UC Davis
UC Davis Electronic Theses and Dissertations

Title
Expanding Microbial Potential for Renewable Chemical Production

Permalink
https://escholarship.org/uc/item/45b2t6xr

Author
Matson, Morgan

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Expanding Microbial Potential for Renewable Chemical Production

By

MORGAN MAYLINDA MATSON

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:

Dr. Shota Atsumi, Chair

Dr. David B. Goodin

Dr. Peter Beal

Committee in Charge

2021

Dedication

To Malinda Matson.

Table of Contents

Dedication ii

Abstract vii

Acknowledgements viii

Chapter 1: Introduction 1

1.1 Introduction: Microbial Production as an Alternative to Fossil Fuels 2

1.2 *E. coli* as a chassis for microbial production 2

1.3 Enabling Non-photosynthetic CO_2 fixation through Electrocatalysis 4

1.4 Alleviating Product Toxicity 10

1.5 References 12

Chapter 2: Electrical-biological hybrid system for CO_2 reduction 17

2.0 Abstract 18

2.1 Introduction 19

Fig. 2.1. Electro-biochemical CO_2 conversion in this study. 21

2.2 Methods 23

2.2.1 Reagents 23

2.2.2 Strains and plasmids 24

2.2.3 Culture conditions 24

2.2.4 Gas chromatography (GC) sample preparation 25

2.2.5 GC- mass spectrometry (MS) analysis 25

2.2.6 Electrochemical cultivation 26

2.3 Results 27

2.3.1. Transferring the formate-assimilation pathway into *E. coli* 27

Table 2.1. List of strains used in this study. 29

Fig. 2.2. Transfer the pathway from formate to CH_2-THF to E. coli. 30

2.3.2 Engineering RGP in glycine auxotroph 31

Fig. 2.3. Glycine production from CO_2 using the reverse reaction of GCS. 32

2.3.3 L-serine synthesis from formate and CO_2 via RGP in L-serine auxotroph 33

Fig. 2.4. L-serine synthesis from formate and CO_2 via RGP 34

2.3.4 Optimization of the L-serine-to-pyruvate pathway 35

2.3.5 Pyruvate synthesis from formate and CO_2 36

2.3.6 Electrochemical cultivation 37

Fig. 2.5. Electrochemical-cultivation system for the RGP Strain. 39
2.5 Conclusion 41
2.6 References 42
2.7 Supplemental Information 52
Table S2.1. E. coli strains used in this study 52
Table S2.2. List of plasmids used in this study 53
Table S2.3. Oligonucleotides used in this study 54
Table S2.4. DNA sequences of synthesized genes used in this study 56
Table S2.5. Set of primers and template for plasmid construction 59
Table S2.6. Conversion efficiency from CO_2 to formate 60
Table S2.7. Conversion efficiency of electricity to formate with the indium electrode in bacterial culture 61
Table S2.8. Abundance of ^{13}C-labeled carbons in L-serine and L-alanine 62
Fig. S2.1. Growth of the formate-assimilation strains on L-serine or glycine 63
Fig. S2.2. Formate and aTc dependency of the Cl operon 64
Fig. S2.3. Growth of the glycine auxotrophic strain on glycine 65
Fig. S2.4. Effects of aTc and IPTG concentrations on the growth of Strain 7 66
Fig. S2.5. Growth of Strain 1 with L-serine or glycine 67
Fig. S2.6. Comparison of L-serine deaminases for L-serine assimilation 68
Fig. S2.7. Replacement of the genomic sdaA promoter (PsdaA) with IPTG-inducible $P_L lacO_1$ 69
Fig. S2.8. Screening electrode for the electro-cultivation system 70
2.7.1 Supplemental References 71
Chapter 3: Adaptive laboratory evolution for improved tolerance of isobutyl acetate in *Escherichia coli* 72
3.1 Introduction 74
Fig. 3.1. IBA Production from glucose and acetate in E. coli. 75
3.2 Methods 77
3.2.1. Reagents 77
3.2.2. Strains and Plasmids 77
3.2.3. Culturing 78
3.2.4. ALE for IBA tolerance 78
3.2.5. Genome sequencing and data analysis 79
3.2.6. Tolerance Assays 79

3.2.7. Fluorescence assays 80
3.2.8. Production 80
3.2.9. Gas Chromatography Analysis 80
3.2.10. Glucose and acetate analysis by High Performance Liquid Chromatography 81
3.3 Results and Discussion 81
3.3.1. Adaptive laboratory evolution and characterization of the evolved mutants 81
Fig. 3.2. Adaptive laboratory evolution to improve IBA tolerance. 83
3.3.2. Further characterization of evolved mutant M20 84
Fig. 3.3. ISO and IBA production from JCL260 and M20 under various conditions. 86
3.3.3. Whole genome sequencing and analysis of evolved strains 87
Fig. 3.4. Overview of mutations in M01-23. 88
3.3.4. CRISPR/Cas9 mediated M20 reversion 89
3.3.5. Construction and characterization of M01-07 mutations 89
Fig. 3.5. Effect of M01-M07 mutations on tolerance and IBA production. 91
3.3.6. Tolerance to other stressors 95
3.3.7. Construction and characterization of M08-23 mutations 95
Fig. 3.6. Toxic effects of various chemicals on JCL260 (left, blue) and the triple mutant metH, rho, ΔarcA (right, black). 96
Fig 3.7. Effect of M01-M23 mutations on IBA production. 98
3.4 Conclusion 99
3.5 References 99
3.6 Supplemental Information 107
Table S3.1: Genotype of E. coli strains in this study 107
Table S3.2: List of plasmids used in this study 109
Table S3.3: Oligonucleotides used in this study 110
Table S3.4: Plasmid construction by SLIC 117
Table S3.5: CRIPSR Donor DNA Fragments and Corresponding Primers 118
Table S3.6: Guide for CRISPR-Cas9 Construction by Q5 site-directed mutagenesis 122
Table S3.7: Details of ALE Mutations 123
Fig. S3.1. mRFP1 expression under $P_L lacO_1$ promoter in JCL260, M01 and M20. 124
Fig. S3.2. 125
Fig. S3.3. Schematic representation of IS30 insertion in yjjY. 126
Fig. S3.4. 127

Fig. S3.5. Effect of mutations on IBA production. 128
Fig. S3.6. Effect of mutations on IBA production in the metH, rho, ΔarcA triple mutant. 129
3.6.1 Supplemental References 130
Chapter 4: Ongoing Work By the Atsumi Lab and Conclusions 131
4.1 Ongoing Work to Optimize our Electrochemical-biological Hybrid System to Fix CO_2 132
4.2 Conclusion 136
4.3 References 137

Abstract

Engineered microorganisms capable of producing chemical commodities have emerged as sustainable alternatives to petroleum-based manufacturing. Expanded feedstock pools and improved product tolerance are two distinct approaches for improving the sustainability and productivity of microbial systems. Plant biomass sugars, which compete with food resources, are typically employed for most heterotrophic fermentation. Bypassing plants and directly utilizing carbon dioxide (CO_2) is a promising alternative. However, biological CO_2 fixation is slow and limits microbial growth and production. Regarding tolerance, engineering strategies and synthetic tools have improved the production of many target chemicals. Consequently, product concentrations often exceed the host tolerance threshold. Toxic products can curb titers and require expensive removal strategies, making economically viable microbial production challenging. This works seeks to address these issues by (i) construction of an electrochemical-biological hybrid system to fix CO_2 in *Escherichia coli* and (ii) application of adaptive laboratory evolution (ALE) for improved isobutyl acetate (IBA) tolerance in *E. coli*. The one pot hybrid system demonstrates the potential for chemical production from CO_2 and renewable electricity while the improved tolerance highlights the benefits overcoming product toxicity can have on microbial production.

Acknowledgements

The love, support, and encouragement I have received while in grad school is surely why I made it through with my sanity intact (although some might argue otherwise). My mother, my puzzle partner, and baking sous-chef has listened to my woes, weathered my rants, and empowered me time and time again. David, my partner of 10 years, has been a source of comfort and joy, which has only become more apparent through grad school. I'm glad we could get through our PhD programs together (albeit separately). But what is grad school without my lab? I can't count the number of times my lab mates have made me laugh, kept it real, and contributed to my ever-growing boba addiction. All of my undergrads have borne my inane ramblings and probably taught me a lot more about mentorship than I taught them about science. Shota's passion, mentorship style, and general calm approach to life has kept me afloat throughout grad school.

I am thankful for every interaction, good or bad, as they have helped me grow into a more well-rounded and thoughtful individual. I owe a lot to UC Davis, and I am truly grateful.

Chapter 1: Introduction

E. coli as a Platform for Biochemical Production

Portions of this chapter are reproduced with permission from:

Gonzales, J. N.; Matson, M. M.; Atsumi, S., Nonphotosynthetic Biological CO_2 Reduction. *Biochem* **2019**, 58,11,1470-1477.

1.1 Introduction: Microbial Production as an Alternative to Fossil Fuels

The current global dependence on fossil fuels for both energy and chemical production has spurred concerns regarding long-term resource security and resulted in devastating climate change. Over the last 10 years alone the fossil fuel industry has produced 35 billion metric tons of carbon dioxide (CO_2) annually.[1] CO_2 emissions, along with the numerous other pollutants released during the burning or processing of fossil fuels, have led to adverse environmental effects that negatively impact human health and disrupt our delicate ecosystems.[2,3] Finding alternatives to fossil fuels requires satisfying both energy and chemical needs while being carbon neutral or negative. To this end, microbes have emerged as renewable and sustainable production platforms.

Microbial production involves engineering a host organism to overproduce one or more chemicals of interest. This often involves installing synthetic pathways to allow the microbe to convert carbon feedstocks into native or non-native products. Carbon feedstocks can be composed of cheap biomass-derived sugars, lignocellulosic biomass, syn and waste gasses, atmospheric CO_2, or any combination of these, making them highly renewable in nature.[4] The immense variety of microbial hosts available and catalog of biosynthetic pathways developed enables a diverse array of products attainable by microbial production. Everything from valuable and complex bioactives, like pharmaceuticals[5], to versatile bulk chemicals, such as fuels, can be produced from microbes.[6] However, reaching industrial titers comparable to petroleum output requires significant optimization of the microbial host.

1.2 *E. coli* as a chassis for microbial production

Each microbial host has its own unique metabolism, feedstocks, and growth conditions. These factors determine which chemicals can be produced and at what titers. However, new

functions and phenotypes can be readily introduced into well-studied organisms with existing synthetic biology toolsets, allowing for diverse product offerings and high titers. One such organism is the extensively researched bacterium, *E. coli.* Its genetic malleability, rapid growth rate, well understood metabolism, and plethora of available genetic tools have kept *E. coli* at the center of microbial production.[7]

As stated above, having a wide breadth of information available is critical for genetic and phenotypic manipulation purposes. The National Center for Biotechnology Innovation (NCBI) houses the full sequences of more than 700 *E. coli* strains.[8] The Keio collection[9], an assemblage of in-frame, single gene deletion mutants, has helped elucidate gene function and genotype-phenotype relationships in *E. coli*. This data has been used extensively for constructing biosynthetic pathways, rewiring metabolic networks, altering regulation, improving cofactor supplies, and removing competitive pathways. These modifications are made possible by the unparalleled number of genetic tools developed for use in *E. coli*.

Since its discovery in 1855, *E. coli* has been continuously used as a test bed for probing novel functionalities and exploring new techniques. As a result of being at the forefront of research, *E. coli* has a constantly updating pipeline of new or improved tools. This is in stark contrast with less studied organisms where researchers must simultaneously develop tools while modifying their host.[10] This highlights an important advantage of *E. coli* as a production platform, where rapid strain engineering is possible because of its extensive library of tools.

Included in the toolsets available for *E. coli* are means to control gene expression and implement genomic modifications. Producing high titers while maintaining host performance requires tight control over gene expression levels and timing. This is facilitated by riboswitches[11], quorum sensing systems[12], and inducible promoters with variable expression strengths.[13] For

genomic modifications, methods with varying efficiencies and uses are available. The most mainstream of these being λ-red[14], multiplex automated genome engineering (MAGE)[15], and clustered regularly interspaced short palindromic repeats (CRISPR).[16] In particular, CRISPR has been useful for high efficiency genomic insertions, deletions, and point mutations. These modifications are critical for re-directing carbon flux, overcoming regulation, expediting transport, to mention a few. Recent innovations have led to new CRSIPR applications for controlling gene expression, namely CRISPRi[17] and CRISPRa.[18] Omics techniques, including genomics, metabolomics, and transcriptomics, have also been invaluable tools in elucidating the key weaknesses of biosynthetic pathways and inherent strain obstacles. All the methods mentioned above are only a small representation of the genetic tools available for modifying *E. coli*.

The deluge of available genetic information and tools has allowed researchers to construct biosynthetic pathways and optimize numerous *E. coli* strains to produce a diverse variety of chemicals. Some notable products include fuels, flavorants, fragrances, pharmaceuticals, nutraceuticals, plastics, and pigments. Many of these chemicals have moved beyond the laboratory bench onto commercial scale production. *E. coli* has been successfully used for the industrial production of amino acids, like Lysine[19], plastic polymers, such as 1,4 butandiol[20], and many more. These and other examples demonstrate the sustainable and lucrative potential of *E. coli* as production chassis. Although *E. coli* has seen considerable success as a microbial production platform, there remains many avenues of untapped potential for further exploration and innovation.

1.3 Enabling Non-photosynthetic CO_2 fixation through Electrocatalysis

In response to alarmingly high atmospheric CO_2 levels,[21,22] there has been an increased focus on investigating and engineering efficient means of converting atmospheric CO_2 into

industrially relevant fuels and commodities. As increasing population further strains the global food supply[23], it has also become critical to secure carbon feedstocks that do not require crop biomass or arable land. Traditional microbial production strategies rely entirely on these feedstocks, such as sugars or secondary metabolites, for their carbon and energy. CO_2 fixing microorganisms have the potential to produce valuable chemicals from atmospheric CO_2 without the need for biomass-derived feedstocks. The vast majority of global CO_2 fixation occurs in autotrophic photosynthetic organisms through the Calvin–Benson–Bassham cycle (CBB).[24] Photosynthetic organisms, such as cyanobacteria and algae, are promising chassis organisms for chemical production from atmospheric CO_2 but are not without their own challenges.[25-27] The cornerstone of CO_2 fixation, the enzyme ribulose-1,5 bisphosphate carboxylase/oxygenase (RuBisCO), has a notoriously low catalysis rate. RuBisCO does not readily differentiate between O_2 and CO_2, requiring energetically wasteful photorespiration processes to avoid accumulation of toxic intermediates.[28] Attempts to engineer RuBisCO to increase its conversion rate have been met with little success, creating a hurdle for photosynthetic organisms as production platforms.[29-31] Thus, there is a need for finding alternative, non-photosynthetic organisms that can utilize CO_2 as a carbon building block.

Electrocatalysis can be used in conjunction with a wide range of novel chemistries to reduce CO_2 into useful one-carbon molecules. Forming C–C bonds electrocatalytically remains a significant challenge.[32,33] Small carbon molecules, such as acetic acid, have been synthesized electrocatalytically, but at extremely low yields.[34] While synthesizing long chain carbon molecules and fuels is not feasible electrocatalytically, a biological–electrocatalytic hybrid system can capitalize on the innate ability of microorganisms to form C–C bonds. Electricity can enable

artificial photosynthesis in microorganisms by providing reducing power in the form of electrons to fix CO_2 into complex carbon commodities.[35]

The proportion of solar power converted to usable energy by photovoltaic technologies is significantly higher than that by plants or photosynthetic microorganisms.[36] Photovoltaic devices have a conversion efficiency of 18%, while photosynthetic microorganisms currently cap around 5–7% in optimized bioreactor conditions. An electrocatalytic–biological production platform may be a more sustainable and efficient method of energy processing than either method alone. Photovoltaic devices are more efficient at energy capture but storing that energy and incorporating it into the global power grid is challenging.[37,38] Organisms coupled to electrocatalysis have the potential to store this energy in the form of chemical bonds or use it to generate valuable commodities.[39-41]

While the utility of an electrocatalytic–biological production scheme is apparent, this integrated approach is relatively new and unexplored. The main difference between the few studies on this topic is the method used to deliver electrons to the host organism. There are several ways to deliver electricity to a production host: an organic molecule such as formate, ammonia, or H_2 can shuttle electrons, or they can be directly transferred to an electrophilic organism from the electron source.[41-44] Early studies utilized organic electron shuttles like neutral red, but issues such as toxicity, complicated product recovery, and cost have made them less favored.[41] Currently, research is focused on two methods: direct electrode-to-microbe electron transfer or indirectly with electrochemically generated electron donors.

The direct bioelectrochemical model requires the production host to be in immediate contact with the electrode, allowing electron transport conduit proteins to channel electrons needed to power reductive metabolism. Indirect microbial electrosynthesis employs an organic molecule

to deliver electrons from an electrode to the host, thereby powering reductive metabolism. The indirect production strategy has just recently begun to draw scientific and commercial interest owing to recent advancements in electrocatalysis and increased efficiency of solar devices. As such, most research is relatively new and applications are still in their infancy.

Formate is a favorable energy carrier because it is synthesized efficiently electrocatalytically and is readily soluble. Formate has a high redox potential compared to other potential donors, allowing it to support growth via the direct reduction of CO_2 to other organic compounds.[41] Compared to other potential electron donors like ammonia or sulfide, it does not require an additional electron acceptor like oxygen, which can add complexity to the production system. Formate does not have volatility safety concerns, such as cultures fed with H_2.

A formative study to the field reported that *Ralstonia eutropha* produced isobutanol and 3-methyl-1-butanol (3MB) from electrochemically fixed CO_2 [45] via formate and H_2 electron donors. When electrocatalytically synthesized formate accumulates at the anode, it decomposes over time, reducing substrate availability. To achieve maximum yields, formate must be consumed quickly after it is produced. To achieve this, a one-pot CO_2-to-alcohol scheme was designed to electrocatalytically synthesize formate in the *R. eutropha* culture producing mixed alcohols. Three major obstacles were identified in this system: (i) engineering *R. eutropha* to produce liquid fuels, (ii) efficient electrochemical synthesis of formate from CO_2 in fermentation medium, and (iii) enabling microbes to grow despite constant electrical current.

Previously established pathways were introduced to *R. eutropha* to produce isobutanol and 3MB.[46,47] An in-foil cathode and a Pt anode were used to produce formate electrocatalytically in 15% CO_2 media. Introducing electricity was previously shown to impede cell growth in *Escherichia coli*, presumably because of the generation of reactive oxygen or nitrogen species

(ROS and RNS, respectively).[48] To determine the presence of these species, three reporter plasmids known to be induced by hydrogen peroxide, superoxide free radicals, and nitric oxide were employed. Electrocatalytic *R. eutropha* cultures were shown to be high in nitrogen oxide and superoxide free radicals, while hydrogen peroxide remained at low concentrations. A porous ceramic cup was placed around the anode to provide a staunch diffusion gradient for electrochemically generated toxic intermediates. This shield ensures any reactive compounds produced are quenched before entering the fermentation media. The results showed that 140 mg L^{-1} of isobutanol and nearly 50 mg L^{-1} of 3MB were produced over 100 h. Final OD_{600} of *R. eutropha* reached 2.0, proving growth could be sustained despite exposure to constant low voltage.

While this study provided insight into conducting electrocatalysis and microbial synthesis in the same bioreactor concurrently, aerobic microbial electrosynthesis has a unique set of limitations that must be overcome. Aerobic cultures are at a disadvantage compared to anaerobic fermentation in terms of production efficiency.[49] While aerobic cultures are efficient at generating ATP compared to anaerobic cultures, they tend to utilize a large proportion of substrate to produce cell mass instead of products. An aerobic *R. eutropha* fermentation was supplied with H_2 and formic acid at 0 voltage. H_2 fermentation reached titers of ~536 mg L^{-1} isobutanol and ~520 mg L^{-1} of 3MB over 120 h at an OD_{600} of ~24.0. Formic acid fermentation reached titers of ~846 mg L^{-1} of isobutanol and ~570 mg L^{-1} of 3MB at an OD_{600} of ~3.5. Compared to H_2 and formic acid fed fermentations, the titer is much lower despite a relatively comparable OD_{600}, suggesting inhibited growth is not the only deleterious effect voltage has on the production hosts.

Another approach to microbial electrosynthesis involves a formate-dependent system which can be established in biological hosts capable of growing anaerobically, such as *E. coli*. Pyruvate, a key intermediate in central metabolism can be synthesized by a novel pathway termed

the reductive glycine pathway (RGP). The RGP is theoretically the most efficient route for formate assimilation to date.[50] The RGP has three major advantages: (i) it is a linear pathway, (ii) its reactions are thermodynamically favorable, (iii) no oxygen-sensitive pathways are involved. The RGP can be broken into three modules. In the first module, formate and CO_2 are converted into glycine, which in the second module is combined with another formate to produce l-serine. In the third and last module, l-serine is converted into the pyruvate. Two formate, one CO_2, three NADH, and two ATP molecules are required to make one pyruvate in this pathway. The RGP consumes five less ATP and two less NAD(P)H than what the CBB cycle requires to synthesize one pyruvate, making it a promising alternative to photosynthetic CO_2 fixation.

The product of the RGP, pyruvate, is an important central metabolite, but is not commercially valuable. However, it is a substrate for the production of many industrially relevant chemicals, easily linking the RGP with commercialization. Therefore, further engineering downstream of the RGP is required to produce any economically valuable compounds. Fortunately, there is a litany of industrially relevant, pyruvate-dependent metabolic pathways that have been developed in *E. coli* that can be applied; a comprehensive list can be found in other reviews.[51-53] The RGP is also a significant consumer of NADH, resulting in a scarcity of cellular reducing power that severely limits growth. If NADH could be regenerated or another reducing source supplemented, the production rate would likely increase as a result. Despite these limitations, the RGP is a promising candidate for efficient non-photosynthetic CO_2 fixation, which would enhance sustainable production in *E. coli*.

1.4 Alleviating Product Toxicity

While microbes have emerged as promising chassis for chemical production, as living creatures they are susceptible to cellular disruptions and toxins. Short-chain carboxylic acids, alcohols, and esters are natural fermentation products produced in small quantities for routine biological functions. These chemicals are also of significant commercial interest for their use as fuels, solvents, flavorants, and fragrances.[51] When these chemicals are overproduced for the purposes of industrial production, they often exert toxic effects on their microbial host. This toxicity is dependent on the physicochemical properties of the chemical, including carbon chain lengths, functional groups, chain branching features, hydrophobicity, and energy density.[54] The metabolism and physiology of the host also dictate the extent of product toxicity. Therefore, when choosing a target chemical to produce microbially, it is important to consider not only the value of that chemical, but its inherent toxicity and any burdens it might place on the host.[55] This is significant because if a host cannot tolerate the toxic effects, it will have reduced production potential, often exhibited by low titers and genetic instability.

Overcoming product toxicity often relies on understanding the affected cellular mechanisms and alleviating the impact of those mechanisms on the host. However, this requires thorough characterization for each chemical and extensive knowledge of both the genetics and physiology of the host. In some cases, installing or overexpressing transporters specific to the target chemical can be a direct solution for toxicity.[56] Unfortunately, the overexpression of membrane bound transporters, whether native or heterologous, can cause significant burden on the cell due to depleted protein resources and membrane saturation.[57] This approach, along with other rational design tactics, tend to be singularly focused and often result in an unbalanced system with its own inhibiting factors for production. This space is where evolution approaches excel.

Laboratory evolution falls under two distinct categories. The first is the directed modification of a specific gene for novel or improved functionality and the second is the directed culturing of a strain towards a desired phenotype. The former, commonly known as directed evolution, is helpful when there is an enzyme candidate for solving a known problem. For example, improving the kinase activity of a rate-limiting enzyme in the lycopene synthetic pathway to boost production.[58] In this case, the problem, the mevalonate kinase, is easily identifiable and the solution, improved activity, is discernable. But in the case of product toxicity, the problem and the solution are not well defined. Toxic effects can impact numerous, hard to elucidate cellular functions. Ergo, nearly the entirety of metabolism can be targeted for potential solutions. This is where the later strategy, dubbed ALE, is immensely beneficial.

Microorganisms have an incredible innate ability to adapt rapidly to dynamic environmental conditions. ALE capitalizes on this facet by using continuous culturing in a chosen environment to select for cells that exhibit increased fitness. In a sense, ALE enriches for beneficial mutations without any prior knowledge of how those mutations result in the desired phenotype.[59] Often, these mutations are non-intuitive or act synergistically in ways we never could have predicted. The major two applications for ALE are enhancing microbial growth on specific feedstocks or improving tolerance against certain chemicals.[60] An *E. coli* strain adapted to successfully grow on glycerol had mutations in the glycerol kinase gene that resulted in reaction rates 130% higher than wild-type.[61] Another strain, adapted for isobutanol-tolerance, saw 3-fold greater growth rates in otherwise toxic conditions to the wild-type.[62]

Although *E. coli* is arguably the most studied and well understood organism, there is so much more to explore and elucidate about the relationship between its genes, metabolism, and physiology. The mutations derived from ALE studies not only increase the commercial potential

of a particular biosynthetic system, but allow us to gain insight into cellular mechanisms that affect metabolism and cellular fitness that can be applied to other products and strains.

1.5 References

1 Friedlingstein, P. *et al.* Global Carbon Budget 2020. *Earth System Science Data* **12**, 3269-3340, doi:10.5194/essd-12-3269-2020 (2020).

2 Agency, U. S. E. P. Report on the Environment: Greenhouse Gases. (2020).

3 Zhang, W., Qian, C. N. & Zeng, Y. X. Air pollution: A smoking gun for cancer. *Chinese Journal of Cancer* **33**, 173-175, doi:10.5732/cjc.014.10034 (2014).

4 Rumbold, K. *et al.* Microbial renewable feedstock utilization: A substrate-oriented approach. *Bioengineered Bugs* **1**, 359-366, doi:10.4161/bbug.1.5.12389 (2010).

5 Pham, J. V. *et al.* A review of the microbial production of bioactive natural products and biologics. *Frontiers in Microbiology* **10**, doi:10.3389/fmicb.2019.01404 (2019).

6 Case, A. E. & Atsumi, S. Cyanobacterial chemical production. *Journal of Biotechnology* **231**, 106-114, doi:10.1016/j.jbiotec.2016.05.023 (2016).

7 Pontrelli, S. *et al. Escherichia coli* as a host for metabolic engineering. *Metab Eng* **50**, 16-46, doi:10.1016/j.ymben.2018.04.008 (2018).

8 Ncbi. Genome Assembly and Annotation report. (2016).

9 Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology* **2**, 1, doi:10.1038/msb4100050 (2006).

10 Carroll, A. L., Case, A. E., Zhang, A. & Atsumi, S. Metabolic engineering tools in model cyanobacteria. *Metabolic Engineering* **50**, 47-56, doi:10.1016/j.ymben.2018.03.014 (2018).

11 Scull, C. E., Dandpat, S. S., Romero, R. A. & Walter, N. G. Transcriptional Riboswitches Integrate Timescales for Bacterial Gene Expression Control. *Frontiers in Molecular Biosciences* **7**, doi:10.3389/fmolb.2020.607158 (2021).

12 Shong, J. & Collins, C. H. Quorum sensing-modulated and-Gate promoters control gene expression in response to a combination of endogenous and exogenous signals. *ACS Synthetic Biology* **3**, 238-246, doi:10.1021/sb4000965 (2014).

13 Yu, T. C. *et al.* Multiplexed characterization of rationally designed promoter architectures deconstructs combinatorial logic for IPTG-inducible systems. *Nature Communications* **12**, 325, doi:10.1038/s41467-020-20094-3 (2021).

14 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).

15 Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898, doi:10.1038/nature08187 (2009).

16 Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology* **20**, 490-507, doi:10.1038/s41580-019-0131-5 (2019).

17 Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8**, 2180-2196, doi:10.1038/nprot.2013.132 (2013).

18 Fontana, J. *et al.* Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nature Communications* **11**, 1618, doi:10.1038/s41467-020-15454-y (2020).

19 Kojima, H. O., Y., Kawamura, K. & Sano, K. Method of producing L-lysine by fermentation. USA patent (2000).

20 Burgard, A., Burk, M. J. & Osterhout, R. a. Development of a commercial scale process for production of 1,4-butanediol from sugar. *Current Opinion in Biotechnology* **42**, 118-125, doi:10.1016/j.copbio.2016.04.016 (2016).

21 May, H. D., Evans, P. J. & LaBelle, E. V. The bioelectrosynthesis of acetate. *Current Opinion in Biotechnology* **42**, 225-233, doi:10.1016/j.copbio.2016.09.004 (2016).

22 Sabine, C. L. *et al.* The oceanic sink for anthropogenic CO2. *Science* **305**, 367-371, doi:10.1126/science.1097403 (2004).

23 Godfray, H. C. J. *et al.* Food security: The challenge of feeding 9 billion people. *Science* **327**, 812-818, doi:10.1126/science.1185383 (2010).

24 Nickelsen, K. The Path of Carbon in Photosynthesis (1937–1954). *History, Philosophy and Theory of the Life Sciences* **8**, 201-250, doi:10.1007/978-94-017-9582-1_6 (2015).

25 Berla, B. M. *et al.* Synthetic biology of cyanobacteria: Unique challenges and opportunities. *Frontiers in Microbiology* **4**, doi:10.3389/fmicb.2013.00246 (2013).

26 Georgianna, D. R. & Mayfield, S. P. Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature* **488**, 329-335, doi:10.1038/nature11479 (2012).

27 Zhang, A., Carroll, A. L. & Atsumi, S. Carbon recycling by cyanobacteria: Improving CO2 fixation through chemical production. *FEMS Microbiology Letters* **364**, 16, doi:10.1093/femsle/fnx165 (2017).

28 Bauwe, H., Hagemann, M. & Fernie, A. R. Photorespiration: players, partners and origin. *Trends in Plant Science* **15**, 330-336, doi:10.1016/j.tplants.2010.03.006 (2010).

29 Greene, D. N., Whitney, S. M. & Matsumura, I. Artificially evolved Synechococcus PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochemical Journal* **404**, 517-524, doi:10.1042/BJ20070071 (2007).

30 Parikh, M. R., Greene, D. N., Woods, K. K. & Matsumura, I. Directed evolution of RuBisCO hypermorphs through genetic selection in engineered E.coli. *Protein Engineering, Design and Selection* **19**, 113-119, doi:10.1093/protein/gzj010 (2006).

31 Whitney, S. M., Houtz, R. L. & Alonso, H. Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. *Plant Physiology* **155**, 27-35, doi:10.1104/pp.110.164814 (2011).

32 Riplinger, C., Sampson, M. D., Ritzmann, A. M., Kubiak, C. P. & Carter, E. A. Mechanistic contrasts between manganese and rhenium bipyridine electrocatalysts for the reduction of carbon dioxide. *Journal of the American Chemical Society* **136**, 16285-16298, doi:10.1021/ja508192y (2014).

33 Taheri, A., Thompson, E. J., Fettinger, J. C. & Berben, L. A. An Iron Electrocatalyst for Selective Reduction of CO2 to Formate in Water: Including Thermochemical Insights. *ACS Catalysis* **5**, 7140-7151, doi:10.1021/acscatal.5b01708 (2015).

34 Genovese, C., Ampelli, C., Perathoner, S. & Centi, G. Mechanism of C-C bond formation in the electrocatalytic reduction of CO2 to acetic acid. A challenging reaction to use renewable energy with chemistry. *Green Chemistry* **19**, 2406-2415, doi:10.1039/c6gc03422e (2017).

35 Rabaey, K. & Rozendal, R. A. Microbial electrosynthesis - Revisiting the electrical route for microbial production. *Nature Reviews Microbiology* **8**, 706-716, doi:10.1038/nrmicro2422 (2010).

36 Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* **332**, 805-809, doi:10.1126/science.1200165 (2011).

37 Dunn, B., Kamath, H. & Tarascon, J. M. Electrical energy storage for the grid: A battery of choices. *Science* **334**, 928-935, doi:10.1126/science.1212741 (2011).

38 Soloveichik, G. L. Battery technologies for large-scale stationary energy storage. *Annual Review of Chemical and Biomolecular Engineering* **2**, 503-527, doi:10.1146/annurev-chembioeng-061010-114116 (2011).

39 Lewis, N. S. & Nocera, D. G. Powering the planet: Chemical challenges in solar energy utilization. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15729-15735, doi:10.1073/pnas.0603395103 (2006).

40 Lovley, D. R. & Nevin, K. P. Electrobiocommodities: Powering microbial production of fuels and commodity chemicals from carbon dioxide with electricity. *Current Opinion in Biotechnology* **24**, 385-390, doi:10.1016/j.copbio.2013.02.012 (2013).

41 Tremblay, P. L. & Zhang, T. Electrifying microbes for the production of chemicals. *Frontiers in Microbiology* **6**, 201, doi:10.3389/fmicb.2015.00201 (2015).

42 Ross, D. E., Flynn, J. M., Baron, D. B., Gralnick, J. A. & Bond, D. R. Towards electrosynthesis in Shewanella: Energetics of reversing the Mtr pathway for reductive metabolism. *PLoS ONE* **6**, 2, doi:10.1371/journal.pone.0016649 (2011).

43 Tashiro, Y., Hirano, S., Matson, M. M., Atsumi, S. & Kondo, A. Electrical-biological hybrid system for CO2 reduction. *Metabolic Engineering* **47**, 211-218, doi:10.1016/j.ymben.2018.03.015 (2018).

44 Khunjar, W. O., Sahin, A., West, A. C., Chandran, K. & Banta, S. Biomass Production from Electricity Using Ammonia as an Electron Carrier in a Reverse Microbial Fuel Cell. *PLoS ONE* **7**, 9, doi:10.1371/journal.pone.0044846 (2012).

45 Li, H. *et al.* Integrated electromicrobial conversion of CO2 to higher alcohols. *Science* **335**, 1596, doi:10.1126/science.1217643 (2012).

46 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).

47 Atsumi, S. *et al.* Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Microbiology and Biotechnology* **85**, 651-657, doi:10.1007/s00253-009-2085-6 (2010).

48 Jeong, J., Kim, J. Y. & Yoon, J. The role of reactive oxygen species in the electrochemical inactivation of microorganisms. *Environmental Science and Technology* **40**, 6117-6122, doi:10.1021/es0604313 (2006).

49 Huang, W. D. a. Analysis of biofuels production from sugar based on three criteria: Thermodynamics, bioenergetics, and product separation. *Energy and Environmental Science* **4**, 784-792, doi:10.1039/c0ee00069h (2011).

50 Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochimica et Biophysica Acta - Bioenergetics* **1827**, 1039-1047, doi:10.1016/j.bbabio.2012.10.013 (2013).

51 Carroll, A. L., Desai, S. H. & Atsumi, S. Microbial production of scent and flavor compounds. *Current Opinion in Biotechnology* **37**, 8-15, doi:10.1016/j.copbio.2015.09.003 (2016).

52 Peralta-Yahya, P. P. & Keasling, J. D. Advanced biofuel production in microbes. *Biotechnology Journal* **5**, 147-162, doi:10.1002/biot.200900220 (2010).

53 Pfleger, B. F., Gossing, M. & Nielsen, J. Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering* **29**, 1-11, doi:10.1016/j.ymben.2015.01.009 (2015).

54 Wilbanks, B. & Trinh, C. T. Comprehensive characterization of toxicity of fermentative metabolites on microbial growth Mike Himmel. *Biotechnology for Biofuels* **10**, 262, doi:10.1186/s13068-017-0952-4 (2017).

55 Ko, Y. S. *et al.* Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chemical Society Reviews* **49**, 4615-4636, doi:10.1039/d0cs00155d (2020).

56 Dunlop, M. J. *et al.* Engineering microbial biofuel tolerance and export using efflux pumps. *Molecular Systems Biology* **7**, 784, doi:10.1038/msb.2011.21 (2011).

57 Mukhopadhyay, A. Tolerance engineering in bacteria for the production of advanced biofuels and chemicals. *Trends in Microbiology* **23**, 498-508, doi:10.1016/j.tim.2015.04.008 (2015).

58 Chen, H. *et al.* Directed evolution of mevalonate kinase in: Escherichia coli by random mutagenesis for improved lycopene. *RSC Advances* **8**, 15021-15028, doi:10.1039/c8ra01783b (2018).

59 Portnoy, V. A., Bezdan, D. & Zengler, K. Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* **22**, 590-594, doi:10.1016/j.copbio.2011.03.007 (2011).

60 Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution - principles and applications for biotechnology. *Microb Cell Fact* **12**, 64, doi:10.1186/1475-2859-12-64 (2013).

61 Herring, C. D. *et al.* Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. *Nature Genetics* **38**, 1406-1412, doi:10.1038/ng1906 (2006).

62 Atsumi, S. *et al.* Evolution, genomic analysis, and reconstruction of isobutanol tolerance in Escherichia coli. *Mol Syst Biol* **6**, 449, doi:10.1038/msb.2010.98 (2010).

Chapter 2: Electrical-biological hybrid system for CO_2 reduction

This chapter is reproduced with permission from:

Tashiro, Y.; Hirano, S.; Matson, M. M.; Atsumi, S.; Kondo, A., Electrical-biological hybrid system for CO_2 reduction. *Metabolic Engineering.* **2018,** 47, 211-218.

2.0 Abstract

Here we have developed an electrochemical-biological hybrid system to fix CO_2. Natural biological CO_2 fixation processes are relatively slow. To increase the speed of fixation we applied electrocatalysts to reduce CO_2 to formate. We chose a user-friendly organism, *Escherichia coli*, as host. Overall, the newly constructed CO_2 and formate fixation pathway converts two formate and one CO_2 to one pyruvate via glycine and L-serine in *E. coli*. First, one formate and one CO_2 are converted to one glycine. Second, L-serine is produced from one glycine and one formate. Lastly, L-serine is converted to pyruvate. *E. coli*'s genetic tractability allowed us to balance various parameters of the pathway. The carbon flux of the pathway was sufficient to compensate L-serine auxotrophy in the strain. In total, we integrated both electrocatalysis and biological systems into a single pot to support *E. coli* growth with CO_2 and electricity. Results show promise for using this hybrid system for chemical production from CO_2 and electricity.

2.1 Introduction

Global energy and environmental problems have stimulated increased efforts towards synthesizing chemicals from renewable resources, reducing greenhouse gas emissions, and utilizing atmospheric CO_2. Employing CO_2 as a substrate creates either a carbon sink or closed loop production system, both of which have numerous environmental and industrial benefits. Biological production systems have been developed to sequester CO_2 indirectly using microbial fermentation with plant biomass-based sugars (e.g. corn ethanol)[63] or directly by photosynthetic microorganisms (e.g. algae and cyanobacteria).[6,64] However, such biological systems need to overcome many issues including low areal productivity, massive water usage, and competition with food production. Biological systems using autotrophic organisms have resulted in moderate improvements to achieve chemical production from CO_2.[27,65] Inorganic methods such as electrocatalysts have been developed to reduce CO_2 to useful one carbon molecules by utilizing a wide range of novel chemistries. However, these electrocatalysts are unable to form carbon-carbon bonds.[32,33,66] Combining biological and electrochemical systems into one approach would help overcome their individual challenges. Here, in the common industrial organism *Escherichia coli*, we designed and constructed a biosynthetic pathway integrated with electrocatalysis to synergistically improve CO_2 fixation. This integration provides a fundamentally new approach for carbon fixation to produce valuable chemical products.

Recently a few electro-biochemical hybrid production systems have been developed with the potential to be more effective compared to their traditional biological counterparts.[45,67,68] A pivotal study was reported by Li et al. (2012) wherein the chemolithotrophic bacterium, *Ralstonia eutropha*, was used to produce isobutanol from electrochemically reduced CO_2 via formate. While the work provided clear proof-of-concept, advancement of hybrid systems has been limited due to

inadequate tools and scarce knowledge of readily engineered host organisms. This is also compounded by *R. eutropha* being an obligate aerobe. Oxygen is a strong electron acceptor so it inhibits the reduction of CO_2 and during electrochemical reactions can be converted to various bacterial toxins such as hydrogen peroxide (H_2O_2).[45] To avoid these issues *E. coli* was chosen as the host for the hybrid system due to its anaerobic metabolic capabilities and native possession of many required enzymes.

Biological CO_2 fixation systems typically use autotrophs, mainly photoautotrophs, to convert CO_2 and light energy into useful chemicals such as fuel, bulk feedstocks, and pharmaceuticals[6,69], but none have achieved commercial-level titers. This is in part due to ribulose bisphosphate carboxylase-oxygenase (RuBisCO), the major autotrophic carboxylase by which > 99.5% of global inorganic carbons are fixed.[70] Although attempts have been made for decades to increase the rate and efficiency of this key carbon fixation enzyme, improvements have been marginal and do not significantly increase the overall rate of carbon fixation.[31,71] To overcome this limitation we explored a RuBisCO-independent CO_2 fixation pathway in *E. coli*.

The constructed pathway from CO_2 and formate to pyruvate was named the reductive glycine pathway (RGP) **(Fig. 2.1)**. It has been suggested that RGP is the most theoretically efficient route for formate assimilation, although RGP has not been constructed or discovered in any organisms.[50] RGP has the following advantages as a carbon fixation pathway: 1) it is a linear pathway, which makes improvements easier compared to circular pathways such as the Calvin-Benson (CB) cycle, 2) its reactions are thermodynamically favorable[50], one of the most important factors in efficiency, and 3) it doesn't utilize oxygen-sensitive enzymes, which represents a robust pathway functional in both aerobic and anaerobic conditions. The main challenge of RGP is its requirement of two different carbon sources, CO_2 and formate. This is why the electrocatalysis

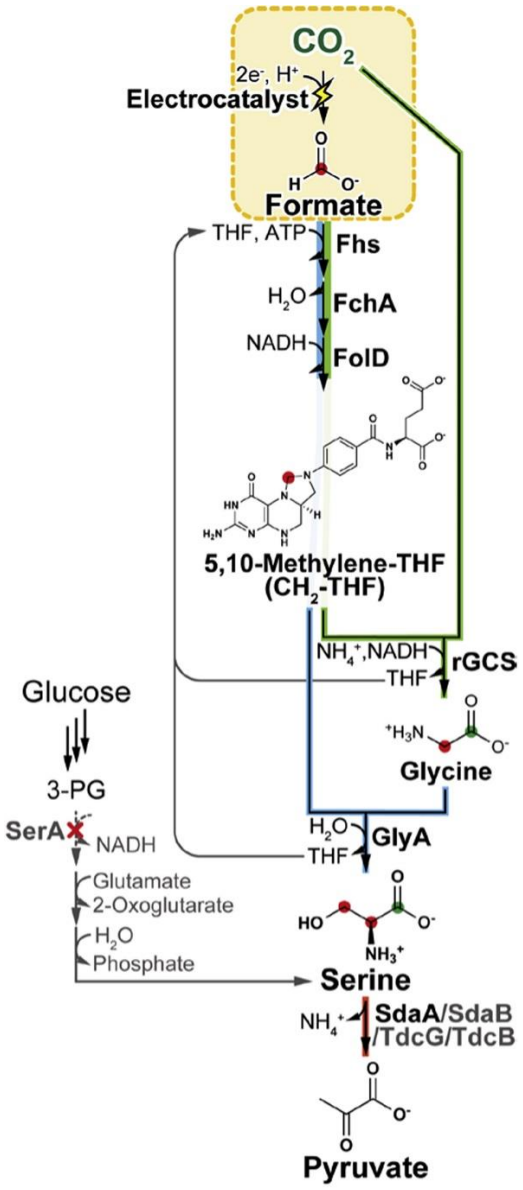


Fig. 2.1. Electro-biochemical CO_2 conversion in this study.

CO_2 is first converted to formate by an electrocatalyst. Formate is then assimilated through the RGP, which is comprised of 3 modules: 1) Formate, via part of the Wood-Ljungdahal pathway through a CH_2-THF intermediate, is combined with another CO_2 by rGCS to synthesize glycine. 2) Formate is again metabolized to CH_2-THF via part of the WLP and is combined with glycine (produced by module 1) by the SGC enzyme GlyA to produce L-serine. 3) Lastly, serine is irreversibly deaminated to pyruvateline).

aspect is critical, as it converts CO_2 to formate, eliminating the need for two carbon sources. Thus, in our hybrid system only CO_2 is fed into the system.

To assimilate formate that is electro-chemically generated from CO_2, exogenous formate-tetrahydrofolate ligase (Fhs)[72] was utilized since *E. coli* lacks a native Fhs. To fix CO_2 the native glycine cleavage system (GCS)[73] was installed. This protein complex contains glycine dehydrogenase (GcvP), aminomethyltransferase (GcvT), lipoic acid-containing protein (GcvH), and lipoamide dehydrogenase (Lpd). GCS catalyzes the cleavage of glycine to 5,10-methylene-THF (CH_2-THF) and CO_2 **(Fig. 2.1)**.[73] We hypothesize that this reaction is reversible, allowing the hybrid to utilize CO_2 to produce glycine, a reversal of GCS (rGCS).

The overall hybrid RGP consists of three modules **(Fig. 2.1)**. The first module is glycine synthesis from formate and CO_2 via a CH_2-THF intermediate **(Fig. 2.1, green line)**. This module utilizes rGCS and a segment of the Wood-Ljungdahal pathway (WLP).[65,74] The second module is L-serine synthesis from CH_2-THF and glycine **(Fig. 2.1, blue line)**. In this module, formate is again converted to CH_2-THF, which is then coupled with glycine to synthesize L-serine by an L-serine hydroxymethyltransferase (GlyA) in the Serine-Glycine cycle (SGC).[75] Recently, Yishai et al. (2017) have shown that the SGC, including GlyA, is functional in *E. coli*.[76] The third module is the irreversible L-serine deamination to produce pyruvate **(Fig. 2.1, red line)**. Overall, one molecule of pyruvate is synthesized from two formate, one CO_2, three NAD(P)H, and two ATP molecules. Thus, RGP requires less ATP and NAD(P)H to produce one pyruvate from CO_2 than the CB cycle, which requires five NADPH molecules and seven ATP molecules to produce one pyruvate.[77,78] This is in part due to formate, a reduced form of CO_2, being more readily fixed than CO_2 thereby decreasing the required energy for carbon fixation.

Here, we present an electrochemical-biological hybrid approach to convert CO_2 to the central metabolite pyruvate. In particular, we employed rGCS for biological CO_2 fixation to avoid the utilization of RuBisCO. We integrated both electrocatalyst and biological systems into a single pot to investigate whether they are compatible with each other. This integration would improve energy efficiency for CO_2 fixation, which is a bottleneck in current biochemical production. In this study we used the engineered system to support *E. coli* growth, but because the product of this approach is a central metabolite, this system could be combined with almost any biochemical production pathway to produce a wide range of chemicals from CO_2 and electricity. As such, this system would be more renewable than those based on fast-growing heterotrophic hosts that use sugar precursors, and more efficient than those based on (photo)autotrophic hosts that use RuBisCO.

2.2 Methods

2.2.1 Reagents

All enzymes were purchased from New England Biolabs. All synthetic oligonucleotides were ordered from Eurofins Genomics. ^{13}C-labeled sodium hydrogen bicarbonate and ^{13}C-labeled sodium formate were purchased from Cambridge Isotope Laboratories Inc. Unlabeled sodium hydrogen carbonate was purchased from Wako Pure Chemical Industries, Ltd. The other chemicals were purchased from Nacalai Tesque, Inc.

2.2.2 Strains and plasmids

All strains in this study are listed on Table S2.1. Gene deletion was carried out as described in Datsenko and Wanner (2000).[14] To remove kan^R cassette, a plasmid, 708-FLPe (Gene Bridge GmbH), was used. YT276 was constructed from YT151 with replacement of the natural promoter for *sdaA* (*PsdaA*) with $P_L lacO_1$[79] by using CRISPR-Cas9 based method.[80] All modifications were confirmed with PCR and sequencing.

All plasmids used in this study are listed on Table S2.2. Oligonucleotides used in this study are listed in Table S2.3. Heterogeneous genes were codon-optimized for *E. coli* and synthesized by Thermo Fisher Scientific. The synthesized DNA sequences are listed in Table S2.4. Table S2.5 shows a set of primers and the template used for plasmid construction. The constructed plasmids were verified by sequencing.

2.2.3 Culture conditions

Overnight cultures were prepared in 3 mL of Luria-Bertani (LB) media containing appropriate antibiotics. Antibiotic concentrations were as follows: tetracycline (5 μg mL^{-1}), gentamicin (5 μg mL^{-1}), kanamycin (25 μg mL^{-1}), ampicillin (100 μg mL^{-1}).

Growth assay was carried out in M9 minimal medium (33.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 9.4 mM NH_4Cl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$) including 1000 × A5 trace metal mix (2.86 g H_3BO_3, 1.81 g $MnCl_2 \cdot 4H_2O$, 0.079 g $CuSO_4 \cdot 5H_2O$, 49.4 mg $Co(NO_3)_2 \cdot 6H_2O$ per liter water) with appropriate carbon sources.

Overnight culture in LB media was spun down at 6000*g* for 1 min and supernatant was discarded. Subsequently, cell pellet was re-suspended with M9 minimal media with or without 4 g L^{-1} glucose. OD_{600} of the solution was adjusted to ~2. The solution was inoculated 1% in 5 mL

M9 minimal media with appropriate a carbon source and additives. Cells were incubated at 37 °C. Cell growth was monitored by measuring OD_{600} in a Synergy HTX Plate Reader (BioTek Instruments, Inc.).

2.2.4 Gas chromatography (GC) sample preparation

Cell pellet samples were re-suspended and filtered with a 0.5 μm filter (Toyo Roshi Kaisha, Ltd) by vacuum filtration. The cell sample on a membrane was dipped in 500 μL methanol. After mixing, 500 μL of chloroform and 200 μL of deionized water were added in order. The sample was centrifuged at 4 °C and 17,000*g* for 5 min. The supernatant was taken and stored at −80 °C. The extract was dried at 30 °C in FreeZone2.5 (Labconco) for overnight. The dried sample was dissolved in 50 μL of *N*,*N*-dimethylformamide. For the GC-MS analysis, the sample was derivatized with 50 μL (*N*-*tert*-butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA) + 1% *N*-methyl-*N*-(tert-butyldimethylsilyl)trifluoroacetamide (TBDMS) (Thermo Scientific) at 80 °C for 60 min.

2.2.5 GC- mass spectrometry (MS) analysis

The produced metabolites were quantified by GC-MS. The system consisted of GCMS-QP2010 Ultra (Shimadzu) equipped with a CP-Sil 8 CB-MS capillary column (30 m length, 0.25-mm diameter, 0.25-μm film thickness) (Agilent). Injector temperature was held at 230 °C. Injection volume was 1 μL, injected at a 10:1 split ratio. Helium was used as the carrier gas. GC oven temperature was held 150 °C for 5 min, then increased at a rate of 10 °C per min until 300 °C and held for 5 min. The ion source temperature and interface temperature were 200 °C. Solvent cut

time was 4.5 min for the solvent. Detector voltage was approximately 0.1 kV. Mass spectra and retention time for samples were compared with external standards.

2.2.6 Electrochemical cultivation

All experiments for electrochemical cultivation of *E. coli* were performed using a bioelectrochemical reactor (BER) comprising a bottle-type single chamber as previously described87 **(Fig. 2.5)**. A three-electrode system, which included a working electrode (cathode), a reference electrode, and a counter electrode (anode), was used. An indium sheet (IN-203321 (Nilaco)) was used as the working electrode (75 × 25 × 0.1 mm) and Pt mesh (PT-358080 (Nilaco)) was used as the counter electrode (50 × 10 × 0.08 mm, 80 mesh $inch^{-1}$). The counter electrode was inserted in the anodic bag formed by a proton exchange membrane (Nafion 117 (DuPont)). The working electrode, the anodic bag, and an Ag/AgCl reference electrode (HS-205C (DKK-TOA)) were inserted into the cathodic chamber. The cathodic chamber and the anodic bag were filled with 250 mL of M9 medium and 5 mL of 100 mM NaCl, respectively. In the BER, three electrodes were connected with a potentiostat (PS-08 (Tohogiken)), and the potential of the working electrode was electrochemically regulated to −1.2 or −2.0 V (vs. Ag/AgCl). All reported potentials on the working electrode pertain to the Ag/AgCl reference electrode (type: saturated KCl). A gas inlet line was set on the side arm of cathodic chamber to supply 100% CO_2 (v/v) by sparging, and an outlet line was connected to cathodic chamber via a silicon cap on the top of the reactor. *E. coli* cells were inoculated in the BER at OD_{600} 0.02 and then incubated at 37 °C with 300 rpm stirring.

Supplied electrons (not shown) from the working electrode was calculated as electric current (A) × time (s)/ Faraday constant (9.45×10^4 C mol^{-1}). If all electrons are used for formate

production, theoretically produced formate amounts to half mol of supplied electrons, since 2 mol of electrons is needed for the production of 1 mol of formate. Electrochemical conversion efficiency of formate was calculated as measured formate in working chamber / theoretical amounts of formate production (taken from the supplied electrons calculation) **(Table S2.6)**.

To optimize the electrochemical potential on the working electrode (indium) for electrochemical cultivation was examined with applied potential at −1.6, −1.8, and −2.0 V **(Table S2.7)**. Same apparatus described above was used except basal medium (L^{-1}) including 0.5 g NH_4Cl, 0.4 g H_2HPO_4, 0.49 g $MgCl_2 \cdot 6H_2O$, 0.05 g $CaCl_2 \cdot 2H_2O$, 0.05 g KCl, 4.2 g $NaHCO_3$, and 20 mL 1 M phosphate buffer was fulfilled instead of 50 mM $NaHCO_3$ solution. The control experiment without electrolysis were also operated in the reactor without electrodes.

2.3 Results

2.3.1. Transferring the formate-assimilation pathway into *E. coli*

In the RGP, formate is converted to CH_2-THF (Fig. 2.1**,** blue line, upper segment). This conversion is catalyzed by three enzymes, Fhs[72], FchA[81], and FolD[82,83] of the WLP[74]. CH_2-THF is utilized as a building block for both glycine and L-serine syntheses. We chose to synthesize and test the well characterized *fhs*, *fchA*, and *folD* genes from *C. ljungdahlii*[84] and *Acetobacterium woodii*[85] **(Table S2.4)**. To evaluate formate-assimilation activity we constructed screening strains by deleting two genes encoding L-serine hydroxymethyl transferase (*serA*) and glycine decarboxylase (*gcvP*) from the parent strain, Strain 1 **(Table 2.1)**. SerA converts 3-phospho-D-glycerate to 3-phosho-hydroxypyruvate in the L-serine biosynthesis pathway **(Fig. 2.1)**. The *serA*-knockout strain **(Table 2.1; Strain 2)** is

auxotrophic for L-serine **(Fig. S2.1a)**. Glycine rescued the growth of Strain 2 **(Fig. S2.1b)** since glycine can be converted to L-serine by the GCS and GlyA **(Fig. 2.1)**. When the *gcvP* gene was deleted in Strain 2 **(Table 2.1; Strain 3)** glycine was unable to rescue the growth of Strain 3 **(Fig. S2.1)**.

The operons of the WLP segments in *C. ljungdahlii* and *A. woodii* (Cl operon and Aw operon, respectively) were separately introduced into Strain 3, and the growth of these strains was compared without and with formate **(Fig. 2.2)**. The growth of the strain with the Cl operon was rescued with formate, but the strain with the Aw operon did not grow in the presence of formate. Next, to confirm whether all three genes of the Cl operon were necessary for function different combinations of the three Cl operon genes were tested **(Fig. 2.2)**. Wild-type *E. coli* has the *ads* gene encoding a bifunctional enzyme having the activities of FchA and FolD[86], but not *fhs*. Formate rescued the growth of the strain with *fhs-folD,* but the growth was much slower than that of the strain with all three genes. Thus, the three genes construct was used for following experiments. To find the optimum concentration of formate three different formate concentrations (0.02, 0.2, and 2.0 $g\,L^{-1}$) were tested on the growth of Strain 3 with the Cl operon **(Fig. S2.2a)**. The best growth was observed with 0.2 $g\,L^{-1}$ formate, therefore we used that concentration for further studies.

Table 2.1. List of strains used in this study.

Strain No.	Host	Gene knockout	Plasmid	Plasmid Contents
1	JCL16	–	–	–
2	YT151	Δ*serA*	–	–
3	YT166	Δ*serA*Δ*gcvP*	–	–
4	YT152	Δ*glyA*	–	–
5	YT152	Δ*glyA*	pYT100, pZE12MCS	$P_L$tetO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*, $P_L$lacO$_1$: -
6	YT152	Δ*glyA*	pYT100, pYT048	$P_L$tetO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*, $P_L$lacO$_1$: *gcvTHP*
7	YT151	Δ*serA*	pYT100, pYT048	$P_L$tetO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*, $P_L$lacO$_1$: *gcvTHP*
8	YT276	Δ*serA*, P_{sdaA}*::*$P_L$lacO$_1$	–	–
9	YT276	Δ*serA*, P_{sdaA}*::*$P_L$lacO$_1$	pYT100, pYT048	$P_L$tetO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*, $P_L$lacO$_1$: *gcvTHP*

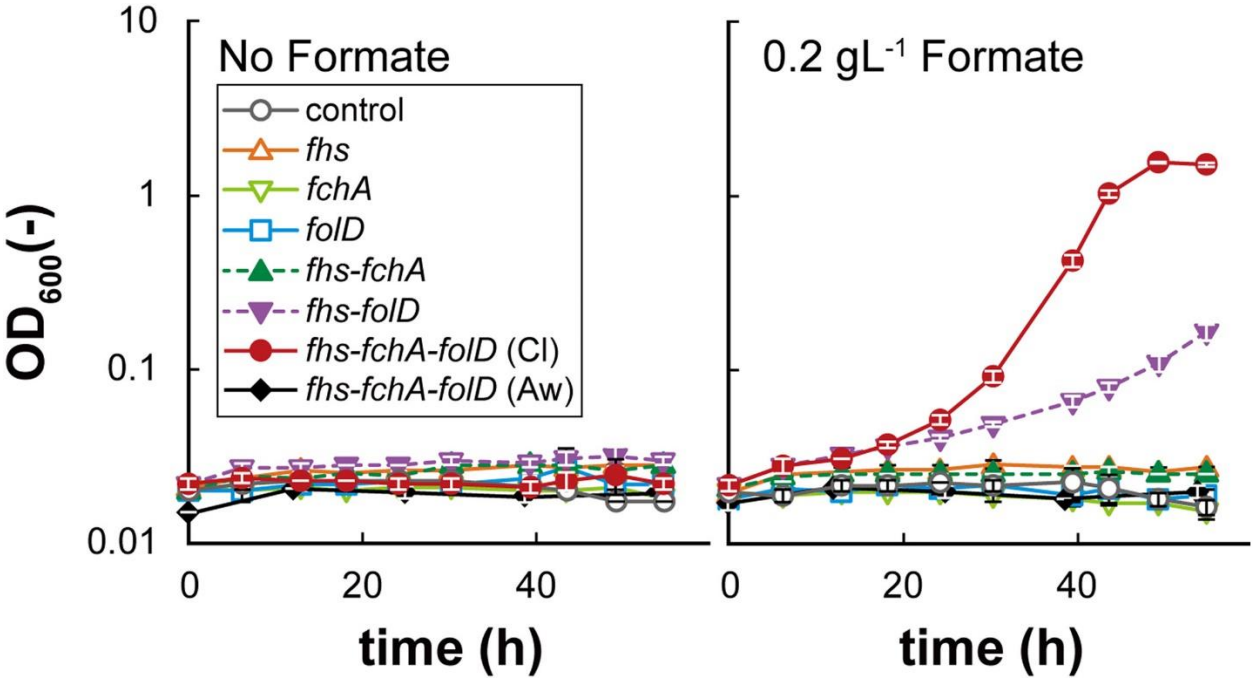


Fig. 2.2. Transfer the pathway from formate to CH_2-THF to E. coli.

The Cl and Aw operon genes were introduced to the L-serine auxotroph strain, Strain 3 (Table 2.1). Cells were cultured in 5 mL of M9 minimal media containing 4 g L^{-1} of glucose without (left) and with (right) 0.2 g L^{-1} formate at 37 °C. Open circles, Strain 3 with empty vector (control); open triangles, Strain 3 with fhs_{Cl}; open down triangles, Strain 3 with $fchA_{Cl}$; squares, Strain 3 with $folD_{Cl}$; closed triangle, Strain 3 with fhs_{Cl}-$fchA_{Cl}$; closed down triangles, Strain 3 with fhs_{Cl}-$folD_{Cl}$; closed circles, Strain 3 with fhs_{Cl}-$fchA_{Cl}$-$folD_{Cl}$; diamonds, Strain 3 with fhs_{Aw}-$fchA_{Aw}$-$folD_{Aw}$. N = 3 biological replicates; error bars represent s.d.

2.3.2 Engineering RGP in glycine auxotroph

GCS catalyzes the reaction from glycine into CO_2, NH_4^+, and CH_2-THF.78 GCS is conserved in a wide range of bacteria, archaea, and plants.[87] In amino acid- and purine-degrading microbes, GCS has carbon exchange activities between glycine and CO_2.[88] To test whether *E. coli* GCS is functional as rGCS the *glyA* gene was deleted from Strain 1 **(Table 2.1; Strain 4)**. Strain 4 is auxotrophic for glycine and growth is rescued by glycine supplementation **(Fig. S2.3)**.

The Cl operon was introduced into Strain 4 **(Table 2.1; Strain 5)**. Strain 5 was cultured with four different formate/bicarbonate conditions **(Fig. 2.3a)**. In the presence of both formate and bicarbonate Strain 5 had the fastest growth. Formate without bicarbonate also slightly rescued the growth of Strain 5, which is likely due to atmospheric CO_2 or CO_2 produced from native *E. coli* decarboxylation reactions. In Strain 5, the endogenous GCS operon containing *gcvTHP* is expressed from the genome. To increase the expression of *gcvTHP* the GCS operon was cloned into a high-copy plasmid **(Table S2.2; pYT048)** which was introduced into Strain 3 with the Cl operon **(Table 2.1; Strain 6)**. Strain 6 grew significantly faster than Strain 5 with both formate and bicarbonate supplementation **(Fig. 2.3b)**.

To confirm that formate and bicarbonate are incorporated to form glycine, ^{13}C-bicarbonate and/or ^{13}C-formate were fed to the strain and the synthesized glycine was analyzed with GC-MS. Position 1 carbon and position 2 carbon in glycine came from CO_2 and formate, respectively **(Fig. 2.3c)**.[73] Glycine synthesized from ^{13}C-labeled bicarbonate or formate has one higher *m/z* value, which indicated incorporation of the labeled substrate. When unlabeled bicarbonate was used M0, M1, and M2 accounted for 76.2%, 16.8%, and 7.0%, respectively **(Fig. 2.3d)**. These values were very similar to data from the NIST library standards (0; 75.7%, 1; 17.2%, and 2; 7.0%). M1 and M2 would be caused by two silicon atoms, which has relatively abundant isotopes (^{29}Si (4.7%)

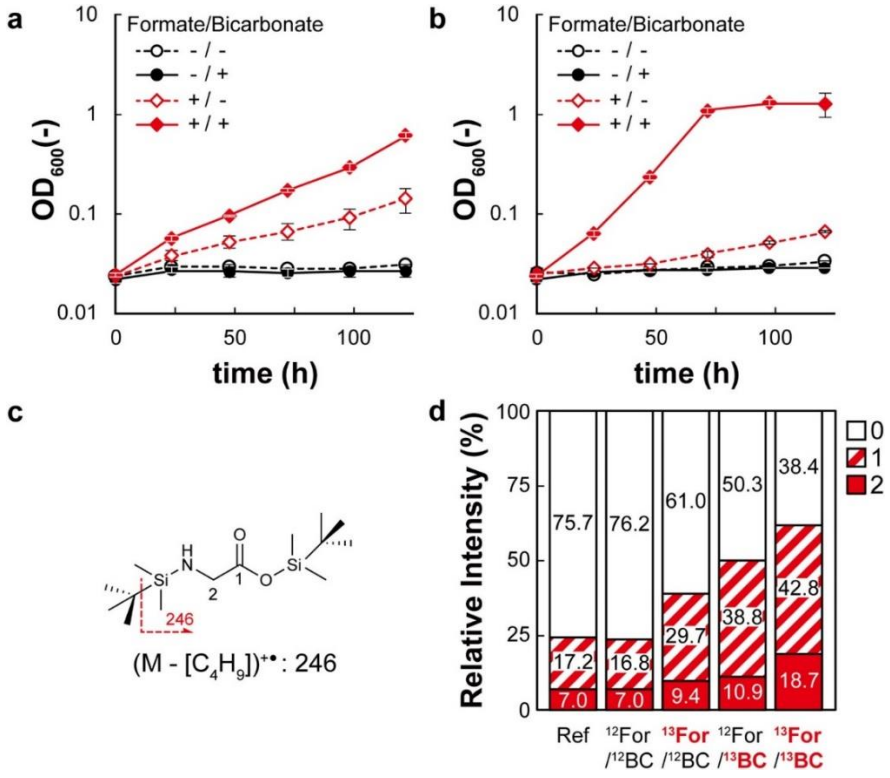


Fig. 2.3. Glycine production from CO_2 using the reverse reaction of GCS.

The Cl operon was introduced to the glycine auxotroph strain, Strain 4 (termed Strain 5 (Table 2.1)). The plasmid containing the E. coli GCS genes (pYT048) was introduced into Strain 5 (termed Strain 6 (Table 2.1)). Cells were cultured in 5 mL of M9 minimal media containing 4 g L^{-1} of glucose without and with 0.2 g L^{-1} formate and 50 mM bicarbonate at 37 °C. (a) Growth of Strain 5 and (b) Strain 6. N = 3 biological replicates; error bars represent s.d. (c) Structure of glycine derivatized with MTBSTFA. If the pathway is active, position 1 and position 2 would be derived from CO_2 and formate, respectively. (d) Relative intensity of m/z = 246 (M0), 247 (M1), and 248 (M2) in the glycine produced in Strain 6. Cells were cultured using unlabeled formate and bicarbonate (^{12}For/^{12}BC), ^{13}C formate and unlabeled bicarbonate (^{13}For/^{12}BC), unlabeled formate and ^{13}C bicarbonate (^{12}For/^{13}BC), or ^{13}C formate and ^{13}C bicarbonate (^{13}For/^{13}BC). Ref indicates data from the NIST library of standards.

and ^{30}Si (3.1%)), in the derivatized molecule **(Fig. 2.3c)**. In ^{13}C-labeled formate or bicarbonate conditions M1 was up to 29.7% or 38.8%, respectively. Existence ratio of M1 was significantly increased by fed ^{13}C-labeled formate and bicarbonate. When both ^{13}C-bicarbonate and ^{13}C-formate were fed M2 was more than twice as high as unlabeled condition. These results indicated that the constructed pathway was functional as designed.

To confirm that formate and bicarbonate are incorporated to form glycine, ^{13}C-bicarbonate and/or ^{13}C-formate were fed to the strain and the synthesized glycine was analyzed with GC-MS. Position 1 carbon and position 2 carbon in glycine came from CO_2 and formate, respectively **(Fig. 2.3c)**.[73] Glycine synthesized from ^{13}C-labeled bicarbonate or formate has one higher *m/z* value, which indicated incorporation of the labeled substrate. When unlabeled bicarbonate was used M0, M1, and M2 accounted for 76.2%, 16.8%, and 7.0%, respectively **(Fig. 2.3d)**. These values were very similar to data from the NIST library standards (0; 75.7%, 1; 17.2%, and 2; 7.0%). M1 and M2 would be caused by two silicon atoms, which has relatively abundant isotopes (^{29}Si (4.7%) and ^{30}Si (3.1%)), in the derivatized molecule **(Fig. 2.3c)**. In ^{13}C-labeled formate or bicarbonate conditions M1 was up to 29.7% or 38.8%, respectively. Existence ratio of M1 was significantly increased by fed ^{13}C-labeled formate and bicarbonate. When both ^{13}C-bicarbonate and ^{13}C-formate were fed M2 was more than twice as high as unlabeled condition. These results indicated that the constructed pathway was functional as designed.

2.3.3 L-serine synthesis from formate and CO_2 via RGP in L-serine auxotroph

To assess L-serine synthesis activities from formate and CO_2 we prepared Strain 2 harboring the Cl operon and GCS operon plasmids **(Table 2.1; Strain 7)**. Strain 2 is auxotrophic for L-serine. Thus, if RGP sufficiently supplies L-serine from formate and bicarbonate the growth

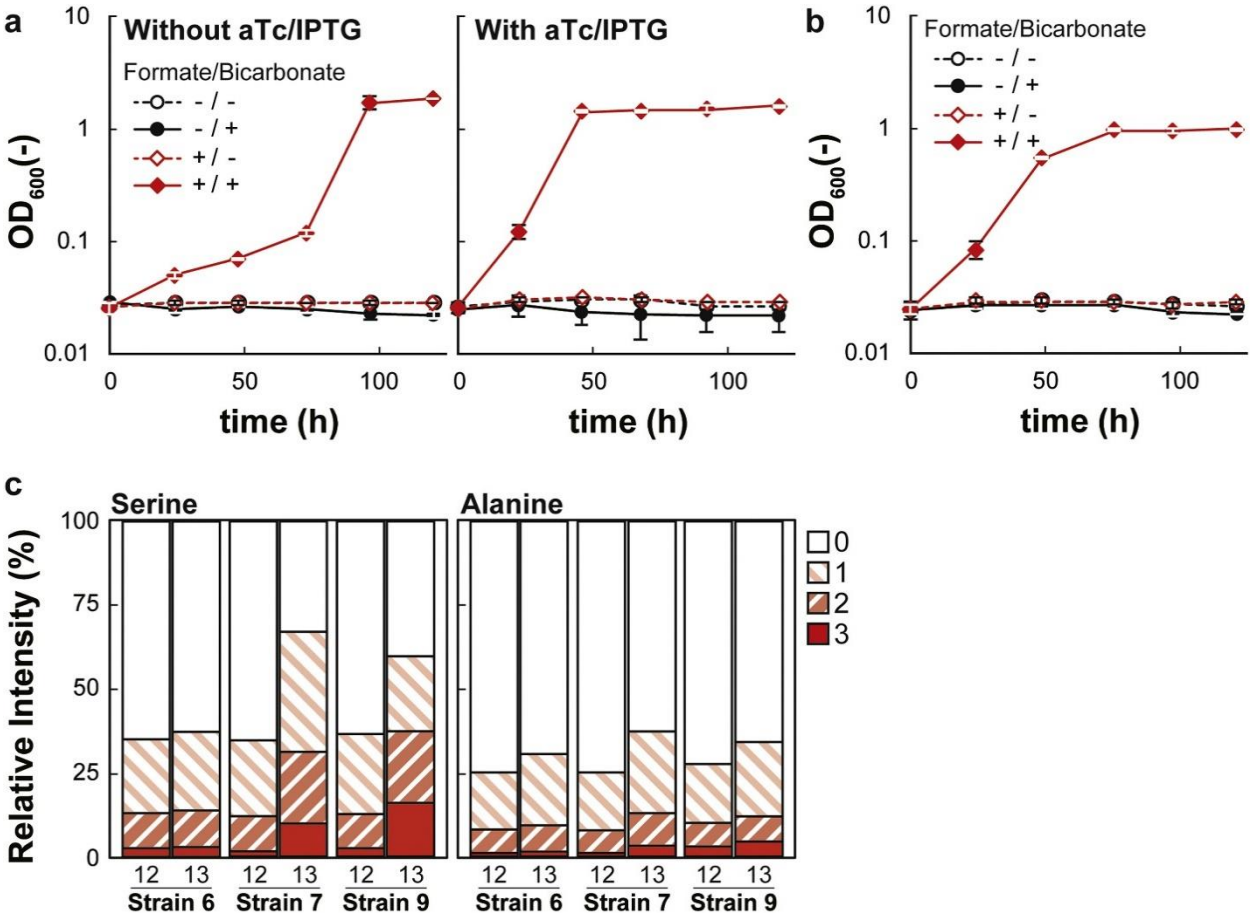


Fig. 2.4. L-serine synthesis from formate and CO_2 via RGP

(a) The Cl operon and GCS plasmids were introduced to the serine auxotroph strain, Strain 2 (termed Strain 7 **(Table 2.1)**). Cells were cultured in 5 mL of M9 minimal media containing 4 g L^{-1} of glucose without and with 0.2 g L^{-1} formate and 50 mM bicarbonate in the absence (left) and presence (right) of the inducers (10 μg L^{-1} aTc and 0.1 mM IPTG) at 37 °C. (b) *P*sdaA was replaced with $P_L lacO_1$ in Strain 7 (termed Strain 9 **(Table 2.1)**). Cells were cultured with the same condition described in (a), including both aTc and IPTG. N = 3 biological replicates; error bars represent s.d. (c) Relative intensity of the M0, M1, M2, and M3 fragments in L-serine and L-alanine produced in Strains 6, 7, and 9 **(Table 2.1)**. The values and errors are shown in Table S2.8. 12, unlabeled substrates; 13, ^{13}C-bicarbonate and ^{13}C-formate.

of Strain 7 would be rescued. To test this, Strain 7 was cultured in four different formate/bicarbonate conditions **(Fig. 2.4a)**. Growth was only observed in the formate (+)/bicarbonate (+) condition. These results indicated that RGP was active enough to produce L-serine from formate and bicarbonate in *E. coli*.

To improve the L-serine biosynthesis from formate and bicarbonate an aTc-inducible Cl operon was constructed **(Fig. S2.2b)** and the expression levels of the Cl and GCS operons were altered **(Fig. S2.4)**. In the construct, the Cl and GCS operons were under control of anhydrotetracyclin (aTc) and isopropyl β-D-1- thiogalactopyranoside (IPTG), respectively. First, four concentrations of aTc (0, 1, 10, and 100 $\mu g\, L^{-1}$) were compared **(Fig. S2.4a)**. The promoter ($P_L tetO_1$) is known as a tightly regulated synthetic promoter, but it is leaky on a high copy plasmid[79], thus the Cl operon might be expressed enough to support the growth of Strain 7 without aTc. Higher aTc concentrations improved the growth of Strain 7 and 10 $\mu g\, L^{-1}$ aTc was used for further study. Next, in the presence of 10 $\mu g\, L^{-1}$ of aTc, various IPTG concentrations were tested **(Fig. S2.4b)**. Strain 7 grew best in the presence of 10 $\mu g\, L^{-1}$ aTc and 0.1 mM IPTG. With the addition of both aTc and IPTG, the growth of Strain 7 was significantly enhanced compared to the non-induced, but optimized conditions **(Fig. 2.4a)**.

2.3.4 Optimization of the L-serine-to-pyruvate pathway

Pyruvate, a central metabolite, is the final product in the RGP **(Fig. 2.1)**. Almost all essential metabolites can be made from pyruvate, making it a prime target for future pathway development to produce high-value chemicals. The growth of Strain 1 was tested on glucose, L-serine, or glycine as a sole carbon source **(Fig. S2.5)**. The results indicated that Strain 1 did not efficiently convert L-serine to pyruvate, as L-serine supplementation suffered significantly

reduced growth compared to glucose. The glycine supplemented cultures did not grow, indicating that there are no native pathways to efficiently convert glycine to pyruvate. To allow for growth on L-serine four endogenous L-serine deaminase genes, *sdaA, sdaB, tdcB* and *tdcG,* were expressed from high copy plasmids in Strain 1 **(Fig. S2.6)**. SdaA had the greatest growth benefit on L-serine and was therefore chosen for further improvements.

The native promoter of *sdaA* (*PsdaA*) on the genome of Strain 2 was replaced with the IPTG-inducible $P_L lacO_1$85 **(Table 2.1; Strain 8 (Fig. S2.7))**. L-serine auxotrophy dictates that growth can only be observed when L-serine is successfully converted to pyruvate, providing a useful platform to screen SdaA. Strain 2 can only grow when both glucose and L-serine are supplemented, while neither glucose nor L-serine individually have significant growth benefits **(Fig. S2.7b)**. The use of the $P_L lacO_1$ promoter in Strain 8 allowed for growth solely on L-serine **(Fig. S2.7b)**. These data indicated that in Strain 8, synthesized L-serine is effectively converted to pyruvate.

2.3.5 Pyruvate synthesis from formate and CO_2

To produce pyruvate from formate and CO_2 via the RGP, the Cl operon and GCS operon plasmids were introduced into Strain 8 **(Table 2.1; Strain 9)**. Strain 9 was cultured in four different formate/bicarbonate conditions and induced using 10 μg L^{-1} aTc and 0.1 mM IPTG **(Fig. 2.4b)**. Only the formate (+)/bicarbonate (+) condition rescued the growth of Strain 9, indicating that the constructed RGP is sufficient to produce the pyruvate required for the growth of Strain 9.

To confirm that the growth rescue was due to the installed RGP, ^{13}C-bicarbonate and ^{13}C-formate were fed to Strain 6, Strain 7, and Strain 9, and the synthesized L-serine and L-alanine were relatively quantified with GC-MS **(Fig. 2.4c and Table S2.8)**. Pyruvate was not measured

because it was not detectable under our conditions. L-alanine is generated from pyruvate by a single reaction catalyzed by alanine transaminase. Thus, the abundance ratio of ^{13}C in L-alanine can be correlated to carbon flux towards pyruvate from CO_2 and formate via RGP. Due to the *glyA*-knockout, L-serine was only synthesized from glucose in Strain 6. In the strain, the abundance ratio of ^{13}C-labeled L-serine was not significantly changed **(Fig. 2.4c)**. On the other hand, the ^{13}C-labeled L-serine ratio in Strain 7 and Strain 9 increased compared to Strain 6. In the M3 fragment, ^{13}C-labeled L-serine ratio increased from 1.5% and 2.5% to 9.9% and 16.0% in Strain 7 and Strain 9, respectively **(Table S2.8)**. These results indicated that L-serine was synthesized from formate and bicarbonate via RGP in Strains 7 and 9. ^{13}C-labeled L-alanine was also slightly increased in Strain 7 compared to Strain 6 **(Fig. 2.4c and Table S2.8)**. Increases in the ^{13}C-labeled L-alanine ratio indicated that carbons from formate and/or bicarbonate were incorporated into alanine via pyruvate. These results indirectly indicated that the RGP was able to produce pyruvate from formate and bicarbonate via an L-serine intermediate in *E. coli*.

2.3.6 Electrochemical cultivation

To eliminate the need to feed formate to the engineered strains, electrocatalysis that converts CO_2 to formate was integrated into the culturing conditions **(Fig. 2.5a)**. In this system, 100% CO_2 gas is directed onto the working electrode to promote formate formation **(Fig. 2.5a)**. To achieve high efficiency in formate formation from electric current, various electrodes were evaluated on the basis of conversion efficiency from electric current to formate **(Fig. S2.8a)**. In the experiment, formate formation occurred by using indium, tin, lead, carbon, or copper. Indium especially had high efficiency for formate formation (> 60%) **(Table S2.6)** and showed high performance in the bacterial culture as well (> 90% at −2.0 V) **(Fig. S2.8b and Table S2.7)**. The

formate-generating rate in the bacterial culture was about 10 mM (= 450 mg L^{-1}) per day and was stable during the 20 day experiment. These results indicated that this system is practical for bacterial cultivation.

Strain 6 was incubated in M9 minimal media in the electrochemical cultivation system **(Fig. 2.5b)**. Strain 6 did not grow without electrolysis since formate was not supplied in the culture. When a potential of −1.2 or −2.0 V was applied on the working electrode, Strain 6 grew after a long lag phase. Under −2.0 V of potential the cell density reached optical density at 600 nm (OD_{600}) 0.47, which is almost 4-times higher than with −1.2 V (OD_{600} 0.13). The growth of Strain 6 was strongly dependent on the applied potential. The results suggested that the electrocatalysts were compatible with the biological system.

The growth rate and the final cell density of the culture grown with −2.0 V were much worse than those of the positive control (no electrolysis with formate supplement) **(Fig. 2.5b)**. Also distinctive was the long lag phase with electrolysis compared to the positive control (5 days vs 1 day). One possible explanation for the difference in lag phase is growth inhibition by reactive oxygen species (ROS) such as H_2O_2. It is known that growth of bacteria is inhibited by ROS in electro-fermentation systems.[45,68] In this experiment the cultures were prepared without using an anaerobic chamber, therefore ROS were likely to be produced in the cathodic chamber. The result suggested that Strain 6 started to grow after ROS were reduced by gas feeding or that the strain adapted to the culturing conditions. However, many other factors may also impair cell growth in the system, necessitating the characterization of these factors for further improvements.

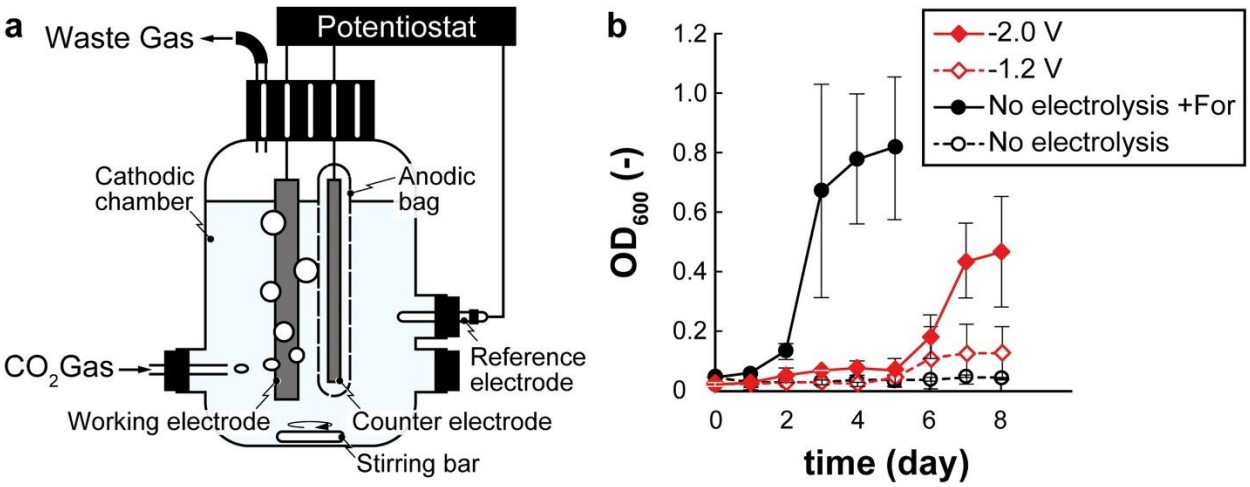


Fig. 2.5. Electrochemical-cultivation system for the RGP Strain.

In the electrochemical-cultivation system (a), Strain 6 was cultured in 250 mL of M9 minimal media containing 4 g L^{-1} of glucose and 50 mM bicarbonate with 100% CO_2 gas supply at 37 °C. The potential of the working electrode was −1.2 (open diamonds) or −2.0 V (closed diamonds). No electrolysis with formate (closed circles) or without formate (open circle) were performed as controls.

2.4 Discussion

Here we constructed an integrated system of biological and electrochemical catalysts for CO_2 fixation. This system has several advantages, such as high energy efficiency and high compatibility with renewable electricity compared to natural CO_2 fixation systems.[67] Almost all CO_2 fixation occurs naturally through the CB cycle. However, the CB cycle is not an ideal CO_2-fixing pathway due to low solar energy utilization efficiency compared to modern photovoltaic solar cells and due to slow enzymatic turnover.[78,89] Although various, more efficient carbon fixation pathways have been designed[78,89,90], none have been constructed *in vivo*. Introducing a new carbon fixing pathway into living cells is challenging since organisms have complex metabolic and enzymatic networks that require informed and precise manipulations. In this study, RGP, one of the most efficient carbon fixing pathways theorized[50], was successfully constructed in *E. coli*.

To successfully construct the RGP there were several key advantages to our hybrid system to overcome the inherent challenges of carbon fixation. First, *E. coli* has all the required enzymatic components for RGP except *fhs*, which may allow RGP construction with minimal perturbation of *E. coli* metabolism. Second, two-third of carbons in the RGP are from formate, which is electrochemically generated from CO_2. Carbon-carbon bonds using formate as a substrate are more readily formed than in CO_2. Additionally, formate is more soluble in water than CO_2. Third, *E. coli* is one of the most well-established host strains, allowing us to balance various parameters of RGP quickly. For example, we were able to engineer knock-out strains to screen RGP enzyme efficiency. Individual enzyme and operon expression levels were also altered readily using inducer-dependent promoters. These optimization processes are crucial to achieve the desired phenotype. Lastly, *E. coli* was compatible with electrocatalysis to produce formate from CO_2.

Our hybrid system employed an electrocatalyst to convert CO_2 to formate. Electrocatalysts have been previously applied to directly transfer electrons to bacteria.[35,42,91,92] However, the direct transfer of electrons to bacteria requires a large electrode surface area. In contrast, approaches that deliver electrons to microbes through an intermediate do not. An intermediate such as formate provides a buffer between electricity generation and the cellular catalysts, so the cells can cope with the intermittent nature of renewable electricity.

To provide electricity for our hybrid production system, renewably-generated electricity would be utilized. While renewably-generated electricity has already far surpassed the efficiency of solar energy utilization compared to natural photosynthesis (plants, algae, or bacteria), there are other advantages to using renewable electricity. The capacity of renewable electricity generation has been drastically increasing in recent years, especially for solar and wind energy. In 2016, the share of renewable electricity in global power generation reached 7.5%.[93] The total capacities of solar and wind energy have surpassed 700 GW, although both are intermittent and non-dispatchable sources of energy. In some cases, renewable electricity pays for itself (negative cost), or could not make money (zero cost) due to excess generation. However, that excess supply presents the problem of requiring construction of larger storage facilities. It has been proposed to utilize that excess electricity for conversion of CO_2 into chemicals.[45] Thus, our hybrid biochemical production system would contribute to solving the general problem of efficient renewable electricity utilization while also accelerating renewable energy growth.

2.5 Conclusion

This study demonstrates that our electrical-biological hybrid system in *E. coli* is capable of fixing carbon and integrating the products into central metabolism. Although the first iteration of

this system was successful, further work is needed to improve the RGP strain, in particular to stream-line metabolic carbon flux. However, considering *E. coli*'s genetic tractability, this system has the potential to improve rapidly. Various biosynthetic pathways have previously been constructed in *E. coli* to produce a variety of valuable chemicals from central metabolites.[94] However, transferring these pathways into autotrophic bacteria such as cyanobacteria and *R. eutropha* remains challenging.[6,65] To avoid those challenges, these pathways could be readily transferred into the engineered RGP *E. coli* strain to produce target chemicals from CO_2 and electricity. This hybrid system therefore has incredible potential for biochemical production that would far exceed the renewability and efficiency of current biological systems.

2.6 References

6 Case, A. E. & Atsumi, S. Cyanobacterial chemical production. *Journal of Biotechnology* **231**, 106-114, doi:10.1016/j.jbiotec.2016.05.023 (2016).

14 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).

27 Zhang, A., Carroll, A. L. & Atsumi, S. Carbon recycling by cyanobacteria: Improving CO2 fixation through chemical production. *FEMS Microbiology Letters* **364**, 16, doi:10.1093/femsle/fnx165 (2017).

31 Whitney, S. M., Houtz, R. L. & Alonso, H. Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. *Plant Physiology* **155**, 27-35, doi:10.1104/pp.110.164814 (2011).

32 Riplinger, C., Sampson, M. D., Ritzmann, A. M., Kubiak, C. P. & Carter, E. A. Mechanistic contrasts between manganese and rhenium bipyridine electrocatalysts for the reduction of carbon dioxide. *Journal of the American Chemical Society* **136**, 16285-16298, doi:10.1021/ja508192y (2014).

33 Taheri, A., Thompson, E. J., Fettinger, J. C. & Berben, L. A. An Iron Electrocatalyst for Selective Reduction of CO2 to Formate in Water: Including Thermochemical Insights. *ACS Catalysis* **5**, 7140-7151, doi:10.1021/acscatal.5b01708 (2015).

34 Genovese, C., Ampelli, C., Perathoner, S. & Centi, G. Mechanism of C-C bond formation in the electrocatalytic reduction of CO2 to acetic acid. A challenging reaction to use renewable energy with chemistry. *Green Chemistry* **19**, 2406-2415, doi:10.1039/c6gc03422e (2017).

35 Rabaey, K. & Rozendal, R. A. Microbial electrosynthesis - Revisiting the electrical route for microbial production. *Nature Reviews Microbiology* **8**, 706-716, doi:10.1038/nrmicro2422 (2010).

36 Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* **332**, 805-809, doi:10.1126/science.1200165 (2011).

37 Dunn, B., Kamath, H. & Tarascon, J. M. Electrical energy storage for the grid: A battery of choices. *Science* **334**, 928-935, doi:10.1126/science.1212741 (2011).

38 Soloveichik, G. L. Battery technologies for large-scale stationary energy storage. *Annual Review of Chemical and Biomolecular Engineering* **2**, 503-527, doi:10.1146/annurev-chembioeng-061010-114116 (2011).

39 Lewis, N. S. & Nocera, D. G. Powering the planet: Chemical challenges in solar energy utilization. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15729-15735, doi:10.1073/pnas.0603395103 (2006).

40 Lovley, D. R. & Nevin, K. P. Electrobiocommodities: Powering microbial production of fuels and commodity chemicals from carbon dioxide with electricity. *Current Opinion in Biotechnology* **24**, 385-390, doi:10.1016/j.copbio.2013.02.012 (2013).

41 Tremblay, P. L. & Zhang, T. Electrifying microbes for the production of chemicals. *Frontiers in Microbiology* **6**, 201, doi:10.3389/fmicb.2015.00201 (2015).

42 Ross, D. E., Flynn, J. M., Baron, D. B., Gralnick, J. A. & Bond, D. R. Towards electrosynthesis in Shewanella: Energetics of reversing the Mtr pathway for reductive metabolism. *PLoS ONE* **6**, 2, doi:10.1371/journal.pone.0016649 (2011).

43 Tashiro, Y., Hirano, S., Matson, M. M., Atsumi, S. & Kondo, A. Electrical-biological hybrid system for CO2 reduction. *Metabolic Engineering* **47**, 211-218, doi:10.1016/j.ymben.2018.03.015 (2018).

44 Khunjar, W. O., Sahin, A., West, A. C., Chandran, K. & Banta, S. Biomass Production from Electricity Using Ammonia as an Electron Carrier in a Reverse Microbial Fuel Cell. *PLoS ONE* **7**, 9, doi:10.1371/journal.pone.0044846 (2012).

45 Li, H. *et al.* Integrated electromicrobial conversion of CO2 to higher alcohols. *Science* **335**, 1596, doi:10.1126/science.1217643 (2012).

46 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).

47 Atsumi, S. *et al.* Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Microbiology and Biotechnology* **85**, 651-657, doi:10.1007/s00253-009-2085-6 (2010).

48 Jeong, J., Kim, J. Y. & Yoon, J. The role of reactive oxygen species in the electrochemical inactivation of microorganisms. *Environmental Science and Technology* **40**, 6117-6122, doi:10.1021/es0604313 (2006).

49 Huang, W. D. a. Analysis of biofuels production from sugar based on three criteria: Thermodynamics, bioenergetics, and product separation. *Energy and Environmental Science* **4**, 784-792, doi:10.1039/c0ee00069h (2011).

50 Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochimica et Biophysica Acta - Bioenergetics* **1827**, 1039-1047, doi:10.1016/j.bbabio.2012.10.013 (2013).

51 Carroll, A. L., Desai, S. H. & Atsumi, S. Microbial production of scent and flavor compounds. *Current Opinion in Biotechnology* **37**, 8-15, doi:10.1016/j.copbio.2015.09.003 (2016).

52 Peralta-Yahya, P. P. & Keasling, J. D. Advanced biofuel production in microbes. *Biotechnology Journal* **5**, 147-162, doi:10.1002/biot.200900220 (2010).

53 Pfleger, B. F., Gossing, M. & Nielsen, J. Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering* **29**, 1-11, doi:10.1016/j.ymben.2015.01.009 (2015).

54 Wilbanks, B. & Trinh, C. T. Comprehensive characterization of toxicity of fermentative metabolites on microbial growth Mike Himmel. *Biotechnology for Biofuels* **10**, 262, doi:10.1186/s13068-017-0952-4 (2017).

55 Ko, Y. S. *et al.* Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chemical Society Reviews* **49**, 4615-4636, doi:10.1039/d0cs00155d (2020).

56 Dunlop, M. J. *et al.* Engineering microbial biofuel tolerance and export using efflux pumps. *Molecular Systems Biology* **7**, 784, doi:10.1038/msb.2011.21 (2011).

57 Mukhopadhyay, A. Tolerance engineering in bacteria for the production of advanced biofuels and chemicals. *Trends in Microbiology* **23**, 498-508, doi:10.1016/j.tim.2015.04.008 (2015).

58 Chen, H. *et al.* Directed evolution of mevalonate kinase in: Escherichia coli by random mutagenesis for improved lycopene. *RSC Advances* **8**, 15021-15028, doi:10.1039/c8ra01783b (2018).

59 Portnoy, V. A., Bezdan, D. & Zengler, K. Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* **22**, 590-594, doi:10.1016/j.copbio.2011.03.007 (2011).

60 Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution - principles and applications for biotechnology. *Microb Cell Fact* **12**, 64, doi:10.1186/1475-2859-12-64 (2013).

61 Herring, C. D. *et al.* Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. *Nature Genetics* **38**, 1406-1412, doi:10.1038/ng1906 (2006).

62 Atsumi, S. *et al.* Evolution, genomic analysis, and reconstruction of isobutanol tolerance in Escherichia coli. *Mol Syst Biol* **6**, 449, doi:10.1038/msb.2010.98 (2010).

63 Lynd, L. R., Cushman, J. H., Nichols, R. J. & Wyman, C. E. Fuel etanol from cellulosic biomass. *Science* **251**, 1318-1323, doi:10.1126/science.251.4999.1318 (1991).

64 Wijffels, R. H. & Barbosa, M. J. An outlook on microalgal biofuels. *Science* **329**, 796-799, doi:10.1126/science.1189003 (2010).

65 Fast, A. G. & Papoutsakis, E. T. Stoichiometric and energetic analyses of non-photosynthetic CO 2-fixation pathways to support synthetic biology strategies for production of fuels and chemicals. *Current Opinion in Chemical Engineering* **1**, 380-395, doi:10.1016/j.coche.2012.07.005 (2012).

66 Goeppert, A., Czaun, M. & Jones, J. P. a. Recycling of carbon dioxide to methanol and derived products-closing the loop. *Chemical Society Reviews* **43**, 7995-8048, doi:10.1039/c4cs00122b (2014).

67 Claassens, N. J. & Sousa, D. Z. a. Harnessing the power of microbial autotrophy. *Nature Reviews Microbiology* **14**, 692-706, doi:10.1038/nrmicro.2016.130 (2016).

68 Liu, C., Coln, B. C., Ziesack, M., Silver, P. A. & Nocera, D. G. Water splitting-biosynthetic system with CO2 reduction efficiencies exceeding photosynthesis. *Science* **352**, 1210-1213, doi:10.1126/science.aaf5039 (2016).

69 Nybo, S. E., Khan, N. E., Woolston, B. M. & Curtis, W. R. Metabolic engineering in chemolithoautotrophic hosts for the production of fuels and chemicals. *Metabolic Engineering* **30**, 105-120, doi:10.1016/j.ymben.2015.04.008 (2015).

70 Raven, J. A. Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquatic Microbial Ecology* **56**, 177-192, doi:10.3354/ame01315 (2009).

71 Raines, C. A. Increasing photosynthetic carbon assimilation in C3 plants to improve crop yield: Current and future strategies. *Plant Physiology* **155**, 36-42, doi:10.1104/pp.110.168559 (2011).

72 Paukert, J. L. & Rabinowitz, J. C. 85 Formyl-Methenyl-Methylenetetrahydrofolate Synthetase (Combined): A Multifunctional Protein in Eukaryotic Folate Metabolism. **66**, 616-626, doi:10.1016/0076-6879(80)66515-X (1980).

73 Kikuchi, G., Motokawa, Y., Yoshida, T. & Hiraga, K. Glycine cleavage system: Reaction mechanism, physiological significance, and hyperglycinemia. *Proceedings of the Japan Academy Series B: Physical and Biological Sciences* **84**, 246-263, doi:10.2183/pjab.84.246 (2008).

74 Schuchmann, K. & Mller, V. Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. *Nature Reviews Microbiology* **12**, 809-821, doi:10.1038/nrmicro3365 (2014).

75 Schirch, V., Hopkins, S., Villar, E. & Angelaccio, S. Serine hydroxymethyltransferase from Escherichia coli: Purification and properties. *Journal of Bacteriology* **163**, 1-7, doi:10.1128/jb.163.1.1-7.1985 (1985).

76 Yishai, O., Goldbach, L., Tenenboim, H., Lindner, S. N. & Bar-Even, A. Engineered Assimilation of Exogenous and Endogenous Formate in Escherichia coli. *ACS Synthetic Biology* **6**, 1722-1731, doi:10.1021/acssynbio.7b00086 (2017).

77 Bar-Even, A. Formate Assimilation: The Metabolic Architecture of Natural and Synthetic Pathways. *Biochemistry* **55**, 3851-3863, doi:10.1021/acs.biochem.6b00495 (2016).

78 Bar-Even, A., Noor, E. & Milo, R. A survey of carbon fixation pathways through a quantitative lens. *Journal of Experimental Botany* **63**, 2325-2342, doi:10.1093/jxb/err417 (2012).

79 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 Research* **25**, 1203-1210, doi:10.1093/nar/25.6.1203 (1997).

80 Jiang, Y. *et al.* Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Applied and Environmental Microbiology* **81**, 2506-2514, doi:10.1128/AEM.04023-14 (2015).

81 Clark, J. & Ljungdahl, L. G. Purification and properties of 5,10-methenyltetrahydrofolate cyclohydrolase from Clostridium formicoaceticum. *Journal of Biological Chemistry* **257**, 3833-3836, doi:10.1016/s0021-9258(18)34857-9 (1982).

82 Formyl-methenyl-methylenetetrahydrofolate synthetase(combined) from yeast. Biochemical characterization of the protein from an ADE3 mutant lacking the formyltetrahydrofolate synthetase function. *Journal of Biological Chemistry* **255**, 2569-2577, doi:10.1016/s0021-9258(19)85930-6 (1980).

83 Ljungdahl, L. G., O'Brien, W. E., Moore, M. R. & Liu, M. T. 83 Methylenetetrahydrofolate Dehydrogenase from Clostridium formicoaceticum and Methylenetetrahydrofolate Dehydrogenase, Methenyltetrahydrofolate Cyclohydrolase (Combined) from Clostridium thermoaceticum. **66**, 599-609, doi:10.1016/0076-6879(80)66513-6 (1980).

84 Kopke, M. *et al.* Clostridium ljungdahlii represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13087-13092, doi:10.1073/pnas.1004716107 (2010).

85 Poehlein, A. *et al.* An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis. *PLoS ONE* **7**, 3, doi:10.1371/journal.pone.0033439 (2012).

86 D'Ari, L. & Rabinowitz, J. C. Purification, characterization, cloning, and amino acid sequence of the bifunctional enzyme 5,10-methylenetetrahydrofolate dehydrogenase/5,10-methenyltetrahydrofolate cyclohydrolase from Escherichia coli. *Journal of Biological Chemistry* **266**, 23953-23958, doi:10.1016/s0021-9258(18)54377-5 (1991).

87 Kikuchi, G. The glycine cleavage system: Composition, reaction mechanism, and physiological significance. *Molecular and Cellular Biochemistry* **1**, 169-187, doi:10.1007/BF01659328 (1973).

88 Gariboldi, R. T. & Drake, H. L. Glycine synthase of the purinolytic bacterium, Clostridium acidiurici. Purification of the glycine-CO2 exchange system. *Journal of Biological Chemistry* **259**, 6085-6089, doi:10.1016/s0021-9258(20)82108-5 (1984).

89 Schwander, T. a. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900-904, doi:10.1126/science.aah5237 (2016).

90 Siegel, J. B. *et al.* Computational protein design enables a novel one-carbon assimilation pathway. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 3704-3709, doi:10.1073/pnas.1500545112 (2015).

91 LaBelle, E. V. & May, H. D. Energy efficiency and productivity enhancement of microbial electrosynthesis of acetate. *Frontiers in Microbiology* **8**, 756, doi:10.3389/fmicb.2017.00756 (2017).

92 Lovley, D. R. Powering microbes with electricity: Direct electron transfer from electrodes to microbes. *Environmental Microbiology Reports* **3**, 27-35, doi:10.1111/j.1758-2229.2010.00211.x (2011).

93 BP statistical review of world energy. (BP, 2017).

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

2.7 Supplemental Information

Table S2.1. E. coli strains used in this study

Strain Name	Genotype	Ref
BW25113	*rrnB*$_{T14}$ Δ*lacZ*WJ16 *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LC78}$	[95]
JCL16	BW25113/F' [*traD*36, *proAB*$^{+}$ *lacI*^{q} ZΔM15 Tn10(tet^{r})]	[96]
YT151	Same as JCL16 but with Δ*serA*	This study
YT152	Same as JCL16 but with Δ*glyA*	This study
YT166	Same as JCL16 but with Δ*serA* Δ*gcvP*	This study
YT276	Same as YT151 but with *PsdaA*::*PLlacO1*	This study

Table S2.2. List of plasmids used in this study

Plasmid name	Gene contents	Origin	Marker	Ref
pAL300	$P_L$lacO$_1$: *alsS$_{Bs}$-alsD$_{Ah}$-adh$_{Cb}$*	ColE1	Amp^R, Gent^R	[97]
pAL420	$P_L$tetO$_1$	p15A	Cm^R	This study
pAL956	$P_L$lacO$_1$: *sfGFP*	Cola	Kan^R	[98]
pAL998	$P_L$lacO$_1$: *sfGFP*	Cola	Gent^R	This study
pYT007	P_{A1}lacO$_1$: *fhs$_{Cl}$*	p15A	Kan^R	This study
pYT036	P_{A1}lacO$_1$: *fhs$_{Cl}$-folD*	p15A	Kan^R	This study
pYT042	$P_L$lacO$_1$	Cola	Gent^R	This study
pYT043	$P_L$lacO$_1$: *sdaA*	Cola	Gent^R	This study
pYT044	$P_L$lacO$_1$: *sdaB*	Cola	Gent^R	This study
pYT045	$P_L$lacO$_1$: *tdcB*	Cola	Gent^R	This study
pYT046	$P_L$lacO$_1$: *tdcG*	Cola	Gent^R	This study
pYT048	$P_L$lacO$_1$: *gcvTHP*	ColE1	Amp^R	This study
pYT079	P_{A1}lacO$_1$: *fhs$_{Aw}$-folD*	p15A	Kan^R	This study
pYT085	P_{A1}lacO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*	p15A	Kan^R	This study
pYT086	P_{A1}lacO$_1$: *fhs$_{Aw}$-fchA$_{Aw}$-folD$_{Aw}$*	p15A	Kan^R	This study
pYT099	$P_L$tetO$_1$	p15A	Kan^R	This study
pYT100	$P_L$tetO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$-folD$_{Cl}$*	p15A	Kan^R	This study
pYT101	P_{A1}lacO$_1$: *fchA$_{Cl}$*	p15A	Kan^R	This study
pYT102	P_{A1}lacO$_1$: *folD$_{Cl}$*	p15A	Kan^R	This study
pYT103	P_{A1}lacO$_1$: *fhs$_{Cl}$-fchA$_{Cl}$*	p15A	Kan^R	This study
pYT104	P_{A1}lacO$_1$: *fhs$_{Cl}$-folD$_{Cl}$*	p15A	Kan^R	This study
pYT171	*gRNA* for *PsdaA::PLlacO1*	pMB1	Spec^R	This study
pZA23MCS	P_{A1}lacO$_1$	p15A	Kan^R	[99]
pZA31-luc	$P_L$tetO$_1$: *luc*	p15A	Cm^R	[99]
pZE12MCS	$P_L$lacO$_1$	ColE1	Amp^R	[99]

Cl, *Clostridium ljungdahalii*; *Aw*, *Acetobacter woodii*; Non-annotated genes are from *Escherichia coli*.

Table S2.3. Oligonucleotides used in this study

Name	Sequence
YT101	GCTTCCCAACCTTACCAGAG
YT167	TAATGACTCTAGAGGCATCAAATAA
YT168	CATGGTACCTTTCTCCTCTTTAA
YT736	GACTCCTGTTGATAGATCCAGTAATG
YT737	GCGCCCTCTGGTAAGGTTGGGAAGCCACCGTGGAAACGGATGAAGG
YT738	ATTACTGGATCTATCAACAGGAGTCGATCTCGGCTTGAACGAATTGTTAGG
YTC001	CATGGTACCTTTCTCCTCTTTAATG
YTC012	GGTACCTTTCTCCTCTTTAATGAATTC
YTC013	ACGCGTGCTAGAGGCATCAAATAAAAC
YTC111	AACCCGGAAGGAGATATACCATGGCAGCAAAGATTATTGACGGTAA
YTC112	GATGCCTCTAGCACGCGTTTACTCATCCTGTGGATCATGATATTCAACG
YTC113	TGCCATGGTATATCTCCTTCCGGGTTAAAACAGACCGCTGATCACACCA
YTC131	CCTACTCAGGAGAGCGTTCAC
YTC134	CATGGTACCTTTCTCCTCTTTAATGAATTCG
YTC135	ACGCGTGCTAGAGGCATCAAATAA
YTC141	AAGAGGAGAAAGGTACCATGATTAGTCTATTCGACATGTTTAAGGTGGG
YTC142	ATTTGATGCCTCTAGCACGCGTTTAGTCACACTGGACTTTGATTGCCAG
YTC143	AAGAGGAGAAAGGTACCATGATTAGCGTATTCGATATTTTCAAAATCGG
YTC144	TTGATGCCTCTAGCACGCGTTTAATCGCAGGCAACGATCTTCA
YTC145	TTAAAGAGGAGAAAGGTACCATGCATATTACATACGATCTGCCGGT
YTC146	TTGATGCCTCTAGCACGCGTTTAAGCGTCAACGAAACCGGTG
YTC147	AAGAGGAGAAAGGTACCATGATTAGTGCATTCGATATTTTCAAAATTGG
YTC148	TTGATGCCTCTAGCACGCGTTTAGCCGCAGACCACTTTAATGGC

YTC154	CATTAAAGAGGAGAAAGGTACCATGGCACAACAGACTCCTTTGTACG
YTC155	ATTTGATGCCTCTAGCACGCGTTTACTGGTATTCGCTAATCGGTACGCA
YTC217	ATTCATTAAAGAGGAGAAAGGTACCATG
YTC230	CATGGTACCTTTCTCCTCTTTAATGAATTC
YTC231	TAAACGCGTGCTAGAGGC
YTC249	CCCGGAAGGAGATATACCATG
YTC330	GTCGACGAGGAATCACCATGAAACTGG
YTC331	ATGGTATATCTCCTTCCGGGTTATTTGCTCA
YTC332	AACCCGGAAGGAGATATACCATGGGC
YTC333	TTTATTTGATGCCTCTAGCACGCGTTTACAG
YTC334	TAAACGCGTGCTAGAGGCATCAAATAAAACG
YTC335	GTGATTCCTCGTCGACTTAAAACAGACCGCTGATCACACCATTTTCATC
YTC336	GTCGACGAGGAATCACCATGGAATTTGA
YTC337	ATGGTATATCTCCTTCCGGGTTACAGCA
YTC338	AACCCGGAAGGAGATATACCATGGC
YTC339	GTGATTCCTCGTCGACTTAGAACAGACCGGCAATAACACCATTTTCATC
YTC388	CGACGTCTAAGAAACCATTATTATCATGACATTAACC
YTC389	GAGCTCGCTTGGACTCCTG
YTC390	TATCAACAGGAGTCCAAGCG
YTC391	TGTCATGATAATAATGGTTTCTTAGACGTC
YTC399	TGATGCCTCTAGCACGCGTTTATTTGCTCAGTGCGCTAACAACTTTTTC
YTC400	GTTTTAAGTCGACGAGGAATCACCATGGGCCAGATTATCAAAGGTAAAC
YTC401	CATTAAAGAGGAGAAAGGTACCATGAAACTGGCCGATAAAAGC
YTC402	CATTAAAGAGGAGAAAGGTACCATGGGCCAGATTATCAAAGG

Table S2.4. DNA sequences of synthesized genes used in this study

Gene	DNA sequence
***fhs*$_{Cl}$**	ATGACCTACAAAAGCGATATTGAAATTGCCCAAGAGTGCACCATGAAAGACATTAAAGA AATCGCCAAAAAACTGAACATCAGCGAAGATGACATTGAGCTGTATGGTAAATACAAAG CCAAAGTGAACTATAACCTGCTGAAAACCACACCGGGTAAAAATGGTAAACTGATTCTG TGTACCGCCATTAATCCGACACCGGCAGGCGAAGGTAAAACCACCACCGCAATTGGTGT TGCAGATGCACTGAATCGTATGGGTAAAAGCGTTGTTGTTGCACTGCGTGAACCGAGCA TGGGTCCGGTTTTTGGTATTAAAGGTGGTGCAGCAGGCGGTGGTTATGCACAGGTTGTT CCGATGGAAGACATCAATCTGCATTTTACCGGTGATATTCATGCACTGACCGCAGCAAAT AATCTGCTGGCAGCAATGATTGATAACCATATTTATCAGGGCAACAAACTGAATATTGAT CCGCGTCGTATTGCATGGCGTCGTTGTGTTGATATGAATGATCGTCAGCTGCGTTTTGTT GTTGATGGTCTGGGTGGTAAAGCAAATGGTACACCGCGTGAAGATGGTTTTGATATTACC GTTGCAAGCGAAATCATGGCCATTTTTTGTCTGAGCAGCGATATTATCGATCTGAAAAAT CGCATTGCCAAAATCGTTGTTGGTTATACCCGTGATGGTAAACCGGTTACCGCACATGAT CTGAAAGCCGAAGGTGCAATGGCAGCACTGCTGAAAGATGCCCTGAAACCGAATCTGG TTCAGACCCTGGAAGGTACACCGGCATTTGTTCATGGTGGTCCGTTTGCAAATATTGCGC ATGGTTGTAATAGCATTATGGCAACCCGTATGGCACTGCATTTTGGTGATTATGTTGTTAC CGAAGCAGGTTTTGGTGCAGATCTGGGTGCAGAAAAATTTCTGGACATTAAATGTCGTA TGGCAGGTCTGAAACCGGATGCAGTTATTATTGTTGCAACCGTTCGTGCGCTGAAATATA ACGGTGGTGTTCCGAAAGCAGATCTGAATAATGAAAACCTGGGTGCACTGGAAAAAGG TCTGCCGAATCTGCTGAAACATGTTGAAAACATCACCAAAGTGTATAAACTGCCTGCCG TTGTTGCGCTGAATGCATTTCCGACCGATACCCAGGCAGAACTGAAACTGGTTGAAGAT AAATGTAAAGAACTGGGCGTGAATGTGAAACTGAGCGAAGTTTGGGCAAAAGGTGGC GAAGGTGGTATTGAAGTTGCAAAAGAAGTTCTGCGCCTGATCAAAGAAGAGAAAAACG ATTTTCAGTTCGCGTACGATGAAAAACTGCCGATTCGTGATAAAATTCGTGCAGTTGCAC AGAAAATCTATGGTGCCGATGATGTTACCTTTACCAATCAGGCAGATAAAGAAATTGATG AGCTGGAAAAACTGGGCTTTGGTAAAACACCGGTTTGTATTGCAAAAACCCAGTATAGC CTGACCGATGATCAGACCAAACTGGGTCGTCCGACCGGTTTTAACATTACCGTGCGTCA GGTTACCATTAGTGCCGGTGCCGGTTTTGTGGTTGCAGTTACCGGTAGTATTATGAAAAT GCCAGGTCTGGGTAAAGTTCCGAGCGCAGAAAAAATTGATGTGGATGAAAATGGTGTG ATCAGCGGTCTGTTTTAA
***fchA*$_{Cl}$**	ATGAAACTGGCCGATAAAAGCTGCACCGATTTTATTGAAGTTCTGGCAAGCAAAGCAGC AACACCGGGTGGTGGTGGCGGTAGCGCAATTACCGGTGCCATTGGTATGGCACTGGGTG GTATGGTTTGTAATCTGACCATTGGCAAAAAAAAATACGCCCAGTATGATGAAAAAGTG AAAGGCATTCTGAAACGTTCCGATGAACTGCAAGCAGAACTGCTGAAAATGATGGATG CAGATGCAGAATGTTTTCTGCCGCTGAGCAAAGCATATGGTATGCCGAAAGATACCGAG GAACAGAAAAAAATCAAAGAAGAAACCCTGGAAAAATGCCTGAAACAGGCATGTAGC GTTCCGGTTAGCATTGTTAAACAGGCCTATGAAGCAATCAAACTGCATGAAGCACTGGT TGATAATTGTAGCAAACTGGCCATTAGTGATGTTGGTGTTGGTGTGCAGTGTCTGCGTGC AGCAATTATTGGTGCACAGCTGAATGTGATTATCAACATCAACAGCATCAAAGACCAAG AATATGTGAAAAAAGTTAAAGCCGAAACCGAACCGCTGGTTGAAGAAGGTATCAAAAT TGCCGATAAAGTGTACGAAAAAGTTGTTAGCGCACTGAGCAAATAA
***folD*$_{Cl}$**	ATGGGCCAGATTATCAAAGGTAAACCGGTTGCAGATGCAATTAGCGAAGCACTGACCAA AGAAGTTAACGACCTGAAAGTTAAAGGCATTACCCCGAAACTGACCCTGGTTCGTGTTG GTGCAAATGGTAGCGATCTGGCATATGAAAAAGGTGCCCTGAAAAAATGCGAAAAAATT GGTATTGAAGCCGTGGTGAAAGAACTGCCTGCCGATATTAGCCAGGATAAATTCATTGA AGAACTGAAAAAAATCAACGCCGATAAAACCGTGAATGCCATTATGGTTTTTCGTCCGT TTCCGAAACAGCTGGATGAAAGCGTGATCAAATATATCATTGCACCGGAAAAAGATGTG GATTGCTTTAGTCCGGTTAATGTTGCAAAACTGATGGAAAAAGACATGACCGGTTTTGC ACCGTGTACCCCGAGCGCAGTTATTGAAATTCTGAAACACTATAAAGTGCCGATGAAAG GCAAAAATGCCGTTATTGTTGGTCGTAGCATGGTTGTTGGCAAACCGGCATGTATGCTGC

	TGCTGAATGAAAATGCAACCGTTACCGTTTGTCATAGCAAAACCACCGATATGCCGAAA GTTTGTAGCCAGGCAGATATTCTGGTTGTGGGTATTGGTAAAGCCAAAATGATCGATAGC AAATATGTGAAAGATGGTGCGGTTGTTATTGATGTGGGCATTAATGTTGATGAGAGCGGT AAACTGTGTGGTGATGTTGATACCGAAGATTGTGAAGCAAAAGCCAGCATGATTACACC GGTTCCGGGTGGTGTTGGTAGCGTTACCAGCAGCATTCTGGCACAGCATATTGTTAAAG CATGTAAACTGCAAAACAATCTGTAA
***fhs**$_{Aw}$*	ATGGGCTTTAAAAGCGATATTGAAATTGCACAAGAAGCCACACCGCAGGATATTCGTGA AATTGCCAAAAAACTGGGTCTGACCGAAGATGATCTGGATCTGTATGGTAAATACAAAG CCAAAGTGGATTATAACCTGCTGAAAAAAAGCACCGGTAAAAAAGCCCGTCTGATTCTG ACCACCGCAATTAATCCGACACCGGCAGGCGAAGGTAAAACCACCACAACCATTGGTG TTGCAGATGGTCTGAGCCGTATTGGTAAAAACACCCTGGTTGCACTGCGTGAACCGAGC CTGGGTCCGGTTTTTGGTGTTAAAGGTGGTGCAGCCGGTGGTGGTTATGCACAGGTTGT TCCGATGGAAGACATTAATCTGCATTTTACGGGTGATTTTCATGCCATTGGTGCAGCAAA TAATCTGCTGGCAGCAATGCTGGATAACCATATTAAACAGGGCAACGAGCTGAAAATCG ACGCCAAAAAAATCACCTGGCGTCGTTGTGTTGATATGAATGATCGTCAGCTGCGTAATA TTGTTGATGGTCTGGGTGGTAGCGGTGATGGTGTTGTTCGTGAAGATGGTTTTGACATTA CCGTTGCAAGCGAAGTTATGGCAGCATTTTGTCTGAGCAGCGATATTAGCGATCTGAAA GAACGTCTGGGTCGTATTATTGTTGCATATAGCTTTACAGGTGAACCGGTTACCGCAGAA CAGCTGAAAGCAAATGGTGCAATGGCTGCACTGCTGAAAGATGCACTGAAACCGAATC TGGTTCAGACCCTGGAAGGTACACCGGCATTTATTCATGGTGGTCCGTTTGCAAATATTG CCCATGGTTGTAATAGCGTTATTGCAACCCGTATGGCAATGCATTTTGCCGATTATGTTGT TACCGAAGGTGGTTTTGGTGCAGATCTGGGTGCAGAAAAATTTCTGGACATTAAATGTC GTATGGCCAACCTGAAACCGGATGCAGTTATTATTGTGGCAACCGTTCGTGCCCTGAAAT ATAACGGTGGTGTTGCCAAAGCAGATCTGAATAATGAAAATCTGGAAGCACTGAAAGC AGGTCTGCCGAATCTGCTGAAACATGTTGAAAACATTACCCAGGTGTTTAAACTGCCTG CCGTTGTTGCAATTAATGAATTTCCGCTGGATACCGAAGCAGAACTGCAACTGGTGAAA AGCGAATGTCAGAAACTGGGTGTTAATGTTGCCATTAGCCAGGTTTGGGCAAAAGGCG GTGAAGGTGGTGAAGAACTGGCAAAAGAAGTTGTTCGTCTGATTGATGAAAGCGAAGG CAACTTTGAGTATTGCTATGACCTGGATATTCCGATCAAAGAAAAAATCGAAACCATTGC CACCCGCATTTATGGTGCCGATGGTGTGGATTTTACTCCGGCAGCAGCAAAAGAAATGG ATCGCCTGACCGCACTGGGTTTTGATAAAGTTCCGATTTGTATGGCGAAAACGCAGTATA GCCTGACCGATGATGCAACCAAACTGGGACGTCCGACCGGTTTCAAAATTACCGTTCGC CAGCTGACCATTAGTGCCGGTGCAGGTTTTATCATTGCACTGACCGGTGAAATTATGAAA ATGCCTGGCCTGCCGAAAGTTCCGGCAGCCGAAAAAATTGATGTTGATGAAAATGGTGT TATTGCCGGTCTGTTCTAA
***fchA**$_{Aw}$*	ATGGAATTTGATGCAGTTGGCACCGAACAGAGCAGCATTAGCGAGAAAAAATGTACCG AATTTGTGGAAGCCCTGTATAGCAAAGCAGCAGTTCCGGGTGGTGGTGGCGCAGCAGC ACTGGTTGGTGCAGTTGGTACAGCACTGGCAGGTATGGTTGGTAATCTGACCACCGGTA AAAAAAAATATGCAGCCTTTGAAGATGACATCCAGCGCATTCTGAAAGAAGCACAGATT CTGCAAGATCGTCTGCTGGCAATGATTGATGAAGATGCAAAAAACTTTCTGCCGCTGAG CAAAGCATATGGTCTGCCGAAAGAAACCGAAGCAGAAAAAGCCTATAAAGAAAAAACC CTGGAAGAGTGTACCAAAGTTGCATGTAGCATTCCGCTGGAAATTGTTGAAGTTTGCTA TAAAGCAGTCCTGCTGCAAGAAGAACTGGTTGGCAAAGGTAGCGCACTGGCAATTAGT GATGTTGCATGTGGTGTTCAGTGTCTGCGTGCAGCAATGATTAGCGGTTGGGTTAATGTT CTGATCAACATCAAAACCATCAAAGATAAAGACTACGTGGATGACGTGAACAAACGTGT TAAACCGATGCTGGAAAAAGGTGTGGAAATTTGCGATCGTGTTTATGCGGATGTTGAAA AACAGCTGCTGTAA
***folD**$_{Aw}$*	ATGGCAGCAAAACTGCTGAGCGGTAAAGAAGTTAGCGAAAGTATGCTGGCCGAAGTTC TGAAAGATGCAAATGAACTGAAAGCCAAAGGCATCACCGTTAAAATGGCAATTATGCGT GTTGGTGAAGATCCGGGTAGCATTAGCTATGAAAAAAGCATTATTACCCGCATGGGCAA AAGCAACATTGAAGTTGAAAGCGTTCAGTTTCCGATTGATGTTACCGAAGCAGATTTTAT TGCCAAACTGCAGAGCATCAACGAGGATAAAAACATTCATAGCGTGCTGATTTTTCAGC CGCTGCCGGATCAGATTGATGCAGAAAAAATCAAATATCTGCTGAGTCCGGAAAAAGAT CCGGATGCACTGAATCCGACCAATCTGGGTAAACTGATGATTGCAGATGAACGTGGTTT TTTTCCGTGTACAGCCGAAGGTGTTATGGAAATGTTCAAATTCTATAACATCGACGTGAA

	AGGCAAAGATGTGGTGGTTATTAACAATAGCAATGTTCTGGGCAAACCGCTGACCATTAT GCTGACCAATGAATTTGCAACCGTTACCATGTGTCACGTGTTTACCAAAGATACCGCAA GCTATACCAAAAAAGCCGATATTGTTGTTACCGCCTGTGGTATTTATGGTCTGGTTAAAC CGGATATGCTGAGCGAAGATTGTATTCTGATTGATGTGGCAATGGCCCAGATGAAAGATG AAAACAAAGAATTTGTGCTGAACGAAGAAGGCAAAAAAATCCGTACCGGTGATGCACA TGTTGATTGCCTGAATAAAGTTGCAATGATTACCAGCGCAACACCGGGTTGTGGTGGTG GCACCGGTCCGATTACCACCGCACTGCTGGCAAAACATGTTATTAAAGCATGTAAAATG CAGAACGGCCTGCTGTAA

Cl, *C. ljungdahalii*; Aw, *A. woodii*

Table S2.5. Set of primers and template for plasmid construction

Plasmid	Backbone			Insert(s)			
	Primer 1	Primer 2	Template	Primer 1	Primer 2	Template	Contents
pAL998	YT101	YT736	pAL956	YT737	YT738	pAL300	*gent*^{R}
pYT007	YTC012	YTC013	pZA12MCS	-	-	-	*fhs*$_{Cl}$
pYT036	YTC113	YTC013	pYT007	YTC111	YTC112	gDNA of AL1050	*folD*$_{Ec}$
pYT043	YTC134	YTC135	pAL998	YTC141	YTC142	gDNA of AL1050	*sdaA*
pYT044	YTC134	YTC135	pAL998	YTC143	YTC144	gDNA of AL1050	*sdaB*
pYT045	YTC134	YTC135	pAL998	YTC145	YTC146	gDNA of AL1050	*tdcB*
pYT046	YTC134	YTC135	pAL998	YTC147	YTC148	gDNA of AL1050	*tdcG*
pYT048	YTC134	YTC135	pZE12MCS	YTC154	YTC155	gDNA of AL1050	*gcvT, gcvH,* & *gcvP*
pYT079	YTC230	YTC249	pYT036	-	-	-	*fhs*$_{Aw}$
pYT085	YTC334	YTC335	pYT036	YTC330 YTC332	YTC331 YTC333	*fchA*$_{Cl}$* *folD*$_{Cl}$*	*fchA*$_{Cl}$ & *folD*$_{Cl}$
pYT086	YTC334	YTC339	pYT079	YTC336 YTC338	YTC337 YTC333	*fchA*$_{Aw}$* *folD*$_{Aw}$*	*fchA*$_{Aw}$ & *folD*$_{Aw}$
pYT099	YTC388	YTC389	pAL420	YTC390	YTC391	pZA23MCS	*kan*^{R}
pYT100	YTC001	YTC334	pYT099	YTC217	YTC131	pYT085	*fch*$_{Cl}$, *fchA*$_{Cl}$, & *folD*$_{Cl}$
pYT101	YTC230	YTC231	pZA23MCS	YTC399	YTC401	*fchA*$_{Cl}$*	*fchA*$_{Cl}$
pYT102	YTC230	YTC231	pZA23MCS	YTC402	YTC333	*folD*$_{Cl}$*	*folD*$_{Cl}$
pYT103	YTC334	YTC335	pYT036	YTC399	YTC330	*fchA*$_{Cl}$*	*fchA*$_{Cl}$
pYT104	YTC334	YTC335	pYT036	YTC400	YTC333	*folD*$_{Cl}$*	*folD*$_{Cl}$

AL1050 is an *E. coli* strain[100]. *synthesized gene (see Supplementary Table 1)

Table S2.6. Conversion efficiency from CO_2 to formate

Electrode	Efficiency (%)
In	**61.4**
Sn	**40.8**
Pb	**37.5**
C	**6.4**
Cu	**2.3**

In; Indium, Sn; Tin, Pd; Lead, C; Carbon, Cu; Copper

Table S2.7. Conversion efficiency of electricity to formate with the indium electrode in bacterial culture

Voltage (V)	Efficiency (%)
-1.6	56.5
-1.8	83.2
-2.0	91.0

Table S2.8. Abundance of ^{13}C-labeled carbons in L-serine and L-alanine

		Strain 6		**Strain 7**		**Strain 9**	
Substrate*		^{12}C	^{13}C	^{12}C	^{13}C	^{12}C	^{13}C
Serine†	M+0	65.0 ±0.3	62.8 ±1.2	65.3 ±0.8	33.0 ±6.0	63.5 ±7.0	40.2 ±1.0
	M+1	22.1 ±0.3	23.4 ±0.6	22.6 ±0.4	35.8 ±3.1	23.9 ±2.1	22.4 ±4.0
	M+2	10.4 ±0.2	11.0 ±0.5	10.5 ±0.5	21.4 ±1.8	10.2 ±3.2	21.4 ±1.4
	M+3	2.5 ±0.2	2.9 ±0.2	1.5 ±1.2	9.9 ±1.3	2.5 ±2.3	16.0 ±3.0
Alanine†	M+0	74.9 ±0.3	69.4 ±0.9	74.9 ±1.8	62.7 ±0.5	72.4 ±4.1	65.8 ±5.3
	M+1	17.0 ±0.3	21.3 ±0.7	17.3 ±1.7	24.4 ±0.2	17.6 ±3.8	22.2 ±3.4
	M+2	7.0 ±0.05	7.8 ±0.1	6.7 ±0.2	9.7 ±0.1	7.0 ±0.3	7.5 ±0.6
	M+3	1.1 ±0.02	1.5 ±0.2	1.1 ±0.06	3.2 ±0.4	3.0 ±1.8	4.5 ±2.5

*^{12}C and ^{13}C indicate unlabeled and ^{13}C-labaled substrates (formate and bicarbonate), respectively. †The samples were derivatized with MTBSTFA. The values are relative intensity of ^{13}C (%). N=3 biological replicates; errors represent s.d.

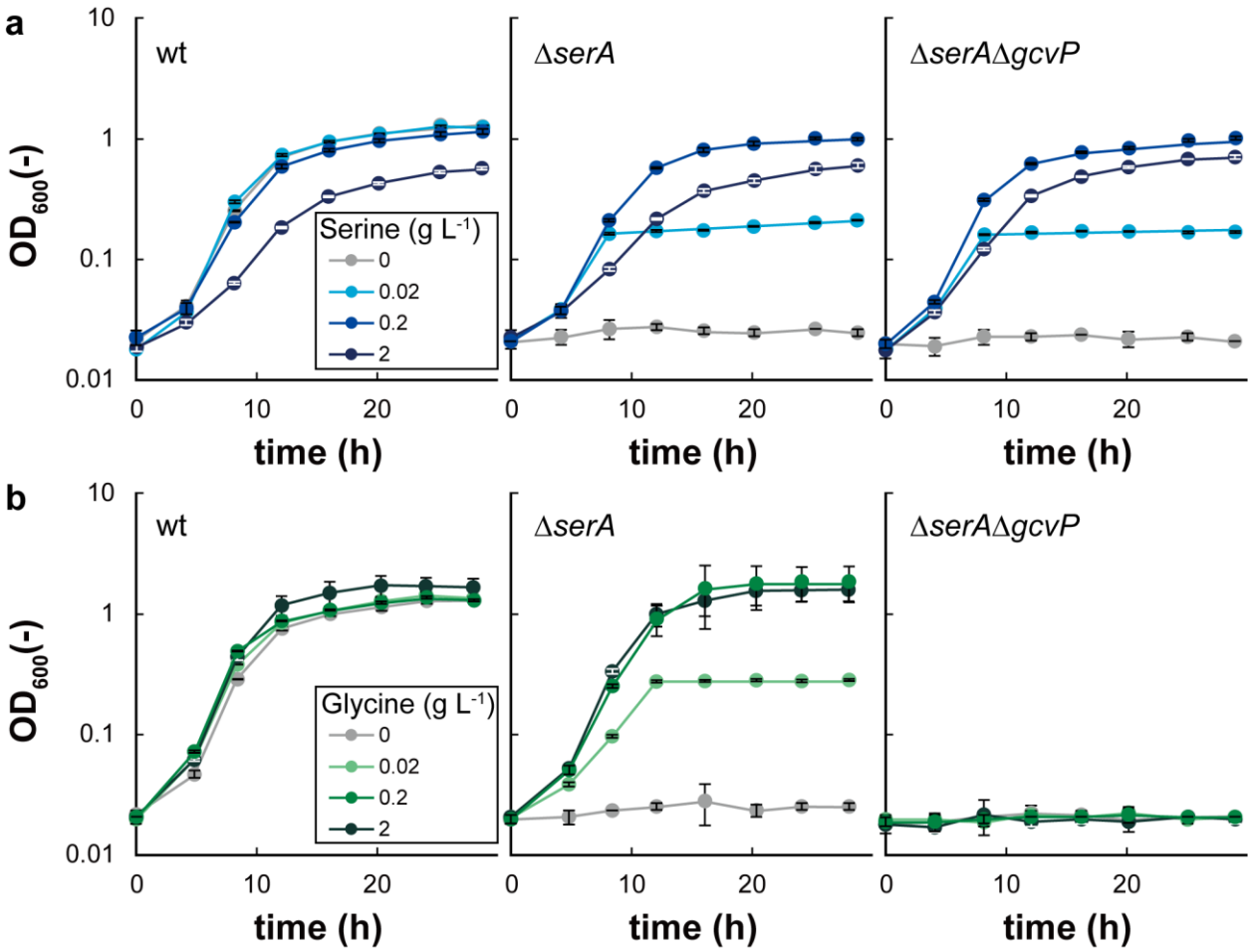


Fig. S2.1. Growth of the formate-assimilation strains on L-serine or glycine

Cells (Strain 1 control, **left**), Strain 2 (Strain 1 with Δ*serA*, **center**), and Strain 3 (Strain 1 with Δ*serA*Δ*gcvP*, **right**); **Table 2.1**) were cultured in 5 mL of M9 minimal media containing 4 g L^{-1} of glucose and various concentrations of L-serine (**a**) or glycine (**b**) at 37˚C. N=3 biological replicates; error bars represent s.d.

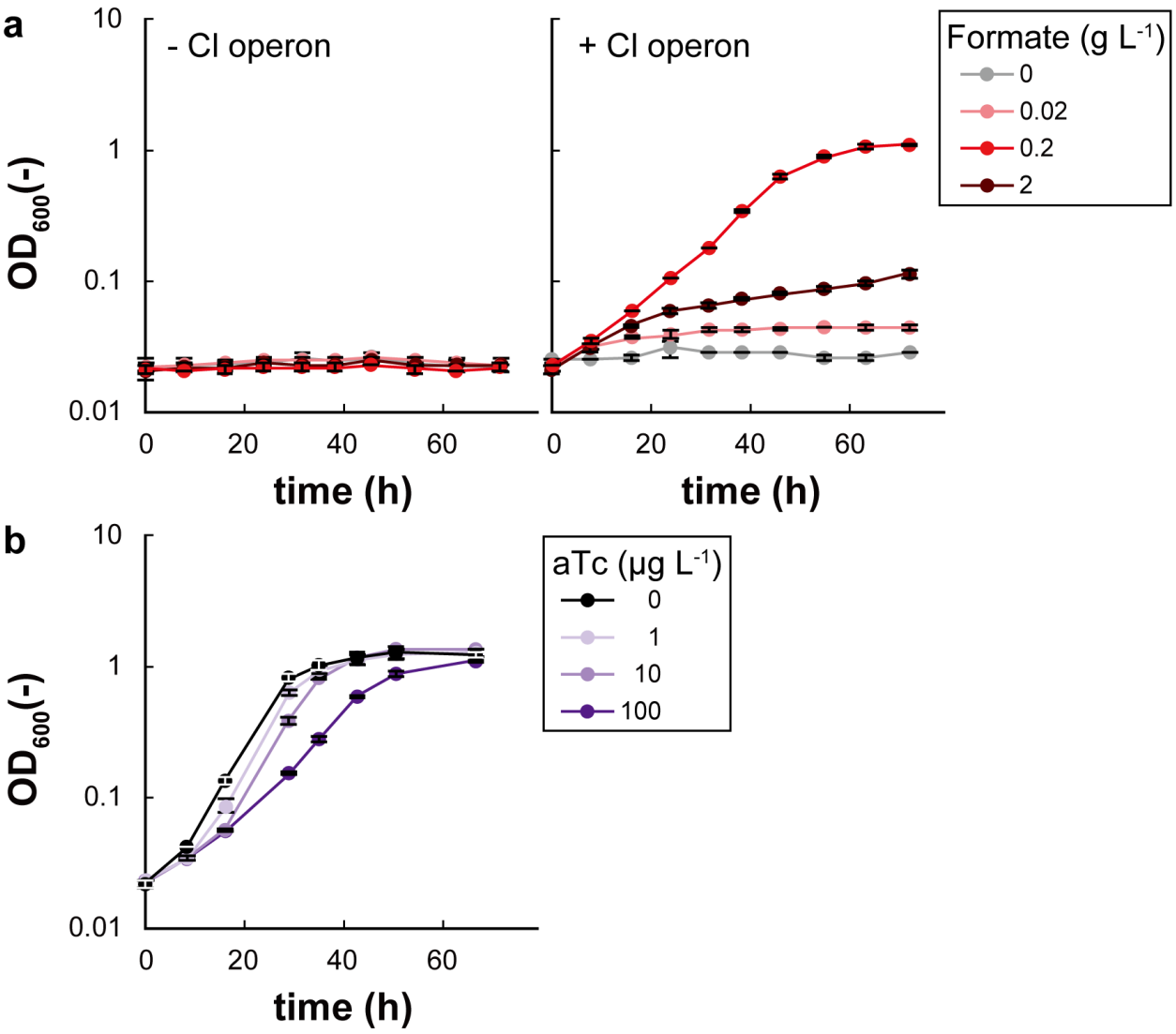


Fig. S2.2. Formate and aTc dependency of the Cl operon

(a) Cells (Strain 3 harboring pYT099 (Empty vector, **left**) and Strain 3 harboring pYT100 (Cl operon, **right**); **Table 2.1 and Supplementary Table 2.2**) were cultured in 5 mL of M9 minimal media including 4 g L^{-1} of glucose, 100 µg L^{-1} aTc and various concentrations of formate (0, 0.02, 0.2 or 2 g L^{-1}) at 37˚C. **(b)** Cells (Strain 3 harboring pYT100) were cultured in 5 mL of M9 minimal media including 4 g L^{-1} of glucose, 0.2 g L^{-1} of formate and various concentrations of aTc (0, 1, 10, or 100 µg L^{-1}) at 37˚C. N=3 biological replicates; error bars represent s.d.

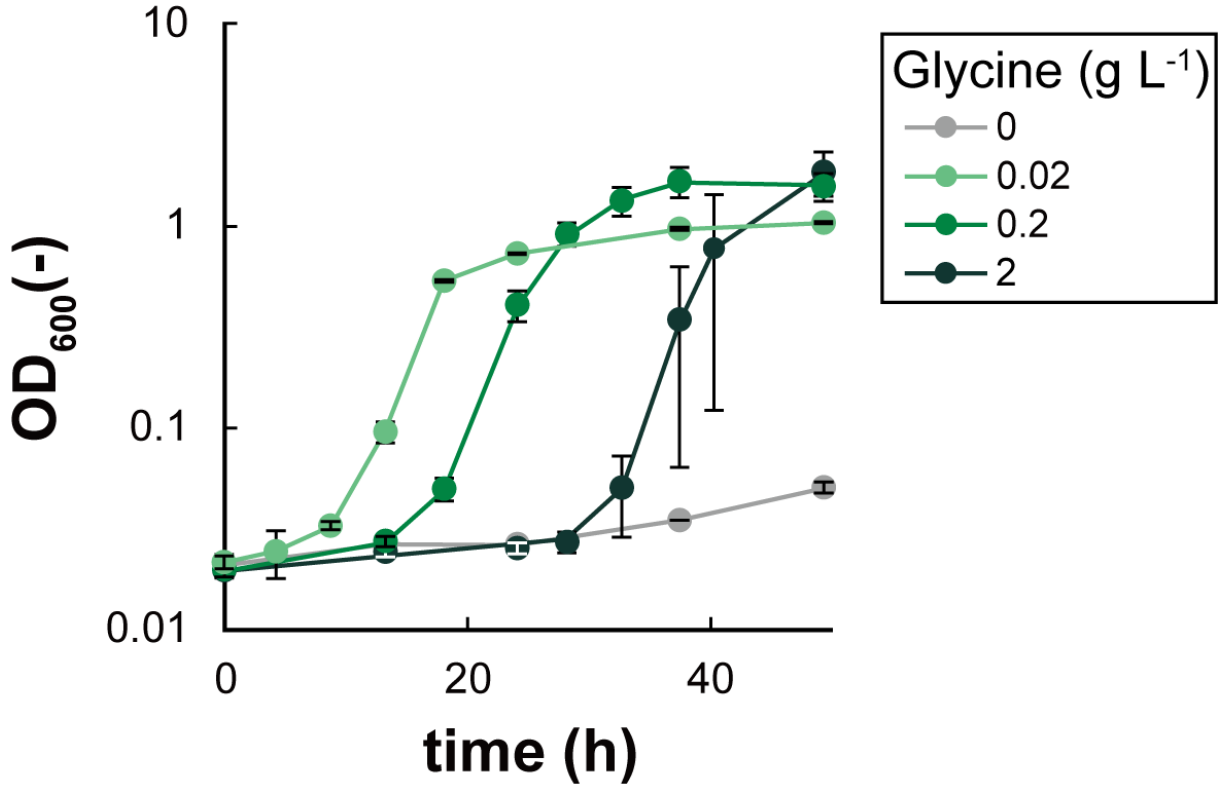


Fig. S2.3. Growth of the glycine auxotrophic strain on glycine

Cells (Strain 4 (Strain 1 with Δ*glyA*; **Table 2.1**)) were cultured in 5 mL of M9 minimal media including 4 g L^{-1} glucose and various concentrations of glycine (0, 0.02, 0.2 and 2 g L^{-1}) at 37˚C. N=3 biological replicates; error bars represent s.d.

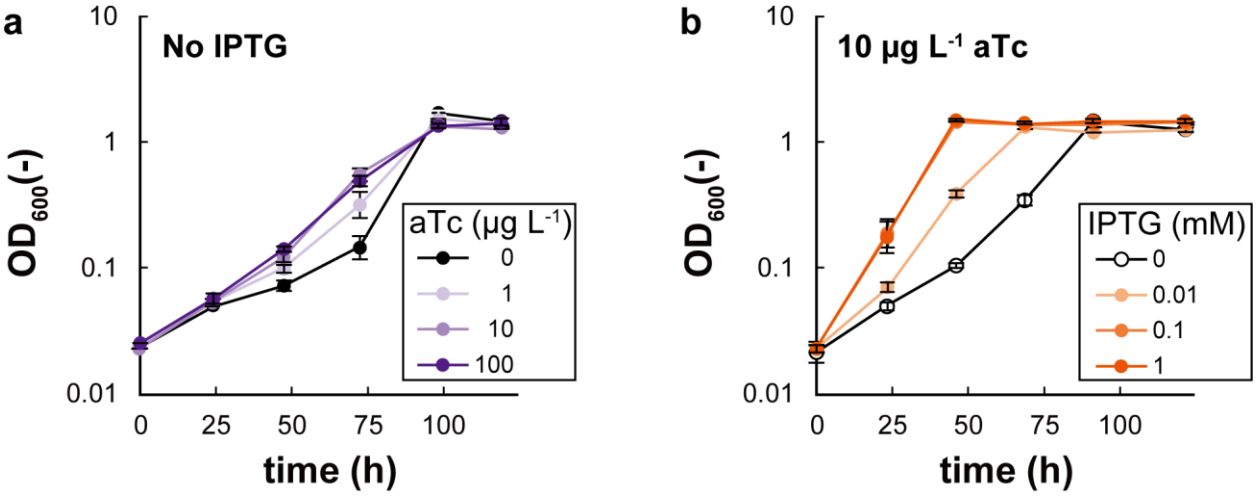


Fig. S2.4. Effects of aTc and IPTG concentrations on the growth of Strain 7

(a) Strain 7 (Strain 2 harboring pYT100 and pYT048; Table 2.1) was cultured in 5 mL of M9 minimal media including glucose (4 g L^{-1}), formate (0.2 g L^{-1}), and bicarbonate (50 mM) with various concentrations of aTc (0, 1, 10, or 100 μg L^{-1}) without IPTG at 37˚C. (b) Strain 7 was cultured with the same condition described in (a) except aTc was 10 μg $L^{-1}$and IPTG was various concentrations (0, 0.01, 0.1, or 1 mM). N=3 biological replicates; error bars represent s.d.

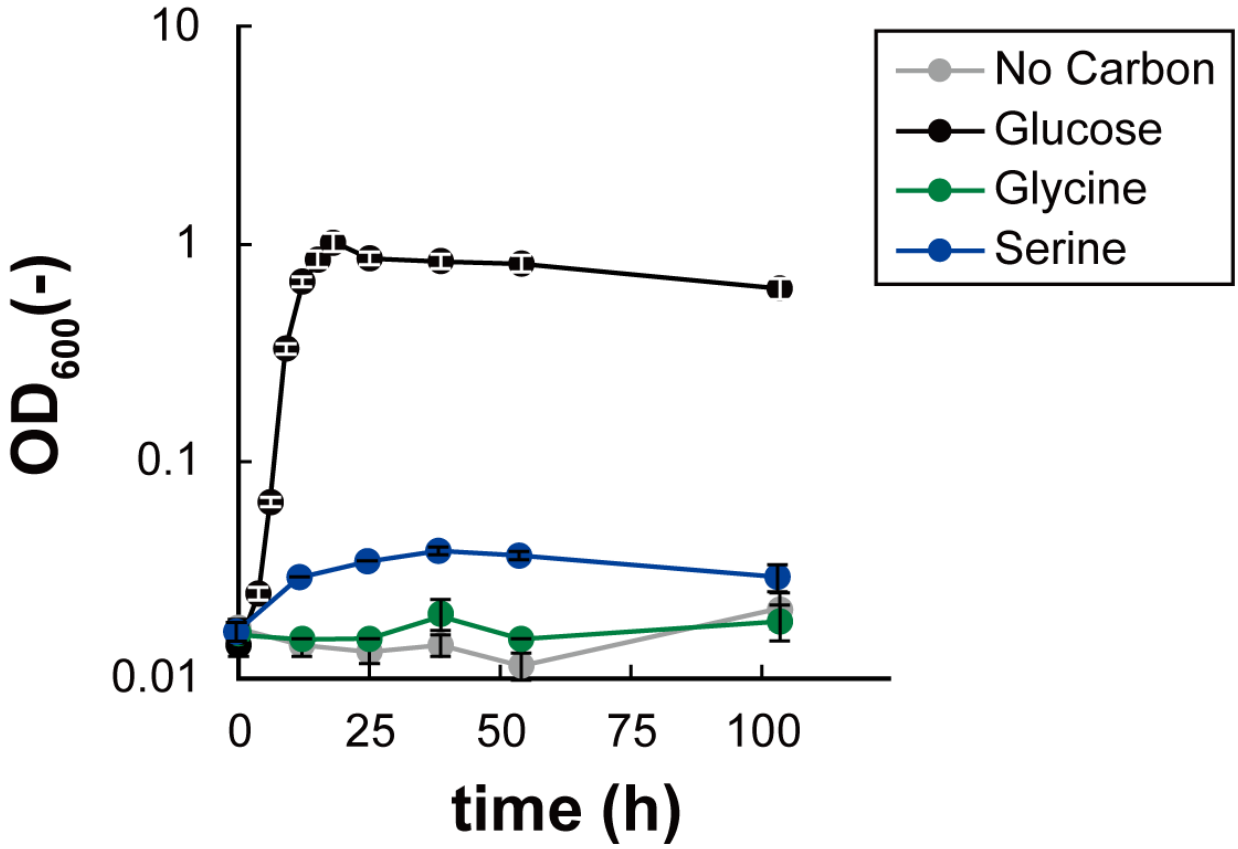


Fig. S2.5. Growth of Strain 1 with L-serine or glycine

Strain 1 was cultured in 5 mL M9 minimal media with 4 g L^{-1} of either glucose, L-serine, glycine or without a carbon source at 37˚C. N=3 biological replicates; error bars represent s.d.

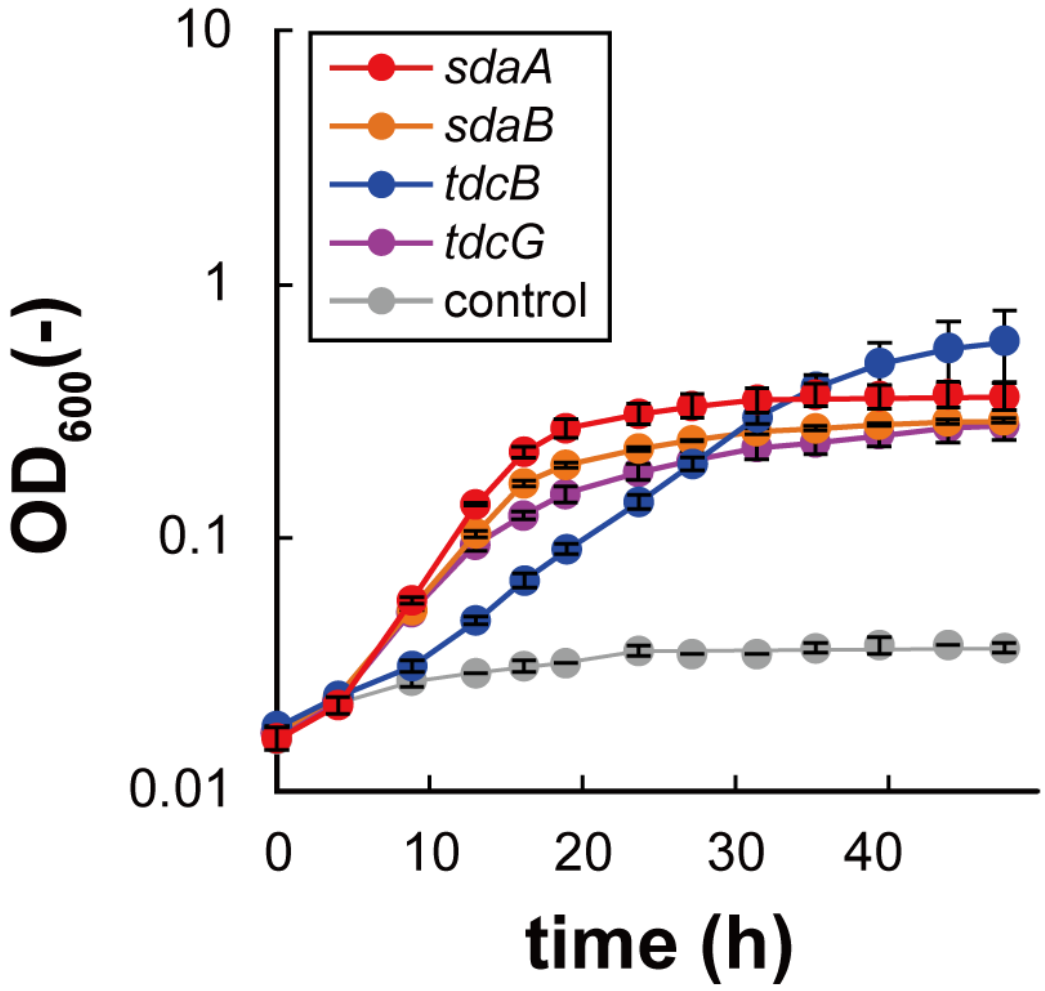


Fig. S2.6. Comparison of L-serine deaminases for L-serine assimilation

Cells (Strain 1 harboring pYT043 (*sdaA* (red)), pYT044 (*sdaB* (orange)), pYT045 (*tdcB* (blue)), pYT046 (*tdcG* (purple)), or pYT042 (control (empty; gray))) were cultured in 5 mL M9 minimal media with 4 g L^{-1} serine as the sole carbon source at 37°C. N=3 biological replicates; error bars represent s.d.

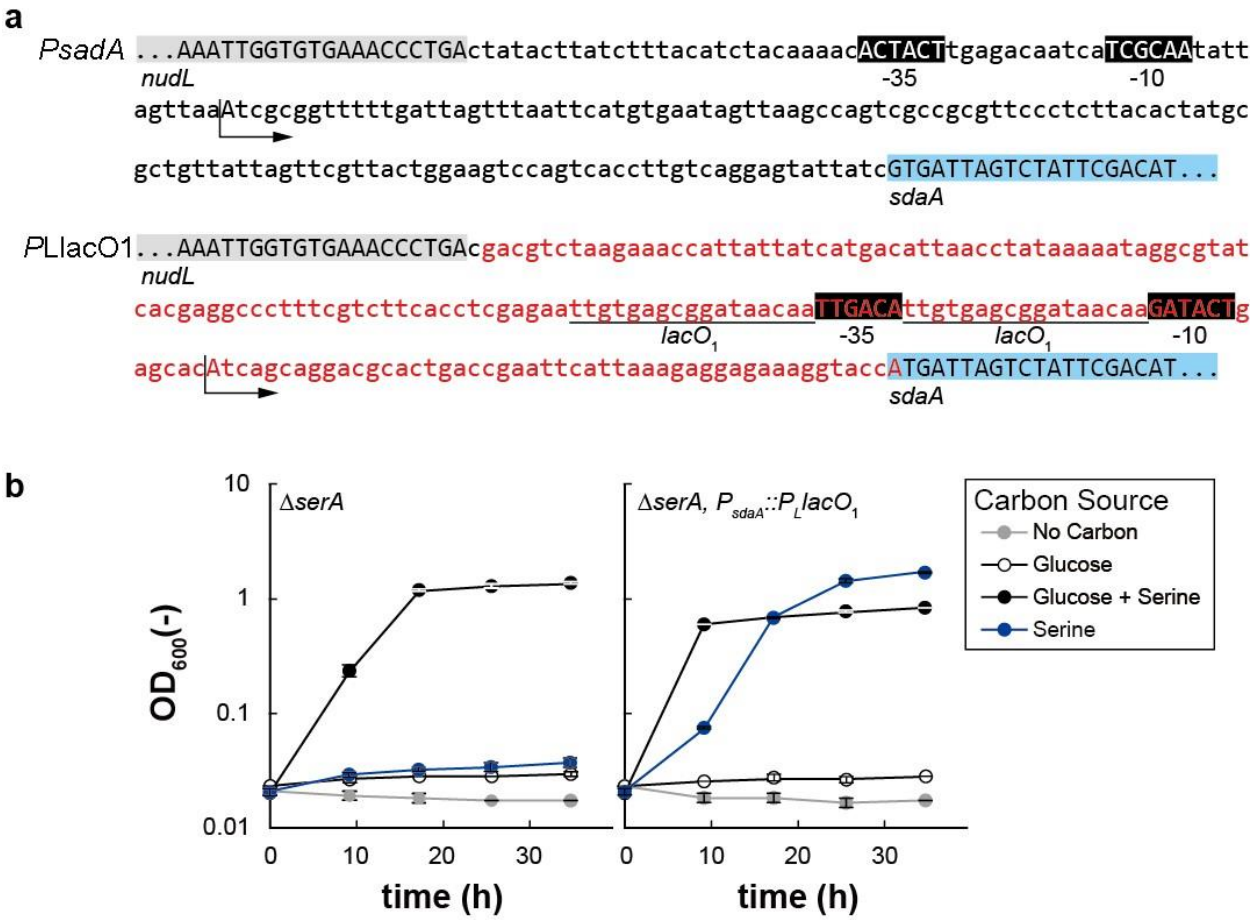


Fig. S2.7. Replacement of the genomic sdaA promoter (PsdaA) with IPTG-inducible *$P_L lacO_1$*

(a) The promoter sequences of the *sdaA* gene in Strain 2 (*P*sdaA$^+$, **top**) and Strain 8 ($P_L$lacO_1, **bottom**) are shown. The replaced regions are shown in red. **(b)** Strain 2 (**left**) and Strain 8 (**right**) were cultured in 5 mL M9 minimal media with 4 g L^{-1} glucose (open circle), 4 g L^{-1} glucose and 0.2 g L^{-1} L-serine (black circle), 4 g L^{-1} serine (blue circle), or no carbon source (grey) at 37˚C. N=3 biological replicates; error bars represent s.d.

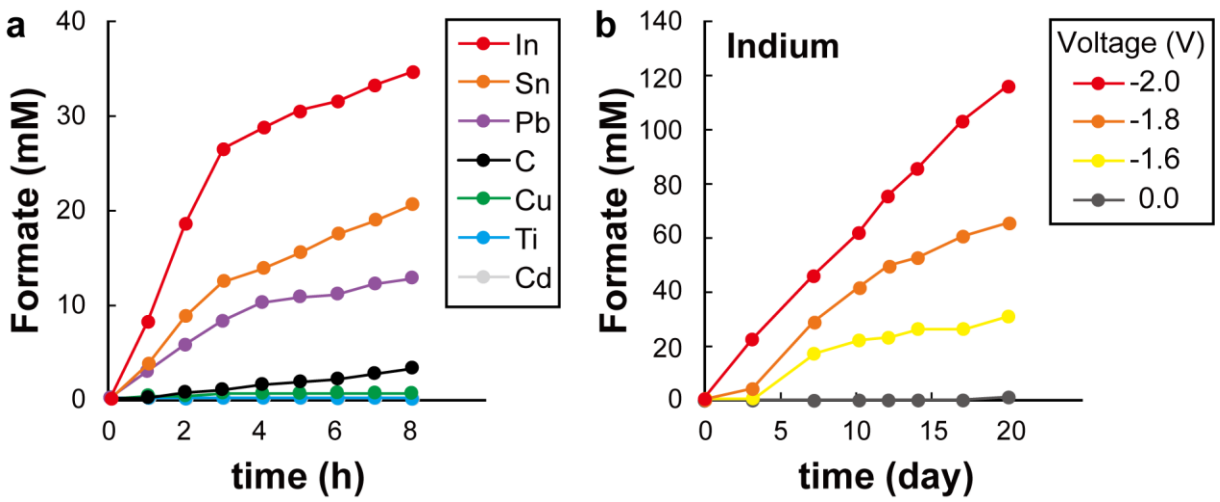


Fig. S2.8. Screening electrode for the electro-cultivation system

(**a**) Seven materials acting as working electrodes were screened at -2.0 V for their ability to form formate from CO_2(**b**) Various voltage conditions for the indium electrode were tested in bacterial culture. Formate production was measured. In; Indium, Sn; Tin, Pb; Lead, C; Carbon, Cu; Copper, Ti; Titanium, Cd; Cadmium.

2.7.1 Supplemental References

95 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645, doi: 10.1073/pnas.120163297 (2000).

96 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).

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

98 Tashiro, Y., Desai, S. H. & Atsumi, S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nature communications* **6**, 7488, doi:10.1038/ncomms8488 (2015).

99 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.org/10.1093/nar/25.6.1203 (1997).

100 Rodriguez, G. M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in *Escherichia coli*. *Nat Chem Biol* **10**, 259-65, doi: 10.1038/nchembio.1476 (2014).

Chapter 3: Adaptive laboratory evolution for improved tolerance of isobutyl acetate in *Escherichia coli*

This chapter is reproduced with permission from:

Matson, M. M., Cepeda, M. M., Zhang, A., Case, A. E., Kavvas, E. S., Wang, X., Carroll, A. L., Tagkopoulos, I., Atsumi, S. Adaptive laboratory evolution for improved tolerance of isobutyl acetate in *Escherichia coli*. Submitted to *Metabolic Engineering* August 14th, 2021. **3.0 Abstract**

Previously, we engineered *Escherichia coli* to produce isobutyl acetate (IBA). Titers (20 g/L) greater than the toxicity threshold (3 g/L) were achieved using layer-assisted production. To avoid this costly and complex method, we applied adaptive laboratory evolution (ALE) to *E. coli* for improved IBA tolerance. Over 37 rounds of selective pressure, 22 IBA-tolerant mutants were isolated. Remarkably, these mutants not only tolerate high IBA concentrations, but also have improved production. Using whole-genome sequencing followed by CRISPR/Cas9 mediated genome editing, we successfully elucidated the mutations (SNPs in *metH*, *rho* and deletion of *arcA*) that confer improved tolerance and production. We found that efficient supply of acetyl-CoA and altered transcriptional machinery improved IBA production in the evolved mutants. By combining select beneficial mutations and without phase separation, we constructed a strain capable of 3.2-fold greater production than the parent strain. These results highlight the impact improved tolerance has on the production capability of a biosynthetic system.

3.1 Introduction

Microbial production has emerged as a promising alternative to traditional petroleum-dependent industries and natural extraction processes. Microbes can be engineered to produce a diverse variety of bioactive and industrially relevant chemicals from renewable feedstocks, such as sugars and CO_2.[7,101-105] The bioproduction of esters has been of particular interest due to their versatile commercial applications as flavorants, fragrances, solvents, coatings, and paints[106]. A promising target ester is isobutyl acetate (IBA), a volatile, fruity smelling compound with great potential as a drop-in biofuel. Compared to similar alcohols fuels, such as isobutanol (ISO), IBA's higher volatility and lower polarity and hygroscopicity allow it to be more easily separated from aqueous cell cultures.[107] These properties, along with its ubiquitous use in the fragrances and food industries, make IBA a valuable commodity chemical and have driven efforts to produce the compound biologically.

Previously, we constructed an efficient IBA production pathway in *E. coli* **(Fig. 3.1).**[98,107] This pathway utilizes an alcohol-*O*-acyltransferase (ATF)[108] enzyme that condenses acetyl-CoA and ISO, derived from the 2-keto acid-based pathway[109], to form IBA.[107,110] A major challenge of biological IBA production is its inherent toxicity. The production strain JCL260[96], previously modified for ISO production, experiences sever growth-inhibition at 3 g/L IBA. This drastically limits production unless IBA is separated from the culture, such as by removal from the headspace via gas-stripping or from the media via extraction using an organic bilayer. Using hexadecane as a bilayer and an orthogonal acetate assimilating pathway for increased acetyl-CoA generation, titers (20 g/L) greater than the toxicity threshold (3 g/L) were achieved.[98] While these systems greatly increase IBA production, they also introduce greater production costs and complexities to large scale production. One potential solution to IBA toxicity lies in its low water solubility. At

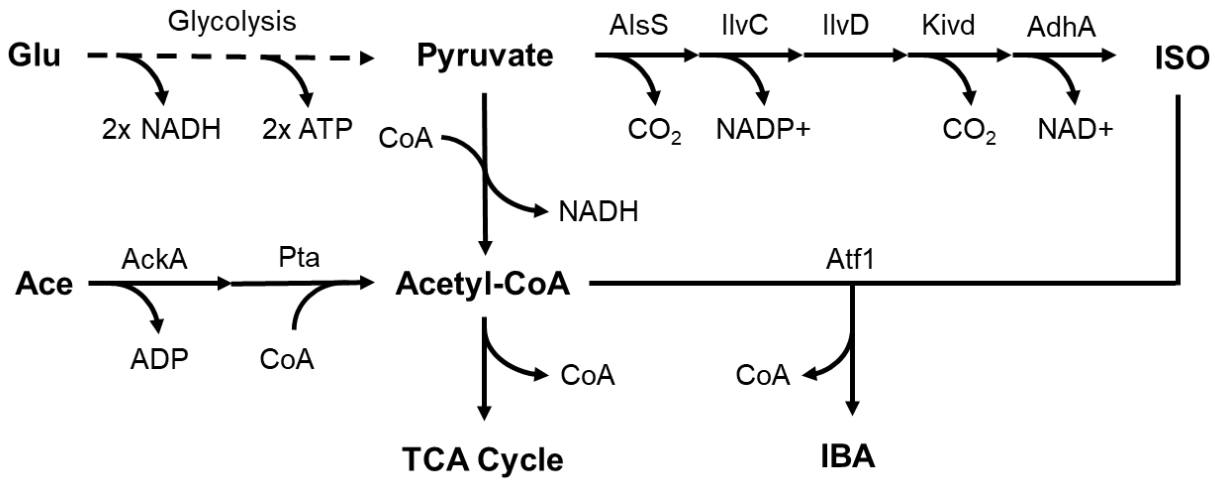


Fig. 3.1. IBA Production from glucose and acetate in E. coli.

Ace, acetate; AckA from *E. coli*, acetate kinase; AdhA, alcohol dehydrogenase from *Lactococcus lactis*; AlsS, acetolactate synthase from *Bacillus subtilis*; Atf1, alcohol-*O*-acetyl transferase from *Saccharomyces cerevisiae*; Glu, D-glucose; IBA; isobutyl acetate; IlvC, 2-hydroxy-3-ketol-acid reductoisomerase from *E. coli*; IlvD, dihydroxy-acid hydratase from *E. coli*; ISO, Isobutanol; Kivd, 2-keto acid decarboxylase from *L. lactis*; Pta, phosphate acetyltransferase from *E. coli.*

~8 g/L, IBA innately forms a bilayer in aqueous solutions.[111] Therefore, production utilizing natural bilayer formation is theoretically possible if *E. coli* can be adapted to grow under saturating IBA conditions.

Adaptive laboratory evolution (ALE) is a powerful tool for optimizing strain tolerance against environmental stressors.[59,60,112] Through adaptation, desired phenotypes are achieved through the accumulation of mutations over time to cope with a selective pressure, such as toxicity. These mutations result in improved cellular fitness under stress typically by changing gene expression, altering enzyme function, or re-wiring regulation.[113] Often, these changes are synergistic combinations that would otherwise be inaccessible or overlooked by rational design. Previously, ALE has been used to alleviate toxicity for several compounds in various organisms, including ethanol[114], 3-hydroxypropionic acid[115] and ISO[62] in *E. coli*, propionic acid[116] in *Saccharomyces cerevisiae*, and *p*-coumaric and ferulic acid[117] in *Pseudomonas putida* KT2440. These cases highlight the power and versatility of ALE for the overproduction of toxic commodities in microbial hosts.

In this study, we employed an ALE serial transfer method to evolve the ISO production host strain, JCL260, and isolated the IBA tolerant mutants M1-23. Excitingly, these evolved mutants also demonstrated enhanced IBA production. To elucidate the mutations beneficial to tolerance and production, the entire genomes of JCL260 and M1-23 were sequenced. The identified mutations were reconstructed individually and in combination in JCL260 to test for enhanced tolerance and production. We demonstrated the *metH, rho,* and *arcA* mutations alleviate IBA toxicity while *metH, rho* alone significantly boost production. These reconstructed strains were capable of better tolerance and higher production, respectively, than the evolved strains.

3.2 Methods

3.2.1. Reagents

All enzymes were purchased from New England Biolabs. All antibiotics were purchased from MilliporeSigma. All synthetic oligonucleotides were ordered from Integrated DNA Technologies. Sanger sequencing was performed by Genewiz. All chemicals were purchased from MilliporeSigma.

3.2.2. Strains and Plasmids

All strains in this study are listed in Table S3.1. JCL260 was used as the base strain for ALE. All plasmids and oligonucleotides used in this study are listed in Tables S3.2 and S3.3, respectively. Plasmids for IBA production and acetate assimilation were constructed using sequence and ligation independent cloning (SLIC, Table S3.4).[118] The constructed plasmids were verified by sequencing.

Genomic modifications were constructed using CRISPR-Cas9-mediated homologous recombination.[119] Donor DNA repair fragments for genomic modifications were constructed by PCR assembly or amplification from genomic DNA using primers listed in Tables S3.3 and S3.5. Plasmids encoding sgRNAs for CRISPR-Cas9-mediated homologous recombination were constructed with Q5 site-directed mutagenesis (NEB) using pTargetF (Addgene plasmid # 62226) as a template. Primer sets and templates used for plasmid construction are listed in Table S3.6. All genomic modifications were PCR and sequence verified.

3.2.3. Culturing

Overnight cultures were grown in 3 ml Luria broth (LB) media containing appropriate antibiotics. Antibiotic concentrations were as follows: kanamycin (50 μg ml^{-1}), ampicillin (200 μg ml^{-1}), tetracycline (20 μg ml^{-1}). IBA selection and tolerance were carried out in M9P medium (33.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 9.4 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$) including 5 g/L yeast extract (Research Products International); 10 g/L glucose (Fisher BioReagents); 1000 × A5 trace metal mix (2.86 g H_3BO_3, 1.81 g $MnCl_2 \cdot 4H_2O$, 0.079 g $CuSO_4 \cdot 5H_2O$, 49.4 mg $Co(NO_3)_2 \cdot 6H_2O$ per liter water). IBA production was carried out in the same M9P medium except supplemented with 50 g/L glucose. Optical densities were measured at 600 nm (OD_{600}) with a Synergy H1 hybrid plate reader (BioTek Instruments, Inc.).

3.2.4. ALE for IBA tolerance

An overnight culture of JCL260 was used to inoculate three 15 mL screw cap tubes at a starting OD_{600} of 0.1 in 5 mL of M9P media supplemented with 1.5 g/L IBA. Cultures were grown at 37 °C, 250 rpm for 24 hours then OD_{600} was taken. Out of three tubes, the culture with the highest OD_{600} was chosen as the inoculum for the next round of selection. If the culture OD_{600} exceeded 1.0, the concentration of IBA used in next round of selection was increased by 0.05-0.1 g/L. To isolate evolved mutants, cultures were struck onto M9P media plates containing 3.2 g/L IBA and the resulting colonies were designated with M#. After 22 rounds of selection with JCL260 as the inoculum, we isolated M01-07. P1 phage transduction[120] was carried out on M01-07 pooled with JCL260, followed by 5 rounds of selection and an additional P1 phage transduction. This resulted in M08. Using M08 as the inoculum, 10 rounds of selection resulted in M09-23.

3.2.5. Genome sequencing and data analysis

The genomic DNA of JCL260 and M01-23 was purified by Qiagen Blood and Cell Culture DNA Mini Kit. Whole-genome sequencing was performed by the Joint Genome Institute (JGI Proposal ID: 503669, NCBI Project ID: JCL260, 537595; EM1-23, 537607-537838). The sequenced JCL260 parental strain was aligned to the BW25113 genome (GenBank: CP009273.1) using *breseq*[121] to generate a reference sequence for downstream alignments. In addition to predicted mutations, the generated reference sequence includes *breseq* mutations described as unassigned junction evidence. The sequences of the 22 evolved mutants were then aligned to the reference sequence using *breseq*, which resulted in mutation predictions for each evolved strain. The sequences of the 22 evolved mutants were also aligned to an *E. coli* F plasmid (NCBI Reference Sequence: NC_002483.1). For all alignments involving evolved strains, *breseq* mutations described as unassigned junction evidence were not included in the set of analyzed mutations.

3.2.6. Tolerance Assays

Overnight cultures grown in LB media were centrifuged at 6,000 g for 1 min and the supernatant was discarded. The cell pellets were resuspended with M9P media. The concentrated cells were inoculated at an OD_{600} 0.1 into 5 mL of M9P media, containing various concentrations of the target chemical in Parafilm-wrapped 15 mL screw-cap test tubes. Media supplemented with acetic acid were pH adjusted to 7 with 5 M KOH. Cells were incubated at 37 °C, 250 rpm. OD_{600} readings were taken at 24 hours.

3.2.7. Fluorescence assays

Overnight cultures were inoculated at 1% in 3 mL of LB media. Cells were grown to an OD_{600} of 0.4-0.6 at 37 °C, 250 rpm then induced with 1 mM IPTG. The cultures were grown at 37 °C, 250 rpm, for 24 hours. Fluorescence emission was measured at 510 nm and with a Synergy H1 hybrid plate reader (BioTek Instruments, Inc.).

3.2.8. Production

Overnight cultures grown in LB media were inoculated at 1% in 20 mL of M9P in 250 mL screw-cap baffle flasks. Cells were grown to an OD_{600} of 0.4-0.6 at 37 °C, 250 rpm then induced with 1 mM IPTG and if applicable, supplemented with 10 g/L acetate, 1.7 g/L IBA, or 5 g/L ISO. The cultures were wrapped in Parafilm then grown at 30 °C, 250 rpm for the duration of the experiment.

3.2.9. Gas Chromatography Analysis

Concentrations of ISO and IBA were analyzed by GC–FID. The GC system is a GC-2010 with an AOC-20 S auto sampler and AOC-20i auto injector (Shimadzu). The column used was a DB-WAX capillary column (30 m length, 0.32 mm diameter and 0.5 μm film thickness; Agilent Technologies). The GC oven temperature was held at 225 °C, and the FID was held at 330 °C. The injection volume was 0.5 μl, injected at a 15:1 split ratio. Helium was used as the carrier gas. Retention times from samples were compared with standards.

To prepare samples for GC analysis, 1 ml of cell culture was centrifuged at 20,000 *g* for 10 min at 4 °C. 100 μL of culture supernatant was diluted with 900 μL MilliQ water.

3.2.10. Glucose and acetate analysis by High Performance Liquid Chromatography

Concentrations of acetate and glucose were analyzed by 20A HPLC (Shimazu) equipped with a differential refractive detector 10A and an Aminex fast acid analysis column (Bio-Rad). The mobile phase was 5 mM of H_2SO_4, maintained at a flow rate of 0.6 ml min^{-1} at 65 °C for 12.5 min.

To prepare samples for HPLC analysis, 1 mL of cell culture was centrifuged at 20,000g for 10 min at 4 °C. 40 μL of filtered culture supernatant was injected into the column for analysis.

3.3 Results and Discussion

3.3.1. Adaptive laboratory evolution and characterization of the evolved mutants

We chose JCL260[109], previously constructed for high ISO production, as the parent strain for evolution. JCL260 experiences severe growth-inhibition at 3 g/L IBA **(Fig. 3.2b)**. To improve JCL260's tolerance towards IBA, we employed a serial transfer method[114] where we inoculated M9P media supplemented with 1.5 g/L IBA and for each subsequent round we diluted the surviving culture into fresh media supplemented with gradually increasing concentrations of IBA **(Fig. 3.2a)**. After 22 rounds, we isolated the seven largest colonies on an LB agar plate containing 3.2 g/L IBA and designated them M01-07. To assess the impact of ALE on tolerance and production in the evolved mutants, we randomly selected M01 for initial screening. After sequencing the genome, we found that M01 is nearly identical to M02-07 (details in 3.3.3).

We cultured JCL260 and M01 in the presence of 2, 3, and if relevant, 4 and 5 g/L IBA to evaluate tolerance. Compared to JCL260, M01 showed increased growth in the absence of IBA and presence of 2 and 3 g/L IBA, slight growth at 4 g/L IBA, and no growth at 5 g /L IBA **(Fig 3.2b)**. To determine if ALE affected IBA production, we cultured JCL260 and M01 harboring the

IBA production plasmids, pAL603 ($P_L$lacO$_1$:*alsS-ilvCD*, $P_L$lacO$_1$:*kivd-adhA*; **Table S3.2**) and pAL685 ($P_L$lacO$_1$:*ATF1*, **Table S3.2**), in the absence of a solvent layer. This resulted in no ISO or IBA production in M01. To test gene expression under the control of $P_L$lacO$_1$, pAL682 ($P_L$lacO$_1$:*mRFP1* encoding a red fluorescence protein; **Table S3.2**) was introduced into JCL260 and M01. In the absence of IPTG, $P_L$lacO$_1$ was tightly repressed in JCL260 but not well repressed in M01 **(Fig. S3.1)**. These results suggest a loss of repression for $P_L$lacO$_1$ in M01 **(Fig. S3.1)**. It is well documented that plasmids with leaky expression suffer from instability, wherein the plasmid is maintained through cell proliferation but accumulates mutations that inactivate its genes.[122,123] This is especially true for plasmids harboring genes responsible for toxic metabolites[124], like the ISO and IBA pathway genes. The insufficient lac repression in M01 resulted in constitutive expression of the $P_L$lacO$_1$ controlled pathway genes, which likely triggered the loss of ISO and IBA production to avoid considerable growth burden or death during culturing.

We incorporated *lacI*^{q} into pAL685, generating pAL1114 **(Table S3.2)**, to re-establish lac repression in M01. With the additional repression, M01 was able to produce 2.6 g/L IBA, a 2.2-fold improvement over JCL260 which produced 1.2 g/L **(Fig. 3.2c)**. Conversely, JCL260 produced more ISO than M01, 8.6 g/L and 4.4 g/L respectively **(Fig. 3.2c)**. IBA production relies on the condensation of ISO and acetyl-CoA **(Fig. 3.1)**. Previously, we demonstrated that a lower ISO to IBA ratio corresponds to enhanced acetyl-CoA generation facilitating increased flux towards IBA.[98,107]

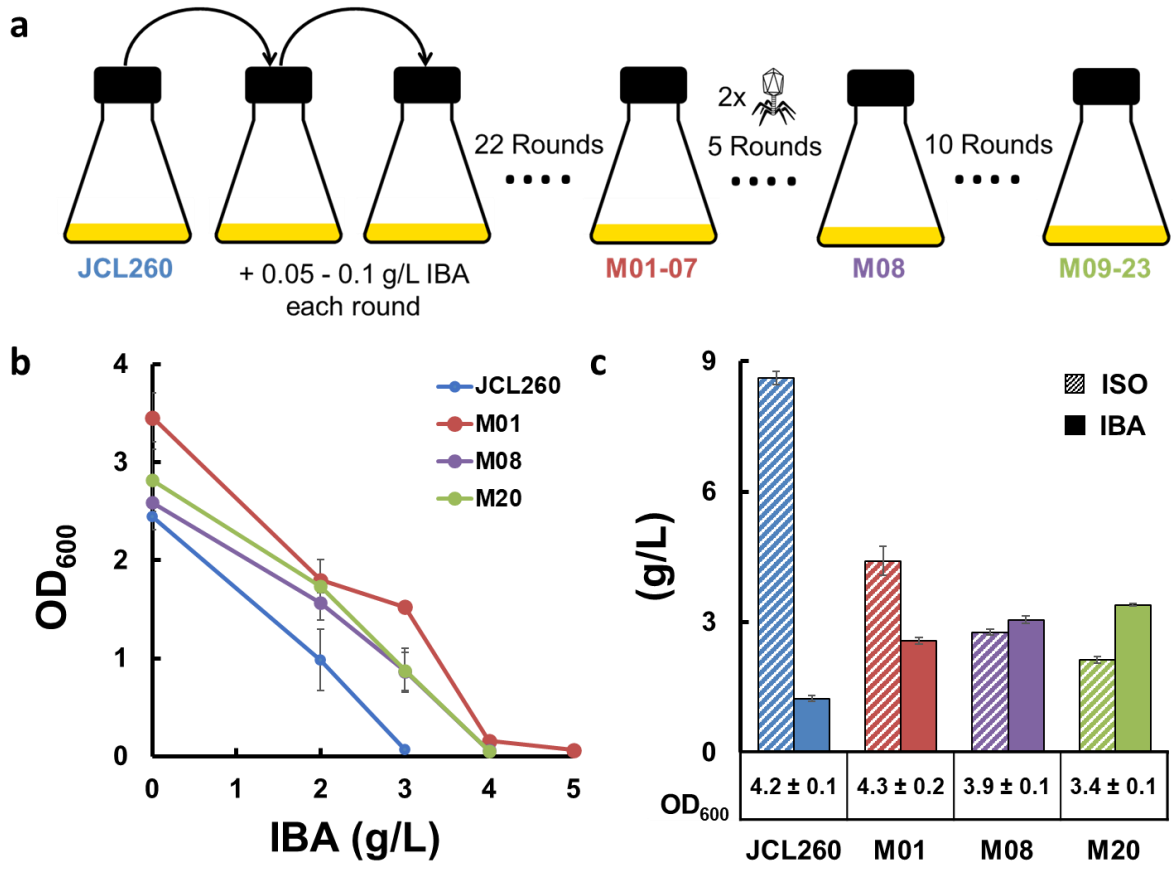


Fig. 3.2. Adaptive laboratory evolution to improve IBA tolerance.

(a) ALE experimental design and chronicle of evolution. Starting at 1.5 g/L IBA in M9P media, surviving cultures were inoculated into fresh media with increasing IBA concentrations, 0.05-0.1 g/L, every round. M01-07 were isolated after 22 rounds. M08 was isolated after two P1 phage transductions and 5 rounds starting at 3.2 g/L IBA. M09-23 were isolated after the final 10 rounds starting at 3.4 g/L IBA. **(b)** JCL260, M01, M08 and M20 were inoculated at OD_{600} 0.1 into 5 mL of M9P media containing 2, 3, and where relevant, 4 and 5 g/L IBA. OD_{600} was measured after 24 h. **(c)** IBA production (solid), ISO production (pattern) and final OD_{600} of JCL260, M01, M08 and M20 cultured in M9P media after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

To continue improving our evolved strains, we implemented genome shuffling[125], a powerful method to remove non-essential or deleterious mutations while generating novel combinations of beneficial mutations. For recombination, we pooled M01-07 with JCL260 and used P1 phage transduction.[120] We then combined the resulting colonies and used them as inoculum for 5 rounds of selection starting at 3.2 g/L IBA, which was followed by another transduction **(Fig 3.2a)**. This resulted in the evolved strain M08. Compared to M01, M08 exhibited similar growth at 2 g/L, reduced growth at 3 g/L and sever growth inhibition at 4 g/L **(Fig. 3.2b)**. Although M08 exhibited worse tolerance than M01, it produced more IBA, 3.1 g/L, and had a lower ISO to IBA ratio **(Fig. 3.2c)**. We resumed ALE to see if we could improve upon M08's tolerance while maintaining high production. After 10 rounds starting at 3.4 g/L IBA, we isolated M09-23 and randomly selected M20 for initial screening. After sequencing the genome, we found that M20 is nearly identical to M09-23 (details in 3.3.3). The growth of M08 and M20 were similar at all concentrations of IBA, but M20 produced 3.4 g/L IBA, a 2.8-fold improvement over JCL260 **(Fig. 3.2c)**. M20 had the lowest ISO to IBA ratio of all strains. In summary, our ALE method was successful at increasing IBA tolerance with the added benefit of enhanced production.

3.3.2. Further characterization of evolved mutant M20

To better understand the phenotypic changes that arose from ALE, we wanted to further probe the production capabilities of our evolved mutants. As the highest producing mutant among the strains initially screened, we chose M20 for these studies. First, we determined that JCL260 and M20 harboring only the ISO pathway produced the same amount of ISO, but M20 had a much higher final cell density **(Fig. S3.2)**. The denser growth of M20 may be a result of increased intracellular acetyl-CoA supply, as acetyl-CoA is the primary metabolite for growth and biomass

formation in *E. coli* **(Fig. 3.1)**. To further probe acetyl-CoA generation in the two strains, we monitored the conversion of 5 g/L ISO to IBA in JCL260 and M20 harboring only *ATF1*. JCL260 plateaued at 38% conversion (2.1 g/L IBA) at 8 h whereas M20 achieved 57% conversion (3.0 g/L IBA) by 10 h **(Fig 3.3a)**. The ability to produce more IBA from the same amount of ISO further supports M20 having a larger acetyl-CoA supply. Additionally, the two strains had the same final cell density, suggesting in the presence of ATF1 M20 utilizes acetyl-CoA for IBA production instead of enhancing growth.

Next, we wanted to gauge how the two strains respond to an enhanced acetyl-CoA supply. However, CoA thioesters are not natively transported across *E. coli*'s outer membrane[126], which prevented us from directly supplementing the cultures with exogenous acetyl-CoA. Instead, we increased intracellular acetyl-CoA via acetate assimilation.[98] We introduced the acetate kinase–phosphotransacetylase pathway plasmid, pAL956 ($P_L$lacO$_1$:*ackA-pta*; **Table S3.2**), along with the IBA production pathway plasmids into JCL260 and M20. After culturing the strains for 24 h with 50 g/L glucose and 10 g/L acetate, JCL260 produced 3.1 g/L IBA and consumed 29% of the glucose and 23% of the acetate whereas M20 produced 3.8 g/L IBA and consumed 28% of the glucose and only 11% of the acetate **(Fig. 3.3bc)**. With acetate feeding, JCL260 closely matched the production profile of M20, suggesting acetyl-CoA was previously limiting for production **(Fig. 3.3bc)**. Interestingly, using less acetate M20 produced slightly more IBA than JCL260. These results strongly indicate that M20 has enhanced acetyl-CoA generation from glucose, facilitating higher IBA production.

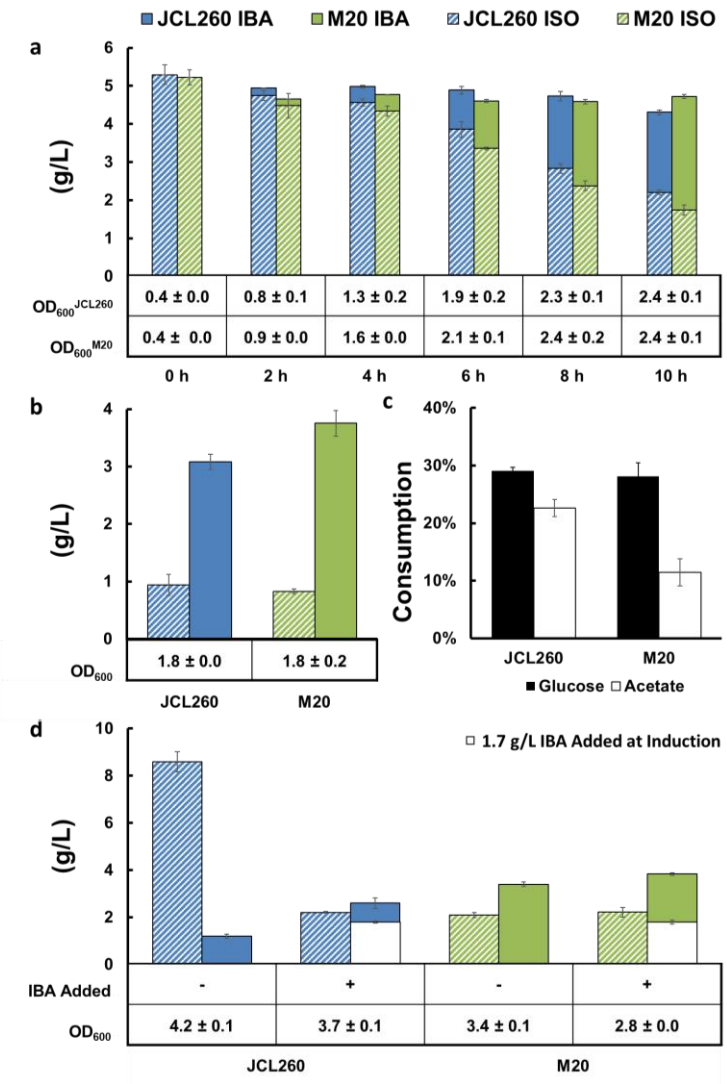


Fig. 3.3. ISO and IBA production from JCL260 and M20 under various conditions.

(a) IBA production (solid), ISO consumption (pattern), and final OD_{600} in M9P media with 5 g/L ISO added at induction. **(b-c)** IBA Production with acetate feeding. JCL260 and M20 with the IBA biosynthetic and acetate assimilation pathways were grown in M9P media with 50 g/L glucose and 10 g/L acetate. **(b)** IBA production (solid), ISO production (pattern) and OD_{600} after 24 h. **(c)** Glucose and acetate consumption. **(d)** IBA production (solid), ISO production (pattern) and OD_{600} after 24 h cultured in M9P media with and without 1.7 g/L IBA added at induction. N = 3, biological replicates; error bars represent s.d.

Lastly, we wanted to test the effect of IBA stress during growth and production for JCL260 and M20. To do so, we added 1.7 g/L IBA into the production media at induction. Under IBA stress, JCL260 showed a severe 68% decrease in ISO production whereas M20 had no decrease **(Fig. 3.3d)**. For both strains, the addition of IBA resulted in comparable decreases to IBA production (~30%) and final cell density (~15%). In the presence and absence of IBA, M20 had higher IBA production compared to JCL260. These results suggest the enhance acetyl-CoA supply seen in M20 is maintained under IBA stress and, in the presence of both the ISO pathway and ATF1, is siphoned from growth and biomass formation for production.

3.3.3. Whole genome sequencing and analysis of evolved strains

To identify the beneficial mutations associated with improved tolerance and production, the genomes of M01-23 and JCL260 were sequenced. Using *breseq*[121], the evolved mutant reads were aligned to both JCL260 and the *E. coli* reference genome sequence, BW25113 **(Table S3.1)**. This enabled us to identify the genetic mutations that arose during ALE **(Fig. 3.4, Table S3.5)**. M01-07 were found to have four primary mutations, two single-nucleotide polymorphisms (SNPs) in *rho* and *metH,* one IS30 insertion into *yjjY*, and one deletion in $lacI^q$ on F'. Two other mutations found, an intergenic mutation and insertion, did not carry forward in ALE, suggesting they do not benefit tolerance and production. M08 was found to have four additional mutations, three SNPs in *cysP*, *dctA*, and *yiiR* and one deletion including and in between *yahI* and *mhpC* **(Fig. 3.4, Table S3.7)**. M09-23 had one additional conserved mutation, an intergenic insertion between *glmU* and *atpC*. Mutations that persisted through ALE (*rho*, *metH*, *cysP*, *dctA*, *yiiR* SNPs, $lacI^q$ and *yahI-mhpC* deletions, *glmU*/*mhpC* indel, and IS30 insertion in *yjjY*) were targeted for further analysis.

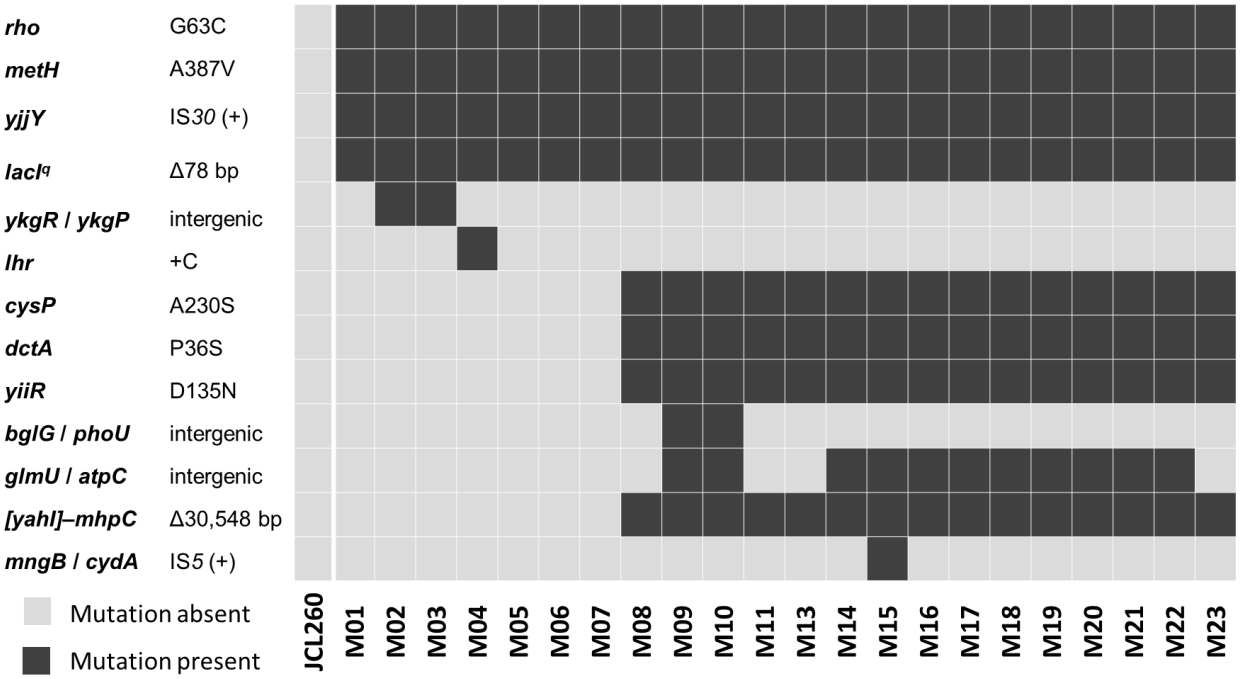


Fig. 3.4. Overview of mutations in M01-23.

Using JCL260 as comparison (grey), we identified mutations present (black) in the evolved strains that arose during ALE. (+) represents forward direction of the IS elements. More detailed information about these mutations is in Table S3.7.

3.3.4. CRISPR/Cas9 mediated M20 reversion

The reversion of a beneficial mutation would result in decreased tolerance and/or production. The M20 mutant contains all the conserved ALE mutations, hence was chosen for reversion. We attempted repair of the SNPs in M20 (*rho*, *metH*, *yiiR*, *cysP*, and *dctA*) using a CRISRP/Cas9 method.[119] Unfortunately, the modifications failed, likely due to the complete loss of lac repression in M20 **(Fig. S3.1)**. The CRISPR system relies on a lac promoter, P_{trc}[127], to control plasmid curing mechanisms **(Table S3.2)**. As an alternative to M20 reversion, we decided to reconstruct the mutations in JCL260.

3.3.5. Construction and characterization of M01-07 mutations

To identify the mutations responsible for improved IBA tolerance and production in strains M01-07, we introduced the *metH*, *rho*, Δ*arcA*, and Δ*lac* mutations individually and in combination in JCL260 using a CRISPR/Cas9 based method[119] **(Table S3.1, S3.5-7)**. The *metH* and *rho* SNPs were reconstructed with the addition of non-codon altering PAM scrambles in which either G in the NGG sequence was changed to avoid recognition by Cas9. The IS30 insertion in *yjjY* and the $lacI^q$ deletion required more nuanced genetic manipulations. Inserted near the start codon of *yjjY*, the IS30 sequence contains terminators at each end that reduce the expression of genes both upstream and downstream of the insertion site.[128] The *yjjY* gene is located between *ParcA* and *arcA*. The IS30 terminators in the *ParcA*:*arcA* reading frame are particularly strong, permitting less than 2% expression in downstream genes from external promoters.[128] Therefore, the IS30 insertion should disrupt *yjjY* expression and strongly reduce the expression of *arcA* from *ParcA* **(Fig. S3.3)**. Although we attempted to duplicate the IS30 insertion with the CRISRP/Cas9 method in JCL260 several times, we were unable to construct the mutant. Instead, to mimic the loss of *yjjY*

and *arcA* expression, we deleted the entire loci (Δ*arcA-yjjY-ParcA*, denoted Δ*arcA*). *lacI*^{q} is located on the F' plasmid, which is nearly identical to the genome in the region of interest. Hence we had to knockout the entire lac operon (Δ*lac*) on both the genome and F' plasmid to emulate Δ*lacI*^{q}.

The constructed individual and combined mutants were treated with 3 g/L IBA for 24 h **(Fig 3.5a)**. JCL260 and M01, which contains all the conserved ALE mutations from M01-M07, were used as benchmarks for improvement. All strains had identical growth in the absence of IBA. No individual mutant showed improved tolerance compared to JCL260. Conversely, all the combined mutants demonstrated better tolerance than JCL260. The *metH*, *rho*, Δ*arcA* triple mutant had the best tolerance phenotype, with a 1.7-fold higher final cell density than M01. Next, we examined how each mutant effected IBA production **(Fig. 3.5b)**. The *metH* and *rho* single mutants showed increased IBA production relative to M01, with the *metH, rho* double mutant producing the most IBA at 4 g/L. This represents a 1.5-fold improvement over M01 (2.6 g/L IBA), 1.2-fold improvement over M20 (3.4 g/L IBA), and a remarkable 3.2-fold improvement over JCL260 (1.2 g/L IBA). All mutant strains containing Δ*arcA* showed a decrease in IBA production.

Overall, when compared to JCL260, the most IBA tolerant strain (*metH*, *rho*, Δ*arcA*) demonstrated no production benefit and the best IBA producing strain (*metH, rho*) displayed a moderate improvement to tolerance. Having successfully identified the beneficial mutations in M01-07, we attempted to understand how these genes influence IBA tolerance and production. Unlike other toxic bioproducts, like ethanol, the mechanistic effects of IBA toxicity on metabolism have not been characterized. However, under a microscope, cells treated with 3 g/L IBA are elongated compared to cells grown in the absence of IBA **(Fig. S3.4)**. Notably, M01 cells appear remarkably longer than JCL260 cells. This suggests IBA inhibits cell division, particularly in M01.

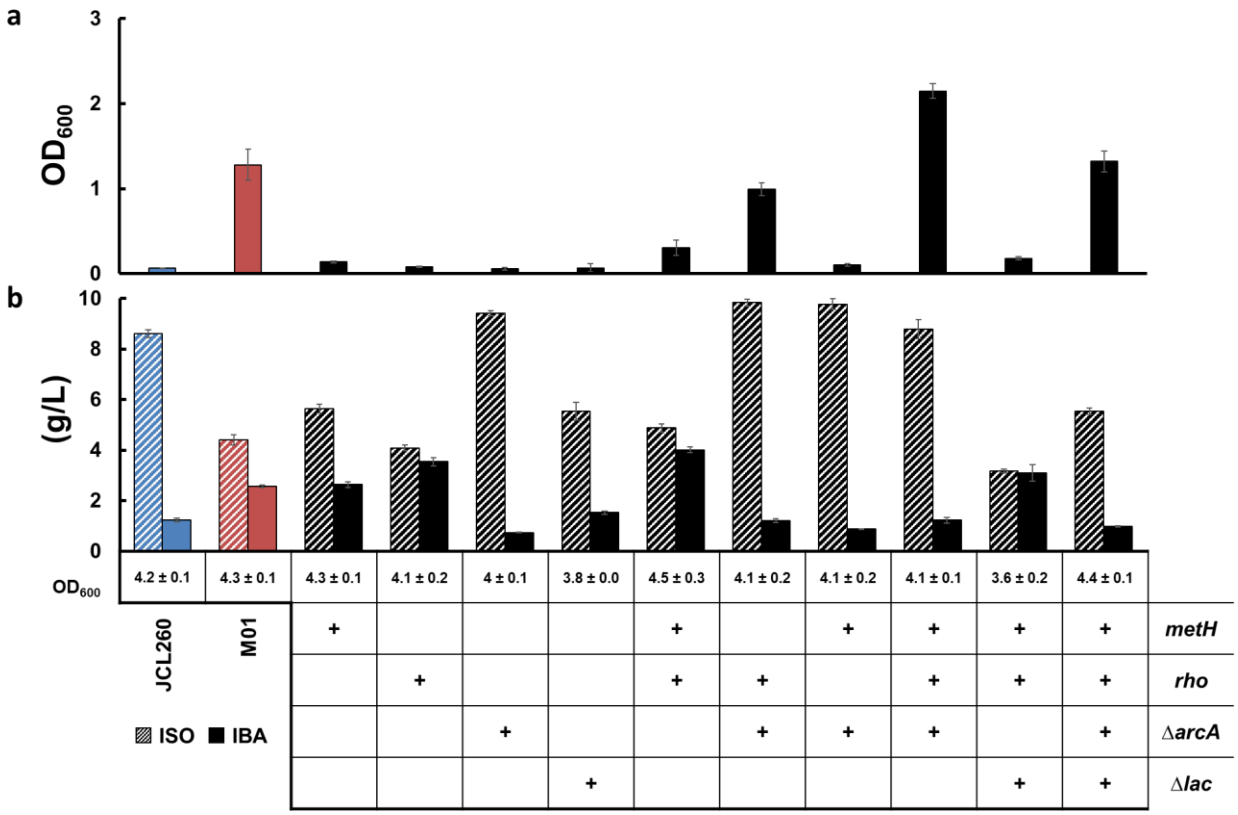


Fig. 3.5. Effect of M01-M07 mutations on tolerance and IBA production.

(a) Strains were inoculated at OD_{600} 0.1 into 5 mL of M9P media containing 3 g/L IBA. OD_{600} was measured after 24 h. **(b)** IBA production (solid), ISO production (pattern) and final OD_{600} of the strains cultured in M9P media after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

Cell membrane disruption is a frequent phenomenon seen with toxic and hydrophobic metabolites.[129-131] Hydrophobic metabolites are more lipid soluble, leading them to interfere with membrane composition and fluidity. This often causes substantial cell morphology modifications, like cell elongation. High concentrations of toxic metabolites are also known to disturb transcriptional machinery and damage DNA[132], which trigger stress responses and lead to cell elongation.[133,134]

ArcA is a transcription regulator that directly or indirectly effects of the expression of over 350 genes.[135,136] Primarily, ArcA represses genes associated with aerobic metabolism, such as the operons for β-oxidation of fatty acids, the TCA cycle, and glycolate/glyoxylate oxidation.[137] It has been shown that Δ*arcA* results in increased flux to the TCA cycle both in aerobic and anerobic conditions, while only in anerobic or microaerobic conditions does the deletion increase flux towards biomass formation.[138,139] This may explain why introducing Δ*arcA* improves tolerance but decreases IBA production in our microaerobic system **(Fig 3.5)**, as more acetyl-CoA is being used for growth or biomass generation rather than being condensed for IBA production. Interestingly, Δ*arcA* was shown to improve the production of acetyl-CoA derived products in BL21 strains.[136] This is attributed to lower catabolite repression and the presence of the acetyl-CoA synthetase pathway in BL21 strains, which allow for improved acetate assimilation compared to K12 strains.

Inactivated *arcA* also explains the higher glucose utilization and lower acetate utilization seen in M20 **(Fig. 3.3b)**. Δ*arcA* has been previously shown to increase glycolytic flux and pyruvate dehydrogenase (PDH) levels, the enzyme responsible for the conversion of pyruvate to acetyl-CoA.[140,141] These factors directly improve glucose utilization, allowing for more biomass formation from the same amount of glucose. In *E. coli*, the small RNA SdhX negatively regulates

ackA, which codes for an enzyme in the reversible acetate kinase–phosphotransacetylase pathway that interconverts acetate and acetyl-CoA **(Fig. 3.1)**. ShdX binds to the start codon of *ackA* mRNA, consequently blocking translation.[142] When *arcA* is deleted, SdhX levels increase which decrease AckA levels, thereby lowering the ability of M20 to assimilate and produce acetate.[138,142] This prevents significant carbon loss from acetyl-CoA during fermentation, furnishing more carbon for growth or biomass formation.

The Rho transcription terminator factor is coded by the *rho* gene. Rho is a global regulator of gene expression and has been shown to acutely inhibit PrpCD and CadA.[143] PrpCD are enzymes involved in propionate catabolism and their inhibition results in increased propionyl-CoA levels that subsequently inhibit PDH.[144] CadA is heavily involved in *E. coli*'s acid stress response.[145] Previous ALE studies have identified *rho* mutations that confer improved tolerance towards ethanol and L-serine[146,147], with one study attributing the improved tolerance to reduced Rho activity.[148] The compromised activity alleviated the effects of premature termination induced by ethanol. If IBA exerts similar effects on *E. coli's* transcriptional machinery, this, along with decrease inhibition of PrpCD and CadA, would explain the role of *rho* in improved tolerance **(Fig 3.5)**.

Additionally, diminished Rho activity was shown to increase membrane permeability and upregulate various small molecule transporters.[149] During production, improved efflux of IBA from the cytosol would reduce its intracellular accumulation, therefore decrease the toxic effects of IBA and allow for greater production. Rho inhibition also rendered DNA damage repair in *E. coli* ineffective and resulted in various physiology changes.[150] Since DNA damage can result in cell elongation[151], this implicates the *rho* mutant in the significant cell elongation seen in M01 under IBA stress. Using a microscope, we were able to confirm *rho* cells appear exceptionally

longer than the other strains **(Fig. S3.4)**. Interestingly, while the *metH*, *rho*, Δ*arcA* triple mutant also displays elongate cells, they do not appear as long as M01 or *rho*. Therefore, *metH* and Δ*arcA* may play a role in recovering the *rho* phenotype. As a global transcription terminator, Rho impacts numerous, non-related areas of metabolism. Thus, further study will be needed to elucidate the precise benefits of the *rho* mutation on IBA tolerance and production.

The *metH* gene encodes for a cobalamin-dependent methionine synthase that catalyzes the synthesis of methionine from homocysteine. Although it is unknown if the *metH* mutation (A387V) improves synthase activity, increased intracellular methionine stimulates *E. coli* growth at various temperatures (37-46°C) and alleviates the toxicity of acetic acid and other weak organic acids.[152] Methionine acts as a form of negative feedback inhibition on its own biosynthetic pathway, preventing the build-up of its toxic precursor, homocysteine.[153] Moreover, various stressors have been shown to halt ribosomes on AUG codons, which limits the availability of methionyl-tRNA^{Met} for protein synthesis.[154] Enhanced methionine generation was able to alleviate this deficiency.[148] This study also demonstrated a cooperative effect of increased methionine and reduced Rho activity on alleviating ethanol toxicity by separation of transcription from translation. These factors may explain the synergy seen between *metH*, *rho* for production and ultimately the role of *metH* in improving IBA tolerance **(Fig 3.5)**.

The *lacI*^{q} deletion on the F' plasmid and the lac operon deletion located within [*yahI*]-*mhpC* explain the stepwise loss of lac repression seen in the mutants **(Fig. S3.1)**. Since we had to reintroduce *lacI*^{q} into the system to re-establish control over $P_L$lacO$_1$ expression, it is hard to gauge the holistic effect of Δ*lac* on production. Using fluorescence as a guide, the incremental loss of lac repression led to a stepwise increase in mRFP1 expression under $P_L$lacO$_1$ compared to JCL260.

This may explain the slight improvement to production seen in Δ*lac* **(Fig 3.5b)**, as the ISO and IBA pathway genes are also under the control of $P_L lacO_1$ **(Table S3.2)**.

3.3.6. Tolerance to other stressors

To test if our most IBA tolerant strain (*metH*, *rho*, Δ*arcA*) also showed improved tolerance to other chemical stressors, we cultured JCL260 and the triple mutant in the presence of propyl acetate, octanoic acid, acetic acid, ISO, and 3-methyl-1-butanol. Interestingly, compared to JCL260 the triple mutant showed improved tolerance towards acetate esters and acids, but not higher chain alcohols **(Fig. 3.6)**. Curiously, a previous study evolving JCL260 demonstrated deletion of the TolC-efflux system improved ISO tolerance.[62] Reduced Rho activity has been shown to cripple TolC-efflux in *E. coli*[149], yet no improvement to ISO tolerance was seen in the triple mutant. Likely, multiple mutations to genes induced by the presence of higher alcohols and heavily involved in membrane composition and permeability are required to alleviate the toxic effects of ISO and 3-methyl-1-butanol.

3.3.7. Construction and characterization of M08-23 mutations

To identify the mutations responsible for improved IBA production in strains M08-23, we introduced the *cysP*, *dctA*, *yiiR*, and Δ*yahI-mhpC* mutations in combination with the *metH*, *rho*, and Δ*arcA* mutations using CRISPR/Cas9[119] **(Fig. 3.4, Table S3.1&3.7)**. The *cysP*, *dctA* and *yiiR* SNPs were reconstructed with the addition of non-codon altering PAM scrambles. The [*yahI*]-*mhpC* deletion (30Kb) was replicated by knocking out the entire region (denoted Δ*yahI-mhpC*). Each operon within the 30Kb region was also knocked out individually (Δ*yahIHJKLMNO*,

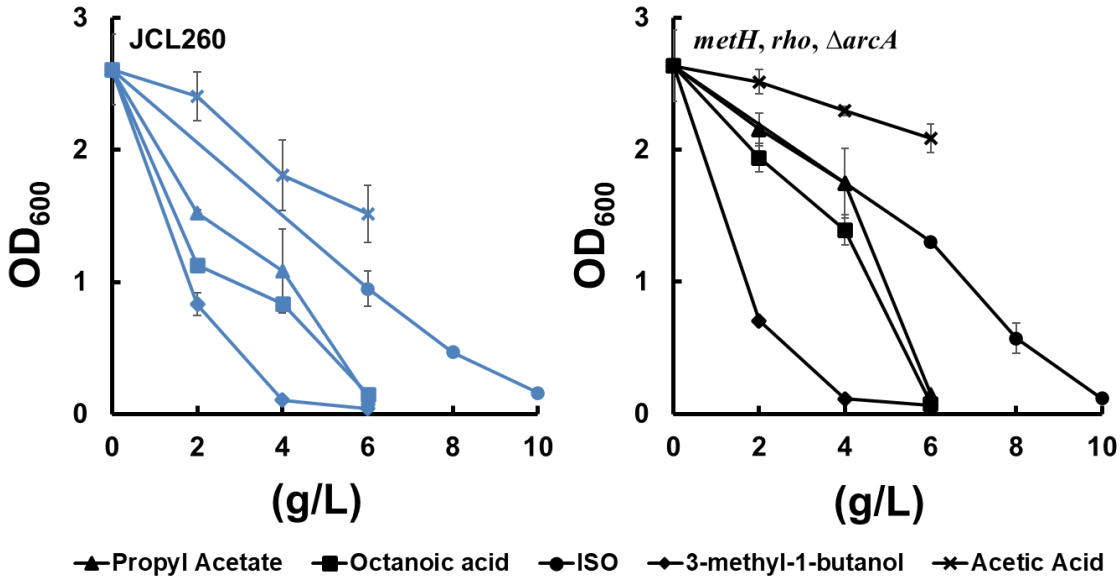


Fig. 3.6. Toxic effects of various chemicals on JCL260 (left, blue) and the triple mutant metH, rho, ΔarcA (right, black).

The strains were inoculated at OD_{600} 0.1 into 5 mL of M9P media containing the various chemicals at the indicated concentrations. OD_{600} was measured after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

denoted Δ*yah*; Δ*mhpRABC*, denoted Δ*mhp*; Δ*prpRBCDE,* denoted Δ*prp*, Δc*odBA*, denoted Δc*od*; Δ*cynRTSX*, denoted Δ*cyn*). The lac operon, also located between *yahI-mhpC,* was previously examined **(Fig. 3.5)**. Despite multiple attempts, we were unable to replicate the SNP between *glmU* and *atpC.*

We were only interested in the effect of the mutations on IBA production since the M09-23 mutants demonstrated better production, but worse tolerance than the M01-07 mutants. JCL260 and M20, which contains all the conserved ALE mutations, were used as benchmarks for improvement. The *cysP*, *dctA*, *yiiR* SNPs combination with *metH*, *rho* demonstrated decreased production compared to *metH, rho* alone. Similar to our previous results **(Fig 3.5)**, the addition of Δ*arcA* significantly reduced production **(Fig. 3.7)**. Neither Δ*yahI-mhpC* or any individual operon deletion, except Δ*lac*, showed improved IBA production compared to JCL260 **(Fig. 3.7, S3.5)**. No combination of Δ*yahI-mhpC* and its operons with *metH*, *rho*, and/or Δ*arcA* improved production either **(Fig. 3.7, S3.6)**. The lack of improvement in the M08-23 mutations may be due to metabolic discrepancies between the reconstructed and evolved strains. It has been shown that IS30 elements do not completely terminate the activity of neighboring genes, suggesting there may be residual activity of ArcA in the evolved mutants.[128]

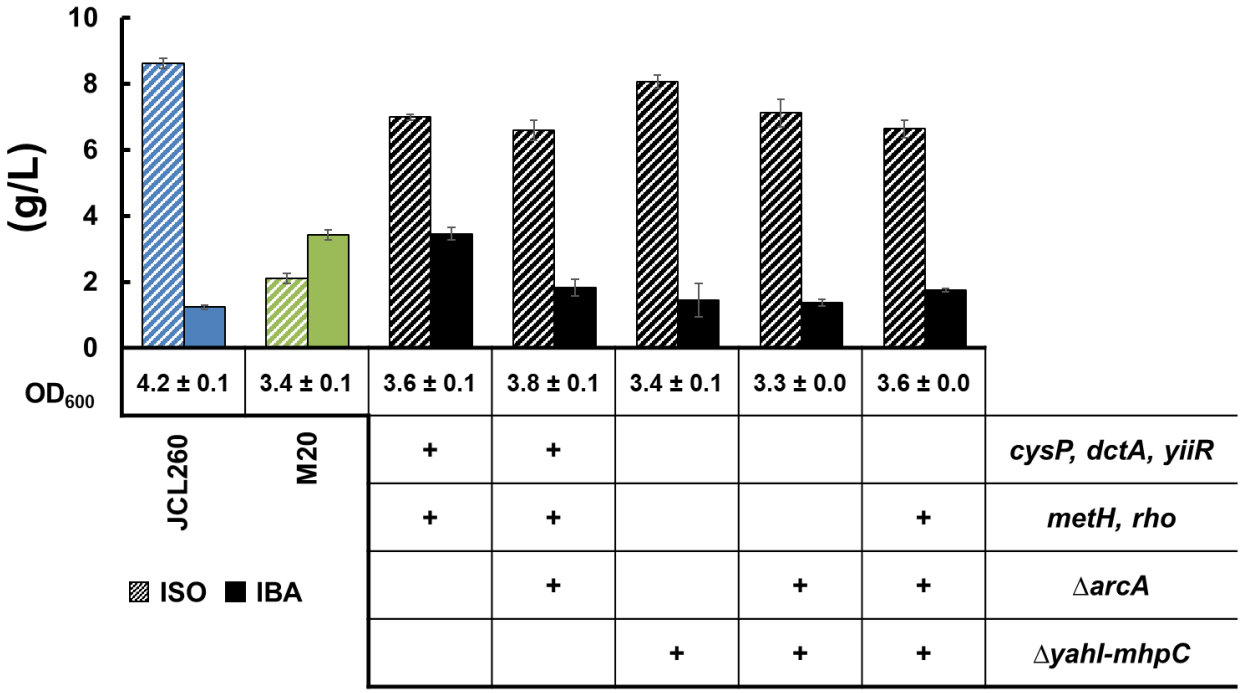


Fig 3.7. Effect of M01-M23 mutations on IBA production.

IBA production (solid), ISO production (pattern) and OD_{600} of the strains cultured in M9P media after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

3.4 Conclusion

Product toxicity plays a critical role in the industrial viability of a microbial production system.[60] Using ALE, we obtained 22 mutant strains capable of increased IBA tolerance and production. Through individual and combined mutation constructs, we were able to determine the synergistic effects of the mutations on tolerance and production. Through these reconstructions, we were able to create strains with higher production (*metH*, *rho*) and better tolerance (*metH*, *rho*, Δ*arcA*) than the evolved mutants. Additionally, we demonstrated our most IBA-tolerant strain also showed improved tolerance towards other acetate esters and acids. While these genes have been previously implicated in improved tolerance or production for other metabolites and organisms, further studies are required to understand their mechanistic roles in *E. coli* IBA metabolism.

3.5 References

7 Pontrelli, S. *et al.* *Escherichia coli* as a host for metabolic engineering. *Metab Eng* **50**, 16-46, doi:10.1016/j.ymben.2018.04.008 (2018).

59 Portnoy, V. A., Bezdan, D. & Zengler, K. Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* **22**, 590-594, doi:10.1016/j.copbio.2011.03.007 (2011).

60 Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution - principles and applications for biotechnology. *Microb Cell Fact* **12**, 64, doi:10.1186/1475-2859-12-64 (2013).

62 Atsumi, S. *et al.* Evolution, genomic analysis, and reconstruction of isobutanol tolerance in Escherichia coli. *Mol Syst Biol* **6**, 449, doi:10.1038/msb.2010.98 (2010).

96 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).

98 Tashiro, Y., Desai, S. H. & Atsumi, S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nature communications* **6**, 7488, doi:10.1038/ncomms8488 (2015).

101 Case, A. E. & Atsumi, S. Cyanobacterial chemical production. *J Biotechnol* **231**, 106-114, doi:10.1016/j.jbiotec.2016.05.023 (2016).

102 Sheng, H. *et al.* Metabolic Engineering of Microorganisms for the Production of Flavonoids. *Frontiers in bioengineering and biotechnology* **8**, 589069, doi:10.3389/fbioe.2020.589069 (2020).

103 Sun, L. & Alper, H. S. Non-conventional hosts for the production of fuels and chemicals. *Curr Opin Chem Biol* **59**, 15-22, doi:10.1016/j.cbpa.2020.03.004 (2020).

104 Keasling, J. *et al.* Microbial production of advanced biofuels. *Nat Rev Microbiol*, doi:10.1038/s41579-021-00577-w (2021).

105 Yang, D. *et al. Escherichia coli* as a platform microbial host for systems metabolic engineering. *Essays Biochem* **65**, 225–246, doi:10.1042/EBC20200172 (2021).

106 Carroll, A. L., Desai, S. H. & Atsumi, S. Microbial production of scent and flavor compounds. *Curr Opin Biotechnol* **37**, 8-15, doi:10.1016/j.copbio.2015.09.003 (2016).

107 Rodriguez, G. M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in *Escherichia coli*. *Nat Chem Biol* **10**, 259-265, doi:10.1038/nchembio.1476 (2014).

108 Verstrepen, K. J. *et al.* Expression levels of the yeast alcohol acetyltransferase genes *ATF1*, *Lg-ATF1*, and *ATF2* control the formation of a broad range of volatile esters. *Appl Environ Microbiol* **69**, 5228-5237, doi:10.1128/AEM.69.9.5228-5237.2003 (2003).

109 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).

110 Tai, Y. S., Xiong, M. & Zhang, K. Engineered biosynthesis of medium-chain esters in *Escherichia coli*. *Metab Eng* **27**, 20-28, doi:10.1016/j.ymben.2014.10.004 (2015).

111 Riemenschneider, W. Esters, Organic. *Ullmann's Encyclopedia of Industrial Chemistry*, doi:10.1002/14356007.a09_565 (2000).

112 Lee, S. & Kim, P. Current Status and Applications of Adaptive Laboratory Evolution in Industrial Microorganisms. *J Microbiol Biotechnol* **30**, 793-803, doi:10.4014/jmb.2003.03072 (2020).

113 Conrad, T. M., Lewis, N. E. & Palsson, B. O. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol* **7**, 509, doi:10.1038/msb.2011.42 (2011).

114 Yomano, L. P., York, S. W. & Ingram, L. O. Isolation and characterization of ethanol-tolerant mutants of *Escherichia coli* KO11 for fuel ethanol production. *J Ind Microbiol Biotechnol* **20**, 132-138, doi: 10.1038/sj.jim.2900496 (1998).

115 Nguyen-Vo, T. P. *et al.* Development of 3-hydroxypropionic-acid-tolerant strain of *Escherichia coli* W and role of minor global regulator *yieP*. *Metab Eng* **53**, 48-58, doi:10.1016/j.ymben.2019.02.001 (2019).

116 Xu, X., Williams, T. C., Divne, C., Pretorius, I. S. & Paulsen, I. T. Evolutionary engineering in *Saccharomyces cerevisiae* reveals a TRK1-dependent potassium influx mechanism for propionic acid tolerance. *Biotechnol Biofuels* **12**, 97, doi:10.1186/s13068-019-1427-6 (2019).

117 Mohamed, E. T. *et al.* Adaptive laboratory evolution of *Pseudomonas putida* KT2440 improves p-coumaric and ferulic acid catabolism and tolerance. *Metab Eng Commun* **11**, e00143, doi:10.1016/j.mec.2020.e00143 (2020).

118 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).

119 Jiang, Y. *et al.* Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **81**, 2506-2514, doi:10.1128/AEM.04023-14 (2015).

120 Thomason, L. C., Costantino, N. & Court, D. L. *E. coli* genome manipulation by P1 transduction. *Curr Protoc Mol Biol* **Chapter 1**, Unit 1 17, doi:10.1002/0471142727.mb0117s79 (2007).

121 Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* **1151**, 165-188, doi:10.1007/978-1-4939-0554-6_12 (2014).

122 Tyo, K. E., Ajikumar, P. K. & Stephanopoulos, G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat Biotechnol* **27**, 760-765, doi:10.1038/nbt.1555 (2009).

123 Saleski, T. E. *et al.* Optimized gene expression from bacterial chromosome by high-,throughput integration and screening. *Sci Adv* **7**, 7, doi:10.1126/sciadv.abe1767 (2021).

124 Gruber, D. F., Pieribone, V. A., Porton, B. & Kao, H. T. Strict regulation of gene expression from a high-copy plasmid utilizing a dual vector system. *Protein Expr Purif* **60**, 53-57, doi:10.1016/j.pep.2008.03.014 (2008).

125 Gong, J., Zheng, H., Wu, Z., Chen, T. & Zhao, X. Genome shuffling: Progress and applications for phenotype improvement. *Biotechnol Adv* **27**, 996-1005, doi:10.1016/j.biotechadv.2009.05.016 (2009).

126 Black, P. N. & DiRusso, C. C. Transmembrane movement of exogenous long-chain fatty acids: proteins, enzymes, and vectorial esterification. *Microbiol Mol Biol Rev* **67**, 454-472, doi:10.1128/MMBR.67.3.454-472.2003 (2003).

127 Brosius, J., Erfle, M. & Storella, J. Spacing of the -10 and -35 regions in the tac promoter. Effect on its *in vivo* activity. *J Biol Chem* **260**, 3539-3541, doi: 10.1073/pnas.88.5.1760 (1985).

128 Dalrymple, B. & Arber, W. The characterization of terminators of RNA transcription on IS30 and an analysis of their role in IS element-mediated polarity. *Gene* **44**, 1-10, doi:10.1016/0378-1119(86)90036-3 (1986).

129 Alakomi, H. L. *et al.* Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Appl Environ Microbiol* **66**, 2001-2005, doi:10.1128/AEM.66.5.2001-2005.2000 (2000).

130 Huffer, S., Clark, M. E., Ning, J. C., Blanch, H. W. & Clark, D. S. Role of alcohols in growth, lipid composition, and membrane fluidity of yeasts, bacteria, and archaea. *Appl Environ Microbiol* **77**, 6400-6408, doi:10.1128/AEM.00694-11 (2011).

131 Royce, L. A., Liu, P., Stebbins, M. J., Hanson, B. C. & Jarboe, L. R. The damaging effects of short chain fatty acids on *Escherichia coli* membranes. *Appl Microbiol Biotechnol* **97**, 8317-8327, doi:10.1007/s00253-013-5113-5 (2013).

132 Kang, Y. *et al.* Analysis of alcohol-induced DNA damage in *Escherichia coli* by visualizing single genomic DNA molecules. *Analyst* **141**, 4326-4331, doi:10.1039/c6an00616g (2016).

133 Wehrens, M. *et al.* Size Laws and Division Ring Dynamics in Filamentous *Escherichia coli* cells. *Curr Biol* **28**, 972-979 e975, doi:10.1016/j.cub.2018.02.006 (2018).

134 Wong, I. *et al.* An agar gel membrane-PDMS hybrid microfluidic device for long term single cell dynamic study. *Lab Chip* **10**, 2710-2719, doi:10.1039/c004719h (2010).

135 Salmon, K. *et al.* Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. *J Biol Chem* **278**, 29837-29855, doi:10.1074/jbc.M213060200 (2003).

136 Liu, X. & De Wulf, P. Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *J Biol Chem* **279**, 12588-12597, doi:10.1074/jbc.M313454200 (2004).

137 Park, D. M., Akhtar, M. S., Ansari, A. Z., Landick, R. & Kiley, P. J. The bacterial response regulator ArcA uses a diverse binding site architecture to regulate carbon oxidation globally. *PLoS genetics* **9**, e1003839, doi:10.1371/journal.pgen.1003839 (2013).

138 Nikel, P. I., Pettinari, M. J., Ramirez, M. C., Galvagno, M. A. & Mendez, B. S. *Escherichia coli arcA* mutants: metabolic profile characterization of microaerobic cultures using glycerol as a carbon source. *J Mol Microbiol Biotechnol* **15**, 48-54, doi:10.1159/000111992 (2008).

139 Perrenoud, A. & Sauer, U. Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in *Escherichia coli*. *J Bacteriol* **187**, 3171-3179, doi:10.1128/JB.187.9.3171-3179.2005 (2005).

140 Nizam, S. A., Zhu, J., Ho, P. Y. & Shimizu, K. Effects of arcA and arcB genes knockout on the metabolism in Escherichia coli under aerobic condition. *Biochemical Engineering Journal* **44**, 240-250, doi:10.1016/j.bej.2008.12.017 (2009).

141 Vemuri, G. N., Eiteman, M. A. & Altman, E. Increased recombinant protein production in *Escherichia coli* strains with overexpressed water-forming NADH oxidase and a deleted ArcA regulatory protein. *Biotechnol Bioeng* **94**, 538-542, doi:10.1002/bit.20853 (2006).

142 De Mets, F., Van Melderen, L. & Gottesman, S. Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc Natl Acad Sci U S A* **116**, 1043-1052, doi:10.1073/pnas.1815288116 (2019).

143 Cardinale, C. J. *et al.* Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**, 935-938, doi:10.1126/science.1152763 (2008).

144 Rocco, C. J. & Escalante-Semerena, J. C. In *Salmonella enterica*, 2-methylcitrate blocks gluconeogenesis. *J Bacteriol* **192**, 771-778, doi:10.1128/JB.01301-09 (2010).

145 Peng, S., Stephan, R., Hummerjohann, J. & Tasara, T. Transcriptional analysis of different stress response genes in *Escherichia coli* strains subjected to sodium chloride and lactic acid stress. *FEMS Microbiol Lett* **361**, 131-137, doi:10.1111/1574-6968.12622 (2014).

146 Goodarzi, H. *et al.* Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in *E. coli*. *Mol Syst Biol* **6**, 378, doi:10.1038/msb.2010.33 (2010).

147 Mundhada, H. *et al.* Increased production of L-serine in *Escherichia coli* through Adaptive Laboratory Evolution. *Metab Eng* **39**, 141-150, doi:10.1016/j.ymben.2016.11.008 (2017).

148 Haft, R. J. *et al.* Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. *Proc Natl Acad Sci U S A* **111**, E2576-2785, doi:10.1073/pnas.1401853111 (2014).

149 Hafeezunnisa, M. & Sen, R. The Rho-Dependent Transcription Termination Is Involved in Broad-Spectrum Antibiotic Susceptibility in *Escherichia coli*. *Front Microbiol* **11**, 605305, doi:10.3389/fmicb.2020.605305 (2020).

150 Jain, S., Gupta, R. & Sen, R. Rho-dependent transcription termination in bacteria recycles RNA polymerases stalled at DNA lesions. *Nature communications* **10**, 1207, doi:10.1038/s41467-019-09146-5 (2019).

151 Mo, A. H. & Burkholder, W. F. YneA, an SOS-induced inhibitor of cell division in *Bacillus subtilis*, is regulated posttranslationally and requires the transmembrane region for activity. *J Bacteriol* **192**, 3159-3173, doi:10.1128/JB.00027-10 (2010).

152 Mordukhova, E. A., Lee, H. S. & Pan, J. G. Improved thermostability and acetic acid tolerance of *Escherichia coli* via directed evolution of homoserine *o*-succinyltransferase. *Appl Environ Microbiol* **74**, 7660-7668, doi:10.1128/AEM.00654-08 (2008).

153 Tuite, N. L., Fraser, K. R. & O'Byrne C, P. Homocysteine toxicity in *Escherichia coli* is caused by a perturbation of branched-chain amino acid biosynthesis. *J Bacteriol* **187**, 4362-4371, doi:10.1128/JB.187.13.4362-4371.2005 (2005).

154 VanBogelen, R. A., Kelley, P. M. & Neidhardt, F. C. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. *J Bacteriol* **169**, 26-32, doi:10.1128/jb.169.1.26-32.1987 (1987).

3.6 Supplemental Information

Table S3.1: Genotype of E. coli strains in this study

Strain	Lab Strain #	Genotype	Ref
BW25113	AL1061	*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$	[95]
JCL16	AL87	BW25113/F′ [*traD36*, *proAB*$^{+}$, *lacI*^{q} *Z*Δ*M*15 *Tn10*(Tet^{R})]	[96]
JCL260	AL17	Same as JCL16 but with Δ*adhE*, Δ*frdBC*, Δ*fnr-ldhA*, Δ*pta*, Δ*pflB*	[96]
rho	AL3538	Same as JCL260, but with *rho*^{G63C}	This study
metH	AL3540	Same as JCL260, but with *metH*^{A387V}	This study
metH, rho, cysP, dctA, yiiR	AL3657	Same as *metH, rho*, but with *cysP*^{A230S}, *dctA*^{P36S}, *yiiR*^{D135N}	This study
Δ*arcA*	AL3887	Same as JCL260, but with Δ*arcA-yjjY-ParcA*	This study
Δ*prp*	AL3894	Same as JCL260, but with Δ*prpRBCDE*	This study
Δ*cyn*	AL3895	Same as JCL260, but with Δ*cynRTSX*	This study
Δ*cod*	AL3898	Same as JCL260, but with Δ*codBA*	This study
Δ*lac*	AL3899	Same as JCL260, but with Δ*lacIZYA*	This study
metH, Δ*arcA*	AL3915	Same as *metH*, but with Δ*arcA-yjjY-ParcA*	This study
metH, rho, Δ*arcA*	AL3919	Same as *metH*, Δ*arcA* but with *rho*^{G63C}	This study
Δ*yah*	AL3921	Same as JCL260, but with Δ*yahIHJKLMNO*	This study
Δ*mhp*	AL3927	Same as JCL260, but with Δ*mhpRABC*	This study
rho, Δ*arcA*	AL3942	Same as *rho*, but with Δ*arcA-yjjY-ParcA*	This study
metH, rho	AL3943	Same as *metH*, but with *rho*^{G63C}	This study
Δ30 kb	AL3946	Same as JCL260, but with Δ*yahI-mhpC*	This study
metH, rho, Δ*arcA* Δ*prp*	AL3951	Same as *metH, rho,* Δ*arcA*, but with Δ*prpRBCDE*	This study

metH, rho, Δ*arcA,* Δ*yah*	AL3592	Same as *metH, rho,* Δ*arcA*, but with Δ*yahIHJKLMNO*	This study
metH, rho, Δ*arcA,* Δ*mhp*	AL3955	Same as *metH, rho,* Δ*arcA*, but with Δ*mhpRABC*	This study
metH, rho, Δ*arcA,* Δ*cod- cyn*	AL3956	Same as *metH, rho,* Δ*arcA*, but with Δ*codB-cynX*	This study
metH, rho, Δ*arcA,* Δ*lac*	AL3957	Same as *metH, rho,* Δ*arcA*, but with Δ*lacIZYA*	This study
metH, rho, cysP, dctA, yiiR, Δ*arcA*	AL3982	Same as *metH, rho, cysP, dctA, yiiR* but with Δ*arcA-yjjY-ParcA*	This study
Δ30 kb, Δ*arcA, metH, rho*	AL3983	Same as Δ30 kb, Δ*arcA,* but with $metH^{A387V}$ and rho^{G63C}	This study
Δ30 kb, Δ*arcA*	AL3984	Same as Δ30 kb, but with Δ*arcA-yjjY-ParcA*	This study
metH, rho, Δ*lac*	AL3996	Same as *metH, rho,* but with Δ*lacIZYA*	This study

Table S3.2: List of plasmids used in this study

Plasmid Name	Genotype	Ref
pCas	pSC101 ori, Kan^R P_{cas}:*cas9*, P_{araB}:*gam-bet-exo*, *lacI*^{q}, P_{trc}:sgRNA-ColE1	[119]
pTargetF	ColE1ori, $Spec^R$, sgRNA-pmB1	[119]
pAL603	ColE1 ori, Amp^R, $P_L lacO_1$:*alsS-ilvCD,* $P_L lacO_1$:*kivd-adhA*	[107]
pAL682	p15A ori, Kan^R, $P_L lacO_1$:*mRFP1*	[107]
pAL685	p15A ori, Kan^R, $P_L lacO_1$:*ATF1*	[107]
pAL953	ColA ori, Kan^R, $P_L lacO_1$:*ackA-pta*	[98]
pAL1114	p15A ori, Kan^R, *lacI*^{q}, $P_L lacO_1$:*ATF1*	This study
pAL1839	ColE1 ori, Amp^R, sgRNA-*metH*	This study
pAL1840	ColE1 ori, Amp^R, sgRNA-*rho*	This study
pAL1841	ColE1 ori, Amp^R, sgRNA-*cysP*	This study
pAL1842	ColE1 ori, Amp^R, sgRNA-*yiiR*	This study
pAL1843	ColE1 ori, Amp^R, sgRNA-*dctA*	This study
pAL1934	p15A ori, $Spec^R$, *lacI*^{q}, $P_L lacO_1$:*ATF1*	This study
pAL2088	ColE1 ori, Amp^R, sgRNA-Δ*codBA*	This study
pAL2089	ColE1 ori, Amp^R, sgRNA-Δ*lacIZYA*	This study
pAL2107	ColE1 ori, Amp^R, sgRNA-Δ*prpRBCDE*	This study
pAL2108	ColE1 ori, Amp^R, sgRNA-Δ*cynRTSX*	This study
pAL2138	ColE1 ori, Amp^R, sgRNA-Δ*yahIJKLMNO*	This study
pAL2139	ColE1 ori, Amp^R, sgRNA-Δ*mhpRABC*	This study
pAL2151	ColE1 ori, Amp^R, sgRNA-Δ*codBA-cynRTSX*	This study

Table S3.3: Oligonucleotides used in this study

Name	DNA sequence
AZ349	GCGCGTCAACGTTTTAGAGCTAGAAATAGC
AZ350	GACATCCAGCACTAGTATTATACCTAGGAC
AZ351	CCTACCTCGCGTTTTAGAGCTAGAAATAGC
AZ352	AGCTGTCTGCACTAGTATTATACCTAGGAC
AZ353	TATGAAGCGCGTTTTAGAGCTAGAAATAGC
AZ354	CTGTTTACGGACTAGTATTATACCTAGGAC
AZ355	AAAGATACTCGTTTTAGAGCTAGAAATAGC
AZ356	ACCATATTTAACTAGTATTATACCTAGGAC
AZ357	TGAAACCGCTGTTTTAGAGCTAGAAATAGC
AZ358	TTTGCTCGCCACTAGTATTATACCTAGGAC
AZ642	TACCCTTGCCGTTTTAGAGCTAGAAATAGC
AZ643	TGCAGCACTTACTAGTATTATACCTAGGAC
AZ662	TCTGCGGAACGTTTTAGAGCTAGAAATAGC
AZ663	CGCCTGAGCGACTAGTATTATACCTAGGAC
AZ664	GATGTAACACGTTTTAGAGCTAGAAATAGC
AZ665	GCTTTTTTGGACTAGTATTATACCTAGGAC
AZ895	AGGCTAAGAGGTTTTAGAGCTAGAAATAGC
AZ896	GTTATGTTGCACTAGTATTATACCTAGGAC
AZ900	ATTCATCTGCGTTTTAGAGCTAGAAATAGC
AZ901	CTGCCACAGCACTAGTATTATACCTAGGAC
AZ902	CCCCACCTTTTTGCACTCATTCATATAAAAAATATATTTCCCCACGAAAACGAT
AZ903	CCGCCGCATTCTATTCATCTGAAGATAAAAAGCAATCGTTTTCGTGGGGAAAT
AZ904	TGAATAGAATGCGGCGGATTTTTTGGGTTTCAAACAGCAAAAAGGGGGAATTTCACGACT
AZ905	CACGCCGCATCCGGCATAAACTAAGCTCGCTGTAACCCAGTCGTGAAATTCCCCCT

AZ906	CGGATGCGGCGTGAACGCCTTATCCGGCCTACGTAGAGCACTGAACTCGTAGGCCTGATA
AZ907	GATCAGCGGCTGGAATTGCCTGATGCGCTACGCTTATCAGGCCTACGAGTTCAGTGCTCT
AZ908	AACGCTGCTTGTTTTAGAGCTAGAAATAGC
AZ909	GTTGCAGTGCACTAGTATTATACCTAGGAC
AZ937	GCCTGGGTTTTCACAGAAGTGAAGCATGGC
AZ938	GGATTGTCGATAGTCAGGCACACCTGGAGC
AZ949	TCACTGCCGAAAATGAAAGCCAGTAAAGAAGTTACAACGGACGATGAGTTACGTATCT
AZ950	CGCTTTTTAGCGCCGTTTTTATTTTTCAACCTTATTTCCAGATACGTAACTCATCGTCCG
AZ951	ACGGCGCTAAAAAGCGCCGTTTTTTTTGACGGTGGTAAAGCCGACAGAAGGATATGT
AZ952	TTCCTGACTGTACTAACGGTTGAGTTGTTAAAAAATGCTACATATCCTTCTGTCGGCT
AZ953	ACCGTTAGTACAGTCAGGAAATAGTTTAGCCTTTTTTAAGCTAAGTAAAGGGCTTTTTCT
AZ954	GGTGCGAATTTACAAATTCTTAACGTAAGTCGCAGAAAAAGCCCTTTACTTAGCTTAAA
AZ976	CAGGCCGGGCGTTTTAGAGCTAGAAATAGC
AZ977	CCAGTGACGTACTAGTATTATACCTAGGAC
MMM40	GAGTCAGTGAGCGAGGAAGC
MMM199	TGAAAAACGTTGAAGTGTTCGATACTGGCGGTCGTGGCGCGACCACCACTTTT
MMM200	CCGATTCGAAGCTAATCAGCACATCGCCCAGGCCGCGCTCGGCAAAAGTGGTGGTCGCGC
MMM201	TGCTGATTAGCTTCGAATCGGAAGTGAACAACATCCGTAAACAGTATGAATCGCAAGGCT
MMM202	ATTCCGCCAGAATGTTGGTTTTCGGAATCACCACTTCAAAGCCTTGCGATTCATACTGTT
MMM203	AACCAACATTCTGGCGGAATTCCCGGTGGCGTGGGTTGATAAAAACGTGCAGGCCAACGG
MMM204	AGCCAGTTCAGATAGGCTTTGGCGGCTTTTTCCGTACCGTTGGCCTGCACGT
MMM205	GTTTTACTAAAGGACACCCTATGAAAACCTCTCTGTTTAAAAGCCTTTACTTTCAG
MMM206	CCAAGGAGAATACCAATGGCTATCGCTGTCAGGACCTGAAAGTAAAGGCTTTTAAACAGA
MMM207	AGCCATTGGTATTCTCCTTGGCCATTTCTATCCTGAAATAGGCGAGCAAATGAAGTCCCT
MMM208	GAGCGATGATCATCTTAATGAGCTTAACGAAGCCGTCGCCAAGGGACTTCATTTGCTCGC

MMM209	GCTCATTAAGATGATCATCGCTCCTGTCATCTTTTGTACCGTCGTAACGGGCATTGCGGG
MMM210	GCGACTGCGCCGGTACGACCGACCGCCTTCATGCTTTCCATGCCCGCAATGCCCGTTAC
MMM211	GGAGCCGCTGAACATTGGCGAAGATAGCCTGTTTGTGAACGT
MMM212	TGAACTTAGCGGAACCGGTGACGTTGGTGCGTTCACCCACGTTCACAAACAGGCTATC TT
MMM213	ACCGGTTCCGCTAAGTTCAAGCGCCTGATCAAAGAAGAGAAATACAGCGAGGCGCTGG AT
MMM214	ATATCGATAATCTGCGCGCCGTTTTCCACTTGTTGACGCACGACATCCAGCGCCTCGCTG
MMM215	GGCGCGCAGATTATCGATATCAACATGGATGAAGGGATGCTCGATGCCGAAGCGGCGAT G
MMM216	GCGAGCGATATCCGGTTCACCGGCAATCAGATTGAGAAAACGCACCATCGCCGCTTCGG C
MMM217	TGGGGCTGGAAAACCTGGCTCGTATGCGTAAGCAGGACATTATTTTTGCCATCCTG
MMM218	CACCAAAGATATCTTCGCCACTCTTTGCGTGCTGCTTCAGGATGGCAAAAATAATGTCCT
MMM219	GTGGCGAAGATATCTTTGGTGATGGCGTACTGGAGATATTGCAGGATGGATTTTGTTTCC
MMM220	CATCAGGACCGGCCAGATAAGAGCTGTCTGCGGAACGGAGGAAACAAAATCCATCCTG CA
MMM221	TCTGGCCGGTCCTGATGACATCTACGTTTCCCCTAGCCAAATCCGCCGTTTCAACCTCCG
MMM222	GGCGGGCGAATCTTACCAGAGATGGTATCACCAGTGCGGAGGTTGAAACGGCG
MMM223	ACCGGGTGTCTGGCAATG
MMM224	CATCATCACCAGAATCAACGTCGGGACAAAACGCCCCACCGCCCATTGCCAGACACCC GG
MMM225	GACGTTGATTCTGGTGATGATGCTTATCGATCTTGGTGCGTTTGTTGGCACACAAGGCG A
MMM226	GGCTTTATACTTCACATTCTGAGTATCTTTACCATATTTATTTTCGCCTTGTGTGCCAAC
MMM227	TACTCAGAATGTGAAGTATAAAGCCGACAATAAATCAAGTAATTAACCTTTAATATATCT
MMM228	ACAATCCCGGAAAAAAATAGCAGTAAGAGATATATTAAAGGTTAATTACTTGATTTATTG
MMM229	TGCTATTTTTTTCCGGGATTGTTATTTCCACGATGCAAATGACAATACCCGGAAAATCCT
MMM230	AGTTCAGAAGGATTTTCCGGGTATTGTCA
MMM231	CCTCACGTACTATGTGTACGCTC
MMM232	CGCCGACAGCAGCGTTAC
MMM233	CTTCCAAATCACCAAACGG

MMM234	GCGGAGAGCGGCAGCAGC
MMM235	GTAGCTACCCCAAACTGATTAC
MMM236	CTGCCCGTTGTCGGTCAGTC
MMM237	GATCCAGGTACATTTCTGTG
MMM238	CATCTGACTGTTAGTGGCC
MMM239	GACCACCGAAGCATTTTAC
MMM240	CAGCGCGTGCGGCAGTTC
MMM241	CTGCGCCAGTACGTGCAGGAG
MMM242	GAAGATGATATCTTCTGGCG
MMM243	GCGAAGTGAACAGATTTCTG
MMM244	CCACGAAGACCTTTATTCAG
MMM245	GACATGAGTCGATGACCG
MMM246	GCACGAAATTCGTTCAGTGG
MMM247	GGAAATTGCCACGCCAATGTG
MMM248	CAGCAGATAAGGCGCGGTCG
MMM249	GCAACTCCATCAAGCTAACG
MMM265	GATGTGCTGATTAGCTTCGAATCGGAAGTGAACAACATCCGTAAACAGTATGAAGCGCAA
MMM266	CGCCAGAATGTTGGTTTTCGGAATCACCACTTCAAAGCCTTGCGCTTCATACTGTTTACG
MMM267	AGCCATTGGTATTCTCCTTGGCCATTTCTATCCTGAAATAGGCGAGCAAATGAAGGCCCT
MMM268	GGAGCGATGATCATCTTAATGAGCTTAACGAAGCCGTCGCCAAGGGCCTTCATTTGCTCG
MMM269	ATATCGATAATCTGCGCGCCGTTTTCCACTTGTTGACGCGCGACATCCAGCGCCTCGCTG
MMM270	GTGGCGAAGATATCTTTGGTGATGGCGTACTGGAGATATTGCAGGATGGATTTGGTTTCC
MMM271	CATCAGGACCGGCCAGATAAGAGCTGTCTGCGGAACGGAGGAAACCAAATCCATCCTGCA
MMM272	GGCTTTATACTTCACATCCTGAGTATCTTTACCATATTTATTTTCGCCTTGTGTGCCAAC
MMM273	TACTCAGGATGTGAAGTATAAAGCCGACAATAAATCAAGTAATTAACCTTTAATATATCT
MMM310	CATGACATTAACCTATAAAAATAGGCG

MMM311	ATAATAATGGTTTCTTAGACGTCGG
MMM312	CCGACGTCTAAGAAACCATTATTATTCACTGCCCGCTTTCCAGTCGG
MMM313	CGCCTATTTTTATAGGTTAATGTCATGATGATAAGCACGCGTCAACTCGAGC
MMM318	CAGGATGAGGATCGTTTCGCATGAGGGAAGCGGTGATCGC
MMM319	CTCGAACCCCAGAGTCCCGCTTATTTGCCGACTACCTTGG
MMM320	GCGGGACTCTGGGGTTCGAG
MMM321	GCGAAACGATCCTCATCCTG
MMM427	GTATTCCGCGACTAGTATTATACCTAGGAC
MMM428	ATATCAGAGCGTTTTAGAGCTAGAAATAGC
MMM429	GAAGTAATTCTGAATCCTATGTAAACATCTCCGATGCGTAAGTTTATCG
MMM430	CGAACGCTTTATCCGGCATAAATTTCAATAGATGATCACCGATAAACTTACGCATCGGAG
MMM431	TGCCGGATAAAGCGTTCGCGCTGCATTCGGCAGTTCAGCTTTTCAGCCGCCGCCACGCCA
MMM432	GGATGCACCGCGCGACGCATCCGACAATAACCTACTCTTCCATCGCCTGGCGTGGCGGCG
MMM433	TCGCGCGGTGCATCCGGCACTGTGTGCCGATGCCTGATGCGACGCTGACGCGTTTTATCA
MMM434	TCCGACCTACGGTTCAGGTCCGTAGGCATGATAAAACGCGTCAGCGT
MMM435	CGTTTATGCCCTGACTGAAC
MMM436	GAAGGTTAATCCCAGCATGAC
MMM437	CTATGCCGCTGAATTGATGACC
MMM438	CTAAAGTTGTTATCTTGCGACACG
MMM442	CGTAGACCCAACTAGTATTATACCTAGGAC
MMM443	ACATTGAAAGGTTTTAGAGCTAGAAATAGC
MMM444	GGATGCGGCGTGAACGCCTTATCCGGCCTACGTAGAGCACTGAACTCGTAG
MMM445	CTGCGGCTGGAATTGCCTGATGCGCTACGCTTATCAGGCCTACGAGTTCAGTGCTCTACG
MMM446	GGCAATTCCAGCCGCAGACCTGTGTCAGCGGCTACCGTGATTCATTTCCGCCCTGTGGGT
MMM447	CATGAACAAAGCGCAGGAACAAGCGTCGCATCATGCCTCTTTGACCCACAGGGCGGAAAT

MMM448	GTTCCTGCGCTTTGTTCATGCCGGATGCGGCTAATGTAGATCGCTGAACTTGTAGGCCTG
MMM449	AAATTATAAAAATTGCCTGATACGCTGCGCTTATCAGGCCTACAAGTTCAGCG
MMM450	ATTCGGTACGTGGCGGCAAGGTGATT
MMM451	GACTACACGGTAACAATCGGTGATAACGTACTG
MMM452	GGTTACAGCGAGCTTAGTTTATGCC
MMM453	CTTTTCCGATAACGATTGGCAATAACG
MMM460	TAGGCGTTGGACTAGTATTATACCTAGGAC
MMM461	AACAGAACGGGTTTTAGAGCTAGAAATAGC
MMM462	CGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAAT
MMM463	CTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAA
MMM464	TCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCAACGTAAATGCATGCCGC
MMM465	GCTTCGTCATGACCATTCCTTAAGGGCGTGCGAGGAAATTAGCGGCATGCATTTACGTTG
MMM466	AGGAATGGTCATGACGAAGCATACTCTTGAGCAACTGGCGGCGGATTTACGCCGCGCCGC
MMM467	ATACCAATCAGATCGCGCAGCGGTGCAATCGCTTCGCCCTGCTCTGCGGCGCGGCGTA
MMM471	TTGCGCGAGAAGATTGTGCACCGCC
MMM472	CACCACTTCCAGCGCCGGAAGTACC
MMM473	CTTCAATGCGGGGTTGCAGAACACG
MMM474	GCGGCGTGAACGCCT
MMM475	ACGCTTATCAGGCCTACAATTCTCGCGCCACTCGTAGGCCGGATAAGGCGTTCACGCCGC
MMM476	AATTGTAGGCCTGATAAGCGTAGCGCATCAGGCAGTCTGGCGTTGGTCATAACCCCATCA
MMM477	TTCTTATAAATATTCATGTTGTTATGTCCGCTACAGAGGGTGATGGGGTTATGACCAACG
MMM478	GGACATAACAACATGAATATTTATAAGAAGAAGTAATTCTGAATCCTATGTAAACATCTC
MMM479	CAATAGATGATCACCGATAAACTTACGCATCGGAGATGTTTACATAGGATTCAGAATTAC
MMM480	GTAAGTTTATCGGTGATCATCTATTGAAATTTATGCCGGATAAAGCGTTCGCGCTGCATT
MMM481	AAGCTGAACTGCCGAATGCAGCGCGAACGC
MMM482	AAGGGAAGTGGTTTTAGAGCTAGAAATAGC

MMM483	TCCACATTGTACTAGTATTATACCTAGGAC
MMM484	GAATAGAATGCGGCGGATTTTTTGGGTTTCAAACAGCAAAAAGGGGGAATTTCCTGTGGG
MMM485	TGAACAAAGCGCAGGAACAAGCGTCGCATCATGCCTCTTTGACCCACAGGAAATTCCCCC
MMM486	GTAAACATCTCCGATGCGTAAGTTTATCGG
MMM487	CTGGATGTCAGGGTGTTGTATTGCC
MMM488	GCGGCGTGAACGCCTTATCCGGCCTACGAGTGGCGCGA
MMM489	ACGCCAGACTGCCTGATGCGCTACGCTTATCAGGCCTACAATTCTCGCGCCACTCGTAGG
MMM490	ATCAGGCAGTCTGGCGTTGGTCATAACCCCATCACCCTCTGTAGCGGACATAACAACGGA
MMM491	CCGCCGCCAGTTGCTCAAGAGTATGCTTCGTCATGACCATTCCGTTGTTATGTCCGCTAC
MMM492	AGCAACTGGCGGCGGATTTACGCCGCGCCGCAGAGCAGGGCGAAGCGATTGCACCGCTGC
MMM493	TAAGCCGCTTCAGCGTTATCGATACCAATCAGATCGCGCAGCGGTGCAATCGC
MMM494	CCTACATGTTCATGCCGGATGCGGC
MMM495	CATTCTTGGCTGGTGCGAACGTTACGG
MMM496	AGCTGTGGGTCAAAGAGGCATGATGCGACGCTTGTTCCTGCGCTTTGTTCATGCCGGA
MMM497	CGCTTATCAGGCCTACAAGTTCAGCGATCTACATTAGCCGCATCCGGCATGAACAAAGCG
MMM498	AACTTGTAGGCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAATATTCACCACCCTGA
MMM499	GTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATTATAAAAATTG
MMM500	GGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCAACGTAAATG
MMM501	GCAGGCTATTCTGGTGGCCGGAAGGCGAAGCGGCATGCATTTACGTTGACACCATCGAA
MMM502	GTTACAACTGGTGGGGTTCTGCGGC
MMM503	TTTGCCCGGAACAAGACCGC

Table S3.4: Plasmid construction by SLIC

Plasmid	Backbone			Insert			
	Primer 1	Primer 2	Template	Primer 1	Primer 2	Template	Contents
pAL1114	MMM310	MMM311	pAL685	MMM312	MMM313	Z1	*lacI*^{q}
pAL1934	MMM320	MMM321	pAL685	MMM318	MMM319	pTargetF	Spec^{R}

Table S3.5: CRIPSR Donor DNA Fragments and Corresponding Primers

Gene Target(s)	Donor DNA Sequence	Primers
codBA	CCCCACCTTTTTGCACTCATTCATATAAAAAATATATTTCCCC ACGAAAACGATTGCTTTTTATCTTCAGATGAATAGAATGCGG CGGATTTTTTGGGTTTCAAACAGCAAAAAGGGGGAATTTCA CGACTGGGTTACAGCGAGCTTAGTTTATGCCGGATGCGGCG TGAACGCCTTATCCGGCCTACGTAGAGCACTGAACTCGTAG GCCTGATAAGCGTAGCGCATCAGGCAATTCCAGCCGCTGATC	AZ902-906
IS30 Insertion in *yjjY*	GCCTGGGTTTTCACAGAAGTGAAGCATGGCGCGGAACTCGC TGCGCGGCAGCTTGTACTGCTCGCCATCAGGGCCGATCAAC GAACGGCTGTTGATGTCCAGTTCCCAACCATTGAACTTGTA GCTTTCAACGCTACGACGTTCTTCGCTGACAGTACCCAGATT CATGGTACGGGACAGTAGGTTGCGTGCACGAATCGTCAGTT CACGCGGGTTGAACGGTTTGGTGATGTAGTCATCTGCACCG ATTTCGAGGCCGAGAATTTTATCGACTTCGTTGTCACGGCCA GTCAGGAACATCAACGCAACATTCGCCTGCTCGCGCAGTTC ACGCGCTAACAGAAGACCGTTCTTACCCGGCAGATTGATATC CATGATCACCAGGTTGATGTCATATTCAGAGAGGATCTGATG CATTTCCGCGCCATCTGTCGCTTCGAAAACATCATAGCCTTC CGCTTCGAAAATACTTTTCAACGTGTTGCGTGTTACCAACTC GTCTTCAACGATAAGAATGTGCGGGGTCTGCATGTTTGCTAC CTAAATTGCCAACTAAATCGAAACAGGAAGTACAAAAGTCC CTGACCTGCCTGATGCATGCTGCAAATTAACATGATCGGCGT AACATGACTAAAGTACGTAATTGTAGATTCAATTGGTCAACG CAACAGTTATGTGAAAACATGGGGTTGCGGAGGTTTTTTGA ATGAGACGAACATTTACAGCAGAGGAAAAAGCCTCTGTTTT TGAACTATGGAAGAACGGAACAGGCTTCAGTGAAATAGCGA ATATCCTGGGTTCAAAACCCGGAACGATCTTCACTATGTTAA GGGATACTGGCGGCATAAAACCCCATGAGCGTAAGCGGGCT GTAGCTCACCTGACACTGTCTGAGCGCGAGGAGATACGAGC TGGTTTGTCAGCCAAAATGAGCATTCGTGCGATAGCTACTGC GCTGAATCGCAGTCCTTCGACGATCTCACGTGAAGTTCAGC GTAATCGGGGCAGACGCTATTACAAAGCTGTTGATGCTAATA ACCGAGCCAACAGAATGGCGAAAAGGCCAAAACCGTGCTT ACTGGATCAAAATTTACCATTGCGAAAGCTTGTTCTGGAAA AGCTGGAGATGAAATGGTCTCCAGAGCAAATATCAGGATGG TTAAGGCGAACAAAACCACGTCAAAAAACGCTGCGAATATC ACCTGAGACAATTTATAAAACGCTGTACTTTCGTAGCCGTGA AGCGCTACACCACCTGAATATACAGCATCTGCGACGGTCGC ATAGCCTTCGCCATGGCAGGCGTCATACCCGCAAAGGCGAA AGAGGTACGATTAACATAGTGAACGGAACACCAATTCACGA ACGTTCCCGAAATATCGATAACAGACGCTCTCTAGGGCATTG GGAGGGCGATTTAGTCTCAGGTACAAAAAACTCTCATATAGC CACACTTGTAGACCGAAAATCACGTTATACGATCATCCTTAG ACTCAGGGGCAAAGATTCTGTCTCAGTAAATCAGGCTCTTA CCGACAAATTCCTGAGTTTACCGTCAGAACTCAGAAAATCA CTGACATGGGACAGAGGAATGGAACTGGCCAGACATCTAGA ATTTACTGTCAGCACCGGCGTTAAAGTTTACTTCTGCGATCC TCAGAGTCCTTGGCAGCGGGGAACAAATGAGAACACAAAT GGGCTAATTCGGCAGTACTTTCCTAAAAAGACATGTCTTGCC CAATATACTCAACATGAACTAGATCTGGTTGCTGCTCAGCTA	AZ937, 938

	AACAACAGACCGAGAAAGACACTGAAGTTCAAAACACCGA AAGAGATAATTGAAAGGGGTGTTGCATTGACAGATTGAATC TACAATTGCGTTCTTGATGCACTTTCCATCAACGTCAACAAC ATCATTAGCTTGGTCGTGGGTACTTTCCCTCAGGACCCGACA GTGTCAAAAACGGCTGTCATCCTAACCATTTTAACAGCAAC ATAACAGGCTAAGAGGGGCCGGACACCCAATAAAACTACGC TTCGTTGACATATATCAAGTTCAATTGTAGCACGTTAACAGTT TGATGAAATCATCGTATCTAAATGCTAGCTTTCGTCACATTAT TTTAATAATCCAACTAGTTGCATCATACAACTAATAAACGTG GTGAATCCAATTGTCGAGATTTATTTTTTATAAAATTATCCTA AGTAAACAGAAGGATATGTAGCATTTTTTAACAACTCAACCG TTAGTACAGTCAGGAAATAGTTTAGCCTTTTTTAAGCTAAGT AAAGGGCTTTTTCTGCGACTTACGTTAAGAATTTGTAAATTC GCACCGCGTAATAAGTTGACAGTGATCACCCGGTTCGCGGT TATTTGATCAAGAAGAGTGGCAATATGCGTATAACGATTATTC TGGTCGCACCCGCCAGAGCAGAAAATATTGGGGCAGCGGC GCGGGCAATGAAAACGATGGGGTTTAGCGATCTGCGGATTG TCGATAGTCAGGCACACCTGGAGC	
arcA-yjjY-ParcA	TCACTGCCGAAAATGAAAGCCAGTAAAGAAGTTACAACGG ACGATGAGTTACGTATCTGGAAATAAGGTTGAAAAATAAAA ACGGCGCTAAAAAGCGCCGTTTTTTTTGACGGTGGTAAAGC CGACAGAAGGATATGTAGCATTTTTTAACAACTCAACCGTTA GTACAGTCAGGAAATAGTTTAGCCTTTTTTAAGCTAAGTAAA GGGCTTTTTCTGCGACTTACGTTAAGAATTTGTAAATTCGCA CC	AZ949-954
cysP	ACCGGGTGTCTGGCAATGGGCGGTGGGGCGTTTTGTCCCGA CGTTGATTCTGGTGATGATGCTTATCGATCTTGGTGCGTTTGT TGGCACACAAGGCGAAAATAAATATGGTAAAGATACTCAGA ATGTGAAGTATAAAGCCGACAATAAATCAAGTAATTAACCTT TAATATATCTCTTACTGCTATTTTTTTCCGGGATTGTTATTTCC ACGATGCAAATGACAATACCCGGAAAATCCTTCTGAACT	MMM199-204
cysP Reversion	TGAAAAACGTTGAAGTGTTCGATACTGGCGGTCGTGGCGCG ACCACCACTTTTGCCGAGCGCGGCCTGGGCGATGTGCTGAT TAGCTTCGAATCGGAAGTGAACAACATCCGTAAACAGTATG AAGCGCAAGGCTTTGAAGTGGTGATTCCGAAAACCAACATT CTGGCGGAATTCCCGGTGGCGTGGGTTGATAAAAACGTGCA GGCCAACGGTACGGAAAAAGCCGCCAAAGCCTATCTGAACT GGCT	MMM199, 200, 265, 266, 203, 204
dctA	GTTTTACTAAAGGACACCCTATGAAAACCTCTCTGTTTAAAA GCCTTTACTTTCAGGTCCTGACAGCGATAGCCATTGGTATTC TCCTTGGCCATTTCTATCCTGAAATAGGCGAGCAAATGAAGT CCCTTGGCGACGGCTTCGTTAAGCTCATTAAGATGATCATCG CTCCTGTCATCTTTTGTACCGTCGTAACGGGCATTGCGGGCA TGGAAAGCATGAAGGCGGTCGGTCGTACCGGCGCAGTCGC	MMM205-210
dctA Reversion	GTTTTACTAAAGGACACCCTATGAAAACCTCTCTGTTTAAAA GCCTTTACTTTCAGGTCCTGACAGCGATAGCCATTGGTATTC TCCTTGGCCATTTCTATCCTGAAATAGGCGAGCAAATGAAGG CCCTTGGCGACGGCTTCGTTAAGCTCATTAAGATGATCATCG CTCCTGTCATCTTTTGTACCGTCGTAACGGGCATTGCGGGCA TGGAAAGCATGAAGGCGGTCGGTCGTACCGGCGCAGTCGC	MMM205, 206, 267, 268, 207, 208
metH	GGAGCCGCTGAACATTGGCGAAGATAGCCTGTTTGTGAACG TGGGTGAACGCACCAACGTCACCGGTTCCGCTAAGTTCAAG	MMM211-216

	CGCCTGATCAAAGAAGAGAAATACAGCGAGGCGCTGGATGT CGTGCGTCAACAAGTGGAAAACGGCGCGCAGATTATCGATA TCAACATGGATGAAGGGATGCTCGATGCCGAAGCGGCGATG GTGCGTTTTCTCAATCTGATTGCCGGTGACCGGATATCGCTC GC	
metH Reversion	GGAGCCGCTGAACATTGGCGAAGATAGCCTGTTTGTGAACG TGGGTGAACGCACCAACGTCACCGGTTCCGCTAAGTTCAAG CGCCTGATCAAAGAAGAGAAATACAGCGAGGCGCTGGATGT CGCGCGTCAACAAGTGGAAAACGGCGCGCAGATTATCGATA TCAACATGGATGAAGGGATGCTCGATGCCGAAGCGGCGATG GTGCGTTTTCTCAATCTGATTGCCGGTGAACCGGATATCGCT CGC	MMM211-13, 269, 215, 216
rho	TGGGGCTGGAAAACCTGGCTCGTATGCGTAAGCAGGACATT ATTTTTGCCATCCTGAAGCAGCACGCAAAGAGTGGCGAAGA TATCTTTGGTGATGGCGTACTGGAGATATTGCAGGATGGATT TTGTTTCCTCCGTTCCGCAGACAGCTCTTATCTGGCCGGTCC TGATGACATCTACGTTTCCCCTACCAAATCCGCCGTTTCAAC CTCCGCACTGGTGATACCATCTCTGGTAAGATTCGCCCGCC	MMM217-222
rho Reversion	TGGGGCTGGAAAACCTGGCTCGTATGCGTAAGCAGGACATT ATTTTTGCCATCCTGAAGCAGCACGCAAAGAGTGGCGAAGA TATCTTTGGTGATGGCGTACTGGAGATATTGCAGGATGGATT TGGTTTCCTCCGTTCCGCAGACAGCTCTTATCTGGCCGGTCC TGATGACATCTACGTTTCCCCTAGCCAAATCCGCCGTTTCAA CCTCCGCACTGGTGATACCATCTCTGGTAAGATTCGCCCGCC	MMM217, 218, 270, 271, 221, 222
yiiR	ACCGGGTGTCTGGCAATGGGCGGTGGGGCGTTTTGTCCCGA CGTTGATTCTGGTGATGATGCTTATCGATCTTGGTGCGTTTGT TGGCACACAAGGCGAAAATAAATATGGTAAAGATACTCAGA ATGTGAAGTATAAAGCCGACAATAAATCAAGTAATTAACCTT TAATATATCTCTTACTGCTATTTTTTTCCGGGATTGTTATTTCC ACGATGCAAATGACAATACCCGGAAAATCCTTCTGAACT	MMM223-230
yiiR Reversion	ACCGGGTGTCTGGCAATGGGCGGTGGGGCGTTTTGTCCCGA CGTTGATTCTGGTGATGATGCTTATCGATCTTGGTGCGTTTGT TGGCACACAAGGCGAAAATAAATATGGTAAAGATACTCAGG ATGTGAAGTATAAAGCCGACAATAAATCAAGTAATTAACCTT TAATATATCTCTTACTGCTATTTTTTTCCGGGATTGTTATTTCC ACGATGCAAATGACAATACCCGGAAAATCCTTCTGAACT	MMM223-225, 272, 273, 228-230
3,853,295$^{C\rightarrow A}$	CTTGAAAAGCACAAAAGCCAGTCTGGAAACAGGCTGGCTT TTTTTTGCGCGTGTGACCCGTCCTGAATAGCGTTCACATAGA TCCTGCTGATATAAAACCCCCCTGTTTTCCTGTTTATTCATTT ATCGAAATAAGAGCAAAAACATCCACCTGACGCTTAAATTA AGGTACTGCCTTAATTTTCTGCAGACAAAAGGCGTGACGAT GGTCGAAAATGGCGCTTTCGTCAGCGGGGATAATCCGTTATT G	MMM327-332
prpRBCDE	GAAGTAATTCTGAATCCTATGTAAACATCTCCGATGCGTAAG TTTATCGGTGATCATCTATTGAAATTTATGCCGGATAAAGCGT TCGCGCTGCATTCGGCAGTTCAGCTTTTCAGCCGCCGCCAC GCCAGGCGATGGAAGAGTAGGTTATTGTCGGATGCGTCGCG CGGTGCATCCGGCACTGTGTGCCGATGCCTGATGCGACGCT GACGCGTTTTATCATGCCTACGGACCTGAACCGTAGGTCGG A	MMM429-434
cynRTSX	GGATGCGGCGTGAACGCCTTATCCGGCCTACGTAGAGCACT GAACTCGTAGGCCTGATAAGCGTAGCGCATCAGGCAATTCC	MMM444-449

	AGCCGCAGACCTGTGTCAGCGGCTACCGTGATTCATTTCCG CCCTGTGGGTCAAAGAGGCATGATGCGACGCTTGTTCCTGC GCTTTGTTCATGCCGGATGCGGCTAATGTAGATCGCTGAACT TGTAGGCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAAT TT	
mhpRABC	CGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGA ATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGG TTTTGCGCCATTCGATGGTGTCAACGTAAATGCATGCCGCTA ATTTCCTCGCACGCCCTTAAGGAATGGTCATGACGAAGCATA CTCTTGAGCAACTGGCGGCGGATTTACGCCGCGCCGCAGAG CAGGGCGAAGCGATTGCACCGCTGCGCGATCTGATTGGTAT	MMM462-467
yahIJKLMNO	GCGGCGTGAACGCCTTATCCGGCCTACGAGTGGCGCGAGAA TTGTAGGCCTGATAAGCGTAGCGCATCAGGCAGTCTGGCGT TGGTCATAACCCCATCACCCTCTGTAGCGGACATAACAACAT GAATATTTATAAGAAGAAGTAATTCTGAATCCTATGTAAACAT CTCCGATGCGTAAGTTTATCGGTGATCATCTATTGAAATTTAT GCCGGATAAAGCGTTCGCGCTGCATTCGGCAGTTCAGCTT	MMM474-481
codB-cynX	CCCCACCTTTTTGCACTCATTCATATAAAAAATATATTTCCCC ACGAAAACGATTGCTTTTTATCTTCAGATGAATAGAATGCGG CGGATTTTTTGGGTTTCAAACAGCAAAAAGGGGGAATTTCC TGTGGGTCAAAGAGGCATGATGCGACGCTTGTTCCTGCGCT TTGTTCATGCCGGATGCGGCTAATGTAGATCGCTGAACTTGT AGGCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAATTT	AZ902-903 MMM484,4854 48,449
yahI-mhpC	GCGGCGTGAACGCCTTATCCGGCCTACGAGTGGCGCGAGAA TTGTAGGCCTGATAAGCGTAGCGCATCAGGCAGTCTGGCGT TGGTCATAACCCCATCACCCTCTGTAGCGGACATAACAACGG AATGGTCATGACGAAGCATACTCTTGAGCAACTGGCGGCGG ATTTACGCCGCGCCGCAGAGCAGGGCGAAGCGATTGCACCG CTGCGCGATCTGATTGGTATCGATAACGCTGAAGCGGCTTA	MMM488-493
lacIZYA	AGCTGTGGGTCAAAGAGGCATGATGCGACGCTTGTTCCTGC GCTTTGTTCATGCCGGATGCGGCTAATGTAGATCGCTGAACT TGTAGGCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAAT ATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGC CATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCAACGTA AATGCATGCCGCTTCGCCTTCCGGCCACCAGAATAGCCTGC	MMM496-501

Table S3.6: Guide for CRISPR-Cas9 Construction by Q5 site-directed mutagenesis

		pTargetF	
Modification	**Plasmid**	**20 bp sgRNA sequence 5' → 3'**	**Primers**
metH^{A387V}	pAL1839	GCTGGATGTCGCGCGTCAAC	AZ3349, AZ350
rho^{G63C}	pAL1840	GCAGACAGCTCCTACCTCGC	AZ351, AZ352
cysP^{A230S}	pAL1841	CCGTAAACAGTATGAAGCGC	AZ353, AZ354
yiiR^{D135N}	pAL1842	TAAATATGGTAAAGATACTC	AZ355, AZ356
dctA^{P36S}	pAL1843	GGCGAGCAAATGAAACCGCT	AZ357, AZ358
3,853,295$^{C\rightarrow A}$	pAL1990	AAGTGCTGCATACCCTTGCC	AZ642, AZ643
3,853,295$^{C\rightarrow A}$	pAL1998	CGCTCAGGCGTCTGCGGAAC	AZ662, AZ663
3,853,295$^{C\rightarrow A}$	pAL1999	CCAAAAAAGCGATGTAACAC	AZ664, AZ665
Δ*arcA-yjjY-ParcA*	pAL2086	GCAACATAACAGGCTAAGAG	AZ896, AZ894
Δ*codBA*	pAL2088	GCTGTGGCAGATTCATCTGC	AZ900, AZ901
Δ*lacIZYA*	pAL2089	GCACTGCAACAACGCTGCTT	AZ908, AZ909
Δ*prpRBCDE*	pAL2107	CGCGGAATACATATCAGAGC	MMM427, MMM428
Δ*cynRTSX*	pAL2108	TGGGTCTACGACATTGAAAG	MMM442, MMM443
Δ*yahIJKLMNO*	pAL2138	ACGTCACTGGCAGGCCGGGC	AZ976, AZ977
Δ*mhpRABC*	pAL2139	CCAACGCCTAAACAGAACGG	MMM460, MMM461
Δ*codBA* Δ*cynRTSX*	pAL2151	ACAATGTGGAAAGGGAAGTG	MMM482, MMM483

Table S3.7: Details of ALE Mutations

Gene(s)	Region	Mutation Type	Mutation[a]	Coordinates
rho	Coding	SNP	G63C (GGT→TGT)	3,904,476
metH	Coding	SNP	A387V (GCG→GTG)	4,159,428
yjjY	Coding	IS	IS30 (+)	4,573,446
lacI^{q}	Coding	Deletion	Δ78 bp	F' plasmid
ykgR / *ykgP*	Intergenic	SNP	G→A	309,254
lhr	Coding	Indel	C Insertion	1,671,324
cysP	Coding	SNP	A230S (GCG→TCG)	2,480,713
dctA	Coding	SNP	P36S (CCG→TCG)	3,621,215
yiiR	Coding	SNP	D135N (GAT→AAT)	4,047,158
bglG /*phoU*	Intergenic	SNP	A→T	3,844,666
glmU / *atpC*	Intergenic	SNP	C→A	3,853,295
[yahI]–mhpC	Coding	Deletion	Δ30,548 bp	336,518
mngB /*cydA*	Intergenic	IS	IS5 (+)	761,425

[a] (+) represents forward direction of the IS elements. Brackets represent partial deletion within that gene. Slats represent mutations in the non-coding regions between two genes.

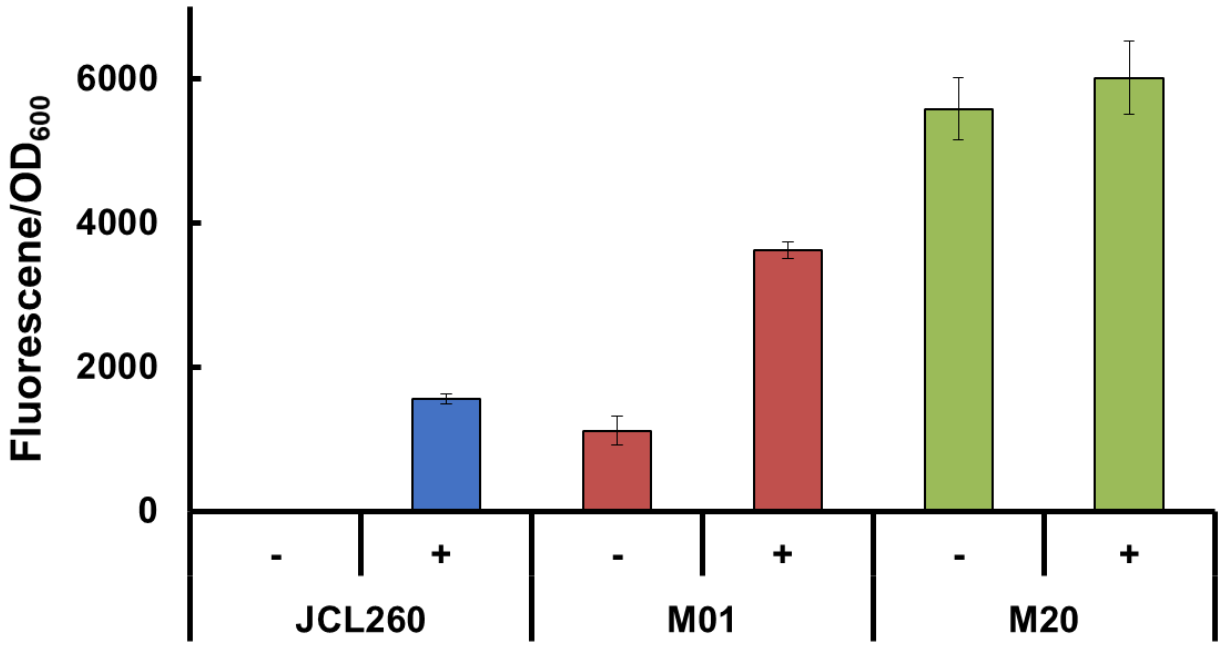


Fig. S3.1. mRFP1 expression under $P_{L}lacO_{1}$ promoter in JCL260, M01 and M20.

In LB media, cultures were induced with (+) or without (-) 1 mM IPTG then grown for 24 h at 37 °C, 250 rpm. N = 3, biological replicates; error bars represent s.d.

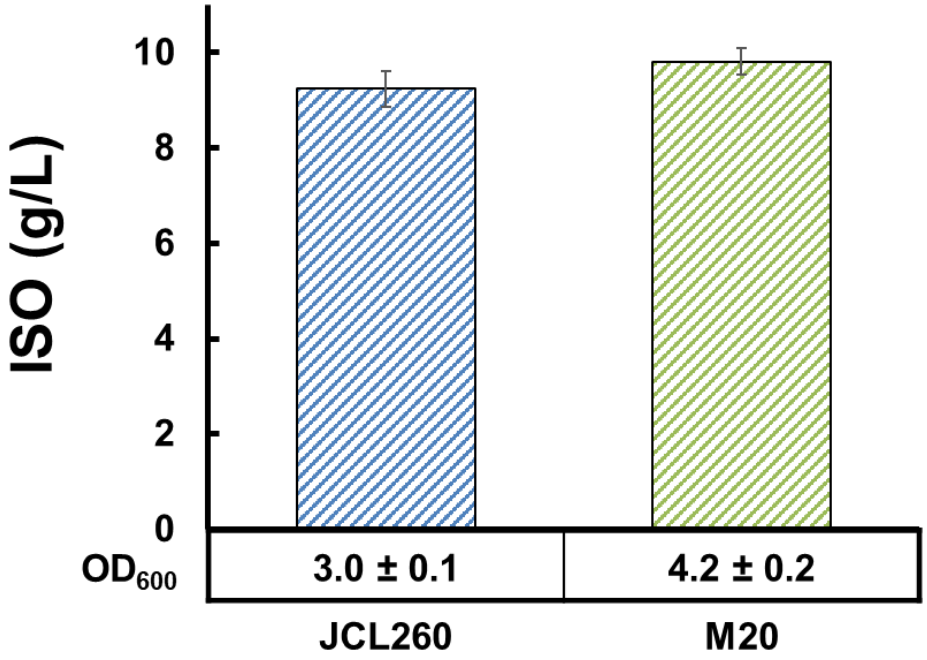


Fig. S3.2.

ISO production and OD600 from JCL260 and M20 harboring pAL603 after 24 h cultured in M9P media. N = 3, biological replicates; error bars represent s.d.

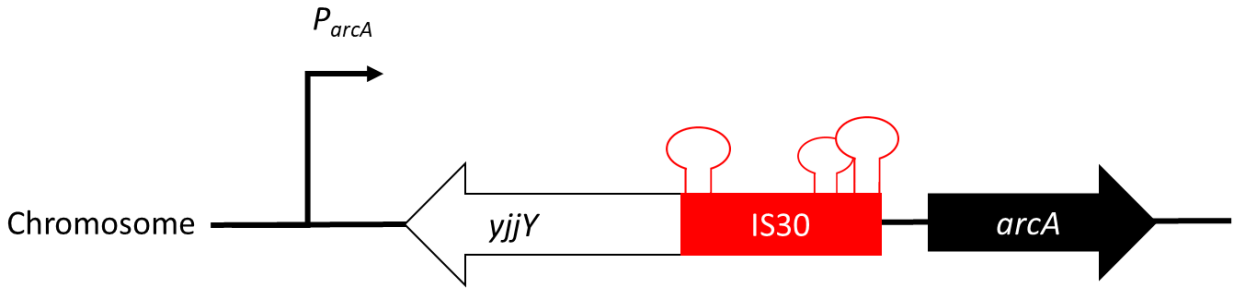


Fig. S3.3. Schematic representation of IS30 insertion in yjjY.

Orientation of the Δ*arcA-yjjY-ParcA* loci and location of the IS30 terminators (red stem loops) that strongly reduce *arcA* expression from *ParcA*.

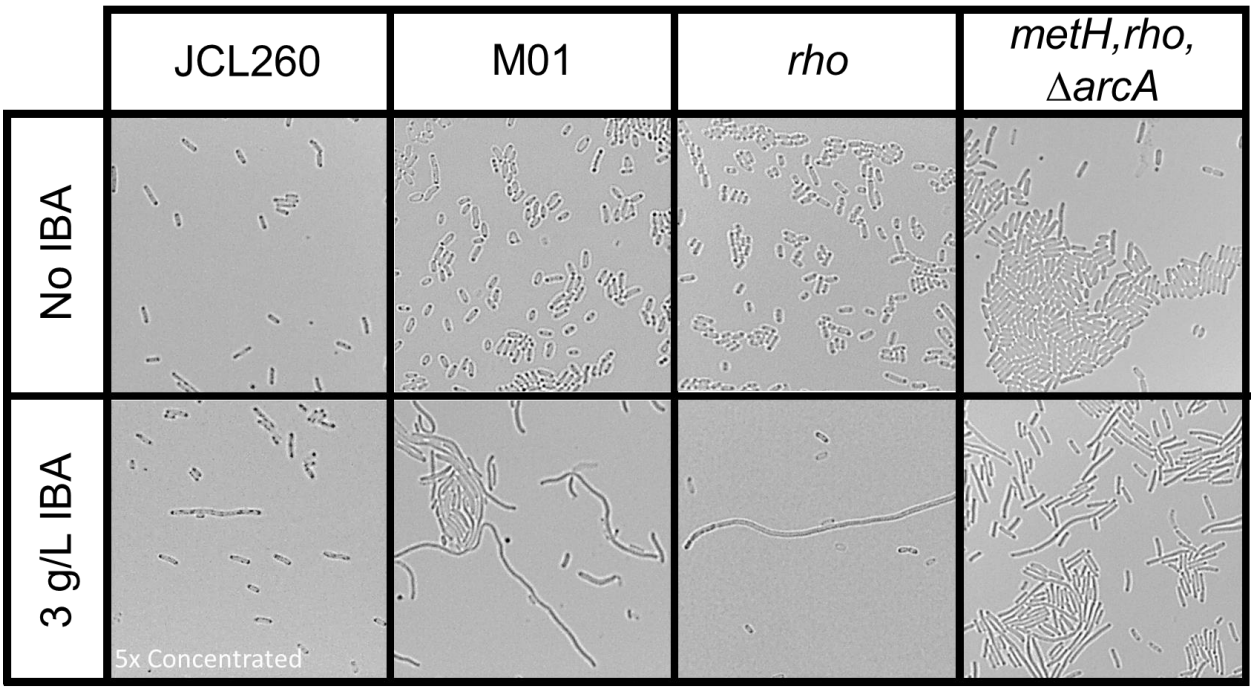


Fig. S3.4.

Microscope images taken at 100x (DM750 Microscope; Leica) of JCL260, M01, *rho*, and the triple mutant *metH, rho,* Δ*arcA* cultured in 5 mL of M9P minimal media containing 0 and 3 $g\,l^{-1}$ IBA after 24 h. 5 mL of the JCL260 culture supplemented with 3 g/L was spun down and resuspended in 500 μL media to allow for ample number of cells for comparison.

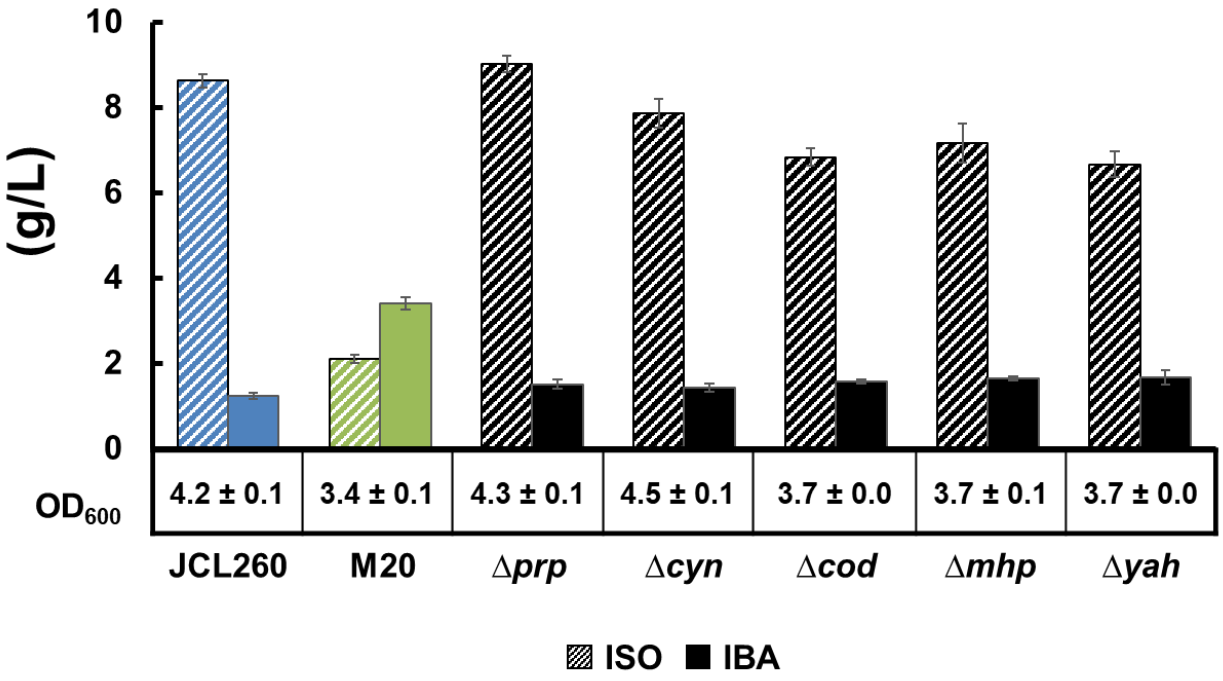


Fig. S3.5. Effect of mutations on IBA production.

IBA production (solid), ISO production (pattern) and final OD_{600} of the strains cultured in M9P media after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

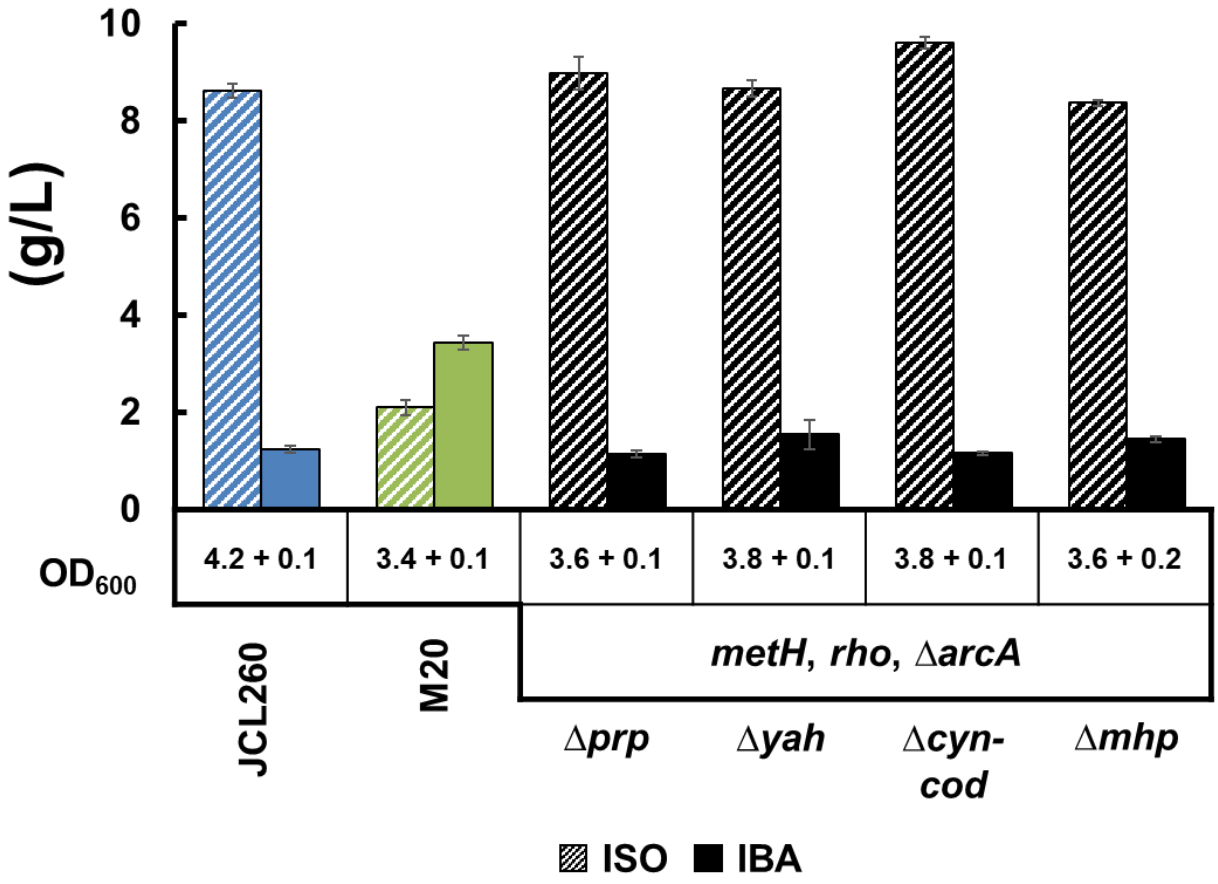


Fig. S3.6. Effect of mutations on IBA production in the metH, rho, ΔarcA triple mutant.

IBA production (solid), ISO production (pattern) and final OD_{600} of the strains cultured in M9P media after 24 h. N = 3 or more, biological replicates; error bars represent s.d.

3.6.1 Supplemental References

95 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645 (2000).

96 Atsumi, S. *et al.* Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab Eng* **10**, 305-311 (2008).

98 Tashiro, Y., Desai, S. H. & Atsumi, S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nature communications* **6**, 7488, doi:10.1038/ncomms8488 (2015).

107 Rodriguez, G. M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in *Escherichia coli*. *Nat Chem Biol* **10**, 259-265, doi:10.1038/nchembio.1476 (2014).

119 Jiang, Y. *et al.* Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **81**, 2506-2514, doi:10.1128/AEM.04023-14 (2015).

Chapter 4: Ongoing Work by the Atsumi Lab and Conclusions

4.1 Ongoing Work to Optimize our Electrochemical-biological Hybrid System to Fix CO_2

The expression of *E. coli's gcvTHP* from a high-copy plasmid significantly improved growth on CO_2 and formate, but the growth rate was still much slower than that with glycine feeding. These results suggest that rGCS is one of the limiting steps of the RGP. To improve the CO_2 fixation capabilities of our system, we will compare the rGCS activity of GCSs from various organisms to find candidate(s) for further improvement. Of particular interest is GcvP, the CO_2 fixing enzyme of the rGCS. Two species, *Arthrobacter globiformis* and *Clostridium acidiurici*, show relatively high specific activity among the reported GCSs and both are speculated to have functional rGCSs.[88,155] However, currently no GCS crystal structures exist for these two organisms or *E. coli*. Structural information is critical for interpreting the results of directed evolution[156] and designing modifications[157] for further improvements. While there are a variety of GcvTH and Lpd crystal structures accessible through the Protein Data Bank, only two bacterial GCSs that include GcvP (*Thermus thermophilus* (PDB 1wyt *et al.*) and *Synechocystis* sp. PCC 6803 (PDB 4lgl *et al.*)[158]) are available. Therefore, the rGCS activity of the four GCSs will be compared against the *E. coli* GCS using a glycine auxotrophic *E. coli* strain.

Having identified the most active rGCS, we will develop an evolutionary method to improve its CO_2 fixation efficiency by focusing on GcvP which catalyzes the synthesis of glycine from CO_2 and CH_2-THF. We believe *gcvP* is a good starting point for improving the rGCS since CO_2 binding and fixation are often the rate limiting steps.[159] We seek to overcome the limitations of rational strain design attributed to uncharacterized reaction mechanisms and unknown structural requirements by applying random mutagenesis. One of the most powerful approaches to improve enzymatic reactions is the creation of a library of variants with randomly altered function and application.[156] These libraries undergo selective pressure to isolate mutants with the desired

phenotype. Growth-dependent screening methods are commonly exploited due to the advantage of a built-in selective pressure, wherein only the desired phenotype is expected to survive and/or thrive.[160] This is beneficial in the face of unknown structure-function relationships, one of the most daunting investigative tasks of rational design. Directed evolution expedites the process of elucidating key residues by highlighting residues or regions of interest that can be further explored and understood through rational mutagenesis. We will apply this strategy to the chosen rGCS, specifically GcvP, to improve CO_2 fixation efficiency. The *gcvP* libraries will be subjected to a high-throughput, growth-based selection method wherein successful variants will see increased growth under glycine auxotrophy due to improvements in CO_2 fixation efficiency. The isolated variants will be used for understanding the reaction mechanism and how to improve CO_2 fixation efficiency.

In preparation for directed evolution, we will start by generating a *gcvP* mutant library. Mutagenesis can also be expanded to other GCS genes if necessary. The *gcvP* gene will be mutated using error-prone PCR[161] and cloned using restriction enzyme ligation into the GCS plasmid previously constructed. Error-prone PCR will be conducted using the GeneMorph II Random Mutagenesis Kit (Agilent). We will use a low mutation rate (1-5 mutations/1 kb) to avoid introducing negative mutations that impair activity. Highly competent (10^9 cfu/μg DNA) cells such as XL10 Gold Ultracompetent cells (Agilent) will be transformed with the ligation reaction. Previously, during work on an ARPA-E project to convert gaseous ethylene to liquid ethanol, I constructed a 1.2 x10^5 variant library of the oleate hydratase from *Elizabethkingia meningoseptica* (OhyA, EC 4.2.1.53). From our experience with this task, we are confident in this library generation method.

The variants generated above will be characterized *in vivo* using the growth-dependent screening method. We will then characterize the individual mutations *in silico* and *in vitro* and create several variant constructs that strategically combine beneficial mutations together. These variants will also be tested *in vivo* and compared against the growth rates of the strain expressing the wild-type GCS or a strain supplemented with glycine. Ultimately, our benchmark for successful improvement of function is increased *in vivo* CO_2 fixation into glycine to enhance growth. By combining mutations, we hope to observe synergistic effects in improving CO_2 fixation.

Once we have developed a more efficient RGP, it is critical to integrate the pathway into the genome. Plasmids, which serve as initial vehicles for gene expression, are unstable, can multimerize, result in varied copy numbers per cell, and often require antibiotics which increase the cost of production.[162] These qualities make plasmids undesirable for robust long-term production, but rarely adversely impact short-term screening of novel pathways or modifications. Enabling chromosomal RGP expression is also important for downstream production of chemicals from pyruvate. Initial screening necessitates production pathways be expressed using plasmids and additional plasmids would increase the burden placed on the cells, potentially decreasing cellular fitness and production.

Genome integration can be accomplished using the same CRISPR/Cas9 system mentioned in Chapter 2 and 3. Many regions in the *E. coli* genome are suitable for stable, high-efficiency integration. Deleted genes, such as *serA* or *glyA*, and intergenic locus SS9[163] will serve as the insertion site(s) for the RGP. To ensure efficient chromosomal expression we will assess the integrations using growth-based screening methods. Different integration sites and promoters can be used to optimize expression to meet or exceed plasmid-based performance.

For efficient production it is important to survey and fix induction issues present in our strains. Previously, it was shown that no induction and low concentration induction conditions were nearly identical, indicating weak or reduced repression. Using plasmids expressing sfGFP under $P_L lacO_1$ and $P_L tetO_1$, we will characterize the lac and tet inducible systems strength and repression in our strain. If a lack of repression is found, the Z1 fragment containing *lacI*^{q}*, tetR, and spec*^{r} can be integrated into the chromosome or *lacI*^{q} and/or tetR can be introduced onto a plasmid. The impact these integrations have on repression can be assessed using the same fluorescence screening. If these do not produce favorable results, promoter segments can be swapped out for stronger counterparts or other inducible systems, like P_{BAD}, can be used.

Installing the RGP into the genome and optimizing repression are critical steps to enabling production in our biological system. Isobutanol is a promising target chemical to screen the production potential of the RGP. Its five-step pyruvate dependent pathway requires only 2 NAD(P)H, which is critical in a system already starved for cofactors.[47] If necessary, we can explore methods to improve NAD(P)H regeneration. A variety of different enzymes and pathways have been used to increased cofactor availability, which in turn have enhanced production.[164] Isobutanol will serve as a potent proof of concept that guides future production improvements in our system.

In addition to the biological improvements, there is significant headway to be made on advancing the electrocatalytic component of our hybrid system. We have a collaboration with the Berben group, who were able to produce formate selectively with 96% Faradaic efficiency at pH 7 using a glass carbon electrode.[33] Current work is being done to create the bioreactor system and optimize the reaction conditions to match the biological system. Our ultimate goal is to deliver an efficient hybrid system with comparable production to traditional *E. coli* fermentation.

4.2 Conclusion

Metabolic engineering is a burgeoning field that continues to grow and expand as new pathways are developed, more hosts are investigated, carbon feedstock pools are expanded, and novel tools are created. Every year, more microbially produced chemicals are brought to market. Still, hundreds more chemicals need to be bio-produced to eradicate petrochemicals in the global industrial landscape. Continued microbial expansion and improvement requires tractable solutions to intricate challenges. The two works described above represent different avenues for improving strain design that can be strategically applied to achieve economically competitive and truly sustainable microbial production.

Our electrical-biological hybrid system demonstrates *E. coli* can fix CO_2 and integrate that carbon into central metabolism. The RGP strives to meet the sustainability of a phototroph while delivering the output of a heterotroph. Yet, coupled electrocatalysis and microbial production remain relatively uncharted. This results in unanswered questions and potential problems that need to be resolved before any hybrid system can become an economically viable production platform. It is still unclear whether direct or indirect electron transfer will triumph as a superior strategy. Highly specialized and expensive bioreactors may be required for direct electron transport scale up strategies. On the other hand, more studies are required to design a microbial system capable of neutralizing the ROS's deleterious effects or a system entirely devoid of ROS. Challenges like these are typical of budding technologies and it will be exciting to see how they are addressed as research into these areas continues.

The success of our IBA-tolerant strains marks an important step in providing new possibilities for overcoming product toxicity. This work highlights how complicated genotype-phenotype relationships are and how evolutionary processes can be exploited to achieve desirable

outcomes. Simultaneously, it exposes the intensive proteomic and metabolomic studies required to understand the causal factors associated with metabolic perturbations. Evolutionary studies tap into unexplored or overlooked areas that can provide powerful insight into how organisms operate and respond to stimulus. More work needs to be done to understand how the elucidated mutations impact *E. coli*'s ability to tolerate IBA and other chemicals.

By delivering sustainable, long-term energy and chemical security, engineered microbes are poised to usurp petroleum as the central producer in many industries. Research continues to drive innovation by enabling a wide variety of chemicals to be produced in increasing renewable ways. Making microbial production economically viable is not only critical for expanding the field, but for the future of humanity.

4.3 References

33 Taheri, A., Thompson, E. J., Fettinger, J. C. & Berben, L. A. An Iron Electrocatalyst for Selective Reduction of CO2 to Formate in Water: Including Thermochemical Insights. *ACS Catalysis* **5**, 7140-7151, doi:10.1021/acscatal.5b01708 (2015).

47 Atsumi, S. *et al.* Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Microbiology and Biotechnology* **85**, 651-657, doi:10.1007/s00253-009-2085-6 (2010).

88 Gariboldi, R. T. & Drake, H. L. Glycine synthase of the purinolytic bacterium, Clostridium acidiurici. Purification of the glycine-CO2 exchange system. *Journal of Biological Chemistry* **259**, 6085-6089, doi:10.1016/s0021-9258(20)82108-5 (1984).

155 Kochi, H. & Kikuchi, G. Mechanism of the reversible glycine cleavage reaction in Arthrobacter globiformis: I. purification and function of protein components required for the reaction. *Journal of Biochemistry* **75**, 1113-1127, doi:10.1093/oxfordjournals.jbchem.a130483 (1974).

156 Rubin-Pitel, S. B., Cho, C. M. H., Chen, W. & Zhao, H. Directed Evolution Tools in Bioproduct and Bioprocess Development. 49-72, doi:10.1016/B978-044452114-9/50004-9 (2007).

157 Schwarte, A. *et al.* NewProt - A protein engineering portal. *Protein Engineering, Design and Selection* **30**, 441-447, doi:10.1093/protein/gzx024 (2017).

158 Hasse, D. *et al.* Structure of the homodimeric glycine decarboxylase P-protein from synechocystis sp. PCC 6803 suggests a mechanism for redox regulation. *Journal of Biological Chemistry* **288**, 35333-35345, doi:10.1074/jbc.M113.509976 (2013).

159 Farazdaghi, H. The single-process biochemical reaction of Rubisco: A unified theory and model with the effects of irradiance, CO2 and rate-limiting step on the kinetics of C3 and C4 photosynthesis from gas exchange. *BioSystems* **103**, 265-284, doi:10.1016/j.biosystems.2010.11.004 (2011).

160 Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nature Reviews Genetics* **16**, 379-394, doi:10.1038/nrg3927 (2015).

161 Cadwell, R. C. & Joyce, G. F. Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**, 28-33 (1992).

162 Saleski, T. E. *et al.* Optimized gene expression from bacterial chromosome by high-throughput integration and screening. *Science Advances* **7**, eabe1767 doi:10.1126/sciadv.abe1767 (2021).

163 Bassalo, M. C. *et al.* Rapid and Efficient One-Step Metabolic Pathway Integration in E. coli. *ACS Synthetic Biology* **5**, 561-568, doi:10.1021/acssynbio.5b00187 (2016).

164 Spaans, S. K., Weusthuis, R. A., van der Oost, J. & Kengen, S. W. M. NADPH-generating systems in bacteria and archaea. *Frontiers in Microbiology* **6**, doi:10.3389/fmicb.2015.00742 (2015).