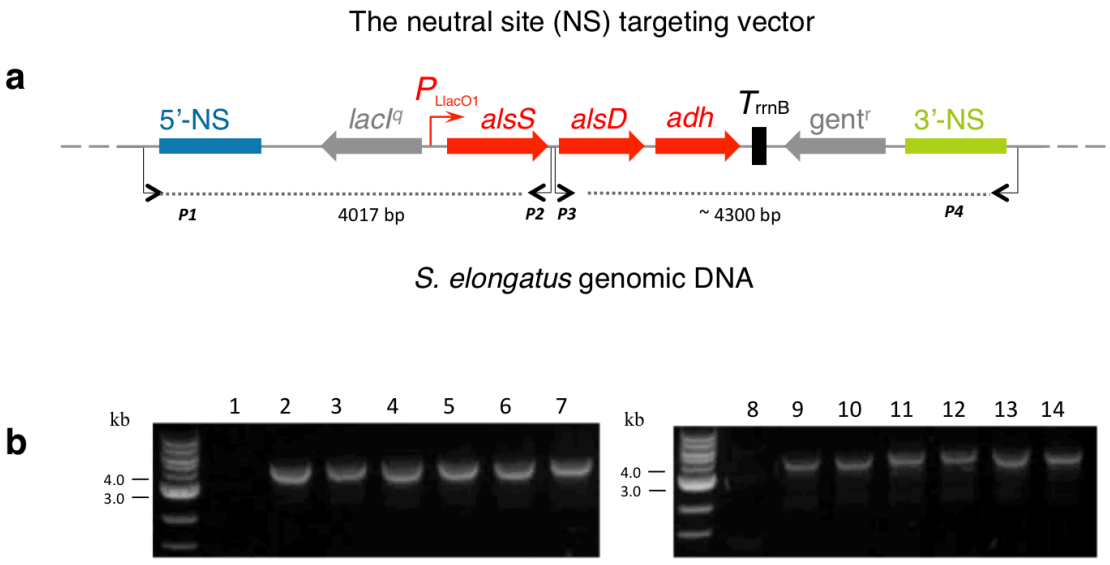


Figure 2-10 PCR confirmation of the integration of 23BD pathway genes into the S. elongatus chromosome. (a) Schematic representation showing the integration of the 23BD pathway genes alsS, alsD and adh into NS on the S. elongatus chromosome. (b) Confirmation of correct recombinants. PCR performed with specific primers listed in Table S2, with genomic DNA of S. elongatus strains (Table S1) as a template. Left Panel: Result of PCR with P1 and P2 to demonstrate integration of lacIq and alsS. Lane 1: wild type, lane 2: AL769, lane 3: AL770, lane 4: AL771, lane5: AL753, lane 6: AL756 and lane 7: AL757. Right panel: Result of PCR with P3 and P4 to demonstrate integration of alsD and adh genes. Lane 8: wild type, lane 9: AL769, lane 10: AL770, lane 11: AL771, lane 12: AL753, lane 13: AL756, lane 14: AL757.

2.12 Acknowledgements

This work was supported by Asahi Kasei Corporation. We thank Jordan McEwen and Mike Connor for characterization of NS used in this study.

2.13 References

1 McFarlane, J. & Robinson, S. Survey of Alternative Feedstocks for Commodity Chemical Manufacturing. *Oak Ridge National Laboratory* (2007).

2 Serferlein, K. E. Annual Energy Review (E.I Administration). (2008).

3 Raupach, M. R. *et al.* Global and regional drivers of accelerating CO_2 emissions. *Proc Natl Acad Sci U S A* **104**, 10288-10293, doi:10.1073/pnas.0700609104 (2007).

4 Herzog, H. & Golomb, D. In Encyclopedia of Energy. Edited by Cleveland CJ. New York. 277-287 (2004).

5 Ruffing, A. M. Engineered cyanobacteria: teaching an old bug new tricks. *Bioeng Bugs* **2**, 136-149, doi:15285 [pii] (2011).
6 Machado, I. M. & Atsumi, S. Cyanobacterial biofuel production. *J Biotechnol* doi:10.1016/j.jbiotec.2012.03.005, doi:10.1016/j.jbiotec.2012.03.005 (2012).
7 Ducat, D. C., Way, J. C. & Silver, P. A. Engineering cyanobacteria to generate high-value products. *Trends Biotechnol* **29**, 95-103, doi:10.1016/j.tibtech.2010.12.003 (2011).
8 Scharlemann, J. P. & Laurance, W. F. Environmental science. How green are biofuels? *Science* **319**, 43-44, doi:10.1126/science.1153103 (2008).
9 Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237-240 (1998).
10 Golden, S. S., Brusslan, J. & Haselkorn, R. Genetic engineering of the cyanobacterial chromosome. *Methods Enzymol* **153**, 215-231 (1987).
11 Heidorn, T. *et al.* Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods Enzymol* **497**, 539-579, doi:10.1016/B978-0-12-385075-1.00024-X (2011).
12 Huang, H. H., Camsund, D., Lindblad, P. & Heidorn, T. Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res* **38**, 2577-2593, doi:10.1093/nar/gkq164 (2010).
13 Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS Chem Biol* **3**, 64-76, doi:10.1021/cb7002434 (2008).
14 Atsumi, S., Higashide, W. & Liao, J. C. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotechnol* **27**, 1177-1180, doi:10.1038/nbt.1586 (2009).
15 Lan, E. I. & Liao, J. C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc Natl Acad Sci U S A* **109**, 6018-6023, doi:10.1073/pnas.1200074109 (2012).
16 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, doi:S1389-1723(03)80034-8 [pii] (2003).
17 Lindberg, P., Park, S. & Melis, A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab Eng* **12**, 70-79, doi:10.1016/j.ymben.2009.10.001 (2010).
18 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, doi:10.1016/j.ymben.2012.03.005 (2012).
19 Liu, X., Sheng, J. & Curtiss, R., 3rd. Fatty acid production in genetically modified cyanobacteria. *Proc Natl Acad Sci U S A* **108**, 6899-6904, doi:10.1073/pnas.1103014108 (2011).
20 Tan, X. *et al.* Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab Eng* **13**, 169-176, doi:10.1016/j.ymben.2011.01.001 (2011).

21 Shen, C. R. *et al.* Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl Environ Microbiol* **77**, 2905-2915, doi:10.1128/AEM.03034-10 (2011).
22 Lan, E. I. & Liao, J. C. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab Eng* **13**, 353-363, doi:10.1016/j.ymben.2011.04.004 (2011).
23 Bond-Watts, B. B., Bellerose, R. J. & Chang, M. C. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat Chem Biol* **7**, 222-227, doi:10.1038/nchembio.537 (2011).
24 Ji, X. J., Huang, H. & Ouyang, P. K. Microbial 2,3-butanediol production: a state-of-the-art review. *Biotechnol Adv* **29**, 351-364, doi:10.1016/j.biotechadv.2011.01.007 (2011).
25 Celinska, E. & Grajek, W. Biotechnological production of 2,3-butanediol--current state and prospects. *Biotechnol Adv* **27**, 715-725, doi:10.1016/j.biotechadv.2009.05.002 (2009).
26 Tran, A. V. & Chambers, R. P. The dehydration of fermentative 2,3-butanediol into methyl ethyl ketone. *Biotechnol Bioeng* **29**, 343-351, doi:10.1002/bit.260290308 (1987).
27 van Haveren, J., Scott, E. L. & Sanders, J. Bulk chemicals from biomass. *Biofuels Bioprod Bioref* **2**, 41-57 (2008).
28 Syu, M. J. Biological production of 2,3-butanediol. *Appl Microbiol Biotechnol* **55**, 10-18 (2001).
29 Wijffels, R. H. & Barbosa, M. J. An outlook on microalgal biofuels. *Science* **329**, 796-799, doi:10.1126/science.1189003 (2010).
30 Greenwell, H. C., Laurens, L. M., Shields, R. J., Lovitt, R. W. & Flynn, K. J. Placing microalgae on the biofuels priority list: a review of the technological challenges. *J R Soc Interface* **7**, 703-726, doi:10.1098/rsif.2009.0322 (2010).
31 Eiteman, M. A. & Gainer, J. L. In situ extraction versus the use of an external column in fermentation. *Appl Microbiol Biotechnol* **30**, 614-618 (1989).
32 Blankenship, R. E. in *Molecular Mechanisms of Photosynthesis* Ch. 9, 172-203 (Blackwell Science Ltd, 2002).
33 Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89, doi:10.1038/nature06450 (2008).
34 Scheer, M. *et al.* BRENDA, the enzyme information system in 2011. *Nucleic Acids Res.* **39**, D670-676 (2011).
35 Atsumi, S., Li, Z. & Liao, J. C. Acetolactate synthase from Bacillus subtilis serves as a 2-ketoisovalerate decarboxylase for isobutanol biosynthesis in Escherichia coli. *Appl Environ Microbiol* **75**, 6306-6311, doi:10.1128/AEM.01160-09 (2009).
36 Godtfredsen, S. E., Lorck, H. & Sigsgaard, P. On the occurrence of a-acetolactate decarboxylases among microorganims. *Carlsberg Res, Commun.* **48**, 239-247 (1983).
37 Zhang, R. *et al.* Crystal structure of a carbonyl reductase from *Candida parapsilosis* with anti-Prelog stereospecificity. *Protein Sci* **17**, 1412-1423, doi:10.1110/ps.035089.108 (2008).

38 Rattray, F. P., Walfridsson, M.' Nilsson, D. Purification and characterization of a diacetyl reductase from *Leuconostoc paseudomesenteroides*. *International Dairy Journal* **10** (2000).
39 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, doi:10.1039/b913501d (2009).
40 Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**, 1203-1210, doi:gka167 [pii] (1997).
41 Aristidou, A. A., San, K. Y. & Bennett, G. N. Modification of central metabolic pathway in *Escherichia coli* to reduce acetate accumulation by heterologous expression of the bacillus subtilis acetolactate synthase gene. *Biotechnol Bioeng* **44**, 944-951, doi:10.1002/bit.260440810 (1994).
42 Park, H. S., Xing, R. & Whitman, W. B. Nonenzymatic acetolactate oxidation to diacetyl by flavin, nicotinamide and quinone coenzymes. *Biochim Biophys Acta* **1245**, 366-370 (1995).
43 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, doi:10.1128/AEM.07901-11 (2012).
44 Mutsuda, M., Michel, K. P., Zhang, X., Montgomery, B. L. & Golden, S. S. Biochemical properties of CikA, an unusual phytochrome-like histidine protein kinase that resets the circadian clock in *Synechococcus elongatus* PCC 7942. *J Biol Chem* **278**, 19102-19110, doi:10.1074/jbc.M213255200 (2003).
45 Ng, W. O., Zentella, R., Wang, Y., Taylor, J. S. & Pakrasi, H. B. PhrA, the major photoreactivating factor in the cyanobacterium *Synechocystis* sp. strain PCC 6803 codes for a cyclobutane-pyrimidine-dimer-specific DNA photolyase. *Arch Microbiol* **173**, 412-417 (2000).
46 Bassham, J. A., Benson, A. A. & Calvin, M. The path of carbon in photosynthesis. *J Biol Chem* **185**, 781-787 (1950).
47 Najmudin, S. *et al.* Purification, crystallization and preliminary X-ray crystallographic studies on acetolactate decarboxylase. *Acta Crystallogr D Biol Crystallogr* **59**, 1073-1075, doi:S0907444903006978 [pii] (2003).
48 Dexter, J. & Pengcheng, F. Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environ. Sci.* **2**, 857-864, doi:10.1039/b811937f (2009).
49 Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* **4**, 251-256, doi:nmeth1010 (2007).

Chapter 3 Combinatorial optimization of cyanobacterial 2,3-butanediol production

This chapter closely resembles published work. Citation: Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Combinatorial optimization of cyanobacterial 2,3-butanediol production. Metab. Eng. 22C, 76-82, doi:10.1016/j.ymben.2014.01.001 (2014).

3.1 Overview

A vital goal of renewable technology is the capture and re-energizing of exhausted CO_2 into usable carbon products. Cyanobacteria fix CO_2 more efficiently than plants, and can be engineered to produce carbon feedstocks useful for making plastics, solvents, and medicines. However, fitness of this technology in the economy is threatened by low yields in engineered strains. Robust engineering of photosynthetic microorganisms is lagging behind model microorganisms that rely on organic (high-energy) carbon, such as *Escherichia coli*, due in part to slower growth rates and increased metabolic complexity involved in phosotynthesis. In this work we show that protein expression from characterized parts is unpredictable in *Synechococcus elongatus* sp. strain PCC 7942, and may contribute to slow development. To overcome this, we apply a combinatorial approach and show that modulation of the 5'-untranslated region (UTR) can produce a range of protein expression sufficient to optimize chemical feedstock production from CO_2.

Efforts are intensifying to bridge the gap between fossil carbon consumption and renewable supply. At the same time, rising consumption of high-energy carbon is driving emissions of inorganic (exhausted) carbon into the atmosphere in the form of CO_2. Cyanobacteria, fast growing prokaryotes and simple photosynthetic organisms, can

directly sequester CO_2 by biological conversion into organic carbon using sunlight. Cyanobacteria can be grown in brine and elevated temperatures, removing competition with food crops for water and land supply. Despite high CO_2 fixation, cyanobacteria do not naturally accumulate useful feedstock chemicals.[1] Recently, the application of synthetic biology – the assembly of disparate genetic elements for novel use – has led to an increasing number of strains in which chemical feedstock production is achieved.[2-5] However, these strains have failed to reach titers representative of their carbon fixation rates. Metabolic imbalance has been shown to contribute to toxicity and poor fitness during production in heterotrophic strains[6-9], and may be particularly important in photosynthetic strains, which must maintain light harvesting complexes and the electron transport chain. However, rationally balancing protein expression requires precise predictive knowledge of the behavior of genetic elements, which even in highly studied model organisms remains a challenge[10-13]. Efforts have been made to balance pathways in *E. coli* by randomly modulating enzyme concentrations through varied translational and transcriptional control elements.[6,7,14,15] These investigations invariably depend on highly visible products such as reporter proteins or carotenoids, allowing for easy identification of a balanced phenotype from a pool of mutants. In contrast, designs for the synthesis of chemical feedstocks from CO_2 have focused on products such as ethanol, butanol, alkanes, and fatty acids, which do not result in an easily identified phenotype. Because feedstock chemical synthesis in the cell draws carbon away from metabolism, the phenotype of high production is lowered growth, which appears very similar to that for toxicity, preventing growth based assays.[16] Attempts to model and rationally optimize whole pathways *in silico* are advancing rapidly but are still limited by the complexity of

metabolic response to heterogeneous pathways.[17] As has been seen in *E. coli*, and as we show in this work for cyanobacteria, characterization of parts with reporter proteins, and replacement with production enzymes changes expression levels unpredictably.[7,18] We aimed to use a combinatorial approach with elements characterized *in situ* to produce a range of protein expression in *S. elongatus* to increase production of the 2,3-butanediol (23BD) pathway.

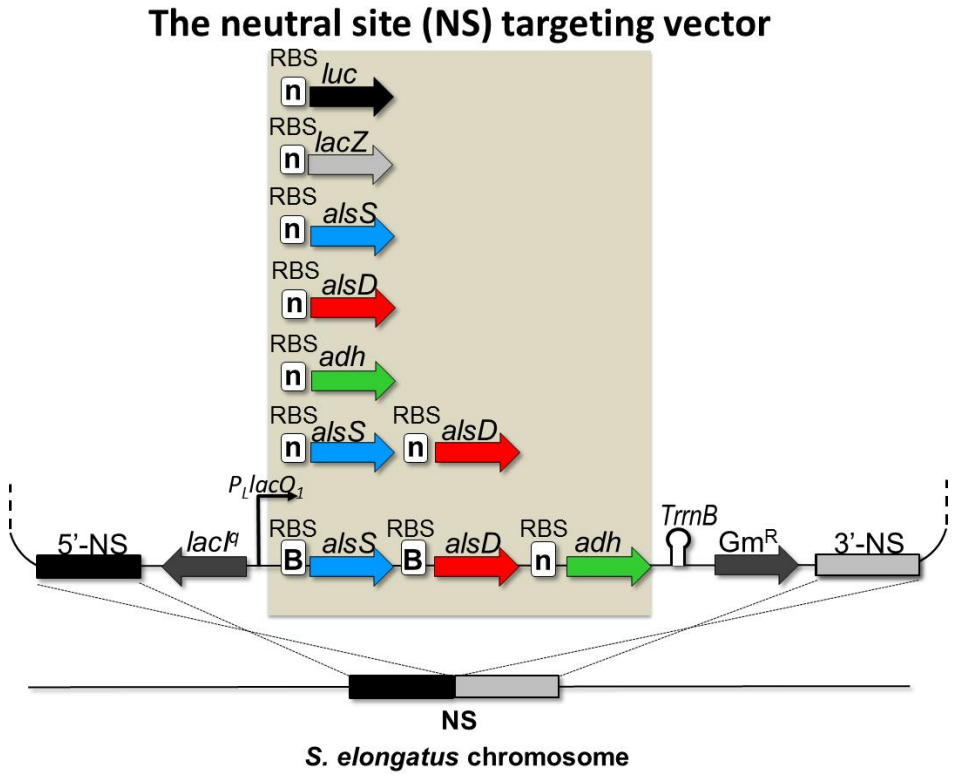


Figure 3-1 Schematic representation of recombination to integrate the expression cassettes into the *S. elongatus* chromosome. Constructs with "n" represent a list of combinatorial strains each with varied RBS sequence of A,B,C,D, or U (Table 3-1).

3.2 Results & Discussion

In previous work we installed 23BD production in *S. elongatus* by characterizing a range of genes to find proteins most suitable to the host[16] (Figure 3-2a). This allowed for production of an industrial chemical feedstock from CO_2. To investigate whether balancing activity could improve production of 23BD we aimed to conduct a combinatorial investigation of protein expression. Although there are many other factors such as the strength of the promoter and the order of the genes that may be used to balance the pathway, this study focused only on translational efficiency to limit the number of parameters. We grouped the 23BD pathway as an operon with a single promoter controlling the transcription of three genes. The use of a single operon prevents variability in transcriptional control[6], and also allows for control of total protein load through variation of the strength of a single promoter.[19] The translation of each gene in the operon from a single mRNA can be modulated by intergenic regions, containing ribosome binding sites (RBS), preceding each gene.[6,20,21] Salis *et al*. have characterized translation initiation rates (TIRs) in *E. coli*.[22] Working from that study we chose 4 RBS sequences (Table 3-1) representing a limited diversity within a range of expression of 6 fold.

Table 3-1 Ribosome binding site sequence used in this study

Name	Sequence 5' → 3'	Reference
RBS U1	gaattcattaaagaggagaaaggtacaATG	23
RBS U2	gaattcgtcgacgaggaatcaccATG	24
RBS A	attaataaggaggaattaagcATG	25
RBS B	attaataaggaggaattattagcATG	25
RBS C	attaatgaggataaattaagcATG	25
RBS D	ctttaatttaacgtaaataaggaagtcattATG	25

The translation start codon is capitalized.

The effect of the sequence on TIRs is predicted to change with the sequence of the gene being translated.[7,22] Secondary structure at the 5' end of mRNA can block translation initiation even through variation of only a few base pairs around the start codon.[26-29] Changing the RBS sequence, which is up to 30 base-pairs adjacent to the start codon in this study, may change the secondary structure significantly, especially in cases where the number of GC base pairs increased.[30] Degradation rates are higher for mRNA with lower TIRs[27]; moreover, the effect of RBS sequences on degradation of free mRNA is complicated by slower translation of free mRNA compared to those still bound to RNA polymerase, hence an increased importance of 5' structure for early ribosome recruitment.[31,32] Factors impacting transcription or protein activity after translation, such as folding, protein degradation, or formation of inclusion bodies, are assumed to have negligible variation due to changes in the UTR within each protein group. We quantified the net outcome of these effects in *S. elongatus* by placing each RBS sequence upstream of the 23BD pathway genes *alsS* (*Bacillus subtilis*), *alsD* (*Enterobacter aerogenes*), and *adh* (*Clostridium beijerinckii*) expressed individually, as well as two common reporters *lacZ* and luciferase *(luc*) genes, and assaying for enzyme activity (Figure 3-1&Figure 3-2). RBS names were denoted in order of acetolactate synthase (ALS) activity with "A" = strongest to "D" = weakest, and "U" for RBS used in the previous (non-optimized) production strain [16] (Figure 3-2b).

Figure 3-2g shows a summary of each RBSs performance modulating enzyme expression in *S. elongatus*. We observed that each enzyme in the 23BD pathway displayed activity varying by RBS with its own pattern, and that ranges differed significantly from those predicted using RFP in *E. coli* (Figure 3-2b-d, Figure 3-3). In

contrast to the *E. coli* range of 6-fold for RBS A through D, acetolactate decarboxylase (ALDC) activity was controlled over a range of >25 fold with activity clustered in the high end (Figure 3-2c). ALS activity was modulated over a range of >60 fold (depending on extremely low activity from the construct with RBS D) (Figure 3-2b). In contrast, two reporter proteins LacZ and Luc exhibited the same pattern as each other over a similar 10 fold range (Figure 3-2e&f).

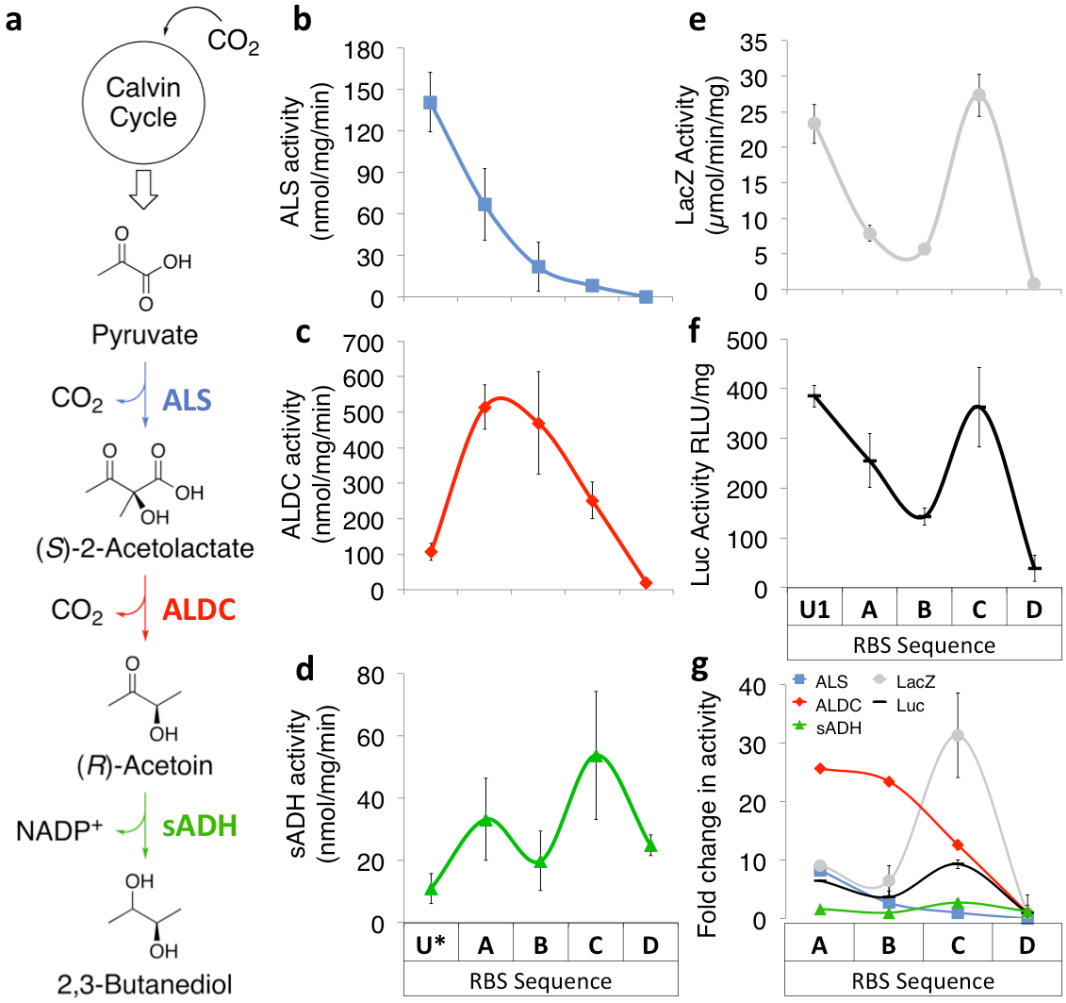


Figure 3-2 RBS modulation of enzyme activity in *S. elongatus*. a, Acetoin and 23BD pathway in *S. elongatus*. b-d, RBS modulation of 23BD pathway proteins in *S. elongatus* after 24 h induction, b, ALS c, ALDC, and d, sADH. e-f, RBS modulation of reporter proteins in *S elongatus* after 24 h induction, e, β-galactosidase (LacZ) and f, luciferase. g, Comparative range of modulation when protein is changed. Error bars indicate SD (n = 3).

The similarity in behavior for elements between reporters, and disparity of this result with genes of interest emphasizes that direct evidence for pathway modulation is still needed on a case by case basis. The pattern for alcohol dehydrogenase (ADH) showed only a 2.5 fold variation (Figure 3-2d). Attempts to predict these effects and estimate TIRs for each protein using up to date software[33] produced patterns that differed from our observation (Figure 3-3). While some patterns displayed decreased range, other RBS sequences inverted their behavior. Overall, we were able to characterize in *S. elongatus* a range of activity for each enzyme under control of different RBS, allowing us to investigate optimization of the pathway.

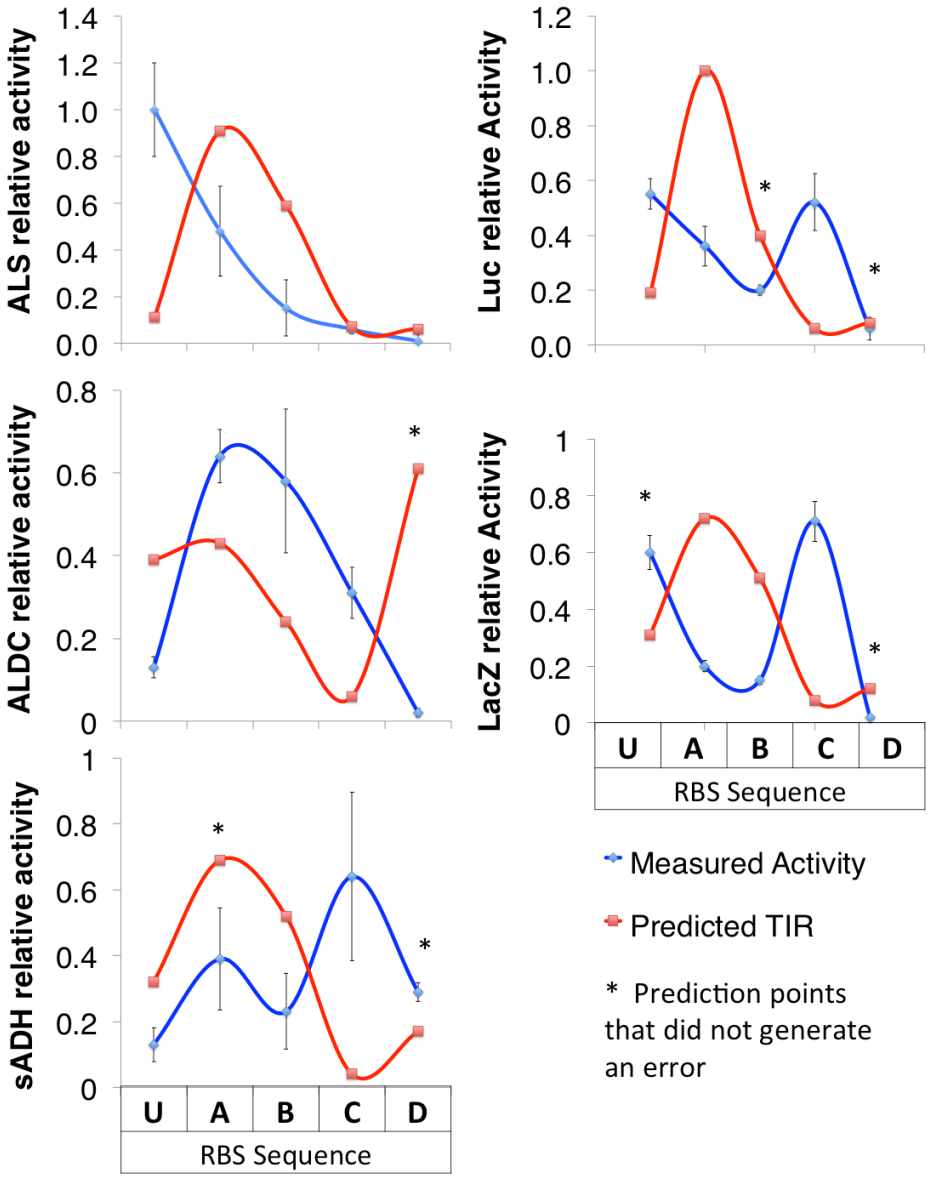
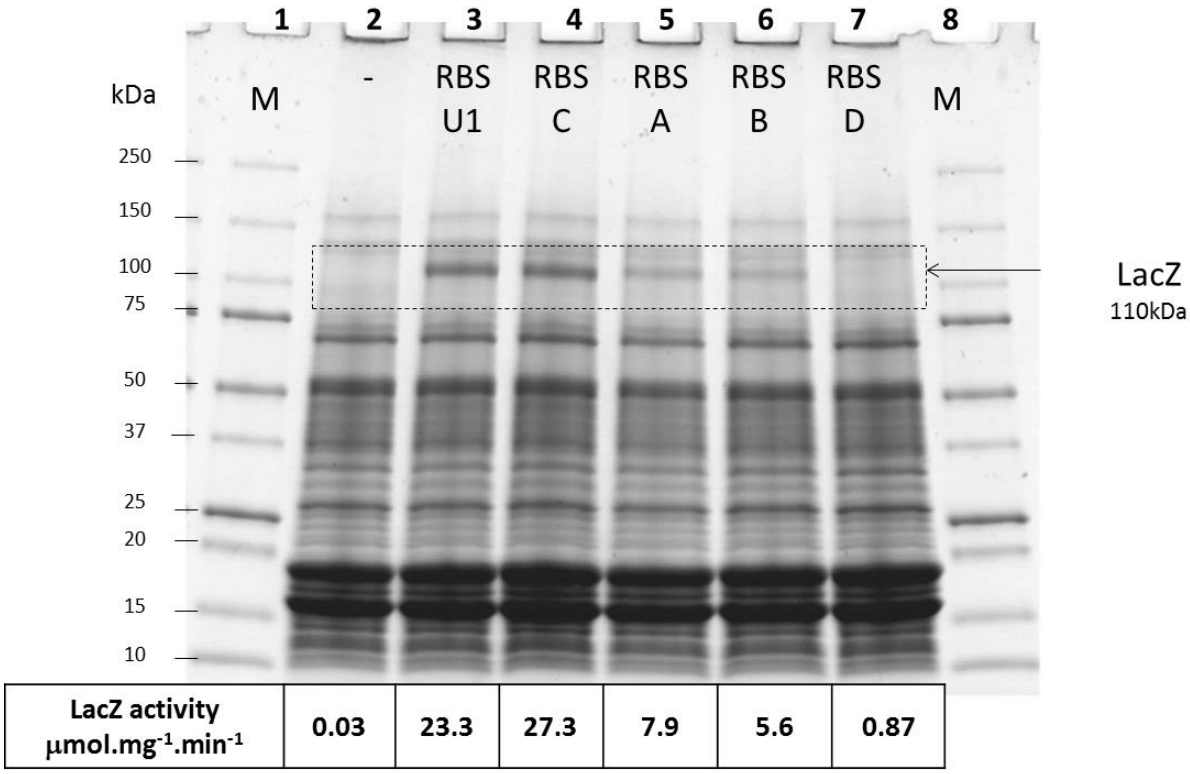


Figure 3-3 Accuracy of reverse engineered prediction in *S. elongatus*. Normalized activity measured from cell lysate (blue) versus TIR predictions (red). Asterisks represent values for TIR prediction that were able to complete the calculation with fully validated assumptions. Error bars indicate SD (n = 3).

To confirm protein expression in *S. elongatus* under RBS A-D and U, soluble fractions of cell extracts from engineered strains containing the *lacZ* gene were analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 3-4). The amount of LacZ was estimated to be about 0.4 (RBS A), 0.32 (RBS B), and 1.08% (RBS

C) of total protein respectively, which correlates well with activity (R^2 = 0.9791). No band was detected for RBS D.



LacZ activity μmol.mg^{-1}.min^{-1}	0.03	23.3	27.3	7.9	5.6	0.87

Figure 3-4 SDS/PAGE analysis of the lacZ expression. The soluble fraction of cell extracts from engineered strains were analyzed with SDS/PAGE. Size ladder (lane 1&8), AL723 (without lacZ) (lane 2), lacZ expressing strain AL1103 (RBS U1) (lane 3), AL1106 (RBS C) (lane 4), AL1104 (RBS A) (lane 5), AL11005 (RBS B) (lane 6) and AL1107 (RBS D) (lane 7). The calculated molecular weight of LacZ is 110 kDa.

3.2.1 Combinatorial optimization of acetoin production

We first explored optimization of the two gene acetoin pathway in *S. elongatus* (Figure 3-2a). Sixteen strains were constructed and tested (Figure 3-5) representing each possible combination of RBS A through D preceding each gene. The paired activities of ALS and ALDC in cell lysate were measured from 7 strains, 4 of them with fixed RBS A upstream of the *alsS* gene (Figure 3-5 purple) and 4 strains with fixed RBS B modulating the *alsD*

gene (Figure 3-5 orange). As the RBS varies for the first gene in the pathway, protein activity is modulated similarly to that when expressed alone. In the paired assay, activity of the second enzyme is dependent on production of 2-acetolactate from the first reaction by ALS (Figure 3-2a). This allows us to estimate the accumulation of intermediates in the pathway. What we observe is that when the RBS is varied for the second gene it results in only slightly modulated activity, consistently converting less than half of the available 2-acetolactate pool (Figure 3-5 purple). Lack of modulation for the second gene could result in part from translational coupling with the first gene, or from uniform low expression from the second place in the operon.[18] From this result we would expect dominant modulation of production as the first RBS is varied, however in production we observe a rescuing effect of a strong RBS in front of the second gene. This may reflect differences in pathway activity *in vivo*, where substrate concentrations are low, versus *in vitro*, where substrate is in excess leading to a larger range of activity. Additional factors such as protein degradation during cell lysis or changes in protein localization could also affect the range of activity observed in assays as compared to production.[32]

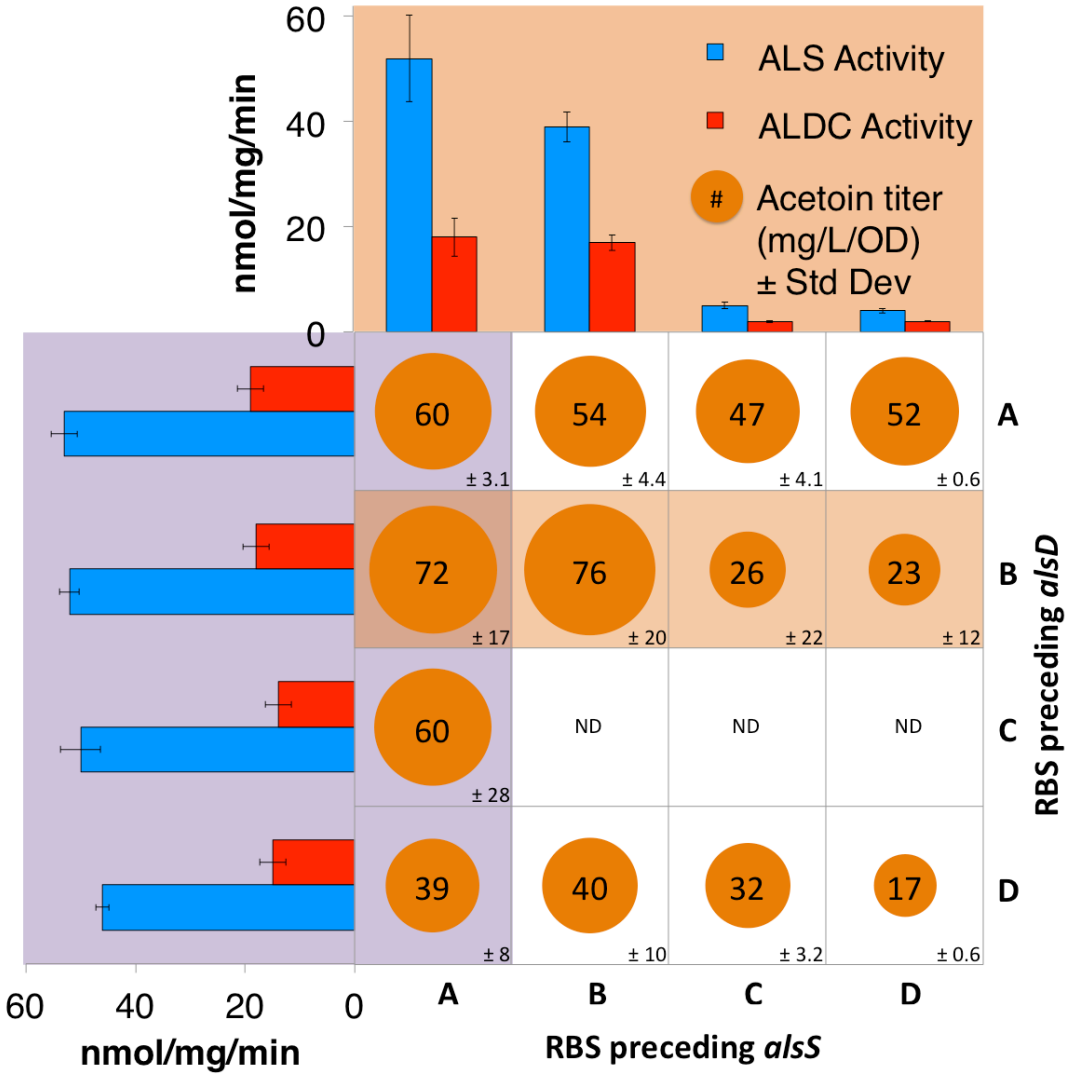


Figure 3-5 Combinatorial optimization of acetoin production. All combinations for RBS-modulated expression of the acetoin pathway in *S. elongatus* using RBS characterized in this study. The column and row in purple and orange correspond to strains whose ALS activity (blue bar) and ALDC (red bar) were measured during acetoin production. Error bars indicate SD (n = 3). ND is not determined, active transformants could not be isolated.

We are unable to obtain efficient acetoin production from three constructs containing *alsD* preceded by RBS C. We have repeatedly isolated these three strains, but each lost acetoin productivity during growth, due to toxicity or an unknown regulatory element. This trend is not significantly affected by variation of the first RBS for combinations B through C, but is completely abolished by RBS A. Additionally strains

with RBS C preceding the first gene (*alsS*) do not suffer the same effect, and single gene *alsD* expression from RBS C is viable (Figure 3-2c). This suggests that mRNA secondary structure is contributing to toxicity within the cell rather than protein imbalance, although additional experiments are needed to elucidate the mechanism. Sharp deviations from overall trends due to unique pairings of regulatory elements have been observed previously even in well characterized systems.[13] Variation of ALS and ALDC activities indicates that the combinatorial assembly of RBS sequences can modulate expression levels of genes within the acetoin operon resulting in a range of production (Figure 3-5). In our system, strong RBS sequences preceding each gene did not result in the strongest production from the pathway with a slight advantage in maximum titer being inferred by pairing of the second strongest RBSs for each gene (B-*alsS*, B-*alsD*) (Figure 3-5). To study the effects of the RBS combinations further, we measured rates of change for acetoin production and growth of each strain in 96-well plates. Under these conditions, the strain with B-*alsS*, B-*alsD* continued to increase acetoin production and growth after 24 h, suggesting that the acetoin pathway is more robust within metabolism compared to constructs containing RBS A or U, which showed decelerating production rates after 24 h (Figure 3-6).

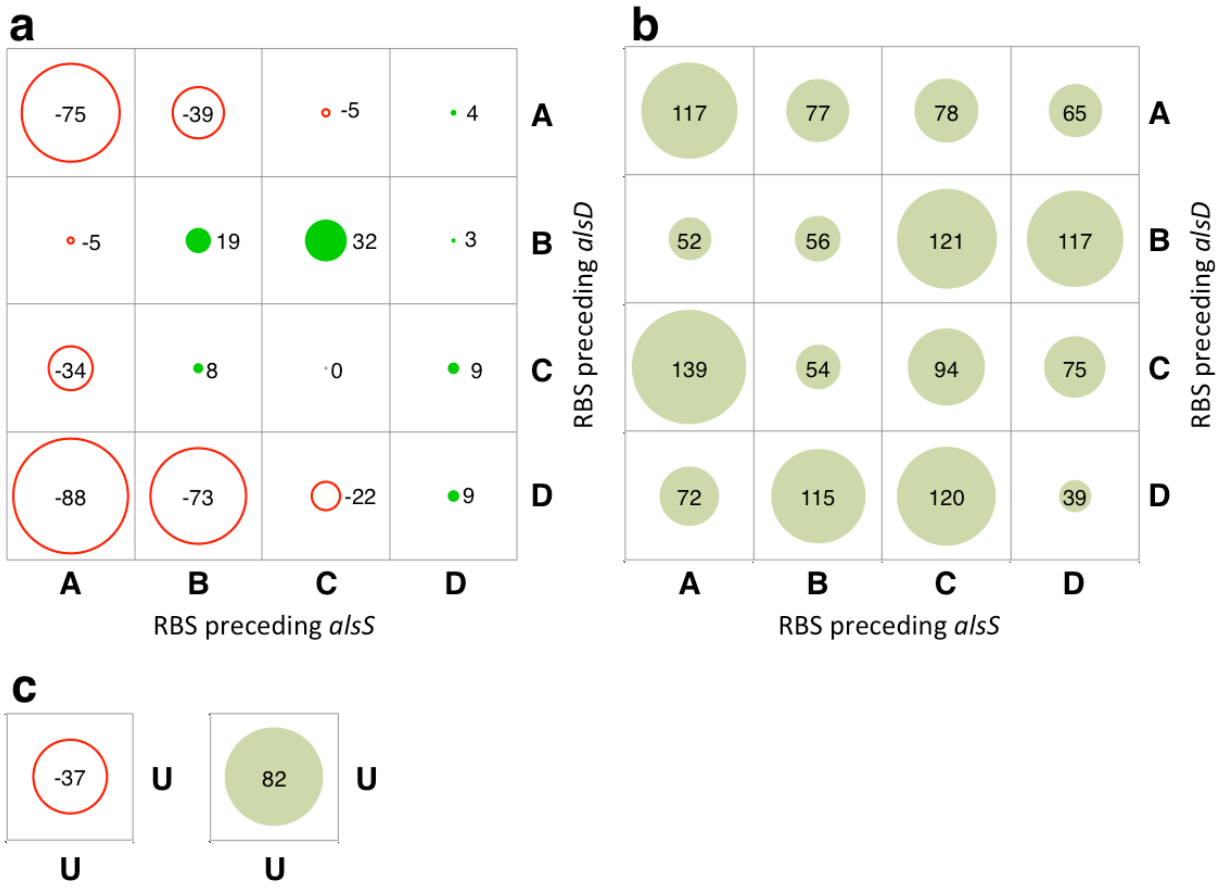


Figure 3-6 Rates of change in production and growth in RBS modulated *S. elongatus* strains. a, Acceleration of acetoin production at 24 h after induction (μg l^{-1} h^{-2}). **b**, Acceleration of growth at 24 h after induction (cells μl^{-1} h^{-2}). **c**, Acceleration of acetoin production and growth for the non-optimized strain. U represents the RBS preceding each gene in a non-optimized strain constructed without *adh*.

3.2.2 A combinatorial approach to optimize the 23BD pathway in S. elongatus

To optimize the final step we used the balanced acetoin construct (B-*alsS*, B-*alsD*) and added the *adh* gene preceded by each of the four RBS A-D (Figure 3-1, Figure 3-2f & Figure 3-7). We expected only a small range of activity change based on individual characterization (Figure 3-2f) and location of the gene at the end of the operon. However, we observed significant modulation by RBS not only of sADH, but also of the other two enzymes expressed from the pathway with ALS activity attenuating in the longer operon.

Activity of ALDC was measured with excess 2-acetolactate, to obtain results irrespective of ALS, and these activities increased and decreased with sADH activity. Strikingly, the construct with RBS D, the second weakest sequence for *adh* as characterized in single gene experiments (Figure 3-2d), resulted in the strongest activity for sADH (Figure 3-7). Conversely the construct with the second strongest sequence RBS A (Figure 3-2d) resulted in the lowest activities (Figure 3-7). Production of 23BD closely followed the trend for ALDC and sADH activity and did not appear to be limited by ALS. The optimum construct (B-*alsS*, B-*alsD*, D-*adh*) produced 496 mg l^{-1} after 72h, achieving a 180% increase over the non-optimized strain (Figure 3-7 & Figure 3-8).

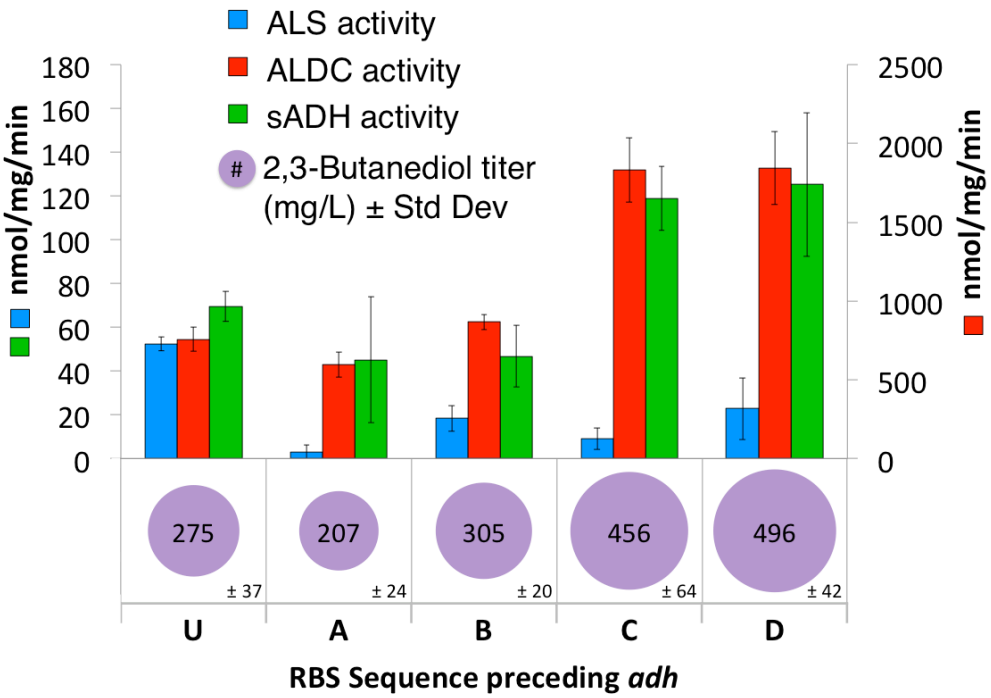


Figure 3-7 23BD production from RBS modulated strains. Bars represent ALS (blue), ALDC (red), and ADH (green) activities, and purple circles represent 23BD production from strains expressing RBS B *alsS*, RBS B *alsD* and every combination of four RBSs sequences upstream of the *adh* gene. U represents the RBS used for each gene in the non-optimized strain. Error bars indicate SD (n = 3).

The weakest construct (B-*alsS*, B-*alsD*, A-*adh*) and the non-optimized construct produced 207 mg l^{-1} and 275 mg l^{-1} respectively. Overall, characterization of the final strains did not match the pattern predicted by RBS characterization in single gene, or two gene operon designs (Figure 3-2d & Figure 3-5). Despite this, varying the RBS sequences preceding each gene generated a range of protein expression from combinatorial strains that was sufficient to produce an optimized 23BD pathway.

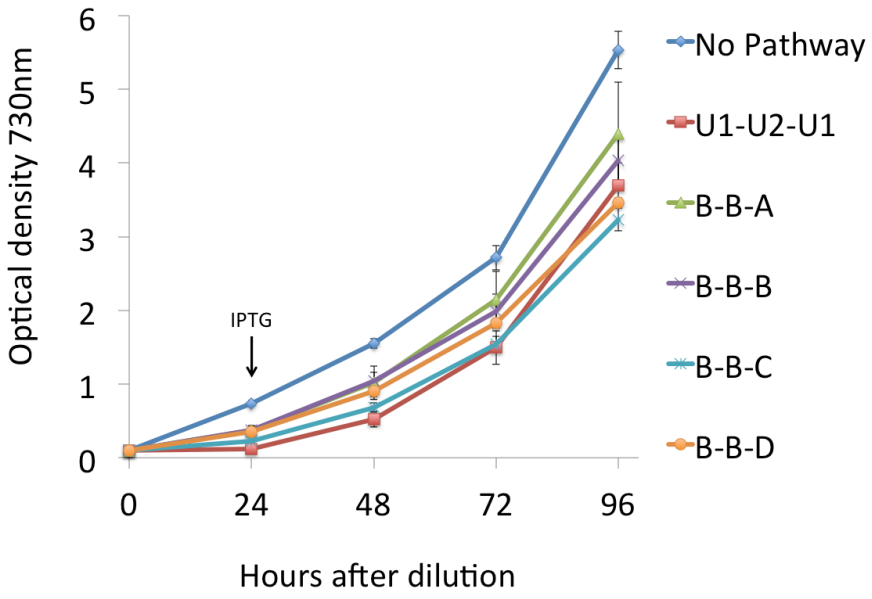


Figure 3-8 Growth curve during 2,3-butanediol production. Letters U1, U2, A, B, C, and D represent RBS sequences preceding *alss*, *alsD*, and *adh* respectively (Figure 3-1 & Table 3-1).

3.3 Conclusion

The combinatorial optimization outlined here shows that even *in situ* characterization of genetic elements produces unpredictable results in *S. elongatus* when combined into an operon for feedstock chemical production. This throws into question the viability of

reporter based characterization of standardized parts.[34] Identifying the causative factors and underlying mathematics for variation of gene expression is an ongoing challenge even in well defined systems such as *E. coli*.[10,22] In photosynthetic strains with desirable qualities, such as tolerance to high temperature and high salt, very little information may be available about variation in their transcriptional and translational systems. Increasing titers of feedstock chemicals from CO_2 requires methods that can optimize pathways in hosts without complete elucidation of their metabolism. The relatively low number of engineered strains used in this study allowed for characterization over a feasible scale for chromatographic measurement, providing a counterpoint to the pressure for highly visible and tractable chemical targets imposed by randomization and high-throughput screening. In the absence of rational design, a guided combinatorial approach using 4 RBS sequences proved sufficient to modulate enzyme activity across a three gene feedstock chemical pathway, and enabled increased 23BD titers.

Specific Methods

Plasmid construction

Plasmids used and constructed are listed in Table 3-2. Oligonucleotides for constructions and verifications are listed in Table 3-3. All plasmids were constructed by DNA assembly techniques using sequence and ligation independent cloning (SLIC)[36,37] in *E. coli* XL1-Blue (Agilent Technologies, Santa Clara, CA).

To construct plasmids with four different ribosome binding site combinations upstream of the *alsS* gene from *B. subtilis* and the *alsD* gene from *E. aerogenes*, we used the plasmid pAL302 (RBS U1 *alsS*-RBS U2 *alsD*)[16] as template. Each RBS was replaced with RBS A-D (Table 3-1).

The *adh* gene from *C. beijerinckii*, which was chemically synthesized by DNA2.0 Inc. (Menlo Park, CA) previously[16], was flanked by the selected RBS sequences (RBS A-D) by PCR. The resulting PCR fragments were inserted to downstream of *alsD* gene on the neutral site (NS) targeting plasmids by SLIC (Table 3-3).

To clone *alsS*, *alsD*, *adh*, *luc,* and *lacZ* genes individually under different RBS sequences (RBS U (U2 for *alsD* and U1 for others), A-D), each gene was flanked by the selected RBS by PCR. The PCR fragments were inserted into pAL312[16] by SLIC.

Table 3-2 Strains and plasmids used in this study

Strains	Genotype	Source
***E. coli* strain**		
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10* (*Tetr*)]	Agilent Technologies
***S. elongatus* strains**		
PCC7942	wild type	S. Golden
AL723	$P_L lacO_1$ and Gent^R integrated at NS	[38]
AL734	Same as AL723 but $P_L lacO_1$ ::RBS D *alsD*	This study
AL735	Same as AL723 but $P_L lacO_1$ ::RBS B *alsS*- RBS A *alsD*	This study
AL736	Same as AL735 but RBS B *alsD*	This study
AL738	Same as AL735 but RBS D *alsD*	This study
AL739	Same as AL723 but $P_L lacO_1$ ::RBS C *alsS*- RBS A *alsD*	This study
AL740	Same as AL739 but RBS C *alsS*-RBS B *alsD*	This study
AL742	Same as AL739 but RBS D *alsD*	This study
AL743	Same as AL723 but $P_L lacO_1$ ::RBS D *alsS*- RBS A *alsD*	This study
AL745	Same as AL743 but RBS C *alsD*	This study
AL746	Same as AL743, but RBS D *alsD*	This study
AL762	Same as AL723 but $P_L lacO_1$ ::RBS U1 *alsS*	[38]
AL763	Same as AL723 but $P_L lacO_1$ ::RBS U1 *alsS*-RBS U2 *alsD*	[38]
AL773	Same as AL723 but $P_L lacO_1$ ::RBS U1 *luc*	This study
AL774	Same as AL773 but RBS A *luc*	This study
AL775	Same as AL773 but RBS B *luc*	This study
AL776	Same as AL773 but RBS C *luc*	This study
AL777	Same as AL773 but RBS D *luc*	This study
AL778	Same as AL723 but $P_L lacO_1$ ::RBS A *alsS*- RBS A *alsD*	This study
AL779	Same as AL778 but RBS B *alsD*	This study
AL780	Same as AL778 but RBS C *alsD*	This study
AL784	Same as AL735 but RBS C *alsD*	This study
AL788	Same as AL739 but RBS C *alsD*	This study
AL791	Same as AL743 but RBS B *alsD*	This study
AL1103	Same as AL723 but $P_L lacO_1$ ::RBS U1 *lacZ*	This study
AL1104	Same as AL1103 but RBS A *lacZ*	This study
AL1105	Same as AL1103 but RBS B *lacZ*	This study
AL1106	Same as AL1103 but RBS C *lacZ*	This study
AL1107	Same as AL1103 but RBS D *lacZ*	This study
AL1350	Same as AL723 but $P_L lacO_1$ ::RBS B *alsS*- RBS B *alsD*-RBS A *adh*	This study
AL1351	Same as AL1350 but RBS B *adh*	This study
AL1352	Same as AL1350 but RBS C *adh*	This study

AL1353	Same as AL1350, but RBS D *adh*	This study
AL1421	Same as AL723 but $P_L lacO_1$::RBS A *alsS*	This study
AL1422	Same as AL1421 but RBS B *alsS*	This study
AL1423	Same as AL1421 but RBS C *alsS*	This study
AL1424	Same as AL1421 but RBS D *alsS*	This study
AL1425	Same as AL723 but $P_L lacO_1$::RBS A *alsD*	This study
AL1426	Same as AL1425 but RBS B	This study
AL1427	Same as AL1425 but RBS C	This study
AL1428	Same as AL1425 but RBS D	This study
AL1429	Same as AL723 but $P_L lacO_1$::RBS A *adh*	This study
AL1430	Same as AL1429 but RBS B *adh*	This study
AL1431	Same as AL1429 but RBS C *adh*	This study
AL1432	Same as AL1429 but RBS D *adh*	This study
AL1449	Same as AL1425 but RBS U2 *alsD*	This study
AL1450	Same as AL1429 but RBS U1 *adh*	This study

Plasmids

pAL243	NS targeting vector; ColE1 ori; $P_L lacO_1$::RBS A *alsS*-RBS A *alsD*; $Gent^R$	This study
pAL244	Same as pAL243 but RBS B *alsD*	This study
pAL245	Same as pAL243 but RBS C *alsD*	This study
pAL246	Same as pAL243 but RBS D *alsD*	This study
pAL247	Same as pAL243 but $P_L lacO_1$::RBS B *alsS*-RBS A *alsD*	This study
pAL248	Same as pAL247 but RBS B *alsD*	This study
pAL249	Same as pAL247 but RBS C *alsD*	This study
pAL250	Same as pAL247 but RBS D *alsD*	This study
pAL251	Same as pAL243 but $P_L lacO_1$::RBS C *alsS*-RBS A *alsD*	This study
pAL252	Same as pAL251 but RBS B *alsD*	This study
pAL253	Same as pAL251 but RBS C *alsD*	This study
pAL254	Same as pAL251 but RBS D *alsD*	This study
pAL255	Same as pAL243 but $P_L lacO_1$::RBS D *alsS*-RBS A *alsD*	This study
pAL256	Same as pAL255 but RBS B *alsD*	This study
pAL257	Same as pAL255 but RBS C *alsD*	This study
pAL258	Same as pAL255 but RBS D *alsD*	This study
pAL279	Same as pAL243 but $P_L lacO_1$::RBS U1 *luc*	This study
pAL280	Same as pAL279 but RBS A	This study
pAL281	Same as pAL279 but RBS B	This study
pAL282	Same as pAL279 but RBS C	This study
pAL283	Same as pAL279 but RBS D	This study
pAL301	Same as pAL243 but $P_L lacO_1$::RBS U1 *alsS*	[38]
pAL302	Same as pAL243 but $P_L lacO_1$::RBS U1 *alsS*-RBS U2 *alsD*	[38]
pAL312	Same as pAL243 but P_{LlacO1}; $Gent^R$	[38]
pAL319	Same as pAL243 but $P_L lacO_1$::RBS A *lacZ*	This study
pAL320	Same as pAL319 but RBS B	This study
pAL325	Same as pAL319 but RBS C	This study
pAL326	Same as pAL319 but RBS D	This study
pAL327	Same as pAL319 but RBS U1 *lacZ*	This study
pAL510	Same as pAL243, but $P_L lacO_1$::RBS B *alsS*-RBS B *alsD*-RBS A *adh*	This study
pAL511	Same as pAL510 but RBS B *adh*	This study
pAL512	Same as pAL510 but RBS C *adh*	This study
pAL513	Same as pAL510 but RBS D *adh*	This study
pAL561	Same as pAL301but RBS A *alsS*	This study
pAL562	Same as pAL301but RBS B *alsS*	This study
pAL563	Same as pAL301 but RBS C *alsS*	This study
pAL564	Same as pAL301 but RBS D *alsS*	This study
pAL565	Same as pAL243 but $P_L lacO_1$::RBS A *alsD*	This study
pAL566	Same as pAL565 but RBS B *alsD*	This study
pAL567	Same as pAL565 but RBS C *alsD*	This study
pAL568	Same as pAL565 but RBS D *alsD*	This study
pAL569	Same as pAL243 but $P_L lacO_1$::RBS A *adh*	This study
pAL570	Same as pAL569 but RBS B *adh*	This study
pAL571	Same as pAL569 but RBS C *adh*	This study
pAL572	Same as pAL569 but RBS *adh*	This study

pAL577	Same as pAL565 but RBS U2 *alsD*	This study
pAL578	Same as pAL569 but RBS U1 *adh*	This study

Table 3-3 Oligonucleotides used in this study

Name	Sequence 5' → 3'	Plasmid constructed
IM11	AGATCCTCTAGAGTCGACCTG	pAL279, 280, 281, 282, 283, 327, 319, 320, 325, 326, 452, 453, 454, 455, 561, 562, 563,564,465, 466, 467, 468, 469, 470, 471, 472, 473,474
IM46	CAGGTCGACTCTAGAGGATCTCTACAGGATTACGACGGCTT	pAL452, 453, 454, 455
IM76	ATTAATAAGGAGGAATTAAGCATGTTGACAAAAGCAACAAA	pAL243, 244, 245, 246, 561
IM77	GCTTAATTCCTCCTTATTAATCTAGAGAGCTTTCGTTTTCA	pAL243, 247, 251, 255
IM78	ATTAATAAGGAGGAATTAAGCATGAATCATGCTTCAGATTG	pAL243, 247, 251, 255, 565
IM79	ATTAATAAGGAGGAATTATTAGCATGAATCATGCTTCAGATTG	pAL244, 248, 252, 256, 566
IM80	ATTAATGAGGATAAATTAAGCATGAATCATGCTTCAGATTG	pAL245, 249, 253, 257, 567
IM81	CTTTAATTTAACGTAAATAAGGAAGTCATTATGAATCATGCTTCAGATTG	pAL246, 250, 254, 258, 568
IM82	ATTAATAAGGAGGAATTATTAGCATGTTGACAAAAGCAACAAA	pAL247, 248, 562, 249, 250
IM83	ATTAATGAGGATAAATTAAGCATGTTGACAAAAGCAACAAA	pAL251, 252, 253, 254, 563
IM84	CTTTAATTTAACGTAAATAAGGAAGTCATTATGTTGACAAAAGCAACAAA	pAL255, 256, 257, 258, 564
IM85	GCTAATAATTCCTCCTTATTAATCTAGAGAGCTTTCGTTTTCA	pAL244, 248, 252, 256
IM86	GCTTAATTTATCCTCATTAATCTAGAGAGCTTTCGTTTTCA	pAL245, 249, 253, 257
IM87	AATGACTTCCTTATTTACGTTAAATTAAAGCTAGAGAGCTTTCGTTTTCA	pAL246, 250, 254, 258
IM94	ATTAATAAGGAGGAATTAAGCATGGAAGACGCCAAAAACA	pAL280
IM96	ATTAATAAGGAGGAATTATTAGCATGGAAGACGCCAAAAACA	pAL281
IM98	ATTAATGAGGATAAATTAAGCATGGAAGACGCCAAAAACAT	pAL282
IM100	CTTTAATTTAACGTAAATAAGGAAGTCATTATGGAAGACGCCAAAAACAT	pAL283
IM106	GCTTAATTCCTCCTTATTAATGGTCAGTGCGTCCTGCTGA	pAL243, 244, 245, 246, 280, 319, 466, 471, 565, 569
IM107	GCTAATAATTCCTCCTTATTAATGGTCAGTGCGTCCTGCTGA	pAL247, 248, 249, 250, 28, 320, 467, pAL472, 562, 566, 570,
IM108	GCTTAATTTATCCTCATTAATGGTCAGTGCGTCCTGCTGAT	pAL251, 252, 253, 254, 282, 325, 468, 563, 567, 571, 473
IM109	AATGACTTCCTTATTTACGTTAAATTAAAGGGTCAGTGCGTCCTGCTGAT	pAL255, 256, 257, 258, 283, 326, 469, 474
IM116	TTTAATGAATTCGGTCAGTG	pAL279, 327, 465, 470
IM123	AGGTCGACTCTAGAGGATCTTTACAATTTGGACTTTCCG	pAL279, 280, 281, 282, 283
IM124	CACTGACCGAATTCATTAAAGAGGAGAAAGGTACAATGGAAGACGCCAAAAACAT	pAL279
IM126	CACTGACCGAATTCATTAAAGAGGAGAAAGGTACAATGACCATGATTACGGATT	pAL327
IM127	CAGGTCGACTCTAGAGGATCTTTATTTTTGACACCAGACCA	pAL327, 319, 320, 325, 326
IM128	ATTAATAAGGAGGAATTAAGCATGACCATGATTACGGATTC	pAL319
IM129	AATAAGGAGGAATTATTAGCATGACCATGATTACGGATTC	pAL320
IM130	ATTAATGAGGATAAATTAAGCATGACCATGATTACGGATTC	pAL325
IM131	ACGTAAATAAGGAAGTCATTATGACCATGATTACGGATTC	pAL326
IM163	GCTTAATTCCTCCTTATTAATCTAACTTTCTACTGAACGGA	pAL452
IM164	ATTAATAAGGAGGAATTAAGCATGAAAGGTTTTGCCATGCT	pAL452, 569
IM165	GCTAATAATTCCTCCTTATTAATCTAACTTTCTACTGAACGGA	pAL453,570
IM166	ATTAATAAGGAGGAATTATTAGCATGAAAGGTTTTGCCATGCT	pAL453, 571
IM167	GCTTAATTTATCCTCATTAATCTAACTTTCTACTGAACGGA	pAL454, 572
IM168	ATTAATGAGGATAAATTAAGCATGAAAGGTTTTGCCATGCT	pAL454
IM169	AATGACTTCCTTATTTACGTTAAATTAAAGCTAACTTTCTACTGAACGGA	pAL455
IM170	CTTTAATTTAACGTAAATAAGGAAGTCATTATGAAAGGTTTTGCCATGCT	pAL455
IM174	CACTGACCGAATTCATTAAAGAGGAGAAAGGTACAATGGGTCATCACCACCATCA	pAL465
IM175	CAGGTCGACTCTAGAGGATCTTTATTTGTACAGTTCGTCCATG	pAL465, 466, 467, 468, 469
IM176	ATTAATAAGGAGGAATTAAGCATGGGTCATCACCACCATCATC	pAL467
IM177	ATTAATAAGGAGGAATTATTAGCATGGGTCATCACCACCATCATC	pAL468
IM178	ATTAATGAGGATAAATTAAGCATGGGTCATCACCACCATCATC	pAL469
IM180	CACTGACCGAATTCATTAAAGAGGAGAAAGGTACAATGTCTAAAGGTGAAGAATT	pAL470
IM181	ATTAATAAGGAGGAATTAAGCATGTCTAAAGGTGAAGAATT	pAL471
IM182	ATTAATAAGGAGGAATTATTAGCATGTCTAAAGGTGAAGAATT	pAL472
IM183	ATTAATGAGGATAAATTAAGCATGTCTAAAGGTGAAGAATT	pAL473
IM184	CTTTAATTTAACGTAAATAAGGAAGTCATTATGTCTAAAGGTGAAGAATT	pAL474
IM185	GCAGGTCGACTCTAGAGGATCTTTATTTGTACAATTCATCCA	pAL470, 471, 472, 473, 474
IM188	TTCGTCGACGAGGAATCACCATGAATCATGCTTCAGATTG	pAL577
IM190	ATTCAAGAGGAGAAAGGTACATGAAAGGTTTTGCCATGCT	pAL578
IM191	GTACCTTTCTCCTCTTGAAT	pAL578
IM192	GGTGATTCCTCGTCGACGAATTCGGTCAGTGCGTCCTGCT	pAL577

3.4 Acknowledgements

This work was supported by the Asahi Kasei Corporation.

3.5 References

1 Machado, I. M. & Atsumi, S. Cyanobacterial biofuel production. *J. Biotechnol.* **162**, 50-56, doi:10.1016/j.jbiotec.2012.03.005 (2012).

2 Ducat, D. C., Way, J. C. & Silver, P. A. Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* **29**, 95-103, doi:10.1016/j.tibtech.2010.12.003 (2011).

3 Rabinovitch-Deere, C. A., Oliver, J. W., Rodriguez, G. M. & Atsumi, S. Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* **113**, 4611-4632, doi:10.1021/cr300361t (2013).

4 Heidorn, T. *et al.* Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods Enzymol.* **497**, 539-579, doi:10.1016/B978-0-12-385075-1.00024-X (2011).

5 Bandyopadhyay, A., Stockel, J., Min, H., Sherman, L. A. & Pakrasi, H. B. High rates of photobiological H2 production by a cyanobacterium under aerobic conditions. *Nat. Commun.* **1**, 139, doi:10.1038/ncomms1139 (2010).

6 Pfleger, B. F., Pitera, D. J., Smolke, C. D. & Keasling, J. D. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol.* **24**, 1027-1032, doi:10.1038/nbt1226 (2006).

7 Zelcbuch, L. *et al.* Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* **41**, e98, doi:10.1093/nar/gkt151 (2013).

8 Barbirato, F., Grivet, J. P., Soucaille, P. & Bories, A. 3-Hydroxypropionaldehyde, an inhibitory metabolite of glycerol fermentation to 1,3-propanediol by enterobacterial species. *Appl. Environ. Microbiol.* **62**, 1448-1451 (1996).

9 Zhu, M. M., Skraly, F. A. & Cameron, D. C. Accumulation of methylglyoxal in anaerobically grown *Escherichia coli* and its detoxification by expression of the *Pseudomonas putida* glyoxalase I gene. *Metab. Eng.* **3**, 218-225, doi:10.1006/mben.2001.0186 (2001).

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

11 Mutalik, V. K. *et al.* Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods* **10**, 347-353, doi:10.1038/nmeth.2403 (2013).

12 Ferreira, J. P., Overton, K. W. & Wang, C. L. Tuning gene expression with synthetic upstream open reading frames. *Proc. Natl. Acad. Sci. USA* **110**, 11284-11289, doi:10.1073/pnas.1305590110 (2013).

13 Kosuri, S. *et al.* Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **110**, 14024-14029, doi:10.1073/pnas.1301301110 (2013).
14 Blazeck, J., Garg, R., Reed, B. & Alper, H. S. Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol. Bioeng.* **109**, 2884-2895, doi:10.1002/bit.24552 (2012).
15 Blazeck, J. *et al.* Generalizing a hybrid synthetic promoter approach in Yarrowia lipolytica. *Appl. Microbiol. Biotechnol.* **97**, 3037-3052, doi:10.1007/s00253-012-4421-5 (2013).
16 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).
17 Lewis, N. E., Nagarajan, H. & Palsson, B. O. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol.* **10**, 291-305, doi:10.1038/nrmicro2737 (2012).
18 Levin-Karp, A. *et al.* Quantifying Translational Coupling in E. coli Synthetic Operons Using RBS Modulation and Fluorescent Reporters. *ACS Synth. Biol.*, doi:10.1021/sb400002n (2013).
19 Glick, B. R. Metabolic load and heterologous gene expression. *Biotechnol. Adv.* **13**, 247-261, doi:http://dx.doi.org/10.1016/0734-9750(95)00004-A (1995).
20 Barrick, D. *et al.* Quantitative analysis of ribosome binding sites in E.coli. *Nucleic Acids Res.* **22**, 1287-1295 (1994).
21 Chen, H., Bjerknes, M., Kumar, R. & Jay, E. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs. *Nucleic Acids Res.* **22**, 4953-4957 (1994).
22 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, doi:10.1038/nbt.1568 (2009).
23 Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**, 1203-1210, doi:gka167 [pii] (1997).
24 Atsumi, S. *et al.* Metabolic engineering of Escherichia coli for 1-butanol production. *Metab Eng* **10**, 305-311, doi:10.1016/j.ymben.2007.08.003 (2008).
25 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, doi:10.1038/nbt.1568 (2009).
26 Hall, M. N., Gabay, J., Debarbouille, M. & Schwartz, M. A role for mRNA secondary structure in the control of translation initiation. *Nature* **295**, 616-618 (1982).
27 de Smit, M. H. & van Duin, J. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. USA* **87**, 7668-7672 (1990).
28 Qing, G., Xia, B. & Inouye, M. Enhancement of Translation Initiation by A/T-Rich Sequences Downstream of the Initiation Codon in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **6**, 133-144, doi:10.1159/000077244 (2003).

29 Lamping, E., Niimi, M. & Cannon, R. D. Small, synthetic, GC-rich mRNA stem-loop modules 5 ' proximal to the AUG start-codon predictably tune gene expression in yeast. *Microb. Cell. Fact.* **12**, doi: 10.1186/1475-2859-12-74 (2013).
30 Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in Escherichia coli. *Science* **324**, 255-258, doi:10.1126/science.1170160 (2009).
31 Lim, H. N., Lee, Y. & Hussein, R. Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. USA* **108**, 10626-10631, doi:Doi 10.1073/Pnas.1105692108 (2011).
32 Montero Llopis, P. *et al.* Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77-81, doi:10.1038/nature09152 (2010).
33 Salis, H. M. The ribosome binding site calculator. *Methods Enzymol.* **498**, 19-42, doi:10.1016/B978-0-12-385120-8.00002-4 (2011).
34 Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **26**, 787-793, doi:10.1038/nbt1413 (2008).
35 Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. & Stanier, R. Y. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *J. Gen. Microbiol.* **111**, 1-61 (1979).
36 Machado, H. B., Dekishima, Y., Luo, H., Lan, E. I. & Liao, J. C. A selection platform for carbon chain elongation using the CoA-dependent pathway to produce linear higher alcohols. *Metab. Eng.* **14**, 504-511, doi:10.1016/j.ymben.2012.07.002 (2012).
37 Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **4**, 251-256, doi:10.1038/nmeth1010 (2007).
38 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).
39 Golden, S. S., Brusslan, J. & Haselkorn, R. Genetic engineering of the cyanobacterial chromosome. *Methods Enzymol.* **153**, 215-231, doi:0076-6879(87)53055-5 [pii] (1987).
40 Voges, O., Proskauer, B. Beitraege zur Ernaehrungsphysiologie und zur Differential Diagnose der Bakterien der hemmorrhagischen Septicamie. *Z. Hyg.* **28** (1898).
41 Yang, Y. T., Peredelchuk, M., Bennett, G. N. & San, K. Y. Effect of variation of Klebsiella pneumoniae acetolactate synthase expression on metabolic flux redistribution in Escherichia coli. *Biotechnol Bioeng* **69**, 150-159, doi:10.1002/(SICI)1097-0290(20000720)69:2<150::AID-BIT4>3.0.CO;2-N [pii] (2000).
42 Bastian, S. & Arnold, F. H. Reversal of NAD(P)H cofactor dependence by protein engineering. *Method. in Mol. Biol.* **834**, 17-31, doi:10.1007/978-1-61779-483-4_2 (2012).
43 Miller, J. H. (Cold Spring Harbor Laboratory Press, 1992).

Chapter 4 Introduction of a carbon sink increases total carbon fixation in *Synechococcus elongatus* PCC7942

4.1 Overview

Metabolic engineering in photosynthetic organisms is a nascent technology, with great promise for renewable energy, specifically for capture of CO_2 to replace dwindling fossil carbon supply. It has been proposed that photosynthetic capacity can be increased by a relaxation of bottlenecks inherent to growth, however, the limits of carbon partitioning within the cell and their effect on carbon fixation are not known. Here I show that expressing genes for the synthesis of pyruvate increases partitioning and leads to a 1.8-fold increase in total carbon yield in the cyanobacterium *Synechococcus elongatus* PCC7942. As partitioning increases, it leads to a crash in the carbon fixation cycle. The data suggest a local maximum for carbon partitioning from the dark reactions that is scalable with light intensity.

Worldwide, carbon dioxide emissions are rising while natural carbon sinks are reaching capacity.[1] In 2012 alone, 5.4 trillion metric tons of CO_2 was released into the atmosphere from the oxidation of fixed carbon.[2] Currently 99% percent of chemicals synthesized in the US originate from non-renewable sources.[3] Technologies for converting CO_2 back to organic carbon are still in developmental phases, and large gains will need to be made to balance carbon production with carbon consumption. To meet this need, interest has been growing in the engineering of photosynthetic organisms for carbon fixation and production of chemicals to replace petroleum derived compounds.[4,5] Cyanobacterial cultures offer a direct process for capturing light and concentrated CO_2

into biomass, and can be installed in locations that do not compete against food for water and land resources. This includes locations that can remediate waste gases from electricity generation and industry, which account for 53% of carbon dioxide emissions in the US.[6] A growing number of engineered pathways converting CO_2 into useful products have been demonstrated in cyanobacterial hosts, and tools for design and genetic manipulation are reaching maturation.[7-9] However, carbon production rates from CO_2 capture remain low. Increasing productivity from engineered cyanobacteria is critical to capturing emissions on a large scale.

Photosynthetic oxygen evolution can be increased up to 3 times its natural rate experimentally by decoupling electron transfer from NADPH turnover.[10] This led to predictions that increasing the rate of NADPH consumption in vivo by the introduction of an exogenous pathway could lead to increases in photosynthetic rates.[11,12] NADPH is primarily consumed through carbon fixation in the dark reactions of the Calvin-Benson (CB) cycle. One hypothesis, is that carbon consuming reactions associated with growth are slower than carbon fixation and that the inability to consume carbon prevents the fixation of more carbon. This bottleneck could be relieved by introduction of a 'carbon sink', defined as a high flux pathway that converts carbon into a final molecule that is secreted from the cell and cannot be utilized by central metabolism.

Early metabolic engineering in cyanobacteria has supported the idea that increased carbon partitioning to secreted products will approach carbon fixation rates associated with biomass production by concurrently reducing growth.[11] However, differences in the paradigm of carbon flux between photosynthetic and fermentative cells make this problematic. While a fermentative cell can be abstracted as a black box that

converts glucose to products at a rate corresponding to protein concentration and cofactor balance, the photosynthetic cell depends also on the concentration of metabolites in the CB cycle, and a large number of proteins and metabolites involved in the light reactions. Each carbon in pyruvate produced from glucose in *E. coli* has passaged through 8 reactions since entering the cell, while in photosynthesis each carbon has undergone at least 27 enzymatic transformations before reaching pyruvate, more than three times as many. Overexpression of pathway enzymes in fermentative cells may asymptotically approach maximum carbon partitioning (until a very high level of expression causes cell burden). However, in a photosynthetic cell, because metabolites must be maintained in cytosolic concentrations required for the CB cycle, too much pathway protein may deplete the metabolite pool and prevent the fixation of more carbon. In this way there will be a local maximum of carbon partitioning corresponding to maximum fixation in the CB cycle.

In this work I aimed to create a carbon sink through expression of a pathway from fixation to pyruvate. In all engineered pathways to date, except for that of 1,2-propanediol expressing *mgsA* and limonene expressing *dxs*, the exogenous pathway has been separated from fixation by at least two steps that remain under native regulation. This allows the cell to respond to perturbations that may affect the pool of CB Metabolites (CBM). In the case of 1,2-propanediol and limonene, expressing genes that target CBM did not significantly increase carbon partitioning, indicating that removal of carbon from the CB cycle was low. I chose the previously published 2,3-butanediol (23BD) pathway in *Synechococcus elongatus* PCC7942 (hereafter *S. elongatus*), for it's

high driving force, and apparent limitation by pyruvate.[12] 23BD is a 4 carbon feedstock chemical, that is valuable for production of plastics, fuels, and antifreeze.

4.1.1 *S. elongatus* displays facultative heterotrophy for pyruvate

S. elongatus has historically been considered an obligate photoautotroph, unable to utilize alternative carbon sources in the absence of genetic modification.[13] It was discovered during the course of this work that under standard illumination and atmospheric carbon, pyruvate can be taken up by cells at rates comparable to growth from excess CO_2 (Figure 4-1a). It was a central aim of this work to increase intracellular levels of pyruvate, thus the addition of extracellular pyruvate was a convenient test for the behavior of metabolism under increased pyruvate. When the 23BD base strain was grown with excess pyruvate, partitioning between 23BD and biomass increased from 45% to 72% (Figure 4-1b). This implies that the 23BD pathway is capable of strong partitioning, and is not limited by competition at pyruvate, but elsewhere in metabolism.

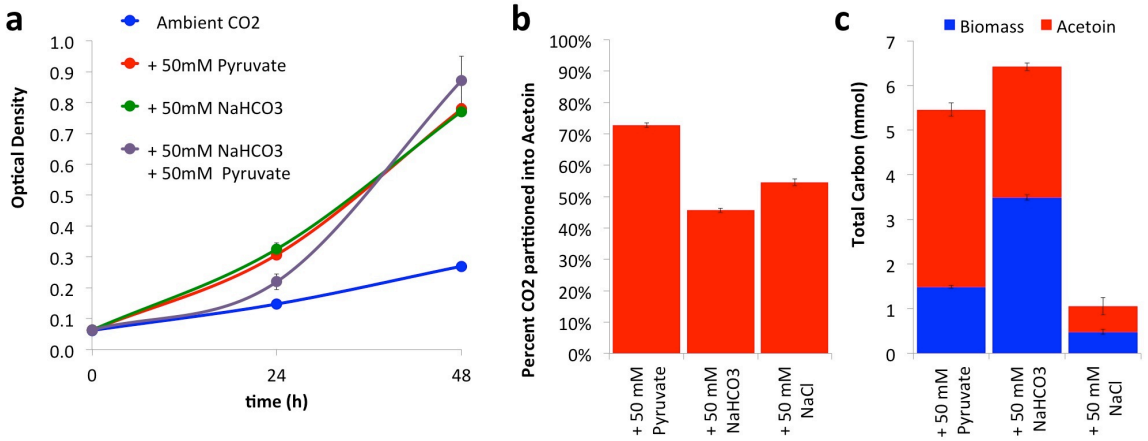


Figure 4-1 Growth and acetoin production during pyruvate feeding. (a) Growth of *Synechococcus elongatus* PCC7942 cells with and without pyruvate, green indicates normal growth on bicarbonate, red indicates growth at the same concentration of pyruvate, blue indicates growth at the same concentration of

sodium salt (ambient CO_2), purple indicates growth on both pyruvate and bicarbonate. (b) Partitioning between acetoin and biomass (c) total carbon fixed into biomass and acetoin.

4.1.2 Construction of single gene pyruvate pathways

In *S. elongatus,* pyruvate is naturally synthesized from PEP which is generated from fixation in two steps (Figure 4-2). Phosphoglycerate-mutase (PGM) converts PGA into 2-phosphoglycerate, which is then isomerized by enolase (ENO) to phospho*enol*pyruvate (PEP). These first two steps are reversible and require no cofactors. PEP is then converted into pyruvate either directly by pyruvate kinase (PYK), which generates ATP, or indirectly through phosphoenolpyruvate kinase (PEPC), which fixes bicarbonate to generate oxaloacetate, forming the start of the partial TCA cycle. In one branch of the TCA cycle, oxaloacetate is reduced by NADH to form malate, which can be oxidized and decarboxylated to produce pyruvate and NADPH. This conversion of PEP to pyruvate has net-zero carbon fixation; however, it may be important in balancing NADH and NADPH ratios, and is thought to be the predominant pathway under some conditions in *Synechocystis sp.* PCC6803.[14]

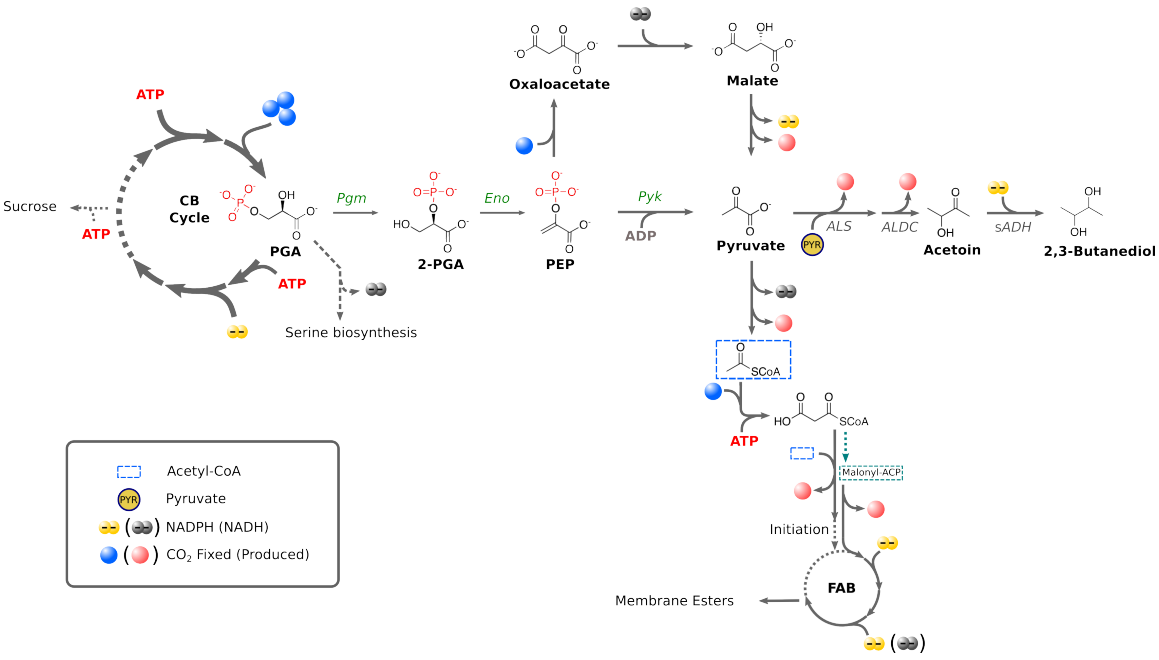


Figure 4-2 Pathways between CO_2 fixation, pyruvate, and 23BD. Molecule names are bolded, enzyme abbreviations are italicized. PGA, 3-phosphoglycerate; 2-PGA, 2-phosphoglycerate; PEP, phospho*enol*pyruvate; *Pgm*, phosphogycerate mutase; *Eno*, enolase; *Pyk*, pyruvate kinase; *Eda*, oxaloacetate decarboxylase.

I sought to increase carbon flux to pyruvate by targeting two major branch-points in pyruvate synthesis. First, the product of fixation, PGA, is immediately consumed in high flux by phosphoglycerate kinase (PGK) with the conversion of ATP to ADP to continue the fixation cycle. PGA is also taken up by D-3-phosphoglycerate dehydrogenase to begin the serine synthesis pathway. To offset carbon loss to serine synthesis, we targeted PGM expression to increase partitioning towards PEP. This also allows for modulation of competition between the fixation cycle and chemical production at the first step. The second major branch-point occurs at PEP, which is consumed in the TCA cycle, and in numerous biosynthesis pathways including those for chorismate and dehydroquinoate in *S. elongatus*. We targeted PYK expression to increase direct conversion of PEP to pyruvate. Two genes coding for each enzyme were chosen from *E.*

coli to minimize native regulation effects (*gpmA* and *gpmM* encoding PGM and *pykA* and *pykF* encoding PYK). One of each was also chosen from *S. elongatus* (*pgm3* encoding for PGM and *pyk* encoding for PYK), to give coverage of three genes at each branch-point. There is a wide variation in PGM activities, which may allow for a balancing effect in native metabolism.[15] The native gene *pgm3* is thought to have stronger directionality for consumption of PGA.

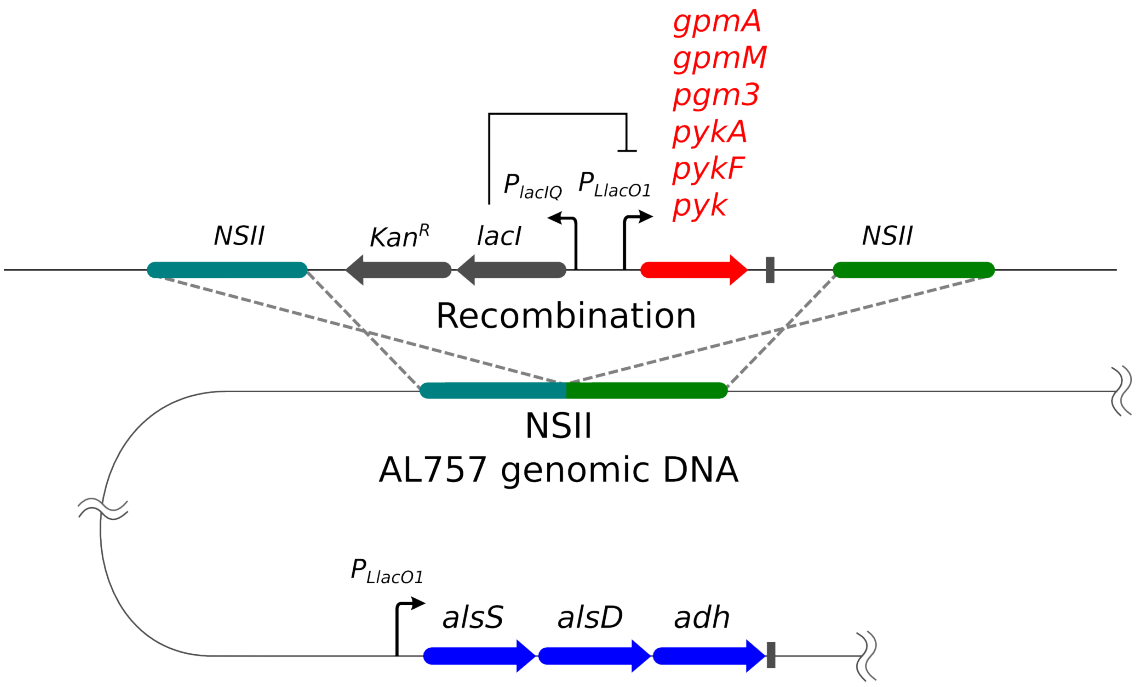


Figure 4-3 Schematic of recombination to integrate pyruvate pathway genes into the chromosome of engineered *S. elongatus*. Stacked names indicate individual strains.

4.1.3 Phenotypic response in single gene strains

To screen the effect of each gene on pathway flux, production of 2,3-butanediol was measured from cyanobacterial cultures grown at 30˚C and 50 umol photons/m^2/s, in BG-11 modified with 50mM bicarbonate (see methods). Carbon partitioning is traditionally

measured in fermentative strains as total carbon in product versus a theoretical maximum calculated from the concentration of carbon source available (most commonly glucose added to the media). In photosynthetic production, the concentration of the carbon source depends on CO_2 fixation, which is dependent on the number of cells undergoing photosynthesis in addition to their individual fixation rates, making metabolically available carbon difficult to accurately quantify (see Appendix A). Partitioning is therefore given as a ratio between two measureable values: mmol CO_2 fixed to produce 23BD and mmol CO_2 fixed to produce observed biomass during a given time period, and data is separated into days for analysis by growth phase. Strains containing genes targeting the first branch-point at PGA produced similarly to the host strain, generating 175 mg/L in 72 hours with a partitioning ratio of 30%(Figure 4-4 a&d). Partitioning increased as cells entered a linear phase above OD ~1.5, and total carbon dropped slightly from 16 mmol/L/day to 14 mmol/L/day as expected for normal growth.[16]

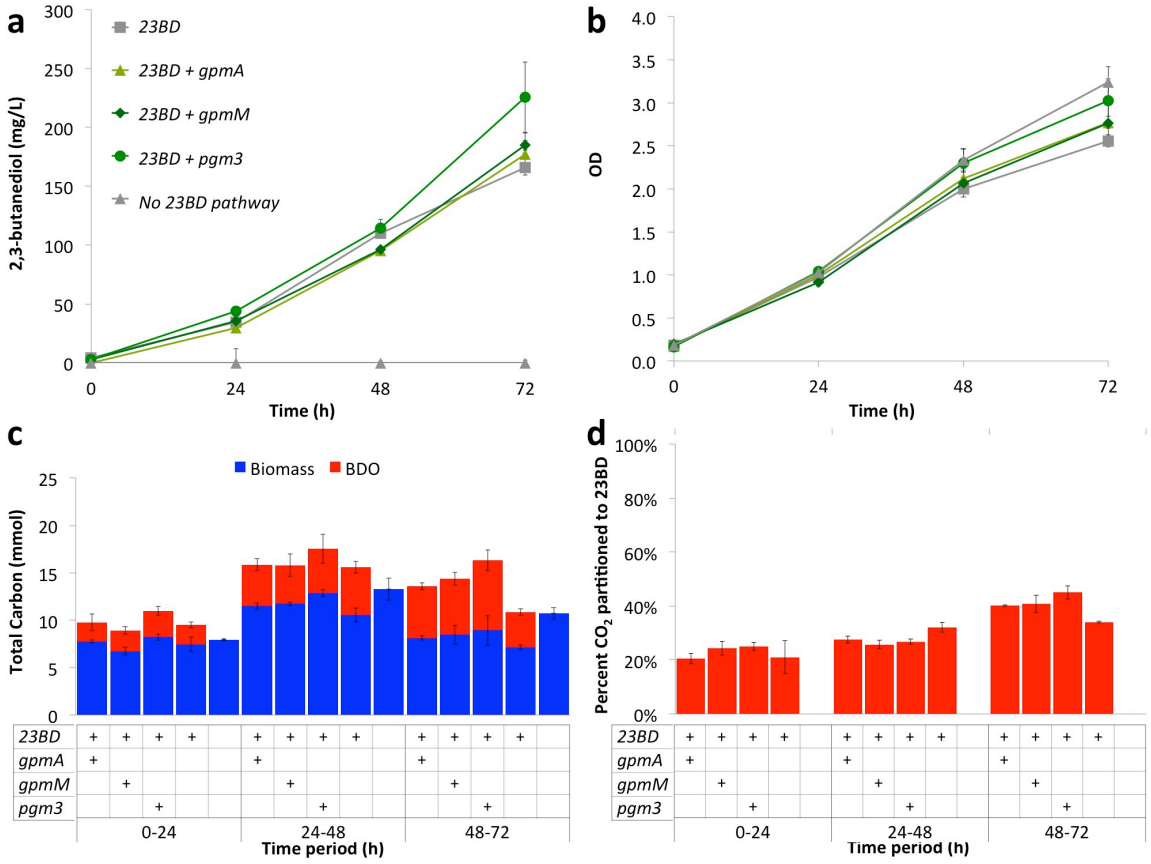


Figure 4-4 Phenotypic response from addition of genes targeting 3-PGA. (a) Accumulative 23BD titer; (b) Time course of growth; (c) Total carbon measured in 23BD and biomass per 24 hour time period; (d) Carbon measured in 23BD as a percentage of total carbon.

Strains containing PYK genes targeting the PEP branch-point produced a comparatively strong response (Figure 4-5). Titers were similar after 72 hours with strains producing ~135 mg/L, however strains expressing *pyk* displayed a marked growth defect and therefore limited capacity for fixation, indicating that to achieve the same titer production per cell was increased (Figure 4-5b). Due to slow growth in modified strains, production was continued to 96 hours. Partitioning for the strain expressing *pyk* increased

to 60% during exponential growth (Figure 4-5d). During the linear phase *pykF* also appears to increase partitioning, reaching 55% with 14 mmol carbon /day. Partitioning also increased in the unmodified strain, however, in this case total carbon is sharply reduced to 6 mmol/day indicating the onset of a stationary phase. Partitioning was not increased measurably in strains expressing *pykA*, although the exponential phase was extended.

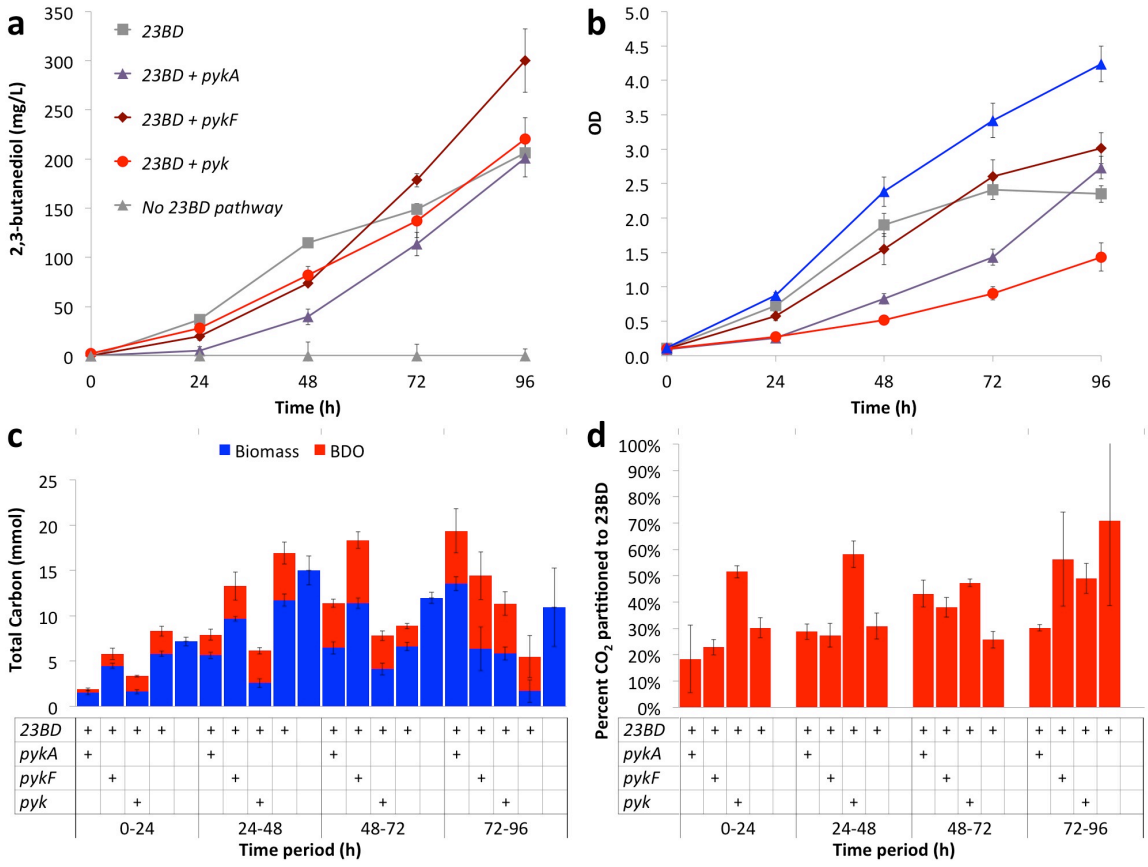


Figure 4-5 Phenotypic response from addition of genes targeting Phosphoenolpyruvate. (a) Accumulative 23BD titre; (b) Time course of growth; (c) Total carbon measured in 23BD and biomass per 24 hour time period; (d) Carbon measured in 23BD as a percentage of total carbon.

4.1.4 Construction of a complete pyruvate pathway

The absence of an effect in partitioning by PGM encoding (Figure 4-4) genes led us to investigate whether flux to pyruvate could be increased by expression of steps between PGM and pyruvate. The enzymatic step controlled by PGM is highly reversible, and can be outcompeted by steps with higher directionality, such as that of phosphoglycerate kinase (PGK) in the CB cycle, which consumes ATP. Expression of an enolase with PGM and PYK genes can allow for decarboxylation of pyruvate to drive carbon flow from fixation (Figure 4-6). Three gene pathways for two combinations were chosen, one with three *E. coli* genes, *gpmA*, *eno*, and *pykA* (strain abbreviated as *Eco*) and one with three endogenous genes, *pgm3*, *eno*, and *pyk* (strain abbreviated as *End*). To allow screening of colonies and assessment of low levels of production, strains for acetoin production, omitting the *adh* gene in the 2,3-butanediol pathway, were constructed in tandem with 23BD strains. Detection of acetoin can be performed by high throughput colorimetric assay, and is much more sensitive than 23BD which is measured by GC (see Chapter 5 - Methods).

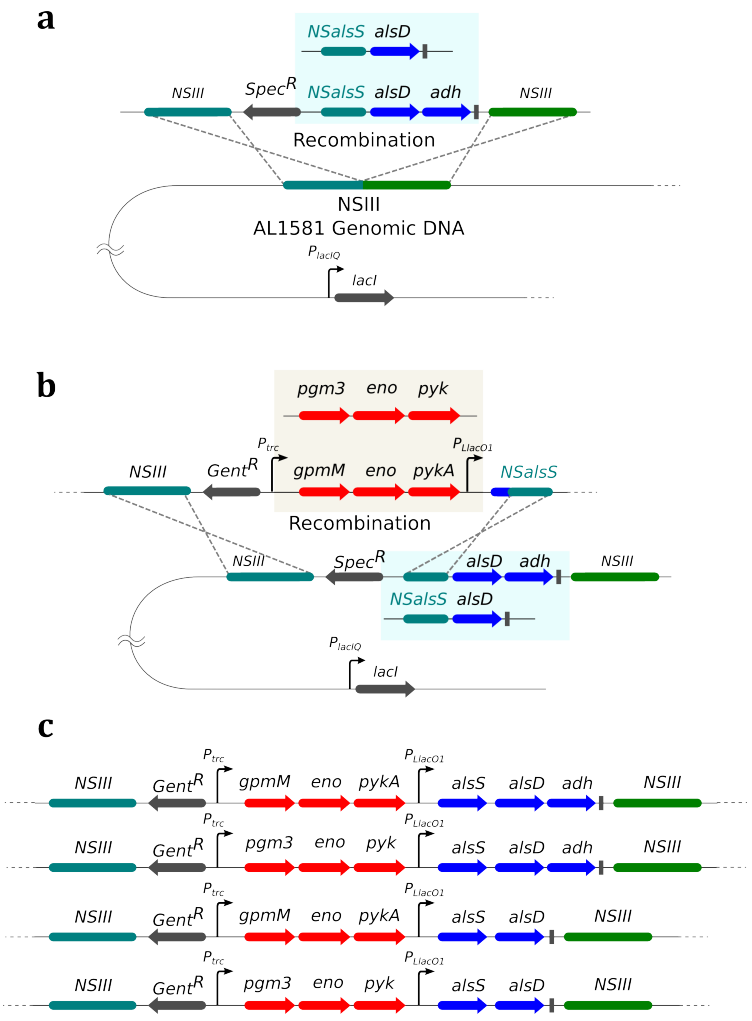


Figure 4-6 Recombination to activate a dormant pathway. (a) schematic representation of recombination to integrate half of the acetoin and 23BD pathways, creating an inactivated (dormant) base strain. (b) schematic representation of recombination to integrate the full pyruvate pathway with the first half of the 23BD or acetoin pathway, activating the dormant pathway genes. (c) final constructs after recombination showing the completed acetoin and 23BD pathways with pyruvate pathway operons.

4.1.5 Carbon capture with a complete pyruvate pathway

Strains containing complete pyruvate pathways grew very slowly in the absence of induction, requiring 2-3 days of growth to reach OD 0.4 after dilution from cells maintained on plates. This is likely due to an inability of the promotional system to completely repress expression of the pyruvate pathway. Despite very low OD, cells

produced measureable amounts of 23BD after induction, reaching final titers of 49 mg/L and 17 mg/L for *End* and *Eco* strains respectively after 96 hours. Strikingly, partitioning to 23BD increased to close to 100% after the first day of induction, due to continued 23BD production in the absence of growth. Total carbon remained low ~1 mmol/L/day due to low cell density. Control strains produced 15-17 mmol/L/day fixed carbon, and partitioned 30%-40% into 23BD as expected.

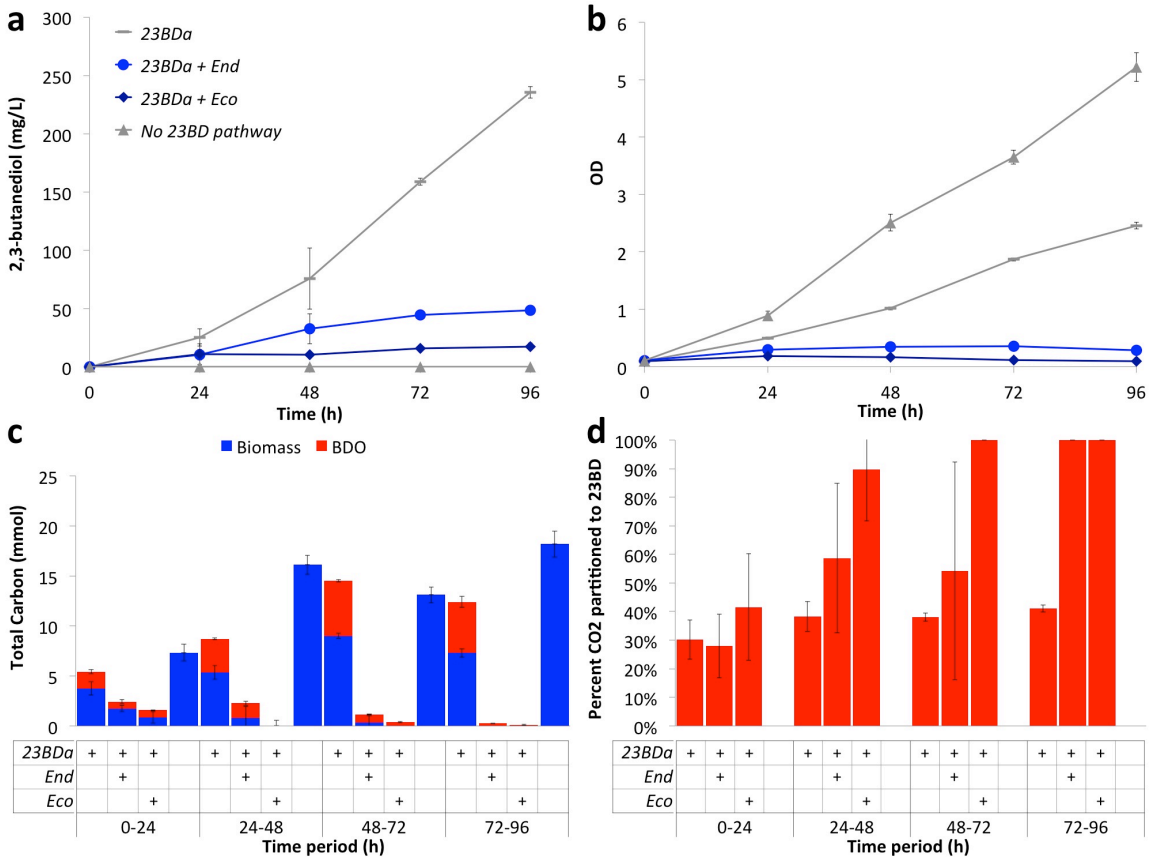


Figure 4-7 Phenotypic response from the 3-gene carbon funnel pathway. (a) Accumulative 23BD titre; (b) Time course of growth; (c) Total carbon measured in 23BD and biomass per 24 hour time period; (d) Carbon measured in 23BD as a percentage of total carbon.

4.1.6 High Light production

The exogenous pyruvate pathway draws carbon from the CB cycle in competition with PGK and Glyceraldehyde-3-phosphate dehydrogenase (GAP), which are ATP and NADPH dependent respectively. If light is limiting, ATP and NADPH turnover could be limiting, slowing PGK and GAPD flux and increasing partitioning towards the pyruvate pathway. Increasing light would be expected to increase ATP and NADPH turnover and rescue PGK and GAP activity. To assess this effect strains were tested in 250 umol photons m^{-2} s^{-1} light. To amplify the response of *End* and *Eco* strains, starting OD was increased from 0.1 to 0.5. All strains increased carbon fixation significantly under these conditions (Figure 4-8). The 23BD control strain and 23BD-*pyk* strains achieved titers of 535 mg/L and 726 mg/L respectively after 60 hours (Figure 4-8a). This was matched by fast growth rates resulting in similar partition rates as for low light, 20-40% for 23BD control strains, and 30-60% for 23BD-*pyk* strain (Figure 4-8d). In comparison the partition rates for *End* and *Eco* strains were significantly lowered compared to low light, fluctuating between 30-70%. This is concurrent with an increase in growth and total carbon to ~10 mmol/day (graphed time periods showing 5 mmol are for 12 hours). This is consistent with a rescuing of carbon fixation. However, total carbon and growth are still significantly lower than control strains grown in the same high-light conditions.

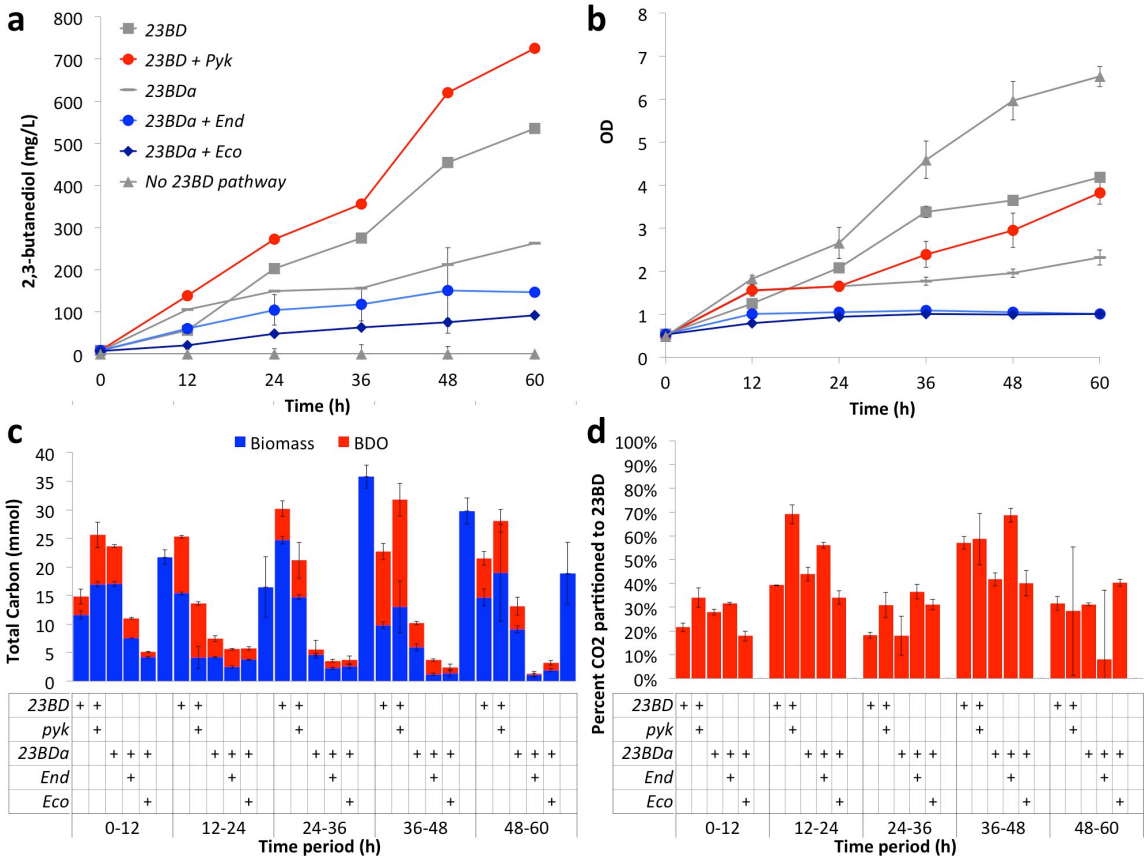


Figure 4-8 Phenotypic response from pathways in high light. (a) Accumulative 23BD titre; (b) Time course of growth; (c) Total carbon measured in 23BD and biomass per 24 hour time period; (d) Carbon measured in 23BD as a percentage of total carbon.

4.1.7 Photosynthetic rates

To assess whether slowed growth is due to carbon partitioning or a loss of fixation, an estimate of productive capacity was made. Significant error can be associated with calculations of capacity due to differences in growth phase between two measured points (see Appendix II). Data was normalized in two separate cases: For overall production per cell (Figure 4-9a&b) accumulative titer and area were used; For total carbon during comparative rates (Figure 4-9c&d) only data during growth phases with an optical density less than 1 were used, and data was not divided by optical density for the time point (see Figure 4-10a). Strains expressing the *pyk* gene in normal light achieve a 2.4x

increase in specific BDO production over the unmodified strain. Strains expressing *pykA* and *pykF* genes produced 1.4 and 1.3 fold increases respectively. PGM genes did not significantly increase productivity. This corresponds to increases in total carbon of 1.8, 1.5, 1.2, and 1.3 fold for *pyk*, *pkA*, *pykF*, and PGM expressing strains respectively. *End* and *Eco* strains produced 1.1 and 0.8 fold 23BD/L/h/OD, compared to the unmodified strain, however, total carbon per cell was diminished to 84% and 64% for *End* and *Eco* strains respectively. This indicates higher carbon partitioning may be affecting the overall rate of carbon fixation.

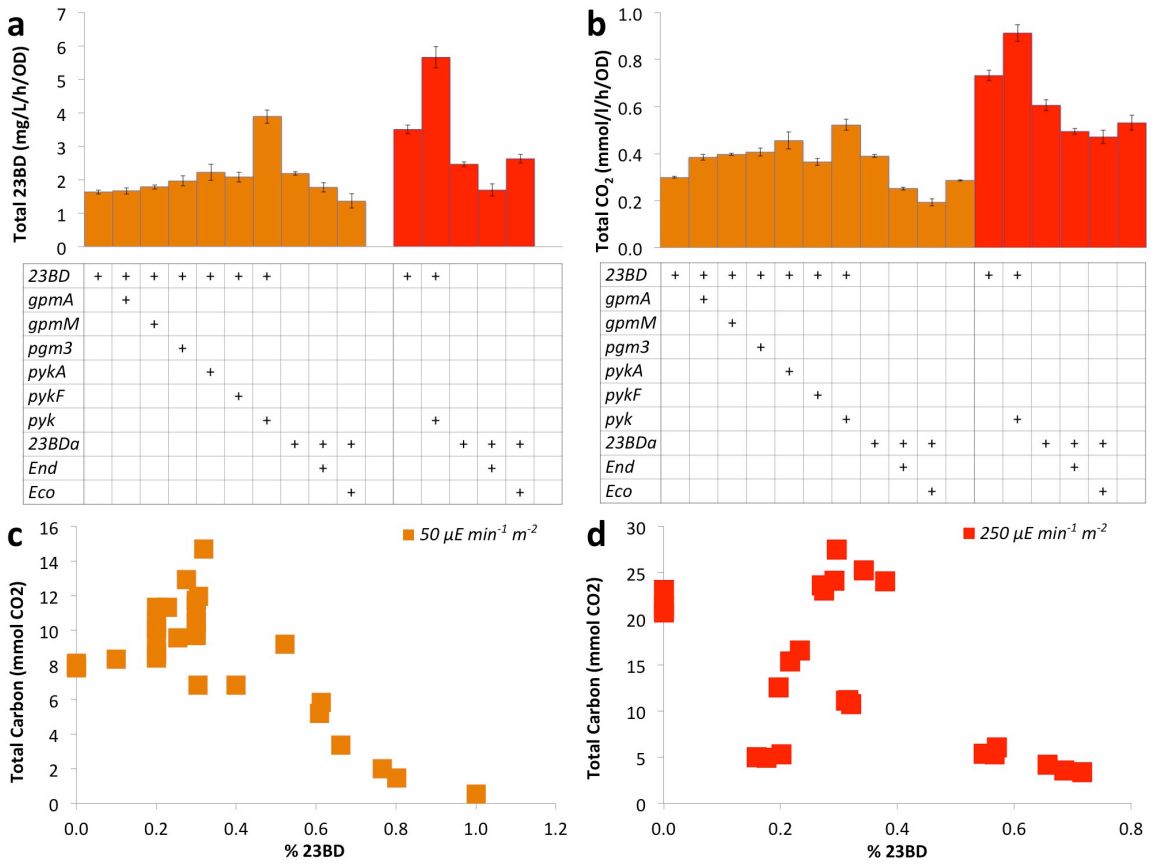


Figure 4-9 Specific Production and Correlations (a) Total BDO titre divided by capacity; (b) Total carbon (biomass and BDO) divided by capacity; (c&d) Total carbon produced per day at growth phases less than OD 1 versus the fraction of carbon made up by BDO, dark blue grown at standard light (50 umol/min/m^2) light blue grown at high light (250 umol/min/m^2, see supporting info for calculations.

In high light, productivity of the *pyk* expressing strain was 1.2 fold higher than the unmodified strain reaching 5.7 mg/L/h/OD overall. Strains expressing *End* and *Eco* pathways achieved 2.5 and 1.7 mg 23BD/L/h/OD respectively, and increase of 1.4- and 1.2-fold over their values at 50 umol/s/m^2 light. This corresponds to production of 0.5 mmol carbon/L/h/OD for both strains, which is an increase of 2.0 and 2.5 fold, respectively. This indicates that increased carbon is being diverted to growth, as was observed in lower partitioning rates. Overall, total carbon is increased in all modified strains expressing single genes, and significantly in the *pyk* expressing strain.

A second productivity analysis was performed by filtering for data obtained during growth phases below OD 1, and plotting total carbon against partitioning, to evaluate the effect of partitioning on carbon fixation (Figure 4-9c&d). Under standard conditions, as partitioning increases there is concurrent increase in total carbon with partitioning until a local maximum corresponding to 15 mmol carbon and 30% partitioning. After this point total carbon declines sharply as partitioning increases. At higher light a similar trend is observed, reaching a maximum of 27 mmol carbon at 30% partitioning before declining sharply to 5 mmol at 60% partitioning.

Oxygen evolution was measured to evaluate whether this increase in carbon correlated with increased activity in the photosystem (Figure 4-10b). During exponential growth, oxygen evolution was very similar between all strains measured, at 9 nmol O_2/min/mg Chl. After 48 hours, oxygen evolution is uniformly decreased. Slightly higher oxygen evolution is observed in the *pyk* expressing strain, corresponding to lower OD and a slower shift into the linear phase.

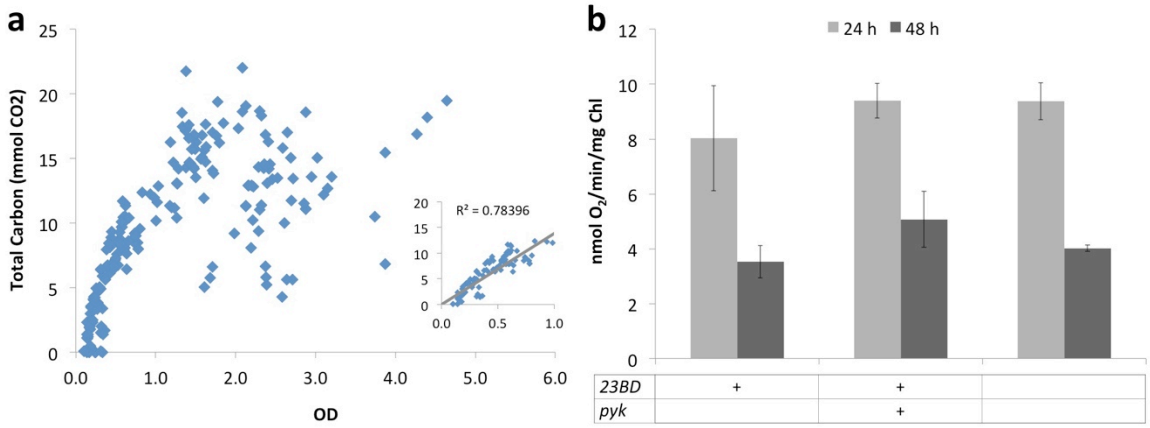


Figure 4-10 Photosynthetic capacity (a) Total carbon per day for all strains and time points plotted against the average OD during the same time period, inset: linear correlation below OD 1; (b) Oxygen evolution for cultures in exponential and linear phases.

4.2 Discussion

In this study, I constructed a carbon sink connecting fixation to 23BD synthesis in the cyanobacterium *S. elongatus*. The top producing strain exhibited a 1.8-fold increase in total carbon fixed, and a 2.4 fold increase in 23BD production per cell overall. Previously it has been hypothesized that the rate of carbon fixation could be faster than the rate of carbon use within the cell, making growth the rate-determining step for photosynthesis[11]. During the testing of strains we generated a range of carbon partitioning, which allowed for a unique insight into the efficiency of fixation during interaction with a carbon sink, uncoupled from growth. As partitioning increased initially, the specific rate of carbon fixation also increased as hypothesized (Figure 4-9c). A local maximum in total carbon uptake during exponential growth was observed at 30% partitioning to 23BD both under standard conditions and at elevated light and temperature indicating that partitioning

impacts the dark reactions irrespective of rate (Figure 4-9c&d). The discovery of a maximum is surprising, as it has been a central goal of metabolic engineering in cyanobacteria to increase partitioning in order to draw more carbon away from growth.[17-19] While we were able to prevent growth entirely through activation of the complete pyruvate pathway, these strains also exhibited severely reduced carbon fixation capacity. This is consistent with the requirement of a minimum pool of carbon in central metabolism to sustain fixation. The correlation between partitioning and fixation observed in this study does not define whether growth is slowed because of competition for carbon, or if competition for carbon is relaxed by inhibition of growth. However, in the former case bottlenecks would be relieved by the carbon sink, and an increase in fixation is predicted. In the second case, bottlenecks would conceivably be increased by inhibited growth, and total fixation would decrease. Further studies are needed to distinguish between these two cases; however, the observation of increased total carbon per cell in this study is encouraging as evidence of carbon being drawn away from growth. Both the PYK and PGM enzymes compete with central metabolism, with PGM limiting synthesis of serine and *pyk* limiting oxaloacetate, a significant source of carbon for the TCA cycle in cyanobacteria[14]. Conversion of oxaloacetate to malate may also be important for the production of NAD^+ to fuel Acetyl-CoA synthesis, an effect that has been seen in the rescue of a slow growth phenotype in mutants harboring a transhydrogenase, through expression of an NADH consuming step[20], as well as rescue of a malic enzyme knockout mutant, by addition of pyruvate in *Synechocysitis* PCC6803.[21,22] However, expression of PGM or genes individually does not severely impact growth, and expression of PYK genes only slightly impairs growth in the 23BD

background (Figure 4-5). It is unlikely that these effects combined are responsible for the loss of growth and carbon fixation with the complete pyruvate pathway (Figure 4-7). In later stages of growth AL757 can continue producing 23BD despite down-regulated growth. This appears to be a global effect at high OD that increases partitioning during the linear and stationary phases. However, in this case total carbon fixation is also decreased in our measurements, as is predicted by slowed growth and falling oxygen evolution (Figure 4-10), which has also been measured previously.[16] An additional effect from PYK is possible through modulation of NADPH/ATP ratios. 23BD production consumes NADPH which impacts the stoichiometry of NADPH/ATP production. Pyruvate kinase produces ATP and may increase this imbalance. However, the cyanobacterial cell has numerous pathways for restoring energetic balance in the presence of high light. Oxygen evolution is relatively constant between strains indicating that the photosystems are not undergoing excessive cycling to maintain energy balance.

Production of lactate from strains expressing a single PYK has recently been demonstrated in *Synechocystis* PCC6803.[17] In this study the authors observe an increase in partitioning and concurrent slight slowing of growth, but not a significant increase in total carbon, indicating that light or carbon may have been limiting in their system. A number of studies have previously reported high carbon partitioning in growing strains, which is incongruous with our current result.[11,23] It is important to clarify that differences in the methods for calculation of partitioning can produce varied results, and that our observation is based on an internally standardized methodology.

4.3 Conclusion

The metabolic carbon sink presented here allows for an increase in total carbon fixation, however at a cost to the stability of fixation overall. Modular expression of genes from the carbon sink pathway allows for varied levels of carbon partitioning, identifying a maximum that is significantly lower than previously envisioned, and a severe inhibition of fixation at high partitioning. Peak specific carbon rates are significantly higher in engineered strains. Theoretical yield can be calculated through either the oxygen evolution or growth rate of wild-type cells in optimized conditions, which give similar estimates of theoretical maximums for carbon fixation at around 3.7 mmol/gDW/h. In our study, 23BD and control strains achieved net specific fixation of 3.18 and 2.5 mmol/gDW/h, which represent a theoretical yield of 86% and 68% respectively. Peak fixation by the best carbon sink strain produced 4.31 mmol/gDW/h, or 117% of the theoretical maximum. This increase in photosynthetic yield is promising for the engineering of efficient carbon capture, and conversion of CO_2 into useful chemicals.

4.4 Construction of plasmids for recombination

Single gene pathways were constructed under control of the P_{LlacO1} IPTG inducible promoter as a single operon. Each pathway was flanked with homologous regions for recombination with neutral site II[24], which allows targeting of genes to the genome of the modified *S. elongatus* strain that has previously been transformed with the 23BD pathway. Because of the high number of doublings, there is a possibility that the transformation process will select against 23BD production during growth from a single cell. Transformations of PYK encoding genes yielded about half the number of positive

colonies as transformations of PGM encoding genes under standard conditions, indicating moderate selective pressure from the gene addition. Colonies of the 23BD parent strain also spontaneously lost productive capacity during transformation, in an unpredictable pattern. The percentage of positive transformants per plate that retained an active 23BD pathway varied from >80% to <40%. Similar losses of activity have been observed in other studies, and the phenomenon is likely to be highly underreported in cyanobacterial studies.[25] After isolation of positive strains, activity remained stable in cultures re-streaked on plates.

It was expected that higher expression of the entire pyruvate pathway may lead to lower transformation efficiencies. To increase chances of positive transformants with an active pathway, transformants were obtained by recombination/activation with the first gene of the 23BD pathway. A base strain was constructed by transforming wild-type with a plasmid containing a truncated 23BD pathway where the promoter and first half of the *alsS* gene are removed. Plasmids were constructed for each 3-gene pathway with the inclusion of a P_{LlacO1} promoter and *alsS* gene, which contains the homologous region for the second half of the *alsS* gene present in the base strain. After transformation, the *alsS* gene and 23BD pathway are completed and activated by the presence of the promoter. Sequencing shows that recombination of the *alsS* gene is complete, reconstituting the 23BD pathway sequence in its original form. An empty vector strain was constructed by recombination of an identical plasmid without inclusion of the pyruvate pathway genes. Active pathways in positive colonies were >90% in all three strains.

Table 4-1 Strains and plasmids used in this study

Strains	Genotype	Source
***E. coli* strain**		
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10* (*Tetr*)]	Agilent Technologies
***S. elongatus* strains**		
	wild type	S. Golden
AL723	$P_L lacO_1$ and Gent^R integrated at NS	[26]
AL757	NSIII*[PLlacO1* :: *alsS-alsD-adh ; Gent^R]*	[26]
AL2064	NSII[*PLlacO1* :: *gpmA* (*E.c.*) ; Kan^R]	This Study
AL2065	NSII[*PLlacO1* :: *gpmM* (*E.c.*) ; Kan^R]	This Study
AL2066	NSII[*PLlacO1* :: *pgm3* (*S.e.*) ; Kan^R]	This Study
AL2067	NSII[*PLlacO1* :: *pykA* (*E.c.*) ; Kan^R]	This Study
AL2068	NSII[*PLlacO1* :: *pykF* (*E.c.*) ; Kan^R]	This Study
AL2069	NSII[*PLlacO1* :: *pyk* (*S.e.*) ; Kan^R]	This Study
AL2081	Same as AL757, with NSII[*PLlacO1* :: *gpmA* (*E.c.*) ; *Kan^R*]	This Study
AL2082	Same as AL757, with NSII[*PLlacO1* :: *gpmM* (*E.c.*) ; *Kan^R*]	This Study
AL2083	Same as AL757, with NSII[*PLlacO1* :: *pgm3* (*S.e.*) ; *Kan^R*]	This Study
AL2084	Same as AL757, with NSII[*PLlacO1* :: *pykA* (*E.c.*) ; *Kan^R*]	This Study
AL2085	Same as AL757, with NSII[*PLlacO1* :: *pykF* (*E.c.*) ; *Kan^R*]	This Study
AL2086	Same as AL757, with NSII[*PLlacO1* :: *pyk* (*S.e.*) ; *Kan^R*]	This Study
AL1792	NSIII*[HRalsS-alsD; Spec^R]*	This Study
AL(pAL711)	NSIII*[P_{trc} :: ; P_{LlacO1} :: alsS-alsD; Gent^R]*	This Study
AL(pAL798)	NSIII*[P_{trc} :: pgm3 - eno (S. elongatus) - pyk; P_{LlacO1} :: alsS-alsD; Gent^R]*	This Study
AL(pAL792)	NSIII*[P_{trc} :: gpmM - eno (E. c.) - pykA; P_{LlacO1} :: alsS-alsD; Gent^R]*	This Study
AL(pAL711)	NSIII*[P_{trc} :: ; P_{LlacO1} :: alsS-alsD-adh; Gent^R]*	This Study
AL(pAL798)	NSIII*[P_{trc} :: pgm3 - eno (S. elongatus) - pyk; P_{LlacO1} :: alsS-alsD-adh; Gent^R]*	This Study
AL(pAL792)	NSIII*[P_{trc} :: gpmM - eno (E. c.) - pykA; P_{LlacO1} :: alsS-alsD-adh; Gent^R]*	This Study
Plasmids		
pAL300	NSIII Targeting vector; *[PLlacO1 :: alsS-alsD-adh; Gent^R]*	[26]
pAM2991	NSI Targeting vector; *[P_{trc} :: ;Spec^R] For source DNA only*	S. Golden
pAL463	NSII targeting vector; *[ColE1*; *P_{LlacO1}*:: ; *Kan^R]*	C.R. Deere
pNozzi	NSII targeting vector; *[ColE1*; P_{LacIq1}*-LacI* ; *Kan^R]*	N. Nozzi
pAL542	Same as pAL463, but with P_{lacIq}*-LacI* integrated upstream of P_{LlacO1}	This study
pAL974	Same as pAL542 but with *[P_{LlacO1}:: gpmA (E. coli MG1655);]*	This study
pAL975	Same as pAL542 but with *[P_{LlacO1}:: gpmM (E. coli MG1655);]*	This study
pAL976	Same as pAL542 but with *[P_{LlacO1}:: pgm3 (S. elongatus);]*	This study
pAL977	Same as pAL542 but with *[P_{LlacO1}:: pykA (E. coli MG1655);]*	This study
pAL978	Same as pAL542 but with *[P_{LlacO1}:: pykF (E. coli MG1655);]*	This study
pAL979	Same as pAL542 but with *[P_{LlacO1}:: pyk (S. elongatus)]*	This study
pAL1015	Same as pAL542 but with *[P_{LlacO1}:: eda (E. coli MG1655)]*	This study
pAL1016	Same as pAL542 but with *[P_{LlacO1}:: pyk - eda;]*	This study
pAL697	NSIII Targeting vector; *[HRalsS-alsD; Kan^R]*	This study
pAL698	NSIII Targeting vector; *[HRalsS-alsD-adh; Kan^R]*	This study
pAL815	NSIII Targeting vector; *[HRalsS-alsD; Spec^R]*	This study
pAL816	NSIII Targeting vector; *[HRalsS-alsD-adh; Spec^R]*	This study
pAL711	NSIII Targeting vector; *[P_{trc} :: ; P_{LlacO1} :: HRalsS; Gent^R]*	This study
pAL798	Same as pAL711 but with *[P_{trc} :: pgm3 - eno (S. elongatus) - pyk;]*	This study
pAL792	Same as pAL711 but with *[P_{trc} :: gpmM - eno (E. c.) - pykA;]*	This study

Table 4-2 Oligonucleotides used in this study

Name	Sequence 5' --> 3'	Plasmid produced
IM66	TTAAGTCTCTATCTCTGCAG	pAL697, pAL978
IM90	CGACTGGAAAGCGGGCAGTG	pAL711
JO14	CAGGAGCTAAGGAAGCTAAACTTGAACGAATTGTTAGACA	pAL815, pAL816
JO15	GCGAGCTCGATATCAAATTAAGCTGTAATGCAAGTAGCGT	pAL815, pAL816
JO38	GCCGGGAAGCCGATCTCGGCTTGAA	pAL711
JO96	AAGCGAGAGTAGGGAACTGCC	pAL792, pAL798
JO100	GCCACCGGTCTCCCGTCTTATTCCACGATGAACAGC	Ligation for pAL792
JO105	GCCACCGGTCTCCGACGATTAAAGAGGAGAAAGGTACAATGTCCAAAATCGTAAAAA	Ligation for pAL792
JO108	GCCACCGGTCTCCAGCATTATGCCTGGCCTTTGATCTC	Ligation for pAL792
JO113	GCCACCGGTCTCCTGCTATTAAAGAGGAGAAAGGTACAATGTCCAGAAGGCTTCGC	Ligation for pAL792
JO116	GGCAGTTCCCTACTCTCGCTTTTACTCTACCGTTAAAATACGCGTGG	pAL792
JO117	AAGCGTAAATCTGGCATCGG	Ligation for pAL792
JO120	AAGCCTCTCGCGAGGAGAGG	Ligation for pAL792
JO124	TAAGGTTGGGAAGCCCTGC	pAL542
JO125	GGTTTCTTAGACGTCGGAATTGTATGGCATGATCGAGCGGG	pAL542
JO126	GCAGGGCTTCCCAACCTTAAGCTGCATTAATGAATCGGCC	pAL542
JO127	CAATTCCGACGTCTAAGAAACC	pAL542
JO172	AAAACGCAGCATTCATGGCC	pAL697, pAL978
JO179	CAGACAGAGCGCACCCTCTTGATGTGAGCAAAAGGCCAGC	pAL711
JO184	CGCTCTCCTGAGTAGGACAAATTCGAGTTGACGCGTGCTTATC	pAL711
JO196	AGCGGGACTCTGGGGTTCGAGTTAGCCGTTGATTGCAGGTGC	pAL697, pAL978
JO197	TGCACCTGCAATCAACGGCTAACTCGAACCCCAGAGTCCCGC	pAL697, pAL978
JO198	TAAACGGCCGACTGCTTGGGCCATGAATGCTGCGTTTTCTCGAGGTGAAGACGAAAGGG	pAL697, pAL978
JO199	TCCTGCAGAGATAGAGACTTAATACTCAGGAGAGCGTTCACCG	pAL697, pAL978
JO200	CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGTGGAAACGGATGAAGGCACG	pAL711
JO202	AATTCGTTCAAGCCGAGATCGGCTTCCCGGCTTAGCCGTTGATTGCAGGTGC	pAL711
JO203	ATAACAATTATTAAAGAGGAGAAAGGTACAATGGTTTTAAGACTCTATCTCG	pAL798
JO204	GCCACCGGTCTCCCGTCAGTTGCTTAAGTTCCAGCG	Ligation for pAL798
JO205	GCCACCGGTCTCCGACGATTAAAGAGGAGAAAGGTACAATGCCAGACGATTACGGGACG	Ligation for pAL798
JO206	GCCACCGGTCTCCAGCAACCTAGCGGCCTTTCGGTCC	Ligation for pAL798
JO207	GCCACCGGTCTCCTGCTATTAAAGAGGAGAAAGGTACAATGCAACCCAACGACTTTCAGC	Ligation for pAL798
JO208	GGCAGTTCCCTACTCTCGCTTTTAGACGGAAATCGGCTGCTCG	pAL798
JO209	GTACCTTTCTCCTCTTTAATAATTGTTATCCGCTCACAATTCCACACATTATACGAGCCG	pAL792, pAL798
JO211	ATAACAATTATTAAAGAGGAGAAAGGTACAATGTTGGTTTCTAAAAAACCTATGG	pAL792
JO213	ATCGCCGCATCAGTACGCC	Ligation for pAL798
JO214	ATTCTCGGGTGGGTCTTCTGCG	Ligation for pAL798
JO215	TTAATTTGATATCGAGCTCGCTTAGCCGTTGATTGCAGGTGC	pAL815, pAL816
JO216	AAGTTTAGCTTCCTTAGCTCCTGAAAACGCAGCATTCATGGCCC	pAL815, pAL816
JO270	ATTAAAGAGGAGAAAGGTACAATGGTTTTAAGACTCTATCTCG	pAL976
JO271	TTTATTTGATGCCTCTAGAACGCGTTTAAGTTCCAGCGCGATCGCGCA	pAL976
JO272	TTAAAGAGGAGAAAGGTACAATGTTGGTTTCTAAAAAACCT	pAL975
JO273	TTTATTTGATGCCTCTAGAACGCGTTTATTCCACGATGAACAGCGGCT	pAL975
JO274	TTAAAGAGGAGAAAGGTACAATGGCTGTAACTAAGCTGGT	pAL974
JO275	TTTATTTGATGCCTCTAGAACGCGTTTACTTCGCTTTACCCTGGT	pAL974
JO284	TTAAAGAGGAGAAAGGTACAATGTCCAGAAGGCTTCGCAG	pAL977
JO285	TTGATGCCTCTAGAACGCGTTTACTCTACCGTTAAAATACGC	pAL977
JO286	ACGCGTTCTAGAGGCATCAAATAAAACGAAAGGCTC	pAL974, pAL975, pAL976, pAL977, pAL978, pAL979, pAL1015, pAL1016
JO287	TGTACCTTTCTCCTCTTTAATGAATTCGG	pAL974, pAL975, pAL976, pAL977, pAL978, pAL979
JO288	TTAAAGAGGAGAAAGGTACAATGAAAAAGACCAAAATTGT	pAL978
JO289	TTGATGCCTCTAGAACGCGTTTACAGGACGTGAACAGAT	pAL978
JO290	TTAAAGAGGAGAAAGGTACAATGCAACCCAACGACTTTCAG	pAL979
JO291	TTGATGCCTCTAGAACGCGTTTAGACGGAAATCGGCTGCTC	pAL979
JO305	TTAAAGAGGAGAAAGGTACAATGAAAAACTGGAAAACAAGTG	pAL1015
JO306	TTTATTTGATGCCTCTAGAACGCGTTTACAGCTTAGCGCCTTCTA	pAL1015, pAL1016

Table 4-3 Cloning guide

Annealed Product	Primer 1	Primer 2	Template
pAL542	JO125	JO126	pNozzi
	JO124	JO127	pAL463
pAL974	JO286	JO287	pAL542
	JO274	JO275	*E. coli* MG1655 gDNA
pAL975	JO286	JO287	pAL542
	JO272	JO273	*E. coli* MG1655 gDNA
pAL976	JO286	JO287	pAL542
	JO270	JO271	*S. elongatus* PCC 7942 gDNA
pAL977	JO286	JO287	pAL542
	JO284	JO285	*E. coli* MG1655 gDNA
pAL978	JO286	JO287	pAL542
	JO288	JO289	*E. coli* MG1655 gDNA
pAL979	JO286	JO287	pAL542
	JO290	JO291	*S. elongatus* PCC 7942 gDNA
pAL1015	JO286	JO287	pAL542
	JO305	JO306	*E. coli* MG1655 gDNA
pAL1016	JO286	JO307	pAL979
	JO306	JO308	*E. coli* MG1655 gDNA
pAL697	JO172	JO199	pAL306
	JO197	JO198	pAL463
	JO196	IM66	pAL300
pAL698	JO172	JO199	pAL300
	JO197	JO198	pAL463
	JO196	IM66	pAL300
pAL815	JO14	JO15	pAM2991
	JO215	JO216	pAL697
pAL816	JO14	JO15	pAM2991
	JO215	JO216	pAL698
pAL711	IM90	MC71	pAM2991
	JO184	NA05	pAL300
	JO179	JO202	pAL300
	JO38	JO200	pAL300
pAL792	JO209	JO96	pAL711
	JO211	JO116	Ligation
pAL798	JO209	JO96	pAL711
	JO203	JO208	Ligation
Ligation (gpmM-eno-pykA)	JO117	JO100	*E. coli* MG1655 gDNA
	JO105	JO108	*E. coli* MG1655 gDNA
	JO113	JO120	*E. coli* MG1655 gDNA
Ligation (pgm3-eno-pyk)	JO213	JO204	*S. elongatus* PCC 7942 gDNA
	JO205	JO206	*S. elongatus* PCC 7942 gDNA
	JO207	JO214	*S. elongatus* PCC 7942 gDNA

4.5 Acknowledgements

This work was supported by the Asahi Kasei Corporation.

4.6 References

1 Hoekstra, A. Y. & Wiedmann, T. O. Humanity's unsustainable environmental footprint. *Science* **344**, 1114-1117, doi:10.1126/science.1248365 (2014).
2 (ed Department of Energy) (U.s. Energy Information Administration).
3 McFarlane, J. & Robinson, S. Survey of Alternative Feedstocks for Commodity Chemical Manufacturing. *Oak Ridge National Laboratory* (2007).
4 Hays, S. G. & Ducat, D. C. Engineering cyanobacteria as photosynthetic feedstock factories. *Photosynth Res*, doi:10.1007/s11120-014-9980-0 (2014).
5 Savakis, P. & Hellingwerf, K. J. Engineering cyanobacteria for direct biofuel production from CO. *Curr. Opin. Biotechnol.* **33C**, 8-14, doi:10.1016/j.copbio.2014.09.007 (2014).
6 Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990 – 2012. (2014).
7 Nozzi, N. E., Oliver, J. W. K. & Atsumi, S. Cyanobacteria as a Platform for Biofuel Production. *Front. Bioeng. Biotechnol.* **1**, doi:10.3389/fbioe.2013.00007 (2013).
8 Camsund, D. & Lindblad, P. Engineered transcriptional systems for cyanobacterial biotechnology. *Front. Bioeng. Biotechnol.* **2**, 40, doi:10.3389/fbioe.2014.00040 (2014).
9 Markley, A. L., Begemann, M. B., Clarke, R. E., Gordon, G. C. & Pfleger, B. F. Synthetic Biology Toolbox for Controlling Gene Expression in the Cyanobacterium Synechococcus sp. strain PCC 7002. *ACS Synth Biol*, doi:10.1021/sb500260k (2014).
10 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. *Photosynthesis Res.* **88**, 287-297, doi:10.1007/s11120-006-9048-x (2006).
11 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, doi:10.1128/AEM.07901-11 (2012).
12 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).
13 McEwen, J. T., Machado, I. M., Connor, M. R. & Atsumi, S. Engineering Synechococcus elongatus PCC 7942 for continuous growth under diurnal conditions. *Appl. Environ. Microbiol.* **79**, 1668-1675, doi:10.1128/AEM.03326-12 (2013).
14 Young, J. D., Shastri, A. A., Stephanopoulos, G. & Morgan, J. A. Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. *Metab. Eng.* **13**, 656-665, doi:10.1016/j.ymben.2011.08.002 (2011).

15 Jablonsky, J., Hagemann, M., Schwarz, D. & Wolkenhauer, O. Phosphoglycerate mutases function as reverse regulated isoenzymes in Synechococcus elongatus PCC 7942. *PLoS One* **8**, e58281, doi:10.1371/journal.pone.0058281 (2013).
16 Ungerer, J. *et al.* Sustained photosynthetic conversion of CO2 to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energ. Environ. Sci.* **5**, 8998-9006, doi:Doi 10.1039/C2ee22555g (2012).
17 Angermayr, S. A. *et al.* Exploring metabolic engineering design principles for the photosynthetic production of lactic acid by Synechocystis sp PCC6803. *Biotechnol. Biofuels.* **7**, doi: Doi 10.1186/1754-6834-7-99 (2014).
18 Melis, A. Carbon partitioning in photosynthesis. *Curr. Opin. Chem. Biol.* **17**, 453-456, doi:10.1016/j.cbpa.2013.03.010 (2013).
19 Yu, Y. *et al.* Development of *Synechocystis sp.* PCC 6803 as a phototrophic cell factory. *Mar. Drugs* **11**, 2894-2916, doi:10.3390/md11082894 (2013).
20 Angermayr, S. A., Paszota, M. & Hellingwerf, K. J. Engineering a cyanobacterial cell factory for production of lactic acid. *Appl. Environ. Microbiol.* **78**, 7098-7106, doi:10.1128/AEM.01587-12 (2012).
21 Yang, C. Metabolic Flux Analysis in Synechocystis Using Isotope Distribution from 13C-Labeled Glucose. *Metab. Eng.* **4**, 202-216, doi:10.1006/mben.2002.0226 (2002).
22 Yoshikawa, K., Hirasawa, T. & Shimizu, H. Effect of malic enzyme on ethanol production by Synechocystis sp. PCC 6803. *J. Biosci. Bioeng.*, doi:10.1016/j.jbiosc.2014.06.001 (2014).
23 Gao, Z. X., Zhao, H., Li, Z. M., Tan, X. M. & Lu, X. F. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energ. Environ. Sci.* **5**, 9857-9865, doi:Doi 10.1039/C2ee22675h (2012).
24 Andersson, C. R. *et al.* Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol.* **305**, 527-542 (2000).
25 Jones, P. R. Genetic instability in cyanobacteria - an elephant in the room? *Front. Bioeng. Biotechnol.* **2**, 12, doi:10.3389/fbioe.2014.00012 (2014).
26 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).

Chapter 5 Methods

5.1 Reagents

The chemicals (*R,R*)-2,3-butanediol , *meso*-2,3-butanediol, (*S,S*)-2,3-butanediol and acetoin were obtained from Sigma-Aldrich (St. Louis, MO). NADH and IPTG were obtained from Fisher Scientific (Hanover Park, IL). Phusion polymerase was purchased from NEB (Ipswich, MA). KOD polymerase and NADPH were purchased from EMD4Biosciences (San Diego, CA). Gentamicin was purchased from Teknova (Hollister, CA). Oligonucleotides were synthesized from Integrated DNA Technologies, Inc. (San Diego, CA).

5.2 Culture conditions

Unless otherwise specified, *S. elongatus* strains were cultured in BG-11 medium[1] with the addition of 50 mM $NaHCO_3$. Cells were grown at 30˚C with rotary shaking (100 rpm) and light (55 μE $s^{-1}m^{-2}$) provided by four 86 cm 20 W fluorescent tubes 5 cm above the cell cultures. Cell growth was monitored by measuring OD_{730}.

For acetoin and 23BD production in *S. elongatus*, cells in exponential phase were diluted to an OD_{730} of 0.1 in 25 ml BG-11 medium including 50 mM $NaHCO_3$, 10 mg/L thiamine, and 10 mg/L gentamicin in 125 ml baffled shake flasks. Cultures were grown to an OD_{730} of 0.2-0.4 before induction with 1 mM IPTG. Every 24 hours, 10% of the culture volume was removed, the pH was adjusted to 7.5 +/- 0.4 with 10N HCl, and volume was replaced with of BG-11 containing 0.5 M $NaHCO_3$, achieving a final concentration of 50 mM $NaHCO_3$ in the culture.

For comparison of 23BD production in cultures with combinatorial variations, 10 ml of exponentially growing cultures was harvested by centrifugation at 1,500 *g* for 10

min, resuspended to OD_{730} 0.1 in production media (BG-11 medium with the addition of 50 mM $NaHCO_3$, 10 μg/ml thiamine, and 10 μg/ml gentamicin), and allowed to recover for 24 h (~ OD_{730} 0.3) before induction with 1 mM IPTG.

For acetoin production in cultures with combinatorial variations, cultures were prepared as for 23BD with the exception that cultures were resuspended to OD_{730} 0.4 and induced with 1 mM IPTG immediately. The cells were grown for 24 h. A 360 μl sample was harvested by centrifugation at 1,500 *g* for 10 min and acetoin was quantified from the supernatant. In cases where production rate was determined acetoin was measured every 12 h by the same method.

For 23BD production in high light experiments, 25ml cultures were shaken in 125ml baffled flasks on clear polycarbonate ~1" above 4 cool white fluorescent bulbs, providing ~250 55 μmol $s^{-1}m^{-2}$ photons in the PAR range. Every 12 hours 5% of media was removed, the pH was adjusted to 7.5 +/- 0.4 with 10N HCl, and volume was replaced with of BG-11 containing 0.5 M $NaHCO_3$, achieving a final concentration of 50 mM $NaHCO_3$ in the culture.

For acetoin and 23BD production in *E. coli,* overnight cultures were diluted 1:100 into 5 ml of modified M9 medium[2] containing 50 g/L glucose, 5 g/L yeast extract, and 5 mg/L gentamicin in 30-ml culture tubes. Cells were grown at 37˚C to an OD of 0.2-0.4 followed by addition of 0.1 mM IPTG. Production was continued at 30°C on a rotary shaker (250 rpm) for 40 hours.

5.3 Transformation of *S. elongatus*

Transformation of *S. elongatus* was performed as described[3]. Strains were segregated several times by transferring colonies to fresh selective plates. Correct recombinants were

confirmed by PCR to verify integration of targeting genes into the chromosome. Neutral Site III (NSIII) corresponds to recombination between Synpcc7942_0893 (903,564 - 904,283 bp) and Synpcc7942_0894 (904,845 - 905,417 bp) on the *S. elongatus* chromosome. It was confirmed that insertion of the gentamicin resistance gene at this site does not affect the growth of cells. Neutral site II (NSII) has been characterized previously.[4]

5.4 Acetoin Quantification

Acetoin was quantified by the method of Voges and Proskauer[5,6], adapted to small volume on 96 well plates. Sample concentration was varied between 1-10% of final volume to achieve a result within the linear range of detection. This was achieved by dilution in H_2O to 100 μl initial volume. For an assay containing 2% sample (most common), 98 μl water and 2 μl of the supernatant were added to wells and mixed. To this was added 100 μl of a solution, prepared at the time of use, consisting of one part 5% Naphthol dissolved in 2.5N NaOH and one part 0.5% Creatine in water. The assay was monitored every 5 minutes and final readings were taken after 40 min, when the slope of the absorbance curve matched the background oxidation rate of Naphthol. Triplicate measurements of no less than 3 standards, including at least one value each above, below, and within the desired range, were included in every assay.

5.5 2,3-Butanediol Quantification

Supernatant samples from cultures were analyzed with gas chromatography (GC) (Shimadzu) equipped with a flame ionization detector and an HP-chiral 20b column (30 m, 0.32-mm internal diameter, 0.25-mm film thickness; Agilent Technologies). Samples

were prepared by mixing 9 parts supernatant (diluted as necessary in H_20) with 1 part internal standard. For each analysis the GC oven temperature was held at 40˚C for 4 min, increased with a gradient of 15˚C min^{-1} until 235˚C, and held for 4min. Ultra high purity Helium was used as the carrier gas. The temperature of the injector and detector were set at 250˚C. The stereoisomers were identified by matching retention time to standards for (*R,R*)-23BD, *meso*-23BD, and (*S,S*)-23BD.

5.6 O_2 evolution

Evolution of O_2 was measured using a clark-type electrode with the Oxygraph system (Hansatech Instruments Ltd, Norfolk, UK). Under ambient light conditions, 1ml of cells was transferred to a 4 ml borosilicate glass chamber and headspace gas was expelled using a center bored contact plunger with rubber cap. Cells were bubbled with ambient air to standardize oxygen content, and stirred at 100 rpm using a magnetic flea for 2 minutes in darkness to allow the cells to equilibrate with the surrounding water jacket to 25˚C. Cells were then subjected to excess light (60 μE s^{-1} m^{-2}) and allowed to equilibrate for at least 2 minutes until a constant rate could be measured over at least 60 s.

5.7 Enzyme assays

S. elongatus cells were collected 72h after induction by centrifugation (4000g, 5 min) washed in 50mM potassium phosphate buffer (pH 7.5) and resuspended in the same buffer. Crude extract were prepared with 0.1-mm glass beads and a Mini bead beater (Mini Bead Beater 8 (BioSpec Products, Inc., Bartlesville, OK)). The total protein determination was performed by Advanced Protein Assay Reagent from Cytoskeleton, Inc. (Denver, CO).

Acetolactate synthase (ALS) activity was determined as described.[7] The concentration of acetoin produced was measured by a standard curve using pure acetoin. One specific unit of AlsS activity corresponds to the formation of 1 nmol of acetoin per mg of protein per minute.

Alcohol dehydrogenase (ADH) activity was determined by measuring the oxidation of NAD(P)H. The reaction mixture contained 50 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS) pH 7.0, 25 mM acetoin and 0.2 mM NAD(P)H. The consumption of NAD(P)H was monitored at 340 nm. One specific unit of ADH activity corresponds to the oxidation of 1 nmol of NAD(P)H per minute per mg of protein. Background activity was determined for each sample by addition of MilliQ water in place of substrate, and this rate was subtracted from final rates measured with substrate. For use of acetoin as a substrate by native enzymes, activity of the host strain is published as a comparison. The rate without substrate across all *S. elongatus* samples was 8.7 ± 3.9 for NADPH, and 4.8 ± 3.0 (nmol min^{-1} mg^{-1}) for NADH. Across *E. coli* samples the rate without substrate was 4.3 ± 5.6 nmol min^{-1} mg^{-1} measured with NADPH. This does not include the *E. coli* strain containing sADH (*C. b.*), which displayed a higher rate of 69 nmol min^{-1} mg^{-1} in the absence of substrate, and is given independently.

The K_m and *kcat* of sADH (*C. p.*) were determined using C-terminal His-tagged sADH (*C. p.*), the reaction mixture contained 50 mM 3-(N-morpholino) propanesulfonic acid buffer, pH 7.0, 0.25 mM NADPH, acetoin as substrate, and C-terminal His-tagged sADH (*C. p.*)(50 nM). The mixture was incubated at 30°C, and NADPH oxidation was deter-mined at 340 nm using a Synergy H1 plate reader (Biotek). The K_m and V_{max} values

of sADH (*C. p.*) were extrapolated after nonlinear regression of the experimental points using Excel Solver software.

For measurements in Chapter 2, 2-Acetolactate decarboxylase (ALDC) was determined in crude lysate by coexpression with ALS. Activity was determined as described for ALS[7] with the exception that acidification is omitted. Absence of acidification prevents conversion of 2-acetolactate to acetoin in the absence of ALDC. The concentration of acetoin produced by ALDC was measured by a standard curve using pure acetoin. A concurrent background rate of 2-acetolactate conversion in the absence of ALDC was substracted from each activity. One specific unit of ALDC activity corresponds to the formation of 1 nmol of acetoin per mg of total protein per minute.

For measurements in Chapter 3, ALDC activity was determined using the protocol for the ALS assay with the following modifications. The substrate was replaced with 2-acetolactate prepared from ethyl-2-acetoxy-2-methylacetoacetate as described.[8] The acidification step was omitted, and reactions were quenched by the transfer of 20 µl of reaction by multichannel pipette to wells in a 96-well plate, each containing 80 µl 2.5M NaOH. Acetoin was immediately quantified by addition of reagent as described above. One specific unit of ALDC activity corresponds to the formation of 1 nmol of acetoin per mg of protein per minute. For the paired ALDC assay pyruvate was used as substrate, and the acidification step was omitted as described previously.[9]

β-galactosidase (LacZ) activity was determined in the cell extracts by colorimetric assay using *o*-Nitrophenyl-β-D-galactoside (ONPG) as substrate as described.[10] The production of *o*-nitrophenol was monitored at 420 nm. One specific unit of LacZ corresponds to the hydrolize of 1 nmole of ONPG per minute per mg of protein.

Luciferase activity was assayed using the Luciferase Assay System (Promega Corp., Madison, WI). Briefly, 100 μl of Luciferase Assay Reagent was added to 20 μl 10x diluted lysate. Total luminescence was measured for 10 seconds per well.

5.8 Protein purification

sADH (*C. p.*) was overexpressed from pAL410 in *E. coli* BL21 star™ (DE3) (Invitrogen). Overexpressed proteins were purified with His-Spin Protein Miniprep™ (Zymo Research). The protein concentration was determined by Advanced Protein Assay Reagent (Cytoskeleton).

5.9 SDS-PAGE analysis

To estimate the amount of the installed enzymes translated in AL757, cell extracts of AL757 were analyzed with Coomassie Brilliant Blue-G250 staining of SDS-PAGE gels. AL757 and the wild-type were cultured for 72h with or without IPTG. The cells were centrifuges, washed, and resuspended in potassium phosphate buffer (pH 7.5). The cells were lysed with a bead beater (BioSpec Products) and centrifuged 16,000 g for 20 minutes at 4°C. The soluble protein concentration was determined by Advanced Protein Assay Reagent (Cytoskeleton) and 50 μg of protein was run per lane on 4-20% Mini-Protean TGX™ acrylamide gel (Bio Rad). ImageJ (NIH) was used to quantify bands on the gel.

To estimate the *lacZ* expression under different RBS sequences in *S. elongatus*, cell extracts of AL723 (without *lacZ*), AL1103 (RBS U1), AL1104 (RBS A), AL1105 (RBS B), AL1106 (RBS C) and AL1107 (RBS D) were analyzed with Coomassie Brilliant Blue-G250 staining of SDS-PAGE gels. The expression of *lacZ* was induced for

16 h with IPTG. The cells were centrifuged, washed, and suspended in potassium phosphate buffer (pH 7.5). The cells were lysed with a bead beater (Mini Bead Beater 8 (BioSpec Products, Inc., Bartlesville, OK)) and centrifuged 14,000 g for 20 min at 4°C. The soluble protein concentration was determined by Advanced Protein Assay Reagent (Cytoskeleton, Inc., Denver, CO) and 30 μg of protein was run per lane on 4-20% Mini-Protean TGX™ acrylamide gel (Bio Rad). ImageJ (NIH) was used to quantify bands on the gel.

5.10 References

1 Rippka, R. D., J.; Waterbury, J.B.; Herdman. M. Stainer, R.Y. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Journal of General Microbiology* **111**, 1-61 (1979).

2 Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89, doi:10.1038/nature06450 (2008).

3 Golden, S. S., Brusslan, J. & Haselkorn, R. Genetic engineering of the cyanobacterial chromosome. *Methods Enzymol* **153**, 215-231 (1987).

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

5 Voges, O., Proskauer, B. Beitraege zur Ernaehrungsphysiologie und zur Differential Diagnose der Bakterien der hemmorrhagischen Septicamie. *Z. Hyg.* **28** (1898).

6 Westerfeld, W. W. A colorimetric determination of blood acetoin. *J Biol Chem* **161**, 495-502 (1945).

7 Yang, Y. T., Peredelchuk, M., Bennett, G. N. & San, K. Y. Effect of variation of Klebsiella pneumoniae acetolactate synthase expression on metabolic flux redistribution in Escherichia coli. *Biotechnol Bioeng* **69**, 150-159, doi:10.1002/(SICI)1097-0290(20000720)69:2<150::AID-BIT4>3.0.CO;2-N [pii] (2000).

8 Bastian, S. & Arnold, F. H. Reversal of NAD(P)H cofactor dependence by protein engineering. *Method. in Mol. Biol.* **834**, 17-31, doi:10.1007/978-1-61779-483-4_2 (2012).

9 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).

10 Miller, J. H. (Cold Spring Harbor Laboratory Press, 1992).

Chapter 6 Epilogue

Photosynthesis provides a unique opportunity to create a sustainable renewable society, and the efficiency and diversification of engine ered photosynthetic metabolism will play a critical role in defining the scale of civilization in the absence of fossil carbon deposits. During my graduate work I have designed a robust pathway for the production of 2,3-butanediol, optimized protein balance through modulation of translation, and explored the limits of photosynthetic carbon flux in the cyanobacterium *S. elongatus*. It became a central goal of mine to understand the complex relationships between photosynthetically growing cells and the titers of chemicals they produce. My best estimates are described in Chapter 4&Appendix 1, and unfortunately still only represent approximations. Precise conclusions are obfuscated by a high number of unknowns, including cell number (approximated from OD), carbon consumption (provided in excess), and specific irradiation per cell (varied by OD) to name a few. To control the regulation of optimized metabolism, and increase chemical production to near the rates of carbon fixation, will require a highly interdisciplinary systems biology approach, marrying synthetic biology with high-throughput analytical measurement, and computational modeling. We can only hope that such an approach will be undertaken, either in academia or the private sphere. Photosynthetic chemical production will never reach the rates of carbon turnover seen in fermentative systems, and it is important for the scientific community and the public to understand cyanobacterial systems as unique and expensive solutions, useful for recycling of resources, for amelioration of wastes, and above all, for slowing the flux of carbon into the atmosphere. As with other forms of recycling, carbon capture will likely never produce more wealth than it costs, and, rather

than being a highly productive part of the economy as oil and coal have, will require a large investment in subsidies to continue the availability of essential carbon product. However, once the paradigm around carbon has been shifted from that of an endless resource, to that of a hard fought for necessity, cyanobacterial technology will become incredibly valuable. I believe that the limits of photosynthesis, if we are to achieve them, are most clearly seen as the oxygen evolution of the fastest growing photosynthetic cells (3.5h doubling time) uncoupled from NADPH turnover (3x rate), which converts to about 21 mmol CO_2/gDW/h. By the definition of optical density, 99.9% of light is deflected during passage through 1cm of cells at a concentration of roughly 6.6 gDW/m^2. Thus, if we make the assumption that deflected light is consumed, the optimum capacity of engineered cells for carbon capture and conversion to chemical products will be close to 140 mmol/m^2/h (1.7 g C/ m^2/h). This means that a square meter of rooftop in the sunlight, could produce a single charcoal briquette, roughly 25 grams of carbon, each day (14 hours). However, an average house with its roof surface layered in a 1cm thick aquarium of cyanobacteria could capture three bags of charcoal a day. The economic value of this charcoal to a company would be small; however the implications of this resource, for power and chemical production in cities, or in space, are profound.

Appendix I - The relationship between growth and productive capacity in a cyanobacterial culture

Theoretically each cell in a cyanobacterial culture has a specific capacity for photosynthesis and carbon fixation. A culture that has more cells will have a larger capacity, will fix a larger amount of carbon, and can produce a larger amount of biomass and chemical product. As a culture produces biomass it increases its cell number, and therefore the total capacity of the culture is in constant flux. This raises an interesting question for photosynthetic chemical production. *If cultures are to partition carbon between biomass and chemical synthesis, and if biomass increases the rate of chemical synthesis, then what is the optimum partition*? The answer is that optimum partition approaches zero if cells are allowed to grow to an infinite density. This makes sense, because the gains from increasing capacity are much greater than the gains of production from one cell. If cells are limited to a maximum optical density (after which production stops), then the optimum partition approaches 100%. This partition also makes sense, given than slow growth allows cells to produce for an infinite amount of time before production is stopped. Actual production from cells in laboratory cultures lies somewhere between these two cases. As optical density increases above OD 1, carbon fixation efficiencies and light capture per chlorophyll drop significantly but do not stop completely[1]. Thus, higher density cultures still produce more 23BD than low density cultures, but not as much as mid-density growing cultures, which achieve a maximum between high fixation rates and cell number. This maximum may vary between experiments and is instantaneous in time for growing cultures. It is easy to see how this

relationship can lead to confusion when analyzing cultures with different partitioning rates. Careful analysis of growth phase and cell number must be considered when comparing data. As described in Appendix II, the careless use of mathematics can also greatly skew measured results.

A perfect solution to the growth vs partition relationship is the overall aim of achieving 100% carbon partitioning, which can be activated at the point of maximum between fixation rates and cell number, freezing growth without affecting the fixation rate. A second solution is overcoming the regulation that causes loss of efficiency at high cell density. Both aims will require significant advances in engineering to achieve.

Appendix II - Calculations for productivity in cultures with differing growth rates

It is highly desirable to compare carbon fixation and chemical production between a culture that grows quickly to a high density, and a culture that grows slowly. However, it is not simple to achieve this comparison accurately. An exact comparison would require knowledge of the precise capacity of each culture for carbon fixation and light harvesting, at each instant over the entire time course of the experiment. This is not technically feasible. Instead, an estimate must be made. Traditionally, it has been acceptable to normalize production by dividing titer by optical density from the same time point to give units of mg/L/OD (Figure 0-1a, hereafter called the "*dividing by OD*" method). This has obvious limitations if cells are growing exponentially, as most of the optical density has been productive for only a very short time and the bulk of the productive time has occurred at lower cell densities. The second most popular estimate is made by dividing by OD and then by time to yield a rate with the units mg/L/h/OD (Figure 0-1b, hereafter called the "*square*" method). This estimate approximates productivity as if all of the cells have been in the culture the entire time period, which is clearly an overestimate for growing cultures. Capacity is defined here as the sum of hours that each cell has existed in culture between two time points, and is calculated from optical density as the area under the growth curve, in units of OD*hours (ODh). This can be approximated as a straight line (Figure 0-1c, hereafter called the "*linear*" method), or an exponential curve (Figure 0-1d, hereafter called the "*exponential*" method), depending on the growth phase

of the culture. Dividing titer (mg/L) by capacity (ODh) yields a rate with the units mg/L/h/OD.

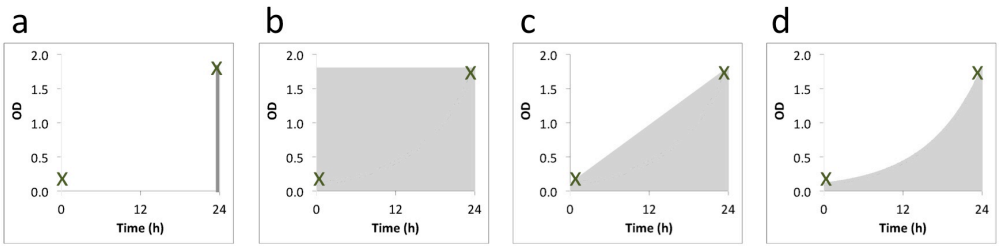


Figure 0-1 Graphical depiction of capacity as estimated by different methods (a) "Divided by OD" - dividing titer by optical density at a single point; (b) "Square" - dividing change in titer by optical density and time at a single point; (c) "Linear" - dividing change in titer by the area under a linear growth curve between two points; (d) "Exponential" - dividing change in titer by the area under an exponential growth curve between two points.

If a number of cultures have grown at different rates between two time points, and the productivity of each cell has been constant and identical, then a perfect normalization would result in equal rates for each culture based on their final titer, final OD, and initial OD. Depending on the type of growth curve between two points, the different analyses yield very different productivities for each culture. This has been modeled for three cases, linear growth, exponential growth, and fast then slow growth. Fast then slow growth can be caused by a number of scenarios, including exposure of a micro-anaerobic culture to increased oxygen during OD sampling, or consumption of carbon source in batch cultures during the first 12 hours of a 24 hour cycle. In such scenarios the actual curve is unknown. Each of the normalization approximations for capacity gives an erroneous relationship between cultures when not perfectly matched to growth (Figure 0-2a-c).

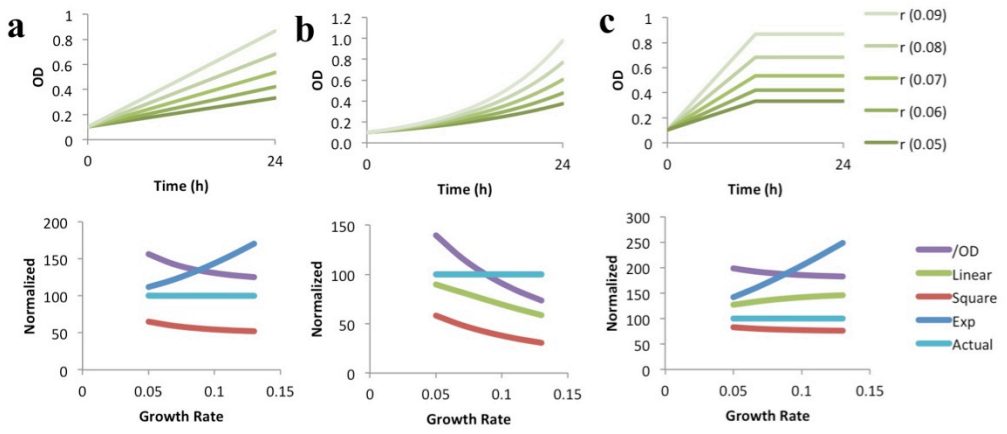


Figure 0-2 Normalized productivity in three modeled growth scenarios (a) Linear growth at five different rates; (b) exponential growth at the same rates as in (a); (c) Fast then slow growth, yielding the same final OD as Linear growth; (d) Normalization of each growth condition using four approximations of capacity, purple indicates the *dividing by OD* method, green indicates the *linear* method, red indicates the *square* method, dark blue indicates the *exponential* method, and light blue indicates the actual relationship between the strains, which all have identical productivity per cell.

In the case of exponential growth, *dividing by OD,* and *linear* capacity both greatly inflate the productivity of slow growing cultures relative to fast growing cultures. In the case of fast then slow growth, or linear growth, using the *exponential* approximation for capacity greatly inflates the productivity of fast growing cultures relative to slow growing cultures. The conclusion is simple: *For growth between two time-points, where the growth curve is unknown, a normalization of productivity per cell, for the comparison of strains with different growth rates, cannot be made.*

On the other hand, for growth with many time-points, where an accurate description of the mode of growth can be made, calculating capacity as the area under the curve yields a very good comparison. For linear and exponential growth (Figure 0-2a&b), when the approximation of capacity matches that of the growth curve (*Linear* and *exponential* methods respectively), normalization yields the correct relationship between

cultures of different growth rates (Figure 0-2d). Using the correct *linear* or *exponential* method for capacity also correctly estimates the scale of productivity. Normalizing by the *square* method yields units of mg/L/h/OD, however it underestimates the quantitative productivity of cells.

In conclusion, productivity should not be estimated between two time-points without characterization of the shape of the growth curve. Productivity should not be estimated using final OD and time from a single final data point. The best estimate of productivity is made by calculating capacity as the area under the growth curve using multiple time-points to define the mode of growth, and can be given as total productivity over the course of production.

Appendix III - Approximations of maximum carbon fixation rates

There are two methods for estimating maximum carbon fixation rates: 1) By growth rate, and 2) By oxygen evolution.

AIII.1 Estimating carbon fixation by growth rate

Growth rates are expressed as a percentage biomass increase per unit time. In the case of cyanobacteria, the unit of time is hours, and a typical maximum growth rate for *S. elongatus* is 0.1, or 10% biomass increased per hour. The amount of carbon in biomass has been measured as 51.34%, or 41 mmol/gDW[2]. Estimating a stoichiometric fixation of carbon per carbon in biomass, the rate becomes 4.1 mmol CO_2/gDW/h. This estimate of fixation is useful for theoretical calculations where the initial dry weight of cells in a culture is known. However, in practice, the only measurement of culture density is the optical density, which must be correlated to a theoretical dry weight.

AIII.2 Estimating carbon fixation by Oxygen evolution

During exponential growth, the rate of oxygen evolution closely matches the rate of carbon fixed by the CB cycle[3] (see also the 'Theoretical yields' section of Chapter 1). Maximum rates of oxygen evolution have been measured at ~200 umol O_2/mg Chl/h with saturating light and CO_2. The concentration of chlorophyll in cells is typically close to 4 ug/ml/OD during production in experiments discussed in this dissertation. Estimating stoichiometric carbon fixation per oxygen evolved gives a rate of 0.8 mmol CO_2/L/h/OD.

Appendix IV - Conversions between OD and biomass

Previously a value of 0.38 g/L/OD has been used for the estimation of biomass from OD in cultures of *Synechococcus elongatus* PCC 7942[4]. In my hands I have measured a value of 0.27 g/L/OD, by rinsing cells in distilled water to remove salts, measuring OD, pelleting and freeze drying before weighing dry cell mass. However, it is likely that the true value for dry weight of cells is even lower.

AIV.1 Theoretical basis for biomass calculations from OD

A growth curve can be drawn with the following equation

a. $OD_0 * 2^{(t/Dt)}$

Where OD_0 is the OD at time=0, t is the time in hours, and Dt is the doubling time.
The following conversions facilitate a relationship between the O_2 evolution rate and maximum doubling time.

Oxygraph rate from Appendix III: *1 mmol CO_2/L/h/OD*

Biomass from rate:

Formal Weight of Carbon: 12 g/mol

1 umol O_2 evolved = 1 umol carbon fixed

1 umol carbon fixed = 12 ug carbon fixed

Contribution of carbon to biomass: *By weight carbon is 51.38% of biomass*

12 ug carbon fixed = 23.4 ug biomass

Therefore: 1 umol CO_2 fixed = 23.4 ug biomass

Biomass generation rate = 0.8 mmol CO_2/L/h/OD * 23.4 = *18.7 mg biomass/L/h/OD*

By multiplying the rate of carbon fixation per OD by the growth curve we get an equation for the increase in O_2 evolution rate over time.

b. $v * OD_0 * 2^{(t/Dt)}$

Where v is the rate, OD_0 is the OD at time=0, t is the time in hours, and Dt is the doubling time.

To calculate the amount of carbon fixed over time (C_{fix}) we integrate this equation.

c. $C_{fix} = ((v * OD_0 * D_t)/(\ln(2)) * (2^{tf/Dt}-2^{ti/Dt})$

Between the same time-points that we integrate we can calculate the increase in OD (dOD) with the following equation.

d.i. $OD_i = OD_0 * 2^{(\mathbf{ti}/Dt)}$

d.ii. $OD_f = OD_0 * 2^{(\mathbf{tf}/Dt)}$

d*.iii.* $dOD = OD_f - OD_i = OD_0 * 2^{(\mathbf{tf}/Dt)} - OD_0 * 2^{(\mathbf{ti}/Dt)}$

Which can be simplified to:

d. $dOD = OD_0 (2^{\mathbf{tf}/Dt} - 2^{\mathbf{ti}/Dt})$

We now have equations that give us the total carbon fixation into biomass in mg/L based on the rate of O_2 evolution over time during growth (***Equation c.***), and we have the total OD increase produced by that fixation over the same period (***Equation d.***). If we divide the biomass produced by fixation by the OD increase we can calculate the maximum theoretical mass of each OD unit (M_{OD}).

e.i $$M_{OD} = \frac{C_{fix}}{dOD} = \frac{((OD_0 * v * D_t)/(\ln(2)) * (2^{tf/Dt} - 2^{ti/Dt})}{OD_0 (2^{tf/Dt} - 2^{ti/Dt})}$$

As you can see, the exponentials and OD_0 cancel out. We are left with an equation depending only on the rate of carbon fixation v, and the doubling time D_t.

e. $$M_{OD} = (v * D_t)/\ln 2$$

Both v and D_t can be determined experimentally.

If we use the converted rate of biomass accumulation from the O_2 linear rate constant (v = 18.7 mg biomass/L/h/OD and the doubling time of 7.5 hours we get:

f. $$M_{OD} = 0.14 \text{ g/L/OD}$$

AIV.2 A Second Proof for Biomass per OD

Definition A: The growth rate describes biomass doubling, while OD is a measurement of

biomass.

i.e. A doubling of OD is assumed to be a doubling of biomass. The growth rate can be expressed in terms of biomass or in terms of OD. However, biomass and OD are related by a constant k_B that varies over time depending on many things including but not limited to variation in chlorophyll concentration.

Definition B: Biomass increases must be equal to or less than total fixation over the same time period.

i.e. Biomass is 51.38% carbon, which is generated purely from fixation of CO_2 during autotrophic growth. The fixation rate is at maximum equal to the O_2 evolution rate. Under normal circumstances biomass cannot be generated from any source other than CO_2 fixation linked to O_2 evolution.

Doubling time and growth rate are related by the following equation.

I.i. $D_t = \ln(2)/r$

I. $r = \ln(2)/D_t$ (units of h^{-1})

Where r is the growth rate i.e. the fraction by which the biomass increases per unit time. The unit time is defined by the unit used in doubling time, in this case it is hours.

From definition 1, the specific growth rate (r_B) depends on the starting biomass (M_B).

II. $r_B = M_B * \ln(2)/D_t$ (units of g biomass /h)

i.e. For a doubling time of 7.5 hours, r = ln(2)/7.5 = 0.092 = 9.2% growth per hour. If initial biomass is 1000 mg dry weight (gDW) and r =0.092 then 0.092 g biomass will be generated after 1 hour. This can be expressed as 0.092 g biomass/gDW/h.

Optical density and biomass can be related by the biomass conversion constant k_B

III. $M_B = k_B * OD \rightarrow OD = M_B / k_B$

Where M_B is biomass in gDW, and k_B is the biomass OD conversion constant.

The specific carbon fixation rate (r_f) is given by:

IV. $r_f = v * OD$ (units of mmol CO_2/L/h)

Where v is the linear fixation rate constant measured by oxygen evolution in units mmol CO_2/L/h/OD.

We can combine equation III and IV to convert the fixation rate to biomass units.

V. $r_f = v * M_B / k_B$ (units of mmol carbon/L/h)

From Definition 2 we know that equations II and V are equal:

VI.i. $r_B = r_f \rightarrow M_B * \ln(2)/D_t = v * M_B / k_B$

VI.ii. $\ln(2)/D_t = v / k_B$

VI. $\mathbf{k_B = (D_t * v) / \ln(2)}$ Which is the same equation as the proof by integration.

Assumption: Maximum carbon fixation rates are concurrent with minimum doubling times.

References

1 Ungerer, J. *et al.* Sustained photosynthetic conversion of CO2 to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energ. Environ. Sci.* **5**, 8998-9006, doi:Doi 10.1039/C2ee22555g (2012).

2 Young, J. D., Shastri, A. A., Stephanopoulos, G. & Morgan, J. A. Mapping photoautotrophic metabolism with isotopically nonstationary (13)C flux analysis. *Metab. Eng.* **13**, 656-665, doi:10.1016/j.ymben.2011.08.002 (2011).

3 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. *Photosynthesis Res.* **88**, 287-297, doi:10.1007/s11120-006-9048-x (2006).

4 Oliver, J. W., Machado, I. M., Yoneda, H. & Atsumi, S. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. USA* **110**, 1249-1254, doi:10.1073/pnas.1213024110 (2013).