UNIVERSITY OF CALIFORNIA

Los Angeles

Design and construction of a synthetic CO_2 fixation cycle *in vitro*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Shanshan Luo

2018

ABSTRACT OF THE DISSERTATION

Design and construction of a synthetic CO_2 fixation cycle *in vitro*

by

Shanshan Luo

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2018

Professor James C. Liao, Chair

The increasing atmospheric carbon dioxide (CO_2) level is calling for more efficient CO_2 fixation systems to re-balance the carbon cycle. At the same time, atmospheric CO_2 can in principle be a cheap and abundant carbon source for synthesizing various organic compounds. Converting CO_2 into useful chemicals can be an ideal solution for a sustainable future. However, due to its high thermodynamic stability and kinetic inertness, the chemical transformation of CO_2 into organic compounds usually requires quite harsh conditions. In contrast, nature has evolved several CO_2 fixation pathways for synthesizing longer carbon chain compounds from CO_2 under ambient conditions. To date, six CO_2 fixation pathways have been reported. However, they all contain key carboxylases with intrinsic properties such as structural complexity and oxygen sensitivity

that hinder their further engineering and applications. To circumvent these problems, we designed a synthetic CO_2 fixation cycle composed of all efficient, simple and oxygen tolerant enzymes and demonstrated its function *in vitro*.

In this work, we utilized the two most active and oxygen-insensitive carboxylases, crotonyl-CoA carboxylase/reductase (Ccr) and phosphoenolpyruvate carboxylase (Ppc) to fix CO_2 and bicarbonate, respectively. Our designed cycle is composed of the reductive glyoxylate synthesis (rGS) pathway and the reductive pyruvate synthesis (rPS) pathway. The rGS pathway converts one pyruvate to one glyoxylate and one acetyl-CoA fixing one bicarbonate, while the rPS pathway converts one acetyl-CoA to pyruvate with one CO_2 fixed. Together, we obtained a CO_2 fixation cycle termed the reductive glyoxylate/pyruvate synthesis (rGPS) cycle.

We first validated the feasibility of rGPS with its stepwise reconstitution *in vitro.* Using a two-pot immobilized enzyme system, we were able to run the rGPS cycle semi-continuously. To further improve its efficiency, we developed electrochemical and enzymatic FAD regeneration systems to run the rGPS cycle continuously in one pot. Using the enzymatic FAD regeneration method increased the CO_2 fixation rate of rGPS to 66.9 nmol min^{-1} mg^{-1} of core cycle proteins.

We further explored the construction of rGPS *in vivo*. The rGS part has been demonstrated *in vivo.* The rPS part has not worked in *E. coli* for growth rescue, although all ten enzymes of rPS were confirmed to be active in crude extract assays. Finally, strategies for constructing a fully functional rGPS *in vivo* have been proposed.

The dissertation of Shanshan Luo is approved.

Yvonne Chen

Yi Tang

Todd O. Yeates

James C. Liao, Committee Chair

University of California, Los Angeles

2018

To my parents,

for always loving and supporting me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. James C. Liao for giving me the opportunity to do what I am passionate about and for his patience, guidance, and support throughout my graduate study. I would also appreciate my other committee members, Dr. Yvonne Chen, Dr. Yi Tang, and Dr. Todd O. Yeates, for their valuable advice on my research.

I would like to extend my sincere appreciation to my colleagues and friends in the Liao lab, especially Dr. Claire Shen and Dr. Paul Lin for teaching me important skill sets; Dr. Ethan Lan, Dr. Tung-Yun Wu, Dr. Igor Bogorad, Dr. Yi-Xin Huo, Dr. Han Li, Dr. Hong Yu, Dr. Xiaoqian Li, and Dr. Sio So Wong for their generous help and valuable insight. The knowledge, inspiration, and strength I gained from them has been an indispensible part in my research training.

The work described here would not be possible without the efforts of my colleague Dr. Hong Yu and my collaborator Dr. Bo Wang. I would also like to thank Dr. Tung-Yun Wu and Dr. Fabienne Duchoud for providing plasmids, Hans Sebastian and Ivan C. Lam for assistance in protein purifications, and Yan Yan for assistance in performing LC-MS.

Finally, I would like to express my deepest gratitude to my parents Binhong Luo and Xiuying Zhao, my boyfriend Ben Li for their unconditional love and support.

VITA

EDUCATION

2011 B.S., Chemical Engineering

Tsinghua University, P.R. China

PUBLICATIONS

Luo S, Yu H, Wong B, Liao JC (2018) A cell-free synthetic CO_2 fixation cycle with specific rate comparable to the Calvin cycle. In preparation.

Liao JC, Mi L, Pontrelli S, Luo S (2016) Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat Rev Microbiol* 14(5):288–304

Chapter 1. Background: current status of CO_2 capture and utilization

1.1 Introduction

The increasing of atmospheric greenhouse gasses (GHGs) from anthropogenic activities has been recognized as a significant contributor to global warming and climate changes (1). Carbon dioxide (CO_2), as the major GHG, accounts for about three-fourths of total global GHG emissions (2). The global average atmospheric CO_2 concentration in 2017 has reached 405 parts per million (ppm) (3), which has increased more than 44%, from 280 ppm during pre-industrial time. However, comprehensive researches have suggested that the safe upper level of atmospheric CO_2 is 350 ppm (4). To avoid the worst effects of climate change occurring, different approaches should be considered and adopted to get a substantial decrease in carbon emission, including enhancing energy efficiency and energy conservation; increasing usage of low carbon fuels and renewable energy; afforestation and reforestation; and CO_2 capture and storage (CCS). This chapter will review current status of CO_2 capture and utilization: mainly on CCS, chemical and biological approaches for CO_2 conversion and examples of CO_2 based biofuels.

1.2 CO_2 capture and storage

Carbon dioxide (CO_2) capture and storage (CCS) (or CO_2 capture and sequestration) is a set of technologies that can capture up to 90% of the waste CO_2 from large point emission sources, such as power plants and industrial processes, preventing large quantities of CO_2 emissions from entering the atmosphere. It is a three-step process, including capturing CO_2 emissions; transporting the captured and compressed CO_2 to a storage site and securely storing it for a long-term isolation from the atmosphere.

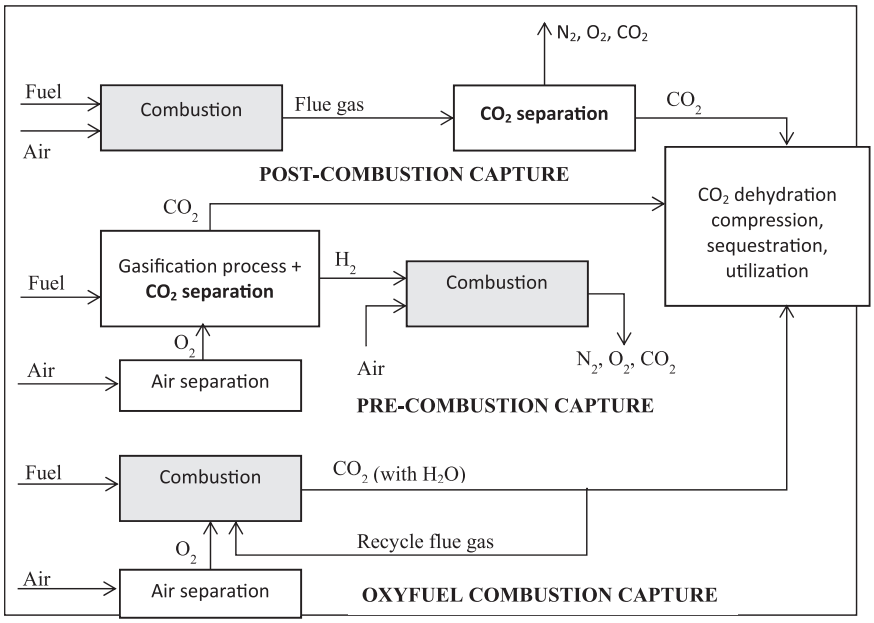


Fig. 1.1. CO_2 capturing processes (6).

There are three main CO_2 capture systems: post-combustion, pre-combustion and oxyfuel combustion (Fig 1.1). Post-combustion process removes CO_2 from the flue gas after the combustion using absorption, adsorption or membrane gas separation technologies. It is currently the most mature process for CO_2 capture and mainly applied to both coal and gas fired plants (5). In the pre-combustion process, the fuel (normally coal or natural gas) goes through pretreatment steps to be converted into a H_2/CO_2 gas mixture before combustion. The high CO_2 concentration in the H_2/CO_2 gas mixture facilitates the further CO_2 separation (5). In the oxyfuel combustion process, only oxygen (O_2) is supplied for the combustion process. This results in a flue gas mixture of mainly CO_2 and condensable water vapor that can be cleaned and separated relatively easily. This process is technically feasible but has increased energy requirements in the separation of oxygen from air (6).

After the CO_2 gets captured, it needs to be transported to the storage sites. Pipelines are

considered to be the most viable method for onshore transporting of high volume of CO_2 for distances up to around 1,000 km (7). For amounts smaller than a few million tons of CO_2 per year or for larger distances overseas, ships could be more economically attractive (7). Rail and road tankers can also be used for small-scale CO_2 transportation.

For CO_2 storage, four main types of geological formations are considered, including depleted oil and gas reservoirs; unmineable coal beds; saline aquifers and besalts (5). Potential CO_2 leakage is a major concern for CO_2 geological sequestration, while risk assessment and monitoring techniques are still being researched (6).

Although CSS is technically feasible, its high overall cost and intensive energy demand still prevent this process from being widely deployed.

1.3 Chemical approaches for CO_2 conversions

Instead of storing CO_2 after capture, reusing captured CO_2 and converting it into useful chemicals and fuels can be a more sustainable and longer-term solution. CO_2, as the ultimate carbon source for all organic molecules on earth, can theoretically be a cheap and abundant feedstock for organic synthesis. However, the total amount of CO_2 used today for producing chemicals is only a very minor fraction of the emitted CO_2 (170 Mt/y with respect to 32,000 Mt/y, ~0.53%) (8). Converting CO_2 into organic compounds is still one of the biggest challenges in synthetic organic chemistry.

CO_2 has the most oxidized state of carbon, making it a thermodynamically stable molecule involving pretty low energy content (ΔG_f^0 = -394 kJ/mol in gas phase). As a result, a substantial energy input is required for reducing CO_2 to energy dense fuels. Structurally, CO_2 is a linear,

nonpolar molecule with a double bond between the carbon and oxygen atoms (O=C=O). Generally, it is kinetically intern. However, owing to the electron deficiency of its carbonyl carbon, CO_2 has a strong affinity toward nucleophiles, in other words, electron-donating reagents. Many nucleophiles, such as organometallics, amines, alkoxides, and even water, readily react with CO_2 and lead to the formation of organic carbonates and carbamates. Thus, CO_2 transformation reactions can be categorized into two main processes (Fig 1.2) (8):

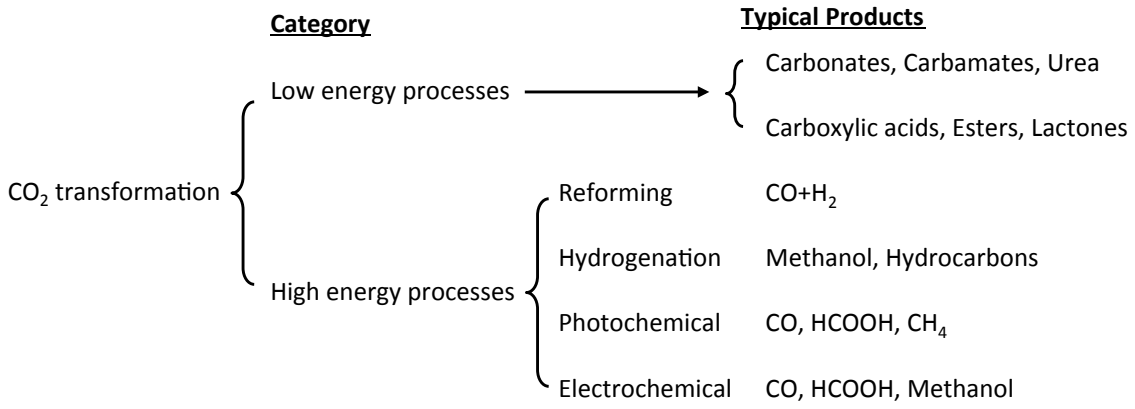


Fig. 1.2. Typical chemical approaches for CO_2 transformation.

1. Low energy processes, in which CO_2 reacts with electron-rich or high-energy compounds (such as amine, epoxides, alkane, alkyne, aromatics, diene, etc), producing low-energy products containing carbon at high oxidation state: +4 (such as carbonates, carbamates, urea and urethane derivatives) or +3 (such as carboxylic acids, esters and lactones) (9). Such reactions can occur without intense energy input.
2. High energy processes, in which CO_2 is reduced to form higher energy products, such as CO, CH_3OH, CH_4, hydrocarbons, and similar compounds. All these processes need to be driven by external energy, such as heat, light (photochemical reduction) or electricity (electrochemical reduction) (10).

Utilizing CO_2 for organic synthesis can date back to the late 1800s. In 1860, salicylic acid was produced from phenol salts and CO_2 using the Kolbe–Schmitt reaction (11). Then, the synthesis of urea from ammonia and CO_2 was achieved in 1922 (12). While, the synthesis of methanol using syngas and CO_2 as well as the carboxylation of epoxides were developed recently in 1970s (8). Despite an impressive amount of CO_2-based synthetic schemes have been developed, only a few processes are industrialized. Among them, the synthesis of urea is the biggest consumer for CO_2, consuming 115 Mt/y (~ 2/3 of the total CO_2 used for chemical production).

The electrochemical and photochemical reduction of CO_2 into energy storage chemicals has attracted extensive research interest in recent years. This approach, mimicking natural photosynthesis, is a promising means of storing the intermittent renewable energy (geothermal, solar, and wind energy) in chemical form. However, usually these processes are only able to produce rather simple molecules such as carbon monoxide, methane, formic acid and methanol (13). Moreover, low energy efficiency and productivity are often associated with electro- or photocatalytic CO_2 transformations (10). Significant advances in developing high energy efficient electro- or photochemical systems with high efficient, high selectivity and high stability electro- or photocatalysts are still required for their large-scale application (14).

1.4 Natural CO_2 fixation pathways

Autotrophic organisms, which can assimilate CO_2 into complex organic molecules via reduction, play a crucial role in the global carbon cycle. Depending on the type of reducing power used, they can be divided into chemolithoautotrophs and photoautotrophs. Chemolithoautotrophs derive electrons for CO_2 reduction from inorganic substrates, such as hydrogen, H_2S, elemental sulfur, CO, ammonia and reduced metal ions. While, photoautotrophs harness energy from

sunlight and get electrons from splitting water with O_2 released. Autotrophs from different ecological niches can employ different pathways to fix CO_2. To date, six CO_2 fixation pathways have been identified: (1) the Calvin-Benson-Bassham cycle (CBB cycle) (15), (2) the reductive citric acid (rTCA) cycle (16), (3) the reductive acetyl-CoA (Wood-Ljungdahl) pathway (17, 18), (4) the 3-hydroxypropionate bicycle (19–21), (5) the 3-hydroxypropionate/4-hydroxybutyrate cycle (22), and (6) the dicarboxylate/4-hydroxybutyrate cycle (23).

1.4.1 Calvin-Benson-Bassham cycle

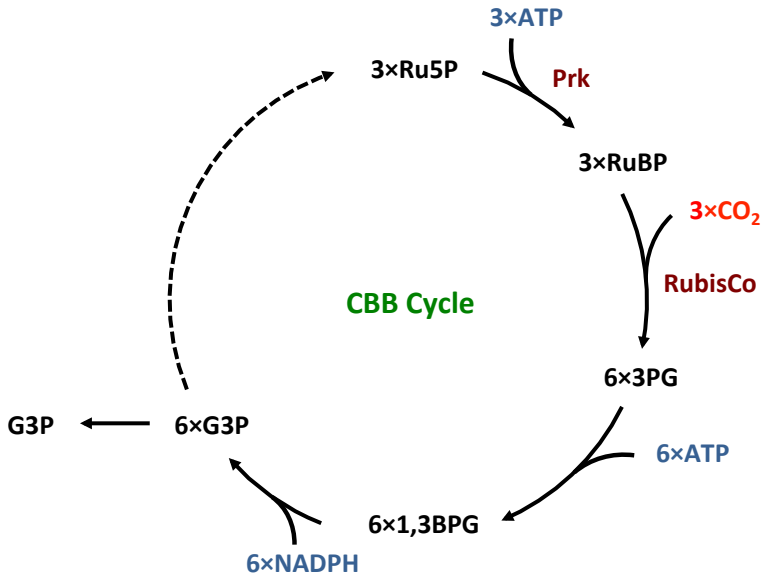


Fig. 1.3. Simplified Calvin-Benson-Bassham cycle.

The dash line presents multiple steps. Ru5P, ribulose 5-phosphate; 3PG, 3-phosphoglycerate; 1,3BPG, 1,3-bisphosphoglycerate; G3P, glyceraldehyde 3-phosphate.

The complete CBB cycle (or reductive pentose phosphate cycle) (Fig 1.3) was identified in 1954 by the research group of Melvin Calvin using the radioactive isotope C14 (15). The majority of atmospheric CO_2 is assimilated through the CBB cycle in the photosynthesis process. All plants and algae use this cycle to fix CO_2. While it is also found to be operating in some eubacteria,

such as cyanobacteria, some aerobic or facultative anaerobic Proteobacteria, CO-oxidizing mycobacteria and representatives of the genera *Sulfobacillus* (iron- and sulphur-oxidizing Firmicutes) and *Oscillochloris* (green sulphur bacteria) (24).

CO_2 is incorporated into the cycle by the key carboxylase, ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO), which catalyzes the carboxylation of ribulose 1,5-bisphosphate (RuBP), yielding two 3-phosphoglycerate. Then, 3-phosphoglycerate is converted to glyceraldehyde-3-phosphate through phosphorylation and reduction steps. Eventually, RuBP is regenerated from glyceraldehyde-3-phosphate through a series of carbon rearrangement reactions.

RubisCo and phosphoribulokinase (Prk) are essential enzymes for a functional CBB cycle. As the characteristic enzyme of CBB cycle, RubisCo has been intensively studied. In nature, four forms of RubisCo have been discovered. Form I, II and III RubisCo catalyzes the carboxylation and oxygenation of RuBP, while form IV, the RubisCo-like protein, does not catalyze either of these reactions (25). RubisCo enables the biosynthesis of organic carbon from atmospheric CO_2, playing a key role in the ecosphere. Despite its obvious importance, RubisCo is not a perfect carboxylase because of two major flaws: 1) RubisCO is a slow enzyme with an average turnover rate of 5 s^{-1}, making the carboxylation reaction the primary rate-limiting step in the CBB cycle. To compensate this sluggish catalytic performance, enormous amounts of RubisCo are produced for carbon fixation. As a result, RubisCo is the most abundant single protein on earth (26). 2) RubisCo is promiscuous, as it also accepts O_2 as a substrate and catalyzes the oxygenation of RuBP, yielding one 3-phosphoglycerate and one 2-phosphoglycolate. The produced 2-phosphoglycolate cannot reenter the CBB cycle and needs to go through a series of detoxification reactions, which inevitably wastes both carbon and energy. Interestingly, for

RubisCo, a negative correlation exists between its activity and selectivity (27). This indicates that RubisCo might already be naturally optimized (28), suggesting the difficulty of its further improvement.

1.4.2 Reductive citric acid (rTCA) cycle

In 1966, the first alternative carbon fixation pathway besides the CBB cycle, the reductive citric acid (rTCA) cycle (Fig 1.4), was proposed in the green sulfur bacterium *Chlorobium limicola* (16). To honor its discoverers, this cycle is also known as the Arnon-Buchanan cycle. Unlike the CBB cycle, the rTCA cycle involves oxygen sensitive enzymes. Thus, this cycle is only found in anaerobes or microaerobes, including some types of Proteobacteria, green sulphur bacteria and members of the microaerophilic *Aquificae* phylum (24).

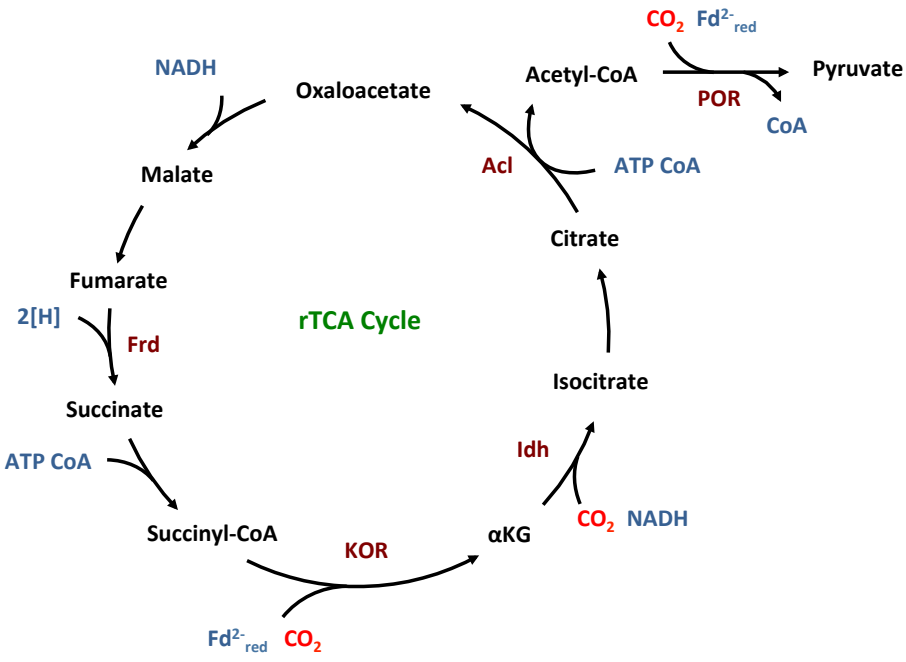


Fig. 1.4. Reductive citric acid (rTCA) cycle.
αKG, α-ketoglutarate; Fd^{2-}_{red}, reduced ferredoxin.

As its name indicates, the rTCA cycle is essentially a reversal (the reductive direction) of the citric acid (TCA) cycle. One acetyl-CoA can be oxidative degraded into two CO_2 molecules through the TCA cycle with reducing power and energy produced. Consequently, the reverse process of the TCA cycle, the rTCA cycle, can reduce two CO_2 into one acetyl-CoA with reducing power and energy consumed. Most enzymes that catalyze reversible reactions in the TCA cycle are also used in the rTCA cycle. Some irreversible reactions are achieved by unique enzymes in the rTCA cycle, including the citrate cleaving enzymes, fumarate reductase and α-ketoglutarate: ferredoxin oxidoreductase (KOR). The ATP-dependent cleavage of citrate to acetyl-CoA and oxaloacetate is one of the key steps in the rTCA cycle. It can be catalyzed by ATP-citrate lyase (Acl) or by the combination of citryl-CoA synthetase and citryl-CoA lyase. Also, several forms of fumarate reductase (Frd) are reported and can be divided into two major groups: membrane bound Frd and soluble Frd (29).

In order to synthesize pyruvate from acetyl-CoA, another key carboxylase, pyruvate: ferredoxin oxidoreductase (POR), is required. Thus, three carboxylases are essential for operating the rTCA cycle for autotrophic growth. Isocitrate dehydrogenase (Idh) is an NAD(P)H-dependent enzyme and catalyzes both carboxylation and decarboxylation reaction. This carboxylation reaction can also be achieved by the combination of 2-oxoglutarate carboxylase and oxalosuccinate reductase, with oxalosuccinate as the intermediate (30). Both KOR and POR are ferredoxin dependent carboxylating enzymes, which are highly oxygen sensitive because they contain the easily oxidized iron sulfur clusters. Interestingly, in *Hydrogenobacter thermophiles*, two distinct forms of KOR are identified: one two-subunit KOR, which is oxygen sensitive and another five-subunit KOR, which is more oxygen tolerant and expressed only in aerobic condition (31). However, the

activity of the oxygen-tolerant KOR is more than one order of magnitude lower than the oxygen-sensitive one (32).

1.4.3 Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway)

The third CO_2 fixation pathway, the reductive acetyl-CoA pathway or Wood-Ljungdahl (WL) pathway (Fig 1.5), was discovered and elucidated in acetogenic bacteria mainly by the laboratories of Harland G. Wood and Lars G. Ljungdahl (17, 18). This pathway is operating in strictly anaerobic organisms. Besides acetogenic bacteria, it is also found in autotrophic sulfate-reducing bacteria and archaea, methanogenic archaea and some planctomycetes (33).

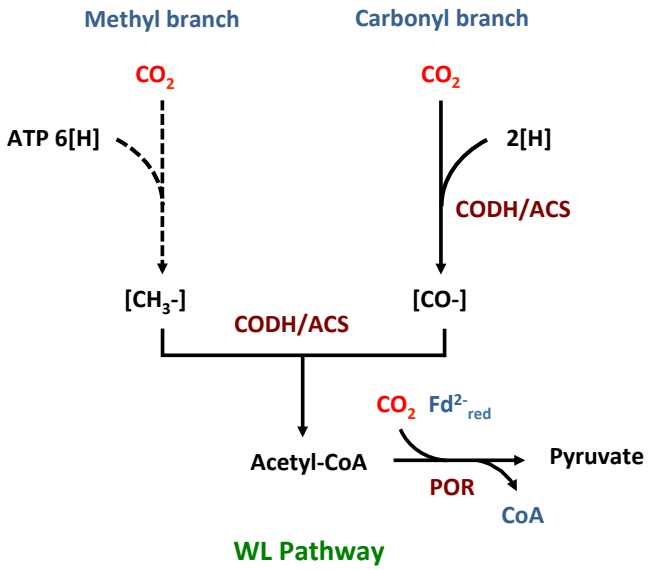


Fig. 1.5. Simplified reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) in acetogens.

Dash lines present multiple steps.

In contrast to the cyclic CBB and rTCA pathways, the WL pathway is a linear pathway, where two molecules of CO_2 are combined directly into one acetyl-CoA. The pathway consists of two

branches: the methyl branch (eastern branch), where one molecule of CO_2 is reduced to a carrier-bound methyl group, and the carbonyl branch (western branch), where another CO_2 molecule is reduced to a enzyme-bound CO. The methyl group and CO are then condensed together with CoA to synthesize one acetyl-CoA through the key enzyme of this pathway, CO dehydrogenase/acetyl-CoA synthase (CODH/ACS), which also catalyzes the reduction of CO_2 to CO.

The methyl branch is composed of a series of reactions and distinct differences exist between bacteria and archaea. In bacteria, the methyl branch starts with reducing CO_2 to formate, which is then condensed with H_4folate at the expense of one ATP. The formed 10-formyl-H_4folate is consecutively reduced to 5-methyl-H_4folate through a series of enzymes. Finally, the methyl group of 5-methyl-H_4folate is transferred to the cobalt center of a corrinoid iron–sulfur-protein with H_4folate released for reentering the pathway. While in archaea, instead of forming formate, CO_2 is first reduced the formyl group of formylmethanofuran. In addition, the C1 carrier used in archaea is H_4methanopterin not H_4folate. On the contrary, the carbonyl branch, a one-step process catalyzed by CODH/ACS, is conserved in bacteria and archaea. The key enzyme, CODH/ACS, probably originates from the same root in bacteria and archaea (24).

The WL pathway is the most energetically efficient CO_2 fixation pathway so far, requiring probably less than one ATP to make one pyruvate from CO_2 (24). It is also presumably the most ancient autotrophic carbon fixation pathway (33). However, the involvement of organometallic enzyme–substrate intermediates, the requirement of special cofactors and extremely sophisticated enzymes, limit the WL pathway to completely oxygen-free ecological niches.

1.4.4 3-Hydroxypropionate (3-HP) bicycle

The 3-hydroxypropionate (3-HP) bicycle (Fig 1.6) was discovered in the phototrophic green non-sulfur bacterium *Chloroflexus aurantiacus* (19–21) and seems to be restricted to members in the family Chloroflexaceae (33). The pathway uses ATP- and biotin-dependent acetyl-CoA and propionyl-CoA carboxylases for carboxylation reactions. All enzymes involved in this bicycle are oxygen insensitive, making it a potential candidate for transferring aerobic heterotrophs to autotrophs (34).

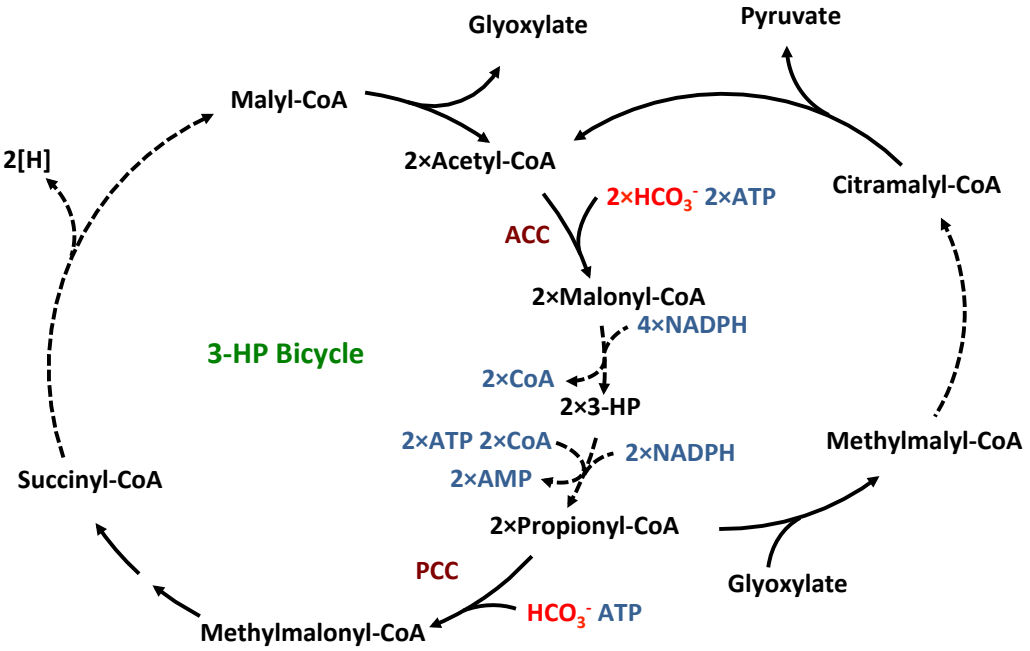


Fig. 1.6. Simplified 3-Hydroxypropionate (3-HP) bicycle.
Dash lines present multiple steps. ACC, acetyl-CoA carboxylase; PCC, propionyl-CoA carboxylase.

The 3-HP bicycle consists of two cycles as its name indicates. The net reaction of the first cycle is fixing two molecules of bicarbonate into one glyoxylate molecule. The first cycle starts from

carboxylating acetyl-CoA to malonyl-CoA using acetyl-CoA carboxylase with one bicarbonate fixed. Malonyl-CoA is then consecutively reduced to propionyl-CoA through two enzymes, the bifunctional malonyl-CoA reductase and the trifunctional propionyl-CoA synthase. Sequentially, propionyl-CoA is carboxylated to methylmalonyl-CoA with propionyl-CoA carboxylase fixing another molecule of bicarbonate. After isomerization, CoA transfer, oxidation and dehydration, methylmalonyl-CoA is converted to (*S*)-malyl-CoA, which is then cleaved into acetyl-CoA and glyoxylate. Acetyl-CoA reenters into the first cycle and glyoxylate is incorporated into the second cycle. In the second cycle, glyoxylate is condensed with propionly-CoA (produced from acetyl-CoA in the first cycle) to form β-methylmalyl-CoA. After a series of reactions, β-methylmalyl-CoA is converted to (*S*)-citramalyl-CoA, which is cleaved into acetyl-CoA and pyruvate. In summary, the 3-HP bicycle fixes three molecules of bicarbonate into one molecule of pyruvate.

Apart from bifunctional malonyl-CoA reductase and trifunctional propionyl-CoA synthase, another key enzyme, malyl-CoA/β-methylmalyl-CoA/citramalyl-CoA lyase, is also a multifunctional enzyme catalyzing three reactions (21). Thus, the 3-HP bicycle includes 19 reactions, but only 13 enzymes are involved.

1.4.5 3-hydroxypropionate/4-hydroxybutyrate (HP/HB) cycle

The 3-hydroxypropionate/4-hydroxybutyrate (HP/HB) cycle (Fig 1.7) was first discovered in autotrophic thermoacidophilic crenarchaeon *Metallosphaera sedula* (22). Later on, this pathway was also found in other Crenarchaeota, including the strictly anaerobic *Stygiolobus azoricus* (35). Recently, a more energy efficient version of the HP/HB cycle was identified in the marine archaeon *Nitrosopumilus maritimus*, a representative of the phylum Thaumarchaeota (36).

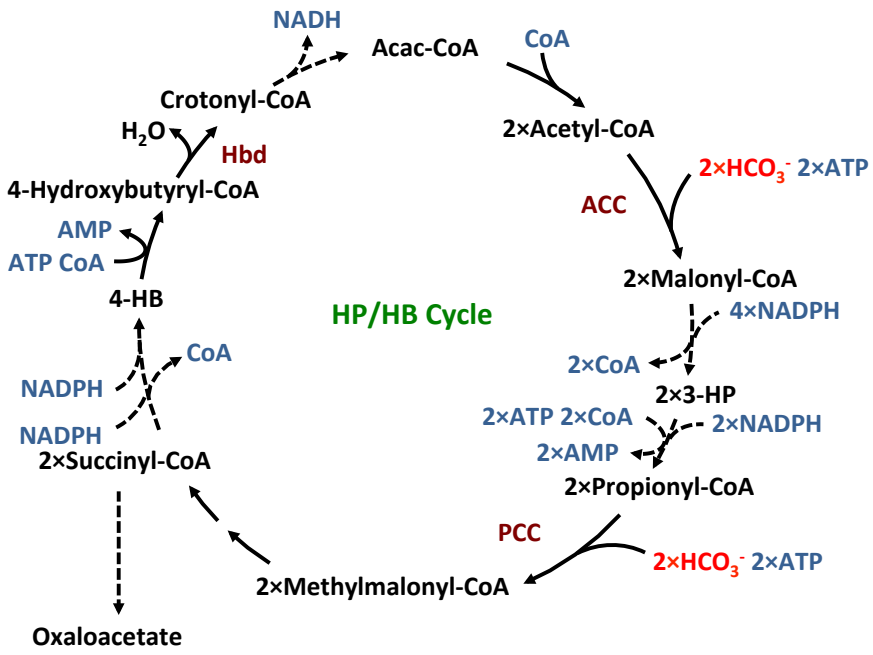


Fig. 1.7. Simplified 3-hydroxypropionate/4-hydroxybutyrate (HP/HB) cycle.
Dash lines present multiple steps. Acac-CoA, acetoacetyl-CoA.

Briefly, the HP/HB cycle fixes two molecules of bicarbonate into one molecule of acetyl-CoA, using the bifunctional ATP- and biotin-dependent acetyl-CoA/propionyl-CoA carboxylase as the carboxylting enzyme. The cycle can be divided into two parts. The first part converts one acetyl-CoA and two bicarbonate molecules to succinyl-CoA through 3-hydroxypropionate. The reaction sequence of this part is identical to the 3-HP bicycle as described above. However, enzymes involved in converting malonyl-CoA to propionyl-CoA are not homologous, suggesting the HP/HB cycle and the 3-HP bicycle have evolved independently (37). The second part converts succinyl-CoA to two molecules of acetyl-CoA via 4-hydroxybutyrate. This part involves the key enzyme of this cycle, the flavin-adenine-dinucleotide (FAD)- and [4Fe-4S]-containing 4-hydroxybutyryl-CoA dehydratase, which converts 4-hydroxybutyryl-CoA to crotonyl-CoA using

a ketyl radical mechanism (38, 39). The product, crotonyl-CoA, is eventually converted into two acetyl-CoA molecules via β-oxidation reactions. 4-hydroxybutyryl-CoA dehydratase (Hbd) has also been reported in anaerobic clostridia, where it is oxygen sensitive. While, 4-hydroxybutyryl-CoA dehydratase from *Metallosphaera sedula* (40) and *Nitrosopumilus maritimus* (36) are robust in the presence of oxygen. The enzymes involved in the HP/HB cycle are oxygen tolerant, making the aerobic operation of this cycle possible.

The recent identified thaumarchaeal HP/HB cycle is a more energy efficient version. Instead of using AMP-producing enzymes for 3-hydroxypropionate and 4-hydroxybutyrate activation, this version uses ADP-producing enzymes, saving two high-energy bonds every turn of this cycle. So far, this version of HP/HB cycle is also the most energy-efficient aerobic natural CO_2 fixation pathway (36). Moreover, the phylogenetic analysis indicates a convergent evolution of the HP/HB cycle in Crenarchaeota and Thaumarcharota (36).

1.4.6 Dicarboxylate/4-hydroxybutyrate (DC/HB) cycle

The most recently discovered CO_2 fixation pathway, the dicarboxylate/4-hydroxybutyrate (DC/HB) cycle (Fig 1.8), was first elucidated in the anaerobic hyperthermophilic crenarchaeon *Ignicoccus hospitalis* (*Desulfurococcales*) (23). Its occurrence has been extended to anaerobic members of *Desulfurococcales* and *Thermoproteales* (41), as well as the facultative aerobic *Pyrolobus fumarii* (*Desulfurococcales*) (35).

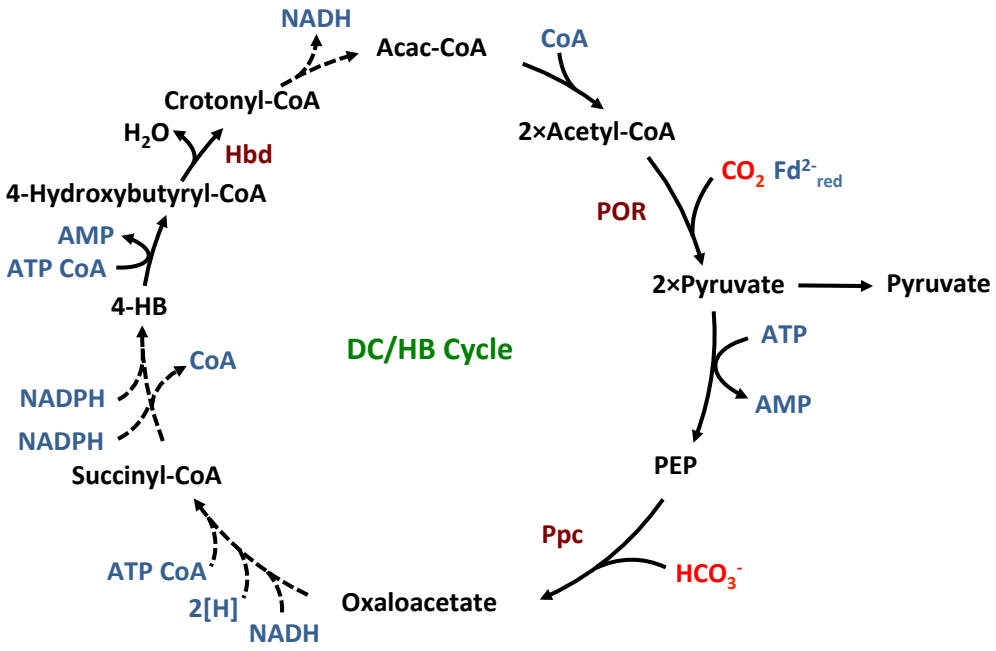


Fig. 1.8. Simplified dicarboxylate/4-hydroxybutyrate (DC/HB) cycle.

Dash lines present multiple steps. Acac-CoA, acetoacetyl-CoA; PEP, phosphoenolpyruvate.

Like the HP/HB cycle, the DC/HB cycle can also be divided into two parts. The first part starts from acetyl-CoA, which is reductively carboxylated to pyruvate using pyruvate: ferredoxin oxidoreductase (POR) at the expense of two equivalents of reduced ferredoxin. The product, pyruvate, is first activated by phosphorylation and further carboxylated to oxaloacetate using phosphoenolpyruvate carboxylase (Ppc). Finally, oxaloacetate is converted to succinyl-CoA using the same reaction sequence and enzymes in the rTCA cycle. The second part regenerates acetyl-CoA from succinyl-CoA through a route identical to the second part of the HP/HB cycle. In summary, the DC/HB cycle fixes one CO_2 and one bicarbonate molecules into one acetyl-CoA molecule using POR and Ppc as carboxylating enzymes. The oxygen sensitivity of POR and ferredoxin makes this cycle limited to anaerobic or microaerobic niches (24).

1.5 Direct conversion of CO_2 to biofuels

The direct conversion of CO_2 into energy-dense liquid fuels can not only mitigate climate change but also supply sustainable energy to meet its growing demand. In nature, autotrophic organisms can fix CO_2 with reducing power harnessed from either sunlight or various electron donors. With tools of metabolic engineering and synthetic biology, these autotrophs can be engineered to produce fuel compounds directly from CO_2 (Fig 1.9).

Phototrophic organisms such as eukaryotic microalgae and prokaryotic cyanobacteria are capable of using sunlight to fix CO_2 into a variety of organic molecules. These organisms can be metabolically engineered to produce fuel molecules directly, bypassing the recalcitrance problem of plant biomass. The capacity of many microalgae species to produce large amounts of lipids (up to 50–70% of lipid per dry weight (42)) as storage products makes them an attractive resource for biodiesel production. To make algae-derived biodiesel economically feasible, improving the efficiency of algae converting CO_2 to energy-dense oil is crucial. Bioprospecting for algae strains with high lipid accumulation, superior growth and harvesting characteristics can provide suitable candidates (43). Next-generation sequencing (44), RNA interference (RNAi) (45), and system biology approaches have provided insights into algal physiological and metabolic processes (46). For example, high throughput RNA-sequencing technologies have helped identify responsive genes for triacylglycerol accumulation under nitrogen deprivation conditions (44), which can provide potential targets for engineering high lipid production strains. Also, RNA interference (RNAi) technology (45), which enables sequence-specific knockdown genes in algal species with the RNAi machinery, is becoming a promising genetic tool for increasing lipid production in algae.

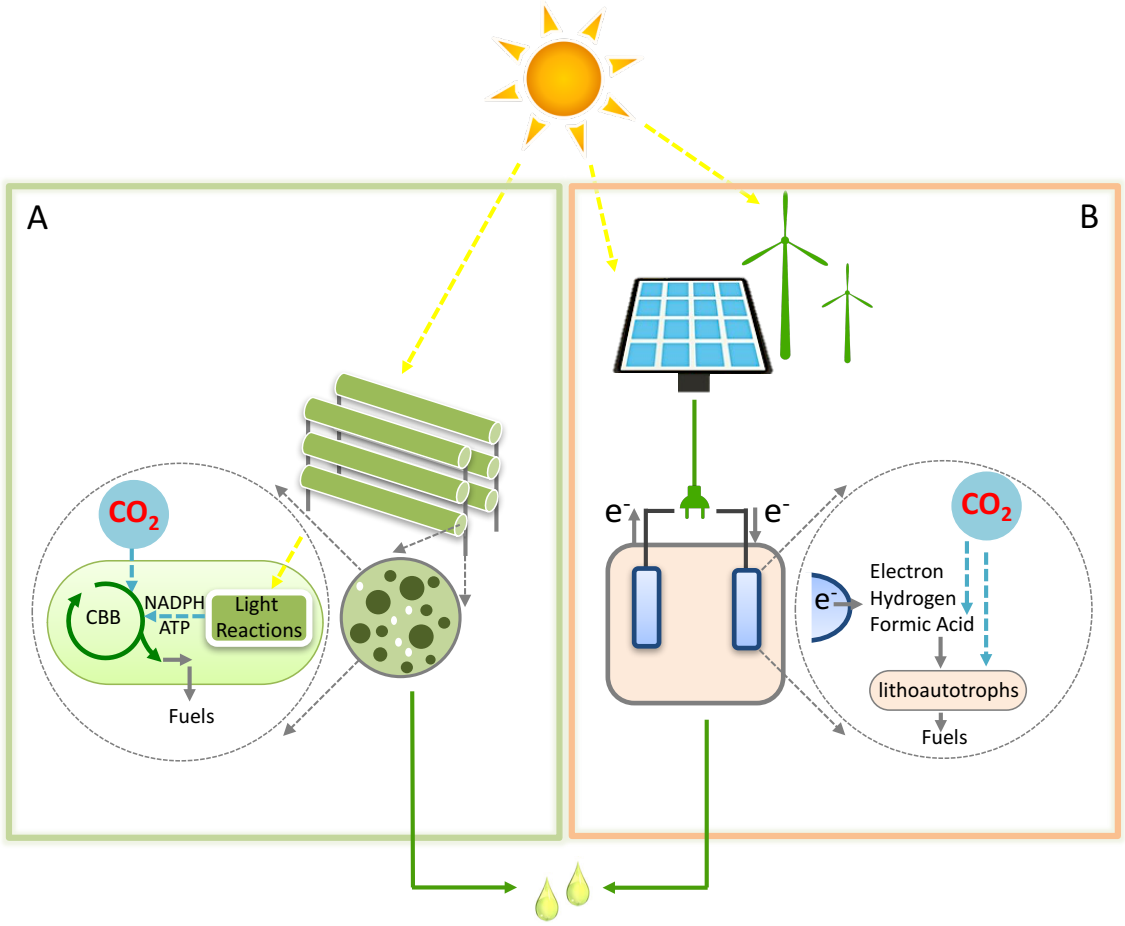


Fig. 1.9. Overview of direct biofuel production from CO_2.

(A) Phototrophic organisms such as microalgae and cyanobacteria can harness energy from sunlight to reduce CO_2 and convert it to liquid fuels. (B) A broad range of lithoautotrophs can fix CO_2 to produce fuels with reducing power from electrons or electrochemically generated electron shuttles, such as H_2 and formic acid.

Cyanobacteria, which are genetically more tractable than microalgae, are more versatile for fuels and chemicals production. Introducing exogenous biosynthetic pathways into *Synechococcus elongatus* PCC 7942 or *Synechocystis* sp. PCC 6803 has led to the production of ethanol (47), isopropanol (48), isobutyraldehyde (49), isobutanol (49), 1-butanol (50, 51), 2-methyl-1-butanol

(52), isoprene (53), ethylene (54), 2,3-butanediol (55), 1,2-propanediol (56) and fatty acids (57) directly from CO_2. Many cyanobacteria strains are capable to produce alkanes naturally (58), and overexpression of alkane biosynthetic genes in cyanobacteria has enhanced the alkane production (59). Besides overexpression of relevant biosynthetic pathways in cyanobacteria, many other strategies have been applied to produce desired compounds more efficiently, such as introducing thermodynamic driving forces (51), aligning cofactor usage (56) and using oxygen-tolerant enzymes (60) in synthetic pathways. Thermodynamic driving forces can be increased by either increasing metabolite pool size through eliminating its draining pathways (61), or by re-routing the path with an ATP expenditure (51). Since cyanobacteria generate NADPH readily through photosynthesis, switching cofactor usage from NADH to NADPH (56) in product formation pathway is helpful. Furthermore, since photosynthesis generates oxygen, using oxygen-tolerant enzymes (60) is fuel production is essential.

Providing large light exposing surfaces is crucial for producing photosynthesis-derived biofuels. To achieve this, open-pond and closed photobioreactors are used to culture photosynthetic microorganisms in large-scale. However, open-pond cultivation systems need large areas of land, and have problems like water loss, contamination and being difficult to control the culture conditions (62, 63). On the other hand, closed photobioreactors is costly for biofuels production (62). In addition, the light capture efficiency is constrained by the nature of photosynthetic apparatus, which can only utilize radiation within a limited spectrum and intensity range, thus making photosynthetic microorganisms access only ~50% of the incident solar energy and be photoinhibited above a certain value of light intensity (64). On the other hand, sensitivity to high light, high temperature and high salt also presents a problem. *Synechococcus* sp. PCC 7002 displays favorable properties in these areas (65). In general, photosynthesis-derived biofuels still

need to compensate light-based constraints that are not encountered in traditional microbial fermentation processes to be economically viable.

Another group of autotrophic organisms, lithoautotrophs, can fix CO_2 with reducing power from electron donors that can be generated from electricity. Using electricity and lithoautotrophs to convert CO_2 into energy-dense liquid fuels, which are referred to as "electrofuels", has many advantages. First, owing to the higher efficiency of photovoltaic technologies than biological photosynthesis (66), electrofuels in theory could have better sun-to-fuel efficiency. Second, separating light and dark reactions circumvents the expensive cultivation systems for photosynthetic microorganisms. Moreover, since the solar or wind electricity is intermittent, converting electricity to liquid biofuel is the most energy-sense approach for energy storage.

Some autotrophic microorganisms, such as *Geobacter* spp., can directly accept electrons from cathodes for anaerobic respiration (67). Recently, a wide range of acetogenic microorganisms including *Sporomusa* spp., *Clostridium* spp., and *Moorella thermoacetica* have shown the capability of accepting electrons from a cathode and fixing CO_2 into acetate and a small amount of 2-oxobutyrate (68, 69), indicating the potential of biofuel production from CO_2 with direct electron feeding. However, bioreactors with large electrode surfaces are needed to establish this process, which will inevitably increase the cost.

Alternatively, a broad range of lithoautotrophic microorganisms are capable to fix CO_2 with reducing power from electron shuttles, such as hydrogen and formic acid. H_2 can be produced on cathodes from water splitting, whereas formic acid can be synthesized electrochemically using water and CO_2. Lithoautotrophic organisms, such as *Cupriavidus necator* (formerly known as *Ralstonia eutropha*), have been engineered to produce biofuels from CO_2 using H_2 (70, 71) or

formic acid (72). Similarly, the H_2-utilizing heterotroph *Pyrococcus furiosus* has been engineered to incorporate CO_2 using a portion of the 3-hydroxypropionate/4-hydroxybutyrate CO_2 fixation pathway from *Metallosphaera sedula* to produce 3-hydroxypropionic acid (73). However, the low solubility of H_2 and low mass transfer rate in these processes may increase the cost of biofuel production.

1.6 Reference

1. IPCC (Intergovernmental Panel on Climate Change) (2013) Climate change 2013: The physical science basis. Working Group I contribution to the IPCC Fifth Assessment Report. *Cambridge, United Kingdom Cambridge Univ Press*. Available at: http://www.ipcc.ch/report/ar5/wg1/.

2. US EPA (United States Environmental Protection Agency) Climate Change Indicators: Global Greenhouse Gas Emissions. Available at: https://www.epa.gov/climate-indicators/climate-change-indicators-global-greenhouse-gas-emissions.

3. ESRL (Earth System Research Laboratory) ESRL Global Monitoring Division - Global Greenhouse Gas Reference Network. Available at: https://www.esrl.noaa.gov/gmd/ccgg/trends/.

4. Hansen J, et al. (2008) Target atmospheric CO2: Where should humanity aim? doi:10.2174/1874282300802010217.

5. Leung DYC, Caramanna G, Maroto-Valer MM (2014) An overview of current status of carbon dioxide capture and storage technologies. *Renew Sustain Energy Rev* 39:426–443.

6. Rahman FA, et al. (2017) Pollution to solution: Capture and sequestration of carbon dioxide (CO2) and its utilization as a renewable energy source for a sustainable future. *Renew Sustain Energy Rev* 71(January):112–126.

7. Metz, B., Davidson, O., De Coninck, H., Loos, M., & Meyer L (2005) *Carbon Dioxide Capture and Storage* doi:10.1002/anie.201000431.

8. Aresta M, Dibenedetto A, Angelini A (2014) Converting "Exhaust" Carbon into "Working" Carbon. *Adv Inorg Chem* 66:259–288.

9. Sakakura T, Choi JC, Yasuda H (2007) Transformation of carbon dioxide. *Chem Rev* 107(6):2365–2387.

10. Kondratenko E V., Mul G, Baltrusaitis J, Larrazábal GO, Pérez-Ramírez J (2013) Status and perspectives of CO2 conversion into fuels and chemicals by catalytic, photocatalytic and electrocatalytic processes. *Energy Environ Sci* 6(11):3112.

11. Lindsey AS, Jeskey H (1957) The Kolbe-Schmitt Reaction. *Chem Rev* 57(4):583–620.

12. Bosch C, Meiser W (1922) Process of manufacturing urea. *United States Pat Off* 209:1–2.

13. Graciani J, et al. (2014) Highly active copper-ceria and copper-ceria-titania catalysts for methanol synthesis from CO 2. *Science* 345(March):546–550.

14. Ma L, Jia I, Guo X, Xiang L (2014) Current Status and perspective of rare earth catalytic materials and catalysis. *Chinese J Catal* 35(2):108–119.

15. Bassham JA, et al. (1954) The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor [1]. *J Am Chem Soc* 76(7):1760–1770.

16. Evans MC, Buchanan BB, Arnon DI (1966) A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. *Proc Natl Acad Sci U S A* 55(4):928–34.

17. Ljungdahl LG, Wood HG (1969) Total Synthesis of Acetate From CO2 by Heterotrophic Bacteria. *Annu Rev Microbiol* 23(1):515–538.

18. Ljungdhal LG (1986) The Autotrophic Pathway of Acetate Synthesis in Acetogenic Bacteria. *Annu Rev Microbiol* 40(1):415–450.

19. Holo H (1989) Chloroflexus aurantiacus secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO2 and acetate. *Arch Microbiol* 151(3):252–256.

20. STRAUSS G, FUCHS G (1993) Enzymes of a novel autotrophic CO2 fixation pathway in the phototrophic bacterium Chloroflexus aurantiacus, the 3-hydroxypropionate cycle. *Eur J Biochem* 215(3):633–643.

21. Zarzycki J, Brecht V, Müller M, Fuchs G (2009) Identifying the missing steps of the autotrophic 3-hydroxypropionate CO2 fixation cycle in Chloroflexus aurantiacus. *Proc Natl Acad Sci U S A* 106(50):21317–22.

22. Berg IA, Kockelkorn D, Buckel W, Fuchs G (2007) A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea. *Science* 318(5857):1782–1786.

23. Huber H, et al. (2008) A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum Ignicoccus hospitalis. *Proc Natl Acad Sci U S A* 105(22):7851–6.

24. Berg I a, et al. (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* 8(6):447–60.

25. Tabita FR, et al. (2007) Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol Mol Biol Rev* 71(4):576–99.

26. Ellis RJ (1979) The most abundant protein in the world. *Trends Biochem Sci* 4(11):241–244.

27. Savir Y, Noor E, Milo R, Tlusty T (2010) Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proc Natl Acad Sci U S A* 107(8):3475–3480.

28. Tcherkez GGB, Farquhar GD, Andrews TJ (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc Natl Acad Sci U S A* 103(19):7246–51.

29. Miura A, Kameya M, Arai H, Ishii M, Igarashi Y (2008) A soluble NADH-dependent fumarate reductase in the reductive tricarboxylic acid cycle of Hydrogenobacter thermophilus TK-6. *J Bacteriol* 190(21):7170–7.

30. Aoshima M, Ishii M, Igarashi Y (2004) A novel biotin protein required for reductive carboxylation of 2-oxoglutarate by isocitrate dehydrogenase in Hydrogenobacter thermophilus TK-6. *Mol Microbiol* 51(3):791–8.

31. Yamamoto M, Arai H, Ishii M, Igarashi Y (2006) Role of two 2-oxoglutarate:ferredoxin oxidoreductases in *Hydrogenobacter thermophilus* under aerobic and anaerobic conditions. *FEMS Microbiol Lett* 263(2):189–193.

32. Yamamoto M, Ikeda T, Arai H, Ishii M, Igarashi Y (2010) Carboxylation reaction catalyzed by 2-oxoglutarate:ferredoxin oxidoreductases from Hydrogenobacter thermophilus. *Extremophiles* 14(1):79–85.

33. Hügler M, Sievert SM (2011) Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Ann Rev Mar Sci* 3:261–89.

34. Mattozzi M d., Ziesack M, Voges MJ, Silver PA, Way JC (2013) Expression of the sub-pathways of the Chloroflexus aurantiacus 3-hydroxypropionate carbon fixation bicycle in E. coli: Toward horizontal transfer of autotrophic growth. *Metab Eng* 16:130–139.

35. Berg IA, Ramos-Vera WH, Petri A, Huber H, Fuchs G (2010) Study of the distribution of autotrophic CO2 fixation cycles in Crenarchaeota. *Microbiology* 156(1):256–269.

36. Konneke M, et al. (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc Natl Acad Sci* 111(22):8239–8244.

37. Fuchs G (2011) *Alternative Pathways of Carbon Dioxide Fixation: Insights into the Early Evolution of Life?* doi:10.1146/annurev-micro-090110-102801.

38. Martins BM, Dobbek H, Cinkaya I, Buckel W, Messerschmidt A (2004) Crystal structure of 4-hydroxybutyryl-CoA dehydratase: radical catalysis involving a [4Fe-4S] cluster and flavin. *Proc Natl Acad Sci U S A* 101(44):15645–9.

39. Buckel W, Golding BT (2006) Radical Enzymes in Anaerobes. *Annu Rev Microbiol* 60(1):27–49.

40. Hawkins AB, Adams MWW, Kelly RM (2014) Conversion of 4-hydroxybutyrate to acetyl coenzyme A and its anapleurosis in the Metallosphaera sedula 3-hydroxypropionate/4-hydroxybutyrate carbon fixation pathway. *Appl Environ Microbiol* 80(8):2536–45.

41. Ramos-Vera WH, Berg IA, Fuchs G (2009) Autotrophic carbon dioxide assimilation in Thermoproteales revisited. *J Bacteriol* 191(13):4286–97.

42. Chisti Y (2007) Biodiesel from microalgae. *Biotechnol Adv* 25(3):294–306.

43. Rodolfi L, et al. (2009) Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnol Bioeng* 102(1):100–112.

44. Boyle NR, et al. (2012) Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in Chlamydomonas. *J Biol Chem* 287(19):15811–15825.

45. Cerutti H, Ma X, Msanne J, Repas T (2011) RNA-mediated silencing in Algae: Biological roles and tools for analysis of gene function. *Eukaryot Cell* 10(9):1164–1172.

46. Gimpel J a., Specht E a., Georgianna DR, Mayfield SP (2013) Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr Opin Chem Biol* 17(3):489–495.

47. Gao Z, Zhao H, Li Z, Tan X, Lu X (2012) Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy Environ Sci* 5(12):9857–9865.

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

49. Atsumi S, Higashide W, Liao JC (2009) Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotechnol* 27(12):1177–80.

50. Lan EI, Liao JC (2011) Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab Eng* 13(4):353–63.

51. Lan EI, Liao JC (2012) ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc Natl Acad Sci* 109(16):6018–6023.

52. Shen CR, Liao JC (2012) Photosynthetic production of 2-methyl-1-butanol from CO2 in cyanobacterium Synechococcus elongatus PCC7942 and characterization of the native acetohydroxyacid synthase. *Energy Environ Sci* 5(11):9574.

53. Lindberg P, Park S, Melis A (2010) Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. *Metab Eng* 12(1):70–79.

54. Ungerer J, et al. (2012) Sustained photosynthetic conversion of CO2 to ethylene in recombinant cyanobacterium Synechocystis 6803. *Energy Environ Sci* 5(10):8998.

55. Oliver JWK, Machado IMP, Yoneda H, Atsumi S (2013) Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc Natl Acad Sci U S A* 110(4):1249–54.

56. Li H, Liao JC (2013) Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO2 to 1,2-propanediol. *Microb Cell Fact* 12(1):4.

57. Liu X, Sheng J, Curtiss R (2011) Fatty acid production in genetically modified cyanobacteria. *Proc Natl Acad Sci U S A* 108(17):6899–6904.

58. Schirmer A, Rude mathew A, Li X, Popova E, Del Cardayre SB (2010) Microbial Biosynthesis of Alkanes. *Science* 329(July):559–562.

59. Wang W, Liu X, Lu X (2013) Engineering cyanobacteria to improve photosynthetic production of alka(e)nes. *Biotechnol Biofuels* 6(1):69.

60. Lan EI, Ro SY, Liao JC (2013) Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy Environ Sci* 6(9):2672.

61. Shen CR, et al. (2011) Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Appl Environ Microbiol* 77(9):2905–2915.

62. Chen CY, Yeh KL, Aisyah R, Lee DJ, Chang JS (2011) Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: A critical review. *Bioresour Technol* 102(1):71–81.

63. Hannon M, Gimpel J, Tran M, Rasala B, Mayfield S (2010) Biofuels from algae: challenges and potential. *Biofuels* 1(5):763–784.

64. Zhu XG, Long SP, Ort DR (2008) What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr Opin Biotechnol* 19(2):153–159.

65. Ludwig M, Bryant D a. (2012) Synechococcus sp. Strain PCC 7002 transcriptome: Acclimation to temperature, salinity, oxidative stress, and mixotrophic growth conditions. *Front Microbiol* 3(OCT):1–14.

66. Blankenship RE, et al. (2011) Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* 332(6031):805–809.

67. Gregory KB, Bond DR, Lovley DR (2004) Graphite electrodes as electron donors for anaerobic respiration. *Environ Microbiol* 6(6):596–604.

68. Nevin KP, Woodard TL, Franks AE, Summers ZM, Lovley DR (2010) Microbial electrosynthesis: feeding microbes electricity to convert carbon dioxide and water to multicarbon extracellular organic compounds. *MBio* 1(2):e00103-10-.

69. Nevin KP, et al. (2011) Electrosynthesis of organic compounds from carbon dioxide is catalyzed by a diversity of acetogenic microorganisms. *Appl Environ Microbiol* 77(9):2882–2886.

70. Torella JP, et al. (2015) Efficient solar-to-fuels production from a hybrid microbial-water-splitting catalyst system. *Proc Natl Acad Sci U S A* 112(8):2337–42.

71. Liu C, Colón BC, Ziesack M, Silver PA, Nocera DG (2016) Water splitting–biosynthetic system with CO2 reduction efficiencies exceeding photosynthesis. *Science* 352(6290): 1210–1213.

72. Li H, et al. (2012) Integrated Electromicrobial Conversion of CO2 to Higher Alcohols. *Science* 335(6076):1596–1596.

73. Keller MW, et al. (2013) Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proc Natl Acad Sci U S A* 110(15):5840–5.

Chapter 2. Design and demonstration of a synthetic CO_2 fixation cycle *in vitro*

2.1 Introduction

The increasing atmospheric carbon dioxide (CO_2) level, if unabated, is threatening to cause disastrous global climate changes and sea level rise (1, 2). Currently, the annual anthropogenic CO_2 emission is approaching 40 $GtCO_2$/year, with cumulative CO_2 emission exceeding 2000 Gt (1). Results of multiple analyses show that cumulative CO_2 emission correlates almost linearly with global temperature rise (1). Thus, to keep the global temperature increase below the acceptable 2 °C limit, negative emission is required to remove legacy CO_2 in the atmosphere (3). Until now, biology is the only demonstrated approach for removing CO_2 from air at scale powered by sunlight. After initial carbon fixation, the fixed carbon can be converted to fuels, chemicals, or storage materials, either to offset the need for fossil fuel or as long-term carbon storage (4).

Nature has evolved several CO_2 fixation pathways for synthesizing organic compounds from CO_2 under ambient conditions. To date, six CO_2 fixation pathways have been reported (5): the Calvin–Benson-Bassham (CBB) cycle, the reductive acetyl-CoA (Wood-Ljungdahl) pathway, the reductive tricarboxylic acid (rTCA) cycle, the 3-hydroxypropionate (3HP) bicycle, the dicarboxylate (DC)/4-hydroxybutyrate (4HB) cycle, and the 3HP/4HB cycle. However, naturally evolved CO_2 fixation pathways all contain key carboxylases with intrinsic problems hindering further engineering (Tables 2.S1 and 2.S2) (6). For example, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo), the key enzyme of CBB, has a relatively low velocity and has a significant oxygenase activity that leads to loss of fixed carbon (7). The negative correlation between its selectivity and velocity indicates that the enzyme might have already been naturally

optimized (8, 9), suggesting the difficulty of its further improvement. Pyruvate-ferredoxin oxidoreductase (PFOR), one of the key enzymes in the Woods-Ljungdahl pathway, the rTCA cycle and the DC/4HB cycle, is oxygen sensitive and requires ferredoxin as the electron carrier (10), making it difficult to use *in vitro* or in non-native organisms. Acetyl-coenzyme A (CoA) carboxylase (ACC), the key enzyme of the 3HP bicycle and the 3HP/4HB cycle is a biotin-dependent enzyme complex containing several different subunits (11), and is difficult to manipulate. Hence, to build a more efficient CO_2 conversion system, alternative CO_2 fixation pathways composed of efficient, oxygen tolerant, and structurally simple enzymes are preferred. Recently, a synthetic CO_2 fixation cycle, the crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle, has been demonstrated *in vitro* for glyoxylate synthesis from CO_2 (12). Despite this success, more synthetic cycles are needed to optimize the CO_2 fixation rate beyond what nature has accomplished.

2.2 Results

2.2.1 Design of the rGPS cycle

The goal of a carbon fixation cycle is to convert CO_2 to the C2 and C3 metabolites, acetyl-CoA and pyruvate, respectively, which serve as precursors for all cellular constituents. We had focused on the simple and oxygen-insensitive carboxylase, phosphoenolpyruvate (PEP) carboxylase (Ppc), which catalyzes the carboxylation of PEP (C3) to form the C4 compound oxaloacetate (OAA). It is one of the most active carboxylases with superior affinity toward HCO_3^- and has no oxygenase activity (13). Ppc is used to replenish intermediates of the tricarboxylic acid (TCA) cycle, or to shuttle CO_2 in C4 or crassulacean acid metabolism (CAM) plants. However, in most cells, OAA cannot be converted to pyruvate or acetyl-CoA, without

carbon loss. To overcome this limitation, we have designed and demonstrated a Ppc-based synthetic pathway to convert one C3 compound (pyruvate) to two C2 metabolites, acetyl-CoA and glyoxylate, with one bicarbonate fixed (14). In this pathway, the C4 compound OAA is reduced and converted to (S)-malyl-CoA with the help of ATP hydrolysis. We termed this pathway the reductive glyoxylate synthesis (rGS) pathway (part of the reverse glyoxylate shunt) (14). In a follow-up design called the malyl-CoA glycerate (MCG) pathway (15), glyoxylate is converted to acetyl-CoA via glycerate, with a net reaction of fixing one carbon to pyruvate to produce two acetyl-CoA (see Fig. 2.S1 for glyoxylate conversion pathways).

To complete a CO_2 fixation cycle, another carboxylation reaction is needed to convert the C2 product back to the C3 compound, pyruvate. In nature, PFOR can fix one CO_2 to acetyl-CoA to produce pyruvate using reduced ferredoxin as the redox cofactor. However, the oxygen sensitivity of PFOR is a major challenge that precludes its function in aerobic organisms. So far, no oxygen tolerant pathway of converting acetyl-CoA to pyruvate with one CO_2 fixed has been demonstrated. Although pathways based on ACC or methylmalonyl-CoA carboxytransferase (MCCT) have been proposed (13), both enzymes are complex and difficult to manipulate. Without converting acetyl-CoA to pyruvate, the designed pathway cannot support autotrophic growth.

To design an oxygen-tolerant pathway that is equivalent to PFOR, we employed crotonyl-CoA carboxylase/reductase (Ccr), which catalyzes the reductive carboxylation of crotonyl-CoA to (2S)-ethylmalonyl-CoA using NADPH as the reducing equivalent (16). Similar to Ppc, Ccr is also a simple enzyme with no oxygen-sensitivity. It catalyzes the carboxylation reaction with a superior activity (Table 2.S1), and is oxygen insensitive. Ccr was first discovered in the ethylmalonyl-CoA pathway for acetyl-CoA assimilation (16). In this pathway, two acetyl-CoA

molecules are condensed to form crotonyl-CoA, which is then carboxylated to the C5 intermediate ethylmalonyl-CoA. If any C5 intermediate of the ethylmalonyl-CoA pathway can turn into pyruvate and acetyl-CoA, then the net reaction of fixing one CO_2 to acetyl-CoA to form pyruvate can be achieved. To this end, we found that the C5 intermediate mesaconyl-C1-CoA could be converted to acetyl-CoA and pyruvate through 3 steps in the 3HP bicycle (17) (Fig. 2.1). Combing 7 reactions of the ethylmalonyl-CoA pathway with 3 steps of the 3HP bicycle leads to converting C2 compound acetyl-CoA back to C3 compound pyruvate with one CO_2 fixed. This pathway is termed the reductive pyruvate synthesis (rPS) pathway (Fig. 2.1A). All 10 enzymes are simple, efficient and oxygen-insensitive, thus avoiding all the challenges involved in PFOR, or ACC and MCCT-based strategies.

The rPS pathway converts one acetyl-CoA to pyruvate with one CO_2 fixed, while the rGS pathway converts one pyruvate to one glyoxylate and one acetyl-CoA fixing one bicarbonate. Together, we obtained a CO_2 fixation cycle with the following overall reaction: $CO_2+HCO_3^-+3ATP+NADPH+2NADH\rightarrow Glyoxylate+FADH_2$. We call this CO_2 fixation cycle the reductive glyoxylate/pyruvate synthesis (rGPS) cycle. The rGPS cycle is thermodynamically feasible as indicated by the Gibbs free energy changes (Fig. 2.S2). The rGPS cycle can be used in conjunction with the MCG pathway to produce either acetyl-CoA or pyruvate (Fig. 2.1BC and Fig. 2.S3). Thus, rGPS is fully capable of supporting cell growth in place of any naturally-evolved CO_2 fixing pathway.

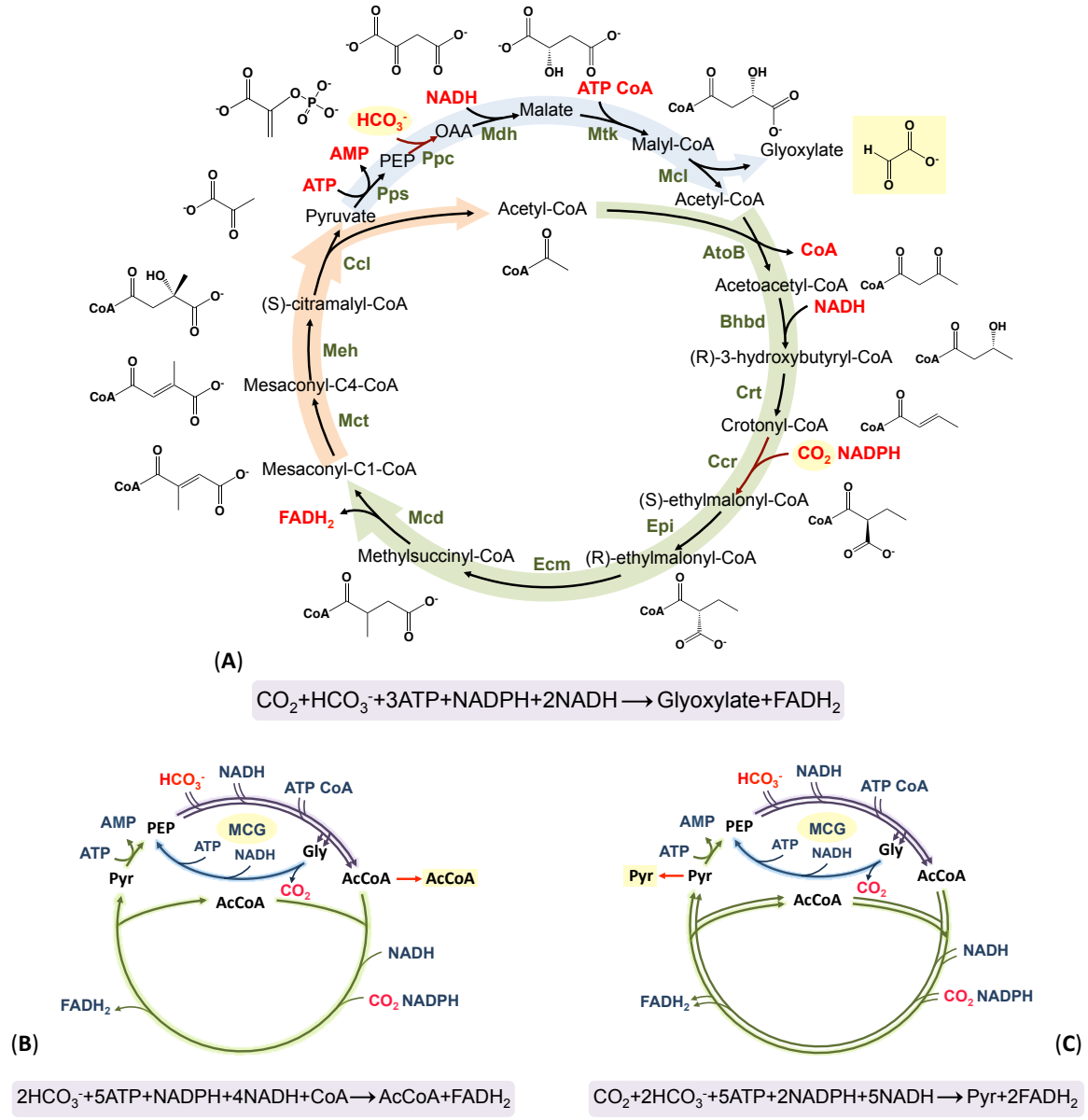


Fig. 2.1. The rGPS cycle.

(A) Topology of the rGPS cycle, which consists of 7 steps from the ethylmalonyl-CoA pathway (green), 3 steps of the 3-hydroxypropionate bicycle (orange), and the reductive glyoxylate synthesis pathway (blue). The red arrow indicates the carboxylation reaction. The sum of reactions in the rGPS cycle is shown in the purple box. See Table 2.S3 for enzyme abbreviations. **(B, C)** Simplified schematic of combing rGPS with MCG to produce acetyl-CoA (AcCoA) (B), or pyruvate (Pyr) (C). Green arrows indicate reactions only in the rGPS cycle, blue arrows indicate reactions only in the MCG pathway, and purple arrows

indicate reactions in both pathways. PEP, phosphoenolpyruvate; OAA, oxaloacetate; Gly, glyoxylate.

2.2.2 Demonstration of one round of rGPS *in vitro*

Direct construction of rGPS *in vivo* is challenging due to the complicated cellular regulation that many leads to unforeseeable problems (18). Additionally, the enzymes, cofactors, and metabolites in the synthetic cycle may be incompatible with each other. To separate the biological problems from the biochemical problems, we choose to demonstrate the rGPS cycle using a cell-free approach (19–21). The well-defined condition of cell-free systems can facilitate detecting problems in the synthetic pathway, including protein and metabolite stability, and enzyme inhibition by metabolites (22, 23). Most importantly, some pathway enzymes and cofactor regeneration systems need to be carefully balanced to avoid depletion or accumulation of certain metabolites that lead to blockage of the pathway (24–26). For naturally evolved pathways, these problems are solved through evolution. For synthetic pathways (12, 24, 27), it is desirable to solve these problems using cell free systems.

To build the cell-free rGPS system, we selected, cloned, expressed in *E. coli*, and his-tag purified 13 enzymes in the rGPS cycle (Fig. 2.S4). The remaining two, Ppc and malate dehydrogenase (Mdh), were acquired. Each enzyme was tested for its activity with enzyme assays (Fig. 2.S5 and Table 2.S3). First, the enzymes were sequentially added to the reaction buffer and the corresponding products were detected from crotonyl-CoA to pyruvate and acetyl-CoA (Fig. 2.S6). The three steps from acetyl-CoA to crotonyl-CoA have been demonstrated in the butanol synthesis (27). Then the rGPS cycle was divided in two pots based on cofactor requirement (Fig. 2.2AB), and each pot was able to accomplish its desired activity. We then demonstrated the

complete cycle in three steps but in one-pot (Fig. 2.2CD). Acetyl-CoA was first converted to (S)-ethylmalonyl-CoA by acetyl-CoA acetyltransferase (AtoB), 3-hydroxybutyryl-CoA dehydrogenase (Bhbd), crotonase (Crt), and Ccr with appropriate co-factors and bicarbonate (Fig. 2.2D). Then, ethylmalonyl-CoA epimerase (Epi), (2R)-ethylmalonyl-CoA mutase (Ecm), and (2S)-methylsuccinyl-CoA dehydrogenase (Mcd) were added to the reaction mixture to convert (S)-ethylmalonyl-CoA to mesaconyl-C1-CoA by using ferrocenium hexafluorophosphate ($FePF_6$) to regenerate flavin adenine dinucleotide (FAD) (Fig. 2.2D). In the last step, the remaining enzymes mesaconyl-C1-CoA-C4-CoA transferase (Mct), mesaconyl-C4-CoA hydratase (Meh), (*S*)-citramalyl-CoA lyase (Ccl), phosphoenolpyruvate synthetase (Pps), Ppc, Mdh, malate thiokinase (Mtk), and malyl coenzyme A lyase (Mcl) were added together with NADH, ATP, and CoA to produce glyoxylate while regenerating acetyl-CoA to complete the cycle (Fig. 2.2D). Starting from acetyl-CoA, after this stepwise reconstitution, the final product glyoxylate was detected, indicating one round of rGPS was completed *in vitro* (Fig. 2.2D).

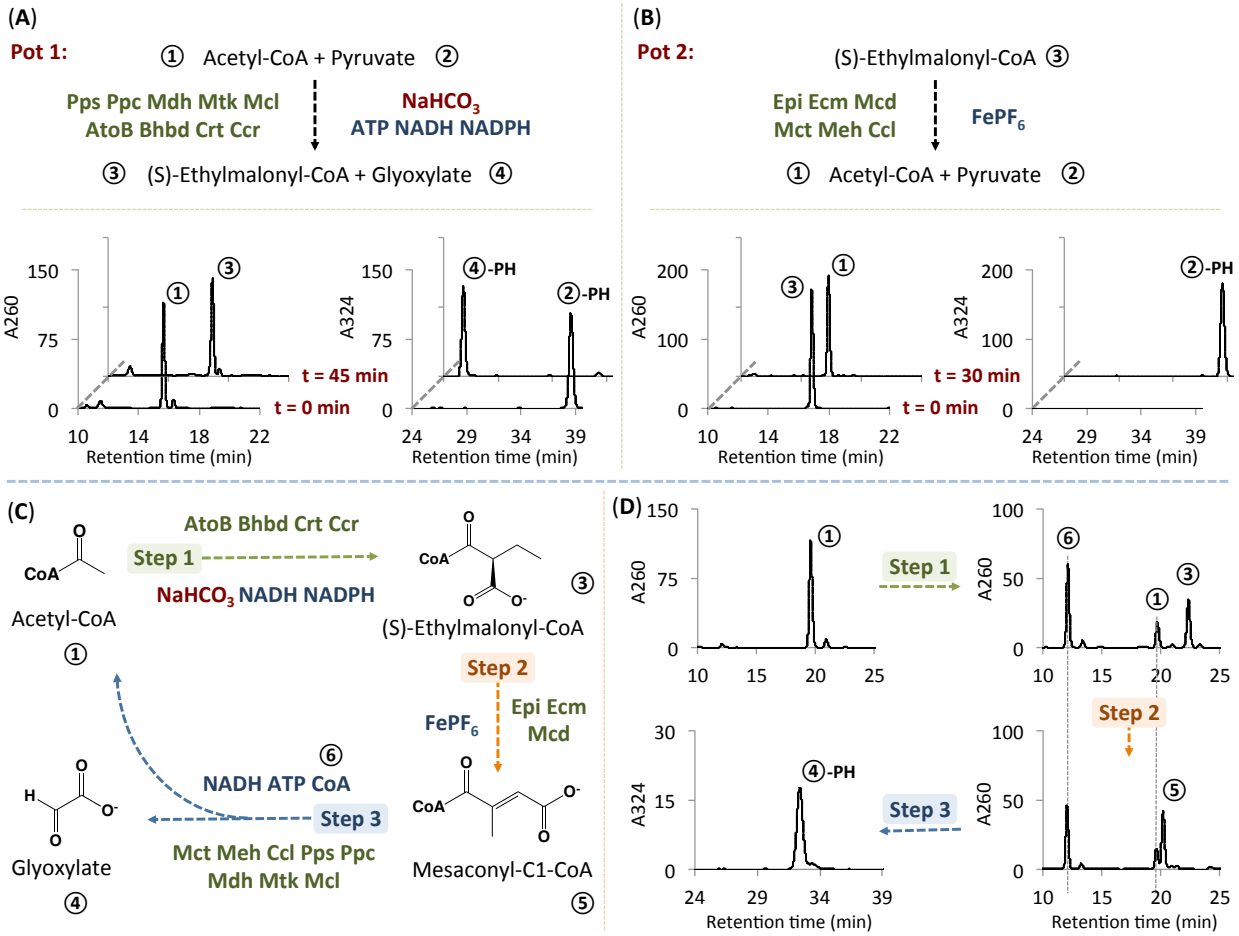


Fig. 2.2. Demonstration of one round of rGPS.

(A, B) Demonstration of the rGPS cycle in two pots. Each pot contained corresponding enzymes (green) and cofactors (blue) together in the reaction buffer containing the corresponding starting substances. (A) The reaction mixture of pot 1 was analyzed with HPLC at time 0 min and time 45 min. (B) The reaction mixture of pot 2 was analyzed with HPLC at time 0 min and time 30 min. **(C)** Demonstration of the rGPS cycle in one pot but three steps through sequentially adding enzymes (green) and cofactors (blue) into the reaction buffer, which contained acetyl-CoA and $NaHCO_3$. **(D)** HPLC analysis of the reaction mixture from step 0 (before adding enzymes) to step 3 (after all enzymes were added). Steps are numbered and colored according to panel (C). Intermediates are numbered according to panel (A), (B), and (C). Free CoA and CoA esters were followed at 260 nm, glyoxylate and pyruvate were derivatized by phenylhydrazine and detected at 324 nm. Note that reaction mixtures in (A), (B) were analyzed using HPLC system 2 method and the

reaction mixture in (D) was analyzed using HPLC system 1 method, resulting in the difference in retention times. See Table 2.S3 for enzyme abbreviations. $FePF_6$, ferrocenium hexafluorophosphate; PH, phenylhydrazone.

2.2.3 Problems of the continuous running of rGPS *in vitro*

To realize continuous operation, high concentrations of cofactors are required in the reaction mixture to drive rGPS. However, we found that the artificial electron acceptor $FePF_6$, which is used for regenerating flavin adenine dinucleotide (FAD) of enzyme Mcd, can also accept electrons from NADH and NADPH. The rate of this spontaneous reaction mainly depends on the concentrations of $FePF_6$, NADH and NADPH in this *in vitro* system. Thus, adding high concentrations of these cofactors at the beginning can cause significant wasteful consumptions. To avoid this problem, we tried to add NADPH and NADH first into the reaction mixture containing all enzymes to convert acetyl-CoA to methylsuccinyl-CoA and then add $FePF_6$ to continue the cycle. However, this strategy did not work since methylsuccinyl-CoA was not stable in our *in vitro* system (Fig. 2.S7). Without supplying $FePF_6$, the cumulated methylsuccinyl-CoA was degraded and the cycle stopped.

2.2.4 Semi-continuous operation of rGPS using a two-pot immobilized enzyme system

To avoid the spontaneous reaction between $FePF_6$ and NAD(P)H as well as the accumulation of the unstable intermediate methylsuccinyl-CoA, we designed a two-pot immobilized enzyme system to run the rGPS cycle semi-continuously (Fig. 2.3). This system divided rGPS into two pots that enabled adding $FePF_6$ and NAD(P)H separately while avoiding the accumulation of methylsuccinyl-CoA. The first pot included Pps, Ppc, Mdh, Mtk, Mcl, AtoB, Bhbd, Crt and Ccr,

converting pyruvate and acetyl-CoA into (S)-ethylmalonyl-CoA and glyoxylate using ATP, NADH and NADPH as cofactors (Fig. 2.3A). The second pot included Epi, Ecm, Mcd, Mct, Meh and Ccl, converting (S)-ethylmalonyl-CoA into acetyl-CoA and pyruvate, using the artificial electron acceptor $FePF_6$ as the cofactor. All enzymes are his-tagged and immobilized with nickel resin beads.

The cycle started from crotonyl-CoA. First, the reaction buffer containing crotonyl-CoA, $NaH^{13}CO3$ and NADPH were added into the first immobilized enzyme pot (Fig. 2.3A). After converting crotonyl-CoA to (S)-ethymalonyl-CoA, the reaction buffer was separated from the immobilized enzymes and added into the second immobilized enzyme pot with $FePF_6$ to convert (S)-ethymalonyl-CoA into pyruvate and acetyl-CoA. After the conversion, the reaction buffer was separated and added into the first pot with ATP, NADH and NADPH to convert pyruvate and acetyl-CoA into (S)-ethylmalonyl-CoA and glyoxylate. Thus, one rGPS cycle was completed (Fig. 2.3B). Then, the reaction buffer was separated and added into the second pot with $FePF_6$ to start the second run of rGPS. By semi-continuously running rGPS for two cycles in the immobilized system, double-labeled glyoxlate was detected (Fig. 2.3C) (see Fig. 2.S8 for ^{13}C-labeling patterns). This success indicated that avoiding the accumulation of methysuccinyl-CoA by keeping a balanced cofactor supply could lead to continuous function of the rGPS cycle.

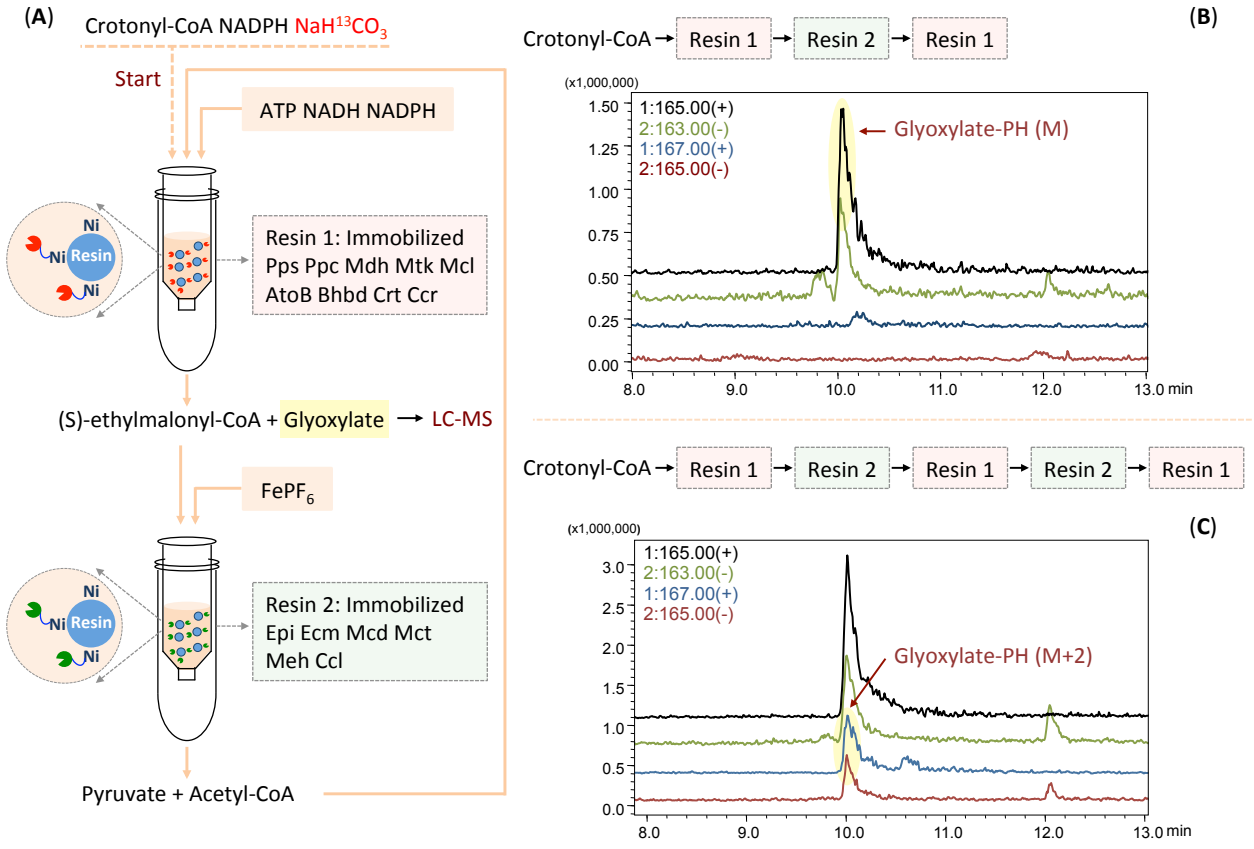


Fig. 2.3. The semi-continuous running of the rGPS cycle in vitro using a two-pot immobilized enzyme system.

(A) Scheme of the immobilized enzyme system, which consists of two pots of corresponding enzymes (red or green) immobilized with nickel resin beads. The first pot converts pyruvate and acetyl-CoA into (S)-ethylmalonyl-CoA and glyoxylate using ATP, NADH and NADPH as cofactors; the second pot converts (S)-ethylmalonyl-CoA into acetyl-CoA and pyruvate, using the artificial electron $FePF_6$ as the cofactor. Reaction mixture containing crotonyl-CoA, $NaH^{13}CO_3$ and NADPH was added to resin 1 to react for 5 min, then was separated and added to resin 2 together with $FePF_6$ for 30 min. Then the reaction mixture was separated and added back to resin 1 again. One cycle can be completed after going through resin 1, resin 2 and resin 1. Two cycles can be completed after going through resin 1 and resin 2 for one more time. **(B)** LC-MS analysis of the reaction mixture after running one rGPS cycle. **(C)** LC-MS analysis of the reaction mixture after running two rGPS cycles. See Fig. 2.S8 for expected labeling patterns. For glyoxylate

phenylhydrazone, $[M+H]^+$ = 165 and $[M-H]^-$ = 163. For double-labeled glyoxylate phenylhydrazone, $[M+H]^+$ = 167 and $[M-H]^-$ = 165.

2.3 Discussion

Nature has evolved thousands of enzymes that serve as a repertoire for designing pathways for a specific purpose. For CO_2 fixation, six natural pathways have been reported, suggesting that multiple pathways are required to meet the specific needs for different organisms (28). More CO_2 fixing pathways may be designed based on the repertoire of the existing enzymes or with modifications (12, 13). Besides thermodynamic and kinetic feasibility, oxygen tolerance and structural simplicity should also be considered as criteria for choosing enzymes for building synthetic CO_2 fixation pathways. The rGPS cycle was designed based on those criteria. In particular, the rPS segment in rGPS is an oxygen-insensitive pathway for fixing one carbon to acetyl-CoA to generate pyruvate, demonstrating the first oxygen-insensitive PFOR-equivalent pathway.

In this chapter, we successfully demonstrated a whole cycle of rGPS *in vitro*. We also identified the main problems hindering the continuous operation of the rGPS cycle. The two-pot immobilized enzyme system enabled its semi-continuous running. However, this semi-continuous operation inevitably compromised the efficiency of the whole cycle. How to operate the oxidation and reduction reactions in rGPS cycle simultaneously to avoid the accumulation of the unstable intermediate methylsuccinyl-CoA is the key point for getting an efficient running rGPS *in vitro*.

2.4 Materials and methods

2.4.1 Materials

All chemicals used were purchased from Sigma-Aldrich unless otherwise specified. Oligonucleotides were purchased from IDT. KOD and KOD Xtreme DNA polymerases were from EMD Millipore. Gibson Assembly Master Mix was purchased from New England Biolabs.

2.4.2 Culture medium and conditions

E. coli strains were grown in Luria-Bertani (LB) medium with appropriate antibiotics at 37 °C (for general molecular biology purposes) or 30 °C (for protein expression). Antibiotics were used at the following concentrations: Carbenicillin, 200 μg/mL; Spectinomycin, 50 μg/mL; and Kanamycin, 50 μg/mL.

2.4.3 Plasmid Construction

All plasmids constructed in this study were assembled using the Gibson isothermal DNA assembly method. pQE9 (Qiagen) or pCDF-Duet1 (Novagen) was used as the vector backbone. Plasmids are list in Table 2.S4. Primers used are listed in Table 2.S5. All plasmids were constructed in *E. coli* strain XL-1 blue for propagation and storage.

2.4.4 Protein purification

E. coli strain XL-1 blue or BL21(DE3) was used as host for protein expression. Induction was accomplished with 0.5 mM IPTG. His-tagged AtoB, Bhbd, Crt, Ccr, Epi, Mct, Meh, Ccl, Ppc, Mdh enzymes were purified by immobilized metal (Ni^{2+}) affinity chromatography using His-Spin Protein Miniprep kit (Zymo Research) with small modifications that the phosphate-based

his-elution buffer was changed to a glycylglycine-based buffer to avoid the inhibition effect of phosphate buffer on Mtk (Fig. 2.S9) and the washing was repeated four times instead of two times. His-tagged Ecm, Mcd, Pps, Mtk, Mcl enzymes were purified by immobilized metal (C_O^{2+}) affinity chromatography using HisPur Cobalt Resin (Thermo Fisher Scientific) with a small modification that the elution buffer was changed to a glycylglycine-based buffer. Enzyme concentrations were determined by the Bradford protein assay using Pierce Coomassie Plus Assay Kit (Thermo Fisher Scientific).

2.4.5 Enzyme Assays

All assays were conducted at room temperature under aerobic conditions. The reaction mixture volume was 0.6 ml unless otherwise indicated. The reaction schemes of each individual enzyme assay are illustrated in Fig. 2.S5.

AtoB Assay. The AtoB activity was measured by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADH ($\varepsilon_{340\ nm}$ = 5.31 mM^{-1} cm^{-1}). The reaction mixture contained 100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.1 mM NADH, 1 mM acetyl-CoA, 15 μg Bhbd, and 0.15 μg AtoB.

Bhbd Assay. The Bhbd activity was measured by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADH. The reaction mixture contained 100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.2 mM NADH, 0.3 mM acetoacetyl-CoA, and 0.19 μg Bhbd.

Crt assay. The Crt activity was measured by monitoring the decrease of absorbance at 290 nm corresponding to the consumption of crotonyl-CoA. The assay mixture contained 100 mM Tris-HCl pH 7.5, 0.2 mM crotonyl-CoA, and 0.008 μg Crt.

Ccr assay. The Ccr activity was measured by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADPH ($\varepsilon_{340\ nm}$ = 5.31 mM^{-1} cm^{-1}). The reaction mixture contained 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 0.2 mM NADPH, 0.2 mM crotonyl-CoA, and 0.5 μg Ccr.

Epi, Ecm, and Mcd assays. Activities of Epi, Ecm, and Mcd were measured in a coupled assay in which the oxidation of (2S)-methylsuccinyl-CoA was monitored at 300 nm with ferrocenium hexafluorophosphate ($FePF_6$) as the artificial electron acceptor. The estimated differential extinction coefficient $\Delta\varepsilon_{300\ nm}$ ($FePF_6$ minus mesaconyl-C1-CoA) was 2.74 mM^{-1} cm^{-1} (29). A substrate mixture contained 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 0.6 mM NADPH, 0.6 mM crotonyl-CoA, and 3 μg Ccr was incubated for 5 minutes to synthesize (2S)-ethylmalonyl-CoA. The reaction mixture contained 100 mM Tris-HCl pH 7.5, 0.2 mM $FePF_6$, 0.1 ml substrate mixture and appropriate amounts of Epi, Ecm, and Mcd. For Epi assay, 0.019 μg Epi, 9 μg Ecm, and 12 μg Mcd were used. For Ecm assay, 3.8 μg Epi, 0.37 μg Ecm, and 12 μg Mcd were used. For Mcd assay, 3.8 μg Epi, 9 μg Ecm, and 0.24 μg Mcd were used. Before assays, Epi was incubated with 0.4 mM $CoCl_2$ on ice for 1 h to enhance its activity, and Ecm was incubated with 0.1 mM Coenzyme B12 in dark on ice for 1 h for activation.

Mct assay. The Mct activity was measured by monitoring the increase of absorbance at 290 nm corresponding to the conversion of mesaconyl-C1-CoA to mesaconyl-C4-CoA. The estimated differential absorption coefficient $\Delta\varepsilon_{290\ nm}$ (mesaconyl-C4-CoA minus mesaconyl-C1-CoA) was 3.3 mM^{-1} cm^{-1} (17). A mixture contained 100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.5 mM propionyl-CoA, 8 mM glyoxylate, 10 μg Mcl, and 35 μg Mch was incubated for 5 minutes to synthesize mesaconyl-C1-CoA. Then Mcl and Mch were removed using HisPur Ni-NTA resin and 1 μg Mct was added into the mixture to start the reaction.

Meh assay. The Meh activity was measured by monitoring the decrease of absorbance at 290 nm corresponding to the consumption of mesaconyl-C4-CoA. The estimated extinction coefficient $\varepsilon_{290\ nm}$ of mesaconyl-C4-CoA was 6.7 mM^{-1} cm^{-1} (17). A mixture contained 100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.5 mM propionyl-CoA, 8 mM glyoxylate, 10 μg Mcl, 35 μg Mch, and 2.5 μg Mct was incubated for 15 minutes to synthesize mesaconyl-C4-CoA. Then Mcl, Mch, and Mct were removed using HisPur Ni-NTA resin and 0.2 μg Meh was added into the mixture to start the reaction.

Ccl assay. The Ccl activity was measured in a coupled assay in which the formation of pyruvate from (*S*)-citramalyl-CoA was followed at 340 nm by monitoring the consumption of NADH with lactate dehydrogenase (Ldh). To synthesize (*S*)-citramalyl-CoA, a mixture contained 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 0.5 mM NADPH, 0.5 mM crotonyl-CoA, 3 μg Ccr was incubated for 5 minutes to convert crotonyl-CoA to (2S)-ethylmalonyl-CoA first, then 1 mM $FePF_6$, 4.5 μg Epi, 3.5 μg Ecm, 4 μg Mcd, 8.5 μg Mct, 5 μg Meh were added and the resulting mixture was incubated for 20 minutes. Precipitations formed from the reduction of $FePF_6$ were removed by centrifugation. 0.5 ml supernatant was taken out and 0.2 mM NADH, 2 U Ldh (Sigma), and 2.25 μg Ccl or 3.10 μg Mcl were added to start the assay.

Pps and Ppc assays. Activities of Pps and Ppc were measured in a coupled assay in which the formation of malate was followed at 340 nm by monitoring the consumption of NADH with Mdh. The reaction mixture contained 100 mM Tris-HCl pH 7.0, 5 mM $MgCl_2$, 50 mM $NaHCO_3$, 1 mM ATP, 1 mM pyruvate, 0.1 mM acetyl-CoA, 0.3 mM NADH and appropriate amounts of Pps, Ppc, and Mdh. For Pps assay, 2 U Ppc (Sigma), 2 U Mdh (Sigma), and 5 μg Pps were used. For Ppc assay, 17 μg Pps, 2 U Mdh (Sigma), and 2.7 μg Ppc were used.

Mdh assay. The Mdh activity was measured by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADH. The reaction mixture contained 100mM Tris-HCl pH 7.5, 0.3 mM NADH, 1 mM oxaloacetate, and 0.06 μg Mdh.

Mtk and Mcl assays. Activities of Mtk and Mcl were measured in a coupled assay in which the formation of glyoxylate was monitored at 324 nm by reacting it with phenylhydazine to form glyoxylate-phenylhydrazone, which had an absorption coefficient $\varepsilon_{324\ nm}$ of 8.08 mM^{-1} cm^{-1}. The reaction mixture contained 100mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 1 mM ATP, 1 mM CoA, 5 mM malate, 10 mM phenylhydazine, and appropriate amounts of Mtk and Mcl. For Mtk assay, 10 μg Mcl and 8.3 μg Mtk were used. For Mcl assay, 8.3 μg Mtk and 0.4 μg Mcl were used.

2.4.6 HPLC Analysis

CoA compounds, glyoxylate phenylhydrazone and pyruvate phenylhydrazone were analyzed by an Agilent 1200 HPLC equipped with a reversed-phase C18 column (LiChrospher 100, end-capped, 5 μm, LiChroCART 125-4 125 × 4 mm). Separation was performed using one of the two gradient systems of 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent A) and 20% acetonitrile (MeCN) in 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent B). Column temperature was maintained at 25 ^oC and the flow rate stayed constant at 1 ml/min.

System 1. Initial conditions were 90% solvent A and 10% solvent B. At 0 min, solution B was increased linearly to 50% over 40 min. At 40 min, the composition of B was increased to 100% over 1 min and held isocratic for 2 min. At 43 min, solution B was decreased to its initial 10% condition over 1 min and held for 16 min to reequilibrate the column.

System 2. Initial conditions were 90% solvent A and 10% solvent B. At 0 min, solution B was increased linearly to 100% over 40 min. At 40 min, the composition of B was held isocratic for 3 min. At 43 min, solution B was decreased to its initial 10% condition over 1 min and held for 16 min to reequilibrate the column.

2.4.7 LC-MS analysis

The labeling pattern of glyoxylate phenylhydrazone was analyzed by a Shimadzu 2020 EV LC-MS equipped with a reversed-phase C18 column (Phenomenex Kinetex, 1.7 μm, 100 Å, 100 × 2.1mm) using positive-and negative-mode electrospray ionization with a linear gradient of 5-95% MeCN-H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 3min with a constant flow rate of 0.3 ml min^{-1}.

2.4.8 Conditions for demonstration of the rGPS cycle in two pots

The rGPS cycle was divided into two parts: 1. from pyruvate and acetyl-CoA to glyoxylate and ethylmalonyl-CoA; 2. from ethylmalonyl-CoA to pyruvate and acetyl-CoA. To demonstrate part 1, a 600 μL reaction mixture containing 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, 4 mM glucose 6-phosphate, 3 mM ATP, 9 mM creatine phosphate, 0.17 mM NADH, 0.17 mM NADPH, 0.2 mM pyruvate, 0.2 mM acetyl-CoA, 19.2 μg Pps, 40 μg Ppc, 2.6 μg Mdh, 42.3 μg Mtk, 22 μg Mcl, 29.7 μg AtoB, 19.4 μg Bhbd, 16.1 μg Crt, 53.9 μg Ccr, 9.5 μg G6PDH, and 13.3 μg CPK was used. The reaction was started with the addition of Pps and AtoB. A sample (200 μL) was withdrawn before adding Pps and AtoB as time 0 min. Another sample (200 μL) was withdrawn at 45 min after adding Pps and AtoB as time 45 min. To demonstrate part 2, a 600 μL reaction mixture containing 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, 0.25 mM ethylmalonyl-CoA (produced from crotonyl-CoA using Ccr), 0.5 mM $FePF_6$, 32.7 μg Epi,

38.5 μg Ecm, 51.6 μg Mcd, 8.6 μg Mct, 8.6 μg Meh, and 33.4 μg Ccl was used. The reaction was started with the addition of Epi. A sample (200 μL) was withdrawn before adding Epi as time 0 min. Another sample (200 μL) was withdrawn at 30 min after adding Epi as time 30 min. The samples were quenched with 5% formic acid, derivatized with 10 mM phenylhydrazine and analyzed with HPLC (system 2).

2.4.9 Conditions for one-pot stepwise reconstitution of the rGPS cycle

The one-pot stepwise reconstitution of rGPS was performed in 600 μL of 100 mM MOPS pH 7.0 containing 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, 0.5 mM NADH, 0.5 mM NADPH, and 0.5 mM acetyl-CoA. A negative control sample (100 μL) was withdrawn before adding enzymes as step 0. 29.7 μg AtoB, 19.4 μg Bhbd, 16.1 μg Crt, and 53.9 μg Ccr were added to start the reactions. After incubating for 30 min, a sample (100 μL) was withdrawn as step 1. Then, 32.7 μg Epi, 38.5 μg Ecm, 51.6 μg Mcd, and 1 mM $FePF_6$ were added and the reaction mixture was further incubated for 20 min before taking another sample (100 μL) as step 2. Finally, 4.3 μg Mct, 3.4 μg Meh, 20 μg Ccl, 9.6 μg Pps, 20 μg Ppc, 1.3 μg Mdh, 21.2 μg Mtk, 11 μg Mcl, 4 mM ATP, 2 mM NADH, and 1 mM CoA were added and the reaction mixture was further incubated for 45 min before the final sample (100 μL) was taken as step 3. The samples were quenched with 5% formic acid, derivatized with 10 mM phenylhydrazine and analyzed with HPLC (system 1).

2.4.10 Semi-continuous operation of the rGPS cycle using a two-pot immobilized enzyme system

To prepare the two-pot immobilized enzyme system, 15 his-tagged enzymes of the rGPS cycle were separated into two groups with each group of enzymes mixed together and bound to his-affinity resin (Zymo Research) to form the immobilized enzyme resin. A 600 μL mixture

containing 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 5 mM $MgCl_2$, 12.9 μg AtoB, 16.1 μg Bhbd, 13.6 μg Crt, 40.8 μg Ccr, 69.2 μg Pps, 54.6 μg Ppc, 31 μg Mdh, 33.3 μg Mtk, and 19.8 μg Mcl was incubated with his-affinity resin drained from 300 μl his-affinity gel for 5 min and centrifuged to remove the liquid to form resin 1. A 600 μL mixture containing 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 5 mM $MgCl_2$, 0.1 mM Coenzyme B12, 0.4 mM $CoCl_2$, 18.9 μg Epi, 91.8 μg Ecm, 122 μg Mcd, 23.1 μg Mct, 16 μg Meh, and 35.1 μg Ccl was incubated with his-affinity resin drained from 300 μl his-affinity gel for 5 min and centrifuged to remove the liquid to form resin 2.

The semi-continuous operation was started by adding a 600 μL reaction mixture composed of 100 mM Tris-HCl pH 7.0, 50 mM $NaH^{13}CO_3$, 5 mM $MgCl_2$, 0.1 mM Coenzyme B12, 0.4 mM $CoCl_2$, 0.5 mM Crotonyl-CoA, and 0.5 mM NADPH into resin 1 to react for 5 min. Then the reaction mixture was separated from resin 1 by centrifugation and added into resin 2 with 1 mM $FePF_6$ to react. After 30 min, the reaction mixture was separated from resin 2 and added back into resin 1 with 2 mM ATP, 3 mM NADPH, and 4 mM NADH to react for 45 min. After this reaction, the resulting mixture was separated from resin 1. 100 μL of the mixture was withdrawn and 300 μL of the rest was added into resin 2 with 2 mM $FePF_6$ to react for 30 min. Then the reaction mixture was separated from resin 2 and added into resin 1 with 1 mM ATP, 2 mM CoA, and 2 mM NADH to convert intermediates to glyoxylate. After 50 min, the reaction mixture was separated from resin 1 and 100 μL of the mixture was withdrawn. The withdrawn samples were quenched in formic acid with a final concentration of 5%. 10 mM phenylhydrazine was used for glyxoylate derivatization. The labeling pattern of glyoxylate phenylhydrazone was analyzed with LC-MS.

2.5 Supplementary figures and tables

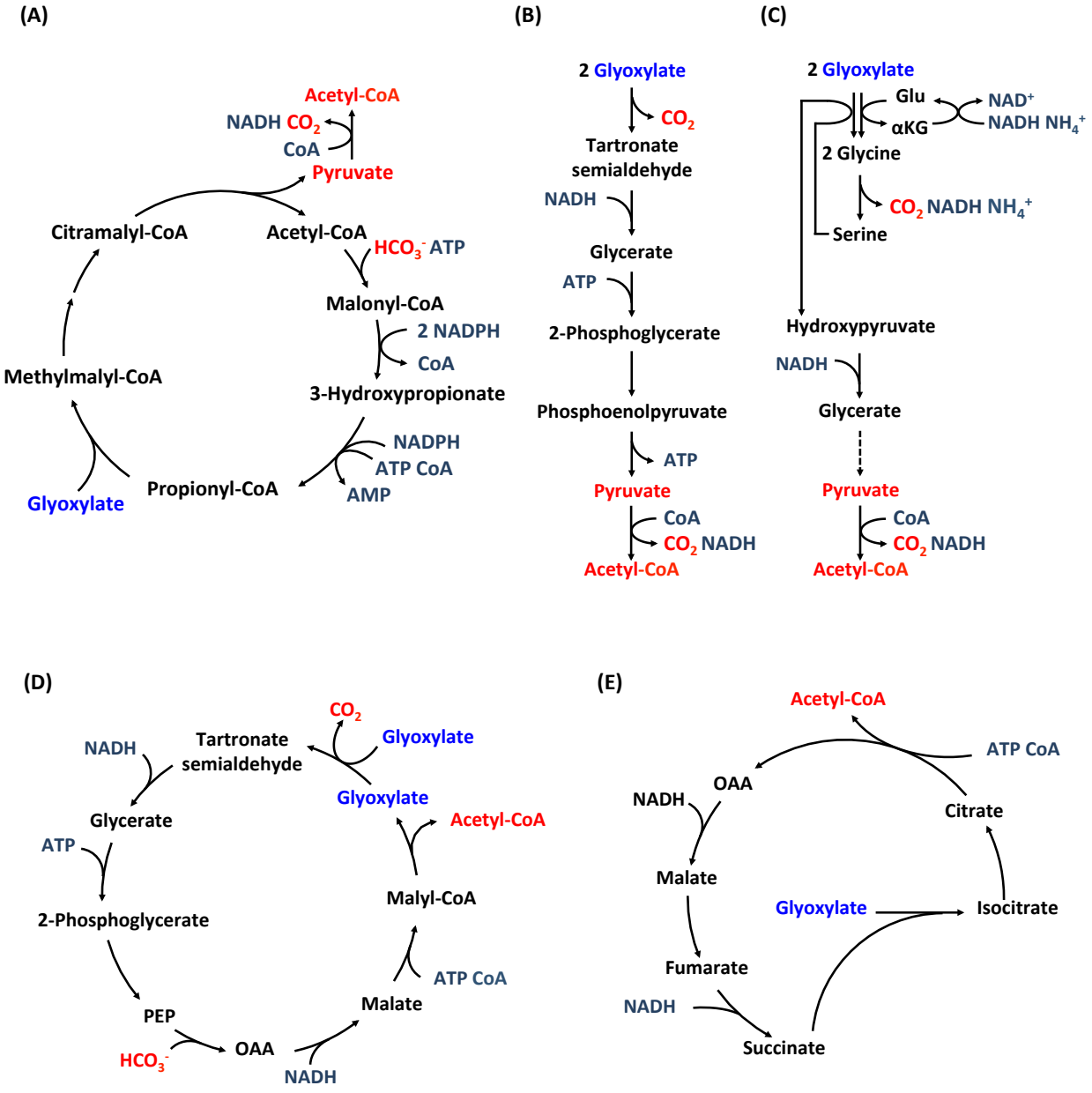
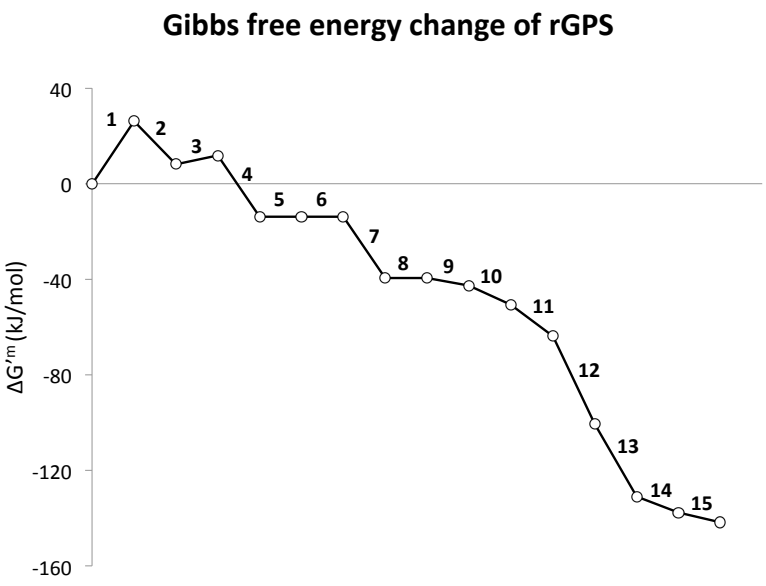


Fig. 2.S1. Pathways of converting glyoxylate to pyruvate and acetyl-CoA.

(A) In the 3HP bicycle, glyoxylate can be converted to pyruvate with one bicarbonate fixed through ACC, and pyruvate is further decarboxylated to acetyl-CoA through pyruvate dehydrogenase (PDH). (B) In the bacterial glycolate catabolic pathway, two glyoxylate are

converted to one pyruvate with one CO2 loss, and pyruvate is further decarboxylated to acetyl-CoA through PDH. (C) In the plant glycolate assimilation pathway, two glyoxylate are converted to one pyruvate with one CO2 loss, and pyruvate is further converted to acetyl-CoA through PDH. (D) In the synthetic MCG pathway, one glyoxylate can be converted to one acetyl-CoA without net carbon loss. (E) In the synthetic reverse glyoxylate shunt pathway, one glyoxylate can be converted to one acetyl-CoA without net carbon loss.



Reaction	Enzyme	$\Delta_r G'^m$ (kJ/mol)
1	AtoB	26.1
2	Bhbd	-18
3	Crt	3.8
4	Ccr	-25.7
5	Epi	0
6	Ecm	0
7	Mcd	-25.7
8	Mct	0
9	Meh	-3.1
10	Ccl	-8
11	Pps	-13.3
12	Ppc	-36.7
13	Mdh	-30.3
14	Mtk	-6.7
15	Mcl	-4.2

Fig. 2.S2. Gibbs free energy change of the rGPS cycle starting from two acetyl-CoA. $\Delta_rG'^m$ and $\Delta G'^m$ were calculated using the online tool-eQuilibrator (http://equilibrator.weizmann.ac.il/) (30). The conditions for $\Delta_rG'^m$ and $\Delta G'^m$ are at 25 ^oC temperature, 1 bar pressure, PH=7.0, Ionic strength=0.1 M and 1 mM concentration used for all reactants.

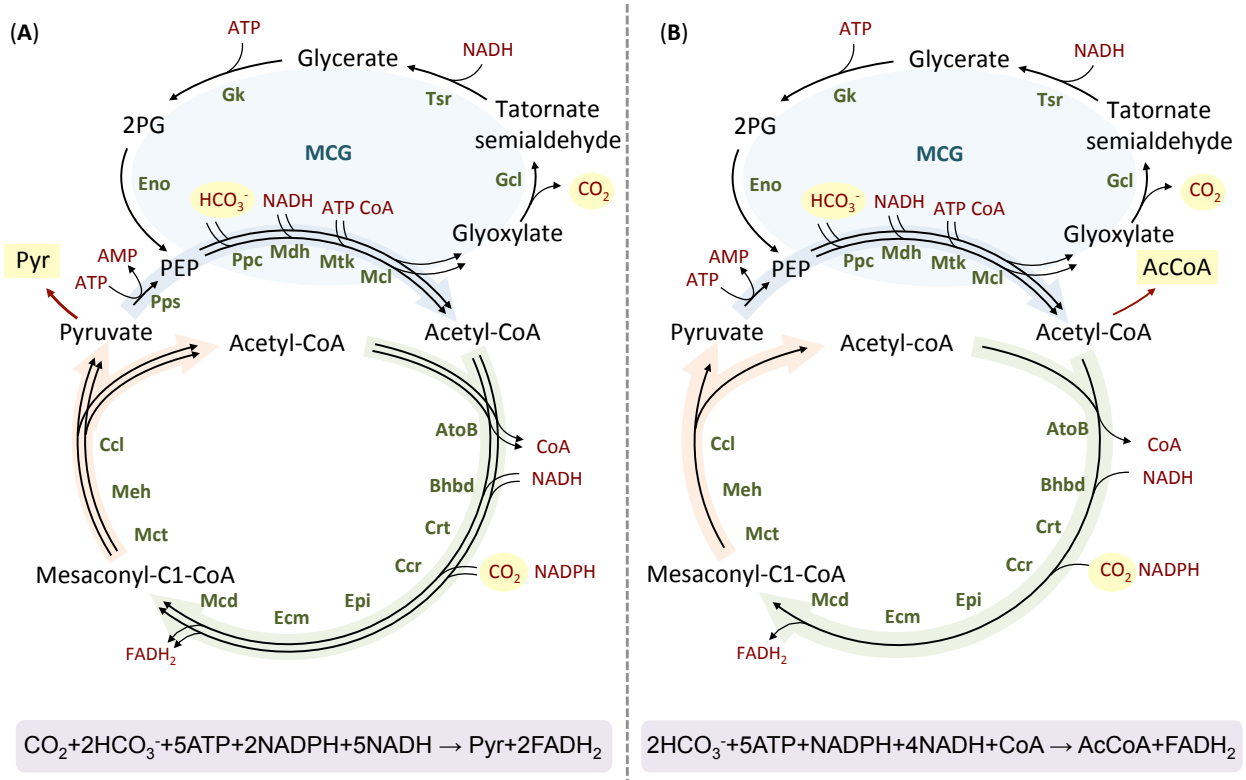
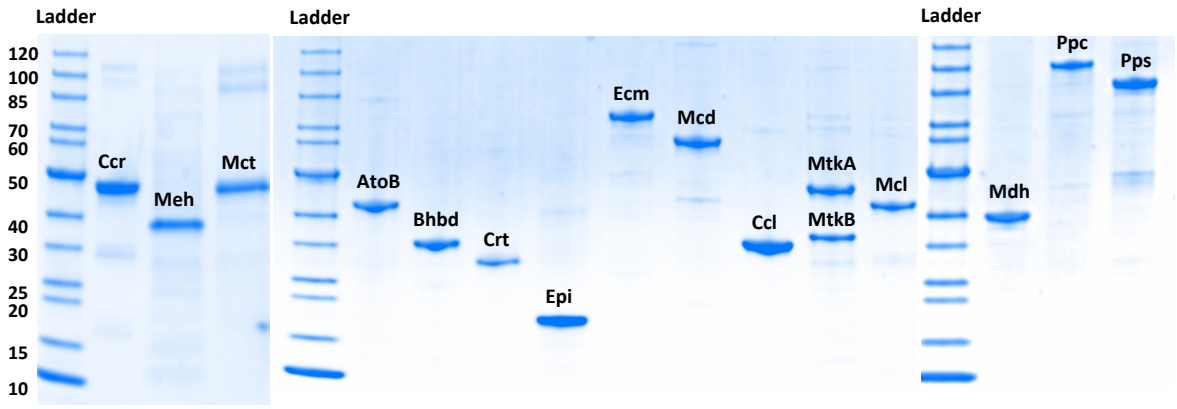


Fig. 2.S3. Detailed Structure of rGPS+MCG.

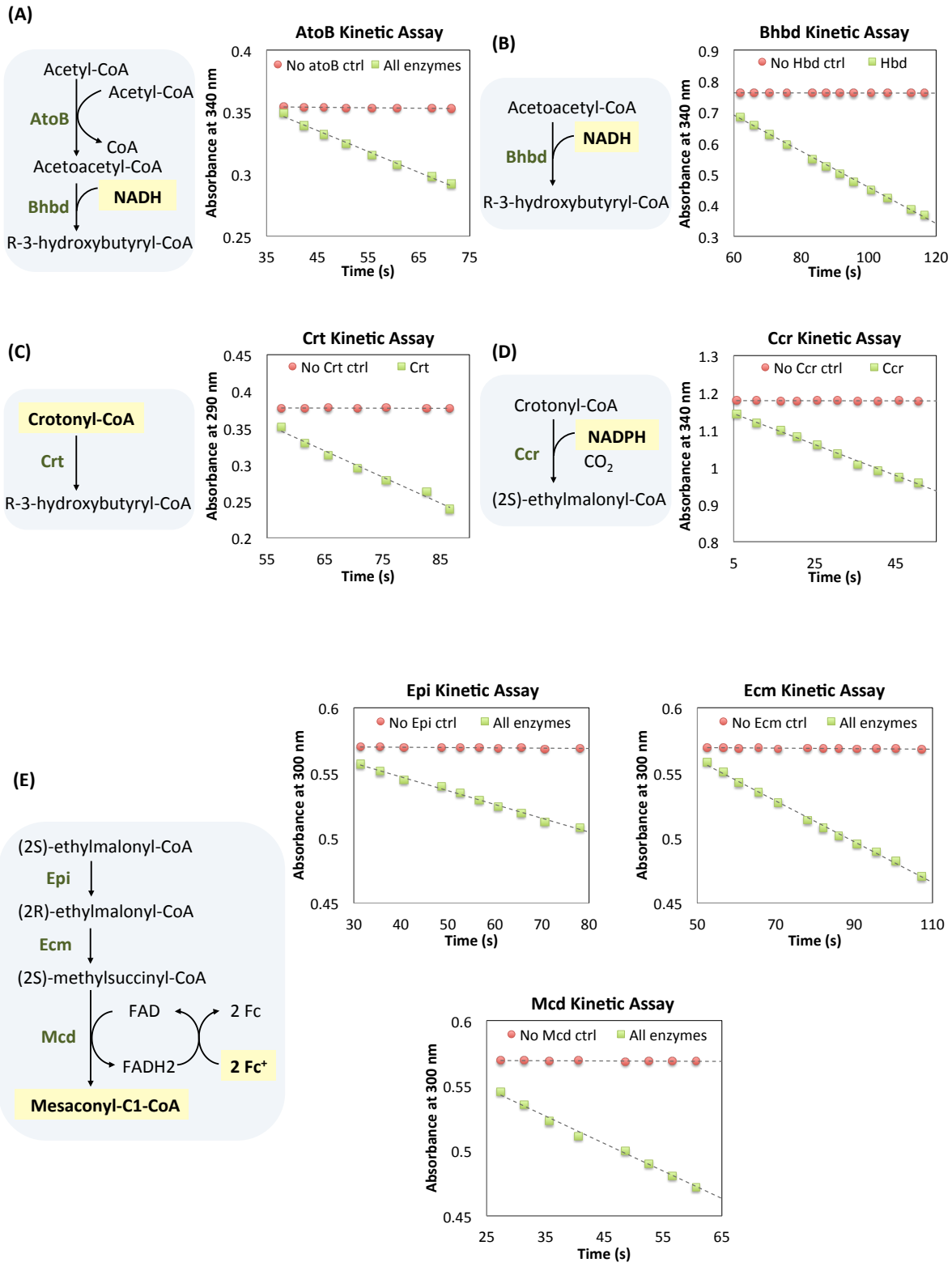
(A) The combination of rGPS with MCG for pyruvate synthesis. (B) The combination of rGPS with MCG for acetyl-CoA synthesis. The sum of reactions is shown in the purple box. See Table 2.S3 for enzyme abbreviations of rGPS. 2PG, 2-phospho-D-glycerate; Gcl, glyoxylate carboligase; Tsr, tartronate semialdehyde reductase; Gk, glycerate 2-kinase; Eno, enolase.



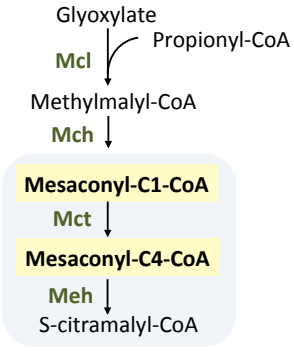
Protein	Ccr	Meh	Mct	AtoB	Bhbd	Crt	Epi	Ecm	Mcd	Ccl	Mtk A	Mtk B	Mcl	Mdh	Ppc	Pps
Theoretical Size (kDa)	48	31	45	40	31	28	14	71	60	29	42	31	36	32	103	87

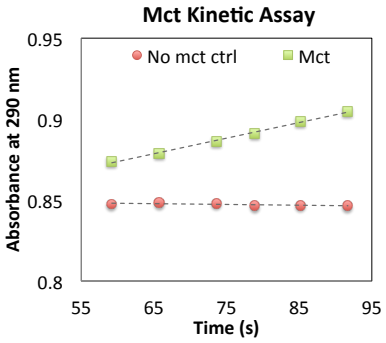
Fig. 2.S4. SDS-PAGE and theoretical sizes of all his-tagged proteins for building rGPS *in vitro*.

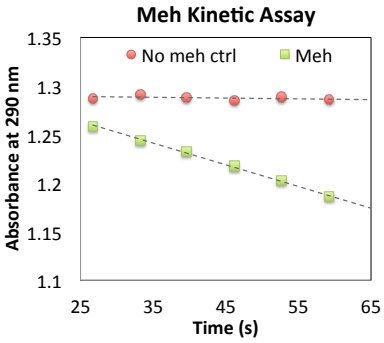
SDS-PAGE denaturing gels of purified Ccr, Meh, Mct, AtoB, Bhbd, Crt, Epi, Ecm, Mcd, Ccl, Mtk, Mcl, Mdh, Ppc and Pps (see Table 2.S3 for enzyme abbreviations). Theoretical sizes were calculated from their corresponding nucleotide sequences (his-tag sequences not included). All proteins were designed with polyhistidine tags and purified by immobilized metal (Ni^{2+} or Co^{2+}) affinity chromatography. The molecular weights of standard proteins are indicated.


(A)
Acetyl-CoA
Acetyl-CoA
AtoB
CoA
Acetoacetyl-CoA
Bhbd
NADH
R-3-hydroxybutyryl-CoA
AtoB Kinetic Assay
No atoB ctrl
All enzymes
Absorbance at 340 nm
Time (s)
(B)
Acetoacetyl-CoA
Bhbd
NADH
R-3-hydroxybutyryl-CoA
Bhbd Kinetic Assay
No Hbd ctrl
Hbd
Absorbance at 340 nm
Time (s)
(C)
Crotonyl-CoA
Crt
R-3-hydroxybutyryl-CoA
Crt Kinetic Assay
No Crt ctrl
Crt
Absorbance at 290 nm
Time (s)
(D)
Crotonyl-CoA
Ccr
NADPH
CO2
(2S)-ethylmalonyl-CoA
Ccr Kinetic Assay
No Ccr ctrl
Ccr
Absorbance at 340 nm
Time (s)
(E)
(2S)-ethylmalonyl-CoA
Epi
(2R)-ethylmalonyl-CoA
Ecm
(2S)-methylsuccinyl-CoA
Mcd
FAD
2 Fc
FADH2
2 Fc+
Mesaconyl-C1-CoA
Epi Kinetic Assay
No Epi ctrl
All enzymes
Absorbance at 300 nm
Time (s)
Ecm Kinetic Assay
No Ecm ctrl
All enzymes
Absorbance at 300 nm
Time (s)
Mcd Kinetic Assay
No Mcd ctrl
All enzymes
Absorbance at 300 nm
Time (s)

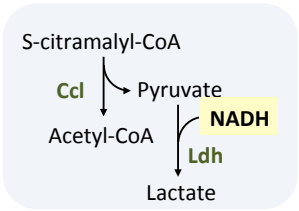
(F)

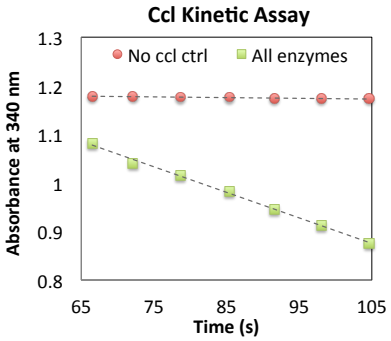

Glyoxylate
Propionyl-CoA
Mcl
Methylmalyl-CoA
Mch
Mesaconyl-C1-CoA
Mct
Mesaconyl-C4-CoA
Meh
S-citramalyl-CoA

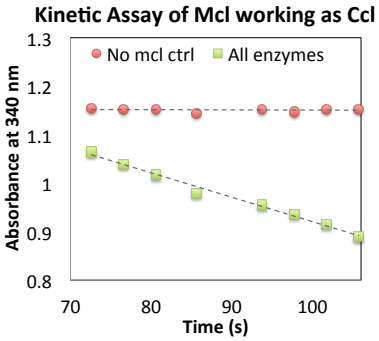

Mct Kinetic Assay
No mct ctrl
Mct
Absorbance at 290 nm
0.95
0.9
0.85
0.8
55
65
75
85
95
Time (s)


Meh Kinetic Assay
No meh ctrl
Meh
Absorbance at 290 nm
1.35
1.3
1.25
1.2
1.15
1.1
25
35
45
55
65
Time (s)

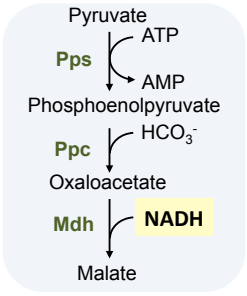
(G)

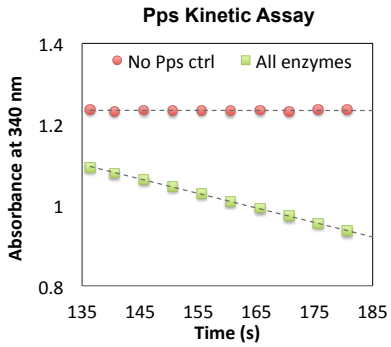

S-citramalyl-CoA
Ccl
Pyruvate
NADH
Acetyl-CoA
Ldh
Lactate

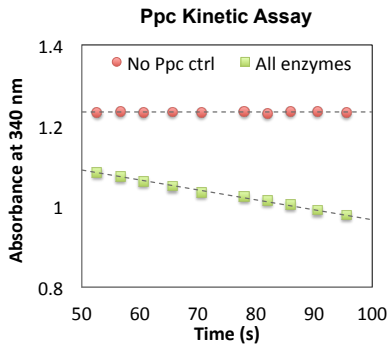

Ccl Kinetic Assay
No ccl ctrl
All enzymes
Absorbance at 340 nm
1.3
1.2
1.1
1
0.9
0.8
65
75
85
95
105
Time (s)


Kinetic Assay of Mcl working as Ccl
No mcl ctrl
All enzymes
Absorbance at 340 nm
1.3
1.2
1.1
1
0.9
0.8
70
80
90
100
Time (s)

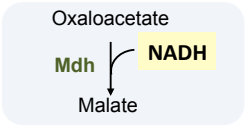
(H)

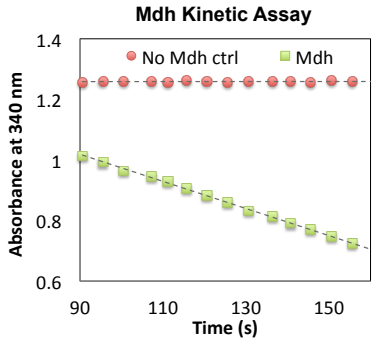

Pyruvate
ATP
Pps
AMP
Phosphoenolpyruvate
HCO3-
Ppc
Oxaloacetate
Mdh
NADH
Malate


Pps Kinetic Assay
No Pps ctrl
All enzymes
Absorbance at 340 nm
1.4
1.2
1
0.8
135
145
155
165
175
185
Time (s)


Ppc Kinetic Assay
No Ppc ctrl
All enzymes
Absorbance at 340 nm
1.4
1.2
1
0.8
50
60
70
80
90
100
Time (s)

(I)


Oxaloacetate
Mdh
NADH
Malate


Mdh Kinetic Assay
No Mdh ctrl
Mdh
Absorbance at 340 nm
1.4
1.2
1
0.8
0.6
90
110
130
150
Time (s)

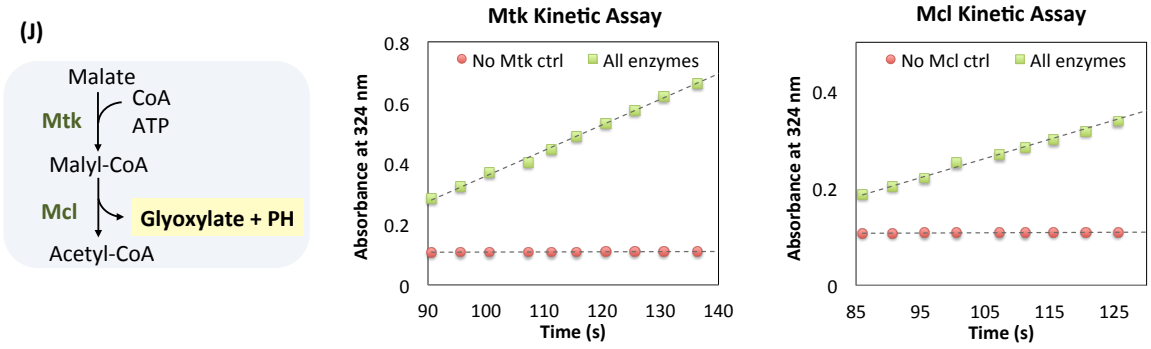


Fig. 2.S5. *In vitro* Enzyme Assay Schemes.

A series of kinetic enzyme assays were performed to test the activity of purified enzymes. (A) AtoB assay. (B) Bhbd assay. (C) Crt assay. (D) Ccr assay. (E) Epi, Ecm, and Mcd assays. (F) Mct and Meh assays. (G) Ccl assay. (H) Pps and Ppc assays. (I) Mdh assay. (J) Mtk and Mcl assays. See "enzyme assays" in methods for details. Mch, methylmalyl-CoA dehydratase; Ldh, lactate dehydrogenase. Specific activities of all enzymes are listed in Table 2.S3.

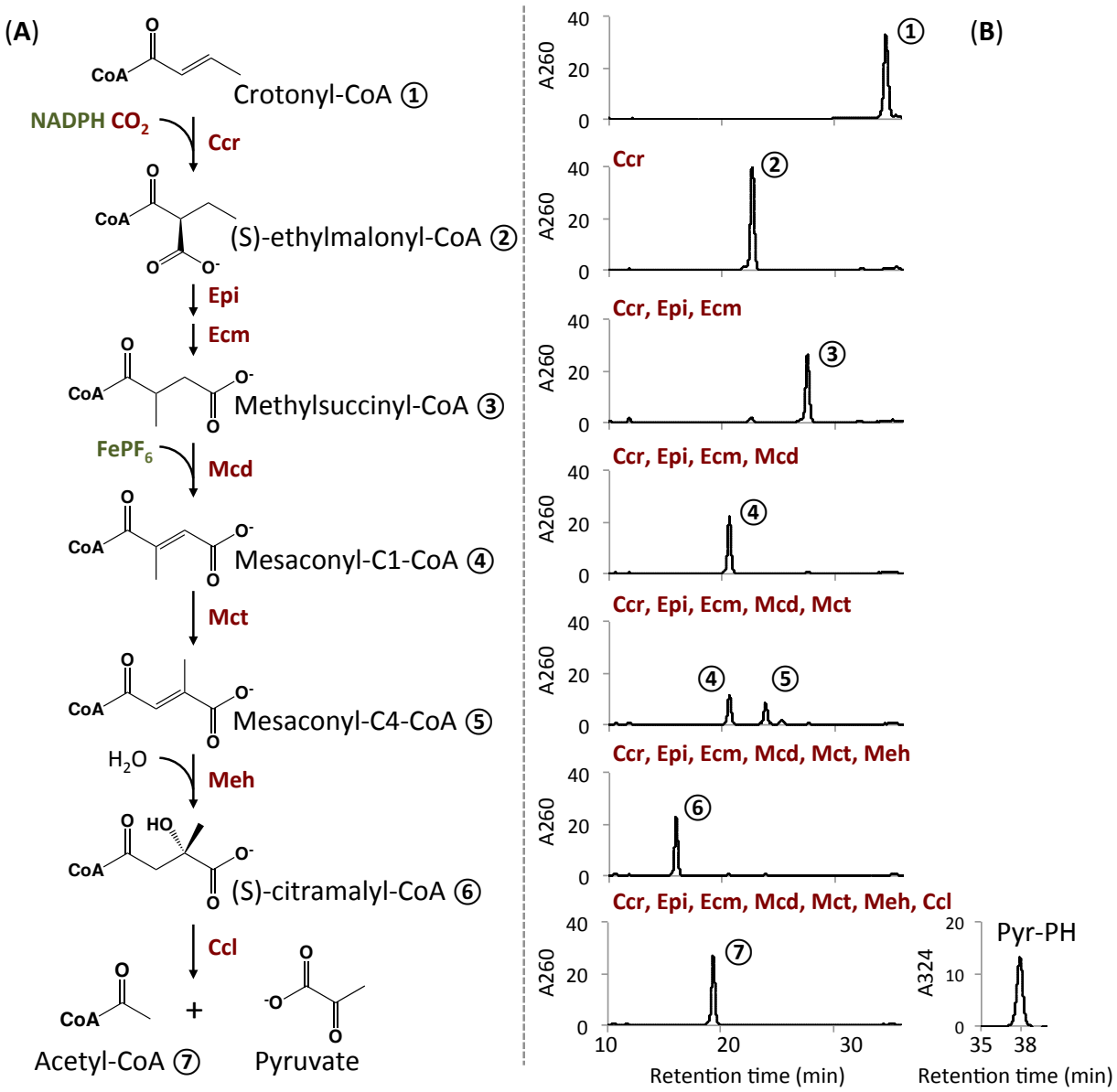


Fig. 2.S6. ***In vitro*** **demonstration of the stepwise conversion of crotonyl-CoA to pyruvate and acetyl-CoA in the rGPS cycle.**

(A) Scheme of converting crotonyl-CoA to acetyl-CoA and pyruvate in 7 steps. **(B)** HPLC analysis of the products of incubating the reaction mixture containing crotonyl-CoA with corresponding enzymes (marked as red) and cofactors. CoA intermediates are numbered according to panel (A). CoA esters were followed at 260 nm, pyruvate was derivatized by phenylhydrazine and detected at 324 nm.

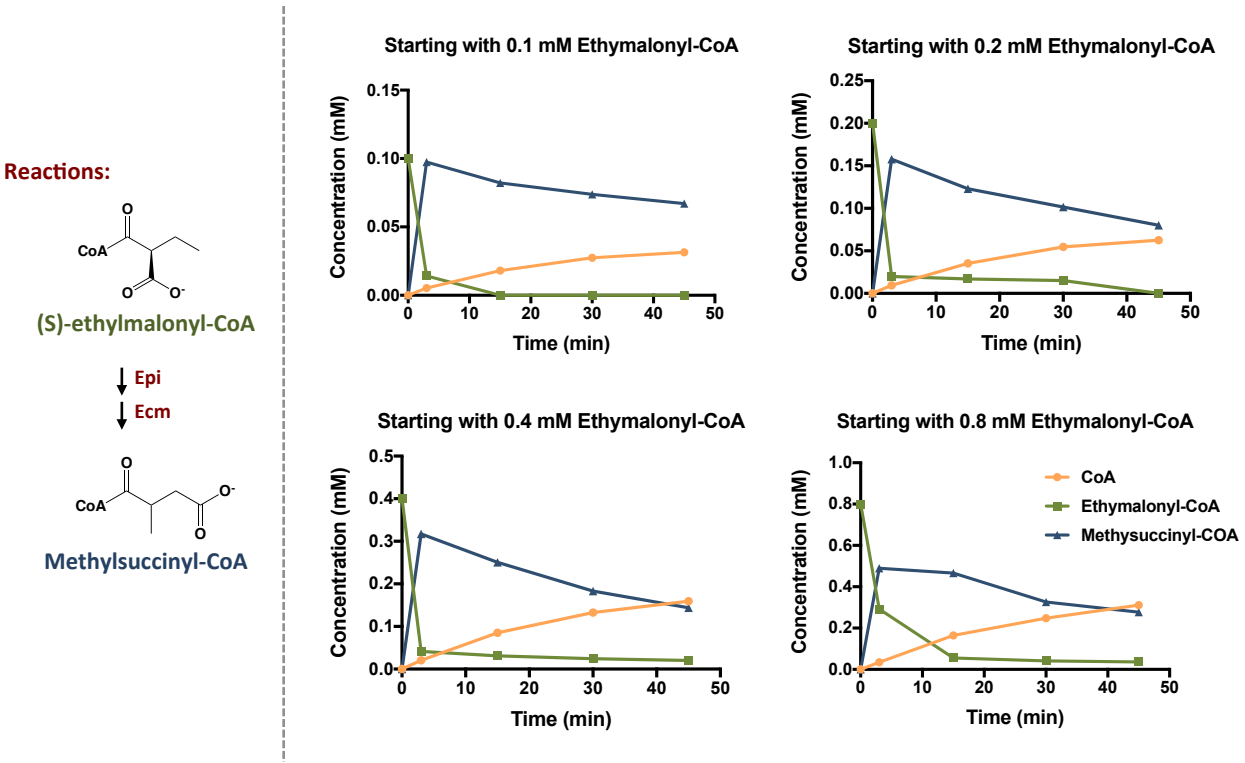


Fig. 2.S7. Stability test of methylsuccinyl-CoA in reaction conditions.

Shown is the time course of producing methylsuccinyl-CoA from 0.1 mM, 0.2 mM, 0.4 mM, or 0.8 mM (S)-ethylmalonyl-CoA using Epi and Ecm in reaction conditions (100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$). The concentrations of free CoA and CoA esters were measured with HPLC (system 2).

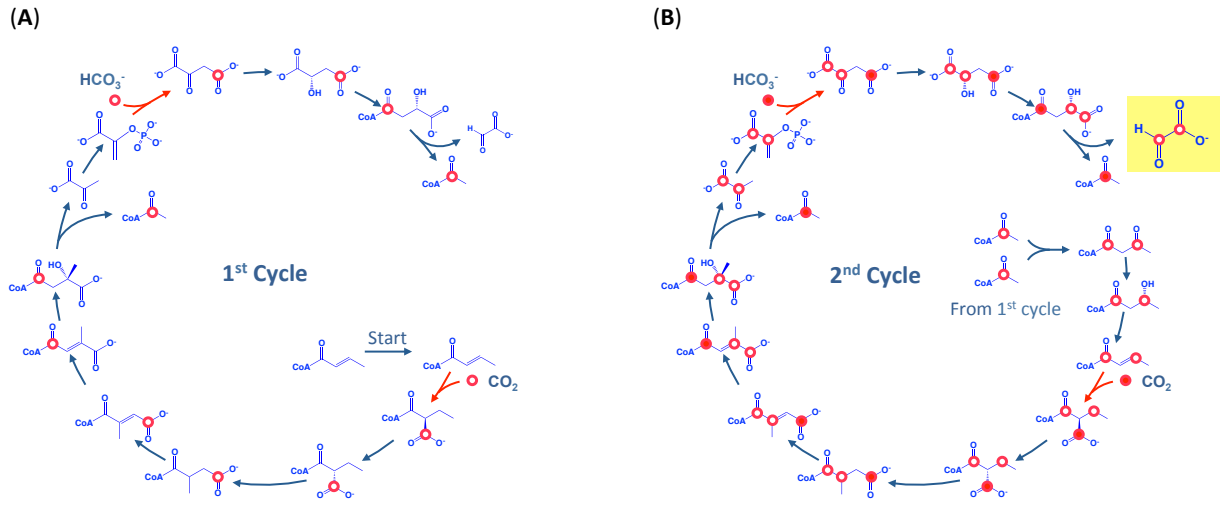


Fig. 2.S8. Expected 13C-labeling patterns of the rGPS cycle with $^{13}CO_2$ and $H^{13}CO_3^-$.

Two carbon fixation steps are featured in red. **(A)** In the first cycle, $^{13}CO_2$ (shown in red circle) is incorporated into crotonyl-CoA and ends up in acetyl-CoA, and $H^{13}CO_3^-$ (shown in red circle) is incorporated into phosphoenolpyruvate and also ends up in acetyl-CoA. **(B)** Two labeled acetyl-CoA molecules enter into the 2nd cycle and finally produce double-labeled glyoxylate. $^{13}CO_2$ and $H^{13}CO_3$ incorporated in the second cycle are shown in red dots.

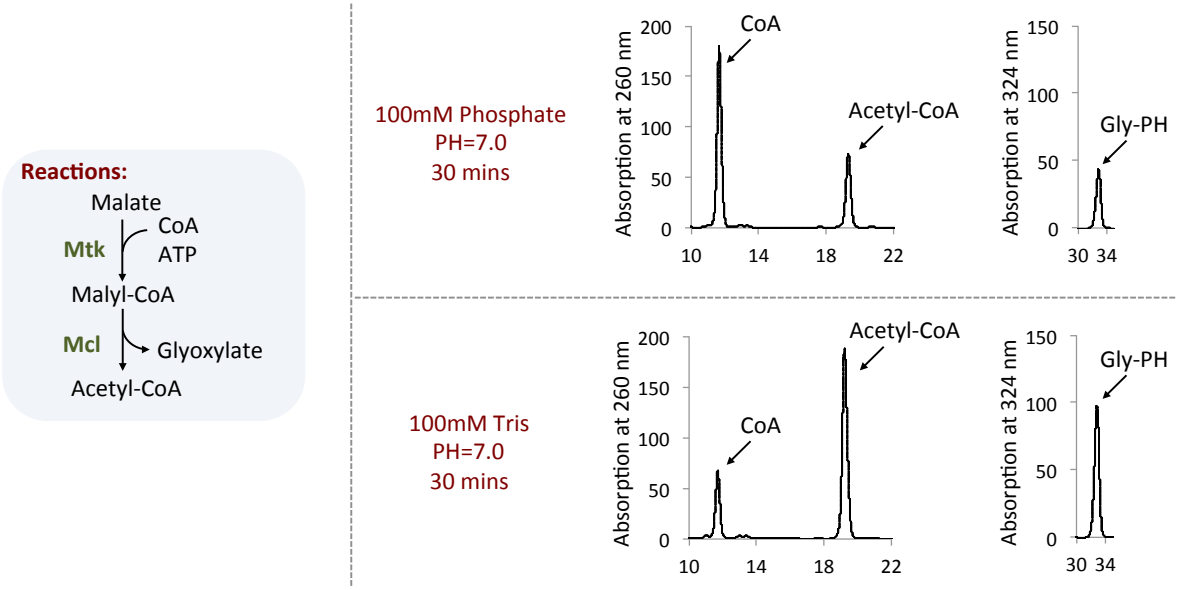


Fig. 2.S9. The inhibition of phosphate buffer on mtk mcl reactions.

Shown are HPLC results of producing glyoxylate and acetyl-CoA from malate and CoA using mtk and mcl in Tris (PH=7.0) buffer or Phosphate (PH=7.0) buffer for 30 minutes. Free CoA and CoA esters were followed at 260 nm, gloxylate was derivatized by phenylhydrazine and detected at 324 nm. Gly-PH, glyoxylate phenylhydrazone.

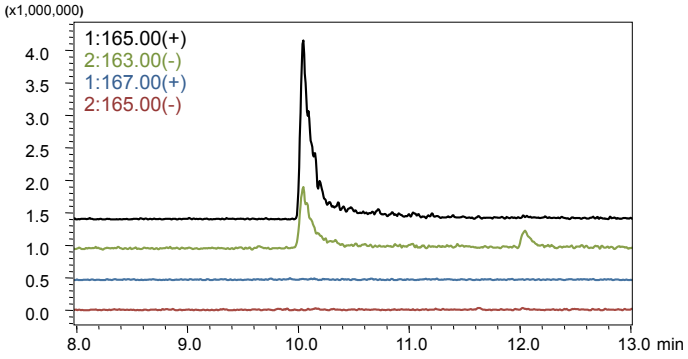


Fig. 2.S10. LC-MS analysis of the glyoxylate phenylhydrazone standard.

For glyoxylate phenylhydrazone, $[M+H]^+$ = 165 and $[M-H]^-$ = 163. For double-labeled glyoxylate phenylhydrazone, $[M+H]^+$ = 167 and $[M-H]^-$ = 165.

Table 2.S1. Comparison of different carboxylases involved in primary metabolism.

Carboxylase abbrev.	Carbon species	Cofactor	Involved in natural CO_2 fixation pathway	Oxygen sensitivity	Specific Activity (μmol/min/mg)[a]
RuBisco	CO_2	No	Yes (CBB)	No, but has oxygenase activity	3.5[b]
ACC	HCO_3^-	ATP, biotion	Yes (HP bi, HP/HB)	No	18[b]
PCC	HCO_3^-	ATP, biotion	Yes (HP bi, HP/HB)	No	29.6[b]
Pyc	HCO_3^-	ATP, biotion	No (mainly in anaplerosis)	No	32.4[b]
Ppc	HCO_3^-	No	Yes (DC/HB)	No	35.2[b]
POR	CO_2	ferredoxin, TPP	Yes (WL, rTCA, DC/HB)	Yes	<1
KOR	CO_2	ferredoxin, TPP	Yes (rTCA)	Yes	<1
CCR	CO_2	NADPH	No (found in ethylmalonyl-CoA pathway)	No	130[b]

[a] from Bar-Even A., et al. PNAS 2010 (13). [b] Under saturating CO_2 or HCO_3^- concentrations. RuBisCo, ribulose 1,5-bisphosphate carboxylase/oxygenase; ACC, acetyl-CoA carboxylase; PCC, propionyl-CoA carboxylase; Pyc, pyruvate carboxylase; Ppc, phosphoenolpyruvate carboxylase; POR, pyruvate: ferredoxin oxidoreductase; KOR, α-Ketoglutarate: ferredoxin oxidoreductase; CCR, Crotonyl-CoA carboxylase/reductase; TPP, thiamine pyrophosphate. CBB, Calvin-Benson-Bassham cycle; rTCA, reductive tricarboxylic acid cycle; WL, Wood-Ljungdahl pathway; HP bi, 3-hydroxypropionate bicycle; HP/HB, 3-hydroxypropionate/4-hydroxybutyrate cycle; DC/HB, dicarboxylate/4-hydroxybutyrate cycle.

Table 2.S2. Comparison of rGPS with natural and realized synthetic CO2 fixation pathways.

To produce one mol of pyruvate

Pathway	Carbon species		ATP	Reducing equivalents			ATP equiv. per CO_2	Oxygen sensitivity	Status
	CO_2	HCO_3^-		NAD(P)H	$FADH_2$	Fd^{2-}_{red}			
CBB cycle	3	0	7	5	0	0	6.5	No, but photorespiration	Natural
rTCA cycle	3	0	2	2	1[a]	2	-	Yes	Natural
WL pathway[b]	3	0	1	2	0	3	-	Yes	Natural
3HP bicycle	0	3	7	6	-1	0	6.8	No	Natural
HP/HB cycle	0	3	9 or 5[c]	6	-1	0	7.5 or 6.2	Yes or No[d]	Natural
DC/HB cycle	2	1	5	2	1[a]	2	-	Yes	Natural
CETCH 5.4	3	0	2	9	-4	0	6.2	No	Synthetic
rGPS	1	2	5	7	-2	0	6.5	No	Synthetic

To produce one mol of acetyl-CoA

Pathway	Carbon species		ATP	Reducing equivalents			ATP equiv. per CO2	Oxygen sensitivity	Status
	CO_2	HCO_3^-		NAD(P)H	$FADH_2$	Fd^{2-}_{red}			
CBB cycle	2	0	7	4	0	0	8.5	No, but photorespiration	Natural
rTCA cycle	2	0	2	2	1[a]	1	-	Yes	Natural
WL pathway[b]	2	0	1	2	0	2	-	Yes	Natural
3HP bicycle	-1	3	7	5	-1	0	9	No	Natural
HP/HB cycle	0	2	6 or 4[c]	4	0	0	8 or 7	Yes or No[d]	Natural
DC/HB cycle	1	1	5	2	1[a]	1	-	Yes	Natural
CETCH 5.4	2	0	2	8	-4	0	8	No	Synthetic
rGPS	0	2	5	5	-1	0	8	No	Synthetic

Shown are the numbers of CO_2 and HCO_3^- fixed and the numbers of ATP, NAD(P)H, $FADH_2$ and Fd^{2-}_{red} molecules consumed (positive numbers) or generated (negative numbers) during the conversion of CO_2 equivalents into one molecule of pyruvate or acetyl-CoA. Every AMP-forming reaction was accounted for two ATP equivalents. For ATP per CO_2 conversions, a P/O

ratio of 2.5 for NAD(P)H and 1.5 for $FADH_2$ was assumed. [a] Fumarate reductase in the pathways uses $FADH_2$ as electron donors. [b] The variant of WL pathway functioning in acetogens. [c] Nine ATP molecules are required in Crenarchaeal HP/HB cycle and five ATP molecules are required in Thaumarchaeal HP/HB cycle. Succinyl-CoA is converted to succinate through succinyl-CoA synthetase. [d] The key enzyme, 4-hydroxybutyryl-CoA dehydratase (Hbd), is oxygen sensitive in clostridia. While, Hbd from *Metallosphaera sedula* and *Nitrosopumilus maritimus* are robust in the presence of oxygen (Half-life of *Metallosphaera sedula* Hbd in aerobic condition is about 4.1 days (31); half-life of *Nitrosopumilus maritimus* Hbd in aerobic condition is 46 hours (32)). CBB cycle, Calvin-Benson-Bassham cycle; rTCA cycle, reductive tricarboxylic acid cycle; WL pathway, Wood-Ljungdahl pathway; 3HP bicycle, 3-hydroxypropionate bicycle; HP/HB cycle, 3-hydroxypropionate/4-hydroxybutyrate cycle; DC/HB cycle, dicarboxylate/4-hydroxybutyrate cycle. Fd^{2-}_{red}, reduced ferredoxin.

Table 2.S3. Summary of specific activities of his-tag purified rGPS enzymes.

Name	Abbrev.	E.C number	Source	Specific Activity (U/mg)
acetyl-CoA acetyltransferase	AtoB	2.3.1.9	*Escherichia coli*	74.8
3-hydroxybutyryl-CoA dehydrogenase	Bhbd	1.1.1.35	*Clostridium acetobutylicum*	203.5
crotonase	Crt	4.2.1.150	*Clostridium acetobutylicum*	3089.3
crotonyl-CoA carboxylase/reductase	Ccr	1.3.1.85	*Methylobacterium extorquens*	58.2
ethylmalonyl-CoA/methylmalonyl-CoA epimerase	Epi	5.1.99.1	*Rhodobacter sphaeroides*	774.5
(2R)-ethylmalonyl-CoA mutase	Ecm	5.4.99.63	*Rhodobacter sphaeroides*	58.0
(2S)-methylsuccinyl-CoA dehydrogenase	Mcd	1.3.8.12	*Rhodobacter sphaeroides*	46.0
mesaconyl-C1-CoA-C4-CoA transferase	Mct	5.4.1.3	*Chloroflexus aurantiacus*	16.3
mesaconyl-C4-CoA hydratase	Meh	4.2.1.153	*Chloroflexus aurantiacus*	97.9
(*S*)-citramalyl-CoA lyase	Ccl	4.1.3.25	*Pseudomonas aeruginosa*	13.4
malyl coenzyme A lyase (working as ccl)	Mcl	4.1.3.25	*Methylobacterium extorquens*	9.17
phosphoenolpyruvate synthetase	Pps	2.7.9.2	*Escherichia coli*	5.1
phosphoenolpyruvate carboxylase	Ppc	4.1.1.31	*Corynebacterium glutamicum*	6.2
malate dehydrogenase	Mdh	1.1.1.37	*Escherichia coli*	493.2
malate thiokinase	Mtk	6.2.1.9	*Methylococcus capsulatus*	4.5
malyl coenzyme A lyase	Mcl	4.1.3.24	*Methylobacterium extorquens*	43.8

One unit of enzyme activity (U) is defined as conversion of one micromole of substrate per minute.

Table 2.S4. Plasmid list.

Plasmid	Description	Reference
pQE9	ColE1 ori; Amp^R; P_{T5}lacO::	Qiagen
pSL75	pQE9, Ccr (*M. e*)	This work
pSL69	pQE9, Ecm (*R. s*)	This work
pSL71	pQE9, Mcd (*R. s*)	This work
pSL72	pQE9, Mct (*C. a*)	This work
pSL78	pQE9, Meh (*C. a*)	This work
pSL120	pQE9, Mdh (*E. c)*	This work
pCDF-Duet1	CDF ori; $Spec^R$; P_{T7}lacO::	Novagen
pSL79	pCDF-Duet1, Epi (*R. s*)	This work
pSL96	pCDF-Duet1, Ccl (*P. a*)	This work
pET-pps	ColE1 ori; Amp^R; P_{T7}lacO:: Pps (*E. c*)	This work
pET-mcl	ColE1 ori; Amp^R; P_{T7}lacO:: Mcl (*M. e*)	This work
pYK-mtk	ColA ori; Kan^R; P_{T7}lacO:: MtkAB (*M. c*)	This work
pTW-atoB	CloDF13 ori; $Spec^R$; P_{T7}lacO:: AtoB (*E. c*)	Gift form Tony Wu
pTW-bhbd	CloDF13 ori; $Spec^R$; P_{T7}lacO:: Bhbd (*C. a*)	Gift form Tony Wu
pTW-crt	CloDF13 ori; $Spec^R$; P_{T7}lacO:: Crt (*C. a*)	Gift form Tony Wu
pET-28a-ppc	F1 ori; Kan^R; P_{T7}lacO:: Ppc (*C. g*)	Gift from Fabienne Duchoud

Table 2.S5. Primer sequences

Construct	Primer names	Primer sequences	Template
pQE9 backbone	pQE9 f	TCTAGAGGCATCAAATAAAACGAAAGGCT	
	pQE9 r	TCCCTGAAAATACAGGTTTTCGGATCC	
pSL75	ccr-Me f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGGCTGCAAGCGCAGCA	*M. extorquens*
	ccr-Me r	GAGCCTTTCGTTTTATTTGATGCCTCTAGATCACATCGCCTTGAGCGGG	
pSL69	ecm-Rs f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGACCCAGAAGGATAGCCCC	*R. sphaeroides*
	ecm-Rs r	GAGCCTTTCGTTTTATTTGATGCCTCTAGACTATTCCGCCGCGAGGGC	
pSL71	mcd-Rs f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGCCGAAGGACGGATCGA	*R. sphaeroides*
	mcd-Rs f	GAGCCTTTCGTTTTATTTGATGCCTCTAGATCAGTCCAGCAGCCGCC	
pSL72	mct-Ca f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGAAGGGTATTCTCCACGGATTGC	*C. acetobutylicum*
	mct-Ca r	GAGCCTTTCGTTTTATTTGATGCCTCTAGACTACGCTGCCCGATCTGGC	
pSL78	meh-Ca f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGAGCAGCGCGGATTGGA	*C. acetobutylicum*
	meh-Ca r	GAGCCTTTCGTTTTATTTGATGCCTCTAGATCATCCCCCCAACTCAACCG	
pSL120	mdh-Ec f	CACGGATCCGAAAACCTGTATTTTCAGGGAATGAAAGTCGCAGTCCTCGGC	*E. coli*
	mdh-Ec r	GAGCCTTTCGTTTTATTTGATGCCTCTAGATTACTTATTAACGAACTCTTCGCCCAGG	
pCDF-Duet1 backbone	pCDE f	TAAGCTTGCGGCCGCATAATG	
	pCDE r	CATCGAATTCGGATCCTGGCTGT	
pSL79	epi-Rs f	CATCATCACCACAGCCAGGATCCGAATTCGATGATCGGACGCCTGAACCA	*R. sphaeroides*
	epi-Rs r	CGACTTAAGCATTATGCGGCCGCAAGCTTATTATACCTGCTCGAGCTCCACG	
pSL96	ccl-Pa f	CATCATCACCACAGCCAGGATCCGAATTCGATGAACCGACAGATCGTGCG	*P. aeruginosa*
	ccl-Pa r	CGACTTAAGCATTATGCGGCCGCAAGCTTATCAGCCACCCTCCCCGG	
pET backbone	pET f	CCTGGCTGAACTGAAGAAATAATCTACTAGCGCAGCTTAATTA	
	pET r	GGTGACGAGCCATTGTTGGACATCGCCGAGCTCGAATTCGGAT	
pET-pps	pps-Ec f	GGATCCGAATTCGAGCTCGGCGATGTCCAACAATGGCTCGTCA	*E. coli*
	pps-Ec r	GTTAATTAAGCTGCGCTAGTAGATTATTTCTTCAGTTCAGCCA	
pET-mcl	mcl-Me f	CGAATTCGAGCTCGGCGATGAGCTTCACCCTGATCC	*M. extorquens*
	mcl-Me r	TAAGCTGCGCTAGTAGATTACTTTCCGCCCATCGCGT	
pYK backbone	pYK f	ATCATCACCACAGCCAGGATCCGTGACTCGAGAAGCTTGAT	
	pYK r	CATGGATATTCATGGTATATCTCCTTCTTAA	
pYK-mtk	mtkA-Mc f	GGAGATATACCATGAATATCCATGAGTACCA	*M. capsulatus*
	mtkA-Mc r	GTTTTCGGATCCGTGGTGATGATGGTGATGGCTGCTTCCCTTGACGATGGCGA	
	mtkB-Mc f	CATCATCACCACGGATCCGAAAACTAAAGGAGATATACCATGAGCGTATT	
	mtkB-Mc r	CACGGATCCTGGCTGTGGTGATGATGGTGATGGCTGCTGAATCTGATTCCGTGTT	

2.6 Reference

1. IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp. Available at: https://www.ipcc.ch/report/ar5/syr/.

2. O'Neill BC, Oppenheimer M (2002) Dangerous Climate Impacts and the Kyoto Protocol. *Science* 296(5575):1971 LP-1972.

3. Sky Scenario | Shell Global Available at: https://www.shell.com/energy-and-innovation/the-energy-future/scenarios/shell-scenario-sky.html.

4. Liao JC, Mi L, Pontrelli S, Luo S (2016) Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat Rev Microbiol* 14(5):288–304.

5. Fuchs G (2011) *Alternative Pathways of Carbon Dioxide Fixation: Insights into the Early Evolution of Life?* doi:10.1146/annurev-micro-090110-102801.

6. Erb TJ (2011) Carboxylases in natural and synthetic microbial pathways. *Appl Environ Microbiol* 77(24):8466–8477.

7. Erb TJ, Zarzycki J (2018) A short history of RubisCO: the rise and fall (?) of Nature's predominant CO2fixing enzyme. *Curr Opin Biotechnol* 49:100–107.

8. Tcherkez GGB, Farquhar GD, Andrews TJ (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc Natl Acad Sci U S A* 103(19):7246–51.

9. Savir Y, Noor E, Milo R, Tlusty T (2010) Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proc Natl Acad Sci U S A* 107(8):3475–3480.

10. Chabriere E, et al. (1999) Crystal structures of the key anaerobic enzyme pyruvate:ferredoxin oxidoreductase, free and in complex with pyruvate. *Nat Struct Biol* 6(2):182–190.

11. Cronan JEJ, Waldrop GL (2002) Multi-subunit acetyl-CoA carboxylases. *Prog Lipid Res* 41(5):407–435.

12. Schwander T, Von Borzyskowski LS, Burgener S, Cortina NS, Erb TJ (2016) A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* 354(6314):900–904.

13. Bar-Even A, Noor E, Lewis NE, Milo R (2010) Design and analysis of synthetic carbon fixation pathways. *Proc Natl Acad Sci* 107(19):8889–8894.

14. Mainguet SE, Gronenberg LS, Wong SS, Liao JC (2013) A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab Eng* 19:116–27.

15. Yu H, Li X, Duchoud F, Chuang DS, Liao JC (2018) Augmenting the calvin-benson-bassham cycle by a synthetic malyl-CoA-glycerate carbon fixation pathway. *Nat Commun* 9(1):1–10.

16. Erb TJ, et al. (2007) Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proc Natl Acad Sci* 104(25):10631–10636.

17. Zarzycki J, Brecht V, Müller M, Fuchs G (2009) Identifying the missing steps of the autotrophic 3-hydroxypropionate CO2 fixation cycle in Chloroflexus aurantiacus. *Proc Natl Acad Sci U S A* 106(50):21317–22.

18. Kwok R (2010) Five hard truths for synthetic biology. *Nature* 463(7279):288–290.

19. Swartz JR (2012) Transforming biochemical engineering with cell-free biology. *AIChE J* 58(1):5–13.

20. Rollin JA, Tam TK, Zhang YHP (2013) New biotechnology paradigm: Cell-free biosystems for biomanufacturing. *Green Chem* 15(7):1708–1719.

21. Wilding KM, Schinn SM, Long EA, Bundy BC (2018) The emerging impact of cell-free chemical biosynthesis. *Curr Opin Biotechnol* 53(1):115–121.

22. Bujara M, Schümperli M, Pellaux R, Heinemann M, Panke S (2011) Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. *Nat Chem Biol* 7(5):271–277.

23. Karim AS, Heggestad JT, Crowe SA, Jewett MC (2018) Controlling cell-free metabolism through physiochemical perturbations. *Metab Eng* 45(November 2017):86–94.

24. Bogorad IW, et al. (2014) Building carbon-carbon bonds using a biocatalytic methanol condensation cycle. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1413470111.

25. Opgenorth PH, Korman TP, Bowie JU (2016) A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nat Chem Biol* 12(6):393–395.

26. Opgenorth PH, Korman TP, Iancu L, Bowie JU (2017) Amolecular rheostat maintains ATP levels to drive a synthetic biochemistry system. *Nat Chem Biol* 13(9):938–942.

27. Bogorad IW, Lin T-S, Liao JC (2013) Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* 502(7473):693–7.

28. Berg IA (2011) Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl Environ Microbiol* 77(6):1925–1936.

29. Erb TJ, Fuchs G, Alber BE (2009) (2 *S*)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol Microbiol* 73(6):992–1008.

30. Flamholz A, Noor E, Bar-Even A, Milo R (2012) EQuilibrator - The biochemical thermodynamics calculator. *Nucleic Acids Res* 40(D1):770–775.

31. Hawkins AB, Adams MWW, Kelly RM (2014) Conversion of 4-hydroxybutyrate to acetyl coenzyme A and its anapleurosis in the Metallosphaera sedula 3-hydroxypropionate/4-hydroxybutyrate carbon fixation pathway. *Appl Environ Microbiol* 80(8):2536–2545.

32. Konneke M, et al. (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc Natl Acad Sci* 111(22):8239–8244.

Chapter 3. Continuous operation of the synthetic CO_2 fixation cycle rGPS *in vitro*

3.1 Introduction

In chapter 2, we already demonstrated the feasibility of the synthetic CO_2 fixation cycle rGPS *in vitro*. Using a two-pot immobilized enzyme system, we were able to run the rGPS cycle semi-continuously. However, this two-pot system inevitably compromised the efficiency of the whole cycle since it operated rGPS in a non-continuous way. To achieve better efficiency of the *in vitro* rGPS system, we need to run this cycle continuously in one-pot.

One big issue for the *in vitro* rGPS is the instability of the intermediate methylsuccinyl-CoA. This means that its generation and consumption flux needs to be balanced so that methylsuccinyl-CoA is consumed as soon as it is generated. Otherwise, the instability of methylsuccinyl-CoA may deplete the intermediate pool and interrupt the rGPS cycle. Methylsuccinyl-CoA is the substrate of Mcd, a FAD dependent dehydrogenase (1). To avoid the degradation of methylsuccinyl-CoA, the FAD should be regenerated efficiently to avoid the accumulation of this unstable intermediate. In cells, electrons of reduced FAD of Mcd eventually transfer to the respiratory chain through electron transfer flavoproteins (1). In the cell-free system, ferrocenium can act as an excellent artificial electron acceptor for reduced FAD (2). However, ferrocenium can also accept electrons from NADH and NADPH (3). Engineering Mcd to directly use molecular oxygen as electron acceptor has been made to avoid using ferrocenium *in vitro*. However, after engineering, the activity dropped significantly (~60 folds less compared to Mcd using ferrocenium hexafluorophosphate ($FePF_6$) as the electron acceptor) (4). Actually, the engineered Mcd became the slowest enzyme in the synthetic CETCH cycle, substantially

lowered the efficiency of the whole cycle. Therefore, we continued using ferrrocenium as the artificial electron acceptor for Mcd in our system.

As mentioned in chapter 2, the reaction rate of oxidizing NADH and NADPH with ferrocenium mainly depends on their concentrations. Thus recycling ferrocenium, NAD(P)H to keep their concentrations low can make this spontaneous side reaction become negligible. Recycling NADH and NADPH is ubiquitous in building *in vitro* systems. However, the method of recycling ferrocenium, to our knowledge, has not been well developed.

3.2 Results

3.2.1 Characterization of the working concentration range of ferrocenium and NAD(P)H

To determine the working concentration range of ferrocenium and NAD(P)H, we first measured the reaction rate of oxidizing NADH with ferrocenium at different concentrations (Fig. 3.1). This spontaneous reaction could be negligible when the reaction rate is below 0.01 $OD_{300}\,min^{-1}$. Thus, we can easily set the upper concentration limits of ferrocenium and NAD(P)H from the heatmap we created (Fig. 3.1).

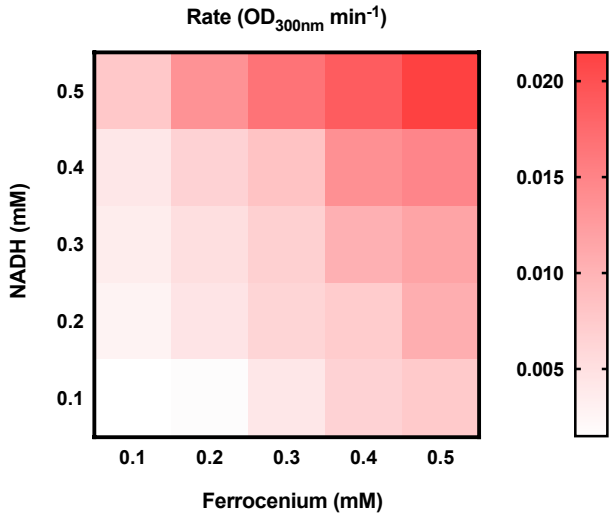


Fig. 3.1. Heatmap of ferrocenium oxidizing NADH.

The reaction rate was measured at 300 nm (both NADH and Ferrocenium have absorbance at 300 nM). The reaction rate below 0.01 OD_{300} min^{-1} is considered as the desired range.

To ensure that Mcd can preform at higher activity, we also characterized the apparent Km of Mcd for ferrocenium (Fig 3.2). The resulting Km value for ferrocenium is about 0.06 mM, which is located inside the desired range on the heatmap. Therefore, we can set the concentration of ferrocenium below 0.2 mM and NAD(P)H concentration below 0.4 mM to keep the spontaneous side reaction minimized as well as have Mcd performed at higher activity.

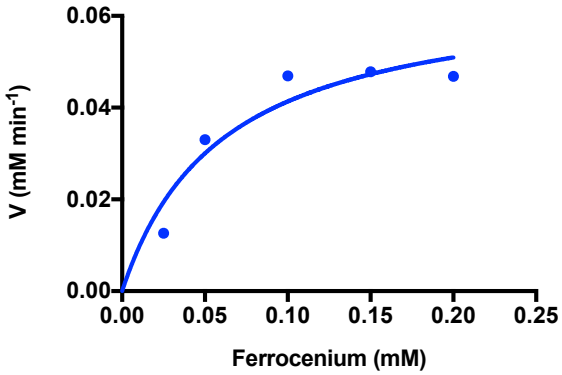


Fig. 3.2. Michaelis-Menten graph of Mcd using ferrocenium as the electron acceptor.

Km for ferrocenium is about 0.06 mM.

3.2.2 Continuous running of the rGPS cycle in vitro using electrochemical FAD regeneration system

To recycle ferrocenium, we first noticed that ferrocenium and its reduced form ferrocene were widely used in biosensors as electron mediators (5). For example, in some glucose biosensors, ferrocenium is reduced to ferrocene by accepting electrons from the reduced FAD in glucose oxidase (GO) (6). Then, ferrocene is oxidized back to ferrocenium by transferring electrons to the anode. Thus, the redox pair ferocene/ferrocenium mediates electron transfer from GO to the electrode under a positive potential. Theoretically, we can use electricity to recycle $FePF_6$. However, ferrocene has very low solubility in water. Since the oxidation rate of ferrocene on the electrode depends on its local concentration on the surface, the low concentration of ferrocene in the aqueous solution definitely limits its oxidation rate. Immobilizing ferrocene and enzymes on the electrode surface is a solution. However, immobilizing enzymes on the electrode may decrease their activity and localizing one reaction to the surface will bring more complexity to our multi-enzyme system. Therefore, we employed a more water-soluble ferrocene derivative, ferrocenylmethanol, to obtain a better oxidation rate in the designed electrochemical system (6). Under +0.4 V potential, ferrocenylmethanol is oxidized and serves as the electron acceptor of Mcd, while NADH and NADPH remain oxidized under this condition. A three-electrode system was integrated with the enzymatic system (Fig. 3.3C). This electricity-integrated system successfully fixed CO_2 at a rate of 37.3 nmol min^{-1} mg^{-1} of core cycle proteins (Fig. 3.3E).

3.2.3 Continuous running of the rGPS cycle *in vitro* using enzymatic FAD regeneration system

We further tried to recycle ferrocenium using horseradish peroxidase (HRP) (Fig. 3.3D), which could utilize H_2O_2 to oxidize ferrocene. The HRP is a heme-containing enzyme that is capable of oxidizing a wide variety of organic and inorganic compounds using H_2O_2 (7). Although HRP can also oxidize NADH and NADPH with H_2O_2, the activity is more than 1000 times lower than its activity with ferrocene (Fig. 3.S1). By controlling the peroxidase amount in our system, the electron acceptor can be regenerated with HRP and H_2O_2 while NAD(P)H will be barely influenced (Fig. 3.S2). To avoid enzyme damage by excess amounts of H_2O_2, H_2O_2 was fed to the system based on the ferrocenium concentration monitored by the absorbance at 300 nm. NADH and NADPH were recycled using Glucose 6-phosphate dehydrogenase (G6PDH) and remained constant in the reaction mixture (Fig. 3.3B). Using this system, the CO_2 fixation rate of the rGPS cycle increased to 66.9 nmol min^{-1} mg^{-1} of core cycle proteins (Fig. 3.3F).

This result is comparable to the *in vivo* carbon fixation rate of the CBB cycle in plant leaves (30-300 nmol CO_2 min^{-1} mg^{-1} CBB protein, assuming that CBB cycle proteins account for 10% of leaf dry weight) (8) and is about 10 times higher than the previously reported synthetic CETCH cycle *in vitro* (5 nmol min^{-1} mg^{-1} core cycle protein) (4) and the reported CBB cycle carbon fixation rate in cell extracts (1 to 3 nmol min^{-1} mg^{-1} CBB cycle protein, assuming that CBB cycle enzymes account for 30% of the total proteins in cell extracts) (9).

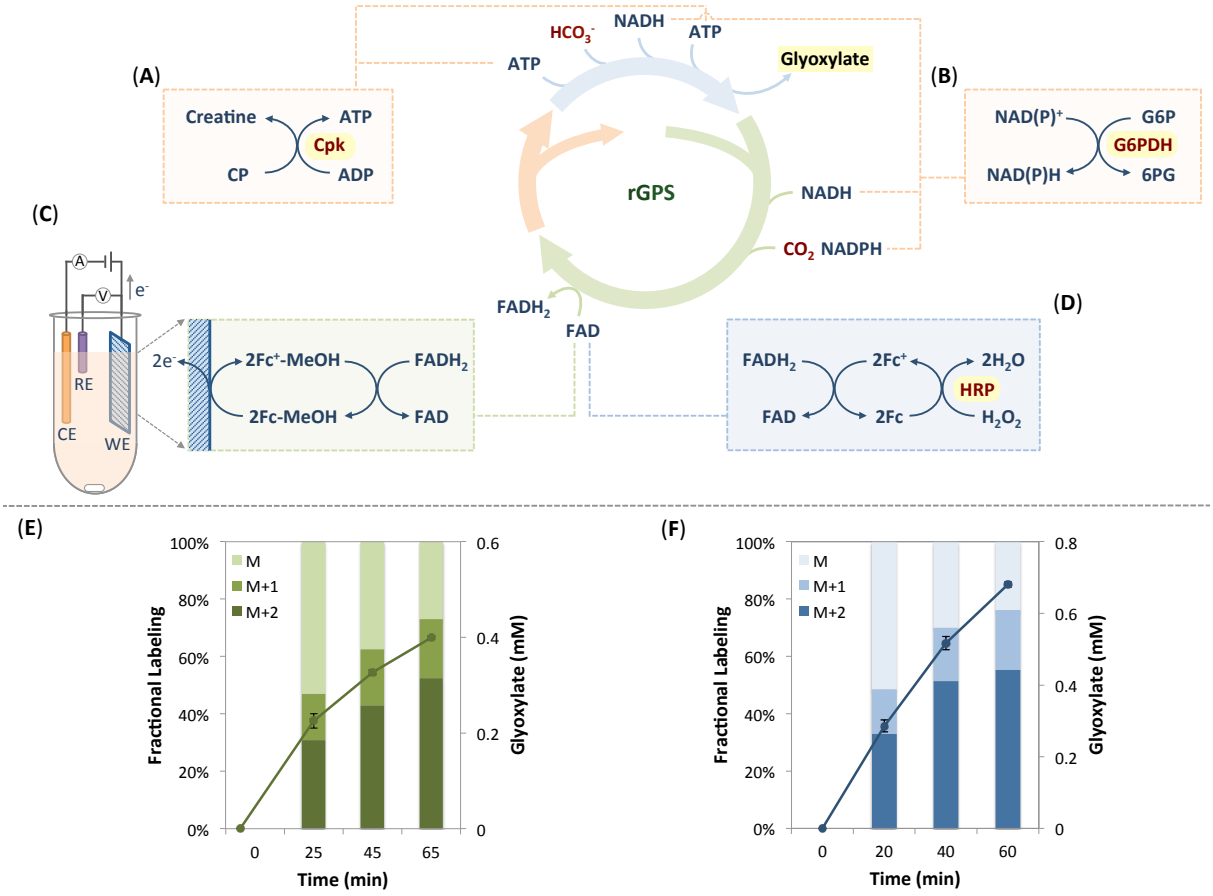


Fig. 3.3. One-pot continuous operation of the rGPS cycle.

(A) The ATP regeneration scheme. **(B)** The NADH/NADPH regeneration scheme. **(C)** The electrochemical FAD regeneration system. **(D)** The enzymatic FAD regeneration system (see methods for details). **(E)** Total glyoxylate produced (right y axis) and its labeling patterns (left y axis) over 65 min of operating rGPS using the electrochemical FAD regeneration (C). **(F)** Total glyoxylate produced (right y axis) and its labeling patterns (left y axis) over 60 min of operating rGPS using the enzymatic FAD regeneration (D). CE, counter electrode; RE, reference electrode; WE, working electrode; CP, creatine phosphate; CPK, creatine phosphokinase; G6PDH, glucose-6-phosphate dehydrogenase; HRP, horseradish peroxidase; Fc-MeOH, ferrocene methanol; Fc^+-MeOH, ferrocenium methanol; Fc, ferrocene; Fc^+, ferrocenium hexafluorophosphate ($FePF_6$).

3.3 Discussion

Despite a few successful cases, constructing synthetic pathways or implanting natural CO_2 fixation pathways in non-native organisms is still a daunting test. The success of building rGPS *in vitro* shows that the cell-free approach can be a powerful tool in synthetic biology. Using the cell free system, we can easily assay components to find potential problems.

The main problem for the cell-free rGPS system is how to operate the oxidation and reduction reactions in rGPS cycle simultaneously to avoid the accumulation unstable intermediate, methylsuccinyl-CoA. The open reaction environment enabled us develop electrochemical and enzymatic methods for FAD regeneration. Using the enzymatic FAD regeneration method even pushed the CO_2 fixation rate to 66.9 nmol min^{-1} mg^{-1}, which is more than 10 folds higher than the performance of *in vitro* CBB cycle. The FAD regeneration methods can be easily applied to building cell-free systems of various metabolic pathways, such as several proposed synthetic CO_2 fixation pathways and fatty acids synthesis.

Using the cell-free system, we showed that the carbon fixation rate of the unevolved rGPS cycle can in fact be comparable to the *in vivo* CBB cycle based on the same amount of protein (8). On the basis of measured specific activity of each enzyme, the rGPS cycle could be twice as efficient as the CBB cycle for pyruvate and acetyl-CoA synthesis (Table 3.S1), suggesting that this synthetic pathway has a high potential to enhance the natural carbon fixation.

The success of the rGPS cycle in cell free systems paves the way for *in vivo* implementations, including enhancing CO_2 fixation in photosynthetic or lithoautotrophic organisms, converting heterotrophic organisms to autotrophs (10, 11). Integrating synthetic pathways in cellular metabolism may rely largely on evolution to tune the cells to use these non-native pathways (10,

12). rGPS provides pyruvate and acetyl-CoA, which in turn provide all components for cell growth. As such, rGPS is evolvable *in vivo*. The rGS segment of the rGPS cycle has already been demonstrated *in vivo* (13, 14). The rPS segment is expected to be compatible with cellular metabolism, provided that Mcd, electron transfer flavoproteins, and the respiratory chain are properly coordinated. Ultimately, using cell-free systems for fixing CO_2 may become feasible, if proteins can be sufficiently stabilized, and cofactors can be regenerated using renewable energy at a deployable scale and sufficiently low cost.

3.4 Materials and methods

3.4.1 Materials

All chemicals used were purchased from Sigma-Aldrich unless otherwise specified.

3.4.2 Enzyme Assays

All assays were conducted at room temperature under aerobic conditions. The reaction mixture volume was 0.6 ml unless otherwise indicated.

HRP assays. The HRP activity of oxidizing NADH was measured by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADH. The reaction mixture contained 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, 0.1 mM NADH, and 126 μg HRP. The HRP activity of oxidizing Fc (ferrocene) was measured by monitoring the increase of absorbance at 300 nm corresponding to the production of Fc^+ ($FePF_6$). The estimated extinction coefficient $\varepsilon_{300\ nm}$ of $FePF_6$ was 4.3 $mM^{-1}\ cm^{-1}$ (1). The reaction mixture contained 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, 0.1 mM Fc (produced from reducing $FePF_6$ with 1,4-dithiothreitol), and 6.3 μg HRP.

3.4.3 HPLC Analysis

CoA compounds and glyoxylate phenylhydrazone were analyzed by an Agilent 1200 HPLC equipped with a reversed-phase C18 column (LiChrospher 100, end-capped, 5 μm, LiChroCART 125-4 125 × 4 mm). Separation was performed using one of the two gradient systems of 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent A) and 20% acetonitrile (MeCN) in 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent B). Column temperature was maintained at 25 ^{o}C and the flow rate stayed constant at 1 ml/min. Initial conditions were 90% solvent A and 10% solvent B. At 0 min, solution B was increased linearly to 50% over 40 min. At 40 min, the composition of B was increased to 100% over 1 min and held isocratic for 2 min. At 43 min, solution B was decreased to its initial 10% condition over 1 min and held for 16 min to reequilibrate the column.

3.4.4 LC-MS analysis

The labeling pattern of glyoxylate phenylhydrazone was analyzed by a Shimadzu 2020 EV LC-MS equipped with a reversed-phase C18 column (Phenomenex Kinetex, 1.7 μm, 100 Å, 100 × 2.1mm) using positive-and negative-mode electrospray ionization with a linear gradient of 5-95% MeCN-H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 3min with a constant flow rate of 0.3 ml min^{-1}.

3.4.5 Assay of continuous running of the rGPS cycle using electrochemical FAD regeneration system

The assay was first performed in a 600 μL reaction mixture containing 100 mM MOPS pH 7.0, 16.7 mM $NaH^{13}CO_3$, 5 mM $MgCl_2$, 4 mM glucose 6-phosphate, 3 mM ATP, 9 mM creatine phosphate, 0.17 mM NADH, 0.17 mM NADPH and all the enzymes except Epi. Enzyme

amounts are listed in Table 3.S2. The reaction was started with the addition of 0.2 mM crotonyl-CoA. After 3 min, Epi and 0.6 mM ferrocenium methanol were added into the reaction mixture. 2 min later, 500 μL of the resulting mixture was added into the electrochemical system described below for 20 min, 40 min, or 60 min at 0.4 V. The rest 100 μL mixture reacted without electricity as the control (results in Fig. 3.S3). The samples were quenched with 5% formic acid, derivatized with 10 mM phenylhydrazine and analyzed with HPLC and LC-MS.

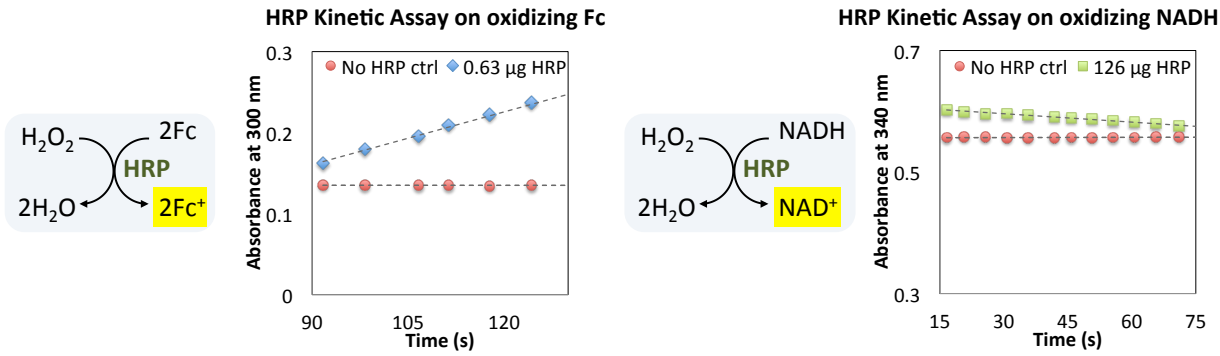
3.4.6 Setup of the electrochemical system

Electrochemical experiments were performed using a Versatile Multichannel Potentiostat (model VMP3) equipped with the 'p' low current option and N'Stat box driven by EC-LAB software (Bio-Logic USA, LLC, Knoxville, TN) in a three electrode configuration consisting of a Pt working electrode, a Pt wire auxiliary electrode, and a Ag/AgCl glass-bodied reference electrode. With a 4-inch silicon wafer (150 μm in thickness), electron beam evaporation was used to deposit 1000 Å of platinum on a 200 Å chromium adhesion layer to get Pt working electrodes. After cutting to a smaller size and soldering with a wire, the surface of Pt working electrode was coated with a nanostructured Pt deposit to increase the surface area as described by Boehler et al. (aqueous solution of 2.5 mM H_2PtCl_6 and 1.5 mM formic acid, -0.1 V vs. Ag/AgCl for 5 min) (15). The Pt working electrode had ~0.78 cm^2 (0.6 cm*1.3cm) surface area immersed into 500 μL reaction mixture. The reactions were conducted in 500 μL stirring solution at 0.4 V[33] vs. Ag/AgCl for 20 min, 40 min, and 60 min with the amperometric currents stabilized at about 35 ± 6.3 μA (n = 7). The ferrocenium methanol was produced by oxidizing 10 mL ferrocene methanol stirring solution at 0.4 V vs. Ag/AgCl for 60 min. Ag/AgCl glass-bodied reference electrodes with 3 M NaCl electrolyte and 0.5-mm-diameter Pt wire auxiliary electrodes were purchased from BASi (West Lafayette, IN).

3.4.7 Assay of continuous running of the rGPS cycle using enzymatic FAD regeneration system

The assay was performed in a 600 μL (final volume) reaction mixture containing 100 mM MOPS pH 7.0, 16.7 mM $NaH^{13}CO_3$, 5 mM $MgCl_2$, 4 mM glucose 6-phosphate, 3 mM ATP, 9 mM creatine phosphate, 0.17 mM NADH, 0.17 mM NADPH, 0.17 mM $FePF_6$, and all the enzymes with the corresponding amounts (Table 3.S2). Note that the concentrations of NAD(P)H and $FePF_6$ were kept low to avoid the spontaneous reaction of $FePF_6$ oxidizing NAD(P)H. The reaction was started with the addition of 0.2 mM crotonyl-CoA. The absorbance at 300 nm of the reaction mixture was monitored. 1 μL 12.5 mM H_2O_2 was added into the reaction mixture when the absorbance at 300 nm decreased to the lowest point. For 600 μL reaction, about 30 μL 12.5 mM H_2O_2 was added in total. Samples (100 μl each) were withdrawn from the reaction mixture at 20 min, 40 min, and 60 min. The samples were quenched with 5% formic acid, derivatized with 10 mM phenylhydrazine and analyzed with HPLC and LC-MS.

3.5 Supplementary figures and tables



Reaction	H_2O_2 + 2 Fc → 2 H_2O+ 2 Fc^+	H_2O_2 + NADH → 2 H_2O+ NAD^+
Specific Activity of HRP (U/mg)	**29.2**	**0.027**

Fig. 3.S1. *In vitro* enzyme assay schemes of HRP and its activities.

HRP, horseradish peroxidase; Fc, ferrocene; Fc^+, ferrocenium hexafluorophosphate ($FePF_6$).

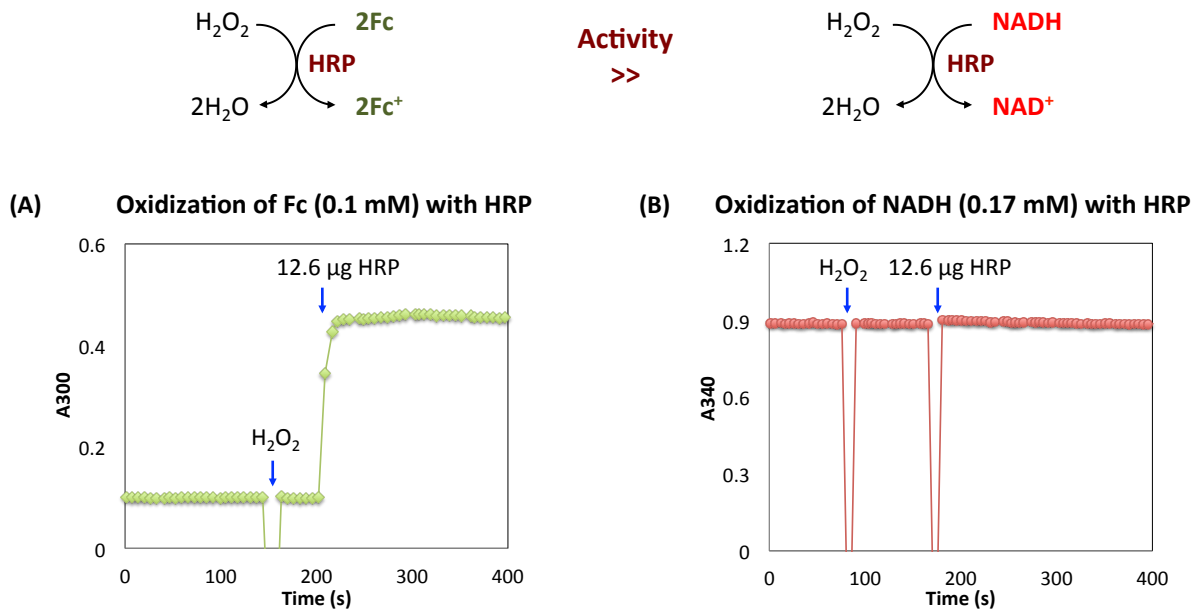


Fig. 3.S2. Controlling the HRP amount in the *in vitro* system enabled efficiently regenerating Fc^+ with HRP and H_2O_2 while keeping NADH barely influenced.

Because the activity of HRP to NAD(P)H is almost 1000 times lower than its activity to Fc. **(A)** Adding 0.1 mM H_2O_2 and 12.6 µg HRP sequentially into the reaction buffer, which contained 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, and 0.1 mM Fc, caused immediate Fc to Fc^+ conversion. **(B)** While, adding 0.1 mM H_2O_2 and 12.6 µg HRP sequentially into the reaction buffer, which contained 100 mM MOPS pH 7.0, 16.7 mM $NaHCO_3$, 5 mM $MgCl_2$, and 0.17 mM NADH, did not influence the NADH concentration. Fc^+ was followed at 300 nm, NADH was followed at 340 nm. HRP, horseradish peroxidase; Fc, ferrocene; Fc^+, ferrocenium hexafluorophosphate ($FePF_6$).

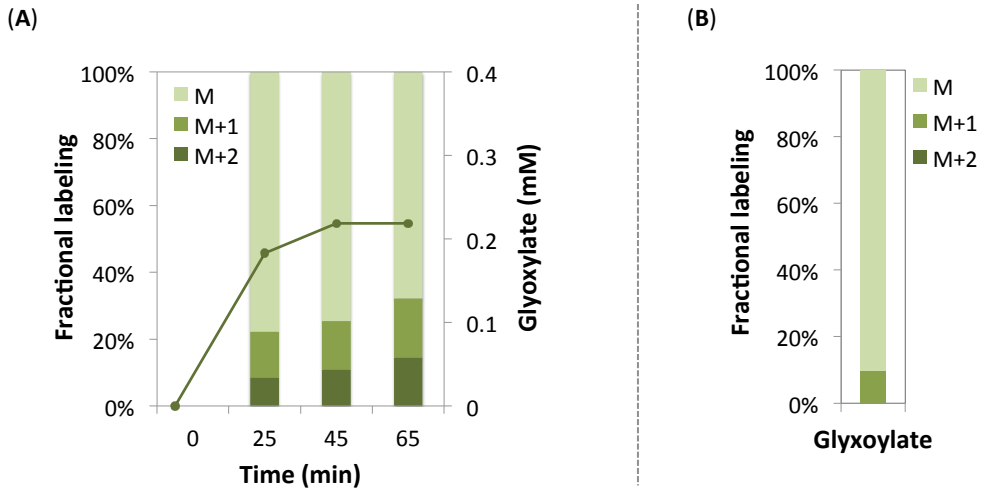


Fig. 3.S3. Controls for the continuous running of rGPS in vitro using electricity.

(A) Total glyoxylate produced (right y axis) and its labeling patterns (left y axis) over 65 min of operating rGPS *in vitro* without electricity (see "assay of continuous running of the rGPS cycle using electricity" in supplementary methods for details). **(B)** The labeling pattern of natural glyoxylate.

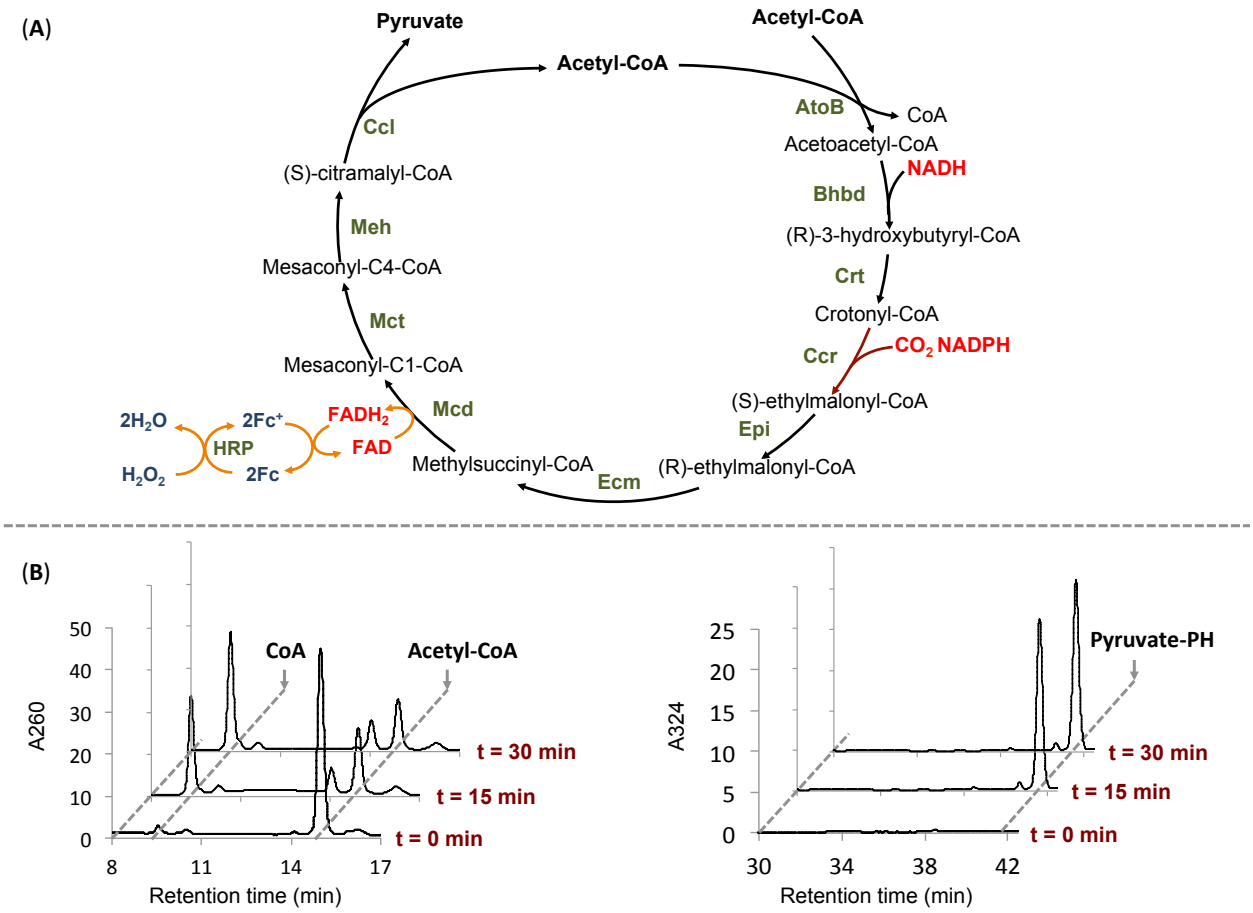


Fig. 3.S4. *In vitro* demonstration of conversion of acetyl-CoA to pyruvate with the rPS part of rGPS.

(A) Scheme of converting acetyl-CoA to pyruvate in 10 steps, using HRP and H_2O_2 for ferrocenium regeneration. (B) HPLC analysis of the products of incubating the reaction mixture containing acetyl-CoA with all enzymes and cofactors for 0 min, 15 min, and 30 min. CoA esters were followed at 260 nm, pyruvate was derivatized by phenylhydrazine and detected at 324 nm. HRP, horseradish peroxidase; Fc, ferrocene; Fc^+, ferrocenium hexafluorophosphate ($FePF_6$); PH, phenylhydrazone).

Table 3.S1. Calculations of theoretical carbon fixation rate of rGPS and CBB.

Summary of theoretical carbon fixation rate

Carbon fixation cycle	Theoretical carbon fixation rate (μmol CO_2/min/mg protein)	
	Pyruvate as the fixation product	Acetyl-CoA as the fixation product
rGPS	1.93	1.50
CBB	1.14	0.73

Summary of enzyme activities for calculating the theoretical carbon fixation rate of rGPS

Name	Specific Activity (U/mg)	Enzyme amounts for producing (μg)	
		Pyruvate at 1 μmol/min	Acetyl-CoA at 1 μmol/min
acetyl-CoA acetyltransferase	74.8[a]	26.7	13.4
3-hydroxybutyryl-CoA dehydrogenase	203.5[a]	9.8	4.9
crotonase	3089.3[a]	0.6	0.32
crotonyl-CoA carboxylase/reductase	58.2[a]	34.4	17.2
ethylmalonyl-CoA/methylmalonyl-CoA epimerase	774.5[a]	2.6	1.29
(2R)-ethylmalonyl-CoA mutase	58.0[a]	34.5	17.2
(2S)-methylsuccinyl-CoA dehydrogenase	46.0[a]	43.5	21.7
mesaconyl-C1-CoA-C4-CoA transferase	16.3[a]	122.7	61.3
mesaconyl-C4-CoA hydratase	97.9[a]	20.4	10.2
(*S*)-citramalyl-CoA lyase	13.4[a]	149.3	74.6
phosphoenolpyruvate synthetase	5.1[a]	196.1	196.1
phosphoenolpyruvate carboxylase	6.2[a]	322.6	322.6
malate dehydrogenase	493.2[a]	4.1	4.1
malate thiokinase	4.5[a]	444.4	444.4
malyl coenzyme A lyase	43.8[a]	45.7	45.7
glyoxylate carboligase	16.5[b]	60.6	60.6
tartronate semialdehyde reductase	107[b]	9.3	9.3
glycerate 2-kinase	181[c]	5.5	5.5
enolase	43[b]	23.3	23.3

Summary of enzyme activities for calculating the theoretical carbon fixation rate of CBB

Name	Specific Activity (U/mg)	Enzyme amounts for producing (μg)	
		Pyruvate at 1 μmol/min	Acetyl-CoA at 1 μmol/min
RuBisCo	1.3[b]	2307.7	2307.7
phosphoglycerate kinase	760[b]	6.6	6.6
glyceraldehyde-3P dehydrogenase	145[b]	34.5	34.5
triose-phosphate isomerase	9030[b]	0.1	0.1
fructose-bisphosphate aldolase	29.5[b]	33.9	33.9
fructose-bisphosphatase	68[b]	14.7	14.7
transketolase	58.9[b]	17.0	17.0
fructose-bisphosphate aldolase	11.8[b]	84.7	84.7
sedoheptulose-bisphosphatase	25.2[b]	39.7	39.7
transketolase	22.4[b]	44.6	44.6
ribose-5-phosphate isomerase	1850[b]	0.5	0.5
ribulose-phosphate 3-epimerase	1200[b]	1.7	1.7
phosphoribulokinase	335[b]	9.0	9.0
phosphoglycerate mutase	415[b]	2.4	2.4
enolase	43[b]	23.3	23.3
pyruvate kinase	360[d]	2.8	2.8
pyruvate dehydrogenase	7.4[e]	0	135.1

[a] from This work. [b] from Bar-Even A., et al. 2010 (16). [c] from Yoshida T., et al. 1992 (17). [d] Aust A., et al. 1975 (18). [e] from Hirabayashi T., et al. 1972 (19).

Table 3.S2. List of enzyme amounts in rGPS assays.

Abbrev.	Source	Concentration (μg ml^{-1})	
		in system with electrochemical FAD regeneration	in system with enzymatic FAD regeneration
AtoB	*Escherichia coli*	29.7	29.7
Bhbd	*Clostridium acetobutylicum*	6.5	6.5
Crt	*Clostridium acetobutylicum*	5.4	5.4
Ccr	*Methylobacterium extorquens*	35.9	35.9
Epi	*Rhodobacter sphaeroides*	10.9	11.1
Ecm	*Rhodobacter sphaeroides*	32.1	36
Mcd	*Rhodobacter sphaeroides*	43.0	34.3
Mct	*Chloroflexus aurantiacus*	14.3	7.3
Meh	*Chloroflexus aurantiacus*	2.9	2.3
Ccl	*Pseudomonas aeruginosa*	0	0
Pps	*Escherichia coli*	31.9	33.8
Ppc	Microbial (Sigma)	33.3	33.3
Mdh	Porcine heart (Sigma)	4.4	4.4
Mtk	*Methylococcus capsulatus*	70.5	75.2
Mcl	*Rhodobacter sphaeroides*	18.4	16.0
G6PDH	*Leuconostoc mesenteroides* (Sigma)	15.7	15.7
CPK	Rabbit muscle (Sigma)	22.2	22.2
HRP	Horseradish roots (Sigma)	0	21

Table 3.S3. Comparison of the in vitro carbon fixation results of rGPS and CETCH 5.4.

Name	CETCH 5.4	rGPS	
		with enzymatic FAD regeneration	with electrochemical FAD regeneration
Final glyoxylate concentration	0.54 mM	0.68 mM	0.4 mM
Starting material	0.2 mM propionyl-CoA	0.2 mM crotonyl-CoA	0.2 mM crotonyl-CoA
Reaction volume	520 µl	600 µl	500 µl
Core enzyme amount	1196 µg	203 µg	165 µg
Reaction time	90 min	60 min	65 min
Cycle turnover number	2.7	3.4	2
CO_2 equivalents per acceptor	5.4	6.8	4
CO_2 fixation rate (nmol/min/mg core enzyme)	**5**	**66.9**	**37.3**

3.6 Reference

1. Erb TJ, Fuchs G, Alber BE (2009) (2 *S*)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol Microbiol* 73(6):992–1008.

2. Khobragade DA, et al. (2012) Acceptor-Substituted Ferrocenium Salts as Strong, Single-Electron Oxidants: Synthesis, Electrochemistry, Theoretical Investigations, and Initial Synthetic Application. *Chem - A Eur J* 18(39):12267–12277.

3. Carlson BW, Miller LL, Neta P, Neta P, Grodkowski J (1984) Oxidation of NADH Involving Rate-Limiting One-Electron Transfer. *J Am Chem Soc* 106(23):7233–7239.

4. Schwander T, Von Borzyskowski LS, Burgener S, Cortina NS, Erb TJ (2016) A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* 354(6314):900–904.

5. Takahashi S, Anzai J (2013) Recent Progress in Ferrocene-Modified Thin Films and Nanoparticles for Biosensors. *Materials (Basel)* 6(12):5742–5762.

6. Kobayashi Y, Hoshi T, Anzai J (2001) Glucose and lactate biosensors prepared by a layer-by-layer deposition of concanavalin A and mannose-labeled enzymes: electrochemical response in the presence of electron mediators. *Chem Pharm Bull (Tokyo)* 49(6):755–7.

7. Veitch NC (2004) Horseradish peroxidase: A modern view of a classic enzyme. *Phytochemistry* 65(3):249–259.

8. DA MATTA FM, LOOS RA, RODRIGUES R, BARROS RS (2001) Actual and potential photosynthetic rates of tropical crop species. *Rev Bras Fisiol Veg* 13(1):24–32.

9. Gale NL, Beck J V. (1967) Evidence for the Calvin cycle and hexose monophosphate pathway in Thiobacillus ferrooxidans. *J Bacteriol* 94(4):1052–1059.

10. Antonovsky N, et al. (2016) Sugar Synthesis from CO2 in Escherichia coli. *Cell* 166(1):115–125.

11. Mattozzi M d., Ziesack M, Voges MJ, Silver PA, Way JC (2013) Expression of the sub-pathways of the Chloroflexus aurantiacus 3-hydroxypropionate carbon fixation bicycle in E. coli: Toward horizontal transfer of autotrophic growth. *Metab Eng* 16:130–139.

12. Lin PP, et al. (2018) Construction and evolution of an Escherichia coli strain relying on nonoxidative glycolysis for sugar catabolism. *Proc Natl Acad Sci*:201802191.

13. Mainguet SE, Gronenberg LS, Wong SS, Liao JC (2013) A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab Eng* 19:116–27.

14. Yu H, Li X, Duchoud F, Chuang DS, Liao JC (2018) Augmenting the calvin-benson-bassham cycle by a synthetic malyl-CoA-glycerate carbon fixation pathway. *Nat Commun* 9(1):1–10.

15. Boehler C, Stieglitz T, Asplund M (2015) Nanostructured platinum grass enables superior impedance reduction for neural microelectrodes. *Biomaterials* 67:346–353.

16. Bar-Even A, Noor E, Lewis NE, Milo R (2010) Design and analysis of synthetic carbon fixation pathways. *Proc Natl Acad Sci* 107(19):8889–8894.

17. YOSHIDA T, FUKUTA K, MITSUNAGA T, YAMADA H, IZUMI Y (1992) Purification and characterization of glycerate kinase from a serine‐producing methylotroph, Hyphomicrobium methylovorum GM2. *Eur J Biochem* 210(3):849–854.

18. Aust A, Yun SL, Suelter CH (1975) Pyruvate kinase from yeast (Saccharomyces cerevisiae). *Methods Enzymol* 42(C):176–182.

19. Hirabayashi T, Harada T (1972) Isolation and properties of a pyruvate dehydrogenase complex from the yeast Hansenula miso. *J Biochem* 71(5):797–804.

Chapter 4. Strategies for constructing a fully functional rGPS *in vivo*

4.1 Introduction

After building the cell-free rGPS system for fixing CO_2 *in vitro*, making this synthetic CO_2 fixation cycle function *in vivo* has became our next milestone. rGPS is composed of efficient, oxygen tolerant and structurally simple enzymes. When transplanted *in vivo,* the expression and activities of rGPS enzymes should not be a big concern. However, problems related to cofactor balance and availability, interactions with the endogenous metabolism, and metabolic regulations, could still hinder the *in vivo* function of rGPS. To test, study and demonstrate the feasibility of parts of and the whole rGPS *in vivo*, we will construct testing platforms using the user-friendly microorganism, *E. coli*. Theoretically, the functional rGPS cycle *in vivo* together with proper energy source can convert *E. coli* to direct CO_2 utilization microorganism. However, converting a heterotroph into an autotroph requires a complete rewiring and remodeling of the central carbon and energy metabolism, making it a daunting task in synthetic biology. Many effects have been made to transplant natural CO_2 fixation pathways into *Escherichia coli* (1, 2)*, Pyrococcus furiosus* (3) or *Methylobacterium extorquens* (4). The *in vivo* function of these pathways has been demonstrated, but the conversion to an autotrophic growth on CO_2 has not been achieved. To realize this goal, we will rely heavily on evolution to tune the rGPS cycle in the cell (5). This chapter will discuss the strategies for constructing rGPS *in vivo* and present some preliminary results.

4.2 Demonstrating the rGPS cycle *in vivo*

To demonstrate the rGPS cycle can function *in vivo*, we will first divide this cycle into two parts and demonstrate each part individually. The fist part includes 5 enzymes (Pps, Ppc, Mdh, Mtk,

Mcl), converting pyruvate to acetyl-CoA and glyoxylate with one HCO_3^- fixed, termed as the reductive glyoxylate synthesis (rGS) pathway (6, 7). The second part consists of 10 enzymes (AtoB, Bhbd, Crt, Ccr, Epi, Ecm, Mcd, Mct, Meh, Ccl), converting acetyl-CoA to pyruvate with one CO_2 fixed, termed as the reductive pyruvate synthesis (rPS) pathway. The "acetyl-CoA auxotrophic" *E. coli* will be used to test the first part. And a "pyruvate auxotrophic" *E. coli* will be used to demonstrate the second part.

4.2.1 Demonstrating "the rGS part" of rGPS *in vivo*

To construct "the rGS part" of rGPS in *E. coli*, it is necessary to overexpress two enzymes: Mtk and Mcl. Other enzymes (Pps, Ppc, and Mdh) are natively expressed in *E. coli*. The acetyl-CoA auxotroph (JCL16ΔaceEFΔpoxBΔpflB) was used as the host. This strain cannot grow on glucose minimal medium unless supplemented with acetate. After expressing Mtk and Mcl from plasmids, the acetyl-CoA auxotroph became able to grow without acetate supplementation (Fig. 4.1).

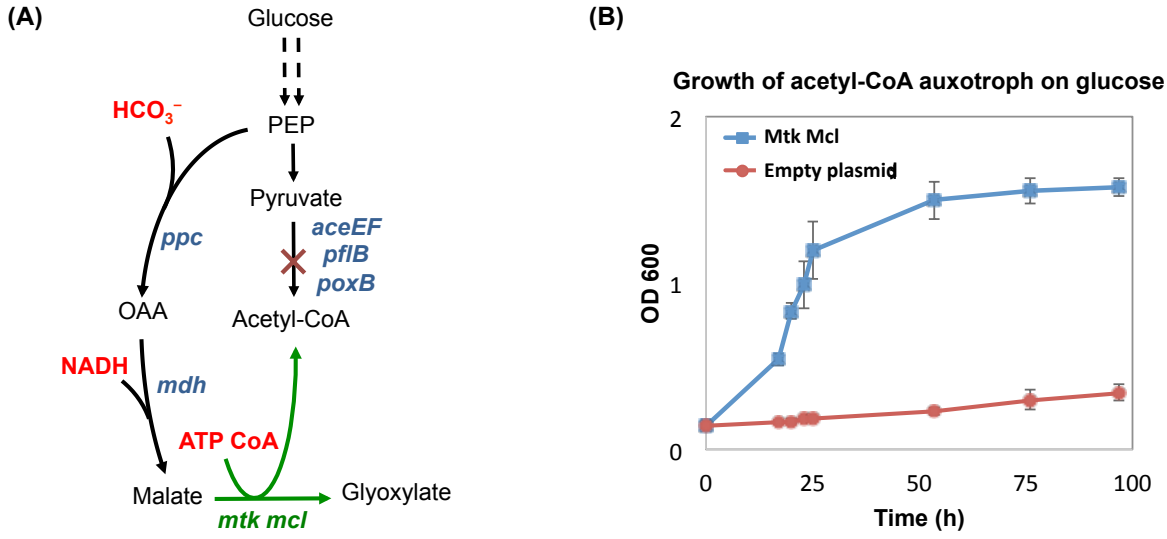


Fig. 4.1. Strategies and results for demonstrating "the rGS part" of rGPS *in vivo*.

(A) Illustration of how "the rGS part" of rGPS rescues the acetyl-CoA auxotrophic strain. The native enzymes are shown in blue. Overexpressed enzymes are shown in green. (B) Growth curve of acetyl-CoA auxotroph with Mtk Mcl overexpressed or empty plasmid in M9 minimal medium with 1% glucose and 0.1mM IPTG (n=3). (Data from Sio Si Wong)

Ideally, the growth of acetyl-CoA auxotroph on pyruvate minimal medium without supplying acetate can demonstrate all five genes of "the rGS part" (Mtk, Mcl, Pps, Ppc, and Mdh) function *in vivo*. However, we found that the acetyl-CoA auxotroph could not grow solely on pyruvate. This phenomenon may be a result of lacking reducing power or metabolic regulations. The growth of acetyl-CoA auxotroph on glucose minimal medium without acetate supplementation already demonstrated four genes of "the rGS part" (Mtk, Mcl, Ppc, and Mdh) could function *in vivo*. Since Pps is natively expressed under growth conditions that require the gluconeogenesis pathway and is highly active in *E. coli*, its expression and activity should not be a concern for constructing rGPS *in vivo*.

4.2.2 Demonstrating "the rPS part" of rGPS *in vivo*

To construct "the rPS part" of rGPS in *E. coli*, we need to overexpress ten enzymes: AtoB, Bhbd, Crt, Ccr, Epi, Ecm, Mcd, Mct, Meh, and Ccl. To test "the rPS part" of rGPS *in vivo,* we constructed a "pyruvate auxotrophic" *E. coli* strain (ΔaceA) (Fig. 4.2). Knocking out isocitrate lyase (aceA) blocked the glyoxylate shunt for acetate assimilation. As a result, acetate cannot be assimilated to pyruvate and can only be burned to CO_2 through TCA cycle. ΔaceA strain cannot grow on acetate and bicarbonate minimal medium unless supplemented with pyruvate.

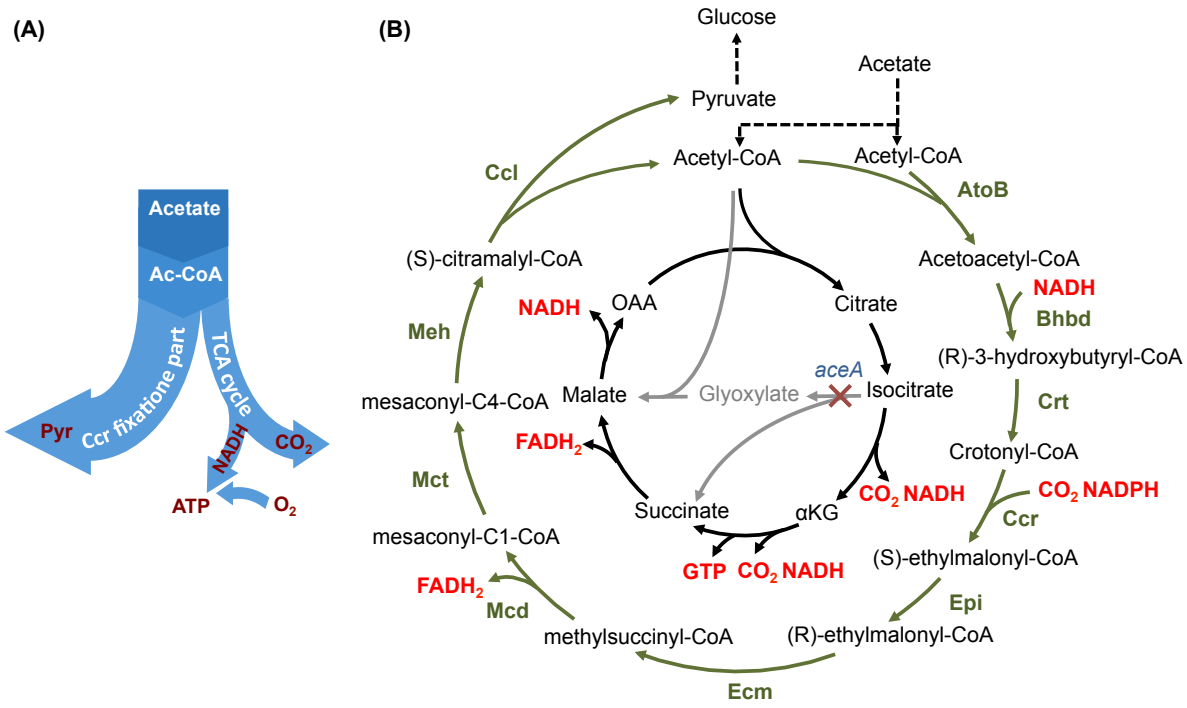


Fig. 4.2. Metabolic strategies for aerobic growth on acetate using "the rPS part" of rGPS in *E. coli*.

(A) Illustration of overall carbon flux. (B) Simplified metabolic map for aerobic growth on acetate using "the rPS part" of rGPS. The native enzymes are shown in blue. Overexpressed enzymes are shown in green.

We grouped ten enzymes of the rPS pathway into two plasmids and overexpressed all ten enzymes in ΔaceA strain to rescue its growth in acetate and bicarbonate minimal medium without supplying pyruvate. If "the rPS part" works *in vivo*, the cell can use this pathway to assimilate acetate to pyruvate and should grow with acetate and bicarbonate as carbon sources (Fig. 4.2). However, this growth rescue failed. To check if all of the enzymes were well expressed, we did crude extract assays to check enzyme activities. Assay results confirmed that ΔaceA strain with both plasmids had all ten enzymes' activities.

The growth rescue failure may result from lacking proper electron transfer proteins to work with Mcd in *E. coli* or other regulation issues (8). Recently, the electron transfer protein for Mcd has been reported (9). We will overexpress this electron transfer protein along with all ten rPS enzymes in ΔaceA strain to check the growth rescue again. If the rescue still fails, we will rely on evolution to get the strain adapted to growing in acetate and bicarbonate minimal medium. Then, we can learn how cells make the rPS pathway work *in vivo* from sequencing, paving the way for constructing the whole rGPS cycle *in vivo*.

4.3 Constructing rGPS in *E. coli* for autotrophic growth

After demonstrating the feasibility of *in vivo* rGPS in two parts, the complete rGPS cycle will be constructed in *E. coli* to convert it to direct CO_2 utilization microorganism. Since rGPS is a non-nature and non-native pathway, implementing such pathway presents daunting challenges related to expression levels, cofactors and regulations, we will mainly utilize evolution to tune the pathway in the cell. First, a selection platform will be constructed to train *E. coli* to use rGPS efficiently. After rGPS is evolved to function well in the cell, other necessary genes will be added and the resulting strain will be further evolved for autotrophic growth.

4.3.1 Evolution of rGPS in *E. coli*

To train *E. coli* to use rGPS pathway efficiently, we constructed a rGPS selection strain (ΔmaeABΔpckΔgcl), which cannot grow on aspartate minimal medium unless the whole rGPS pathway (except Ppc) works.

Phosphoenoylpyrvate carboxykinase (pck) and the malic enzymes (maeA, maeB) were knocked out so that pyruvate or PEP cannot be synthesized from TCA intermediates or acetyl-CoA (Fig.

4.3). After overexpressing malate thiokinase (Mtk) and malyl-CoA lyase (Mcl), ΔmaeABΔpck is able to grow on aspartate minimal medium since glyoxylate produced through Mtk Mcl can be converted to 3-phosphoglycerate for gluconeogenesis.

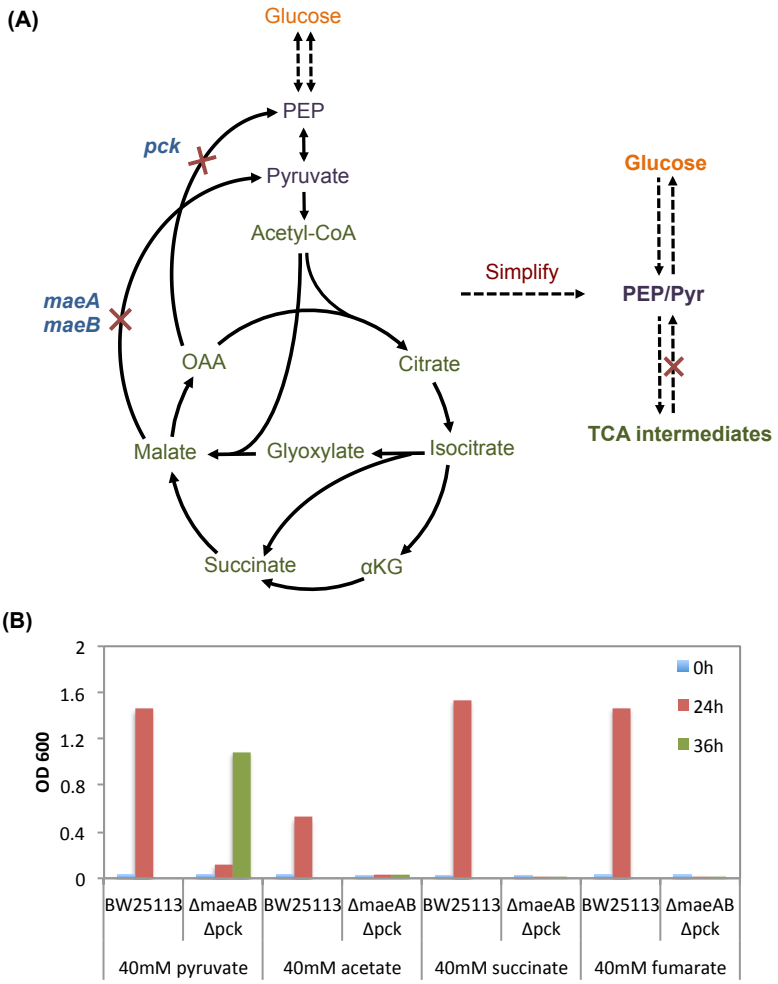


Fig. 4.3. Illustration of ΔmaeABΔpck strain.

(A) Simplified central metabolic map of ΔmaeABΔpck strain. (B) The growth status of ΔmaeABΔpck strain on various carbon sources compared with wild type *E. coli* (BW25113). (Data from Hong Yu)

After knocking out the glyoxylate carboligase (gcl), the resulting strain ΔmaeABΔpckΔgcl cannot grow on aspartate minimal medium with only Mtk and Mcl overexpressed (Fig. 4.4).

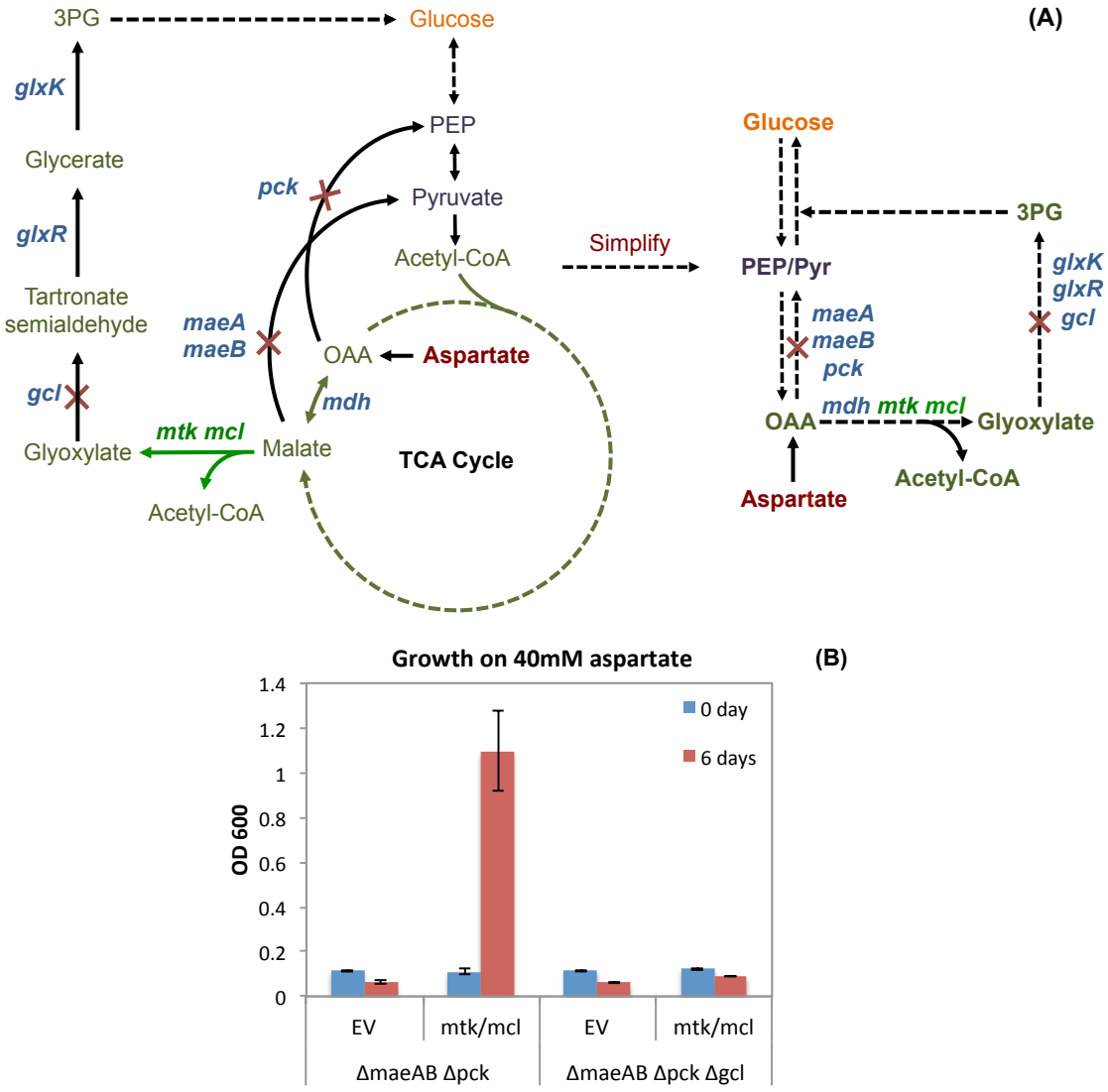


Fig. 4.4. Illustration of ΔmaeABΔpckΔgcl strain.

(A) Simplified central metabolic map of ΔmaeABΔpckΔgcl strain. (B) The growth status of ΔmaeABΔpck or ΔmaeABΔpckΔgcl strain with Mtk Mcl overexpressed or empty vector in M9 minimal medium with 40 mM aspartate and 0.1 mM IPTG (n=3). (Data from Hong Yu)

To rescue ΔmaeABΔpckΔgcl on aspartate, the acetyl-CoA to pyruvate part of rGPS pathway should also be overexpressed. Which means the whole rGPS pathway (except ppc) works, the ΔmaeABΔpckΔgcl strain can grow on aspartate minimal medium (Fig. 4.5). Since ppc is highly

active in *E. coli*, its expression and activity should not be a problem for completing the whole rGPS pathway.

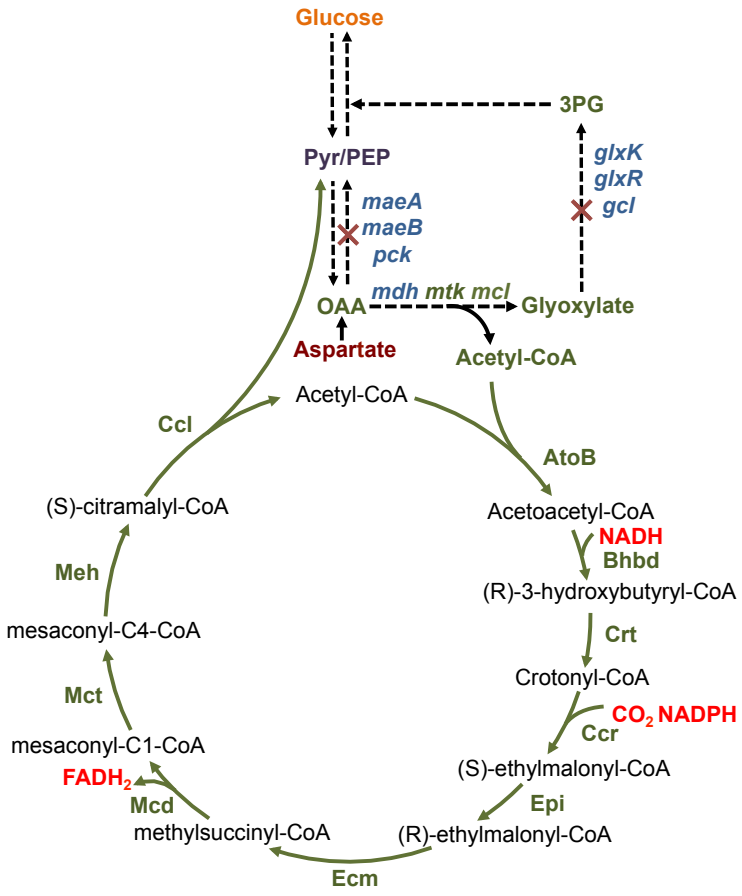


Fig. 4.5. Metabolic strategies for aerobic growth on aspartate using rGPS in ΔmaeABΔpckΔgcl strain.

The native enzymes are shown in blue. Overexpressed enzymes are shown in green.

We will overexpress 12 enzymes of rGPS from plasmids (Mtk, Mcl, AtoB, Bhbd, Crt, Ccr, Epi, Ecm, Mcd, Mct, Meh, Ccl) in the rGPS selection strain (ΔmaeABΔpckΔgcl). After those enzymes are properly overexpressed, ΔmaeABΔpckΔgcl should be able to grow on aspartate minimal medium. To optimize the flux through the constructed pathway, we will evolve the strain for better growth. In aspartate minimal medium, the cells which use the whole rGPS

pathway (except Ppc) will have the highest growth rate and will be enriched in repeated dilution and growth.

4.3.2 Evolution of *E. coli* for autotrophic growth using rGPS

After the whole rGPS pathway (except Ppc) is optimized in the cell, we will overexpress other enzymes that are necessary for autotrophic growth, such as gcl for glyoxylate assimilation, formate dehydrogenase (fdh) for utilizing formate as the electron donor. And the resulting strain will be evolved for autotrophic growth with gradually reducing the amount of aspartate and increasing the amount of bicarbonate and formate. Finally, the strain will grow on CO_2 as the carbon source with formate providing the reducing power (Fig. 4.6).

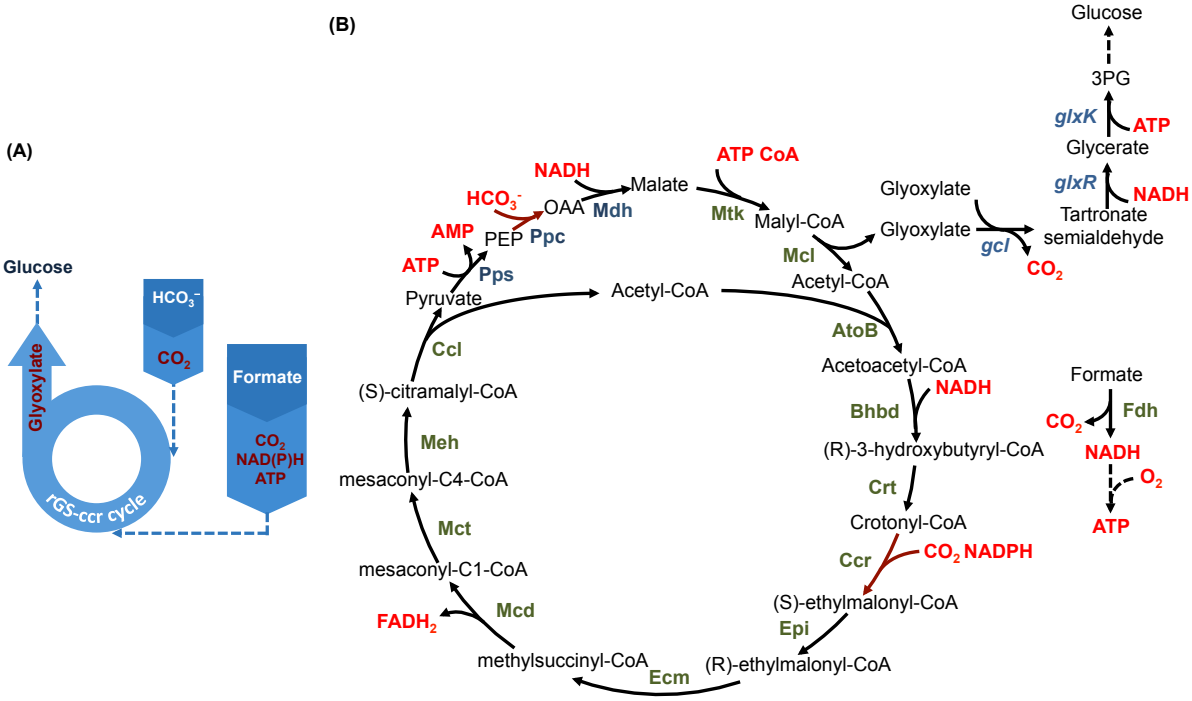


Fig. 4.6. Metabolic strategies for aerobic growth on bicarbonate and formate using rGPS in *E. coli*.

(A) Illustration of overall carbon flux. (B) Simplified metabolic map for aerobic growth on bicarbonate and formate using rGPS. The native enzymes are shown in blue. Overexpressed enzymes are shown in green.

4.4 Materials and methods

4.4.1 Chemicals and reagents

All chemicals used were purchased from Sigma-Aldrich unless otherwise specified. KOD and KOD Xtreme DNA polymerases were from EMD Millipore. Oligonucleotides were obtained from IDT. Gibson Assembly Master Mix was purchased from New England Biolabs.

4.4.2 Culture medium and conditions

E. coli strains were grown in Luria-Bertani (LB) medium or M9 minimal medium with appropriate antibiotics at 37 ^{o}C. M9 minimal medium contains M9 salts (12.8 g/L $Na_2HPO_4 \cdot 7H_2O$, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl), 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.1 mg/mL thiamine hydrochloride, and carbon sources as noted in the work. Antibiotics were used at the following concentrations: carbenicillin, 200 μg/mL and chloramphenicol, 50 μg/mL.

4.4.3 Strain construction

All *E. coli* strains used are listed in Table 4.1. JCL16 was used to create the acetyl-CoA auxotroph ΔaceEFΔpoxBΔpflB. *E. coli* (BW25113) was used to construct other strains (ΔaceA, ΔmaeABΔpck, and ΔmaeABΔpckΔgcl). Gene deletion was performed by P1 transduction with single knockout strain from the Keio collection.

4.4.4 Plasmid Construction

All plasmids used and constructed in this work are list in Table 4.1. Plasmids were assembled using the Gibson isothermal DNA assembly method and constructed in *E. coli* strain XL-1 blue for propagation and storage.

4.4.5 Preparation of cell extract for enzyme assays

Overnight culture was inoculated (1% vol/vol) into fresh LB medium with 5 μM/mL hydroxocobalamin in a foil-wrapped flask. Cells were grown in the dark at 37 °C to OD600 of 0.4-0.6, and induced by 0.1 mM IPTG for an additional 6 h at 30 °C. Then, cells were collected by centrifugation, and stored in the dark at -80 □C. For soluble extract preparation, the pellet of 50 ml cell culture was thawed, and resuspended in 0.6 mL of 100 mM Tris-HCL buffer PH=7.5. Following sonication on ice, the extract was clarified by centrifugation and the supernatant was collected to use for enzyme assays. Protein concentrations were determined by the Bradford protein assay using Pierce Coomassie Plus Assay Kit (Thermo Fisher Scientific).

4.4.6 Enzyme assays

All assays were conducted at room temperature under aerobic conditions. The reaction mixture volume was 0.6 ml unless otherwise indicated.

AtoB assay. The AtoB activity in cell crude extract was tested via the thiolysis direction by monitoring the decrease of absorbance at 303 nm corresponding to the result of acetoacetyl-CoA cleavage (10). The reaction was initiated by the addition of cell crude extract. The reaction mixture contained 100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 200 μM acetoacetyl-CoA, 200 μM CoA, and 5 μl cell crude extract.

Bhbd assay. The Bhbd activity in cell crude extract was tested by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADH ($\varepsilon_{340\ nm}$ = 5.31 mM^{-1} cm^{-1}). The reaction was initiated by the addition of cell crude extract. The reaction mixture contained

100 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.2 mM NADH, 0.3 mM acetoacetyl-CoA, and 5 µl cell crude extract.

Crt assay. The Crt activity in cell crude extract was tested by monitoring the decrease of absorbance at 290 nm corresponding to the consumption of crotonyl-CoA. The reaction was initiated by the addition of cell crude extract. The assay mixture contained 100 mM Tris-HCl pH 7.5, 0.2 mM crotonyl-CoA, and 5 µl cell crude extract.

Ccr assay. The Ccr activity in cell crude extract was tested by monitoring the decrease of absorbance at 340 nm corresponding to the consumption of NADPH ($\varepsilon_{340\ nm}$ = 5.31 mM^{-1} cm^{-1}). The reaction was initiated by the addition of cell crude extract. The reaction mixture contained 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 0.2 mM NADPH, 0.2 mM crotonyl-CoA, and 5 µl cell crude extract.

Epi, Ecm, Mcd, Mct, Meh, and Ccl assays. Activities of Epi, Ecm, Mcd, Mct, Meh, and Ccl in cell crude extract were tested with HPLC by monitoring the production of pyruvate and acetyl-CoA from (2S)-methylsuccinyl-CoA. A substrate mixture contained 100 mM Tris-HCl pH 7.0, 50 mM $NaHCO_3$, 0.6 mM NADPH, 0.6 mM crotonyl-CoA, and 3 µg his-tag purified Ccr was incubated for 5 minutes to synthesize (2S)-ethylmalonyl-CoA. The reaction mixture contained 100 mM Tris-HCl pH 7.5, 0.2 mM ferrocenium hexafluorophosphate ($FePF_6$), 0.1 ml substrate mixture and appropriate amounts of his-tag purified Epi, Ecm, Mcd, Mct, Meh, Ccl, and cell crude extract. For Epi assay, 6 µg Ecm, 10 µg Mcd, 2.5 µg Mct, 5 µg Meh, 9 µg Ccl and 5 µl cell crude extract were used. For Ecm assay, 1.5 µg Epi, 10 µg Mcd, 2.5 µg Mct, 5 µg Meh, 9 µg Ccl and 5 µl cell crude extract were used. For Mcd assay, 1.5 µg Epi, 6 µg Ecm, 2.5 µg Mct, 5 µg Meh, 9 µg Ccl and 5 µl cell crude extract were used. For Mct assay, 1.5 µg Epi, 6 µg Ecm, 10 µg

Mcd, 5 μg Meh, 9 μg Ccl and 5 μl cell crude extract were used. For Meh assay, 1.5 μg Epi, 6 μg Ecm, 10 μg Mcd, 2.5 μg Mct, 9 μg Ccl and 5 μl cell crude extract were used. For Ccl assay, 1.5 μg Epi, 6 μg Ecm, 10 μg Mcd, 2.5 μg Mct, 5 μg Meh, and 5 μl cell crude extract were used. Before assays, cell crude extract was incubated with 0.1 mM Coenzyme B12 in the dark on ice for 1 h for Ecm activation. The reaction mixtures were incubated in room temperature for 30 mins, then were quenched with 5% formic acid, derivatized with 10 mM phenylhydrazine and analyzed with HPLC.

4.4.7 HPLC analysis

CoA compounds and pyruvate phenylhydrazone were analyzed by an Agilent 1200 HPLC equipped with a reversed-phase C18 column (LiChrospher 100, end-capped, 5 μm, LiChroCART 125-4 125 × 4 mm). Separation was performed using one of the two gradient systems of 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent A) and 20% acetonitrile (MeCN) in 40 mM K_2HPO_4/HCOOH buffer at pH 4.2 (solvent B). Column temperature was maintained at 25 ^oC and the flow rate stayed constant at 1 ml/min. Initial conditions were 90% solvent A and 10% solvent B. At 0 min, solution B was increased linearly to 100% over 40 min. At 40 min, the composition of B was held isocratic for 3 min. At 43 min, solution B was decreased to its initial 10% condition over 1 min and held for 16 min to reequilibrate the column.

4.4.8 Growth rescue of *E. coli* strains

Overnight culture was inoculated (1% vol/vol) into fresh LB medium with 5 μM/mL hydroxocobalamin in a foil-wrapped tube. Cells were grown in the dark at 37 °C to OD600 of 0.4-0.6, and induced by 0.1 mM IPTG for an additional 6 h at 30 °C. 1 ml of cell culture was harvested and washed three times with 1 ml of M9 minimal medium. Then the washed culture

was inoculate (1% vol/vol) into 3 mL M9 minimal medium with 5 μM/mL hydroxocobalamin for growth testing in the dark at 37 °C. The growth tests were performed in aerobic conditions.

Table 4.1. Strains and plasmids used.

Strain	Relevant Genotypes	Reference
JCL16	*rrnB*$_{T14}$ *ΔlacZ*$_{WJ16}$ *hsdR514 ΔaraBAD*$_{AH33}$ *ΔrhaBAD*$_{LD78}$ / F' [*traD36 proAB* + *lacI*^{q}*ZΔM15*]	(11)
BW25113	*rrnB*$_{T14}$ *ΔlacZ*$_{WJ16}$ *hsdR514 ΔaraBAD*$_{AH33}$ *ΔrhaBAD*$_{LD78}$	(12)
SS108	Δ*aceEF* Δ*poxB* Δ*pflB* in JCL16	(7)
ΔaceA	Δ*aceA* in BW25113	This work
BWΔ3	Δ*maeAB* Δ*pck* in BW25113	This work
BWΔ4	Δ*maeAB* Δ*pck* Δ*gcl* in BW25113	This work
Plasmid	**Description**	**Reference**
pXL133	ColE1 ori; Amp^{R}; P$_{L}$lacO$_{1}$:: *atoB(E.c)/crt(C.a)/bhbd(C.a)*; P$_{L}$lacO$_{1}$:: *epi(R.s)/ecm(R.s)*	This work
pHYL	p15A ori; Cm^{R}; P$_{L}$lacO$_{1}$:: *ccr(R.s)/mcd (R.s)*; P$_{L}$lacO$_{1}$:: *ccl (P.a)/ meh(C.a)/mct(C.a)*	This work

4.5 Reference

1. Mattozzi M d., Ziesack M, Voges MJ, Silver PA, Way JC (2013) Expression of the sub-pathways of the Chloroflexus aurantiacus 3-hydroxypropionate carbon fixation bicycle in E. coli: Toward horizontal transfer of autotrophic growth. *Metab Eng* 16(1):130–139.

2. Antonovsky N, et al. (2016) Sugar Synthesis from CO2in Escherichia coli. *Cell* 166(1):115–125.

3. Keller MW, et al. (2013) Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proc Natl Acad Sci* 110(15):5840–5845.

4. Schada von Borzyskowski L, et al. (2018) An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in Methylobacterium extorquens AM1. *Metab Eng* 47(April):423–433.

5. Lin PP, et al. (2018) Construction and evolution of an Escherichia coli strain relying on nonoxidative glycolysis for sugar catabolism. *Proc Natl Acad Sci*:201802191.

6. Mainguet SE, Gronenberg LS, Wong SS, Liao JC (2013) A reverse glyoxylate shunt to build a non-native route from C4 to C2 in Escherichia coli. *Metab Eng* 19:116–27.

7. Yu H, Li X, Duchoud F, Chuang DS, Liao JC (2018) Augmenting the calvin-benson-bassham cycle by a synthetic malyl-CoA-glycerate carbon fixation pathway. *Nat Commun* 9(1):1–10.

8. Erb TJ, Fuchs G, Alber BE (2009) (2 *S*)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol Microbiol* 73(6):992–1008.

9. Schwander T, McLean R, Zarzycki J, Erb TJ (2018) Structural basis for substrate specificity of methylsuccinyl-CoA dehydrogenase, an unusual member of the acyl-CoA dehydrogenase family. *J Biol Chem* 293(5):1702–1712.

10. Shen CR, et al. (2011) Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Appl Environ Microbiol* 77(9):2905–15.

11. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451(7174):86–89.

12. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 97(12):6640–5.