Realization of a New-to-Nature Carboxylation Pathway

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Biologie
der Philipps-Universität Marburg

vorgelegt von

Marieke Scheffen

aus Simmerath

Marburg/Lahn, April 2020

Die Untersuchungen zur vorliegenden Arbeit wurden von Januar 2016 bis Dezember 2019 unter der Betreuung von Herrn Prof. Dr. Tobias J. Erb in Marburg am Max-Planck-Institut für terrestrische Mikrobiologie in der Abteilung Biochemie und Synthetischer Metabolismus, sowie im Centre de Recherche Paul Pascal in Bordeaux, Frankreich, durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am 02.07.2020

Erstgutachter: Prof. Dr. Tobias J. Erb
Zweitgutachter: Prof. Dr. Martin Thanbichler

Weitere Mitglieder der Prüfungskommission:
Prof. Dr. Stefan Rensing
Prof. Dr. Lars-Oliver Essen

Tag der Disputation: 15.07.2020

Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „Realization of a New-to-Nature Carboxylation Pathway“ selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 06.04.2020

Marieke Scheffen

Teile dieser Arbeit sind im Rahmen von Abschlussarbeiten entstanden, die von mir betreut wurden. Die Ergebnisse dieser Bachelor- und Masterarbeiten sind teilweise in diese Dissertation eingeflossen. Im Folgenden werden die beteiligten Personen sowie deren experimentelle Beiträge genannt:

Daniel Marchal (Masterstudent)

- Messungen zur Bestimmung der kinetischen Parameter der Glykolyl-CoA Carboxylasen (GCC)
- Screening der Enzym-Bibliotheken der GCC
- Durchführung des Avidin-Gel-Shift Assays der GCC

Katharina Happel (Masterstudentin)

- Erstellung der *Pseudomonas putida* Knockout-Stämme
- Erstellung der *Pseudomonas putida* Expressionsplasmide
- Durchführung der *Pseudomonas putida* Wachstumsversuche

Jessica Lehmann (Bachelorstudentin)

- Erstellung des *E. coli* BL21 (DE3) AI Δ*patZ* Stammes zur Expression von Acyl-CoA Synthetasen

Teile dieser Arbeit sind im Rahmen von Kollaborationen entstanden. Im Folgenden werden die beteiligten Personen sowie deren experimentelle Beiträge genannt:

Thomas Beneyton (Centre de Recherche Paul Pascal, Bordeaux, Frankreich)

- Herstellen der mikrofluidischen Chips und technische Unterstützung während der mikrofluidischen Experimente mit GCC

Sandra und Jan Schuller (MPI für Biochemie, Martinsried)

- Probenaufbereitung und Datenerhebung für die kryo-EM Analyse der GCC

Jan Zarzycki (Max-Planck-Institut für terrestrische Mikrobiologie, Marburg)

- Kristallstrukturbestimmung der Malonyl-CoA Reduktase
- Fitting und Refinement der GCC cryo-EM Struktur

Christoph Diehl (Max-Planck-Institut für terrestrische Mikrobiologie, Marburg)

- Aufreinigung der CETCH-Enzyme und UPLC-MS/MS Messungen für die TaCo/CETCH Experimente

Parts of this thesis that have been published or are in preparation for publication:

Marieke Scheffen, Daniel G. Marchal, Thomas Beneyton, Sandra K. Schuller, Melanie Klose, Christoph Diehl, Jessica Lehmann, Pascal Pfister, Martina Carrillo, Selçuk Aslan, Niña S. Cortina, Peter Claus, Daniel Bollschweiler, Jean-Christophe Baret, Jan M. Schuller, Jan Zarzycki, Arren Bar-Even, Tobias J. Erb (2020).
A new-to-nature carboxylation module improves natural and synthetic CO_2 fixation. *Submitted.*

Marieke Scheffen, Katharina Happel, Tobias J. Erb.
In vivo metabolic engineering of *Pseudomonas putida* for ethylene glycol conversion via the carbon-fixing tartronyl-CoA pathway. *In preparation.*

Publications that are not discussed in this thesis:

Devin L. Trudeau, Christian Edlich-Muth, Jan Zarzycki, **Marieke Scheffen**, Moshe Goldsmith, Olga Khersonsky, Ziv Avizemer, Sarel J. Fleishman, Charles A. R. Cotton, Tobias J. Erb, Dan S. Tawfik, Arren Bar-Even (2018).
Design and in vitro realization of carbon-conserving photorespiration. *PNAS* 115(49), E11455-E11464.

Stefanie Böhnke, **Marieke Scheffen**, Niña S. Cortina, Nicole Paczia, Tobias J. Erb, Mirjam Perner
An activity-based metagenomic screen uncovers novel perception of enzymatically catalyzed decrease of ribulose-1,5-bisphosphate. *In preparation.*

Marieke Scheffen (2019).
Mikroorganismus in den Schlagzeilen. Schreib's noch einmal, Sam: mit 37 Stücken zum minimalen Genom. *Biospektrum* 25: 410.

"*You cannot get through a single day without*
having an impact on the world around you.
What you do makes a difference,
and you have to decide
what kind of difference you want to make."

- Jane Goodall

Contents

Summary

Most inorganic carbon enters the biosphere via the Calvin-Benson-Bassham (CBB) cycle by its key enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). An unproductive side reaction of RuBisCO with oxygen leads to the formation of 2-phosphoglycolate (2-PG), which is recycled via complex pathways into 3-phospho-glycerate (3-PGA), releasing CO_2 in the process. The tartronyl-CoA pathway represents a synthetic pathway that was designed to recycle 2-PG more efficiently, avoiding the release of CO_2, and fixing CO_2 instead. It consists of four main reactions steps, which are not known to take part in any natural metabolic pathway. These steps are the activation of glycolate to glycolyl-CoA, the carboxylation of glycolyl-CoA to tartronyl-CoA as its key reaction, and the subsequent two reductions giving rise to glycerate.

In this work, all required enzymes were identified or established by engineering and the tartronyl-CoA pathway was realized *in vitro*. Promiscuous enzyme candidates performing analogous reactions with similar substrates were screened and further improved to perform their desired functions. These include engineered glycolyl-CoA synthetase and glycolyl-CoA carboxylase (GCC), as well as a tartronyl-CoA reductase. For the engineering of GCC, rational design as well as high-throughput directed evolution was applied resulting in a new-to-nature carboxylase that matches the kinetic properties of natural carboxylases. Moreover, a 1.96 Å resolution cryogenic electron microscopy (cryo-EM) structure of GCC was obtained, highlighting and corroborating the effects of the introduced mutations. The concerted function of all tartronyl-CoA pathway enzymes was confirmed in the context of photorespiration *in vitro*. The *in vitro* reconstitution also included the optimization of reaction parameters as well as efficient cofactor recycling. Besides its function as photorespiratory bypass, the tartronyl-CoA pathway was shown to be functional as an additional carbon fixing module, able to connect a synthetic CO_2 fixation cycle to central carbon metabolism. Furthermore, the tartronyl-CoA pathway was successfully employed for the *in vitro* conversion of the plastic waste component ethylene glycol into the central carbon metabolite glycerate. In an initial attempt of an *in vivo* implementation of the tartronyl-CoA pathway for ethylene glycol assimilation, it was shown that GCC, the key enzyme of the tartronyl-CoA pathway, can be functionally produced in *Pseudomonas putida*.

Zusammenfassung

Der größte Teil des anorganischen Kohlenstoffs gelangt über den Calvin-Zyklus durch das Schlüsselenzym Ribulose-1,5-bisphosphat-carboxylase/-oxygenase (RuBisCO) in die Biosphäre. Eine unproduktive Nebenreaktion von RuBisCO mit Sauerstoff führt zur Bildung von 2-Phosphoglykolat (2-PG), das über komplexe Reaktionen in 3-Phosphoglycerat (3-PGA) zurückgeführt wird, wobei CO_2 freigesetzt wird. Der Tartronyl-CoA-Weg stellt einen synthetischen Stoffwechselweg dar, der entwickelt wurde, um 2-PG effizienter zu recyceln, die Freisetzung von CO_2 zu vermeiden und stattdessen CO_2 zu fixieren. Er besteht aus vier Hauptreaktionsschritten, von denen nicht bekannt ist, dass sie an natürlichen Stoffwechselwegen beteiligt sind. Diese Schritte sind die Aktivierung von Glykolat zu Glykolyl-CoA, die Carboxylierung von Glykolyl-CoA zu Tartronyl-CoA als Schlüsselreaktion und die nachfolgenden zwei Reduktionen, die zu Glycerat führen.

In dieser Arbeit wurden alle erforderlichen Enzyme identifiziert oder durch Engineering etabliert und der Tartronyl-CoA-Weg *in vitro* realisiert. Enzymkandidaten, die anologe Reaktionen mit ähnlichen Substraten katalysieren, wurden getestet und weiter verbessert, um die gewünschten Reaktionen zu katalysieren. Dazu gehören die engineerten Glykolyl-CoA-Synthetase und Glykolyl-CoA-Carboxylase (GCC) sowie eine Tartronyl-CoA-Reduktase. Für das Engineering von GCC wurde rationales Design sowie Hochdurchsatz-Screening von Varianten angewendet. Dies führte zu einer neuen Carboxylase, die den kinetischen Eigenschaften natürlicher Carboxylasen entspricht. Darüber hinaus wurde eine Kryoelektronenmikroskopie-Struktur von GCC mit einer Auflösung von 1,96 Å erhalten, die die Auswirkungen der eingeführten Mutationen hervorhebt und bestätigt. Die konzertierte Funktion aller Enzyme des Tartronyl-CoA-Weges wurde im Kontext von Photorespiration *in vitro* bestätigt. Die *In-vitro*-Rekonstitution umfasste auch die Optimierung der Reaktionsparameter sowie ein effizientes Cofaktorrecycling. Neben seiner Funktion als photorespiratorischer Bypass, funktioniert der Tartronyl-CoA-Weg als zusätzliches Kohlenstoff-Fixierungsmodul, das einen synthetischen CO_2-Fixierungszyklus mit dem zentralen Kohlenstoffmetabolismus verbinden kann. Darüber hinaus wurde der Tartronyl-CoA-Weg erfolgreich für die *In-vitro*-Umwandlung von Ethylenglykol (ein Bestandteil von Kunststoffabfall) zu Glycerat eingesetzt. In einem ersten Versuch einer *In-vivo*-Implementierung des Tartronyl-CoA-Weges zur Ethylenglykol-Assimilation wurde gezeigt, dass GCC, das Schlüsselenzym des Tartronyl-CoA-Weges, in *Pseudomonas putida* funktionell hergestellt werden kann.

1. Introduction

1.1 Global carbon cycle and rising atmospheric CO_2 concentration

Through the global carbon cycle, an amount of 210 Gt organic carbon (corresponding to 770 Gt CO_2) is circulated each year within the atmosphere and the terrestrial and oceanic biosphere (Friedlingstein et al., 2019). The quantitavely most important biochemical process for the fixation of CO_2 into the biosphere is the Calvin-Benson-Bassham (CBB) cycle (Tabita et al., 2007; Berg, 2011), which is found in land plants as well as in marine species like cyanobacteria and eukaryotic phytoplankton (e.g. diatoms) (Hügler and Sievert, 2011). Besides the CBB cycle, five alternative pathways have been discovered in bacteria and archaea: The reductive citric acid (Arnon-Buchanan) cycle, the reductive acetyl-CoA (Wood-Ljungdahl) pathway, the 3-hydroxypropionate (Fuchs-Holo) bi-cycle, as well as the dicarboxylate/4-hydroxybutyrate cycle and the 3-hydroxy-propionate/4-hydroxybutyrate cycle (Berg, 2011; Fuchs, 2011). The organic carbon produced by these autotrophic processes (primary production) on land and in the oceans is taken up by heterotrophic organisms like animals or microbes and converted back into CO_2 by respiration. Perturbing this functioning carbon cycle, anthropogenic CO_2 emissions of 41 Gt per year (averaged globally 2009 - 2018) (Friedlingstein et al., 2019), caused by emissions from fossil fuel burning and land use change such as deforestation, add up on top. Some of this additional CO_2 is taken up by carbon sinks within the biosphere, but an amount of 18 Gt per year remains as additional CO_2 within the atmosphere. As a consequence, the atmospheric CO_2 has been constantly rising from around 275 ppm in 1750 to over 400 ppm in 2018 and contributes to global warming and climate change (Pachauri et al., 2014; Friedlingstein et al., 2019). One of the most challenging tasks in today's world is to stop global warming by reducing CO_2 emissions into the atmosphere. A complementary approach for reducing atmospheric CO_2 would be, to take advantage of this readily available carbon source and develop biotechnological tools to convert CO_2 into value-added compounds. By this, economy would become more independent of fossil fuels and would help to reduce CO_2 emissions at the same time by directly recycling it.

1.2 RuBisCO: The key enzyme of the CBB cycle

As the CBB cycle is the most important natural carbon fixation pathway on earth, it has also been claimed that its key enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), is the most abundant protein on earth with 0.04 Gt in total, which corresponds to 5 kg per person on earth (Ellis, 1979; Phillips and Milo, 2009). Recent calculations suggest that the numbers might even be an order of magnitude higher, in the range of 0.7 Gt (Bar-On and Milo, 2019), which underlines the enormous significance of this enzyme even more. Besides its importance, the high abundancy can be explained by its rather slow carboxylation rate (k_{cat}^{c}), which is usually in the range of 1-10 s^{-1} (Flamholz et al., 2019) and its dual substrate specificity for CO_2 and O_2. In the CBB cycle, RuBisCO catalyzes the carboxylation of ribulose-1,5-bisphosphate (RuBP) with CO_2 to form two molecules of 3-phophoglycerate (3-PGA). If RuBisCO performs its adverse oxygenation reaction with O_2, one molecule of 3-PGA and one molecule of 2-phosphoglycolate (2-PG) are formed. This oxygenation reaction is an unfavourable side reaction of RuBisCO, as 2-PG is a toxic intermediate to central carbon metabolism; it inhibits triosephosphate isomerase and phosphofructokinase in a micromolar range (Anderson, 1971; Kelly and Latzko, 1976). The side reaction with O_2 is most probably a relict of RuBisCO's evolutionary origin. It evolved in an atmosphere without O_2 and therefore in the beginning there was no evolutionary pressure for a selection against oxygenation. Only after the first and second Great Oxygenation Events, which were caused by oxygenic photosynthesis itself, RuBisCOs had to face an atmosphere containing O_2. At this time point, photosynthesis and the CBB cycle were already widely distributed and instead of "inventing" a new way of photosynthesis, rather carbon concentrating mechanisms (CCMs) to enhance carboxylation over oxygenation were developed. These include carboxysomes in cyanobacteria, pyrenoids in algae and hornworts, crassulacean acid metabolism (CAM) and C_4 photosynthesis (Ranson and Thomas, 1960; Slack and Hatch, 1967; Giordano et al., 2005; Villarreal and Renner, 2012; Hagemann et al., 2016; Erb and Zarzycki, 2018). This phenomenon can be observed for several metabolic functions and it has been described as "frozen metabolic accidents" (Leister, 2019). The concept argues that evolution is rather conservative. Once a solution for a specific function has developed, evolution sticks with it and all the necessary functions evolve together, rather then coming up with a completely new solution at a later time point (Leister, 2019). In the case of RuBisCO, a complete "RuBisCOsome" has evolved, including large and small

subunits of RuBisCO, several chaperones for removal of inhibtors and reactivation (Mueller-Cajar, 2017), carboxysomes in cyanobacteria and pyrenoids in eukaryotes (Erb and Zarzycki, 2018).

There are several different types of RuBisCOs known today. Form I RuBisCOs are present in higher plants, eukaryotic algae, cyanobacteria and proteobacteria. They consist of eight large (RbcL) and eight small (RbcS) subunits (L_8S_8), while forms II and III are comprised of only RbcL subunits ($(L_2)_n$, with n between 2 and 5) (Tabita et al., 2008; Erb and Zarzycki, 2018). Form II RuBisCOs are found in proteobacteria and eukaryotic algae, while form III and II/III RuBisCOs were first discovered in archaea and are now known to occur also in a class of bacteria called candidate phyla radiation (CPR) (Liu et al., 2017; Erb and Zarzycki, 2018).

For activation of the enzyme it is first carbamylated by a non-substrate CO_2 molecule at a conserved active site lysine and then binds a Mg^{2+} ion (Lorimer et al., 1976; Stec, 2012). The active complex binds RuBP, which is converted into a reactive enolate intermediate. This then attacks a CO_2 molecule, followed by the formation of a hydrated carboxyketone, which is subsequently cleaved into two molecules of 3-PGA (Tcherkez, 2013). The enolate intermediate can also attack an O_2 molecule, leading to the formation of a peroxyketone intermediate with subsequent cleavage into one molecule of 3-PGA and one molecule of 2-PG. Several studies suggest, that there is an inherent trade-off between velocity and specificity for CO_2: More specific enzymes seem to be rather slow, while faster enzymes have a lower specificity for CO_2 (Tcherkez et al., 2006; Savir et al., 2010; Flamholz et al., 2019). The specificity for CO_2 versus O_2 is described by the CO_2/O_2 specificity factor ($S_{C/O}$). This correlation seems to be reflected in the evolutionary context of the respective enzymes. RuBisCOs of C_3 plants, which experience a relatively high atmospheric O_2 concentration, show a higher $S_{C/O}$ factor than RuBisCOs of C_4 plants, cyanobacteria or anaerobic bacteria, which have elaborate CCMs in the form of concentrating CO_2 in the bundle sheath cells (C_4 plants) or carboxysomes (cyanobacteria) or live in an environment with no available O_2 (anaerobic bacteria).

From a mechanistic point of view, one hypothesis for the observed trade-off between $S_{C/O}$ and carboxylation turnover number (k_{cat}^c) is that enzymes with high $S_{C/O}$ (which also usually have a lower K_M for CO_2), bind the carboxylation transition state very tightly, which helps to discriminate between O_2 and CO_2. However, the transition state resembles

the subsequent carboxyketone intermediate, so that this in turn is also quite tightly bound. This strong binding slows the following cleavage into 3-PGA, and therefore total turnover number (Tcherkez et al., 2006). Another hypothesis claims that a higher catalytic efficiency for carboxylation (k_{cat}^{C}/K_M^{C}) comes unevitably with a higher catalytic efficiency for oxygenation (k_{cat}^{O}/K_M^{O}) because these two factors are strongly correlated. This is based on the assumption, that lowering the energetic barrier for CO_2 addition also lowers the barrier for O_2 addition (Savir et al., 2010; Flamholz et al., 2019). Furthermore, an additional trade-off between function and stability can be observed, limiting the evolution of RuBisCO within these contraints (Studer et al., 2014).

Further knowledge about the reaction mechanism and the evalutation of kinetic data of more different RuBisCOs is necessary to fully resolve the question, whether all existing RuBisCOs are subject to inherent catalytic restraints or if it might be possible, to evolve a RuBisCO with enhanced catalytic properties for carboxylation and at the same time diminished oxygenation activity. Up to now, efforts to evolve faster RuBisCOs with better discrimation between CO_2 and O_2 were of limited success (Mueller-Cajar et al., 2007; Shih et al., 2016) and might potentially never be successful, considering the above-mentioned limitations of the enzyme.

1.3 Photorespiration

The unfruitful oxygenation side reaction of RuBisCO leads to the formation of the highly toxic 2-PG, which is recycled through photorespiration into the CBB cycle intermediate 3-PGA (Figure 1). Photorespiration is a complex and energy-demanding process. It requires eleven different enzymes and cofactors like tetrahydrofolate and ferredoxin, ammonia is released and needs to be refixed, and as its main disadvantage, already fixed carbon is released in form of CO_2 (Bauwe et al., 2010). In higher plants, in addition to that, the intermediates of the pathway are transported between cell compartments. Glycolate is formed within the chloroplast and then transported out into the peroxisome, where it is transformed into glycine. Glycine is then transported into the mitochondria and two molecules of glycine are converted into serine under the release of CO_2. Serine, in turn, is transported back to the peroxisome, where its conversion into glycerate takes place. Glycerate is then transported into the chloroplast and phosphorylated to form 3-phosphoglycerate, which can be fed back into the CBB cycle.

In addition to this C2 cycle found in higher plants, cyanobacteria metabolize glycolate via the glycerate pathway as well as via a complete decarboxylation pathway (Eisenhut et al., 2006; Eisenhut et al., 2008).

Photorespiration coevolved together with oxygenic photosynthesis and is essential for all organisms performing oxygenic photosynthesis, including organisms with efficient CCMs like C_4 plants, eukaryotic algae, and cyanobacteria (Nakamura et al., 2005; Eisenhut et al., 2008; Zelitch et al., 2009; Hagemann et al., 2016). The fact that even organisms with CCMs require photorespiration underlines the importance of fast and efficient removal of even trace amounts of 2-PG for a functioning photosynthesis and central carbon metabolism.

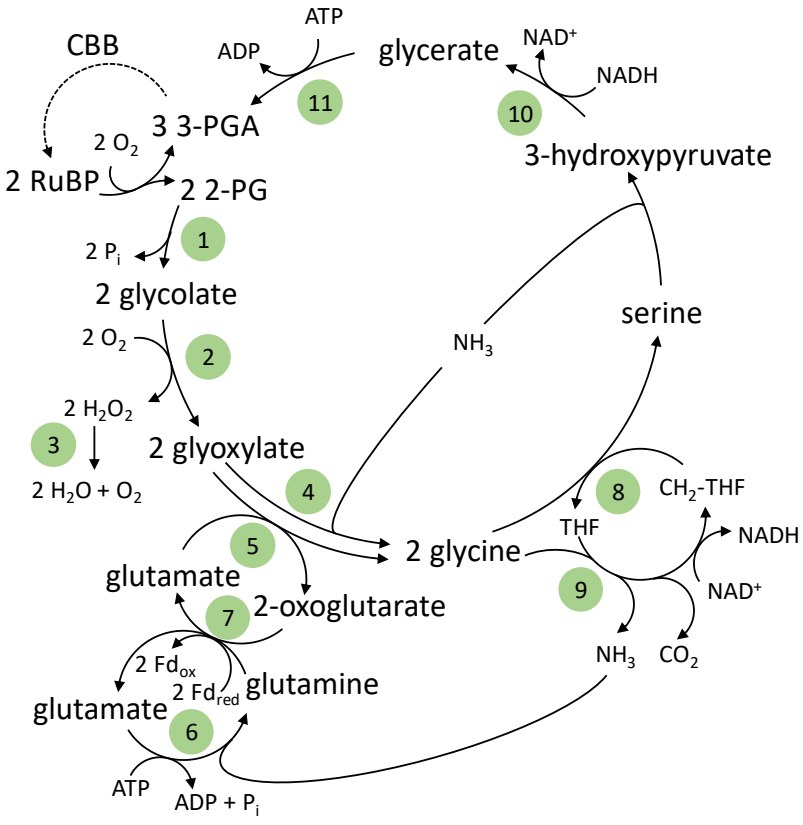


Figure 1. Photorespiratory pathway in higher plants.
1: phosphoglycolate phosphatase, 2: glycolate oxidase, 3: catalase, 4: serine/glyoxylate aminotransferase, 5: glutamate/glyoxylate aminotransferase, 6: glutamine synthetase, 7: glutamine/oxoglutarate aminotransferase, 8: serine hydroxymethyltransferase, 9: glycine decarboxylase, 10: hydroxypyruvate reductase, 11: glycerate kinase. CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 3-PGA: 3-phosphoglycerate, 2-PG: 2-phosphoglycolate, Fd_{ox}: oxidized ferredoxin, Fd_{red}: reduced ferredoxin, THF: tetrahydrofolate

The metabolic intertwining of photorespiration with central carbon metabolism is manifold. It is involved in regulatory interactions with the CBB cycle, photoinhibition, and metabolite shuttles, and furthermore it plays a significant role in C_1- and N-metabolism, like the synthesis of serine and glycine (Hagemann et al., 2016; Fernie and Bauwe, 2020).

Photorespiration is one of the main factors limiting agricultural productivity today. It was estimated that the rate of photorespiratory CO_2 release under current atmospheric conditions is 26 % of the rate of net CO_2 assimilation, meaning that for every five carboxylation reactions of RuBisCO approximately two oxygenation reactions take place (Sharkey, 1988; Walker et al., 2016). Furthermore, approximately one-third of the total ATP and NADPH produced by a photosynthetic leaf is consumed in photorespiration (Walker et al., 2016; Weber and Bar-Even, 2019). In the future, the oxygenation reaction of RuBisCO will eventually play an even more significant role in agricultural productivity, because a temperature increase of 3°C will lead to a 10 % decrease in specificity for CO_2 in RuBisCO (Peterhansel and Maurino, 2011). Models predict, that yields of certain species (e.g. soybean) might decrease by up to 50 % due to the climate change (Walker et al., 2016), which would be a direct impact of photorespiration.

1.4 Synthetic biology approaches for improvement of agricultural productivity

In order to feed approximately 10 billion people in 2050, crop yields have to increase significantly (Hickey et al., 2019). Several approaches have been proposed and worked on to improve agricultural productivity with biotechnology and synthetic biology (Ort et al., 2015; Weber and Bar-Even, 2019; Wurtzel et al., 2019). As already discussed above, engineering of a more efficient RuBisCO has not been successful so far and might actually never be achieved, due to several constraints (see chapter 1.2). Therefore, other attempts are needed that circumvent the inefficiencies of RuBisCOs. Certainly one of the most challenging ones is the C_4 rice project that started in 2008 (www.c4rice.com). It aims at converting the C_3 plant rice (*Oryza sativa*) into a C_4 plant, resulting in less photorespiration and therefore expectedly higher yields of this worldwide important crop. In brief, the transformation of a C_3 into a C_4 plant requires two major changes, which are of anatomical and biochemical nature. The anatomical part requires the introduction of the so-called Kranz anatomy, which is typical for C_4 plants. Here, the leaf vascular

bundles show a certain distribution pattern, which differs from C_3 to C_4 plants, as well as the distribution of specialized cells around the vascular bundles, which are located as concentric rings of mesophyll and bundle sheath cells around the veines (Sedelnikova et al., 2018). A major advance has been made recently by achieving a proto-Kranz anatomy in *O. sativa* by constitutive expression of GOLDEN2-like transcription factors from maize (*Zea mays*). In addition, most of the biochemical factors determining C_4 biochemistry are known by now and several sets of enzymes have been tested in *O. sativa* (Ermakova et al., 2020). Although these achievements are encouraging that a true C_4 rice might be possible in the future, there is still a long way to go (Sedelnikova et al., 2018; Ermakova et al., 2020).

Other attempts aim at introducing algal and cyanobacterial CCMs into higher plants. It has been shown, that ten essential components of *Chlamydomonas reinhardtii* pyrenoids can be expressed in *Arabidopsis thaliana* and that they localize at the relavant compartments (Atkinson et al., 2016). In a similar approach, a cyanobacterial carboxysome was functionally introduced into tobacco (*Nicotiana tabacum*) including the replacement of the endogenous RuBisCO with the cyanobacterial one. The encapsulation of the heterologous RuBisCO into carboxysomes *in planta* was demonstrated, as well as the growth of the plants under elevated CO_2 concentrations (Lin et al., 2014; Occhialini et al., 2016; Long et al., 2018).

More radical approaches even suggest the replacement of the complete CBB cycle with alternative carbon fixation pathways (Bar-Even et al., 2010; Schwander et al., 2016). One of these pathways is the crotonyl-coenzyme A (CoA)/ethylmalonyl-CoA/hydroxy-butyryl-CoA (CETCH) cycle, which was realized *in vitro* and performs at a similar rate as the CBB cycle (Schwander et al., 2016). Furthermore, the engineering of a new form of RuBisCO on a completely new scaffold has been proposed: “neo-RuBisCO”. This new (to be engineered) enzyme would not show the constraints of natural RuBisCOs and could potentially involve a complete new reaction mechanism without an oxygenation side reaction (Ort et al., 2015; Erb and Zarzycki, 2018).

1.5 The tartronyl-CoA pathway

Another approach to improve crop yields is the development of alternative, more energy-efficient pathways for photorespiration. Several pathways have been suggested and have

been tested *in vitro* and *in vivo*. These include the introduction of the glycerate pathway into *A. thaliana*, which led to faster growth and production of more biomass (Kebeish et al., 2007) or the introduction of a glycolate oxidation pathway into *N. tabacum* (South et al., 2019), which also had the effect of increased biomass in plants. Other photorespiratory bypasses and their implementations are discussed in detail in chapter 3.

New photorespiratory bypasses could be derived from computational approaches as described in Trudeau et al. (2018). The systematic search for photosynthetic bypasses included all reactions starting from 2-PG and leading to a CBB cycle intermediate, that are possible if general biochemical reaction rules are considered. These included for example the activation of an organic acid to its corresponding phosphoanhydride or acyl-CoA ester or reduction of a carbonyl to a hydroxycarbon. All possible reactions were considered, even if they led to the formation of new-to-nature chemical compounds. Avoided were compounds, that are highly reactive (e.g. two carbonyl groups adjacent to each other) or compounds that contain both a phosphate and a CoA group. Furthermore, only oxygen-tolerant reactions were considered. The most important prerequisite, however, was that none of the reactions release CO_2. The computational approach yielded a number of different photorespirational bypasses, including the tartronyl-CoA pathway (Figure 2). It consists of six reactions (from 2-PG to 3-PGA), of which most are not known to occur in nature. First, 2-PG is dephosphorylated to yield glycolate, which is subsequently activated to glycolyl-CoA, its corresponding CoA ester. Glycolyl-CoA is carboxylated to form tartronyl-CoA, which represents the key reaction of the tartronyl-CoA pathway. Due to the incorporation of CO_2 in this step, the tartronyl-CoA pathway becomes a carbon-positive photorespiratory bypass. Tartronyl-CoA is then reduced in a two-step reduction via tartronic semialdehyde to glycerate. The phosphorylation of glycerate yields 3-phosphoglycerate, which is a CBB cycle intermediate and closes the cycle (Figure 2). The realization of the tartronyl-CoA pathway is challenging, because several reactions are non-natural. While the dephosphorylation of 2-PG and phosphorylation of 3-PGA are also part of the natural photorespiration pathway (Bauwe et al., 2010), all other reactions, except for the reduction of tartronic semialdehyde (Gotto and Kornberg, 1961), represent synthetic reactions. This means, that for the realization of the tartronyl-CoA pathway it was necessary to identify enzymes, that are potentially promiscuous for the non-natural substrates of the tartronyl-CoA pathway and that are able to catalyze the required reactions.

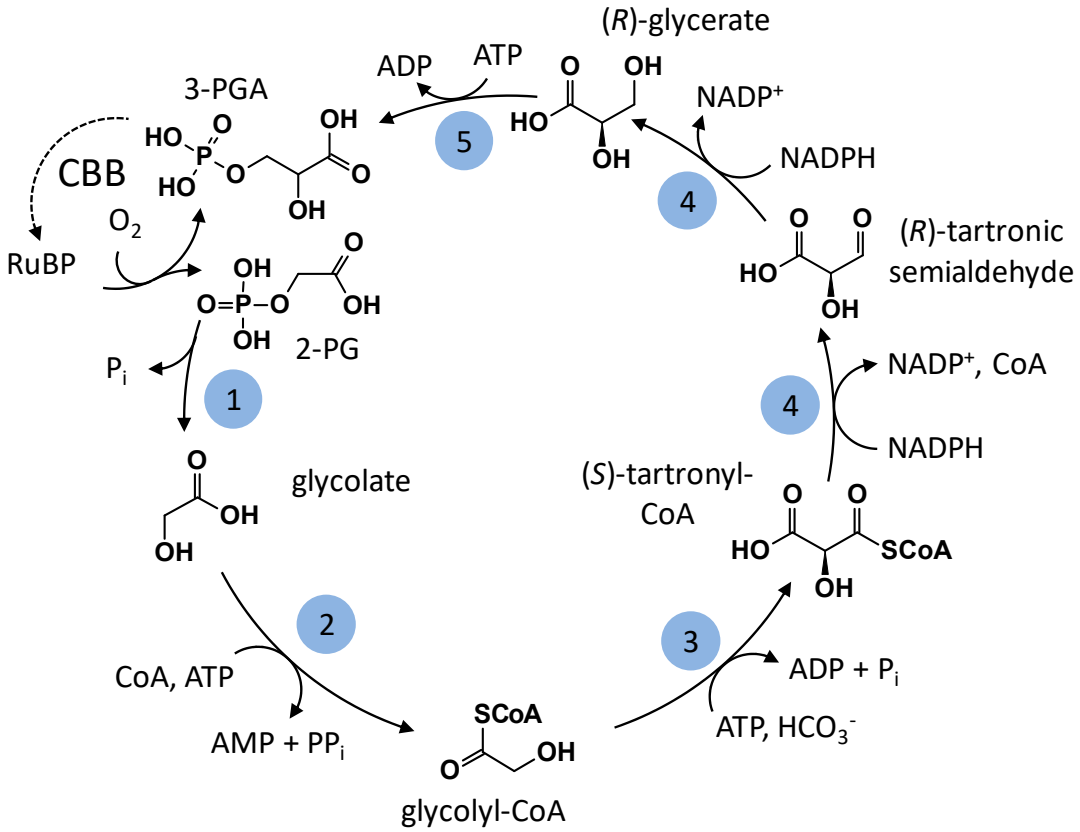


Figure 2. The tartronyl-CoA pathway.
2-phoshpoglycolate (2-PG), RuBisCO's oxygenation product, is dephosphorylated to give glycolate. The latter is activated to its corresponding CoA ester, glycolyl-CoA, which is carboxylated to form tartronyl-CoA. In a two-step reduction, glycerate is formed and subsequently phosphorylated to give 3-phosphoglycerate (3-PGA), which can be fed back into the Calvin (CBB) cycle.
1: phoshphoglycolate phosphatase, 2: glycolyl-CoA synthetase (GCS), 3: glycolyl-CoA carboxylase (GCC), 4: tartronyl-CoA reductase (TCR), 5: glycerate kinase. CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 3-PGA: 3-phosphoglycerate, 2-PG: 2-phosphoglycolate.

1.6 Biotin-dependent carboxylases

As a suitable starting point for the realization of glycolyl-CoA carboxylation, propionyl-CoA carboxylases (PCCs) were considered, because of the structural similarity of glycolyl-CoA and propionyl-CoA. PCCs belong to the family of biotin-dependent carboxylases, which are distributed among all domains of life. They catalyze the carboxylation at the α-position of saturated CoA esters, namely acetyl-CoA (ACCs), propionyl-CoA (PCCs) or long-chain acyl-CoAs (LCCs). Geranyl-CoA carboxylases and 3-methylcrotonyl-CoA carboxylases add the carboxyl group at the γ-position of of α-β unsaturated CoA esters. Furthermore, two classes of biotin-dependent carboxylases are

known, which carboxylate pyruvate at the β-position (pyruvate carboxylases) or urea at one of the two N-atoms (urea carboxylases). Biotin-dependent carboxylases play an important role in many cellular processes, e.g. biosynthesis of fatty acids (ACCs), metabolism of fatty acids and amino acids (PCCs), and the formation of anaplerotic oxaloacetate, which is an intermediate of the tricarboxylic acid (TCA) cycle and a gluconeogenesis precursor (pyruvate carboxylase) (Tong, 2013, 2017).

The overall structure of biotin-dependent saturated acyl-CoA carboxylases differs between the different types. They have a core of three β-subunit dimers, which are decorated by six α-subunits. In the case of PCCs, the β-subunit core forms a short cylinder, which is decorated with three monomeric α-subunits at each end of the cylinder (Huang et al., 2010). Usually, biotin-dependent carboxylases consist of three different domains, of which two are catalytically active. The biotin cofactor is covalently bound to a conserved lysine residue at the non-catalytic biotin carboxyl carrier protein (BCCP) domain. The biotin is carboxylated by the biotin carboxylase (BC; α) domain, which is dependent on MgATP and bicarbonate as the CO_2 donor. Bicarbonate is deprotonated by a glutamate residue and subsequently acts as a nucleophile to attack the γ-phosphate of ATP. The emerging carboxyphosphate intermediate decomposes to CO_2 and phosphate and the phosphate deprotonates the biotin. The biotin enolate intermediate then attacks the CO_2 which leads to the formation of the carboxybiotin (Chou et al., 2009). The transfer of the carboxybiotin from the BC to the carboxyltransferase (CT, β) domain is not completely understood yet. Most probably, the BCCP undergoes a vast conformational change, leading to the translocation of the whole BCCP domain from the BC domain to the CT domain (swinging-domain model) (Tong, 2013). The mechanism of carboxylation of the acyl-CoA at the CT domain is poorly understood as well. The most common hypothesis is that CO_2 is released from the biotin within the active site of the CT domain. The biotin itself then acts as a base and abstracts a proton from the α-carbon of the acyl-CoA. The emerging enolate intermediate attacks the CO_2, which leads to the formation of the carboxylated acyl-CoA product (Knowles, 1989; Attwood and Wallace, 2002).

1.7 Aims of this thesis

The central objective of the present work was the realization of the tartronyl-CoA pathway. As pointed out, the theoretical pathway consists of several reactions that are not known to exist in nature. Therefore, initially the tartronyl-CoA pathway only existed "on paper". In order to achieve its realization, the aims of this thesis can be subdivided into the following points:

i. The enzymes for the catalysis of three different non-natural biochemical reactions had to be identified. These were the conversion of glycolate into glycolyl-CoA, the carboxylation of glycolyl-CoA to form tartronyl-CoA, and the reduction of tartronyl-CoA to give tartronic semialdehyde, and finally glycerate as a second reduction step. As a means for the identification of the required enzymatic activities, databases (e.g. KEGG, BRENDA or UniProt) were to be searched for enzymes performing similar reactions with similar substrates. A number of candidate enzymes for each reaction was to be selected and kinetically characterized. To this end, appropriate activity assays should be established for each enzyme, as well as protocols for the synthesis of the pathway intermediates glycolyl-CoA and tartronyl-CoA. Additionally, after the identification of enzyme canditates able to perform the required reactions to a certain extent, the enzymes were to be engineered towards higher catalytic efficiencies. There was a particular focus on the design and evolution of the carboxylase, as it represents the key enzyme of the tartronyl-CoA pathway.

ii. After the successful identification and engineering of enzyme candidates for the tartronyl-CoA pathway, an additional goal of the present work was to test these enzymes in concert to demonstrate the pathway's functionality as photorespiratory bypass *in vitro*. This included the initial assembly of all enzymatic components and the subsequent optimization of reaction conditions, along with the optimization of cofactor regeneration systems. Furthermore, the performance of the tartronyl-CoA pathway in the context of other applications and metabolic pathways was to be demonstrated, as it represents a new metabolic pathway for the direct conversion of glycolate into glycerate without carbon loss (and even being carbon-positive). For the conversion of the plastic waste component ethylene glycol (EG), another metabolic module should be developed for the conversion of EG into glycolyl-CoA (the entry point to the tartronyl-CoA pathway). In addition, the

tartronyl-CoA pathway was to be tested in conjuction with the synthetic CETCH cycle for carbon fixation.

iii. In order to harness the full potential of the tartronyl-CoA pathway, its implementation into living systems will be required. One aim of this thesis was the initial construction of an *in vivo* system of the tartronyl-CoA pathway for the further evolution of its enzymatic components and for the application of the tartronyl-CoA pathway as EG conversion module. Furthermore, the gene sequences of the (evolved) enzymes of the tartronyl-CoA pathway were to be handed over to the collaboration partners at the MPI of Molecular Plant Physiology (Potsdam-Golm, Germany) and at Evogene (Rehovot, Israel) for their implementation into *Escherichia coli* and *A. thaliana*, respectively.

iv. The mechanistic understanding as well as the rational engineering of enzymes requires the availability of molecular structures of these enzymes or close homologs. Therefore, one of the goals of this thesis was to obtain structures of glycolyl-CoA carboxylase (GCC) and tartronyl-CoA reductase (TCR).

2. Results

2.1 Finding the enzymes for the tartronyl-CoA pathway

For the realization of the tartronyl-CoA pathway it was necessary, to identify enzymes that could perform the required biochemical reactions for each step of the pathway. The general approach here was to search within literature or databases (e.g. BRENDA, KEGG or UniProt) for enzymes that perform analogous reactions with similar substrates and subsequently test these enzymes for their promiscuous activity with the desired substrate. When promising candidate enzymes were identified, they were heterologously produced in *E. coli* and purified using His- or Strep-tag affinity chromatography. The purity of the purified enzymes was routinely checked using SDS polyacrylamide gel electrophoresis (SDS-PAGE) (Figure S1). For each type of enzyme, suitable enzymatic assays were established and the enzymes were kinetically characterized with the desired non-natural substrates. Upon the identification of the required enzymatic activity, these enzymes could be further engineered if a crystal structure of close homologs was available.

2.1.1 Glycolyl-CoA synthetase

For the synthesis of glycolyl-CoA, the first step of the tartronyl-CoA pathway, two possible options were investigated: a CoA transfer from another acyl-CoA donor, as well as the direct ligation of glycolate to CoA. Three different transferases were screened with acetyl-CoA as the CoA donor and glycolate as CoA acceptor and the time-dependent formation of glycolyl-CoA was measured using ultra-high performance liquid chromatography coupled high resolution mass spectrometry (UPLC-hrMS). All of the tested enzymes catalyzed the activation of glycolate. The best candidate, a transferase from *Clostridium aminobutyricum* (AbfT), showed a catalytic efficiency of 129 M^{-1} s^{-1} (Table 1, Figure S2).

Table 1. Acyl-CoA transferases characterized in this study.

Abbr.	mutation(s)	Full name	Organism	v_{max} [nmol min^{-1} mg^{-1}]	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]
*Cp*PCT	n/a	propionate CoA transferase	*Clostridium propionicum*	32 ± 3	149 ± 35 (glycolate)	0.2
*Re*PCT	n/a	propionate CoA transferase	*Ralstonia eutropha*	1,300 ± 70	52 ± 10 (glycolate)	25
AbfT	n/a	4-OH-butyrate CoA transferase	*Clostridium aminobutyricum*	5,500 ± 1,500	34 ± 11 (glycolate)	129

n/a not applicable

In parallel, different acyl-CoA synthetases (ACSs) were tested for their activity with glycolate. In this reaction, free CoA is directly transfered to glycolate by hydrolyzing ATP to AMP and PP_i. The latter is usually directly cleaved in the cell, making this reaction virtually irreversible *in vivo*.

The enzyme activity was measured spectrophotometrically. The ATP hydrolysis reaction was measured coupled to ADP regeneration by myokinase and ATP regeneration by pyruvate kinase (PK) with phosphoenolpyruvate (PEP) and subsequent reduction of pyruvate to lactate by lactic dehydrogenase (LDH). The oxidation of NADH by LDH was followed spectrophotometrically at 340 nm. Among others, two ACSs of *Erythrobacter* sp. NAP1 were tested for their activity with glycolate. The best one of these two was *Ery*ACS1 with a catalytic efficiency of 20 M^{-1} s^{-1} (Table 2, Figure S3).

Table 2. Acyl-CoA synthetases characterized in this study.

Abbr.	mutation(s)	Full name	Organism	v_{max} [nmol min^{-1} mg^{-1}]	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]
ACS19*,[1]	Trudeau et al. (2018)	acetyl-CoA synthetase	*Escherichia coli*	4,670 ± 650	50 ± 15	74
*Ery*PCS_lig	n/a	PCS ligase domain	*Erythrobacter* sp. NAP1	192 ± 14	252 ± 45 (glycolate)	1.3
*Ery*ACS2	n/a	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	311 ± 16	92 ± 12 (glycolate)	4
*Ery*ACS1	n/a	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	450 ± 10	30 ± 2 (glycolate)	20
*Ery*ACS1*	n/a	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	16,320 ± 1,116	35 ± 8 (glycolate)	540
*Ery*ACS1_VA_LP	V379A, L641P	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	984 ± 54	300 ± 30 (glycolate)	4
_Ery_ACS1_VA*, GCS	**V379A**	**acyl-CoA synthetase**	***Erythrobacter* sp. NAP1**	**9,111 ± 508**	**13 ± 3 (glycolate)**	**837**
*Ery*ACS1_VG*	V379G	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	2,666 ± 111	35 ± 5 (glycolate)	89
*Ery*ACS1_VS*	V379S	acyl-CoA synthetase	*Erythrobacter* sp. NAP1	n.d.	n/a	n/a

The enzyme highlighted in bold (GCS) represents the candidate used for further *in vitro* reconstruction of the tartronyl-CoA pathway. n/a not applicable, n.d. not detectable, PCS propionyl-CoA synthetase, [1]enzyme published in Trudeau et al. (2018) and kinetics shown were measured in the present study, *enzyme expressed in *E. coli* BL21 AI (DE3) Δ*patZ*.

As part of the physiological regulation of acetyl-CoA synthetases, these enzymes get acetylated and thereby inactivated *in vivo* (Starai et al., 2002). As a result, the ACSs heterologously produced in *E. coli* were (partially) acetylated and therefore not fully active. It has been reported that acetylation of the *Salmonella* acetyl-CoA synthetase was prevented by mutating the surface loop residue L641 to a proline (Starai et al., 2005). The effect of this mutation was tested for the *Erythrobacter* ACS and indeed, a 2-times higher specific activity was detected, but also a dramatically increased app. K_M for glycolate of 300 ± 30 mM (Table 2). Therefore, other possibilities were investigated to prevent deleterious acetylation of ACS. In *E. coli*, which serves as the production strain for the synthetases, the main acetylase is encoded by the *patZ* gene (Starai and Escalante-Semerena, 2004; de Diego Puente et al., 2015). To prevent possible inactivation of ACS directly upon heterologous expression in *E. coli*, a Δ*patZ* expression strain (BL21 AI (DE3) Δ*patZ*) for overexpression of ACS was constructed. When produced in this acetylase knockout strain, *Ery*ACS1 showed an almost 40-times higher specific activity of 16,320 ± 1,116 nmol min^{-1} mg^{-1} than the enzyme expressed in *E. coli* BL21 AI (DE3)

(Table 2). Compared to another glycolyl-CoA synthetase that was recently described (Trudeau et al., 2018) and had a catalytic efficiency of 74 M^{-1} s^{-1} (Table 2), *Ery*ACS1 showed a more than 7-times higher catalytic efficiency of 540 M^{-1} s^{-1} and was therefore chosen for further enzyme engineering. A homology model was created using acetyl-CoA synthetase of *Salmonella enterica* (PDB 2P2B, 62 % sequence identity, (Arnold et al., 2006; Reger et al., 2007)) as template. The desired substrate, glycolate, is slightly bigger and more polar than acetate. Therefore, for a following design approach, it was attempted to open up the active site pocket slightly, to allow for better accommodation of glycolate, and residue V379 was identified as a target for directed mutagenesis (Figure 3).

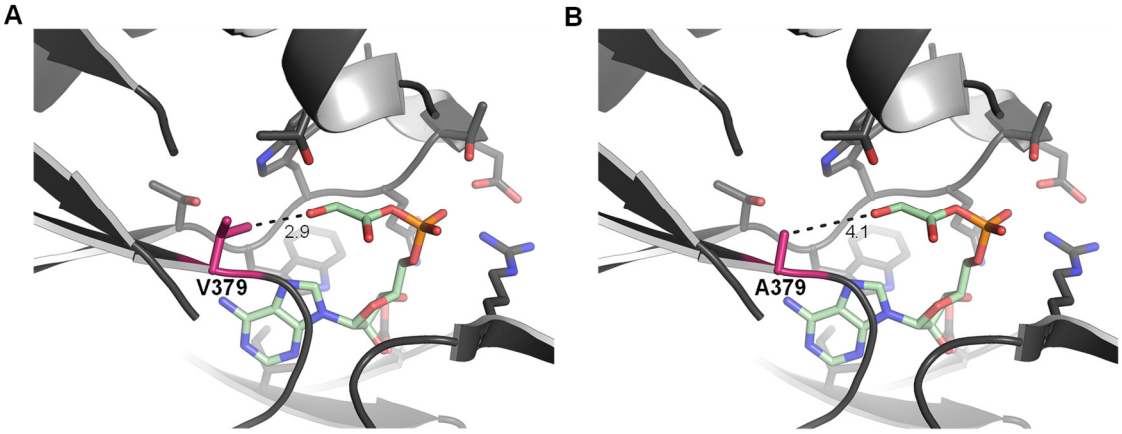


Figure 3. Active site homology model for acetyl-CoA synthetase 1 of *Erythrobacter* sp. NAP1.
The homology model is based on the crystal structure of an acetyl-CoA synthetase from *S. enterica* (PDB 2P2B). The residue targeted for mutagenesis is shown in pink. The reaction intermediate glycolyl-AMP was modeled into the active site based on the position of an adenosine-5'-monophosphate-propyl ester (mimicking the acetyl-AMP reaction intermediate of *S. enterica* acetyl-CoA synthetase) bound in PDB 2P2B. Distances are given in Å. (**A**) In the wildtype *Ery*ACS1 the hydroxyl group of the glycolyl moiety would be in an unfavorable close contact to the non-polar V379. (**B**) In the V379A mutant of *Ery*ACS1 the active site provides more space to presumably better accommodate glycolyl-AMP.

Three different V379 substitutions (Table 2) were tested. The V379A variant showed improved kinetic parameters for the desired reaction with an app. K_M for glycolate of 13 ± 3 mM and a catalytic efficiency of 837 M^{-1} s^{-1} (Table 2). This enzyme will be referred to as glycolyl-CoA synthetase (GCS) from here on.

The best tested transferase for glycolyl-CoA synthesis (AbfT, Table 1) and GCS (Table 2) were tested under realistic physiological conditions regarding substrate concentrations in the plant chloroplast. It was assumed that the concentration of acetyl-CoA is around 50 µM, while the concentration of free CoA is around 5 µM. Both enzymes were tested

with 2 mM of glycolate and the corresponding concentrations of acetyl-CoA and free CoA. Under these conditions tested, the synthetase showed a specific activity of 280 nmol min^{-1} mg^{-1} and thus performed twice as good as the transferase, which showed a specific activity of 140 nmol min^{-1} mg^{-1}. Based on the more favorable kinetic parameters, as well as thermodynamic considerations (irreversibility of the reaction) and independence of other acyl-CoA pools *in vivo*, it was decided to further rely on the engineered GCS for glycolyl-CoA synthesis.

2.1.2 Tartronyl-CoA reductase

The final steps of the tartronyl-CoA pathway are the reduction reactions of tartronyl-CoA to tartronic semialdehyde and subsequently to glycerate. This last reaction of the tartronyl-CoA pathway, the reduction of tartronic semialdehyde, is the only reaction that is found in natural metabolic pathways. It is part of the glycerate pathway and the conversion of glycolate, where tartronic semialdehyde is the product of glyoxylate carboligase (Krakow et al., 1961; Hansen and Hayashi, 1962). For the formation of tartronic semialdehyde from tartronyl-CoA, a tartronyl-CoA reductase was required. In order to perform enzymatic measurements of tartronyl-CoA reductases, the substrate tartronyl-CoA needed to be synthesized. The malonyl-CoA synthetase MatB from *Rhodopseudomonas palustris* (Vagstad et al., 2012; Peter et al., 2016) was tested for activity with tartronate, which was ~ 230 nmol min^{-1} mg^{-1} at a tartronate concentration of 10 mM and proved to be sufficient to synthesize tartronyl-CoA for subsequent purification via preparative high performance liquid chromatography coupled mass spectrometry (HPLC-MS).

The promiscuous activity of different reductases with tartronyl-CoA was measured spectrophotometrically, following the direct oxidation of NADPH at 340 nm. The malonyl-CoA reductase (MCR) of *Sulfolobus tokodaii* (Alber et al., 2006) performs the one-step reduction of malonyl-CoA to malonic semialdehyde and showed a promiscuous specific activity of 140 nmol min^{-1} mg^{-1} with 1 mM tartronyl-CoA (Table 3).

Table 3. Acyl-CoA reductases characterized in this study.

Abbr.	mutation(s)	Full name	Organism	v_{max} [nmol min^{-1} mg^{-}	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]
*St*MCR	n/a	malonyl-CoA reductase	*Sulfolobus tokodaii*	140[1]	n/a	n/a
*Ery*MCR	n/a	malonyl-CoA reductase	*Erythrobacter* sp. NAP1	230 ± 13	0.18 ± 0.04 (tartronyl-CoA)	2,787
_Ca_MCR, TCR	**n/a**	**tartronyl-CoA reductase**	***Chloroflexus aurantiacus***	**635 ± 14**	**0.026 ± 0.003 (tartronyl-CoA)**	**53,745**

The enzyme highlighted in bold (TCR) represents the candidate used for further *in vitro* reconstruction of the tartronyl-CoA pathway. n/a not applicable, [1]measured at 1 mM tartronyl-CoA

In parallel, also bifunctional MCRs that perform both reduction steps from malonyl-CoA to malonic semialdehyde as well as the subsequent reduction to 3-hydroxypropionate, were tested for their activities with tartronyl-CoA (Table 3, Figure S4). The homolog employed in the 3-hydroxypropionate bicycle from *Chloroflexus aurantiacus* (Hügler et al., 2002) showed promiscuous activity with tartronyl-CoA at 635 nmol min^{-1} mg^{-1} and a very favorable app. K_M of 26 µM and was chosen for further use in the tartronyl-CoA pathway.

2.1.2.1 Structure of malonyl-CoA/tartronyl-CoA reductase

The rational engineering of MCR towards higher activity with its non-natural substrate tartronyl-CoA and lower activity with its natural substrate malonyl-CoA required a structure of this enzyme to identify the residues that are involved in binding of the CoA thioester substrate. Initially, there were no structures of this enzyme or close homologs available, making rational enzyme engineering not feasible. With the intention of producing protein crystals of *Ca*MCR, several thousand of different crystallization conditions were screened. After two years of incubation at 4 °C, protein crystals could be observed in two different conditions. The crystals were soaked with the phasing compound Tb-Xo4 (Engilberge et al., 2017) to enable the use of the single-wavelength anomalous scattering (SAD) method. These crystals diffracted to a resolution of 2.9 Å (Table 4) and exhibited an anomalous signal to about 3.5 Å.

Table 4. Data collection, refinement, and model statistics for crystal structure of C-terminal domain of *Ca*MCR.

	MCR with $NADP^+$, malonate	MCR with Tb-Xo4 (SAD data)
Data collection		
Space group	*C* 1 2 1	*C* 1 2 1
Cell dimensions		
a, *b*, *c* (Å)	84.76, 141.71, 73.40	87.68, 141.13, 74.18
α, β, γ (°)	90.00, 96.81, 90.00	90.00, 98.04, 90.00
Resolution (Å)	48.92 - 2.34 (2.46 - 2.34)	55.52 – 2.90 (3.07 – 2.90)
R_{merge}	0.095 (0.405)	0.179 (0.743)
$I / \sigma I$	8.5 (2.7)	7.0 (2.2)
$CC_{1/2}$ (%)	99.3 (86.9)	98.2 (69.0)
Completeness (%)	98.3 (93.5)	99.2 (97.5) / ano. 96.0 (93.7)
Redundancy	3.6 (3.3)	5.2 (4.9) / ano. 2.6 (2.5)
Refinement		Not fully refined
No. unique reflections	35689 (3403)	NA
R_{work} / R_{free}	0.1899 / 0.2083	NA
No. atoms	5221	NA
Protein	4996	NA
Ligands	80	NA
Water	145	NA
Ramachandran (%)		
favored	98.76	NA
allowed	1.24	NA
outliers	0.00	NA
B-factors		
Protein	43.70	NA
Ligands	56.90	NA
Water	44.09	NA
R.m.s. deviations		
Bond lengths (Å)	0.004	NA
Bond angles (°)	0.64	NA

A preliminary structure could be obtained, which only comprised the C-terminal part of the enzyme starting at W571, probably due to degradation of the N-terminal part during the 2-year incubation of the plate before crystallization took place.

The preliminary model was used to solve a native crystal structure with molecular replacement, leading to a model of the C-terminal domain of *Ca*MCR at 2.34 Å (Figure 4, Table 4).

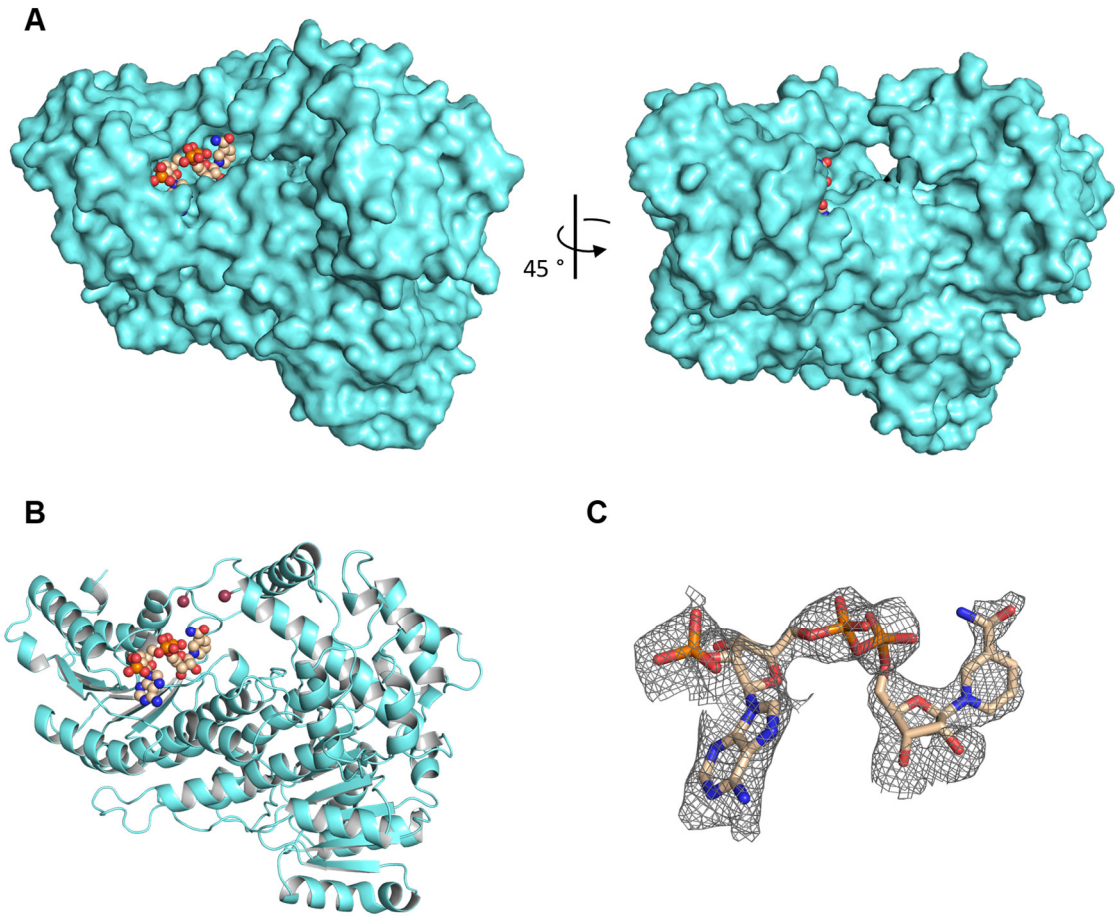


Figure 4. Crystal structure of C-terminal domain of *Ca*MCR/TCR.
(**A**) Surface representation with bound $NADP^+$. In the 45° rotational view, a potential substrate channel for the CoA thioester is observable. (**B**) Secondary structure representation. The purple spheres mark the start and end of an eleven amino acid stretch, which is missing in the structure. (**C**) Simulated annealing F_o-F_c omit map at 2σ of $NADP^+$.

$NADP^+$ was bound in a cavity close to the N-terminus of the structure (Figure 4). In the immediate proximity to $NADP^+$, a channel could be observed, which might serve as the binding site of the CoA thioester (Figure 4). The C-terminal domain is responsible for the

first half reaction of the reduction, the conversion of the CoA thioester to its corresponding semialdehyde (Liu et al., 2013). For efficient electron transfer, it can be expected that the CoA thioester is positioned close to the NADP cosubstrate in the enzyme. In a very recently published structure of the C-terminal domain of MCR of *Porphyrobacter dokdonensis* (46 % sequence identity with *Ca*MCR) with bound free CoA (PDB 6K8T) (Son et al., 2020), the CoA is positioned within the channel that is also present in *Ca*MCR. This supports the hypothesis, that this is also the site of CoA binding in *Ca*MCR.

However, a stretch of eleven amino acids (between G780 and L792) could not be modeled in the final *Ca*MCR structure, because electron density for these residues was missing. Interestingly, their position right next to the NADP cofactor implied they might be part of a very flexible region that potentially closes the active site during catalysis as part of an 'induced fit' system (Figure 4). It could be speculated, that a conformational change takes place only if both cosubstrates (NADPH and CoA thioester) are bound, allowing subsequent catalysis. This missing stretch is present in the above-mentioned structure from *P. dokdonensis* (PDB 6K8T) as well as in another structure from the same publication with bound $NADP^+$ (PDB 6K8U) (Son et al., 2020). However, all structures only have either $NADP^+$ or free CoA bound, respectively, and none of the structures have the actual CoA thioester substrate bound. Most probably, the possible function of the eleven amino acid stretch in CoA thioester coordination would only become clear in case a structure with both cosubstrates would be available.

In conclusion, the structure of the C-terminal domain of *Ca*MCR does not allow the prediction of amino acid residues important for binding of malony-CoA/tartronyl-CoA. Therefore, further studies are necessary in order to obtain a crystal structure with both $NADP^+$ and a CoA thioester bound, as it is expected that the binding of both substrates leads to a conformational change (probably including the eleven amino acid stretch missing in the structure). A structure with both cosubstrates bound would probably enable the identification of important amino acid residues for binding of the CoA ester substrate and therefore enable the rational engineering of *Ca*MCR towards a true TCR. It would be preferable, if this structure would cover the whole enzyme, as it was not possible so far, to obtain a structure of MCR, containing both C- and N-terminal domain (Son et al., 2020). Besides enabling rational engineering, such structure would provide insights into

the reaction mechanism of MCR. It involves the formation of a highly reactive aldehyde intermediate, which might be internally channeled from the active site of the C-domain to the active site of the N-domain, where the second reduction step takes place.

2.1.3 Glycolyl-CoA carboxylase

The key reaction of the tartronyl-CoA pathway is the carboxylation of glycolyl-CoA to tartronyl-CoA. Because of the structural similarity between glycolyl-CoA and propionyl-CoA, four different biotin-dependent PCCs were screened for substrate promiscuity (Table 5). In order to develop an enzymatic assay for glycolyl-CoA carboxylation, it was necessary to synthesize the substrate glycolyl-CoA. To this end, glycolate was activated using carbonyldiimidazole (CDI) for the subsequent direct thioesterification of the CDI-activated carboxylic acid with free CoA (Kawaguchi et al., 1981; Peter et al., 2016) and glycolyl-CoA was purified using preparative HPLC-MS.

Activity for glycolyl-CoA carboxylation was determined spectrophotometrically using two different assays. The ATP hydrolysis reaction was measured coupled to ATP regeneration by PK with PEP and subsequent reduction of pyruvate to lactate by LDH. The oxidation of NADH by LDH was followed spectrophotometrically at 340 nm.

To measure carboxylation activity, spectrophotometric assays at 365 nm were performed coupling the product formation (tartronyl-CoA) to its subsequent reduction via tartronic semialdehyde to glycerate by use of TCR. This coupling reaction was analogous to an activity assay for ACC (Kroeger et al., 2011).

For determination of kinetic parameters, the time-dependent formation of tartronyl-CoA was analyzed using UPLC-hrMS. This allowed the side-by-side comparability of kinetic parameters obtained for glycolyl-CoA and propionyl-CoA, because no simple spectrophotometric read-out for methylmalonyl-CoA (the carboxylation product of PCC) was available.

Of the tested carboxylases, only PCC from *Methylorubrum extorquens* AM1 (*Me*PCC) had minuscule activity with glycolyl-CoA as assessed by UPLC-hrMS measurements of tartronyl-CoA after overnight incubation of the enzymatic assay. Over the course of working with biotin-dependent carboxylases, it became obvious that for full activity it was necessary to coexpress the cognate biotin ligase gene (*birA*) of *M. extorquens* (results in collaboration with Pascal Pfister). Coexpression of this cognate biotin ligase gene resulted in very low but measurable specific activity of 6 ± 0.4 nmol min^{-1} mg^{-1} (k_{cat} = 0.01 s^{-s}) for *Me*PCC with glycolyl-CoA (Table 5, Figure S5).

Table 5. Acyl-CoA carboxylases characterized in this study.

Abbr.	mutation(s)	Organism	v_{max} [nmol min^{-1} mg^{-1}]	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]
*Ca*PCC1	n/a	*Chloroflexus aurantiacus*	not detectable	n/a	n/a
*Ca*PCC2	n/a	*Chloroflexus aurantiacus*	not detectable	n/a	n/a
*Ery*PCC	n/a	*Erythrobacter* sp. NAP1	not detectable	n/a	n/a
*Me*PCC‡	n/a	*Methylorubrum extorquens* AM1	6 ± 0.4	0.33 ± 0.06 (gCoA)	41
*Me*PCC_DI	D407I	*Methylorubrum extorquens* AM1	30†	n/a	n/a
*Me*PCC_DK	D407K	*Methylorubrum extorquens* AM1	instable complex	n/a	n/a
*Me*PCC_GA	G170A	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_GS	G170S	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_GT	G170T	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_LD	L140D	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_LN	L140N	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_YE	Y143E	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
*Me*PCC_YH	Y143H	*Methylorubrum extorquens* AM1	20	n/a	n/a
*Me*PCC_YQ	Y143Q	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
GCC M2‡	Y143H, D407I	*Methylorubrum extorquens* AM1	257 ± 5	0.33 ± 0.02 (gCoA)	1,687
GCC_M2_LD	L100D, Y143H, D407I	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
GCC_M2_LN	L100N, Y143H, D407I	*Methylorubrum extorquens* AM1	41	n/a	n/a
GCC M3‡	L100S, Y143H, D407I	*Methylorubrum extorquens* AM1	416 ± 28	0.39 ± 0.07 (gCoA)	2,310
GCC_M3_DA	L100S, Y143H, D171A, D407I	*Methylorubrum extorquens* AM1	instable complex	n/a	n/a
GCC_M3_DV	L100S, Y143H, D171V, D407I	*Methylorubrum extorquens* AM1	not detectable	n/a	n/a
GCC M4‡	**L100S, Y143H, D407I, I450V**	***Methylorubrum extorquens* AM1**	**1,841 ± 117**	**0.35 ± 0.06 (gCoA)**	**11,387**
GCC M5‡	**L100S, Y143H, D407I, I450V, W502R**	***Methylorubrum extorquens* AM1**	**2,590 ± 130**	**0.15 ± 0.03 (gCoA)**	**36,275**

The enzymes highlighted in bold represent the candidates used for further *in vitro* reconstruction of the tartronyl-CoA pathway. Given specific activities are carboxylation activities with glycolyl-CoA measured coupled to product (tartronyl-CoA) reduction by tartronyl-CoA reductase (TCR). n/a not applicable, gCoA glycolyl-CoA, †measured PK/LDH coupled, ‡coexpressed with *M. extorquens* biotin ligase birA.

It could be observed that the specific activity of product formation as assessed using the coupled spectrophotometric assay with TCR was always much lower than the specific activity measured using a PK/LDH coupled assay that follows ADP formation. Depletion

assays using the TCR-coupled assay and limiting amounts of ATP showed that the carboxylation of glycolyl-CoA consumed 102 ± 16 ATP per carboxylation in *Me*PCC WT. This futile ATP hydrolysis as well as the low specific carboxylation activity of *Me*PCC required the engineering of this enzyme towards higher carboxylation and less futile ATP hydrolysis rates.

2.1.3.2 Rational engineering of glycolyl-CoA carboxylase

In order to engineer *Me*PCC towards a GCC, in a first approach, structure-guided rational design was applied. Based on a homology model of *Me*PCC several residues were targeted in the first shell of the active site of the CT subunit (L140, Y143, G170, D171, D407 Figure 5A, Table 5) towards accommodation of glycolyl-CoA.

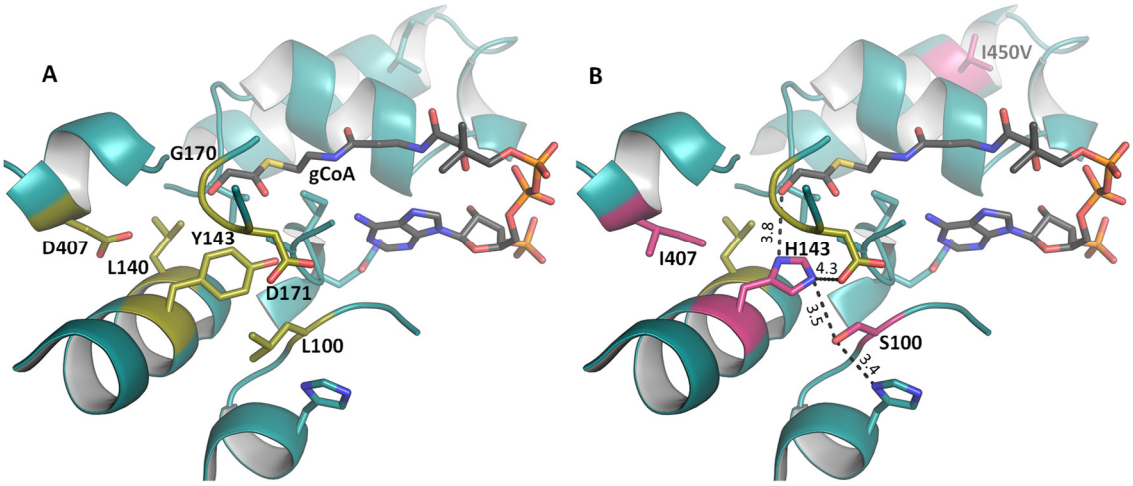


Figure 5. Active site homology models for propionyl-CoA carboxylase of *M. extorquens* (*Me*PCC, A) and variants 4 and 5 of engineered glycolyl-CoA carboxylase (GCC M4 & M5, B).
The active site homology models are based on PDB 3N6R. Glycolyl-CoA (gCoA) was modeled according to the position of methylmalonyl-CoA in PDB 1ON3. Residues targeted during rational design of the active site are colored in yellow. Actual active site changes in the GCC M4 and M5 mutants are colored in pink. Potential new hydrogen bonds are depicted as dashed lines. Distances are given in Å. The Y143H mutation is likely engaging in hydrogen bonding with the hydroxyl group of glycolyl-CoA. The L100S mutation is assumed to fix H143 in a rotamer conformation facilitating hydrogen bonding interaction with glycolyl-CoA. The I450V mutation is located in an α-helix nearby and may impact distances within the active site or at the interface with the biotin carboxylase subunit. In addition to the four mutations in and around the active site, GCC M5 features an additional mutation outside of the active site (W502R), which is located at the rotational symmetry of the enzyme oligomer.

A double mutant with a D407I and an Y143H mutation (GCC M2) showed a more than 40-fold improved specific activity with glycolyl-CoA of 257 ± 5 nmol min^{-1} mg^{-1} ($k_{cat} = 0.6$ s^{-1}, Table 5, Table 6, Figure S5). Furthermore, GCC M2 showed a much improved ratio of ATP hydrolysis:carboxylation of 8.3 ± 0.3. The Y143H mutation is likely engaging in hydrogen bonding with the hydroxyl group of glycolyl-CoA, while the D407I mutation likely directs the hydroxyl group of glycolyl-CoA towards H143.

For further improvement of the glycolyl-CoA carboxylation activity, the second shell around the active site was targeted at residue L100 (Figure 5, Table 5). A triple mutant (GCC M3) with an L100S substitution showed a further improved specific activity of 416 ± 28 nmol min^{-1} mg^{-1} (k_{cat} = 0.9 s^{-1}) and a further improved ATP hydrolysis to carboxylation ratio of 6.5 ± 0.4 (Table 5, Table 6, Figure S5). The L100S mutation was assumed to fix H143 in a certain rotamer conformation to facilitate the formation of a hydrogen bond network and strengthen the interaction of H143 with glycolyl-CoA.

At this stage, from a rational engineering point of view, there was not much more that could be improved, as mutating residues in the first and second shell of the active site had been exploited. Therefore, the next step was the design and high-throughput screening of libraries of GCC variants.

Table 6. Apparent kinetic parameters of different glycolyl-CoA carboxylase variants.

	k_{cat} [s^{-1}]	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]	ratio*	k_{cat} [s^{-1}]	app. K_M [mM]	k_{cat}/K_M [M^{-1} s^{-1}]
	glycolyl-CoA				propionyl-CoA		
WT	0.013 ± 0.001	0.33 ± 0.06	40.7	102 ± 16	26.2 ± 1.9	0.24 ± 0.04	1.09 x 10^5
M2	0.56 ± 0.01	0.33 ± 0.02	1.69 x 10^3	8.3 ± 0.3	0.099 ± 0.002	0.42 ± 0.03	2.36 x 10^2
M3	0.9 ± 0.1	0.39 ± 0.07	2.31 x 10^3	6.5 ± 0.4	0.23 ± 0.01	0.44 ± 0.06	5.17 x 10^2
M4	4.0 ± 0.3	0.35 ± 0.06	1.14 x 10^4	4.6 ± 0.2	1.12 ± 0.04	0.45 ± 0.05	2.49 x 10^3
M5	5.6 ± 0.3	0.15 ± 0.03	3.63 x 10^4	3.9 ± 0.1	11.9 ± 0.4	0.18 ± 0.02	6.63 x 10^4

*ATP hydrolysed per glycolyl-CoA carboxylation.

2.1.3.3 Directed evolution of glycolyl-CoA carboxylase variants

For further engineering of GCC, directed evolution using high-throughput approaches to screen for variants with further decreased futile ATP hydrolysis and higher glycolyl-CoA carboxylation activities was applied. The CT subunit of GCC M3 was diversified using error-prone PCR and libraries with mutation rates between 0.2 and 3.6 mutations/kbp were created. To assess whether the libraries contained enzymes with decreased futile ATP hydrolysis, a high-throughput microfluidic screen was developed to transfer the TCR-coupled spectrophotometric assay of GCC into picoliter-sized water-in-oil (w/o) droplets under ATP-limiting conditions (Figure 6).

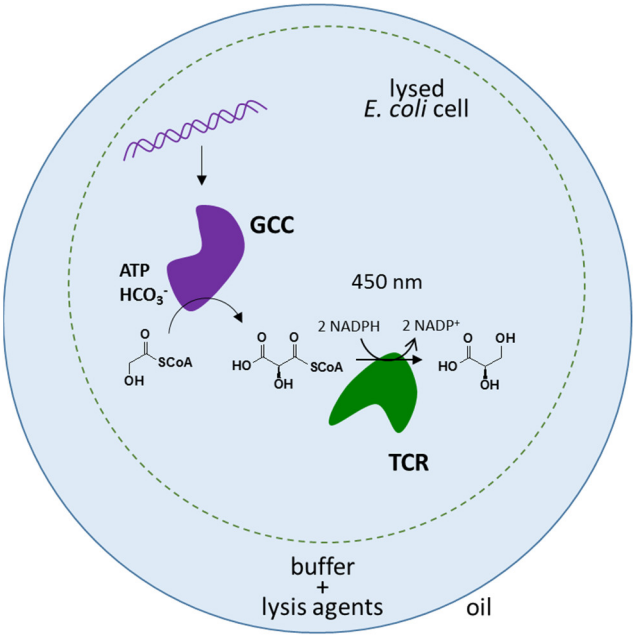


Figure 6. Scheme of microfluidic assay for activity measurements of glycolyl-CoA carboxylase (GCC).
Either purified GCC or single *E. coli* cells expressing variants of GCC were encapsulated in water-in-oil droplets. ATP, $KHCO_3$ and glycolyl-CoA were provided as substrates. The formation of the carboxylation product tartronyl-CoA was coupled to a decrease in NADPH fluorescence by its reduction to glycerate by the coupling enzyme TCR. In order to screen for GCC variants with decreased ATP hydrolysis to carboxylation ratios, the amount of ATP within the droplets was limited.

The NADPH fluorescence within the droplets could then be measured using a laser and photomultiplier set-up. To assure that the starting concentration of NADPH within the droplets was within the linear range, a calibration curve for NADPH was generated using eight different populations of droplets within the same solution that were identifiable by

different concentrations of the coding dye sulforhodamine B. It was found that a concentration of up to 1 mM NADPH was still within the linear range and was therefore chosen as the NADPH concentration for further experiments (Figure 7).

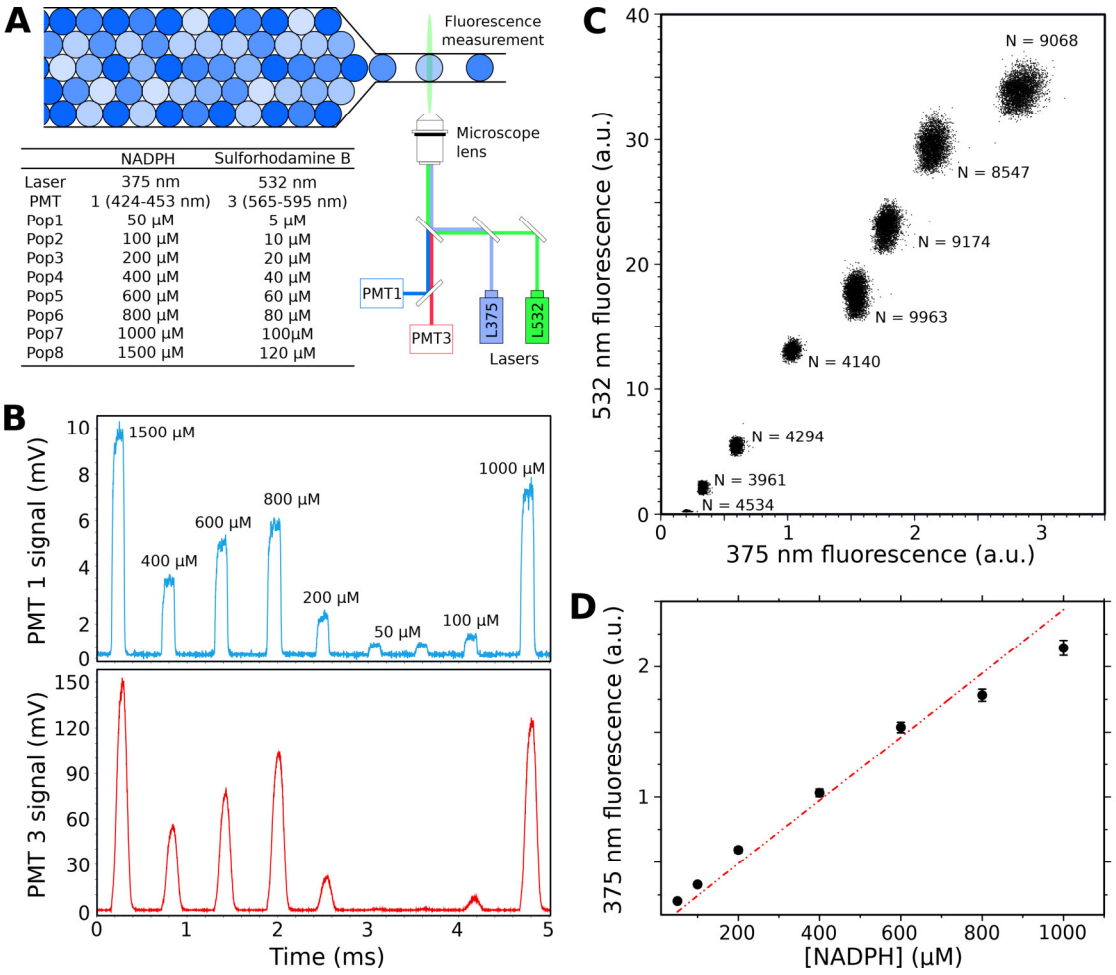


Figure 7. Microfluidic fluorescence measurements.
(**A**) Scheme of the microfluidic optical set up. Single droplet fluorescence was measured at kHz frequencies inside microchannels using the combination of two lasers (375 and 532 nm) with two photomultiplier tubes (PMT1 and PMT3). The optical set up allowed the detection of either NADPH fluorescence (enzymatic read-out) or sulforhodamine B fluorescence (barcoding). To make a NADPH fluorescence calibration, an 8-bit emulsion of water-in-oil (w/o) droplets containing different concentrations of NADPH and sulforhodamine B was produced and analyzed. (**B**) Time sequence of the signal of PMT1 (NADPH) and PMT3 (sulforhodamine B) for the corresponding 8-bit emulsion. (**C**) 2D plot showing the fluorescence distribution of the corresponding 8-bit emulsion. (**D**) NADPH concentration versus fluorescence level using the microfluidic set up.

To assure that the TCR-coupled GCC assay was working within droplets, experiments with purified GCC M3 were performed. The enzyme was encapsulated in 30 pL w/o droplets together with all necessary substrates (except for glycolyl-CoA) and the coupling enzyme. Glycolyl-CoA was injected into the droplets using picoinjection and the time-dependent decrease in NADPH fluorescence was followed during on-chip incubation (Figure 8).

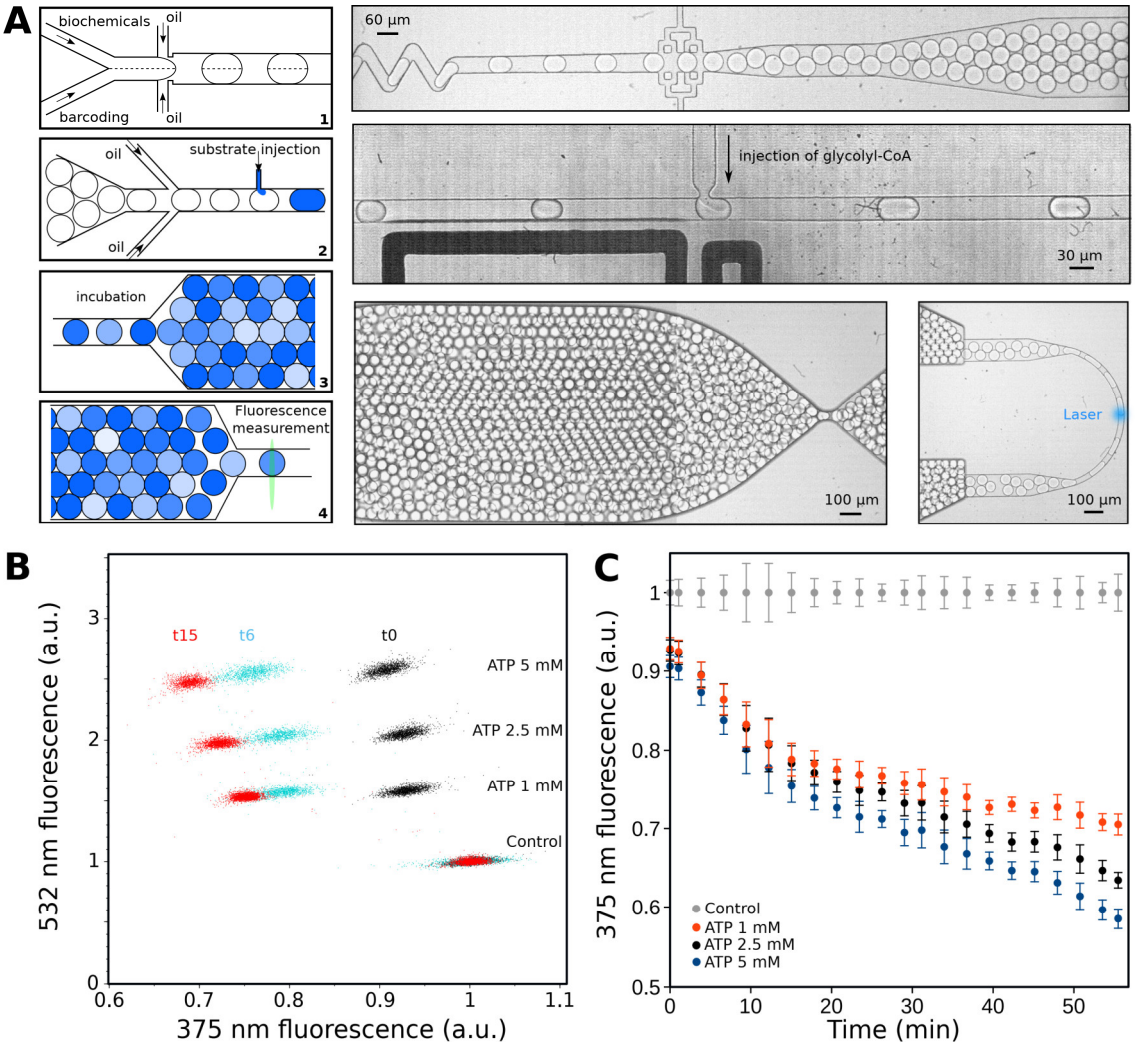


Figure 8. Microfluidic assay for activity measurements of variant 3 of glycolyl-CoA carboxylase (GCC M3) with on chip incubation.
(**A**) Microfluidic workflow: droplets were produced (1), collected off-chip and reloaded in a kinetic device (2). Glycolyl-CoA was picoinjected (2) and droplets were incubated on-chip (3). NADPH fluorescence was measured over time at different measurement loops (4). (**B**) 2D plot showing NADPH oxidation over time in four different populations of droplets. Except for population 1, all populations contained 53 µg mL^{-1} GCC M3. Population 1: 1 mM ATP, 30 µM sulforhodamine B (no enzyme); population 2: 1 mM ATP, 60 µM sulforhodamine B; population 3: 2.5 mM ATP, 90 µM sulforhodamine B; population 4: 5 mM ATP, 120 µM sulforhodamine B. (**C**) The NADPH oxidation over time was dependent on the ATP concentration within the droplets. Control: droplets without GCC M3. Data points represent mean ± SD (n=1500).

It was observed that the GCC assay was functional in w/o droplets as could be seen by the decrease in NADPH fluorescence caused by NADPH oxidation by TCR with tartronyl-CoA as a substrate (the carboxylation product of GCC). It was demonstrated as well, that the decrease in NADPH fluorescence signal was dependent on the ATP concentration (Figure 8B-C). This showed the potential of the microfluidic setup to screen for variants with a more efficient use of ATP for carboxylation than GCC M3. With the available picoinjection chips it was only possible to do the on-chip incubation in a reaction time of up to 1 hour. The enzymatic activity of GCC M3 was not high enough (especially in a single-cell assay where even less activity was expected due to a lower enzyme concentration in the droplets) to allow the use of the picoinjection method for the screening using single *E. coli* cells expressing enzyme variants. For further experiments it was therefore necessary to work with a two-step approach: (i) the encapsulation of single cells with lysis agents and all substrates, and (ii) the off-chip incubation of the encapsulated cells with subsequent reinjection and fluorescence measurement after a freely chosen incubation time.

The efficient lysis of *E. coli* cells within droplets was expected to be a limiting factor in single-cell microfluidics. Therefore, lysis conditions needed to be established first and several lysis conditions were tested. In the end, CelLytic™ B (Sigma-Aldrich) at a final dilution of 0.1x was determined to be the best method, as it did not interfere with the enzymatic assay and was able to efficiently lyse the cells as assessed by SDS-PAGE.

To establish the microfluidic assay for single lysed *E. coli* cells, experiments with cells expressing *Me*PCC WT or GCC M3 were performed. The cells were diluted to a final in-droplet OD_{600nm} of 0.05 to achieve a droplet occupancy by cells of 5 % to 10 %, assuring that droplets containing cells only contain a single cell (following a Poisson distribution). After droplet production, the droplets were incubated at 37 °C and the NADPH fluorescence in the droplets was measured after 1 h, 2 h, 3 h and 21 h. As expected, no enzyme activity was observed within the droplets containing cells expressing *Me*PCC WT, whereas GCC activity was measurable in the droplets containing GCC M3 expressing cells. This activity was already visible after 3 h of incubation, but the maximal activity could only be measured after an incubation time of 21 h (Figure 9). Therefore, all further experiments were conducted with an overnight incubation of droplets at 37 °C.

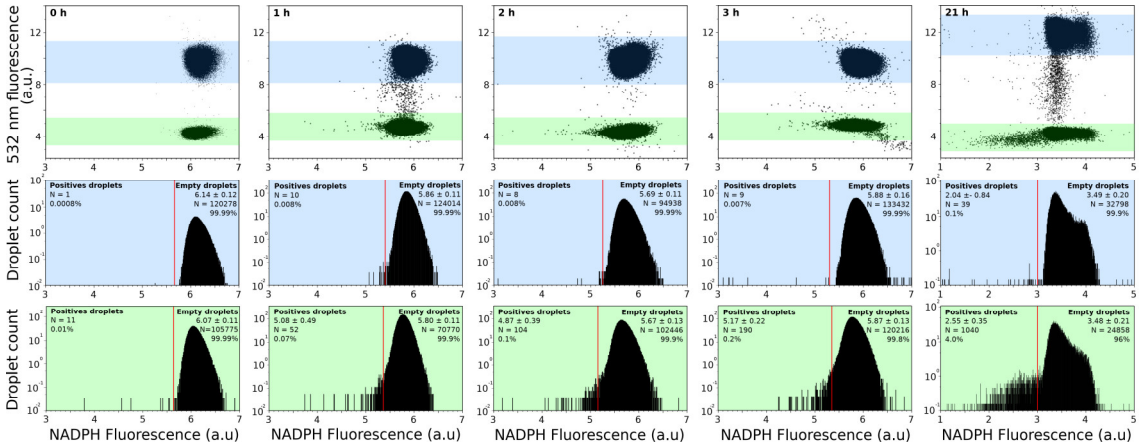


Figure 9. Microfluidic assay with droplets containing single cells.
E. coli cells expressing either *Me*PCC or GCC M3 were encapsulated into droplets with all assay components. The droplets with *Me*PCC contained 60 µM sulforhodamine B and the droplets with GCC M3 contained 20 µM sulforhodamine B as a barcoding. After encapsulation of single cells, the droplets were collected, incubated at 37 °C and fluorescence was measured directly after encapsulation, and 1 h, 2 h, 3 h and 21 h after encapsulation. 2D and 1D histograms show the fluorescence distribution of the 2-bit emulsion over time. The blue population corresponds to *Me*PCC expressing cells and the green population corresponds to GCC M3 expressing cells. The red line threshold gates negative and positive droplets and is defined as the mean value of the main population (empty droplets) minus 4 σ (standard deviation).

With the assay established in droplets with single *E. coli* cells, it was now possible to screen the libraries for more active GCC variants with less futile ATP hydrolysis using the microfluidic setup. Three libraries with diversified β-subunit of GCC with varying mutation rates between 0.2 and 3.6 mutations/kbp and a total size of 150,000 variants were transformed into *E. coli*. The cells were encapsulated next to a GCC M3 control (with different concentrations of the coding dye sulforhodamine B) under ATP limiting conditions and the NADPH fluorescence was measured after overnight incubation of the droplets. The analysis showed that out of the population of library cell droplets, 4.7 % exhibited a higher decrease in NADPH fluorescence than droplets containing GCC M3 cells. This implied that the libraries contained variants with a more efficient ATP usage than GCC M3 (Figure 10).

In an initial approach, the droplets in the orange gate (Figure 10A) were sorted using fluorescence-activated droplet sorting (FADS) (Baret et al., 2009) at a rate of 400 droplets s^{-1} and collected for plasmid recovery and identification of the phenotype corresponding genotypes.

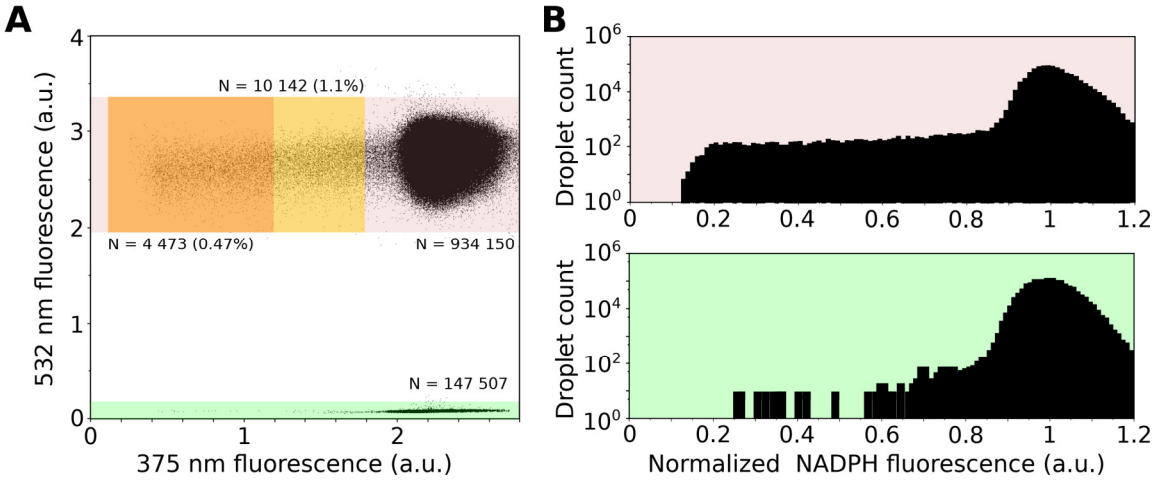


Figure 10. Microfluidic assay for cells expressing variants of glycolyl-CoA carboxylase (GCC).
E. coli cells expressing variants of GCC were encapsulated into water-in-oil droplets including lysis agents, reaction substrates and sulforhodamine B as coding dye (droplets with cells expressing GCC M3 2 µM and with cells expressing GCC variants 20 µM). The library contained variants that produce more carboxylation product (tartronyl-CoA) which led to more NADPH oxidation and therefore higher decrease in NADPH fluorescence signal. 2D (A) and 1D (B) histograms show the fluorescence distribution of the 2-bit emulsion after overnight incubation. The green population corresponds to GCC M3 cells and the pink population corresponds to cells expressing GCC variants. The yellow gate corresponds to active variants (defined as the mean of the main population minus 4 σ) and the orange gate corresponds to variants showing a higher activity than GCC M3.

For a recovery of plasmid DNA from the droplets it was necessary to find suitable PCR conditions including a highly sensitive DNA polymerase that was able to amplify small amounts of DNA. Over the course of one sorting experiment, approx. 6000 droplets from the sorting gate were collected, broken and the content released into 100 µL of water. The used plasmid for *gcc* expression (pCDFduet with CloDF13 replicon) has a copy number of approx. 30 per cell. This means that 180,000 plasmid copies are present in 100 µL, making 45,000 plasmids per one PCR template (25 µL). This corresponds to as little as 0.34 pg of DNA. Different polymerases were screened for their ability to amplify small amounts of DNA. The KAPA HiFi polymerase (Roche) was able to amplify even 0.1 pg of DNA (Figure 11) and was therefore used for nested PCRs on the sorted droplet contents. Unfortunately, it was not possible to amplify DNA out of the droplets under any condition tested, which was probably due to activity of nucleases from the cell lysate during the overnight incubation of droplets at 37 °C, degrading the plasmids present in the droplet content.

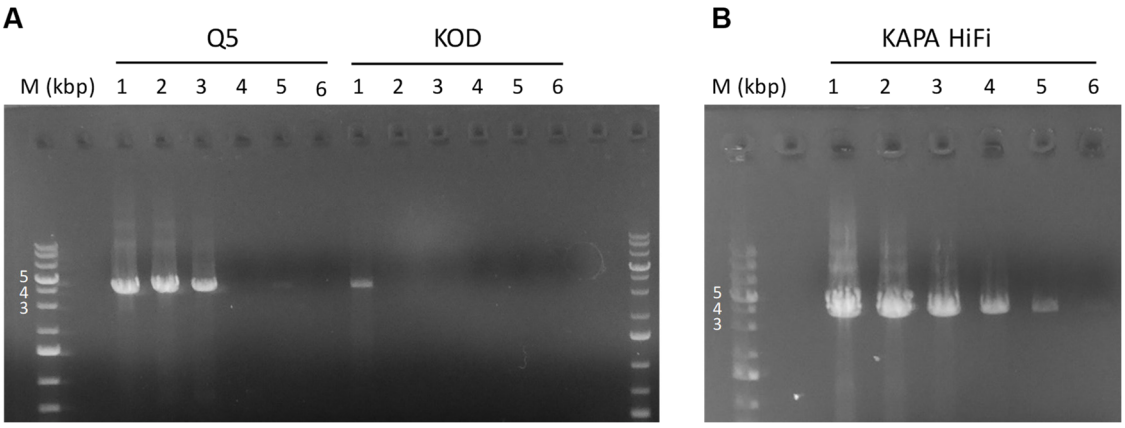


Figure 11. Optimization of PCR conditions for low template amounts.
(**A**) PCR with different template amounts with Q5 and KOD polymerase. (**B**) PCR with different template amounts with KAPA HiFi polymerase.
M: marker, 1: 10 ng, 2: 1 ng, 3: 100 pg, 4: 10 pg, 5: 1 pg, 6: 0.1 pg (template amounts).

With the knowledge that the created GCC libraries indeed contained better variants (Figure 10) it seemed feasible to screen the libraries separately in a microtiter plate screen. This approach had the additional benefit that time-dependent kinetics of single variants could be directly obtained. For the screen, it was intended to use the randomized libraries that had been assessed using microfluidics, as well as saturation mutagenesis libraries with reduced codon redundancy, saturated in the residues L140 and Y143 (Figure 12). As a quality control check, transformations of the saturation mutagenesis libraries were pooled and sequenced after plasmid preparation. The randomization of the desired triplets was confirmed by the presence of signals for all four bases in these triplets, one example is given in Figure 12A. For each of the four libraries, 1500 variants were screened to ensure 95 % coverage of the library. None of the variants showed a higher activity or less futile ATP hydrolysis than GCC M3. Interestingly, the double mutant GCC M2 was rediscovered in the saturation mutagenesis libraries, underscoring the validity of the employed rational design.

For the random mutagenesis libraries of GCC M3, in total 800 variants were screened from libraries 1_1 (3.6 mutations/kbp), 1_2 (1.1 mutations/kbp) and 1_3 (0.2 mutations per kbp). In one of the screens of library 1_2, one variant with highly increased initial slope (indicating faster carboxylation) and higher delta absorption (indicating less futile ATP hydrolysis) than the GCC M3 controls was identified (Figure 13A and Figure 13B). The enzyme was isolated, purified and subsequently characterized. It was found that this

variant, that was named GCC M4, had an almost 5-times higher catalytic efficiency and a 30 % decreased ATP hydrolysis per carboxylation rate than GCC M3 of 4.6 ± 0.2 (Table 6, Figure 14).

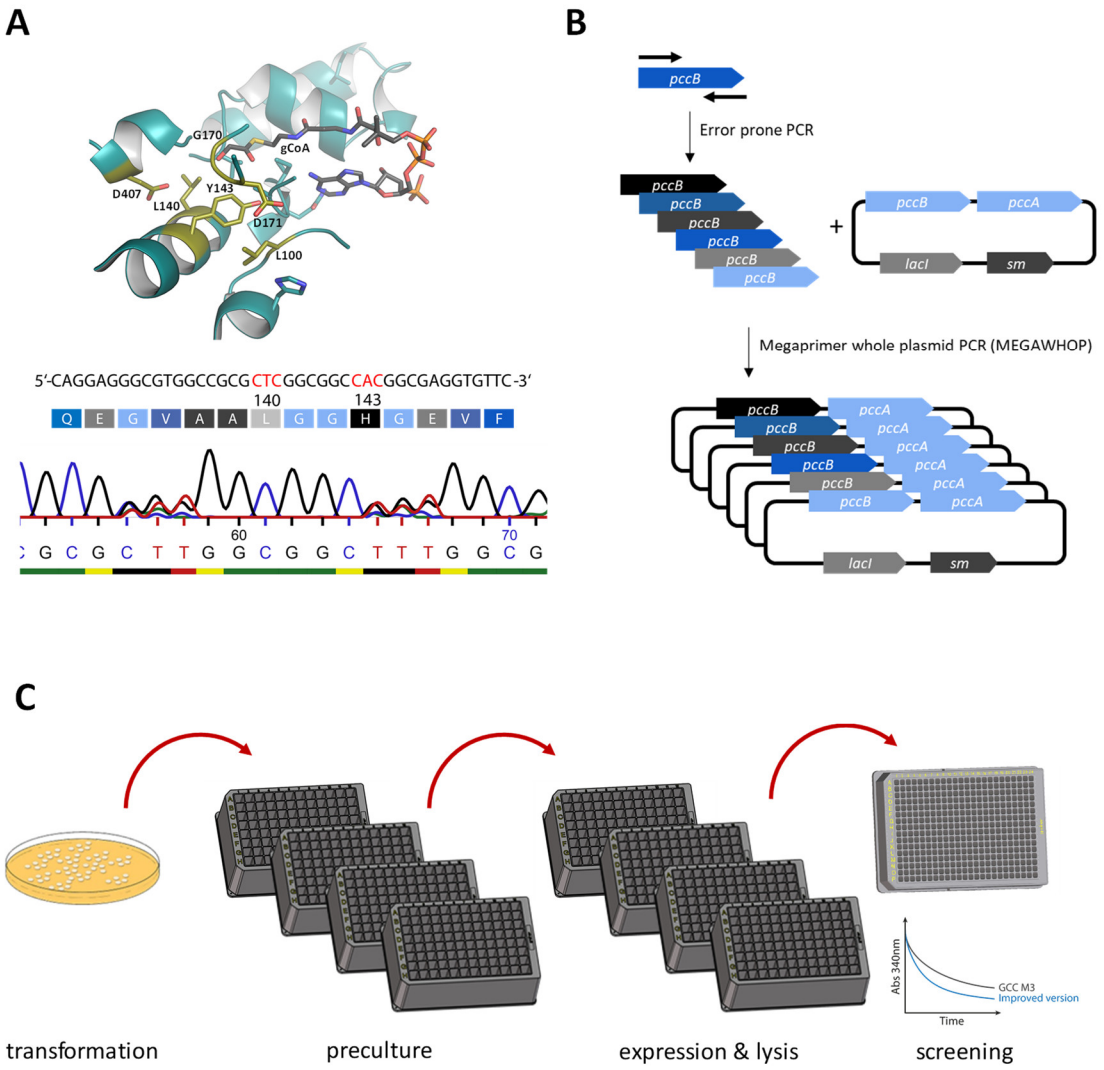


Figure 12. Creation and screening of GCC libraries.
(**A**) Saturation mutagenesis of GCC. The amino acid positions 140 and 143 were randomized by saturation mutagenesis applying a method for reduced codon redundancy (Kille et al., 2013) on four different templates, spanning different combinations of GCC M3 residues (L100_D407, L100_I407, S100_D407, S100_I407). Randomization of the desired codons was checked by a pooled plasmid preparation of the whole library. (**B**) Creation of randomized libraries of the β-subunit of GCC M3. The gene of the β-subunit of GCC M3 was amplified using error-prone PCR and subsequently reintroduced into the plasmid using megaprimer whole plasmid PCR (megaWHOP). (**C**) Scheme for microtiter plate screens of libraries. After transformation of the libraries into *E. coli*, single colonies were picked into precultures and propagated until lysis of the cells. Then 384-well plates were used for screening. (Modified from Daniel Marchal).

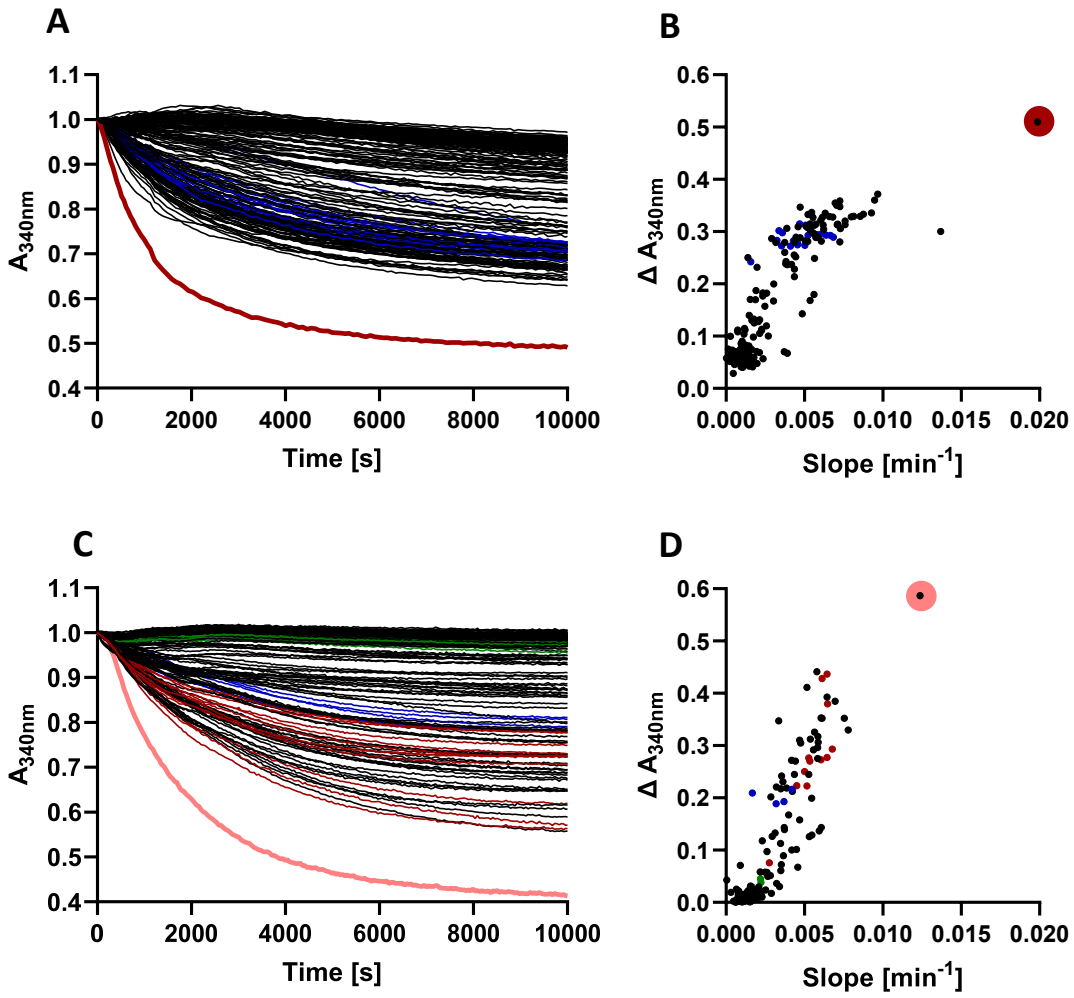


Figure 13. Microtiter plate screens for identification of GCC M4 and GCC M5.
(**A**) 384-well microtiter plate screen of random library 1_2 (GCC M3 variants) with 1.1 mutations per kbp. The time-dependent decrease in NADPH absorption of the identified variant GCC M4 is shown in red. The GCC M3 controls are shown in blue. (**B**) The overall delta absorbance of the first 10,000 s of the screen of library 1_2 (GCC M3 variants) is plotted against the initial slope of absorbance decrease during the first 500 s of the reaction. GCC M4 is shown in red, GCC M3 controls are shown in blue. (**C**) 384-well microtiter plate screen of random library 2_2 (GCC M4 variants) with 1.1 mutations per kbp. The time-dependent decrease in NADPH absorption of the identified variant GCC M5 is shown in salmon. The GCC M3 controls are shown in blue, the PCC WT controls are shown in green and the GCC M4 controls in red. (**D**) The overall delta absorbance of the first 10,000 s of the screen of library 2_2 (GCC M4 variants) is plotted against the initial slope of absorbance decrease during the first 500 s of the reaction. GCC M5 is shown in salmon, GCC M4 controls are shown in red, GCC M3 controls are shown in blue and PCC WT controls are shown in green.

For further evolution of GCC, a subsequent round of random mutagenesis on GCC M4 was performed. Two libraries with 0.2 mutations/kbp (library 2_1) and 1.1 mutations/kbp (library 2_2), respectively, were created with GCC M4 as a template. Of these libraries, a total number of 1600 variants was screened. In one screen of library 2_2, a variant with higher decrease in absorption and a steeper initial slope was identified (Figure 13C and Figure 13D). The variant (GCC M5) was isolated and characterized after purification, and showed a 3-times higher catalytic efficiency and a 15 % decreased ATP hydrolysis to carboxylation ratio of 3.9 ± 0.1 than GCC M4 (Figure 14, Table 6).

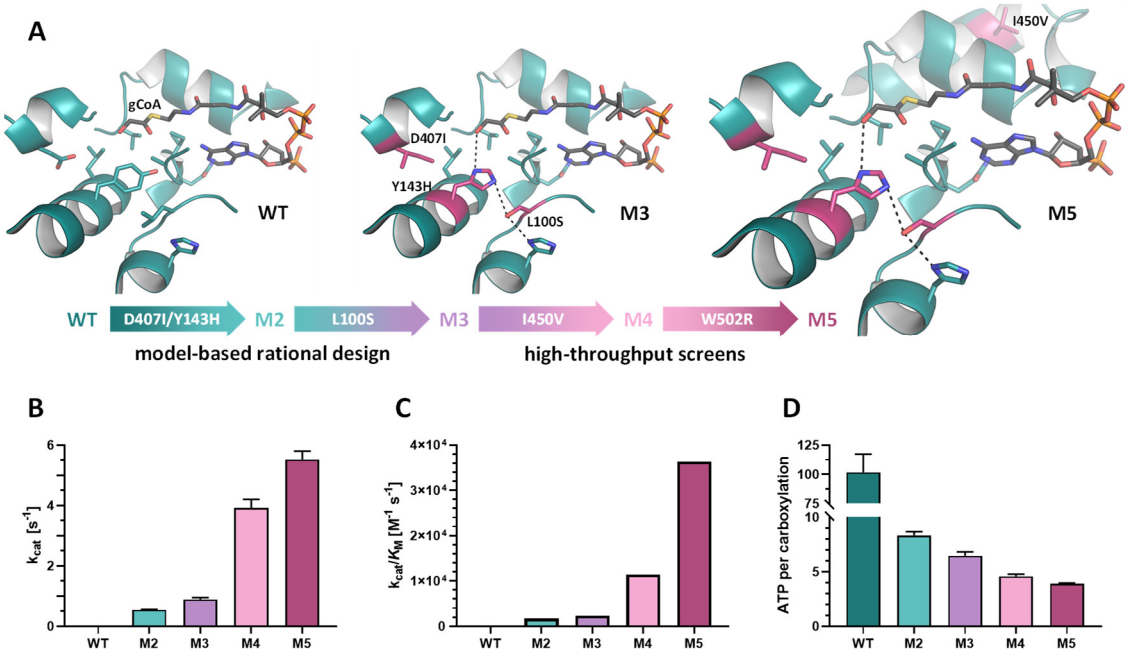


Figure 14. Rational Engineering and directed evolution of GCC leading to variants with higher catalytic efficiencies.
(**A**) Active site homology models of propionyl-CoA carboxylase of *M. extorquens* WT (based on PDB 3N6R) and engineered GCC variants. Variants M2 and M3 were obtained by rational design, M4 and M5 were obtained from subsequent rounds of random mutagenesis. (**B**) Glycolyl-CoA carboxylation activities of GCC variants. Data are presented as mean ± SD. (**C**) Catalytic efficiencies of GCC variants. Data represent mean calculated from results of nonlinear regression analysis. (**D**) Ratio of ATP hydrolysed per carboxylation reaction for MePCC WT and engineered GCC variants. Data represent mean ± SD.

To exclude effects caused by potential incomplete biotinylation of some of the variants, full biotinylation of variants was confirmed by an avidin-gel shift assay (Figure 15).

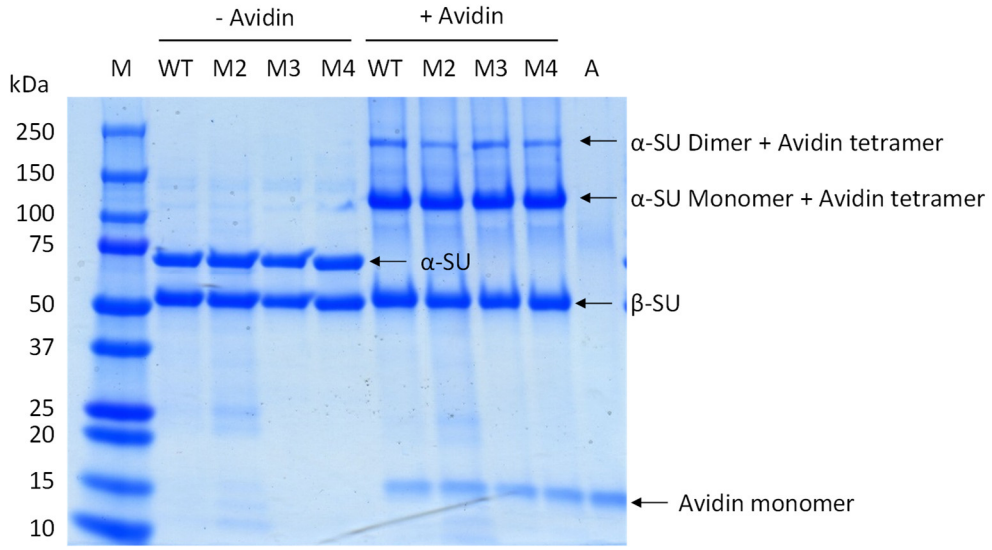


Figure 15. Avidin gel shift assay of to test for full biotinylation of propionyl-CoA carboxylase of *M. extorquens* (WT) and different engineered glycolyl-CoA carboxylase (GCC) variants.
M: molecular weight standard, WT: PCC *M. extorquens* wildtype, M2: PCC *M. extorquens* M2 variant, M3: PCC *M. extorquens* M3 variant, M4: PCC *M. extorquens* M4 variant, A: avidin, α-SU: α-subunit of PCC/GCC, β-SU: β-subunit of PCC/GCC. Theoretical molecular weight: α-SU 72 kDa, β-SU: 58 kDa, α-SU monomer + avidin tetramer 140 kDa, α-SU dimer + avidin tetramer 212 kDa, avidin monomer 17 kDa.

2.1.3.4 Structure of glycolyl-CoA carboxylase

For initial rational design of GCC, the structure of PCC from *M. extorquens* AM1 (*Me*PCC) was modeled (Figure 5) based on the structure of a chimeric PCC (PDB 3N6R) of the α-subunit of *Ruegeria pomeroyi* (*Rp*PCC) and the β-subunit of *Roseobacter denitrificans* (*Rd*PCC) (Huang et al., 2010). The α-subunits of *Me*PCC and *Rp*PCC share a 60 % sequence identity, while the β-subunits of *Me*PCC and *Rd*PCC have 72 % sequence identity. Despite of this homology present, it was of interest, to obtain a structure of *Me*PCC and the created mutants to corroborate the rational design of mutants and get insights into the effect of the two randomly introduced mutations of GCC M4 and M5 (I450V and W502R, respectively). First, it was intended to obtain crystal structures of *Me*PCC and GCC. Several thousand of different crystallization conditions were screened, but sufficiently diffracting protein crystals were never obtained.

Therefore, cryogenic electron microscopy (cryo-EM) was performed with *Me*PCC and GCC M5. We obtained a 3.48 Å resolution structure of *Me*PCC and a 1.96 Å resolution structure of GCC M5 (Figure 16, Figure 17, Figure 18, Table 7). The overall molecular architecture was as also found in the chimeric PCC and the human homolog (Huang et al., 2010). The enzyme consists of a cylindric hexameric core of β-subunits in a donut shape, which is surrounded by single α-subunits, three each at the top and bottom of the β-core. The β-core was well resolved in *Me*PCC as well as in GCC M5, while the decorating α-subunits were not completely resolved. Only the BCCP domain and another anchoring part of the α-subnunit were well resolved. This led to the assumption that the α-subunit might be quite flexible, which would be in accordance with the swinging-domain model of catalysis by PCC (Tong, 2013). It was postulated, that the whole BCCP domain moves during catalysis, potentially also involving a movement of the BC domain itself.

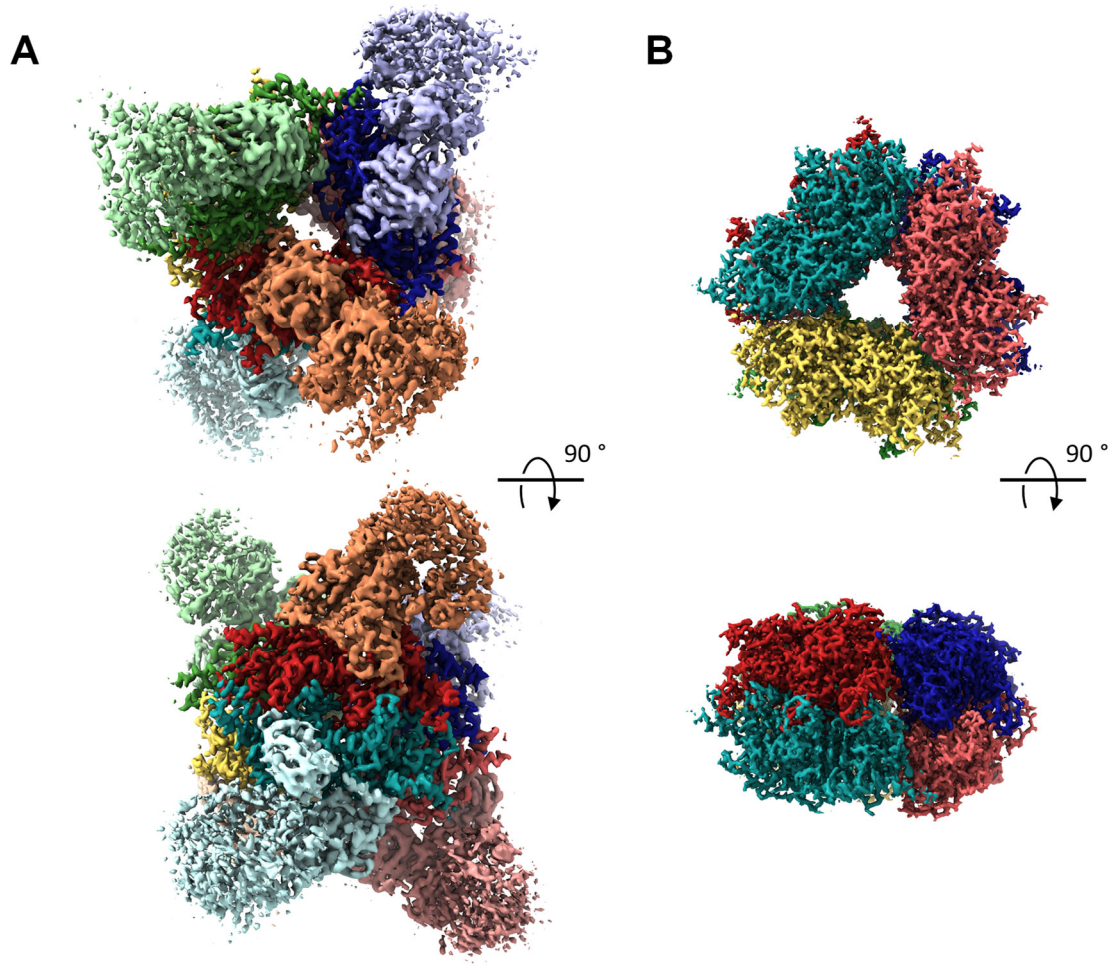


Figure 16. Cryogenic electron microscopy (cryo-EM) structure of *M. extorquens* PCC and GCC M5. (**A**) 3.48 Å resolution cryo-EM structure of holo propionyl-CoA carboxylase of *M. extorquens* WT. The hexameric β-core is surrounded by six single α-subunits (not completely resolved) (PDB 6YBP). (**B**) 1.96 Å resolution cryo-EM structure of the β-core of GCC M5 (PDB 6YBQ).

Table 7. Data collection, refinement, and model statistics for cryo-EM structures of *M. extorquens* PCC (*Me*PCC) and GCC M5.

	***Me*PCC**	**GCC M5**
Data collection		
Microscope	Glacios	Titan Krios G3
Camera	K2	K3 (CDS)
Magnification	36.000 x	105.000 x
Voltage (kV)	200	300
Electron exposure (e^{-}/Å^2)	50.2	55.0
Defocus range (μm)	0.5-3.5	0.5-2.7
Pixel size (Å)	1.18	0.851
Image processing		
Symmetry imposed	C1	C1
Initial particle images (no.)	168.488	5.398.283
Final particle images (no.)	77.714	2.181.317
Applied B-factor (Å^2)	-86.6	-71.7
Final resolution (Å)	3.48	1.96
Refinement Statistics		
Modeling software	COOT, Phenix	COOT, Phenix
Protein residues	4702	4289
Ligands	6× CoA, 6× biotin	6× CoA, 6× biotin
Water	-	871
Map CC (volume)	0.83	0.86
RMS deviations		
Bond lengths (Å)	0.008	0.007
Bond angles (°)	0.761	0.972
ADP (B-factors)		
min / max / mean		
Protein	50.92 / 184.73 / 99.60	3.64 / 71.26 / 24.93
Ligand	102.25 / 136.29 / 124.46	30.41 / 40.40 / 36.62
Water	-	5.86 / 29.25 / 11.77
Ramachandran plot (%)		
Outliers	0.00	0.00
Allowed	6.36	2.70
Favored	93.64	97.30
Rotamer outliers (%)	0.00	0.00
MolProbity score	1.85	1.11
Clash score	7.95	2.07
Resolution (Å) at FSC of 0.143		
masked / unmasked	3.44 / 3.48	1.94 / 1.95
PDB Accession	6YBP	6YBQ
EMDB Accession	EMD-10770	EMD-10771

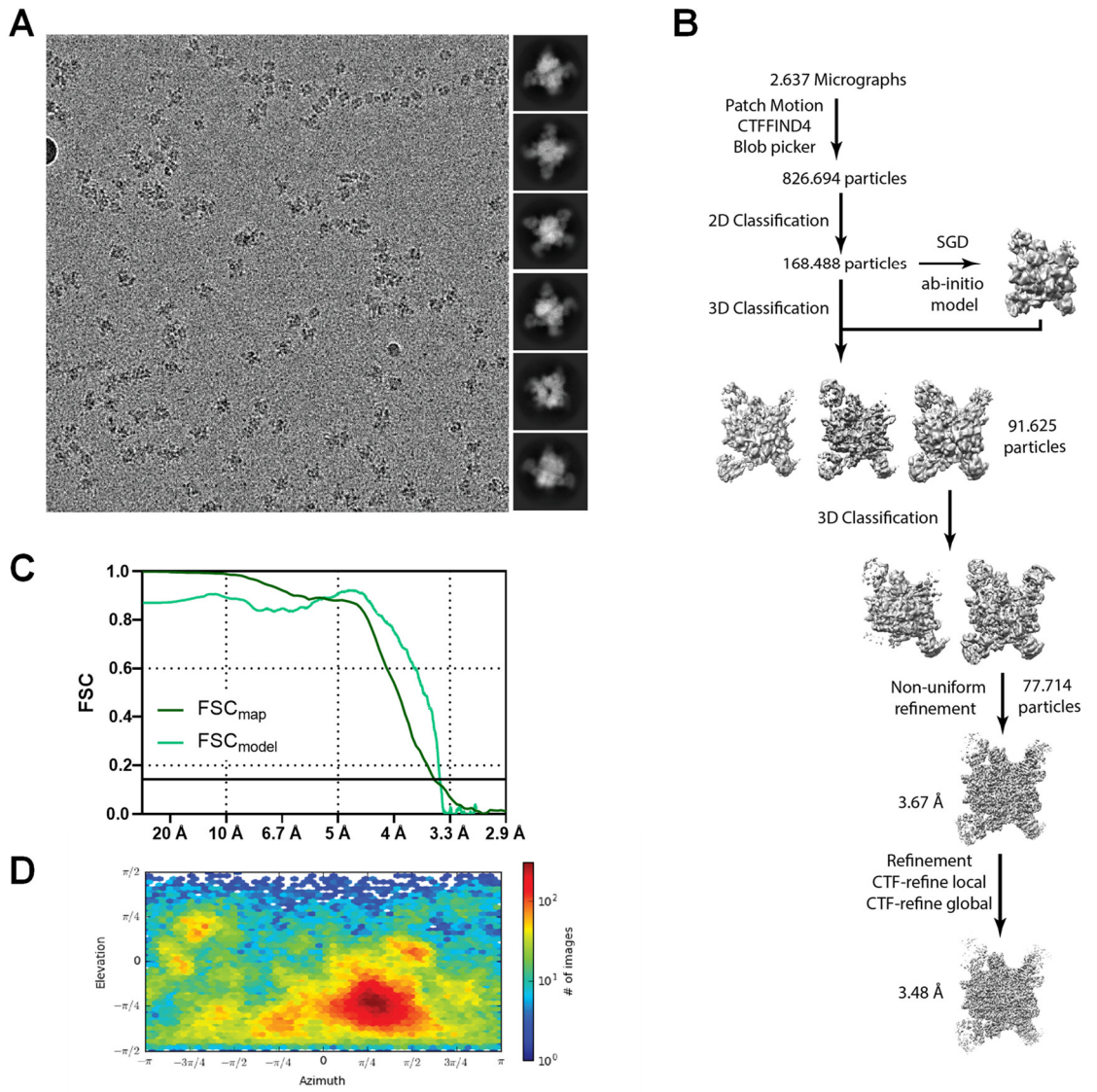


Figure 17. Cryogenic electron microscopy (Cryo-EM) data collection and analysis for *Me*PCC.
(**A**) A representative cryo-EM micrograph collected on an FEI Titan Krios microscope, operated at 300 kV and equipped with a K3 camera and representative reference-free 2D class averages. (**B**) Overview of the cryo-EM data processing scheme. (**C**) FSC_{map}: Gold-standard fourier shell correlation (FSC) plot from the final round of refinement in cryoSPARC. FSC_{model}: Model vs. map FSC for the final Phenix real space refined model. (**D**) Angular distribution of the particles used for the final round of refinement.

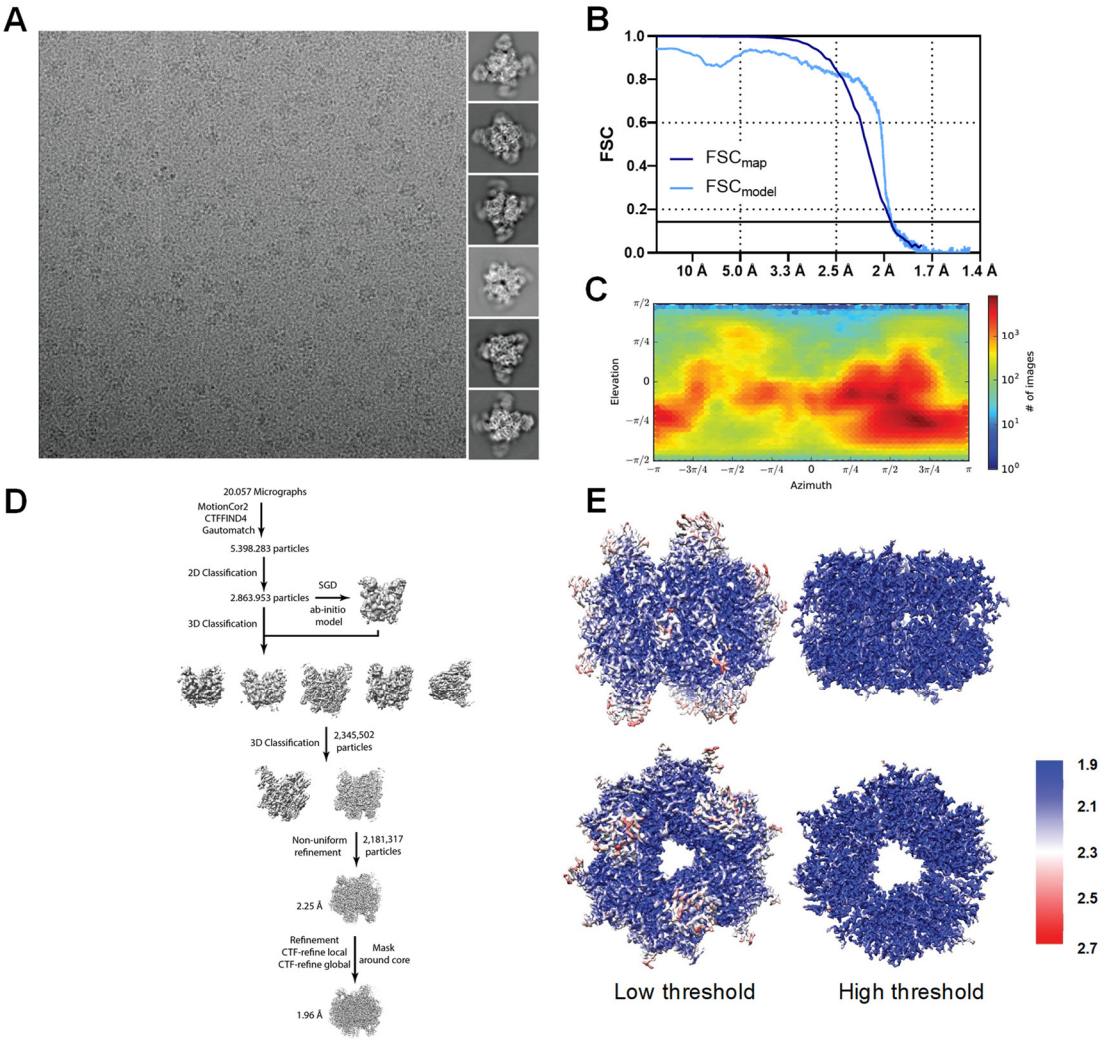


Figure 18. Cryogenic electron microscopy (Cryo-EM) data collection and analysis for GCC M5. (**A**) A representative cryo-EM micrograph collected on an FEI Titan Krios microscope, operated at 300 kV and equipped with a K3 camera and representative reference-free 2D class averages. (**B**) Tight: Gold-standard fourier shell correlation (FSC) plot from the final round of refinement in cryoSPARC. Masked: Model vs. map FSC for the final Phenix real space refined model. (**C**) Angular distribution of the particles used for the final round of refinement. (**D**) Overview of the cryo-EM data processing scheme. (**E**) Local resolution analysis of GCC M5. Maps show variation in local resolution of the central β-core of the complex, as estimated by cryoSPARC. Left: Low density threshold image, showing the presence of the BCCP domain. Right: High-density threshold image, showing the high resolved β-domains of the molecule.

The high resolution of the structure of GCC M5 becomes especially evident, when looking at the electron densities obtained for representative amino acid residues. At a contour level of up to 2 Å around the atoms, the structures of the single amino acids are clearly defined. This can be seen e.g. at the holes in the aromatic rings of phenylalanine, tyrosine and tryptophane, as well as the higher densities around the sulfur atoms of cysteine and methionine, or the "zigzag" structure of the extended side chains of lysine or arginine (Figure 19).

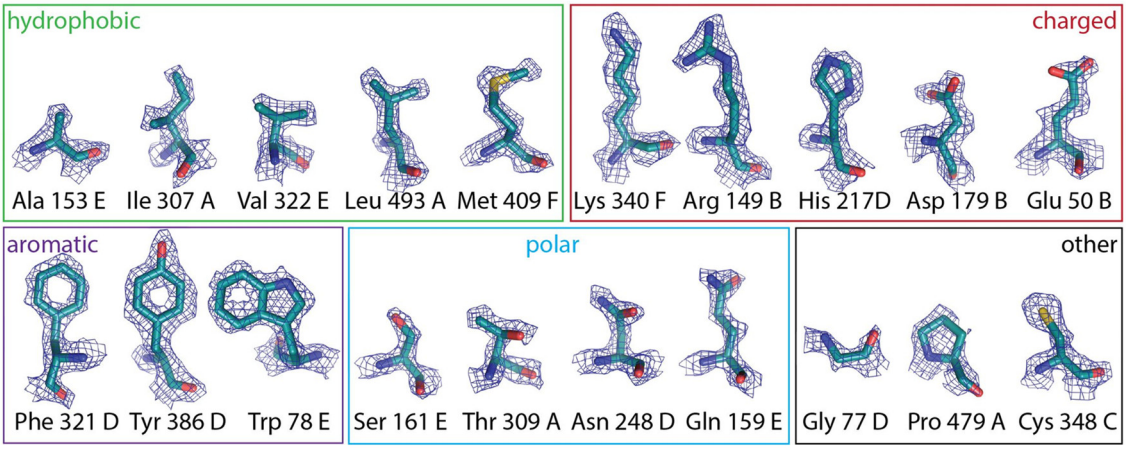


Figure 19. Representative amino acid residues of GCC M5 cryo-EM structure.
Representative side-chains of the complex (in stick representation) and surrounding electron density map. Maps are displayed as a mesh using a contour level of up to 2 Å around the atoms. Residue numbers and subunits (A-F) are specified.

The active sites of *Me*PCC as well as GCC M5 were also well resolved. The electron densities for the rationally altered amino acids as well as for CoA and biotin are shown in Figure 20. The biotin is covalently bound to K633 of the α-subunit and located in a "parking position" between the BC and the CT active sites, which is the same as in the chimeric PCC (PDB 3N6R). This position thus seems to appear to be of biological function and not just an artifact of crystallography or cryo-EM.

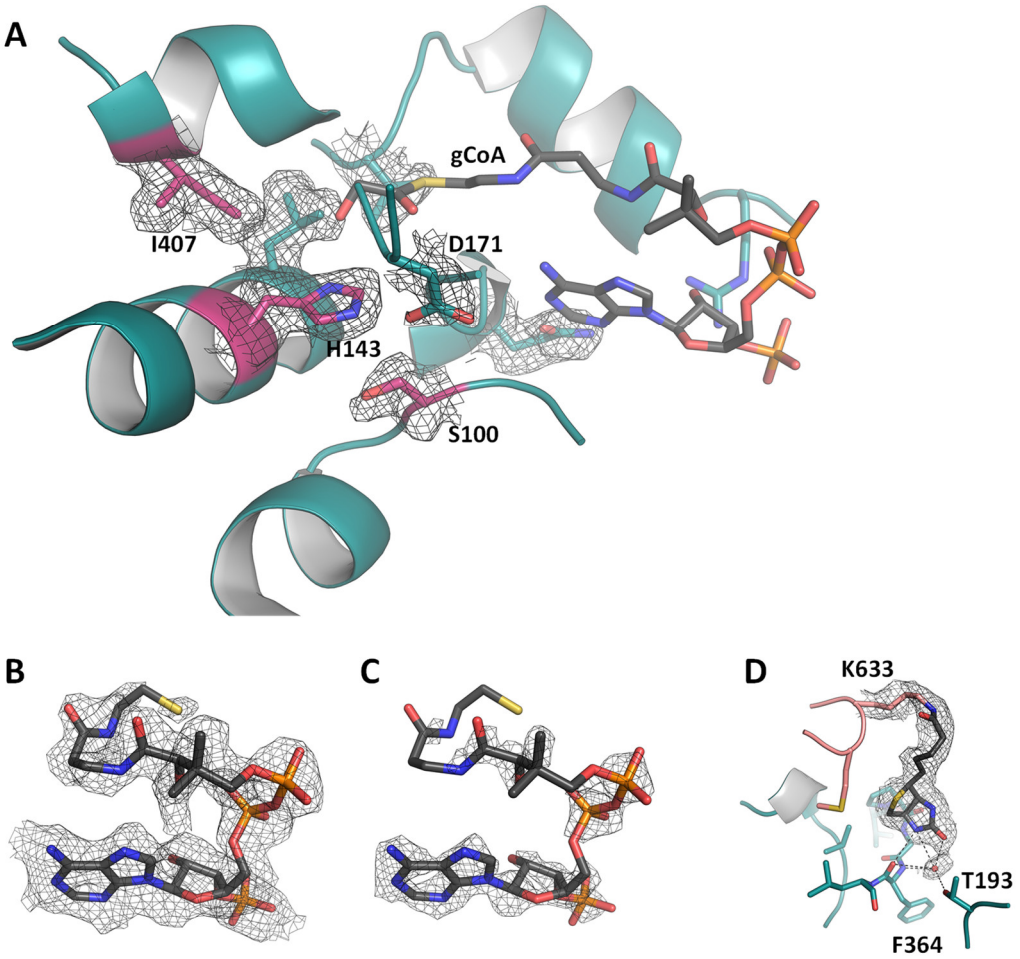


Figure 20. Electron density maps for GCC M5 as obtained by cryo-EM.
(**A**) Active site of GCC M5. Rationally introduced mutations are highlighted in pink. Glycolyl-CoA (gCoA) was modeled based on the position of methylmalonyl-CoA in PDB 1ON3 with additional manual fitting that reflects the GCC M5 active site architecture and coordination distances. (**B**) Electron density map at low contour level for CoA as present in the cryo-EM structure. (**C**) Electron density map at high contour level for CoA emphasizing the tight binding of the adenosyl moiety and the flexibility of the pantetheine arm. (**D**) Biotin cofactor covalently bound to Lys633 of GCC M5 α-subunit (salmon). β-subunits shown in teal. Possible hydrogen bonds are depicted as dashed lines.

The positions of amino acid residues in the CT active sites of *Me*PCC and GCC M5 were very similar to the homology models that were made based on PDB 3N6R (Figure 5, Figure 21), corroborating the rational design of mutations chosen for GCC. The L100S mutation was assumed to fix H143 in a rotamer conformation facilitating hydrogen bonding interaction with glycolyl-CoA. The actual distance of H143 and S100 is 3.8 Å (3.5 Å in the homology model), making a direct hydrogen bond more unlikely than initially thought. Probably more important than anticipated based on the homology model is D171. Its distance to H143 in the homology model was 4.3 Å, while the actual distance

in the structure is 2.8 Å, making a hydrogen bond between those two residues highly likely. The effect of the L100S mutation might therefore be more indirect by providing more space for H143 and D171, facilitating their interaction.

The I450V mutation of GCC M4 and M5 is located in an α-helix nearby and may impact distances within the active site or at the interface with the biotin carboxylase subunit (Figure 21).

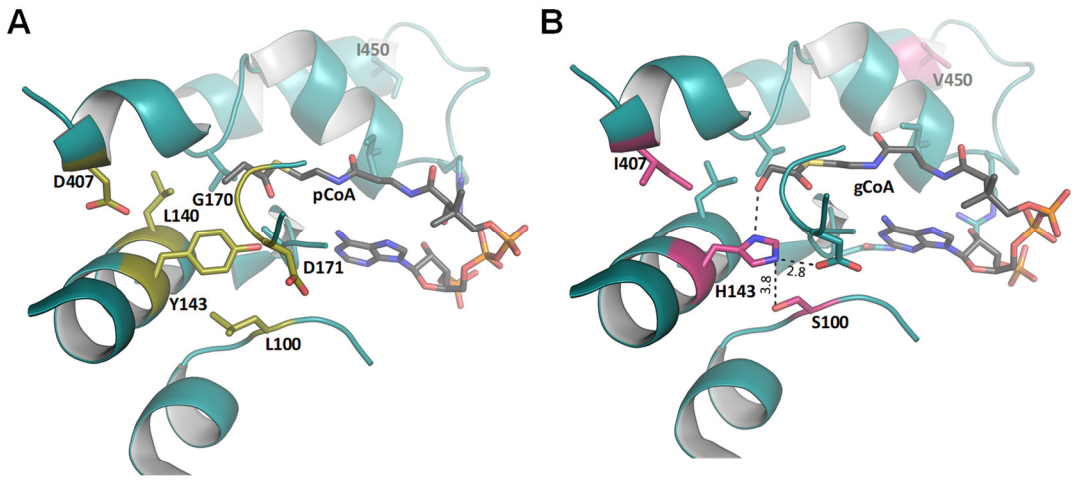


Figure 21. Active site comparison of propionyl-CoA carboxylase of *M. extorquens* and engineered glycolyl-CoA carboxylase.
(**A**) Cryo-EM structure of propionyl-CoA carboxylase of *M. extorquens* (*Me*PCC) active site (PDB 6YBP). Propionyl-CoA (pCoA) was modeled according to the position of methylmalonyl-CoA in PDB 1ON3 (**B**) Cryo-EM structure of engineered glycolyl-CoA carboxylase active site (PDB 6YBQ). Glycolyl-CoA (gCoA) was modeled according to the position of methylmalonyl-CoA in PDB 1ON3 with additional manual fitting that reflects the new active site architecture and coordination distances.
Residues targeted during rational design of the active site are colored in yellow. Actual active site changes in the GCC M4 and M5 mutants are colored in pink. Potential new hydrogen bonds are depicted as dashed lines. Distances are given in Å. The Y143H mutation is likely engaging in hydrogen bonding with the hydroxyl group of glycolyl-CoA.

In addition to the four mutations in and around the active site, GCC M5 features an additional mutation far away from the active site (W502R), which is located at the rotational symmetry of the enzyme oligomer (Figure 22). It was initially thought that the introduced arginine might form polar contacts to other residues in the central core. According to the structure, this is most probably not the case. Instead, it seems that the W502R mutation leads to a shift of the loop, on which it is located (Figure 22B). This shift has no obvious effect on the rest of the enzyme, as the helices around are almost

perfectly aligned in *Me*PCC and GCC M5. It might be speculated that the effect of this mutation becomes more obvious in a different catalytic state of the enzyme during catalysis, as the structure only represents a single non-catalytic state of the enzyme.

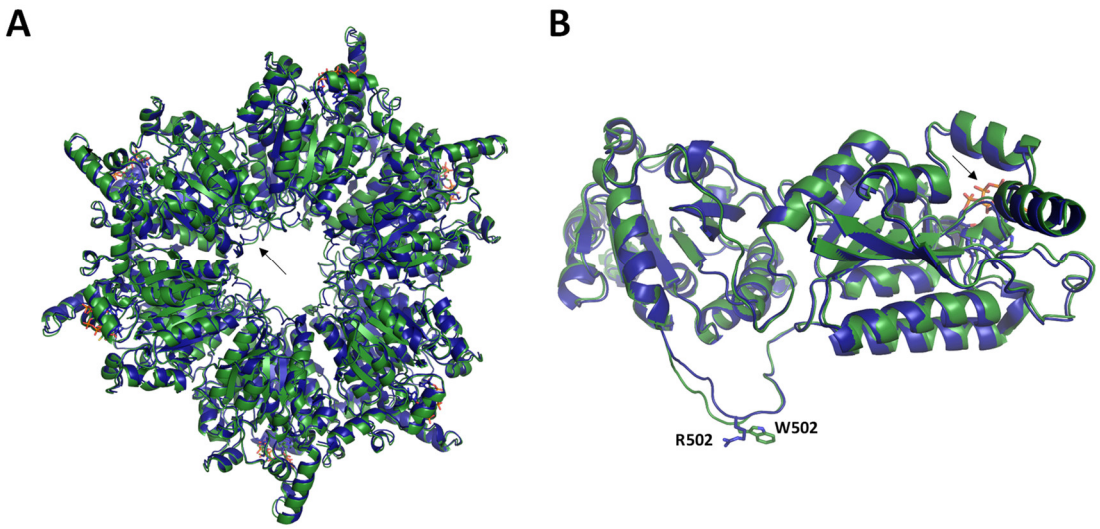


Figure 22. Structural comparison of *Me*PCC and GCC M5 regarding to W502R mutation.
(**A**) Overlay of *Me*PCC and GCC M5 hexameric β-cores. The arrow marks the position of W502 (WT) and R502 (M5). (**B**) Closer look at a selected β-subunit. The arrow marks the position of the CoA (active site of β-subunit).

2.2 *In vitro* reconstruction of the tartronyl-CoA pathway

The successful identification and engineering of GCS, GCC and TCR made it possible, to reconstruct the whole tartronyl-CoA pathway (Figure 23) *in vitro*. Initially, the tartronyl-CoA pathway had been designed as a more efficient and carbon positive photorespiratory bypass when interfaced with the CBB cycle (Figure 23; green box). However, it was obvious that the tartronyl-CoA pathway could as well be interfaced with other metabolic modules to enable the direct carboxylation of glycolate to form glycerate.

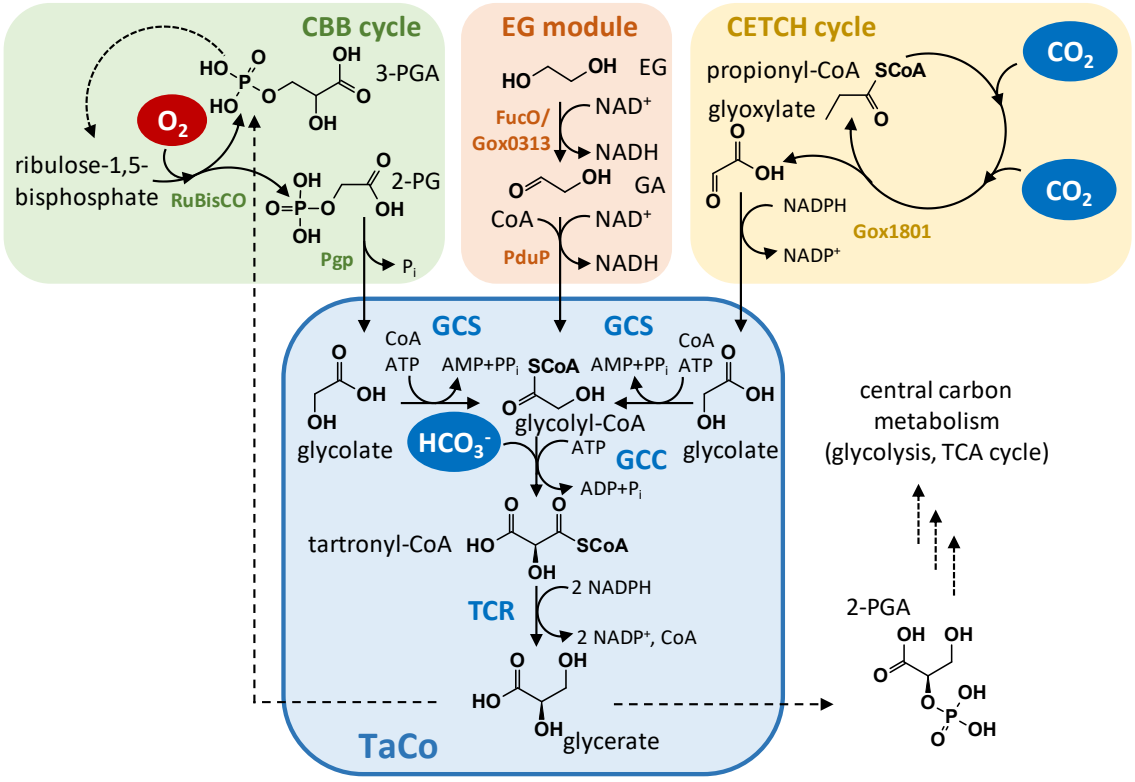


Figure 23. The tartronyl-CoA pathway (TaCo, blue box) and its applications.
(**A**) In the tartronyl-CoA (TaCo) pathway, glycolate is activated by glycolyl-CoA synthetase (GCS) to glycolyl-CoA, which is carboxylated via the key enzyme glycolyl-CoA carboxylase (GCC) into tatronyl-CoA. Tartronyl-CoA is reduced to glycerate by tartronyl-CoA reductase (TCR). The CBB (green) and the CETCH (yellow) cycle can be interfaced with the TaCo pathway on the level of gylcolate, ethylene glycol (EG) assimilation (red) can be interfaced through glycolyl-CoA. GA: glycolaldehyde

For instance, it would be possible to utilize the tartronyl-CoA pathway for the conversion of the plastic waste component and environmental pollutant EG into glycerate for further

incorporation into central carbon metabolism (Figure 23; EG module, red box). Furthermore, the tartronyl-CoA pathway is an attractive way to connect the CETCH cycle with central carbon metabolism and fix a third CO_2 molecule in the process (Figure 23; CETCH cycle, yellow box).

As measured spectrophotometrically, the tartronyl-CoA pathway was functional as a whole *in vitro* and converted glycolate into glycerate at a rate of 27 ± 1 nmol min^{-1} mg^{-1} (Figure 24).

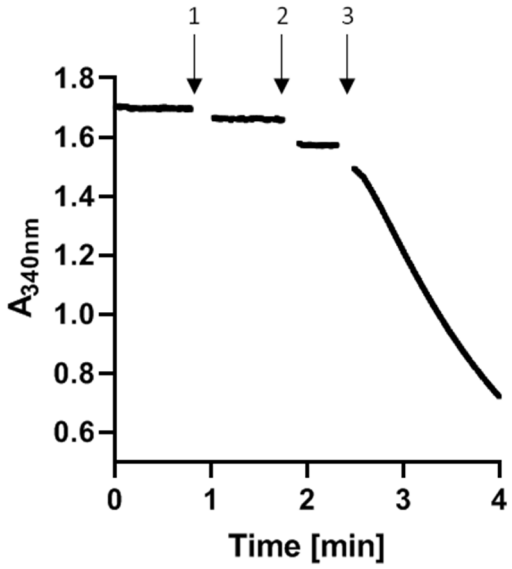


Figure 24. Spectrophotometric assay for activity assessment of the complete tartronyl-CoA pathway. The assay contained 100 mM MOPS/KOH pH 7.8, 10 mM $MgCl_2$, 0.4 mM NADPH, 50 mM $KHCO_3$, 0.15 mg mL^{-1} GCS, 0.56 mg mL^{-1} GCC M4, 1.1 mg mL^{-1} TCR. 1: addition of 20 mM glycolate, 2: addition of 1 mM CoA, 3: addition of 5 mM ATP. Data shown are representative of three independent experiments.

To continuously operate the tartronyl-CoA pathway, it was coupled to different ATP regeneration modules (Nocek et al., 2008; Schwander et al., 2016; Mordhorst et al., 2017). While the use of a polyphosphate kinase-based system was limited due to precipitation of polyphosphate at concentrations above 20 mM, a phosphocreatine-based system proved 3-times more effective (Figure 25).

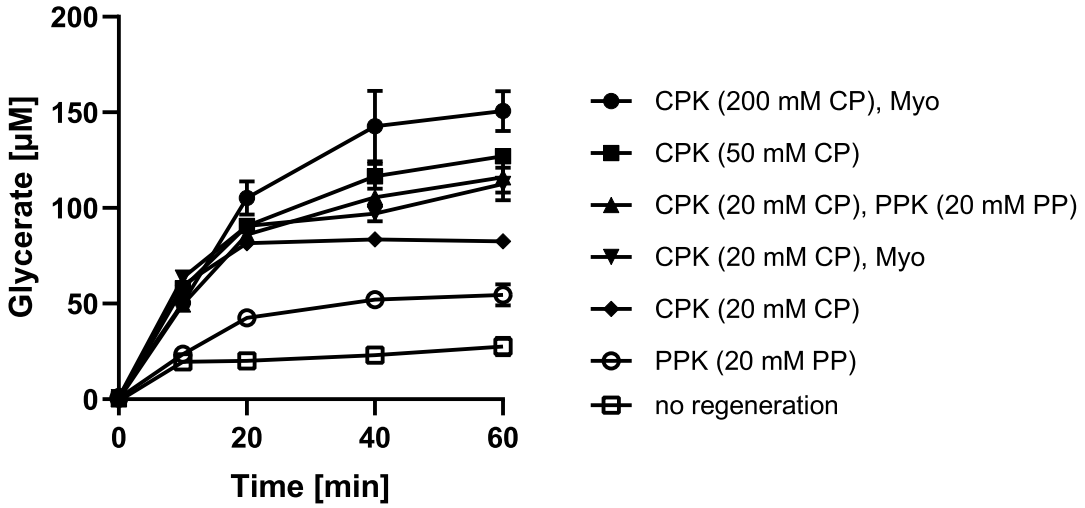


Figure 25. Optimization of ATP regeneration for *in vitro* operation of the tartronyl-CoA pathway.
The assays contained 100 mM MOPS/KOH pH 7.8, 5 mM ATP, 10 mM $MgCl_2$, 2 mM CoA, 2 mM NADPH, 50 mM $KHCO_3$, 1 mM glycolate, 150 µg mL^{-1} GCS, 608 µg mL^{-1} GCC and 1120 µg mL^{-1} TCR. Additionally, the assays contained 20 - 500 mM phosphocreatine (CP), 6.7 U mL^{-1} creatine phosphokinase (CPK), 20 mM polyphosphate (PP), 33 µg mL^{-1} polyphosphate kinase 2-I (PPK2-I) in combination with 133 µg mL^{-1} polyphosphate kinase 2-II (PPK2-II) or 67 µg mL^{-1} myokinase (Myo) (and variations thereof as indicated in the legend). Shown are mean ± range of values for two independent experiments each.

The operation of the core sequence of the tartronyl-CoA pathway was further confirmed by sequentially reconstituting the pathway's reaction sequence (Figure 26). ^{13}C-Labelling was used to confirm carboxylation activity.

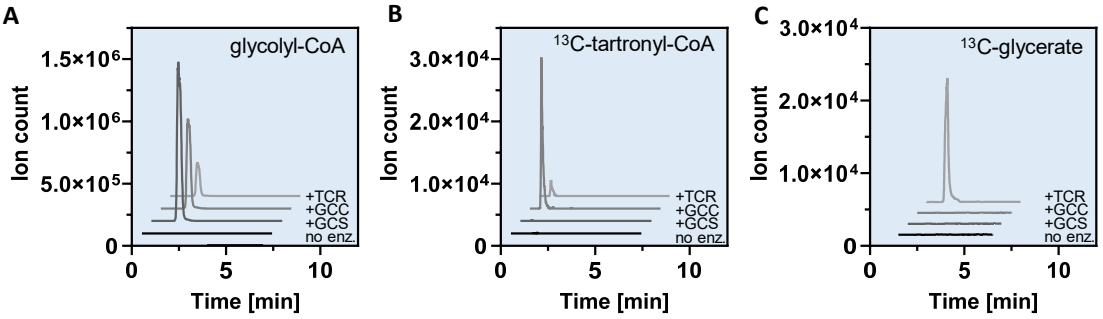


Figure 26. Stepwise reconstitution of the tartronyl-CoA pathway *in vitro*, starting from glycolate and ^{13}C-bicarbonate.
Shown are extracted ion count chromatograms for (**A**) glycolyl-CoA ($[M+H]^+$ 826.1 m/z), (**B**) ^{13}C-labelled tartronyl-CoA ($[M+H]^+$ 871.1 m/z) and (**C**) ^{13}C-labelled, 3-NPH-derivatized glycerate ($[M-H]^-$ 241.1 m/z) at t_0 (no enz.), after addition of GCS, addition of GCC and addition of TCR. The shown chromatograms are representatives of three independent experiments.

2.2.1 The tartronyl-CoA pathway as photorespiratory bypass

Starting from 2-PG, the tartronyl-CoA pathway was tested together with 2-PG phosphatase (Pgp) and glycerate kinase (GlxK) using ^{13}C-labelled bicarbonate. The rate of ^{13}C-labelled phosphoglycerate production was 12.3 ± 0.7 nmol min^{-1} mg^{-1}, indicating that the tartronyl-CoA pathway could be interfaced with photorespiration (Figure 27).

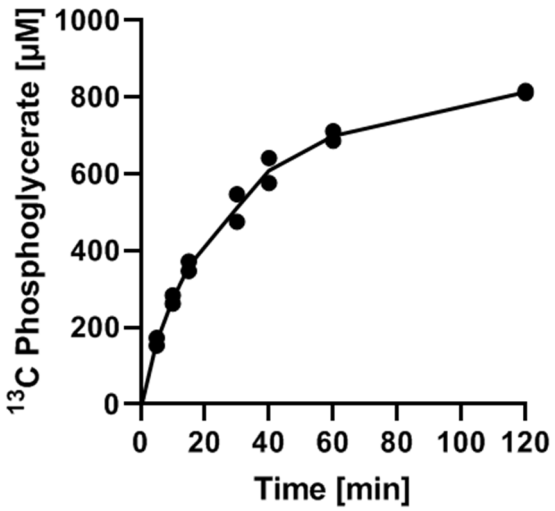


Figure 27. ^{13}C-phosphoglycerate produced by the tartronyl-CoA pathway from 2-phosphoglycolate and ^{13}C-bicarbonate.
The tartronyl-CoA pathway was able to convert RuBisCO's oxygenation product, 2-phosphoglycolate, into phosphoglycerate through direct carboxylation (indicated by the incorporation of $H^{13}CO_3^-$). Shown are the results of two representative independent experiments.

To mimic photorespiratory conditions, experiments with equal amounts of 3-PGA and 2-PG (1 mM each) as would be produced by RuBisCO under 100 % oxygenation reaction, were performed. A malate read-out module was developed that converts 3-PGA into malate, while 2-PG is converted only to malate when the tartronyl-CoA pathway is active (Figure 28A). Isotopic labelling with ^{13}C-bicarbonate allowed to distinguish the fraction of malate derived from 3-PGA (single-label malate) from that derived through the tartronyl-CoA pathway from 2-PG (double-label malate). 3-PGA could be converted into malate by only one carboxylation step, being the carboxylation of PEP to form oxaloacetate by PEP carboxylase (Figure 28A). The carboxylation introduces one ^{13}C label, and therefore the formed malate is single-labeled.

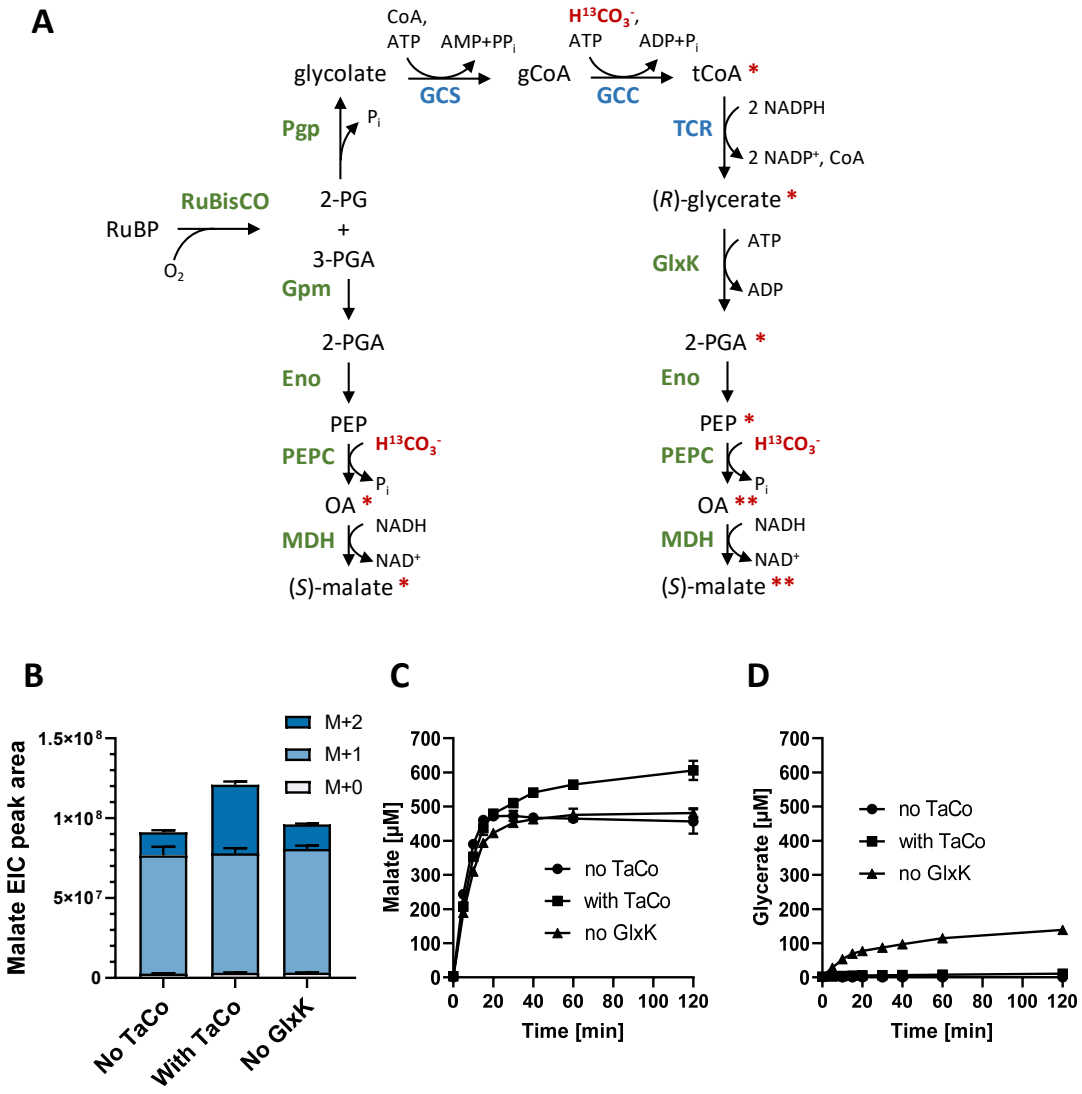


Figure 28. Malate read-out to measure *in vitro* operation of the tartronyl-CoA pathway as photorespiration bypass.

(**A**) Schematic representation of the malate read-out developed in this study. Enzymes in blue represent the tartronyl-CoA pathway. Enzymes in green represent naturally occurring enzymes. Red asterisks indicate ^{13}C incorporation into intermediates and products from labelled bicarbonate. RuBP: ribulose-1,5-bisphosphate, 2-PG: 2-phosphoglycolate, 3-PGA: 3-phosphoglycerate, 2-PGA: 2-phosphoglycerate, gCoA: glycolyl-CoA, tCoA: tartronyl-CoA, PEP: phosphoenolpyruvate, OA: oxaloacetate, Gpm: phosphoglycerate mutase, Eno: enolase, PEPC: PEP carboxylase, MDH: malate dehydrogenase, GCS: glycolyl-CoA synthetase, GCC: glycolyl-CoA carboxylase, TCR: tartronyl-CoA reductase, GlxK: glycerate kinase, TaCo: tartronyl-CoA pathway. (**B**) Isotopic fractions of malate produced by the tartronyl-CoA pathway and the read-out module. The assay for the negative control ("No TaCo") contained GlxK, but no GCS, GCC and TCR. The assay containing the TaCo enzymes ("With TaCo") contained GlxK, GCS, GCC M4 and TCR. The assays for the "no GlxK" control contained GCS, GCC M4 and TCR, but no GlxK. All assays were started by the addition of 5 mM ATP, 1 mM 2-PG and 1 mM 3-PGA. Shown are mean ± range of values for two independent experiments. (**C**) Malate concentrations (all isotopic fractions added up). (**D**) Glycerate concentrations (all isotopic fractions added up).

The conversion of 2-PG could only proceed in the presence of the tartronyl-CoA pathway. Two carboxylation reactions (and therefore two times incorporation of $H^{13}CO_3^-$) were necessary for 2-PG conversion into malate: the first was the carboxylation of glycolyl-CoA to form tartronyl-CoA by GCC in the tartronyl-CoA pathway, the second was carboxylation of PEP to form oxaloacetate by PEP carboxylase (part of the read-out module, Figure 28A). Thus, the malate formed via the tartronyl-CoA pathway was double-labelled while malate being formed from 3-PGA was single-labelled. Malate formation increased in presence of the tartronyl-CoA pathway by 130 µM, which makes 33 %. ^{13}C-labelling confirmed that the malate surplus was provided by the tartronyl-CoA pathway (Figure 28B-C). In the control experiment without GlxK the final concentration of glycerate was 139 µM (Figure 28D), further undergirding that the amount of 130 µM additional malate in the presence of the tartronyl-CoA pathway was indeed formed by the tartronyl-CoA pathway.

Furthermore, glycerate accumulated only in the "no GlxK" control, which implies that all glycerate was converted when GlxK was present. The employed Glxk from *E. coli* only accepts the (*R*)-stereoisomer of glycerate (Hansen and Hayashi, 1962; Doughty et al., 1966). Thus, this experiment confirmed that the physiologically relevant stereoisomer (*R*)-glycerate is the only product of the tartronyl-CoA pathway.

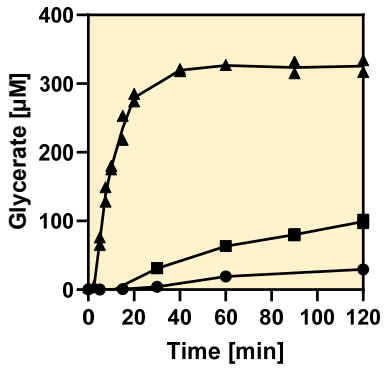
Notably, the reaction of *E. coli* GlxK forms 2-PGA instead of 3-PGA (Bartsch et al., 2008; Zelcbuch et al., 2015), which would be the actual direct intermediate of the CBB cycle. By using the 2-PGA forming glycerate kinase, the product could directly be channeled into central carbon metabolism like glycolysis and the TCA cycle. Furthermore, it is expected that a conversion of the (*R*)-glycerate formed by the tartronyl-CoA pathway into 3-PGA would be readily possible when using a plant-derived phosphoglycerate kinase, which produces 3-PGA (Bartsch et al., 2008; Zelcbuch et al., 2015).

Taken together, the malate read-out experiments demonstrated the potential of the tartonyl-CoA pathway as synthetic, carbon-positive photorespiratory bypass that is able to fix CO_2 even at 100 % RuBisCO oxygenation reaction.

2.2.2 The tartronyl-CoA pathway coupled to the CETCH cycle

The CETCH cycle is a synthetic pathway for the fixation of CO_2 *in vitro* with a 20 % higher efficiency than the CBB cycle of photosynthesis (Schwander et al., 2016). The output molecule of the CETCH cycle is glyoxylate, which can be easily converted into glycolate by using the succinic semialdehyde/glyoxylate reductase (Gox1801) from *Gluconobacter oxydans*. The tartronyl-CoA pathway, in turn, could assimilate glycolate. Coupling of the CETCH cycle with the tartronyl-CoA pathway would further increase the CO_2 fixation efficiency due to the additional CO_2 fixation reaction of GCC.

The combination of the 17 enzymes of CETCH v5.4 (Schwander et al., 2016) with Gox1801 and the tartronyl-CoA pathway initially produced glycerate at a rate of 0.1 nmol min^{-1} mg^{-1} (CT_A, Figure 29). In a second attempt with the addition of the tartronyl-CoA pathway enzymes after 120 min, the system produced glycerate at a rate of 0.4 nmol min^{-1} mg^{-1} (CT_B, Figure 29).



	●	■	▲
Experiment	**CT_A**	**CT_B**	**CT_C**
NADPH regen.	formate DH	formate DH	glucose DH
ATP regen.	PPK2-I, PPK2-II	PPK2-I, PPK2-II	CPK, Myo
polyphosphate conc. [mM]	20	20	n/a
phosphocreatine conc. [mM]	n/a	n/a	100
TaCo addition	@start	after 120 min	after 120 min
SucD	*C. kluyveri*	*C. kluyveri*	*C. difficile*
propCoA (start) [µM]	200	1000	170
final glycerate conc. [µM]	31	105 ± 7	331 ± 21
rate of TaCo [nmol min^{-1} mg^{-1}]*	0.1	0.4 ± 0.0	4.8 ± 0.1

Figure 29. Optimization of experiments for coupling of the tartronyl-CoA pathway to the CETCH cycle for the production of glycerate.
In the course of optimization, the NADPH regeneration was changed from using formate dehydrogenase (formate DH) to glucose dehydrogenase (glucose DH), the ATP regeneration was changed from using polyphosphate kinase (PPK2-I and PPK2-II) to creatine phosphokinase (CPK). The enzyme for succinyl-CoA reduction in the CETCH cycle was changed from SucD from *C. kluyveri* to the homolog from *C. difficile* to reduce side reactivity with glycolyl-CoA. Shown are the results of two representative independent experiments each. * The rate is based on the concentration of the enzymes of the tartronyl-CoA pathway (TaCo). n/a - not applicable**.**

Further analysis identified two potential bottlenecks of the overall system. First, the use of formate dehydrogenase to regenerate NADPH led to the production of the dead-end

metabolite formyl-CoA due to a side reaction of GCS, trapping CoA and consuming ATP. Second, succinyl-CoA reductase of *Clostridium kluyveri* (*Ck*SucD), which was used in the CETCH cycle, had a side reactivity with glycolyl-CoA of 311 ± 21 nmol min^{-1} mg^{-1} (Figure 30), depleting it from the tartronyl-CoA pathway.

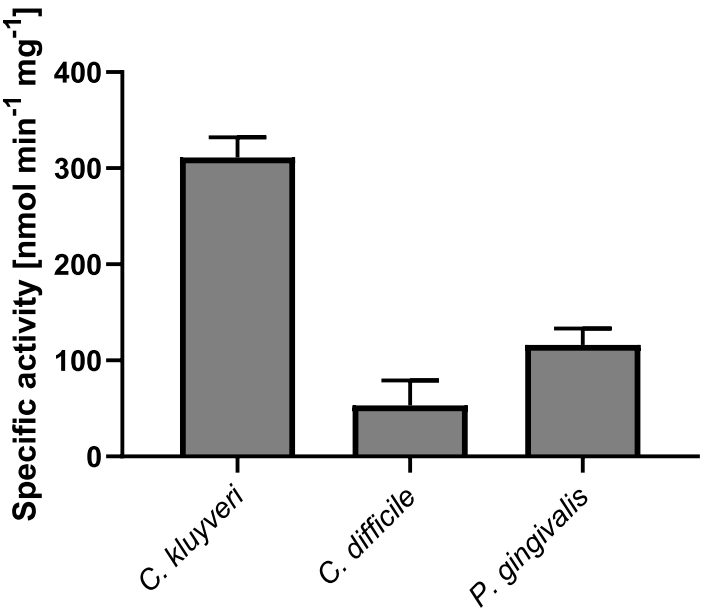


Figure 30. Specific activities of different SucDs (succinyl-CoA reductases) with glycolyl-CoA. Activities of *C. kluyveri* and *C. difficile* SucDs were measured with NADPH as cofactor, *P. gingivalis* SucD was measured with NADH as cofactor. Data represent mean ± SD of n=3 replicates.

To optimize the system, it was decided to switch from formate dehydrogenase to glucose dehydrogenase for NADPH regeneration. Furthermore, different homologs of succinyl-CoA reductase were tested for their activities with glycolyl-CoA to reduce the loss of the intermediate due to this futile side reactivity. It was found that both the homologs from *Clostrioides difficile* (53 ± 21 nmol min^{-1} mg^{-1} measured with NADPH as cosubstrate) as well as from *Porphyromonas gingivalis* (116 ± 17 nmol min^{-1} mg^{-1} measured with NADH as cosubstrate) had lower activities with glycolyl-CoA than the *C. kluyveri* enzyme (Figure 30) while their specific activities with succinyl-CoA were comparable with the *C. kluyveri* homolog.

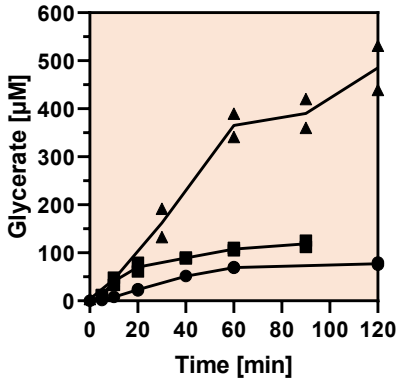
Because of the lower reactivity with glycolyl-CoA and the preference for NADPH as a cosubstrate, the *C. difficile* homolog was used instead of the one from *C. kluyveri* for

further experiments. In addition, higher concentrations of phosphocreatine (CP) were used for enhanced ATP regeneration. Overall, these improvements resulted in an almost 50-fold higher glycerate production rate of 4.8 nmol min^{-1} mg^{-1} (CT_C, Figure 29), which was comparable to production rates of the CETCH cycle (5 nmol min^{-1} mg^{-1}) (Schwander et al., 2016). This demonstrated that the tartronyl-CoA pathway is compatible with synthetic CO_2 fixation pathways to further boost their carbon capture efficiency and yield.

2.2.3 The tartronyl-CoA pathway for ethylene glycol conversion

Ethylene glycol (EG) is a constituent of plastic waste, an important organic solvent and a defrosting agent. Naturally, EG can be assimilated aerobically via the glycerate pathway e.g. in *E. coli* and *P. putida* (Boronat et al., 1983; Mückschel et al., 2012), whereby CO_2 is released by the enzyme glyoxylate carboligase (Gcl). Replacing natural EG conversion with the tartronyl-CoA pathway would provide a carbon-fixing EG degradation route. A pathway for the conversion of EG into the central metabolite glycerate was designed (Figure 23). EG is first converted into glycolaldehyde, which is further oxidized to glycolyl-CoA, the entry point into the tartronyl-CoA pathway. First, *L*-lactaldehyde dehydrogenase of *E. coli* (FucO) that is able to oxidize EG into glycolaldehyde (Sridhara et al., 1969; Baldoma and Aguilar, 1988) was tested under assay conditions for the tartronyl-CoA pathway (pH 7.8, 37 °C). Under these conditions, v_{max} of the reaction was only 83 ± 4 nmol min^{-1} mg^{-1} with an app. K_M of 68 ± 7 mM (Figure S6). The pH optimum for the dehydrogenase reaction of FucO had been determined previously to be 9.5 (Boronat and Aguilar, 1979), explaining the low residual activity at pH 7.8. Moreover, it has been reported that the enzyme is also able to oxidize glycerol to glyceraldehyde (Boronat and Aguilar, 1979), which represents an unwanted side reaction. Usually, all enzyme preparations are stored in buffer containing 20 % glycerol. The use of FucO made it therefore necessary to exchange the buffer of all enzymes used prior to experiments.

For the further oxidation of glycolaldehyde to directly form glycolyl-CoA, the CoA-acylating propionaldehyde dehydrogenase PduP of *Rhodospseudomonas palustris* BisB18 was employed. The promiscuity of this enzyme for various substrates has been shown before (Zarzycki et al., 2017). The activity for the reduction of glycolyl-CoA to form glycolaldehyde (the reverse reaction of the here preferred one) has also been shown (Trudeau et al., 2018). In the present work, a specific activity of 23 µmol min^{-1} mg^{-1} (pH 7.8, 37°C) for the oxidation of glycolaldehyde to form glycolyl-CoA was determined. When both FucO and PduP were combined with GCC and TCR of the tartronyl-CoA pathway, glycerate was produced from EG to a final concentration of 77 ± 3 µM at a rate of 0.4 nmol min^{-1} mg^{-1} (EG_A, Figure 31), showing the concerted function of EG oxidation module and the tartronyl-CoA pathway.



	●	■	▲
Experiment	**EG_A**	**EG_B**	**EG_C**
EG oxidation	FucO	Gox0313	Gox0313
ATP regeneration	n/a	CPK	CPK
phosphocreatine conc. [mM]	n/a	10	200
NADH oxidase [µg µL^{-1}]	n/a	0.04	0.08
final glycerate conc. [µM]	77 ± 3	119 ± 9	485 ± 65
production rate (60 min) [nmol min^{-1} mg^{-1}]*	0.4 ± 0.0	0.6 ± 0.0	2.1 ± 0.2

Figure 31. Optimization of experiments for ethylene glycol (EG) conversion to glycerate by the ethylene glycol module coupled to the tartronyl-CoA pathway.
In the course of optimization, the enzyme for ethylene glycol oxidation was changed from FucO to Gox0313 and ATP regeneration by creatine phosphokinase (CPK) as well as NAD^+ regeneration by a water-forming NADH oxidase was included. The conditions for each experiment are depicted in the table. Shown are the results of two representative independent experiments each. * Production rate is based on the concentration of TaCo enzymes. n/a - not applicable.

To improve initial EG oxidation, FucO was replaced by an alcohol dehydrogenase from *G. oxydans* (Gox0313) which was reported to have an app. K_M of 2.4 ± 0.2 mM and a V_{max} of 4.8 ± 0.2 µmol min^{-1} mg^{-1} for ethylene glycol (determined using a Lineweaver-Burk plot under non-saturated substrate concentrations). The pH optimum of Gox0313 was 8.5 (Zhang et al., 2015), which was closer to the tartronyl-CoA pathway assay conditions than the one of FucO. When the enzyme was tested in the present work and the kinetic parameters were determined following Michaelis-Menten kinetics, it was found that the actual app. K_M of 964 ± 84 mM was much higher than reported, while the V_{max} of 7.1 ± 0.3 µmol min^{-1} mg^{-1} was close to the reported value. Nevertheless, because of its much higher specific activity compared to FucO at 100 mM EG, Gox0313 was chosen for subsequent experiments.

For further improved performance of the pathway for EG conversion, the use of an NADH oxidase was investigated. The thermodynamic feasibility of both EG and glycolaldehyde oxidation is highly dependent on the ratio of NAD^+ over NADH. To ensure a high NAD^+ concentration over the course of the whole experiment, a water-forming NADH oxidase from *Lactobacillus pentosus* (Nowak et al., 2015) was tested. The enzyme had a specific activity for NADH oxidation of 28 µmol min^{-1} mg^{-1} at 0.4 mM NADH, while the specific activity for NADPH (at 0.4 mM) was 0.1 µmol min^{-1} mg^{-1} (only 0.03 % of its NADH

oxidation activity). The high discrimination between NADH and NADPH was important, because NADPH was needed in its reduced form for tartronyl-CoA reduction in the tartronyl-CoA pathway. In addition, ATP regeneration by creatine phosphokinase (CPK) was included to improve the pathway efficiency. Together with the above-mentioned exchange of FucO with Gox0313 and the use of NADH oxidase, the production rate could be slightly improved to 0.6 nmol min^{-1} mg^{-1} with a final glycerate concentration of 119 ± 9 µM (EG_B, Figure 31). Together with the integration of more efficient ATP regeneration using higher concentrations of CP, the rate of glycerate production was enhanced by a factor of 5 to 2.1 nmol min^{-1} mg^{-1} with a final glycerate concentration of 485 ± 65 µM, demonstrating the CO_2-dependent conversion of EG into the central carbon metabolite glycerate.

2.2.4 Towards *in vivo* implementation of the tartronyl-CoA pathway

It has been shown that the tartronyl-CoA pathway can be reconstituted from its single enzymes and that it can be used for several different applications, including the conversion of EG into glycerate. To harness its full potential it is necessary to introduce the tartronyl-CoA pathway into living organisms. Here, an initial attempt to use the tartronyl-CoA pathway for the conversion of EG into the central carbon metabolite glycerate, ultimately enabling growth on EG and CO_2 in *P. putida* KT2440, is presented. An engineered strain harbouring the tartronyl-CoA pathway could be used for the growth-dependent screening for better enzyme variants of enzymes of the tartronyl-CoA pathway. Furthermore, such a strain could be used for biotechnological purposes, to produce chemical building blocks or other high-value compounds from the plastic-waste component EG and CO_2.

P. putida KT2440 is naturally not able to degrade EG, whereas mutants able to use EG as sole carbon source were isolated after adaptive laboratory evolution (Li et al., 2019). It was shown that *P. putida* KT2440 mutants metabolized EG via the glycerate pathway, with glyoxylate carboligase as the key enzyme (Franden et al., 2018) (Figure 32). The regulator GclR usually prevents transcription of the *gcl* operon, containing also hydroxypyruvate isomerase (*hyi*), tartronate semialdehyde reductase (*glxR*), hydroxypyruvate reductase (*ttuD*) and pyruvate kinase (*pykF*). Strains with disrupted GclR showed growth on EG as sole carbon source, demonstrating the role of this regulator as repressor of the *gcl* operon (Li et al., 2019).

In the present work, a *P. putida* KT2440 Δ*gclR* strain has been constructed as a positive control for the growth on EG. Furthermore, a Δ*gcl-hyi-glxR* strain was constructed in order to implement the tartronyl-CoA pathway into this strain (Figure 32). This strain would only be able to grow on EG, if the tartronyl-CoA pathway is active, as the native route for EG conversion is disrupted. As additional control, also a double knockout strain was constructed (*P. putida* KT2440 Δ*gclR* Δ*gcl-hyi-glxR*).

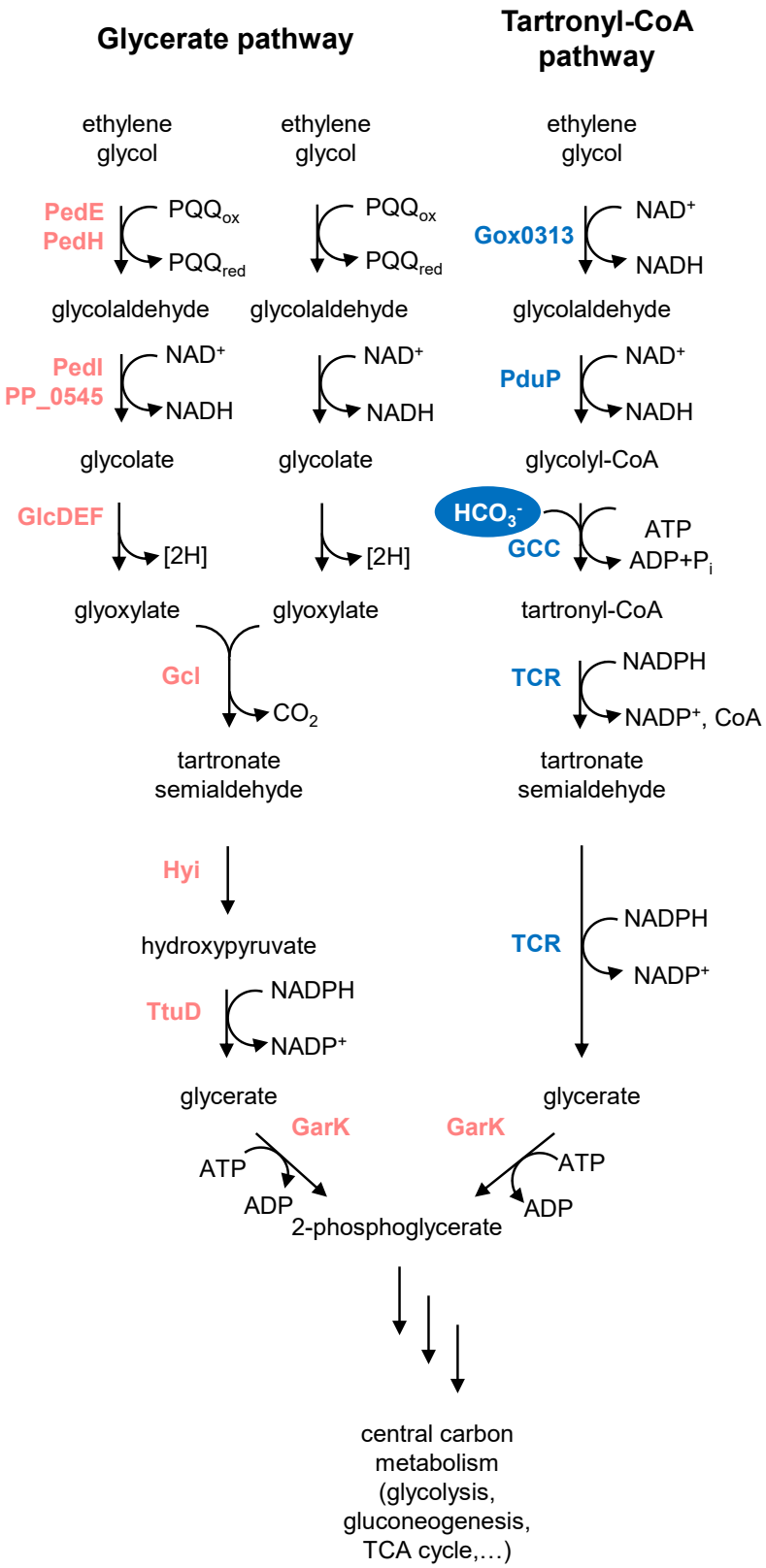


Figure 32. Ethylene glycol metabolism in *P. putida*.
On the left side, proposed ethylene glycol degradation in *P. putida* is shown (Mückschel et al., 2012; Franden et al., 2018; Li et al., 2019). The right side shows the proposed conversion of ethylene glycol via the tartronyl-CoA pathway. Enzyme names in salmon represent enzymes of *P. putida*, names in blue are enzymes of the tartronyl-CoA pathway.

In preliminary growth experiments, all three knockout strains as well as the wt strain showed growth on 20 mM glucose to relative A_{600nm} values between 0.83 and 0.96 after 20 h (Figure 33A). As expected, the Δ*gclR* strain showed growth on 20 mM EG as well to an A_{600nm} value of 0.57 after 20 h (Figure 33B).

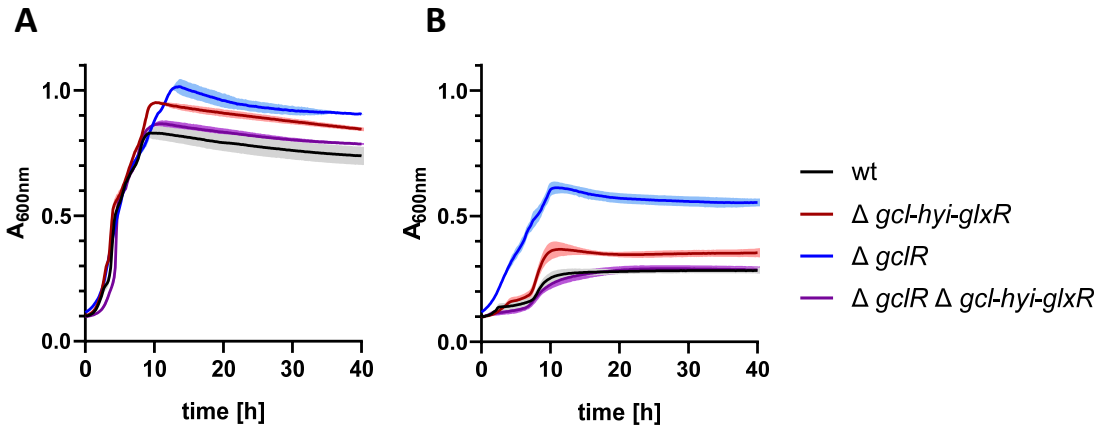


Figure 33. Growth of *P. putida* KT2440 strains in M9 minimal medium.
The shown A_{600nm} values do not represent OD_{600nm} values as they were not measured at a path length of 1 cm in a plate reader. (**A**) 20 mM glucose as sole carbon source. (**B**) 20 mM ethylene glycol as sole carbon source.

Unexpectedly on the other hand, the wt strain as well as the double knockout strain showed some growth on EG to an A_{600nm} value of 0.28 after 20 h. This might as well be explained with impurities in the used EG or technical problems with the plate reader, and this result should be confirmed in further experiments. The Δ*gcl* operon knockout strain showed even a little bit more growth to an A_{600nm} value of 0.34 after 20 h. This suggested that there might be another pathway involved in the degradation of EG in *P. putida* KT2440 than the glycerate pathway, because *P. putida* KT2440 has only one gene copy for the key enzyme of the glyoxylate carboligase (*gcl*) that had been knocked-out in the respective strain. This hypothesis should be investigated further.

To test the expression and activity of the enzymes of the tartronyl-CoA pathway in *P. putida* KT2440, an inducible plasmid-based system was chosen. To this end, a first expression plasmid was constructed, containing the *gcc* M5 and *birA* genes under the control of the 3-methylbenzoate inducible *xylS-Pm* promoter (pTE3608). After

transformation into *P. putida* KT2440 and induction with 0.5 mM 3-methylbenzoate for 4 h cell extracts were prepared. On an SDS gel of the cell lysate, the bands for GccA (72 kDa) and GccB (55 kDa) were visible and the presence of the respective proteins was confirmed by MALDI peptide mass fingerprinting (J. Kahnt, Core Facility for Mass spectrometry and Proteomics, MPI Marburg) (Figure 34A). In the cell lysate of cells containing pTE3608, a specific GCC activity of 11 ± 3 nmol min^{-1} mg^{-1} was detected, which was not present in the cell lysate of cells without the plasmid (Figure 34B). This demonstrated that the key enzyme of the tartronyl-CoA pathway, GCC, could be functionally expressed in *P. putida*.

Taken together, these first preliminary results form the basis for metabolic engineering using the tartronyl-CoA pathway.

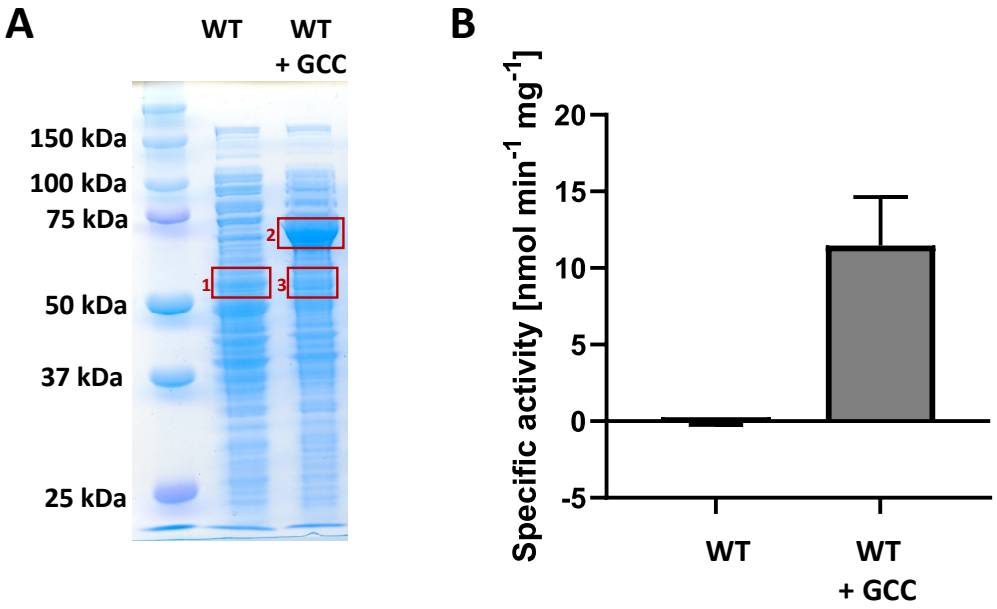


Figure 34. Expression and activity of glycolyl-CoA carboxylase in *P. putida* KT2440.
(**A**) SDS gel of *P. putida* KT2440 cell extracts with (WT + GCC) or without (WT) a plasmid (pTE3608) containing the GCC M5 complex (α- and β-subunit) and biotin ligase (BirA). Bands marked with a red box were analyzed by MALDI peptide fingerprinting (J. Kahnt, Core Facility for Mass spectrometry and Proteomics, MPI Marburg). Band 1 served as a negative control, band 2 was identified as *M. extorquens* PccA (α) subunit (72 kDa), band 3 was identified as *M. extorquens* PccB (β) subunit (55 kDa). (**B**) Specific activity of GCC M5 in *P. putida* KT2440 cell extracts measured spectrophotometrically coupled to TCR.

3. Discussion

3.1 Engineering enzymes for synthetic pathways

One of the most important tasks in the present work was to find enzymes that were capable of performing the required non-natural reactions for the tartronyl-CoA pathway, and furthermore, engineer these enzymes towards higher catalytic efficiencies. As there were no enzymes described for the necessary activities, the engineering of enzymes was dependent on the identification of substrate promiscuity as a starting point. Enzymes are often described as specialized catalysts with a highly specific function in metabolism. Nevertheless, the concept of enzyme promiscuity was first described already in 1976. It is based on the assumption that all enzymes evolved from few ancestors with broad substrate ranges that became more and more specialized over the course of evolution and the emergence of more metabolic pathways. Nevertheless, enzymes still retained the capability of performing more than just their specific physiological reaction (Jensen, 1976; Khersonsky and Tawfik, 2010). In the case of the realization of the non-natural reactions of the tartronyl-CoA pathway, enzymes performing analogous reactions with similar substrates were tested for their promiscuous activities. Most of the intermediates of the tartronyl-CoA pathway are non-natural and do not take part in known metabolic reactions. Tartronyl-CoA is not known as a natural metabolite, and for glycolyl-CoA there was only indirect evidence that it might be present in rat liver, as an intermediate of fatty acid metabolism (Vamecq and Poupaert, 1990). The advantage of engineering enzymes towards reactions with non-natural substrates is that during evolution, there was no selective pressure to discriminate against reactivity with these substrates (Khersonsky and Tawfik, 2010). This is especially obvious in the case of TCR. This enzyme is physiologically a malonyl-CoA reductase, but shows an affinity for tartronyl-CoA, which is comparable with the affinity to its native substrate, malonyl-CoA. On the contrary, malonyl-CoA reductase is not active with methylmalonyl-CoA (J. Zarzycki, personal correspondence), a naturally occurring intermediate, although it is structurally very similar to tartronyl-CoA. This shows that during evolution the enzyme “learned” to discriminate between malonyl-CoA and methylmalonyl-CoA, while the discrimination against tartronyl-CoA was simply not necessary and is therefore not present in the enzyme.

Several methods for laboratory evolution have been developed, which are all built on nature's ability to evolve new functions through the introduction of mutations. It has been shown in several cases that already single exchanges of amino acids within the active site can change an enzyme's specificity (Toscano et al., 2007). This is also the case of the GCS developed in this work. The exchange of an active site valine to alanine in an acetyl-CoA synthetase opened up the active site to allow better accommodation of the slightly bigger glycolate, leading to an overall improvement in app. K_M of almost 3-fold. In comparison to a recently published glycolyl-CoA synthetase (Trudeau et al., 2018), GCS showed more than an order of magnitude improvement in catalytic efficiency. Other methods for directed evolution include site saturation mutagenesis or random mutagenesis, which have been widely used for the realization of non-natural reactions. Recent examples include the evolution of a cytochrome c for the catalysis of carbon-silicon bond formation (Kan et al., 2016), an artificial metalloenzyme for olefin metathesis (Jeschek et al., 2016), a "formolase" enzyme for the carboligation of three formaldehyde molecules into dihydroxyacetone (Siegel et al., 2015), an artificial aldolase with a novel catalytic tetrad (Obexer et al., 2017) or an amine oxidase for the conversion of a novel substrate (Debon et al., 2019). The latter two enzymes have been evolved using microfluidics, a technique that was also employed for the evolution of GCC. The most important prerequisite for high-throughput screens of enzyme libraries is the coupling of genotype and phenotype. It was shown, that this can be achieved by compartmentalization of transcription and translation in the aqueous phase of water-in-oil emulsions (Tawfik and Griffiths, 1998). This principle was successfully applied to screen for more active horseradish peroxidase variants that convert a non-fluorescent substrate into its fluorescent form in a microfluidic lab-on-the-chip approach, applying FADS (Baret et al., 2009; Agresti et al., 2010). Microfluidics approaches have several advantages over high-throughput screenings using microtiter plates and robots. While by state-of-the-art robotic screenings, about one assay per second can be performed, thousands of droplets can be processed by using microfluidic chips in the same time. Furthermore, only very small reaction volumes are required. To screen a library of $5x10^7$ enzyme variants in the microfluidics device, only 150 µL of reagent solution are required, while in a robotic microtiter assay for the same amount of reactions, 5000 liters would be needed. The total cost savings in microfluidics compared to robotic microtiter assays are million-fold (Agresti et al., 2010). Despite these obvious advantages and the above mentioned

examples of successful application of microfluidics for enzyme evolution, it is still not yet a “standard” laboratory technique. This is probably first due to the technical equipment that is needed and that is at least in part not commercially available. Furthermore, for each enzyme to be evolved, a complete system of assay conditions, sample preparation, and downstream processing needs to be developed. This was also the case for the screening of GCC variants using microfluidics, for which many hurdles needed to be overcome. The enzymatic assay of GCC coupled to subsequent reduction of the carboxylation product by TCR was the first example of a coupled enzymatic assay to be performed in microfluidic droplets with an NADPH fluorescence read-out. Therefore, it was first necessary to establish measurements of NADPH fluorescence in droplets and transfer the whole enzymatic assay into droplets. Furthermore, the initial carboxylation activity of GCC M3 (the evolutionary starting point for the microfluidics screen) was rather low, which meant that an overnight incubation of the droplets was needed. Despite great efforts, it was not possible to recover the plasmid DNA of improved variants from the sorted droplets. This was probably due to the long overnight incubation of the droplets containing lysed cells, during which nucleases might have destroyed the plasmids at least partially. Now, with GCC M5, that is more than 6-times faster than GCC M3, the droplet incubation time could be substantially shortened, meaning that a successful recovery of the plasmid DNA becomes more likely. Furthermore, it would probably make sense for further attempts of GCC evolution employing microfluidics, to use a higher copy number plasmid than pCDFDuet (~ 30 copies per cell). In conclusion, microfluidics has a huge potential for enzyme evolution and more studies will help to standardize protocols and speed up the development of new microfluidic assays. For GCC, microfluidics was very helpful to screen the full libraries and confirm the presence of useful mutations in these libraries.

Taken together, only the combination of the initial rational design of the active site together with the use of high-throughput techniques led to the successful engineering of a new carboxylase, making both approaches equally valuable for the evolution of GCC. For further engineering of GCC, it might be advantageous to test more PCC homologs systematically. Introducing the mutations of GCC M5 into other PCCs, might lead to the discovery of a GCC with even better kinetic properties.

3.2 The tartronyl-CoA pathway as photorespiratory bypass

In this work the tartronyl-CoA pathway was realized *in vitro* as a novel pathway for carbon fixation. Initially, it was developed as a carbon-positive bypass for photorespiration. The latter is highly energy-demanding; it requires eleven different enzymes and consumes a total number of ATP and reducing equivalents of 13 per net fixation of one CO_2 (Figure 35; for calculations, see Table 8 and Table S1). Furthermore, natural photorespiration is only 75 % carbon efficient, as one already fixed carbon is released per turn of the pathway (Figure 35A).

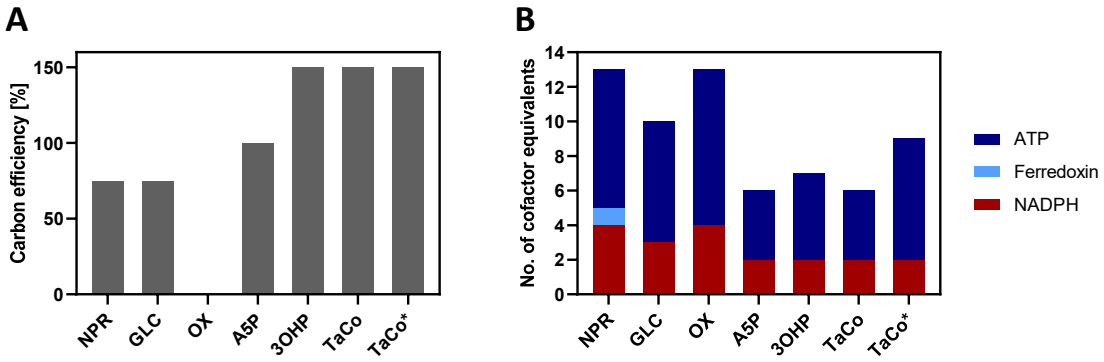


Figure 35. Carbon efficiency and energy requirements of different photorespiratory pathways.
(**A**) Carbon efficiency relates to one complete turn of each of the photorespiratory pathways. For the calculations, it was assumed that natural (canonical) photorespiration, the glycerate bypass and the glycolate pathway start with 2 molecules of 2-phosphoglycolate (i.e. with two oxygenation events), while the arabinose-5-phosphate shunt, the 3-OH-propionate bypass and the tartronyl-CoA pathway start with 1 molecule of 2-phosphoglycolate (i.e. with one oxygenation event). (**B**) Cofactor requirements of photorespiratory pathways per net fixation of one CO_2 (see Table S1 for calculations).
NPR: natural photorespiration, GLC: glycerate bypass, OX: glycolate oxidation pathway, A5P: arabinose-5-phosphate shunt, 3OHP: 3-hydroxypropionate bypass, TaCo: tartronyl-CoA pathway, TaCo*: tartronyl-CoA pathway including futile ATP hydrolysis.

The tartronyl-CoA pathway requires less than half of the cofactors in comparison to natural photorespiration per net fixation of one CO_2, in total only 2 NADPH and 4 ATP equivalents. If the futile ATP hydrolysis by GCC is taken into account (~ 4 ATP equivalents are used per carboxylation), the total number of cofactors is still 40 % less than in natural photorespiration (Figure 35, Table 8).

Furthermore, only five instead of eleven enzymes are involved in the conversion of 2-PG to 3-PGA via the tartronyl-CoA pathway. This is an importan point, which needs to be considered when thinking about energy requirements for a certain metabolic pathway,

because the cost for protein synthesis can be estimated to 4.5 to 5.9 ATP equivalents per amino acid (Zerihun et al., 1998; Amthor, 2000).

Table 8. Comparison of natural photorespiration and different photorespirational bypasses.

	NPR	**GLC (Kebeish et al., 2007)**	**OX (Maier et al., 2012)**	**A5P (Trudeau et al., 2018)**	**3OHP (Shih et al., 2014)**	**TaCo** (this work)
CO_2 release	yes	yes	yes	no	no	no
Place of CO_2 release	mitochondria (far RuBisCO)	chloroplast (near RuBisCO)	chloroplast (near RuBisCO)	no CO_2 released	no CO_2 released	no CO_2 released
Carbon efficiency	**75 %**	**75 %**	**0 %**	**100 %**	**150 %**	**150 %**
Turns of Calvin cycle‡	2 turns	2 turns	3 turns	none	none	none
Way of CO_2-refixation	complete CBB cycle	complete CBB cycle	complete CBB cycle	part of CBB cycle	included in bypass	included in bypass
NH_3 release	yes	no	no	no	no	no
no. of enzymes needed	11	5	6	5	12	5
ATP consumed‡	8	7	9	4	5	4 / 7*
Redox power consumed‡	4 NAD(P)H + 2 Fd_{red}	3 NAD(P)H	4 NAD(P)H	2 NAD(P)H	2 NAD(P)H	2 NAD(P)H

NPR: natural photorespiration, GLC: glycerate bypass, OX: glycolate oxidation pathway, A5P: arabinose-5-phosphate shunt, 3OHP: 3-hydroxypropionate bypass, TaCo: tartronyl-CoA pathway.
Carbon efficiency relates to one complete turn of each of the photorespiratory pathways. For the calculations, it was assumed that natural (canonical) photorespiration, the glycerate bypass and the glycolate pathway start with 2 molecules of 2-phosphoglycolate (i.e. with two oxygenation events), while the arabinose-5-phosphate shunt, the 3-OH-propionate bypass and the tartronyl-CoA pathway start with 1 molecule of 2-phosphoglycolate (i.e. with one oxygenation event).
* The values include the unfruitful ATP hydrolysis of the GCC M5 variant. ‡ per net fixation of one CO_2

Several photorespiratory bypasses have been proposed and realized previously (Table 8). One of the first approaches was the introduction of the glycerate pathway, including glycolate dehydrogenase, glyoxylate carboligase and tartronate semialdehyde reductase (Figure 36) into plants. The overall energy requirement for the glycerate bypass per net fixation of one CO_2 is 7 ATP and 3 NADPH equivalents, only 70 % compared to natural photorespiration, while the carbon efficiency is the same for both pathways (Figure 35, Table 8). The bypass was implemented into *A. thaliana* (Kebeish et al., 2007) and later

in the oilseed crop *Camelina sativa* (Dalal et al., 2015), as well as in *N. tabacum* (tobacco) (South et al., 2019).

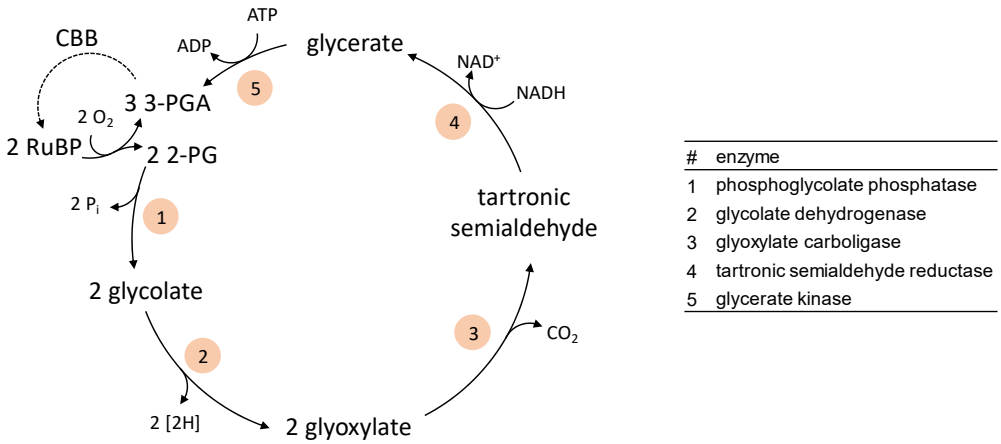


Figure 36. Glycerate bypass.
The physiological electron acceptor of glycolate dehydrogenase (2) is not known. The bypass was implemented into *A. thaliana* (Kebeish et al., 2007), *C. sativa* (an oilseed crop) (Dalal et al., 2015) and *N. tabacum* (South et al., 2019). CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 2-PG: 2-phosphogylcolate, 3-PGA: 3-phosphoglycerate.

In all three studies, the expression of the glycerate bypass *in planta* led to increased biomass compared to wildtype plants. In tobacco, a biomass increase of 13 % was reported (South et al., 2019), while in *C. sativa*, the seed yield increased by 57 - 73 % (Dalal et al., 2015). Remarkably, also the introduction of only glycolate dehydrogenase led to increase in biomass in several studies (Kebeish et al., 2007; Nölke et al., 2014; Dalal et al., 2015). Therefore, the question remains which mechanism exactly led to the biomass increase, if the conversion of glycolate into glyoxylate was apparently sufficent to improve plant performance under photorespiratory conditions.

Another pathway was developed that completely oxidizes glycolate into CO_2 (Fahnenstich et al., 2008; Maier et al., 2012), leading to a complete loss of 2-PG, but higher CO_2 concentrations around RuBisCO. The cofactor costs for the glycolate oxidation pathway per net fixation of one CO_2 are 7 ATP and 3 NADPH equivalents, almost the same as for natural photorespiration. The carbon efficiency is 0 %, because of the complete oxidation of glycolate (Figure 35, Table 8). For the functional implementation of the glycolate oxidation pathway (Figure 37), it was necessary to

introduce glycolate oxidase, malate synthase and catalase into *A. thaliana*. The other required enzymatic activities (phosphoglycolate phosphatase, NADP-malic enzyme and pyruvate dehydrogenase) are already present in the plant chloroplast. In *A. thaliana*, the pathway led to a rosette fresh weight increase of 28 % under photorespiratory conditions in the greenhouse (Maier et al., 2012). Notably, this pathway was also introduced into *N. tabacum* and tested in the field in two subsequent growth seasons. Plants expressing the glycolate oxidation pathway in combination with suppression of the native glycolate transporter using RNA interference (RNAi), showed a biomass increase of 20 - 24 % (South et al., 2019). This showcased that photorespiratory bypasses have the potential to improve plant performance in an agricultural environment, even if they require only slightly less energy than natural photorespiration or completely oxidize 2-PG into CO_2 without carbon recycling.

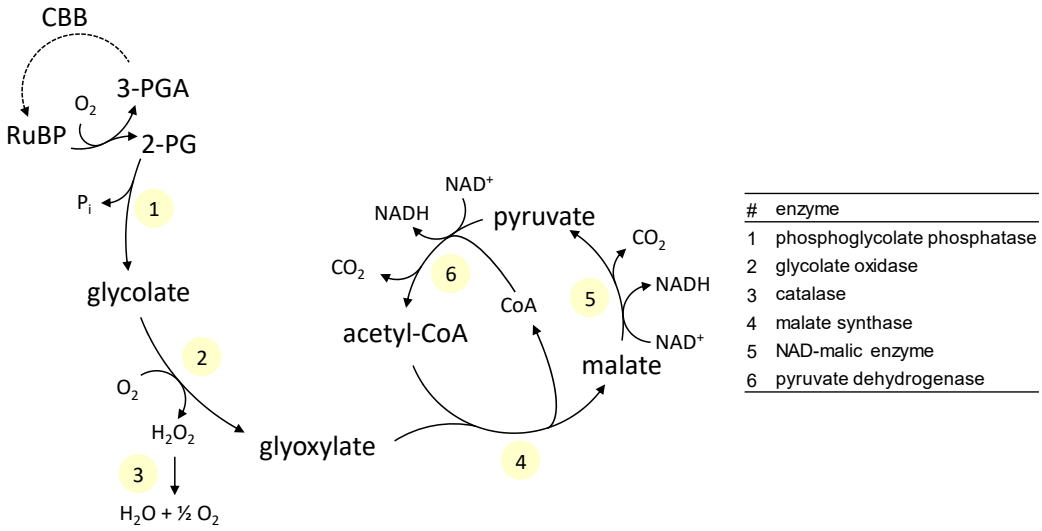
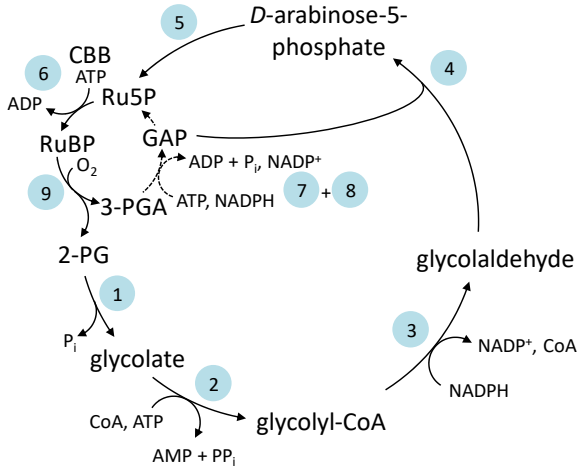


Figure 37. Glycolate oxidation pathway.
The pathway was first implemented into *Arabidopsis thalina* (Maier et al., 2012). A variation was introduced into *N. tabacum* with glycolate dehydrogenase instead of glycolate oxidase (South et al., 2019). CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 2-PG: 2-phosphogylcolate, 3-PGA: 3-phosphoglycerate.

Both the glycerate bypass as well as the glycolate oxidation pathway have the disadvantage that CO_2 is released in the process of 2-PG recycling (Figure 35). In an ideal photorespiratory bypass, the release of CO_2 should be avoided. A possible carbon-neutral alternative for photorespiration was recently proposed: the arabinose-5-phosphate shunt

(Trudeau et al., 2018). In this synthetic photorespiratory pathway that is not known to occur in nature, 2-PG is converted into glycolate, which is activated into glycolyl-CoA by an engineered glycolyl-CoA synthetase (ACS19). This is also the first non-natural step of the tartronyl-CoA pathway. Glycolyl-CoA is then reduced to glycolaldehyde by an engineered glycolyl-CoA reductase (GCR) and glycolaldehyde is condensed with glyceraldehyde-3-phosphate (GAP) to form arabinose-5-phosphate (A5P). A5P is then converted into ribulose-5-phosphate (Ru5P), and thereby fed back as an intermediate into the CBB cycle (Figure 38). The arabinose-5-phosphate shunt was the first carbon-neutral photorespiratory bypass that was realized *in vitro* from new-to-nature engineered enzymatic activities. It was shown that the arabinose-5-phosphate shunt was able to produce RuBP from glycolate *in vitro* at a rate of 0.05 s^{-1}, demonstrating its potential as photorespiratory bypass (Trudeau et al., 2018).

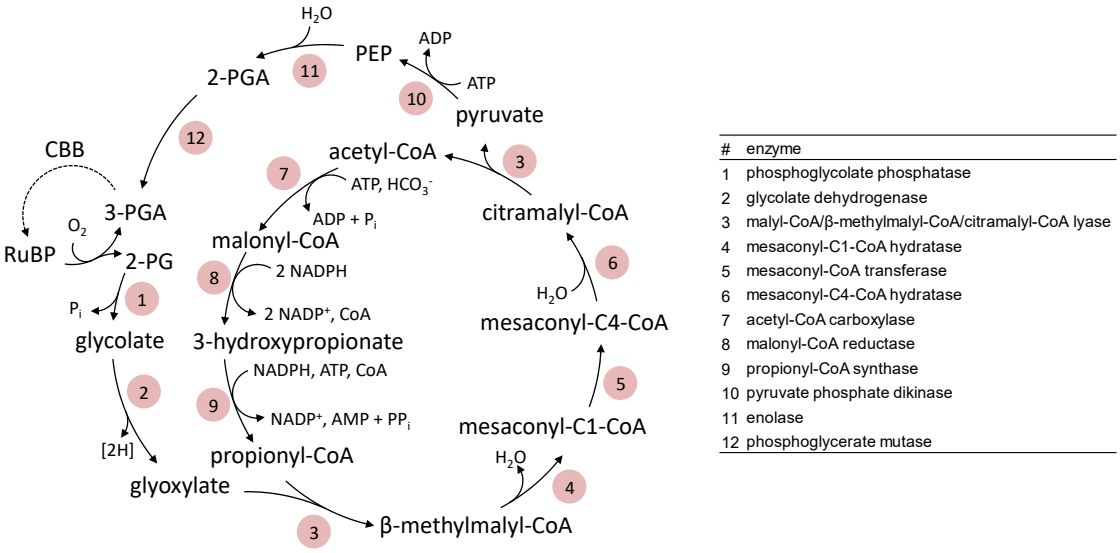


#	enzyme
1	phosphoglycolate phosphatase
2	glycolyl-CoA synthetase
3	glycolyl-CoA reductase
4	fructose-6-phosphate aldolase
5	arabinose-5-phosphate isomerase
6	phosphoribulokinase
7	phosphoglycerate kinase
8	glyceraldehyde-3-phosphate dehydrogenase
9	RuBisCO

Figure 38. Arabinose-5-phosphate shunt.
The pathway contains two engineered enzymes ((2) and (3)) and was shown to operate *in vitro* (Trudeau et al., 2018). CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 2-PG: 2-phosphogylcolate, 3-PGA: 3-phosphoglycerate, GAP: glyceraldehyde 3-phosphate, Ru5P: ribulose-5-phosphate.

The first carbon-fixing photorespiratory bypass was the 3-OH-propionate bypass (Shih et al., 2014). It consists of one half-cycle of the natural CO_2 fixation 3-OH-propionate bicycle that was first described in the green non-sulfur bacterium *C. aurantiacus* (Holo and Sirevåg, 1986; Zarzycki et al., 2009). The whole pathway for the conversion of 2-PG to 3-PGA consists of 12 enzymes (Figure 39), which is in the range of the enzymes

required for natural photorespiration. Because of its CO_2 fixation step via acetyl-CoA carboxylase, the energy requirement for the net fixation of one CO_2 is only 5 ATP and 2 NADPH equivalents, which is only 50 % of that of natural photorespiration. The 3-OH-propionate bypass has been introduced into the cyanobacterium *Synechococcus elongatus* and activities for all six introduced enzymes could be measured. However, cyanobacteria expressing the pathway did not show any growth advantage compared to the wildtype strain (Shih et al., 2014), which was likely due to only low levels of endogenous acetyl-CoA carboxylase activity.



#	enzyme
1	phosphoglycolate phosphatase
2	glycolate dehydrogenase
3	malyl-CoA/β-methylmalyl-CoA/citramalyl-CoA lyase
4	mesaconyl-C1-CoA hydratase
5	mesaconyl-CoA transferase
6	mesaconyl-C4-CoA hydratase
7	acetyl-CoA carboxylase
8	malonyl-CoA reductase
9	propionyl-CoA synthase
10	pyruvate phosphate dikinase
11	enolase
12	phosphoglycerate mutase

Figure 39. 3-hydroxypropionate bypass.
The pathway consists of one half-cycle of the 3-hydroxypropionate bicycle (Zarzycki et al., 2009) and was implemented into the cyanobacterium *Synechococcus elongatus* (Shih et al., 2014). CBB: Calvin-Benson-Bassham cycle, RuBP: ribulose-1,5-bisphosphate, 2-PG: 2-phosphogylcolate, 3-PGA: 3-phosphoglycerate, PEP: phosphoenolpyruvate.

In comparison, all previously proposed and tested photorespiratory bypasses have decreased energy requirements in form of ATP and NAD(P)H than natural photorespiration when normalized to the net fixation of one CO_2 (Figure 35, Table 8). Notably, the tartronyl-CoA pathway and the arabinose-5-phosphate shunt (Figure 38) have the lowest energy requirement of 4 ATP and 2 NADPH equivalents, which is only half of the energy requirement for natural photorespiration (Figure 35). The additional benefit of the tartronyl-CoA pathway over the arabinose-5-phosphate shunt is its carbon-fixation step by GCC, turning photorespiration into a CO_2-fixing pathway that could

support carbon fixation by the CBB cycle (Figure 35A). Both the arabinose-5-phosphate shunt and the tartronyl-CoA pathway have been realized *in vitro*, and their introduction into living organisms (especially plants) will reveal, if their advantages in the *in vitro* system can be transferred *in planta*. However, the yield-enhancing effects of the glycerate bypass as well as the glycolate oxidation pathway in different species and even under field conditions give rise to the expectation, that also the tartronyl-CoA pathway has the potential to improve agricultural yields, especially because of its low energetic requirements and its net carbon fixation.

However, what also needs to be taken into account, are the cofactor requirements of the tartronyl-CoA pathway. First, it involves CoA esters of organic acids, whereas natural photorespiration is independent of CoA ester intermediates. The concentration of all CoA (including free CoA and CoA esters) has been determined to be 31 - 54 µM in chloroplasts, mostly in the form of acetyl-CoA. Only a fraction of the total CoA pool comprises free CoA (< 2 µM) (Post-Beittenmiller et al., 1992). It was shown in this work that GCS is functional at a CoA concentration of 5 µM, which is supposedly close to the expected chloroplast CoA concentration. Nevertheless, the question remains, if shifting of CoA pools towards the synthesis of glycolyl-CoA will affect the normal operation of the plant cell. Furthermore, GCC requires biotin as a cofactor. In plants, the biosynthesis of fatty acids takes place within the chloroplast (Brown et al., 2009) and requires the carboxylation of acetyl-CoA to form malonyl-CoA, which is carried out by a biotin-dependent acetyl-CoA carboxylase (Alban et al., 2000). On the one hand, this supports the assumption that the carboxylation of glycolyl-CoA through the action of a biotin-dependent carboxylase is in principle possible within plants, which also have an estimated cytosolic pool of free biotin of 11 µM (Baldet et al., 1993; Alban et al., 2000). On the other hand, the possibility arises that the heterologous expression of GCC *in planta* leads to an alteration or depletion of the biotin pool within the plant cell, which is deeply intertwined with the essential synthesis of fatty acids, including its regulation (Alban et al., 2000). Along the same lines, the introduction of TCR might also interfere with plant fatty acid metabolism, because this enzyme has a native malonyl-CoA reductase activity, which is more than 10-fold higher than its tartronyl-CoA reductase activity (10 µmol min^{1} mg^{-1} (Hügler et al., 2002)). This activity might lead to the accumulation of the reaction product, 3-hydroxypropionate, as it is a dead-end metabolite in plants.

Taken together, these considerations led to the conclusion, that a successful implementation of the tartronyl-CoA pathway *in planta* is conceivable, as all required cofactors are present. In order to achieve minimal influence on physiological plant vitality, while still enabling a yield enhancement, the implementation might require careful balancing of gene expression. For instance, the gene for biotin ligase (*birA*) would probably only require lower expression levels, because it would only be needed in lower amounts for biotinylation of GCC. Furthermore, low expression levels of *birA* would ensure minimal influence on the native biotin pool. Similarly, GCC and TCR should be produced in levels high enough to support flux through the tartronyl-CoA pathway, while at the same time in levels low enough to avoid detrimental effects on plant fitness as discussed above. On the other hand, the *gcs* gene should be expressed in higher levels, as it is currently the enzyme with the lowest catalytic efficiency of the tartronyl-CoA pathway.

3.3 *In vivo* implementation of the tartronyl-CoA pathway

After the successful *in vitro* demonstration of the tartronyl-CoA pathway in this work, the next step is the *in vivo* implementation of the tartronyl-CoA pathway in order to harness its full potential.

For an *in vivo* implementation of synthetic CO_2 fixation pathways, it is necessary to connect them to central carbon metabolism. The primary output of the CETCH cycle is glyoxylate, as it is for other proposed synthetic CO_2 fixation cycles (Bar-Even et al., 2010; Schwander et al., 2016). To channel glyoxylate into central carbon metabolism (e.g. gluconeogenesis) it is essential to convert it to glycerate. However, there are only very few natural metabolic routes that allow the direct conversion of glyoxylate or glycolate into glycerate and all of them result in the loss of carbon. The recently described β-hydroxyaspartate cycle (Schada von Borzyskowski et al., 2019) yields the C_4 compound oxaloacetate, which needs to be decarboxylated to generate C_3 metabolites. Similarly, photorespiration and the glycerate pathway convert two glyoxylate molecules into glycerate under the release of CO_2 (Krakow et al., 1961; Bauwe et al., 2010). The loss of CO_2 strongly limits the carbon efficiency of these natural pathways. The formation of 3-PGA from CO_2 by the CETCH cycle coupled to the glycerate pathway requires two turns of the CETCH cycle and leads to the loss of already fixed CO_2 through the Gcl reaction. This makes the overall process only 75 % carbon efficient. The conversion of the produced glyoxylate via the tartronyl-CoA pathway requires only one turn of the CETCH cycle and is furthermore carbon preserving, as no CO_2 is released (Table 9). Thus, reducing energy demand and circumventing carbon loss during glycolate assimilation by using the tartronyl-CoA pathway instead could enhance productivity of synthetic CO_2 fixation by the CETCH or other synthetic CO_2 fixation cycles significantly. In this study, the successful coupling of the CETCH and tartronyl-CoA pathway was demonstrated, a promising basis for future *in vivo* implementations

Table 9. Comparison of energy and CO_2 fixation efficiencies of the CETCH cycle coupled to the glycerate pathway or the tartronyl-CoA pathway (TaCo).

	CETCH (v5.4) + glycerate pathway	CETCH (v5.4) + TaCo
CO_2 fixed	4	3
CO_2 release	1	0
CO_2 fixation efficiency	**75 %**	**100 %**
turns of CETCH required	2	1
ATP consumed	3	5
Redox power consumed	9 NAD(P)H	7 NAD(P)H

The values correspond to the formation of 3-phosphoglycerate (3-PGA).

Another *in vivo* application of the tartronyl-CoA pathway would be its introduction into *P. putida*, as initially started in the present work. It could serve as alternative, carbon-fixing option for the assimilation of EG, a plastic waste component and environmental pollutant. Plastic waste has become a worldwide problem within the last decades, with dramatic effects on terrestrial and marine ecosystems. It can accumulate in the form of microplastic particles, which pose a threat especially in the oceans, entering the food chain (Eriksen et al., 2014; Salvador et al., 2019). Recently, polyethylene terephthalate (PET) degrading hydrolases from the bacterium *Ideonella sakaiensis* have been described and characterized (Yoshida et al., 2016; Han et al., 2017). They are able to decompose PET into its monomer components terephthalic acid and EG and have been heterologously produced in different bacteria and eukaryotic algae, with the prospect of serving as systems for efficient PET remediation (Huang et al., 2018; Moog et al., 2019; Seo et al., 2019). It would be conceivable, to co-culture such strains together with engineered *P. putida* (equipped with the tartronyl-CoA pathway) for assimilation of the derived EG. Even thinking one step further, it would then be advantageous to convert EG into value-added compounds, such as biofuels, pharmaceuticals, chemical building blocks or bioplastics such as polyhydroxyalkanoates, as has been shown in *P. putida* using the glycerate pathway (Franden et al., 2018). Employing the tartronyl-CoA pathway for plastic degradation would therefore provide a sustainable solution for the recycling of PET, while additionally fixing CO_2.

Furthermore, the implementation of the tartronyl-CoA pathway into *P. putida* opens up more possibilities for the further evolution of the pathway enzymes. When growth of the engineered strain on EG is dependent on the tartronyl-CoA pathway, it will be possible to introduce enzyme libraries of GCC or TCR and screen for faster growth. This will be especially helpful to screen for GCC variants with higher ATP efficiency, as the availability of ATP for the strain is directly coupled to a growth advantage. Moreover, the growth assay would probably also help to screen for TCR variants with a better discrimination between malonyl-CoA and tartronyl-CoA, because the activity of TCR with malonyl-CoA leads to the accumulation of the dead-end product 3-hydroxypropionate, while malonyl-CoA could be depleted from fatty acid biosynthesis. This implies that variants with a better discrimination between malonyl-CoA and tartronyl-CoA would lead to a faster growth of the strain, as the malonyl-CoA pool would remain available for fatty acid biosynthesis to a higher degree, while the flux through the tartronyl-CoA pathway would be enhanced at the same time.

Another option for the *in vivo* evolution of pathway enzymes was opened up by the recent development of several *E. coli* glycerate biosensor strains (Aslan et al., 2020). Through a series of gene knockouts, central carbon metabolism was dissected in order to achieve different degrees of dependency on glycerate for growth, as in each strain glycerate serves as a carbon source for a different fraction of cellular biomass. Using these strains for the further evolution of the tartronyl-CoA pathway enzymes could prove be a powerful tool for stepwise increasing the evolutionary pressure in order to evolve the enzymes for higher fluxes through the pathway.

3.4 Outlook

The *in vitro* and *in vivo* implementation of the tartronyl-CoA pathway is a good example of how engineering of new biological systems can help to better understand the existing systems. Every problem that needs to be overcome in the process of biological engineering gives us further insights into the biochemistry of life. For example, the work on the new-to-nature GCC led to a very high-resolution cryo-EM structure of this enzyme that allowed us to rationalize the introduced mutations. It will possibly also help to understand the mechanism of substrate carboxylation by biotin-dependent carboxylases better in general, as the exact mode of action is only hypothesized to date. Most probably, the carboxybiotin releases the CO_2 within the CT active site and the biotin itself then serves as the base for proton abstraction to form the enolate intermediate of the CoA thioester. Understanding the mechanism will also be important for understanding the futile ATP hydrolysis of GCC. It is not clear, why the engineered GCC is not completely efficient in carboxylation and hydrolyzes about four molecules of ATP per glycolyl-CoA carboxylation reaction. One possible reason might be the still incorrect positioning of glycolyl-CoA or the enolate intermediate within the active site. Supposedly, through the incorporation of the active site mutations (especially Y143H and L100S) in GCC, the glycolyl-CoA substrate is more tightly bound. Nevertheless, it might still be in a position that is not perfect for efficient carboxylation. Therefore, after the release of CO_2 within the CT active site, the biotin cofactor might not always be able to abstract the proton to form the enolate intermediate, and the CO_2 is lost. Another explanation might be that the enolate is always formed, but is simply not well stabilized, and gets quenched by reaction with e.g. a nearby water molecule before the CO_2 can be attacked. Furthermore, the CO_2 itself might be more loosely or incorrectly bound in GCC and may be too far away for an efficient attack of the enolate, because the overall architecture of the active site has changed. In order to solve the open question of the futile ATP hydrolysis, the now available high-resolution structure of GCC could be the basis for quantum mechanics/molecular mechanics (QM/MM) modelling as a tool for reconstructing the reaction mechanism of GCC and to indicate which of the discussed scenarios are most likely.

The development of new tools for the efficient conversion of CO_2 into value-added compounds is a crucial step towards a more sustainable future. The tartronyl-CoA

pathway, which was realized *in vitro* in the present work, could contribute to that in several areas. First, its introduction into agricultural relevant plants could enhance plant yields, in order to feed the ever-growing world population. Furthermore, it could serve as an additional carbon-fixing connection between other proposed carbon fixation pathways and central metabolism. In that way, the tartronyl-CoA pathway could support the biotechnological production of biomass or value-added compounds like chemical building blocks, pharmaceuticals, biofuels, food additives or bioplastics from CO_2. Along the same lines, the tartronyl-CoA pathway could be employed for the conversion of the plastic waste component EG, either into biomass or any of the above-mentioned biotechnological products.

These applications will rely on the *in vivo* implementation of the tartronyl-CoA pathway, for which copious challenges will have to be overcome, as the pathway includes non-natural metabolites and engineered enzymes that could lead to unwanted side reactions in living systems. Nevertheless, the tartronyl-CoA pathway exhibits considerable advantages as a carbon fixation pathway, which make its incorporation into living systems feasible. These include e.g. its insensivity to O_2 and its utilization of bicarbonate instead of CO_2, the former being more readily available due to its much higher solubility in water. Furthermore, the required enzymatic cofactors and cosubstrates (like biotin or NADPH) as well as CoA thioesters (intermediates of the tartronyl-CoA pathway) are ubiquitous in all domains of life. To sum up, careful analyses of the engineered systems will be needed in order to enable growth and carbon assimilation via the tartronyl-CoA pathway, but the obvious advantages will make the effort worth it.

In conclusion, the tartronyl-CoA pathway opens up the solution space of metabolism for the conversion of CO_2 into value-added compounds and provides new routes for numerous biotechnological applications, some of which we probably not even have considered yet.

4. Materials & Methods

4.1 Materials

Chemicals were obtained from Sigma-Aldrich, Carl Roth GmbH + Co. KG, Santa Cruz Biotechnology Inc. and Merck. $NaH^{13}CO_3$ was obtained from Cambridge Isotope Laboratories Inc. Biochemicals and materials for cloning and protein expression were obtained from Thermo Fisher Scientific, New England Biolabs GmbH and Macherey-Nagel GmbH. CoA was bought from Roche Diagnostics. Materials and equipment for protein purification were obtained from GE Healthcare, BioRad and Merck Millipore GmbH. Pyruvate Kinase/Lactic Dehydrogenase, Malic Dehydrogenase, Glucose-6-Phosphate Dehydrogenase, Glucose Dehydrogenase and Phosphoenolpyruvate carboxylase were bought from Sigma-Aldrich.

4.2 Synthesis of CoA esters

Glycolyl-CoA was synthesized as previously described (Trudeau et al., 2018). Tartronyl-CoA was enzymatically synthesized using the malonyl-CoA synthetase MatB (Peter et al., 2016). The preparation contained in a total volume of 10 mL 200 mM MOPS/KOH pH 7.8, 20 mM tartronate, 7.5 mM ATP, 10 mM $MgCl_2$, 5 mM CoA and 600 µg MatB. The reaction was performed for 18.5 h at room temperature, quenched with 1 % HCl and centrifuged at 17.000 × *g* for 20 min. CoA esters were purified using preparative HPLC-MS. The synthesis mixtures were applied to a preparative Agilent 1260 Infinity HPLC with a Gemini 10µm NX-C18 110 Å column (Phenomenex). Chromatographic separation was achieved using 25 mM ammonium formate pH 4.2 with a methanol gradient from 1 % to 30 % at a flow rate of 25 mL min^{-1}. Glycolyl-CoA or tartronyl-CoA fractions, respectively, were pooled, flash-frozen in liquid nitrogen, and lyophilized. The concentration of CoA-esters was quantified by determining the absorption at 260 nm (ε=16.4 mM^{-1} cm^{-1}).

4.3 Cloning of expression vectors

All *in silico* cloning was performed with Clone Manager 9 (Scientific & Educational Software). For purification, preparation, cloning, transformation and amplification of DNA, standard protocols were used (Sambrook and Russell, 2001). Plasmid isolation and

PCR-product purification was performed according to the manufacturer's protocols using the kits NucleoSpin® Plasmid (Macherey Nagel) and NucleoSpin® Gel and PCR Clean-up (Macherey Nagel). All primers used in this study were synthesized by Eurofins Genomics Germany GmbH. All primers used in this study can be found in Table 10 and all plasmids are listed in Table 11.

Table 10. Primers used in this study.

name	sequence (5' → 3')
ACS_E_L641P	CGGACCCGTCTCTGGTGGATCGTCCGATTGAGGGCCGTCAG
ACS_Ery_V379A	CTGCGTCTGCTGGGCTCTGTAGGTGAGCCGATCAAC
ACS_Ery_V379G	CCCTGCGTCTGCTGGGCTCTGGAGGTGAGCCGATCAAC
ACS_Ery_V379S	GTTCCCTGCGTCTGCTGGGCTCTTCAGGTGAGCCGATCAAC
BirA_Me_bb_for	TATAGAATTCAGATCTAGGAGATATACATATGCAGTTCCG
BirA_Me_bb_rev	TATACTCGAGGGATCCCTATCCAGTCTCAGGCCG
BirA_Me_EcoRI_mut	CTGCTCGAATTTCTGTCGGAAC
EG_birA_fw	GAATTCGCCTGACACACAGGAAACAGACCATGCAGTTCCGGCTAAGTCAG
EG_birA_rv	CATGCCTGCAGGTCGACTCTAGATCAGGCCGCCGTCGCGGCAC
EG_pccA_fw	CGCTCTGAATTAAAGAGGAGAAATTAACTATGTTCGATAAGATCCTGATTGC
EG_pccA_rv	GTTTCCTGTGTGTCAGGCGAATTCCAGGATCACGGCATC
EG_pccB_fw	CGGCCGCGCGAATTCGAGCTCTTAATAAGGAGATATACCATGAAGGACATCCTCGAGAAGCTTGAGG
EG_pccB_rv	ATTTCTCCTCTTTAATTCAGAGCGGGATGTTGTCGTGCTTCTTCCTG
gcl-hyi-glxR_TS1fw	GAGCTCGGTACCCAGGCTAGCGCCATCCTCGCGCTGCAGCG
gcl-hyi-glxR_TS1rv	CCCAGGGCCTGGCAGGTGTTCAACGGGTTGATGGCAGCCCCCGGGATG
gcl-hyi-glxR_TS2fw	AACACCTGCCAGGCCCTGGGCGGCGGCAACTGGGACCAC
gcl-hyi-glxR_TS2rv	GTCGACTCTAGACTCGACCACTTCGGCCATGGCAGCTGCGGCCTTG
gclR_check_fw	GATGCGAGAACACGGCATGCAAATC
gclR_check_out_fw	CGAACAGGGCCGGCCCCGTGCTAC
gclR_check_out_rv	CGTTTCCCAGTTGTAGCCCATGGTG
gclR_check_rv	GAACGAACAGCTGCAACCTCTGAAG
gclR_TS1fw	GCTCGGTACCCCGCAGACAGCTCGAAACCG
gclR_TS1rv	CTTCGTGCCTCGAGAGCCCTCGTTTGCCTGCGTGATCG
gclR_TS2fw	GGGCTCTCGAGGCACGAAGAGAAAGTATGGCGCCATTGTAC
gclR_TS2rv	GTCGACTCTAGAAGCGCGGCGGCGATGCAGGTGGCGAC
hyi_check_fw	CCATCTCGCTGCTGGACTGATCG
hyi_check_out_fw	GCACCCAGCACCTTGCCGTCCTGATC
hyi_check_out_rv	CGCGGTCTTCTTCGTTGAGGTTGCTG
hyi_check_rv	CCATGATGCCGGTGCCGATGAAAC
MC1	CTGACGGTACCTGCGACTAGTTCATTTATGGGGAAGATTGGCAACG
MC60	TGCACCATATGTCGAGTCCAAAGATCGG
patZ_KO_fw	GAACAGTTAGAAAGCGTTTAAAATCATTCGGTCACTTCTGCGGGAGACCGGTAATTAACCCTCACTAAAGGGCG
patZ_KO_rv	CAATAATAACACCAGTCCCGATTAAGTGGTCAACATTTCCAGTACCTTACTAATACGACTCACTATAGGGCTC
PCC_G170A	CGTGCGCGGGCGCCGACGTGTACTCGCCG
PCC_G170S	CGTGCGCGGGCAGCGACGTGTACTCGC
PCC_G170T	CGTGCGCGGGCACCGACGTGTACTCGC
PCC_L100D	GTTCGGCGGCTCGGACTCCGAGGCGCACGCAG
PCC_L100N	GTTCGGCGGCTCGAACTCCGAGGCGCACGCAG
PCC_L100S	GTTCGGCGGCTCGTCCTCCGAGGCGCACGCAG
PCC_satmut_22c_1	CCAGGAGGGCGTGGCCGCG**NDT**GGCGGC**NDT**GGCGAGGTGTTCCG
PCC_satmut_22c_2	CCAGGAGGGCGTGGCCGCG**VHG**GGCGGC**VHG**GGCGAGGTGTTCCG
PCC_satmut_22c_3	CCAGGAGGGCGTGGCCGCG**NDT**GGCGGC**VHG**GGCGAGGTGTTCCG
PCC_satmut_22c_4	CCAGGAGGGCGTGGCCGCG**VHG**GGCGGC**NDT**GGCGAGGTGTTCCG
PCC_satmut_22c_5	CCAGGAGGGCGTGGCCGCG**NDT**GGCGGC**TGG**GGCGAGGTGTTCCG
PCC_satmut_22c_6	CCAGGAGGGCGTGGCCGCG**TGG**GGCGGC**NDT**GGCGAGGTGTTCCG
PCC_satmut_22c_7	CCAGGAGGGCGTGGCCGCG**VHG**GGCGGC**TGG**GGCGAGGTGTTCCG
PCC_satmut_22c_8	CCAGGAGGGCGTGGCCGCG**TGG**GGCGGC**VHG**GGCGAGGTGTTCCG
PCC_satmut_22c_9	CCAGGAGGGCGTGGCCGCG**TGG**GGCGGC**TGG**GGCGAGGTGTTCCG
PCC_satmut_22c_rv	CGCGGCCACGCCCTCCTGGATGCGCGCACCGC

name	sequence (5' → 3')
PCC_Y143H	GCGCTCGGCGGCCACGGCGAGGTGTTCCGC
PccB_fw_P1	GTTTAACTTTAATAAGGAGATATACCATGGGCAGCAGCCATC
PccB_rv_P1	GATTACTTTCTGTTCGACTTAAGCATTATGCGGCCGCAAG
PCCMe_D171A	CGTGCGCGGGCGGCGCCGTGTACTCGC
PCCMe_D171V	CGTGCGCGGGCGGCGTCGTGTACTCGC
PCCtoACC_D407I	CAAGGCCTTCGGCGGCGCCTACATCGTCATGGCCTCCAAGCATG
PCCtoGCC_D407K	CAAGGCCTTCGGCGGCGCCTACAAAGTCATGGCCTCCAAGCATG
PCCtoGCC_L140D	GAGGGCGTGGCCGCGGACGGCGGCTACGGC
PCCtoGCC_L140N	GAGGGCGTGGCCGCGAACGGCGGCTACGGC
PCCtoGCC_Y143E	CGCGCTCGGCGGCGAAGGCGAGGTGTTCCGCCG
PCCtoGCC_Y143Q	CGCGCTCGGCGGCCAAGGCGAGGTGTTCCGCCG
pET16b_NcoI2NdeI	CTTTAAGAAGGAGATATACATATGGCCATCATCATCATC

D: A, G or T; **H**: A, C or T; **M**: A or C, **N**: random base, **V**: A, C or G

The plasmids of propionate CoA transferase from *Clostridium propionicum* (pET-16b_PCT) and *Ralstonia eutropha* (pET-19b::pct) were kind gifts from I. Berg, Münster and A. Steinbüchel, Münster, respectively. The plasmid of the water-forming NADH oxidase from *L. pentosus* (pET28a-NH-nox-L.p.) was a kind gift from V. Sieber, Straubing. The plasmid for the AMP-regenerating polyphosphate kinase (pET28a-AjPKK2) was a kind gift from J. Andexer, Freiburg. For the construction of the expression plasmid for 4-OH-butyrate CoA transferase (AbfT) from *C. aminobutyricum* (pTE1138), the gene was ordered codon-optimized and cloned into a pET16b derivative using XbaI and KpnI for restriction (cloning performed by Martina Carrillo). The genes encoding for *Ery*ACS1 and *Ery*ACS2 were synthesized codon-optimized for *E. coli* by the DOE Joint Genome Institute (JGI) and cloned into the backbone pSEVA141 (ACS1, pTE1007) or pSEVA261 (ACS2, pTE1008), respectively. For the construction of the expression vectors for the ACS1 mutants ACS1_V379G (pTE1417), ACS1_V379S (pTE1418), ACS1_V379A (pTE1434) and ACS1_V379A_L641P (pTE1427), single-oligo directed mutagenesis was performed (Shenoy and Visweswariah, 2003) using the primers ACS_Ery_V379G, ACS_Ery_V379S, ACS_Ery_V379A and ACS_Ery_L641P, respectively. The gene encoding for the ligase domain of the propionyl-CoA synthase (*Ery*PCS_lig) was synthesized by the JGI and subcloned into the vector pET16b, using NdeI and BamHI restriction sites, resulting in plasmid pJZ26. The genes encoding for the two subunits of PCCs from *C. aurantiacus* (*Ca*PCC1 and *Ca*PCC2) were synthesized by the JGI and subcloned as dicistrons into the vector pT7-7 (Tabor and Richardson, 1985), using NdeI and BamHI restriction sites, resulting in plasmids pJZ95 and pJZ96 (cloning performed by Pascal Pfister). The genes encoding for the two subunits of PCC from *Erythrobacter sp.* NAP1 (*Ery*PCC) were synthesized by the JGI and subcloned into the

vector pCDFDuet-1 (Novagen, Merck), using NcoI/HindIII and NdeI/XhoI restriction sites, resulting in plasmid pJZ151 (cloning performed by Jan Zarzycki). For the expression of PCC from *M. extorquens* AM1, plasmid pTE615 was used (Schwander et al., 2016). The PCC plasmids containing the PCC mutants PCC_LD (pJZ99), PCC_LN (pJZ100), PCC_YE (pJZ101), PCC_YH (pJZ141), PCC_YQ (pJZ102), PCC_GA (pJZ127), PCC_GS (pJZ117), PCC_GT (pJZ118), PCC_DI (pJZ105) and PCC_DK (pJZ103), were obtained by performing single-oligo directed mutagenesis (Shenoy and Visweswariah, 2003) using the primers PCCtoGCC_L140D, PCCtoGCC_L140N, PCCtoGCC_Y143E, PCC_Y143H, PCCtoGCC_Y143Q, PCC_G170A, PCC_G170S, PCC_G170T, PCCtoACC_D407I and PCCtoGCC_D407K, respectively, using pTE615 as a template (cloning performed by Jan Zarzycki). The plasmid for GCC M2 (pJZ133) was obtained by performing single-oligo mutagenesis (Shenoy and Visweswariah, 2003) on pJZ105 using the primer PCC_Y143H (cloning performed by Jan Zarzycki). The plasmids containing GCC_M2_LD (pTE1411), GCC_M2_LN (pTE1413) and GCC M3 (pTE1412) were created by single-oligo mutagenesis (Shenoy and Visweswariah, 2003) using pJZ133 as a template and PCC_L100D, PCC_L100N, PCC_L100S as primers, respectively. The plasmids pTE1422 and pTE1425 were created by performing single-oligo mutagenesis (Shenoy and Visweswariah, 2003) on pTE1412 using the primers PCCMe_D171A and PCCMe_D171V, respectively. The plasmid for GCC M4 was obtained from isolation after screening the random mutagenesis library of pTE1412 and was termed pTE3100. The plasmid for GCC M5 was obtained from isolation after screening the random mutagenesis library of pTE3100 and was termed pTE3101. The gene of MCR from *Erythrobacter* sp. NAP1 was synthesized codon-optimized for *E. coli* by JGI and obtained subcloned into pSEVA 321 (pTE1010). The *birA* gene from *M. extorquens* AM1 was amplified from isolated genomic DNA, via flanking primers including restrictions sites for NdeI (BirA_Me_bb_for) and BamHI (BirA_Me_bb_rev). The PCR product was inserted into the expression vector pET16b via restriction digestion by the respective enzymes and ligation. Single-oligo directed mutagenesis (Shenoy and Visweswariah, 2003) was used to remove an EcoRI site from the construct. The Polyhistidin-tag, flanked with NdeI and NcoI sites, was removed by mutating the NcoI site to NdeI by site directed mutagenesis and restriction digestion with NdeI and religation of the construct (cloning performed by Pascal Pfister). The plasmids pCA24N-adk, pCA24N-eno, pCA24N-fucO, pCA24N-glxK, pCA24N-gph and pCA24N-gpmI were

obtained from the ASKA collection (Kitagawa et al., 2005). The gene *gox0313* was ordered readily cloned in the vector pET16b (BaseClear B.V.), using NdeI and BamHI restriction sites (pTE1453). For the construction of pTE1125, the insert of pBBR1p264-gox1801-ST (Meyer et al., 2015) (which was a kind gift of U. Deppenmeier, Bonn) was PCR amplified using the primers MC60 and MC61 and cloned into a pET16b derivative using NdeI and KpnI restriction sites (cloning performed by Martina Carrillo). To obtain the expression plasmid for *Cd*SucD (pTE1816), the gene was synthesized codon-optimized (Eurofins Genomics Germany GmbH) and cloned into pTE380 (Schwander et al., 2016) using NcoI and SalI restriction sites. Plasmids for the expression of all CETCH-enzymes were as described earlier (Schwander et al., 2016).

The vectors pSEVA228-gccB(M4)-gccA-birA (pTE3607) and pSEVA228-gccB(M5)-gccA-birA (pTE3708) were constructed using Gibson Assembly of PCR fragments generated from pTE3100, pTE3101 and pJZ152 using the primer pairs EG_pccA_fw and EG_pccA_rv, EG_pccB_fw and EG_pccB_rv, EG_birA_fw and EG_birA_rv, respectively (Table 10). The backbone vector pSEVA228 was cut using SacI and XbaI prior to Gibson Assembly using the Gibson Assembly Master Mix (New England Biolabs).

Table 11. List of expression plasmids used in this study.

Abbr.	full name	source organism of insert	mutation(s)	plasmid	GenBank accession	reference
AbfT	4-hydroxybutyrate CoA transferase	*Clostridium aminobutyricum*	n/a	pTE1138	CAB60036.2	this work
BirA	biotin ligase	*Methylorubrum extorquens* AM1	n/a	pJZ152	WP_003598178.1	this work
***Ca*MCR, TCR**	malonyl-CoA/ tartronyl-CoA reductase	*Chloroflexus aurantiacus*	n/a	pTrc–McrCa	WP_012258473.1	(a)
***Ca*PCC1**	propionyl-CoA carboxylase	*Chloroflexus aurantiacus*	n/a	pJZ95	YP_001636422.1 YP_001635958.1	this work
***Ca*PCC2**	methylcrotonyl-CoA carboxylase	*Chloroflexus aurantiacus*	n/a	pJZ96	YP_001634995.1 YP_001635635.1	this work
***Cp*PCT**	propionate CoA transferase	*Clostridium propionicum*	n/a	pET-16b_PCT	CAB77207.1	(b)
Eno	enolase	*Escherichia coli*	n/a	pCA24N-eno	APC52982.1	ASKA (c)
***Ery*_ACS2**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	n/a	pTE1008	WP_007166279.1	this work
***Ery*ACS1**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	n/a	pTE1007	WP_007165025.1	this work
***Ery*ACS1_VA, GCS**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	V379A	pTE1417	n/a	this work
***Ery*ACS1_VA_LP**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	V379A,L641P	pTE1427	n/a	this work
***Ery*ACS1_VG**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	V379G	pTE1434	n/a	this work
***Ery*ACS1_VS**	acyl-CoA synthetase	*Erythrobacter sp.* NAP1	V379S	pTE1418	n/a	this work
***Ery*MCR**	malonyl-CoA/ tartronyl-CoA reductase	*Erythrobacter sp.* NAP1	n/a	pTE1010	WP_007163680.1	this work

Abbr.	full name	source organism of insert	mutation(s)	plasmid	GenBank accession	reference
***Ery*PCC**	propionyl-CoA carboxylase	*Erythrobacter sp.* NAP1	n/a	pJZ151	WP_007163953.1 WP_007163945.1	this work
***Ery*PCS_lig**	ligase domain of propionyl-CoA synthase	*Erythrobacter sp.* NAP1	n/a	pJZ26	WP_007163681.1	this work
FucO	lactaldehyde reductase	*Escherichia coli*	n/a	pCA24N-fucO	WP_000013588.1	ASKA (c)
GCC M2	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	Y143H, D407I	pJZ133	n/a	this work
GCC M3	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100S, Y143H, D407I	pTE1412	n/a	this work
GCC M4	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100S, Y143H, D407I, I450V	pTE3100	n/a	this work
GCC M5	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100S, Y143H, D407I, I450V, W502R	pTE3101	n/a	this work
GCC_M2_LD	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100D, Y143H, D407I	pTE1411	n/a	this work
GCC_M2_LN	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100N, Y143H, D407I	pTE1413	n/a	this work
GCC_M3_DA	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100S, Y143H, D171A, D407I	pTE1422	n/a	this work
GCC_M3_DV	glycolyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L100S, Y143H, D171V, D407I	pTE1425	n/a	this work
GlxK	glycerate kinase	*Escherichia coli*	n/a	pCA24N-glxK	WP_062863124.1	ASKA (c)
Gox0313	alcohol dehydrogenase	*Gluconobacter oxydans*	n/a	pTE1453	AAW60096.1	this work
Gox1801	succinic semialdehyde reductase	*Gluconobacter oxydans*	n/a	pTE1125	WP_011253321.1	this work
Gpm	phosphoglycerate mutase	*Escherichia coli*	n/a	pCA24N-gpmI	APC53800.1	ASKA (c)
MatB	malonyl-CoA synthetase	*Rhizobium leguminosarum*	n/a	pEMatB-nHis	AAC83455.1	(d)
Myo	myokinase (adenylate kinase)	*Escherichia coli*	n/a	pCA24N-adk	WP_127790717.1	ASKA (c)
Nox	NADH Oxidase	*Lactobacillus pentosus*	n/a	pET28a-NH-nox-L.p.	WP_050338260.1	(e)
PCC	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	n/a	pTE615	WP_003599287.1 WP_003597263.1	(f)
PCC_DI	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	D407I	pJZ105	n/a	this work
PCC_DK	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	D407K	pJZ103	n/a	this work
PCC_GA	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	G170A	pJZ127	n/a	this work
PCC_GS	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	G170S	pJZ117	n/a	this work
PCC_GT	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	G170T	pJZ118	n/a	this work
PCC_LD	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L140D	pJZ99	n/a	this work
PCC_LN	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	L140N	pJZ100	n/a	this work
PCC_YE	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	Y143E	pJZ101	n/a	this work
PCC_YH	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	Y143H	pJZ141	n/a	this work
PCC_YQ	propionyl-CoA carboxylase	*Methylorubrum extorquens* AM1	Y143Q	pJZ102	n/a	this work
PduP	propionaldehyde dehydrogenase	*Rhodo-pseudomonas palustris* BisB18	n/a	pJZ73	WP_041801450.1	(g)
Pgp	phosphoglycolate phosphatase	*Escherichia coli*	n/a	pCA24N-gph	APC54031.1	ASKA (c)
PPK2-I	polyphosphate kinase	*Sinorhizobium meliloti*	n/a	pET28a-SmPPK2	n/a	(h)
PPK2-II	polyphosphate kinase	*Acinetobacter johnsonii*	n/a	pET28a-AjPPK2	n/a	(h)
***Re*PCT**	propionate CoA transferase	*Ralstonia eutropha*	n/a	pET-19b::pct	CAJ93797.1	(i)
***Cd*SucD**	succinyl-CoA reductase	*Clostridioides difficile*	n/a	pTE1816	WP_004454646	this work

(a) Kroeger et al. (2011), (b) Selmer et al. (2002) (c) Kitagawa et al. (2005), (d) Vagstad et al. (2012), (e) Nowak et al. (2015), (f) Schwander et al. (2016), (g) Zarzycki et al. (2017), (h) Mordhorst et al. (2017), (i) Lindenkamp et al. (2013)

4.4 Library generation of GCC variants

4.4.1 Saturation mutagenesis libraries

For saturation mutagenesis of GCC, the amino acid positions 140 and 143 of the β-subunit were randomized by saturation mutagenesis applying a method for reduced codon redundancy (Kille et al., 2013) on four different templates, spanning different combinations of GCC M3 residues (L100_D407, L100_I407, S100_D407, S100_I407). For the whole plasmid PCR, forward primers (Table 10) were mixed according to Table 12 to achieve an equal coverage of all possible triplets in the saturated positions. The PCR products were purified and digested with DpnI prior to transformation into *E. coli* ElectroMAX DH5α. To ensure a 95 % library coverage of the variant space, the minimal number of transformants for each library was 1450 (Kille et al., 2013). Randomization of the desired triplets was verified by sequencing a plasmid preparation of the pooled libraries.

Table 12. Mixing ratio of primers for saturation mutagenesis.

Primer	Mixing ratio
PCC_satmut_22c_1	144
PCC_satmut_22c_2	81
PCC_satmut_22c_3	108
PCC_satmut_22c_4	108
PCC_satmut_22c_5	12
PCC_satmut_22c_6	12
PCC_satmut_22c_7	9
PCC_satmut_22c_8	9
PCC_satmut_22c_9	1
total:	484

4.4.2 Random mutagenesis libraries

Plasmid libraries of randomly mutagenized GCC were created by megaprimer-based whole-plasmid PCR (Miyazaki, 2011). To generate randomized fragments of the β-subunit of GCC M3 (pTE1412) for the first round of mutagenesis, error-prone PCR (Cirino et al., 2003) was performed using 2.5 U Taq-polymerase with Mg-free buffer (New England Biolabs M0320S), 7 mM $MgCl_2$, 0.4 mM dGTP and dATP each, 2 mM dCTP and dTTP each, 0.4 µM primer PccB_fw_P1, 0.4 µM primer PccB_rev_P2 (Table

10), 10 % (v/v) DMSO, 50 ng template DNA of pTE1412 and 200 - 500 µM $MnCl_2$ in a 50 µL reaction. The randomized fragments were digested with DpnI, purified by agarose gelelectrophoresis and used as mega primers for a whole-plasmid PCR (megaWHOP) as described elsewhere (Miyazaki, 2011) or subjected to another error-prone PCR reaction to further increase the mutation rate. The megaWHOP reaction (50 µL) contained 1 x KOD Hot Start reaction buffer (Novagen), 0.2 mM dNTPs, 1.5 mM $MgSO_4$, 500 ng megaprimer, 50 ng template plasmid (GCC M3, pTE1412) and 2.5 U KOD Hot Start DNA polymerase (Novagen). The megaWHOP product was purified, digested with DpnI, again purified, and transformed into *E. coli* ElectroMAX DH5α (Thermo Fisher Scientific) to ensure a high number of transformants in the resulting libraries.

To estimate the mutation rate for different concentrations of $MnCl_2$ used in the error-prone PCR, the plasmids of six randomly picked clones after megaWHOP were purified, sequenced and analyzed for nucleotide exchanges. Library 1_1 contained 3.6 mutations/kpb (2-times 500 µM $MnCl_2$ in subsequent error-prone PCR), library 1_2 contained 1.1 mutations/kbp (500 µM $MnCl_2$) and library 1_3 contained 0.2 mutations/kbp (200 µM $MnCl_2$).

For the second round of random mutagenesis, plasmid pTE3100 (GCC M4) was used as a template for error-prone PCR as well as for megaWHOP following the procedure described above. Here, it was aimed at creating libraries with mutation rates of 0.2 (library 2_1) and 1.1 per kbp (library 2_2) and 200 µM $MnCl_2$ and 500 µM $MnCl_2$ were used, respectively.

4.5 Bacterial strains

For cloning purposes, *E. coli* DH5α (Thermo Fisher Scientific) was used. For transformation of the megaWHOP products, *E. coli* ElectroMAX DH5α (Thermo Fisher Scientific) was used. For expression of recombinant proteins *E. coli* BL21 (DE3), *E. coli* BL21-AI (Thermo Fisher Scientific) or *E. coli* Rosetta (DE3) pLysS (Novagen Merck) were used. For the expression of biotin-dependent carboxylases, *E. coli* BL21 (DE3) containing the pJZ152 plasmid (encoding for *M. extorquens* biotin ligase) was used (BL21_BirA). For the expression of 4-hydroxybutyryl-CoA synthetase (Hbs), BL21_GroES/GroEL was used (Schwander et al., 2016). For the heterologous expression of acyl-CoA synthetases, an *E. coli* BL21 (DE3) AI Δ*patZ* strain was constructed by using

the Quick & Easy *E. coli* Gene Deletion Kit (Gene Bridges GmbH) following the manufacturer's instructions. In brief, in a first step a PCR product from the functional cassette including a kanamycin resistance gene and flanking homology arms was generated using the primers patZ_KO_fw and patZ_KO_rv (Table 10). In the next step, the pRedET plasmid containing the recombination genes and a temperature-sensitive origin of replication was transformed into BL21 AI (DE3). To start the recombination process, the expression of the Red/ET recombination genes was induced and the linear PCR product transformed into the cells. To confirm the integration of the Kan-cassette and knock-out of *patZ* in the genome of BL21-AI, single colonies were picked and analyzed by colony PCR using the primers patZ_KO_P1 and patZ_KO_P4 (Table 10).

P. putida KT2440 knockout strains (*P. putida* KT2440 Δ*gclR*, *P. putida* KT2440 Δ*gcl-hyi-glxR* and *P. putida* KT2440 Δ*gclR* Δ*gcl-hyi-glxR*) were generated using a scarless deletion method (Martínez-García and de Lorenzo, 2011). The method is based on homologous recombination induced by the appearance of double-strand breaks introduced by the endonuclease I-SceI. In the case of Δ*gclR*, the complete coding sequence from start to stop codon (including both) was removed. In the case of Δ*gcl-hyi-glxR*, the first 99 DNA bases of gcl and the last 99 DNA bases of glxR were left in the knockout strain (leading to the formation of a nonsense peptide) in case they contain regulatory elements. Target sequences upstream (TS1) and downstream (TS2) the regions to be deleted were identified and PCR was performed on a fresh *P. putida* colony using the primers gcl-hyi-glxR_TS1fw and gcl-hyi-glxR_TS1rv for TS1_gcl-hyi-glxR, gcl-hyi-glxR_TS2fw and gcl-hyi-glxR_TS2rv for TS2_ gcl-hyi-glxR, gclR_TS1fw and gclR_TS1rv for TS1_gclR and gclR_TS2fw and gclR_TS2rv for TS2_gclR (Table 10). Subsequently, the corresponding target sequences (TS1 and TS2) were combined using SOEing PCR and cloned into pEMG (Martínez-García and de Lorenzo, 2011) using KpnI and XbaI restriction sites and transformation into *E. coli* DH5α*λpir*. The resulting plasmid was transformed into *P. putida* KT2440 by electroporation (Martínez-García and de Lorenzo, 2012) for cointegration of the TS1-TS2 construct. Successful cointegration was verified by colony PCR using TS1fw and TS2rv primers and electrocompetent cells were prepared of cointegrated strains for transformation of pSW-I, containing the gene for I-SceI (Martínez-García and de Lorenzo, 2011). I-SceI gene expression was induced with 15 mM 3-methylbenzoate for 16 h. The strains were subcultured and after loss of the pSW-I plasmid, the correct gene deletions were verified by colony PCR using primers

binding internally (gclR_check_fw and gclR_check_rv; hyi_check_fw and hyi_check_rv) or externally (gclR_check_out_fw and gclR_check_out_rv; hyi_check_out_fw and hyi_check_out_rv) of the target sequence (Table 10).

4.6 Heterologous expression and purification of recombinant proteins

For the overexpression of recombinant proteins, the corresponding plasmids (Table 11) were transformed into chemically competent *E. coli* cells. See Table 13 for details on the expression strain. The cells were grown on LB agar plates containing the selective antibiotics (Table 13) at 37 °C overnight. Typically, 1 L selective medium was inoculated with the colonies obtained from the overnight culture and grown at 37 °C to an OD_{600} of 0.4 to 1.0, induced with IPTG (and arabinose in the case of ACSs) and grown over night at 21 °C or 25 °C (Table 13). Cells were harvested at 6000 × *g* for 12 min at 4 °C and cell pellets were stored at -80 °C until purification of enzymes. Hbs and 4-hydroxybutyryl-CoA dehydratase (Hbd) were expressed as described earlier (Schwander et al., 2016). Proteins were purified as described in the following (Table 13).

Table 13. Conditions for heterologous expression in *E. coli* and purification of enzymes.

Enzyme abbr.	full name	tag	strain[1]	medium[2]	resistance[3]	induction	expres.	pur.
AKR7a2	succinic semialdehyde reductase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
AbfT	4-hydroxybutyrate CoA transferase	His	Rosetta	GLB	Amp	0.5 mM IPTG	25 °C O/N	b
***Ca*MCR, TCR**	malonyl-CoA/ tartronyl-CoA reductase	Strep	BL21	GLB	Amp	0.5 mM IPTG	25 °C 4h	c
***Ca*PCC1**	propionyl-CoA carboxylase	His	BL21 DE3	GLB	Amp	0.4 mM IPTG	37°C 4h	f
***Ca*PCC2**	propionyl-CoA carboxylase	His	BL21 DE3	GLB	Amp	0.4mM IPTG	37°C 4h	f
Ccr	Crotonyl-CoA carboxylase/ reductase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
***Cp*PCT**	propionate CoA transferase	His	BL21	TB	Amp	0.5 mM IPTG	25 °C O/N	a
Ecm	ethylmalonyl-CoA mutase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
Epi	ethylmalonyl-CoA/ methylmalonyl-CoA epimerase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
***Ery*ACS, GCS**	acyl-CoA synthetase	His	AI	TB	Amp	0.5 mM IPTG	25 °C 4h	a
***Ery*ACS, GCS**	acyl-CoA synthetase	His	AI Δ*patZ*	TB	Amp + Kan	0.5 mM IPTG + 0.02% ara	25 °C 4h	a
***Ery*MCR**	malonyl-CoA/ tartronyl-CoA reductase	Strep	BL21	TB	Amp	0.5 mM IPTG	25 °C O/N	c
***Ery*PCC**	propionyl-CoA carboxylase	His	BL21 DE3	GLB	Spec	not induced	25°C O/N	a
Fdh	formate dehydrogenase (D221A)	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
FucO	lactaldehyde reductase	His	BL21	TB	Cam	0.5 mM IPTG	25 °C O/N	a

Enyzme abbr.	full name	tag	strain[1]	medium[2]	resistance[3]	induction	expres.	pur.
GCC	glycolyl-CoA carboxylase	His	BL21_BirA	GLB[4]	Spec + Amp	0.5 mM IPTG	25 °C O/N	a
GlxK	glycerate kinase	His	BL21	TB	Cam	0.5 mM IPTG	25 °C O/N	a
Gox0313	alcohol dehydrogenase	His	BL21	TB	Amp	0.5 mM IPTG	25 °C O/N	a
Gox1801	succinic semialdehyde/ glyoxylate reductase	Strep	Rosetta	GLB	Amp	0.5 mM IPTG	25 °C O/N	d
Gpm	phosphoglycerate mutase	His	BL21	TB	Cam	0.5 mM IPTG	25 °C O/N	a
KatE	katalase	His	BL21	TB	Cam	0.25 mM IPTG	21°C O/N	e
Mch	mesaconyl-CoA hydratase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
Mcl1	β-methylmalyl-CoA lyase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
Mcm	methylmalonyl-CoA mutase	His	BL21	TB	Amp	0.25 mM IPTG	21°C O/N	e
Mco	methylsuccinyl-CoA oxidase	His	Rosetta	TB	Amp	0.25 mM IPTG	21°C O/N	e
MePCC	propionyl-CoA carboxylase	His	BL21_BirA	GLB[4]	Spec + Amp	0.5 mM IPTG	25 °C O/N	a
Myo	myokinase (adenylate kinase)	His	BL21	TB	Cam	0.5 mM IPTG	25 °C O/N	a
Hbs	4-hydroxybutyryl-CoA synthetase	His	BL21 GroES/GroEL	TB	Amp	0.25 mM IPTG	21°C O/N	e
Hbd	4-hydroxybutyryl-CoA dehydrogenase	His	BL21	TB[5]	Amp	0.25 mM IPTG	21°C O/N	e
Nox	NADH Oxidase	His	BL21	GLB	Kan	0.5 mM IPTG	25 °C O/N	a
Pco	propionyl-CoA oxidase	His	Rosetta	TB	Amp	0.25 mM IPTG	25°C 4h	e
PduP	propionaldehyde dehydrogenase	Strep	BL21	TB	Amp	not induced	25 °C O/N	c
Pgp	phosphoglycolate phosphatase	His	BL21	TB	Cam	0.5 mM IPTG	25 °C O/N	a
Ppk2-I	polyphosphate kinase	His	BL21	TB	Kan	0.25 mM IPTG	21°C O/N	e
Ppk2-II	polyphosphate kinase	His	BL21	TB	Kan	0.25 mM IPTG	21°C O/N	e
***Re*PCT**	propionate CoA transferase	His	BL21	TB	Amp	0.5 mM IPTG	25 °C O/N	a
***Cd*SucD**	succinyl-CoA reductase	His	Rosetta	GLB	Spec	0.5 mM IPTG	25 °C O/N	b, e
***Ck*SucD**	succinyl-CoA reductase	His	BL21	TB	Spec	0.25 mM IPTG	21 °C O/N	e
***Ery*PCS_lig**	ligase domain of propionyl-CoA synthase	His	BL21	TB	Amp	0.5 mM IPTG	25 °C O/N	a

expres.: expression, pur.: purification

[1] see text "Bacterial Strains" in Materials and Methods for details

[2] TB: terrific broth (12 g L^{-1} tryptone, 24 g L^{-1} yeast extract, 0.4 % glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4), GLB: golden LB (LB medium with 0.4 % glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4)

[3] concentrations of antibiotics in medium: ampicillin (Amp) 100 µg mL^{-1}, kanamycin (Kan) 50 µg mL^{-1}, chloramphenicol (Cam) 34 µg mL^{-1}, spectinomycin (Spec) 50 µg mL^{-1}

[4] PCC/GCC medium contained 2 µg mL^{-1} biotin

[5] Hbd medium additives: 100 µM Fe(II)SO_4, 100 µM Fe(III)citrate, 20 mM fumarate

Purification protocol a: Cell pellets were resuspended in two-fold volume of buffer A_{His} (50 mM HEPES, 500 mM NaCl, pH 7.8) containing 0.1 mg mL^{-1} DNase I. The cell suspension was passed through a French pressure cell twice at a pressure of 137 MPa and centrifuged at 100,000 × *g* and 4 °C for 1 h. The supernatant was filtered through a 0.45 µm syringe filter and loaded at a flow rate of 1 mL min^{-1} onto a 1 mL HisTrap™ FF column (GE Healthcare) which had previously been equilibrated with 5 column volumes

(cv) of buffer A_{His}. The column was washed with 20 cv of 90 % buffer A_{His} and 10 % buffer B_{His} (50 mM Tris/HCl, 500 mM NaCl, 500 mM imidazole, pH 7.8) and the protein was eluted with buffer B_{His}. The fraction containing the eluted enzyme was desalted using 5 mL HiTrap™ Desalting columns (GE Healthcare) and buffer A_{Strep} (50 mM HEPES, 150 mM NaCl, pH 7.8) .

Purification protocol b: Cell pellets were resuspended in two-fold volume lysis buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 % glycerol). Resuspended cells were stored at -20 °C until lysis. Cells were lysed by ultrasonication and the lysate was cleared by ultracentrifugation at 50,000 × *g* for 45 min at 4 °C and subsequently filtered through a 0.45 µm filter. The cleared lysate was loaded onto a 1 mL HisTrap™ FF column (GE Healthcare) and unbound protein was removed with 20 mL of 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 75 mM Imidazole. The protein was eluted in 20 mM Tris-HCl pH 7.5, 500 mM NaCl and 500 mM Imidazole. The protein from elution fractions was desalted with a HiTrap™ Desalting column (GE Healthcare) column in 20 mM Tris-HCl pH 7.5, 200 mM NaCl (purification performed by Martina Carrillo).

Purification protocol c: Cell pellets were resuspended in two-fold volume of buffer A_{Strep} (50 mM HEPES, 150 mM NaCl, pH 7.8) containing 0.1 mg mL^{-1} DNase I. The cell suspension was passed through a French pressure cell twice at a pressure of 137 MPa and centrifuged at 100,000 × *g* and 4 °C for 1 h. The supernatant was filtered through a 0.45 µm syringe filter and loaded at a flow rate of 1 mL min^{-1} onto a 1 mL StrepTrap™ HP column (GE Healthcare) which had previously been equilibrated with 5 cv of buffer A_{Strep}. The column was washed with 20 cv of buffer A_{Strep} and the protein was eluted with buffer A_{Strep} containing 3 mM desthiobiotin.

Purification protocol d: Cell pellets were resuspended in two-fold volume lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 % glycerol). Resuspended cells were stored at -20 °C until lysis. Cells were lysed by ultrasonication and the lysate was cleared by ultracentrifugation at 50,000 × *g* for 45 min at 4 °C and subsequently filtered through a 0.45 µm filter. The cleared lysate was loaded onto a 1 mL StrepTrap HP (GE Healthcare) and unbound protein was removed with 20 mL of 50 mM Tris-HCl pH 7.5, 150 mM NaCl. The protein was eluted in 50 mM Tris-HCl pH 7.5, 150 mM NaCl and 2.5 mM of desthiobiotin (purification performed by Martina Carrillo).

Purification protocol e: Cells were resuspended in two-fold volume of lysis buffer (500 mM NaCl, 50 mM HEPES, 10 % glycerol, pH 7.8) containing 5 mM $MgCl_2$, 10 µg ml^{-1} DNase and one tablet of Sigma*FAST* Protease Inhibitor Cocktail (Sigma-Aldrich). The cells were lysed by microfluidization (two times at 16,000 psi). The cell lysate was centrifuged at 50,000 × *g* for 1 h at 4 °C. The supernatant was filtered through a 0.45 µm filter. The lysate was mixed with 3 ml Protino Ni-NTA agarose beads (Macherey-Nagel) and incubated on ice for 30 - 45 min (70 rpm). Afterwards the beads were collected in a gravity column and washed with three cv of lysis buffer. For the removal of unspecific bound proteins, the beads were washed with three cv of lysis buffer containing 50 mM imidazole and three cv with lysis buffer containing 75 mM imidazole. The enzymes were eluted in two cv of lysis buffer containing 500 mM imidazole. The collected recombinant protein was concentrated in Amicon Ultra 15 mL Centrifugal Filters (Merck). For desalting, the protein solution was loaded on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare). The desalting/storage buffer contained 200 mM NaCl, 50 mM HEPES and 10 % glycerol. For Hbs and Hbd, the buffer contained 500 mM NaCl (purification performed by Christoph Diehl).

Purification protocol f: Cell pellets were resuspended in two-fold volume of lysis buffer (200 mM MOPS, 100 mM NaCl, pH 7.8) containing 0.1 mg mL^{-1} DNase I. The cell suspension was passed through a French pressure cell at 137 MPa. The lysate was heat precipitated at 65 °C for 10 min and then cooled on ice and centrifuged at 100,000 × *g* and 4 °C for 1 h. The supernatant was filtered through a 0.45 µm syringe filter and used for enzyme assays (purification performed by Pascal Pfister).

All enzymes were concentrated with Amicon Ultra centrifugal filters (Merck Millipore) with appropriate size cut-offs and concentration was determined using a Nanodrop 2000 spectrophotometer (Thermo Scientific) and the enzyme specific extinction coefficients for calulcation. Enzymes were stored in 20 - 50 % glycerol in the corresponding purification buffer at -20 °C or -80 °C. The storage buffer for propionyl-CoA oxidase (Pco) and methylsuccinyl-CoA oxidase (Mco) additionally contained FAD (1 µM per µM of enzyme). The storage buffer of methylmalonyl-CoA mutase (Mcm) and ethylmalonyl-CoA mutase (Ecm) additionally contained 2 mM Coenzyme B_{12}. The storage buffer of Hbs additionally contained 2 mM $MgCl_2$.

The purity of enzymes was assessed using SDS-PAGE.

4.7 Avidin gel shift assay

To confirm full biotinylation of biotin-dependent carboxylases, avidin gel shift assays were performed as described previously for streptavidin (Fairhead and Howarth, 2015). In brief, 10 µM of carboxylase was mixed with non-reducing SDS-PAGE buffer (final concentrations 2 % SDS, 10 % glycerol, 60 mM Tris-HCl pH 6.8) and incubated for 5 min at 100 °C. 10 µM of avidin (Santa Cruz Biotechnology) in PBS was added and loaded onto a gradient SDS gel (Mini-PROTEAN® TGX™ 4-20 %, Bio-Rad) after 5 min incubation at RT. The gel was run for 90 min at 100 V in Rotiphorese® SDS-PAGE buffer (Roth) and stained afterwards with GelCode™ Blue Safe Protein Stain (Thermo Scientific).

4.8 Activity assays

4.8.1 CoA-transferases

Specific activites of acyl-CoA:propionate transferases (*Re*PCT and *Cp*PCT) and 4-OH-butyrate CoA transferase (AbfT) were determined via analysis of the time-dependent formation of glycolyl-CoA using UPLC-hrMS. The enzyme assay was performed at 37 °C in a total volume of 60 µL. The reaction mixture contained 200 mM MOPS/KOH pH 7.5, 2 mM acetyl-CoA, 100 µg enzyme and 10 mM - 800 mM glycolate. Aliquots were taken at time points 0.5, 1.0, 1.5 and 2.0 minutes and the reaction was immediately stopped by HCl (1 % total concentration). The samples were centrifuged at 17,000 × *g* at 4 °C and the supernatant used for UPLC-hrMS analysis.

4.8.2 Acyl-CoA synthetases

To determine specific activities of acyl-CoA synthetases, a coupled spectrophotometric assay to measure AMP formation was used. The ATP hydrolysis reaction was measured coupled to ADP regeneration by myokinase and ATP regeneration by PK with PEP and subsequent reduction of pyruvate to lactate by lactic dehydrogenase LDH. The oxidation of NADH by LDH was followed spectrophotometrically at 340 nm. The assay was performed at 30 °C and typically contained 100 mM MOPS/KOH (pH 7.8), 5 mM $MgCl_2$, 2.5 mM ATP, 15 U mL^{-1} PK, 23 U mL^{-1} LDH, 2.5 mM PEP, 200 µg mL^{-1} myokinase, 1 mM CoA, 0.4 mM NADH, 2 µg mL^{-1} ACS/GCS and 1 - 200 mM glycolate.

4.8.3 Carboxylases

Activity of glycolyl-CoA carboxylase (GCC) was determined spectrophotometrically at 37 °C in three different ways.

The ATP hydrolysis reaction was measured coupled to ATP regeneration by PK with PEP and subsequent reduction of pyruvate to lactate by LDH. The oxidation of NADH by LDH was followed spectrophotometrically at 340 nm. The reaction mixture (200 µL) typically contained 100 mM MOPS/KOH pH 7.8, 5 mM ATP, 50 mM $KHCO_3$, 2.5 mM PEP, 15 U mL^{-1} PK, 23 U mL^{-1} LDH, 0.4 mM NADH, 7.5 mM $MgCl_2$, 2 mM glycolyl-CoA and 40 µg mL^{-1} GCC.

To measure carboxylation activity, spectrophotometric assays at 365 nm were performed coupling the product formation (tartronyl-CoA) to its subsequent reduction via tartronic semialdehyde to glycerate by use of TCR. This coupling reaction was analogous to an activity assay for ACC (Kroeger et al., 2011). The reaction mixture (200 µL) typically contained 100 mM MOPS/KOH pH 7.8, 1.7 mg mL^{-1} TCR, 50 mM $KHCO_3$, 0.6 mM NADPH, 5 mM $MgCl_2$, 100 µg mL^{-1} GCC, 0.8 - 2 mM glycolyl-CoA and 0.8 - 5 mM ATP.

For determination of the ratio of ATP hydrolysis to carboxylation, the TCR-coupled assay was used under ATP-limiting conditions (0.8 mM) and the ratio of ATP-hydrolysis to carboxylation was calculated from the delta absorption values.

For determination of kinetic parameters of GCC variants, the time-dependent formation of tartronyl-CoA was analyzed using UPLC-hrMS. The assays contained 100 mM MOPS/KOH pH 7.8, 50 mM $KHCO_3$, 6 mM $MgCl_2$, 3 mM ATP, 0.03 - 0.3 mg mL^{-1} GCC and 0.1 - 1.6 mM glycolyl-CoA. Aliquot samples were withdrawn, quenched with 1 % HCl, centrifuged at 17,000 × *g* for 20 min at 4 °C and analyzed by UPLC-hrMS.

4.8.4 Acyl-CoA reductases

The kinetics of tartronyl-CoA reduction of *Ca*MCR (TCR) and *Ery*MCR were determined spectrophotometrically at 365 nm following the oxidation of NADPH. Per molecule tartronyl-CoA, two molecules of NADPH are oxidized by MCR, which was taken into account for the calculation of specific activities. The measurements were carried out at

37 °C in 100 mM MOPS/KOH pH 7.5, 5 mM $MgCl_2$, 0.4 mM NADPH, 115 µg mL^{-1} *Ca*MCR or 88 µg mL^{-1} *Ery*MCR and 0.01 - 2 mM tartronyl-CoA.

4.8.5 Dehydrogenases

The specific activities of FucO and Gox0313 with EG were determined spectrophotometrically at 37 °C and 340 nm. The assay contained 100 mM MOPS/KOH pH 7.8, 155 µg mL^{-1} FucO or 35 µg mL^{-1} Gox0313, 2 mM NAD^+ and 10 - 2000 mM EG. The specific activity of PduP with glycol aldehyde was measured spectrophotometrically at 37 °C and 340 nm and the assay contained 100 mM MOPS/KOH pH 7.8, 11 µg mL^{-1} PduP, 1 mM NAD^+, 0.5 mM CoA and 50 mM glycolaldehyde dimer (corresponding to 100 mM glycolaldehyde monomer).

4.8.6 NADH oxidase

The specific activity of Nox with NADH and NADPH was determined spectrophotometrically at 37 °C and 340 nm. The assay contained 100 mM MOPS/KOH pH 7.8, 0.4 mM NAD(P)H and 12 µg Nox.

4.8.7 Phosphoglycolate phosphatase

Phosphoglycolate phosphatase (Pgp) was assayed in 100 mM MOPS/KOH pH 7.8, 5 mM $MgCl_2$ and 2 mM 2-PG with 7 µg mL^{-1} Pgp at 37 °C. Aliquot samples were taken at time points 0 min, 1 min and 10 min and quenched with HCl (1 % end concentration). The samples were analyzed using the Phosphate colorimetric kit (Sigma-Aldrich) following the manufacturer's instructions and the specific activity was calculated from the enzyme- and time-dependent release of P_i.

4.8.8 Glycerate kinase

The activity of glycerate kinase (GlxK) was measured spectrophotometrically at 340 nm and 37 °C. The assay contained 100 mM MOPS/KOH pH 7.8, 1 mM ATP, 5 mM $MgCl_2$, 0.4 mM NADH, 2.5 mM PEP, 0.5 mM glycerate, 15 U mL^{-1} PK, 23 U mL^{-1} LDH and 4 µg mL^{-1} GlxK.

4.8.9 Phosphoglycerate mutase

Activity of phosphoglycerate mutase (Gpm) was measured spectrophotometrically at 340 nm and 37 °C. The assay contained 100 mM MOPS/KOH pH 7.8, 1 mM ADP, 5 mM $MgCl_2$, 0.4 mM NADH, 66 µg mL^{-1} enolase, 15 U mL^{-1} PK, 23 U mL^{-1} LDH, 1 mM 3-PGA and 5 µg mL^{-1} Gpm.

4.8.10 Succinyl-CoA reductase

Succinyl-CoA reductases were assayed spectrophotemetrically following the decrease of NADPH/NADH absorption at 340 nm at 37 °C. The homologs of *C. kluyveri* and *C. difficile* were measured using NADPH, the *P. gingivalis* homolog was measured using NADH. The assay mixture contained 100 mM MOPS/KOH pH 7.8, 0.4 mM NADH or NADPH, 0.5 mM glycolyl-CoA and 0.1 - 0.4 mg mL^{-1} enzyme.

4.9 Michaelis-Menten kinetics

Michaelis-Menten kinetics were typically calculated from 18 independent activity measurements (6 different substrate concentrations in triplicates) using non-linear regression analysis in GraphPad Prism 8.0.

4.10 Microfluidic experiments

4.10.1 Microfluidic device fabrication

Microfluidic chips were prepared from poly-(dimethylsiloxane) (PDMS, Sylgard 184) from SU8-3000 negative photoresist (MicroChem Corp) molds produced using standard soft-lithography procedures (Beneyton et al., 2014). The surfaces of the microfluidic channels were treated with fluoro-silane (Aquapel, Aquapel) before use.

4.10.2 Microfluidic device operation

All the microfluidic workflows (devices and flow rates) for performing short-term or long-term multiplexed kinetics were as previously described (Beneyton et al., 2018). The microfluidic devices were placed and held on an inverted microscope (IX71, Olympus) and connected to either Nemesys syringe pumps (Cetoni) or a pressure driven pump (Fluigent, MFCS-4C) via PTFE tubing (Fisher Scientific; inner diameter (ID) of 0.3 mm;

outer diameter (OD) of 0.76 mm) to control the flow in the microfluidic devices. Droplets were produced in fluorinated oil (Novec7500, 3M) and stabilized against coalescence by a perfluoropolyether-polyethyleneglycol block-copolymer surfactant (PFPE-PEG-PFPE), synthesized as previously described (Beneyton et al., 2016). The optical set up for fluorescence detection was similar to that reported previously (Beneyton et al., 2018). Data acquisition (DAQ) and control were performed by a DAQ card (National Instruments) executing a program written in LabVIEW (National Instruments). The data acquisition rate for the system was 200 kHz.

4.10.3 Microfluidic assays with purified GCC

To establish a microfluidic assay to screen for better GCC variants, the coupled spectrophotometric assay of GCC with TCR to monitor carboxylation activity was transferred into w/o droplets with purified GCC using a short-term multiplexed kinetics workflow (Beneyton et al., 2018). Substrate solutions containing all substrates except for glycolyl-CoA and the purified GCC were encapsulated in 30 pL w/o droplets and a glycolyl-CoA solution was injected into these droplets to start the reaction. The final concentrations in the droplets were 100 mM MOPS/KOH pH 7.8, 1 - 5 mM ATP, 2.5 - 7.5 mM $MgCl_2$, 1 mM NADPH, 50 mM $KHCO_3$, 1 mM glycolyl-CoA, 80 µg mL^{-1} GCC, 400 µg mL^{-1} *Ca*MCR and 30, 60, 90 or 120 µM sulforhodamine B.

4.10.4 Microfluidic assays with single encapsulated *E. coli* cells

For single-cell microfluidic encapsulation, *E. coli* BL21 (DE3) cells were transformed with PCC, GCC M3 or the randomized libraries 1_1, 1_2 and 1_3, and grown on selective agar plates overnight. 25 mL of selective LB medium was inoculated to an OD_{600nm} of approx. 0.1 and the cells were grown to an OD_{600nm} of 0.4 to 0.6. Protein expression was induced by the addition of 0.2 mM IPTG. The cells were grown for 3 h or overnight, subsequently harvested by centrifugation (10 min at 2500 × *g*) and washed three times with 100 mM MOPS/KOH pH 7.8. In the case of experiments with GCC libraries, the library expressing cells were pooled before centrifugation. After determination of OD_{600nm} the cells were diluted to a final in-droplet OD_{600nm} of 0.05 to achieve a droplet occupancy by cells of 5 % to 10 %, assuring that droplets containing cells only contain a single cell (following a Poisson distribution). Sulforhodamine B was added at varying

concentrations (10 µM to 150 µM) to the solutions. A separate substrate and cell lysis solution was prepared containing (final in-droplet concentration) 100 mM MOPS/KOH pH 7.8, 0 - 5 mM ATP, 5 - 10 mM $MgCl_2$, 1 mM NADPH, 50 mM $KHCO_3$, 0.1 x CelLytic B (Sigma Aldrich), 1 mM glycolyl-CoA and 400 µg mL^{-1} *Ca*MCR. The conditions for cell lysis were as described previously (Fischlechner et al., 2014). Cells were encapsulated in 30 pL droplets, incubated at 37 °C and reinjected for analysis over time using a long-term multiplexed kinetics workflow (Beneyton et al., 2018).

4.10.5 Recovery of DNA out of microfluidic droplets

For recovery of plasmid DNA out of microfluidic droplets, different polymerases and conditions were tested. The tested polymerases included Q5 High-Fidelity DNA polymerase (New England Biolabs), KOD Hot Start DNA polymerase (Merck Millipore) and Kapa Hifi HotStart Ready Mix (Roche). The three polymerases were tested with (purified plasmid) template amounts of 10 ng, 1 ng, 100 pg, 10 pg, 1 pg and 0.1 pg. The PCR products were purified using AMPure Beads (Beckman Coulter) and used as template for a nested PCR. The emulsion of sorted droplets from microfluidic experiments was broken by adding 100 µL of 1H,1H,2H,2H-perfluorooctanol and 100 µL of ddH_2O. The aqueous phase was taken after 10 min of incubation and frozen at -20 °C until use as template for a nested PCR using Kapa Hifi HotStart Ready Mix (Roche).

4.11 Screen of GCC libraries in microtiter plates

The libraries 1_1, 1_2 and 1_3 were transformed into *E. coli* BL21_BirA and colonies were picked into 96-deepwell plates (PlateOne) with LB containing 50 µg mL^{-1} spectinomycin and 100 µg mL^{-1} ampicillin. Plates were incubated over night at 37 °C with subsequent transfer into fresh 96-deepwell plates with LB, 50 µg mL^{-1} spectinomycin, 100 µg mL^{-1} ampicillin and 2µg mL^{-1} biotin to an OD_{600} = 0.1. Protein expression was induced with 0.25 mM IPTG at OD_{600} = 0.4 - 0.6 and incubated over night at 25 °C. Cells were lysed using Celllytic B (Sigma-Aldrich) and stored in 20 % glycerol at -80 °C. Enzyme activity was measured in a plate reader by the coupled enzyme assay with purified *Ca*MCR (TCR). Small-volume 384-well plates (Greiner Bio-One) were used and filled with 2 µl cell extract, 100 mM MOPS/KOH pH 7.8, 1 mM ATP, 50 mM $KHCO_3$, 500 µg ml^{-1} *Ca*MCR, 1 mM NADPH, 10 mM $MgCl_2$ and 1 mM glycolyl-CoA

in a reaction volume of 10 µl. Absorbance of NADPH was measured at 340 nm and 37 °C for 10 h with intervals of 47 s in a plate reader (TECAN Infinite M Plex).

4.12 Optimization of ATP regeneration

For the optimization of ATP regeneration for *in vitro* reconstruction of the tartronyl-CoA pathway different systems were tested, including two types of polyphosphate kinase: PPK2-I from *Sinorhizobium meliloti* for regeneration of ATP from ADP and polyphosphate, and PPK2-II from *Acinetobacter johnsonii* for regeneration of ADP from AMP and polyphosphate (Mordhorst et al., 2017), myokinase and CPK. The assays contained 100 mM MOPS/KOH pH 7.8, 5 mM ATP, 10 mM $MgCl_2$, 2 mM CoA, 2 mM NADPH, 50 mM $KHCO_3$, 1 mM glycolate, 150 µg mL^{-1} GCS, 608 µg mL^{-1} GCC and 1120 µg mL^{-1} TCR. Additionally, the assays contained 20 - 500 mM CP, 6.7 U mL^{-1} CPK, 20 mM polyphosphate (PP), 33 µg mL^{-1} polyphosphate kinase 2-I (PPK2-I) in combination with 133 µg mL^{-1} polyphosphate kinase 2-II (PPK2-II) or 67 µg mL^{-1} myokinase (and variations thereof). The assay was started with the addition of ATP. Samples were withdrawn at different time points and immediately quenched with 1 % HCl. The samples were centrifuged at 17,000 × *g* for 20 min at 4 °C and analyzed via UPLC-MS/MS for glycerate using the derivatization method.

4.13 Sequential reconstruction of the tartronyl-CoA pathway

For the sequential reconstruction of the tartronyl-CoA pathway, the enzymes were added one after the other to enable the detection of the intermediates of the reaction. The assay was run at 37 °C and contained 100 mM MOPS/KOH pH 7.8, 5 mM ATP, 10 mM $MgCl_2$, 2 mM CoA, 4 mM NADPH, 50 mM $NaH^{13}CO_3$, 2 mM glycolate, 200 mM CP, 10 U mL^{1} CPK, 50 µg mL^{-1} myokinase. Before starting the reaction, a t_0 sample was taken and immediately quenched with 1 % HCl. All subsequent samples were treated in the same way. To start the reaction, 0.5 mg mL^{-1} GCS was added. After 20 min, the next sample was taken, followed by the addition of 0.6 mg mL^{-1} GCC M4. After 20 min, the next sample was taken, followed by the addition of 0.7 mg mL^{-1} TCR. After 20 min, the final sample was taken. Samples were analyzed for glycolyl-CoA and tartronyl-CoA by UPLC-hrMS and for glycerate using the derivatization method via UPLC-MS/MS.

4.14 Spectrophotometric assay of the tartronyl-CoA pathway

The initial assay contained 100 mM MOPS/KOH pH 7.8, 10 mM $MgCl_2$, 0.4 mM NADPH, 50 mM $KHCO_3$, 0.15 mg mL^{-1} GCS, 0.56 mg mL^{-1} GCC M4 and 1.1 mg mL^{-1} TCR. Then, first 20 mM glycolate were added, then 1 mM CoA and the assay was then started with the addition of 5 mM ATP. The decrease in NADPH absorption at 340 nm was followed spectrophotometrically.

4.15 *In vitro* reconstruction of the tartronyl-CoA pathway from 2-PG

For the reconstruction of the tartronyl-CoA pathway starting from 2-PG, assays were run at 37 °C. The assays initially contained 100 mM MOPS/KOH pH 7.8, 10 mM $MgCl_2$, 2 mM NADPH, 50 mM $NaH^{13}CO_3$, 2 mM CoA, 200 mM CP, 20 mM polyphosphate, 10 mM glucose-6-phosphate, 4.8 U mL^{-1} glucose-6-phosphate dehydrogenase (G6PDH), 0.19 mg mL^{-1} PPKII-2, 4.8 U mL^{-1} CPK, 0.03 mg mL^{-1} Pgp, 0.58 mg mL^{-1} GlxK, 0.46 mg mL^{-1} GCS, 0.69 mg mL^{-1} GCC M4 and 1.07 mg mL^{-1} TCR. The assay was started with the addition of 2.5 mM 2-PG and 5 mM ATP, samples were withdrawn at different time points and immediately quenched with 1 % HCl. The samples were centrifuged at 17,000 × *g* for 20 min at 4 °C and analyzed via UPLC-MS/MS for ^{13}C-labelled 2-PGA.

4.16 *In vitro* reconstruction of the tartronyl-CoA pathway with malate read-out

The assays of the *in vitro* reconstruction of the tartronyl-CoA pathway starting from 2 -PG and 3-PGA were run at 37 °C and initially contained 100 mM MOPS/KOH pH 7.8, 10 mM $MgCl_2$, 2 mM NADPH, 2 mM NADH, 50 mM $NaH^{13}CO_3$, 2 mM CoA, 200 mM CP, 20 mM polyphosphate, 10 mM glucose-6-phosphate, 0.26 mg mL^{-1} enolase, 0.29 mg mL^{-1} Gpm, 0.03 mg mL^{-1} Pgp, 4.8 U mL^{-1} CPK, 4.8 U mL^{-1} malate dehydrogenase, 4.8 U mL^{-1} G6PDH and 0.19 mg mL^{-1} PPKII-2. The assay for the negative control additionally contained 0.58 mg mL^{-1} GlxK. The assay containing the TaCo enzymes additionally contained 0.58 mg mL^{-1} GlxK, 0.46 mg mL^{-1} GCS, 0.69 mg mL^{-1} GCC M4 and 1.07 mg mL^{-1} TCR. The assays for the “no GlxK” control additionally contained 0.46 mg mL^{-1} GCS, 0.69 mg mL^{-1} GCC M4 and 1.07 mg mL^{-1} TCR. All assays were started with the addition of 5 mM ATP, 1 mM 2-PG and 1 mM

3-PGA. Aliquot samples were withdrawn at different time points and immediately quenched with 1 % HCl. The samples were centrifuged at 17,000 × *g* for 20 min at 4 °C, derivatized with 3-NPH and analyzed via UPLC-MS/MS for malate and glycerate.

4.17 Tartronyl-CoA pathway coupled to the CETCH cycle

For the coupling of the tartronyl-CoA pathway to the CETCH cycle, three experiments were performed in the course of optimization (CT_A, CT_B and CT_C).

CT_A: The assay was run as described earlier for CETCH 5.4 (Schwander et al., 2016) with the following changes: The assay mix did not contain β-methylmalyl-CoA lyase (Mcl) and malate synthase (GlcB) and additionally contained 63 µg mL^{-1} Ppk2-II, 100 mU Gox1801, 0.4 mg ml^{-1} GCS, 0.6 mg ml^{-1} GCC M3 and 0.2 mg ml^{-1} TCR. Aliquot samples were withdrawn at different time points, quenched in 4 % formic acid and centrifuged at 17,000 × *g* for 20 min at 4 °C. The glycerate concentration of the samples was measured using derivatization following analysis via UPLC-MS/MS.

CT_B: The assay was run at 30 °C and 450 rpm and contained 200 mM MOPS/KOH pH 7.5, 10 mM $MgCl_2$, 1 mM CoA, 20 mM ATP, 10 mM NADPH, 40 mM formate, 40 mM polyphosphate, 100 mM $NaHCO_3$ and 1.6 µg ml^{-1} carbonic anhydrase. We used double the amounts of enzymes as described previously for CETCH 5.4 (Schwander et al., 2016). Instead of GlcB and Mcl we added 200 mU Gox1801. The reaction was started with the addition of 2 mM propionyl-CoA. After 120 min, the TaCo assay mix (final concentrations: 10 mM ATP, 0.5 mM CoA, 2 mg ml^{-1} GCS, 2 mg ml^{-1} GCC M4 and 1 mg ml^{-1} TCR) was added in a 1:2 ratio and the reactions were transferred to 37 °C. For the negative control, ddH_2O and buffer were added, corresponding to the concentrations in the TaCo mix. Aliquot samples were withdrawn at different time points, quenched in 4 % formic acid and centrifuged at 17,000 × *g* for 20 min at 4 °C. The glycerate concentration of the samples was measured using derivatization following analysis via UPLC-MS/MS.

CT_C: The assay contained 100 mM HEPES pH 7.5, 5 mM $MgCl_2$, 0.5 mM CoA, 2 mM ATP, 5 mM NADPH, 20 mM glucose, 20 mM polyphosphate, 50 mM $NaHCO_3$, 250 µM propionyl-CoA, 0.8 µg ml^{-1} carbonic anhydrase and the CETCH enzymes in the amounts previously described (Schwander et al., 2016) (CETCH 5.4). Instead of GlcB, Mcl and

Fdh, we added 100 mU Gox1801 and glucose dehydrogenase (Sigma-Aldrich). The assay was run for 3 h at 30 °C and then diluted in a 3:1 ratio with the TaCo master mix (final concentration: 100 mM HEPES-HCl pH 7.5, 5 mM ATP, 10 mM $MgCl_2$, 2 mM CoA, 2 mM NADPH, 30 mM bicarbonate, 100 mM CP, 6.6 U mL^{-1} CPK, 22 µg mL^{-1} myokinase, 0.9 mg ml^{-1} GCS, 1.4 mg ml^{-1} GCC M4 and 1.4 mg ml^{-1} TCR). Aliquot samples were withdrawn at different time points, quenched in 4 % formic acid and centrifuged at 17,000 × *g* for 20 min at 4 °C. The glycerate concentration of the samples was measured using derivatization following analysis via UPLC-MS/MS.

4.18 Ethylene glycol conversion by the tartronyl-CoA pathway

The conversion of EG by the tartronyl-CoA pathway was determined measuring glycerate formation. All assays were run at 37 °C for 2 h. Samples were withdrawn at different time points and immediately quenched with 1 % HCl. The samples were centrifuged at 17,000 × *g* for 20 min at 4 °C, derivatized with 3-NPH and analyzed via UPLC-MS/MS for glycerate.

Assays of experiment **EG_A** contained 100 mM MOPS/KOH pH 7.8, 2 mM NAD^+, 2 mM CoA, 50 mM $KHCO_3$, 10 mM ATP, 13.3 mM $MgCl_2$, 5 mM NADPH, 100 mM EG, 1.7 mg mL^{-1} FucO, 0.4 mg mL^{-1} PduP, 0.4 mg mL^{-1} GCC M4 and 0.7 mg mL^{-1} TCR. Due to a side reactivity of FucO with glycerol, the buffer of all enzymes used was exchanged three times using Amicon Ultra-0.5 mL Centrifugal Filters (Merck Millipore).

Assays of experiment **EG_B** contained 100 mM MOPS/KOH pH 7.8, 2 mM NAD^+, 2 mM CoA, 50 mM $KHCO_3$, 10 mM ATP, 20 mM $MgCl_2$, 4 mM NADPH, 100 mM EG, 10 mM CP, 0.5 mg mL^{-1} Gox0313, 0.3 mg mL^{-1} PduP, 0.7 mg mL^{-1} GCC M4, 1.5 mg mL^{-1} TCR, 3.3 U mL^{-1} CPK and 0.04 mg mL^{-1} Nox.

Assays of experiment **EG_C** contained 100 mM MOPS/KOH pH 7.8, 2 mM NAD^+, 2 mM CoA, 50 mM $KHCO_3$, 5 mM ATP, 20 mM $MgCl_2$, 4 mM NADPH, 119 mM EG, 200 mM CP, 0.5 mg mL^{-1} Gox0313, 0.3 mg mL^{-1} PduP, 0.7 mg mL^{-1} GCC M4, 1.5 mg mL^{-1} TCR, 6.6 U mL^{-1} CPK and 0.08 mg mL^{-1} Nox.

4.19 Mass spectrometry

4.19.1 CoA esters

The measurements for CoA esters were done using an Agilent 6550 iFunnel Q-TOF LC-MS system equipped with an electrospray ionization source set to positive ionization mode. Compounds were separated on a RP C18 column (50 mm x 2.1 mm, particle size 1.7 µm, Kinetex EVO C18, Phenomenex) using a mobile phase system comprised of 50 mM ammonium formate pH 8.1 (A) and methanol (B). Chromatographic separation was carried out using the following gradient condition at a flow rate of 250 µL min^{-1}: 0 min 2.5 % B; 1 min 2.5 % B, 6 min 95 % B; 6.25 min 95 % B; 6.35 min 2.5 % B; 7 min 2.5 % B. Capillary voltage was set at 3.5 kV and nitrogen gas was used for nebulizing (20 psig), drying (13 L min^{-1}, 225 °C) and sheath gas (12 L min^{-1}, 400 °C). The TOF was calibrated using an ESI-L Low Concentration Tuning Mix (Agilent) before measurement (residuals less than 2 ppm for five reference ions) and was recalibrated during a run using 922 *m/z* as reference mass. MS data were acquired with a scan range of 500-1200 *m/z*. LC-MS data were analyzed using MassHunter Qualitative Analysis and TOF Quantifative Analysis (Agilent).

4.19.2 Organic acids

Glycerate and malate were derivatized according to a method described earlier (Han et al., 2013). In brief, 10 µL of sample was mixed with 10 µL 150 mM 1-(3-Dimethyl-aminopropyl)-3-ethylcarbodiimide (EDC), 10 µL 250 mM 3-nitrophenyl-hydrazine (3-NPH), and 10 µL of 7.5 % pyridine in methanol in a 1.5 mL reaction tube. The reaction was incubated at 30 °C for 30 min. After incubation, the samples were centrifuged at 13,000 x *g* for 1 min and the supernatant was transferred into HPLC vials. UPLC-MS/MS analyses were performed on an Agilent 6495B Triple Quad LC/MS system equipped with an electrospray ionization source. The analytes were separated on a RP-18 column (50 mm x 2.1 mm, particle size 1.8 µm, Kinetex EVO C18, Phenomenex) kept at 40 °C using a mobile phase system comprised of 0.1 % formic acid in water (A) and acetonitrile (B). The gradient was as follows: 0 min 5 % B; 1 min 5 % B, 6 min 95 % B; 6.5 min 95 % B; 7 min 5 % B at a flow rate of 250 µL min^{-1}. Samples were held at 15 °C and injection volume was 5 µL. MS/MS data were acquired in negative multiple reaction monitoring (MRM) mode. Capillary voltage was set at 3 kV and nitrogen gas was used as nebulizing

(25 psig), drying (11 L min^{-1}, 130 °C) and sheath gas (12 L min^{-1}, 400 °C). The dwell time and fragmentor voltage were 20 ms and 380 V, respectively. Optimized collision energy used for the 3-NPH derivatized-glycerate (240 *m/z* → 137 *m/z*) and malate (403 *m/z* → 137 *m/z*) was 22 V and 40 V respectively. LC-MS data were analyzed and quantified using MassHunter Qualitative Navigator and QQQ Quantitative Analysis softwares (Agilent).

4.19.3 Phosphoglycerate

Phosphoglycerate was measured using UPLC-hrMS. The analysis was performed using an Agilent 6550 iFunnel Q-TOF LC-MS system equipped with an electrospray ionization source set to negative ionization mode. Compounds were separated on a Phenomenex Luna NH_2 (30 mm x 2 mm, particle size 3 µm, 100 Å) and eluted by a binary gradient mobile phase system comprised of 20 mM ammonium acetate pH 9.3/acetonitrile 95:5 (solvent A) and acetonitrile (solvent B). Chromatographic separation was carried out using the following gradient condition at a flow rate of 250 µL min^{-1}: 0 min 85 % B; 3.5 min 0 % B, 7 min 0 % B; 7.5 min 85 % B; 8 min 85 % B. The column oven and autosampler were maintained at 10 °C. Injection volume was set to 1 µL. The ESI-MS was operated with the following parameters: Capillary voltage: 3.5 kV, Nebulizer: 20 psig, drying gas (N_2): 13 L min^{-1} at 225 °C, sheath gas (N_2): 12 L min^{-1} at 400 °C). The QTOF mass detector was calibrated prior to measurement using an ESI-L Low Concentration Tuning Mix (Agilent) with residuals and corrected residuals less than 2 ppm and 1 ppm respectively. MS data were acquired with a scan range of 50-600 *m/z* and peaks were analyzed and integrated using the Agilent MassHunter Qualitative Analysis software.

4.20 Crystallography and cryo-EM

4.20.1 Crystallization and structure determination of *Ca*MCR

The structure of the C-terminal domain of *Ca*MCR was solved using the single-wavelength anomalous scattering (SAD) method. The sitting-drop vapor-diffusion method was used for crystallisation at 4 °C. *Ca*MCR was purified using Strep-Tag affinity chromatography and subsequent gelfiltration on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) in buffer containing 20 mM MOPS and 50 mM KCl (pH 7.8).

The solution containing 7 mg mL^{-1} purified *Ca*MCR and 5 mM $NADP^+$ was mixed in a 1:1 ratio with a solution containing 200 mM MgCl2, 100 mM Tris (pH 7.0) and 10 % PEG 8000. For phasing, these crystals were briefly soaked in mother liquor supplemented with 40 % glycerol for cryoprotection and 5 mM Tb-Xo4 (Engilberge et al., 2017) before freezing in liquid nitrogen. X-ray diffraction data were collected at beamline P14 of DESY (Hamburg, Germany) at the Tb LIII absorption edge (wavelength of 1.649 Å).

The data was processed with the XDS (Kabsch, 2010) and CCP4 v7.0 (Winn et al., 2011) software packages. A 2.9 Å resolution data set with a strong enough anomalous signal was obtained for SAD processing with CRANK2 of the CCP4 software. The resulting initial model was then used as template for molecular replacement with a dataset of a native crystal. Native crystals were obtained from purified *Ca*MCR from a solution containing 7 mg mL^{-1} purified *Ca*MCR and 5 mM $NADP^+$ that was mixed in a 1:1 ratio with a solution containing 200 mM K_2SO_4 and 20 % PEG 3350. The crystals were soaked with 35 % EG for cryoprotection, 20 mM $NADP^+$ and 8 mM tartronyl-CoA before freezing in liquid nitrogen. The native dataset was collected (wavelength of 0.972 Å) at beamline P14 of DESY (Hamburg, Germany). The data was processed with the XDS and CCP4 7.0 software packages. The molecular replacement with a 2.34 Å dataset was carried out using Phaser of the PHENIX software package (Liebschner et al., 2019) and refined with Phenix.Refine. Additional modeling, manual refining and ligand fitting was done in COOT 0.8.9 (Emsley et al., 2010). Final positional and B-factor refinements, as well as water-picking, were performed using Phenix.Refine.

4.20.2 Cryogenic electron microscopy (cryo-EM) sample preparation and data collection of PCC/GCC

For cryo-EM sample preparation, 4.0 µl of the purified complex (1 mg mL^{-1}) containing 1 mM β,γ-imidoadenosine 5′-triphosphate, 2 mM $MgCl_2$, and 4 mM glycolyl-CoA were applied to glow discharged Quantifoil 2/1 grids, blotted for 3.5 s with force “4” in a Vitrobot Mark IV (Thermo Fisher) at 100% humidity and 4 °C, and plunge frozen in liquid ethane, cooled by liquid nitrogen.

For *Me*PCC, cryo-EM data was collected on an FEI Glacios transmission electron microscope operated at 200 keV using the SerialEM software package (Mastronarde, 2005). A total of 2,637 movie frames were recorded at a calibrated pixelsize of 1.18 Å

using a magnification of 36.000 x. The total dose was 50.2 $e^-/Å^2$ distributed over 40 frames. Micrographs were recorded in a defocus range of -0.5 to -3.5 µm.

All subsequent imaging processing steps were performed in the cryoSPARC software package. The dose-fractionated movies were gain-normalized, aligned and dose-weighted using the Patchmotion algorithm. Defocus values were estimated using the CtfFind4 implementation, and 168,488 particles were automatically chosen using a reference-free blob picker. Particle sorting and reference-free two-dimensional classification were performed to remove non-particle candidates and damaged particles. *Ab initio* model generation using the stochastic gradient descent algorithm was used to prevent any model bias using a completely data-driven starting model. The particles were further classified using heterogeneous refinement into 3 classes. The best-aligning class was subsequently subjected to 3D refinement yielding 3.67 Å global resolution and a B-factor of -99 $Å^2$. After per-particle defocus and higher-order CTF correction in the *Refinement_New* algorithm the resolution improved to 3.48 Å and a B-factor of -86.8 $Å^2$.

For GCC M5, cryo-EM data was acquired with an FEI Titan Krios transmission electron microscope using SerialEM software (Mastronarde, 2005). Movie frames were recorded at a nominal magnification of 105,000 x (calibrated physical pixel size: 0.8512 Å/px) using a K3 direct electron detector (Gatan) in correlated double sampling (CDS) mode and a GIF quantum energy filter (Gatan) at 20 eV slit width. The total electron dose of ~55 electrons per $Å^2$ was distributed over 35 frames. Micrographs were recorded in a defocus range of -0.5 to -2.7 µm.

Cryo-EM micrographs were processed on the fly using the Focus software package (Biyani et al., 2017) if they passed the selection criteria (iciness < 1.05, drift 0.4 Å < x < 70 Å, defocus 0.5 um < x < 5.5 um, estimated CTF resolution < 5 Å). Micrograph frames were aligned using MotionCor2 (Zheng et al., 2017) and the contrast transfer function (CTF) was determined using GCTF (Li et al., 2010). Using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) 5,398,283 particles were picked template-free on 20,057 acquired micrographs. Candidate particles were extracted with a pixel box size of 340 using RELION 3.1 (Scheres, 2012) and cleaned using reference free 2D classification. A total of 2,863,953 particles were imported into CryoSparc (Punjani et al., 2017) and used for *ab initio* construction of initial models and subjected to multiple rounds of heterogeneous refinement to obtain the best-aligning 2,181,317 particles. Non-uniform

refinement resulted in a reconstruction with an estimated resolution of 2.25 Å with a temperature factor of -83.9 $Å^2$. By using several rounds of per-particle defocus estimated and higher-order CTF-refinements the final refinement yielded a global resolution of 1.96 Å and an improved temperature factor of 71.7 $Å^2$ (Gold standard FSC analysis of two independent half-sets at the 0.143 cutoff).

4.21 Structural modeling and analysis

Homology modeling was performed using SWISS-MODEL (Arnold et al., 2006). As template for homology modeling of *Ery*ACS1 a structure of an acetyl-CoA synthetase mutant of *S. enterica* (Reger et al., 2007) (PDB 2P2B, 62 % amino acid identity) was used, which had an adenosine-5'-monophosphate-propyl ester (propyl-AMP) bound, an analog of the reaction intermediate propionyl-AMP. Based on the position of the propyl-AMP in the *S. enterica* ACS, a glycolyl-AMP was modeled into the active site of the *Ery*ACS1 homology model.

For initial homology modeling of the β-subunit of *Me*PCC the structure of the β-subunit of PCC from *R. pomeroyi* (PDB 3N6R, 72 % amino acid identity) (Huang et al., 2010) was used.

Initial cryo-EM map fitting was performed in UCSF-Chimera 1.14 (Pettersen et al., 2004) using homology models based on PDB 3N6R (Huang et al., 2010). Automatic refinement of the structure was done using phenix.real_space_refine of the PHENIX 1.17.1-3660 suite (Liebschner et al., 2019). Manual refinements as well as water picking were performed with COOT 0.89 (Emsley et al., 2010).

All structural depictions were created using PyMol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). The modeling of propionyl-CoA and glycolyl-CoA into the active sites of the *Me*PCC and GCC, respectively, was based on the position of methylmalonyl-CoA in the structure of a methylmalonyl-CoA carboxytransferase from *Propionibacterium freudenreichii* (Hall et al., 2003) (PDB 1ON3, 52 % amino acid identity). Manual fitting and adjustments of the CoA thioesters reflecting differences in active site architectures were done with COOT and Pymol.

4.22 *P. putida* growth experiments

For growth experiments of *P. putida* KT2440 wt and knockout strains, 20 mL over night cultures in LB were grown at 30 °C. The cells were harvested at 4000 × *g* for 20 min and washed three times in M9 medium without carbon source (42.3 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl, 9.4 mM NH_4Cl, 170 µM EDTA, 31 µM $FeCl_3$, 6 µM $ZnCl_2$, 0.8 µM $CuCl_2$, 9.4 µM $CoCl_2$, 1.6 µM H_3BO_3, 68 nM $MnCl_2$, 1 mM $MgSO_4$, 300 µM $CaCl_2$, 4 µM biotin, 4 µM thiamine). The starting OD_{600nm} for the growth experiment was adjusted to 0.1. The experiments were performed in two biological replicates with three technical replicates each in microtiter plates in a total volume of 180 µL of M9 medium containing 50 µg mL^{-1} kanamycin, 0.5 mM 3-methylbenzoate and 20 mM of glucose or EG. The absorption at 600 nm was recorded over 40 h using an Infinite 200 PRO plate reader (Tecan) at 30 °C and constant shaking (orbital) at 1 mm amplitude.

References

1. **Agresti, J.J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J.-C., Marquez, M., Klibanov, A.M., Griffiths, A.D., and Weitz, D.A.** (2010). Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. Proc Natl Acad Sci USA *107*, 4004-4009.

2. **Alban, C., Job, D., and Douce, R.** (2000). Biotin metabolism in plants. Annu Rev Plant Biol *51*, 17-47.

3. **Alber, B., Olinger, M., Rieder, A., Kockelkorn, D., Jobst, B., Hügler, M., and Fuchs, G.** (2006). Malonyl-coenzyme A reductase in the modified 3-hydroxypropionate cycle for autotrophic carbon fixation in archaeal *Metallosphaera* and *Sulfolobus* spp. J Bacteriol *188*, 8551-8559.

4. **Amthor, J.S.** (2000). The McCree–de Wit–Penning de Vries–Thornley respiration paradigms: 30 years later. Ann Bot *86*, 1-20.

5. **Anderson, L.E.** (1971). Chloroplast and cytoplasmic enzymes II. Pea leaf triose phosphate isomerases. Biochim Biophys Acta Enzymol *235*, 237-244.

6. **Arnold, K., Bordoli, L., Kopp, J., and Schwede, T.** (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195-201.

7. **Aslan, S., Noor, E., Vaquerizo, S.B., Lindner, S.N., and Bar-Even, A.** (2020). Design and engineering of *E. coli* metabolic sensor strains with a wide sensitivity range for glycerate. Metab Eng *57*, 96-109.

8. **Atkinson, N., Feike, D., Mackinder, L.C., Meyer, M.T., Griffiths, H., Jonikas, M.C., Smith, A.M., and McCormick, A.J.** (2016). Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. Plant Biotechnol J *14*, 1302-1315.

9. **Attwood, P.V., and Wallace, J.C.** (2002). Chemical and catalytic mechanisms of carboxyl transfer reactions in biotin-dependent enzymes. Acc Chem Res *35*, 113-120.

10. **Baldet, P., Alban, C., Axiotis, S., and Douce, R.** (1993). Localization of free and bound biotin in cells from green pea leaves. Arch Biochem Biophys *303*, 67-73.

11. **Baldoma, L., and Aguilar, J.** (1988). Metabolism of *L*-fucose and *L*-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of *L*-lactaldehyde dissimilation. J Bacteriol *170*, 416-421.

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

13. **Bar-On, Y.M., and Milo, R.** (2019). The global mass and average rate of rubisco. Proc Natl Acad Sci USA *116*, 4738-4743.

14. **Baret, J.C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M.L., Hutchison, J.B., Agresti, J.J., *et al.*** (2009). Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab Chip *9*, 1850-1858.

15. **Bartsch, O., Hagemann, M., and Bauwe, H.** (2008). Only plant-type (GLYK) glycerate kinases produce *D*-glycerate 3-phosphate. FEBS Lett *582*, 3025-3028.

16. **Bauwe, H., Hagemann, M., and Fernie, A.R.** (2010). Photorespiration: players, partners and origin. Trends Plant Sci *15*, 330-336.

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

18. **Beneyton, T., Krafft, D., Bednarz, C., Kleineberg, C., Woelfer, C., Ivanov, I., Vidaković-Koch, T., Sundmacher, K., and Baret, J.-C.** (2018). Out-of-equilibrium microcompartments for the bottom-up integration of metabolic functions. Nat Commun *9*, 2391.

19. **Beneyton, T., Wijaya, I.P.M., Postros, P., Najah, M., Leblond, P., Couvent, A., Mayot, E., Griffiths, A.D., and Drevelle, A.** (2016). High-throughput screening of filamentous fungi using nanoliter-range droplet-based microfluidics. Sci Rep *6*, 27223.

20. **Berg, I.A.** (2011). Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. Appl Environ Microbiol *77*, 1925-1936.

21. **Biyani, N., Righetto, R.D., McLeod, R., Caujolle-Bert, D., Castano-Diez, D., Goldie, K.N., and Stahlberg, H.** (2017). Focus: The interface between data collection and data processing in cryo-EM. J Struct Biol *198*, 124-133.

22. **Boronat, A., and Aguilar, J.** (1979). Rhamnose-induced propanediol oxidoreductase in *Escherichia coli*: purification, properties, and comparison with the fucose-induced enzyme. J Bacteriol *140*, 320-326.

23. **Boronat, A., Caballero, E., and Aguilar, J.** (1983). Experimental Evolution of a Metabolic Pathway for Ethylene Glycol Utilization by *Escherichia coli*. J Bacteriol *153*, 134-139.

24. **Brown, A.P., Slabas, A.R., and Rafferty, J.B.** (2009). Fatty acid biosynthesis in plants - metabolic pathways, structure and organization. In Lipids in photosynthesis (Springer), pp. 11-34.

25. **Chou, C.-Y., Linda, P., and Tong, L.** (2009). Crystal structure of biotin carboxylase in complex with substrates and implications for its catalytic mechanism. J Biol Chem *284*, 11690-11697.

26. **Cirino, P.C., Mayer, K.M., and Umeno, D.** (2003). Generating mutant libraries using error-prone PCR. Methods Mol Biol *231*, 3-9.

27. **Dalal, J., Lopez, H., Vasani, N.B., Hu, Z., Swift, J.E., Yalamanchili, R., Dvora, M., Lin, X., Xie, D., and Qu, R.** (2015). A photorespiratory bypass increases plant growth and seed yield in biofuel crop *Camelina sativa*. Biotechnol Biofuels *8*, 175.

28. **de Diego Puente, T., Gallego-Jara, J., Castaño-Cerezo, S., Sánchez, V.B., Espín, V.F., de la Torre, J.G., Rubio, A.M., and Díaz, M.C.** (2015). The protein acetyltransferase PatZ from *Escherichia coli* is regulated by autoacetylation-induced oligomerization. J Biol Chem *290*, 23077-23093.

29. **Debon, A., Pott, M., Obexer, R., Green, A.P., Friedrich, L., Griffiths, A.D., and Hilvert, D.** (2019). Ultrahigh-throughput screening enables efficient single-round oxidase remodelling. Nat Catal *2*, 740-747.

30. **Doughty, C.C., Hayashi, J.A., and Guenther, H.L.** (1966). Purification and properties of *D*-glycerate 3-kinase from *Escherichia coli*. J Biol Chem *241*, 568-572.

31. **Eisenhut, M., Kahlon, S., Hasse, D., Ewald, R., Lieman-Hurwitz, J., Ogawa, T., Ruth, W., Bauwe, H., Kaplan, A., and Hagemann, M.** (2006). The plant-like C_2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. Plant Physiol *142*, 333-342.

32. **Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M.** (2008). The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proc Natl Acad Sci USA *105*, 17199-17204.

33. **Ellis, R.J.** (1979). The most abundant protein in the world. Trends Biochem Sci *4*, 241-244.

34. **Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K.** (2010). Features and development of Coot. Acta Crystallogr Sect D Biol Crystallogr *66*, 486-501.

35. **Engilberge, S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C.-A., Pitrat, D., Mulatier, J.-C., Madern, D., and Breyton, C.** (2017). Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. Chem Sci *8*, 5909-5917.

36. **Erb, T.J., and Zarzycki, J.** (2018). A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. Curr Opin Biotechnol *49*, 100-107.

37. **Eriksen, M., Lebreton, L.C., Carson, H.S., Thiel, M., Moore, C.J., Borerro, J.C., Galgani, F., Ryan, P.G., and Reisser, J.** (2014). Plastic pollution in the world's oceans: more than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea. PloS one *9*, e111913.

38. **Ermakova, M., Danila, F.R., Furbank, R.T., and Von Caemmerer, S.** (2020). On the road to C_4 rice: advances and perspectives. Plant J *101*, 940–950.

39. **Fahnenstich, H., Scarpeci, T.E., Valle, E.M., Flügge, U.-I., and Maurino, V.G.** (2008). Generation of hydrogen peroxide in chloroplasts of Arabidopsis overexpressing glycolate oxidase as an inducible system to study oxidative stress. Plant Physiol *148*, 719-729.

40. **Fairhead, M., and Howarth, M.** (2015). Site-specific biotinylation of purified proteins using BirA. Methods Mol Biol *1266*, 171-184.

41. **Fernie, A.R., and Bauwe, H.** (2020). Wasteful, essential, evolutionary stepping stone? The multiple personalities of the photorespiratory pathway. Plant J.

42. **Fischlechner, M., Schaerli, Y., Mohamed, M.F., Patil, S., Abell, C., and Hollfelder, F.** (2014). Evolution of enzyme catalysts caged in biomimetic gel-shell beads. Nat Chem *6*, 791-796.

43. **Flamholz, A.I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y.M., Oltrogge, L.M., Alves, R., Savage, D., and Milo, R.** (2019). Revisiting trade-offs between Rubisco kinetic parameters. Biochemistry *58*, 3365-3376.

44. **Franden, M.A., Jayakody, L.N., Li, W.-J., Wagner, N.J., Cleveland, N.S., Michener, W.E., Hauer, B., Blank, L.M., Wierckx, N., and Klebensberger, J.** (2018). Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. Metab Eng *48*, 197-207.

45. **Friedlingstein, P., Jones, M., O'Sullivan, M., Andrew, R., Hauck, J., Peters, G., Peters, W., Pongratz, J., Sitch, S., and Le Quéré, C.** (2019). Global carbon budget 2019. Earth Syst Sci Data *11*, 1783-1838.

46. **Fuchs, G.** (2011). Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu Rev Microbiol *65*, 631-658.

47. **Giordano, M., Beardall, J., and Raven, J.A.** (2005). CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. Annu Rev Plant Biol *56*, 99-131.

48. **Gotto, A., and Kornberg, H.** (1961). The metabolism of C_2 compounds in micro-organisms. 7. Preparation and properties of crystalline tartronic semialdehyde reductase. Biochem J *81*, 273.

49. **Hagemann, M., Kern, R., Maurino, V.G., Hanson, D.T., Weber, A.P., Sage, R.F., and Bauwe, H.** (2016). Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. J Exp Bot *67*, 2963-2976.

50. **Hall, P.R., Wang, Y.F., Rivera-Hainaj, R.E., Zheng, X., Pustai-Carey, M., Carey, P.R., and Yee, V.C.** (2003). Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core. EMBO J *22*, 2334-2347.

51. **Han, J., Gagnon, S., Eckle, T., and Borchers, C.H.** (2013). Metabolomic analysis of key central carbon metabolism carboxylic acids as their 3-nitrophenylhydrazones by UPLC/ESI-MS. Electrophoresis *34*, 2891-2900.

52. **Han, X., Liu, W., Huang, J.-W., Ma, J., Zheng, Y., Ko, T.-P., Xu, L., Cheng, Y.-S., Chen, C.-C., and Guo, R.-T.** (2017). Structural insight into catalytic mechanism of PET hydrolase. Nat Commun *8*, 1-6.

53. **Hansen, R.W., and Hayashi, J.A.** (1962). Glycolate metabolism in *Escherichia coli*. J Bacteriol *83*, 679-687.

54. **Hickey, L.T., Hafeez, A.N., Robinson, H., Jackson, S.A., Leal-Bertioli, S.C., Tester, M., Gao, C., Godwin, I.D., Hayes, B.J., and Wulff, B.B.** (2019). Breeding crops to feed 10 billion. Nat Biotechnol *37*, 744-754.

55. **Holo, H., and Sirevåg, R.** (1986). Autotrophic growth and CO_2 fixation of *Chloroflexus aurantiacus*. Arch Microbiol *145*, 173-180.

56. **Huang, C.S., Sadre-Bazzaz, K., Shen, Y., Deng, B., Zhou, Z.H., and Tong, L.** (2010). Crystal structure of the $\alpha_6\beta_6$ holoenzyme of propionyl-coenzyme A carboxylase. Nature *466*, 1001-1005.

57. **Huang, X., Cao, L., Qin, Z., Li, S., Kong, W., and Liu, Y.** (2018). Tat-independent secretion of polyethylene terephthalate hydrolase PETase in *Bacillus subtilis* 168 mediated by its native signal peptide. J Agric Food Chem *66*, 13217-13227.

58. **Hügler, M., Menendez, C., Schägger, H., and Fuchs, G.** (2002). Malonyl-Coenzyme A Reductase from *Chloroflexus aurantiacus*, a Key Enzyme of the 3-Hydroxypropionate Cycle for Autotrophic CO_2 Fixation. J Bacteriol *184*, 2404-2410.

59. **Hügler, M., and Sievert, S.M.** (2011). Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. Annu Rev Mar Sci *3*, 261-289.

60. **Jensen, R.A.** (1976). Enzyme recruitment in evolution of new function. Annu Rev Microbiol *30*, 409-425.

61. **Jeschek, M., Reuter, R., Heinisch, T., Trindler, C., Klehr, J., Panke, S., and Ward, T.R.** (2016). Directed evolution of artificial metalloenzymes for *in vivo* metathesis. Nature *537*, 661-665.

62. **Kabsch, W.** (2010). Xds. Acta Crystallogr Sect D Biol Crystallogr *66*, 125-132.

63. **Kan, S.J., Lewis, R.D., Chen, K., and Arnold, F.H.** (2016). Directed evolution of cytochrome c for carbon–silicon bond formation: Bringing silicon to life. Science *354*, 1048-1051.

64. **Kawaguchi, A., Yoshimura, T., and Okuda, S.** (1981). A new method for the preparation of acyl-CoA thioesters. J Biochem *89*, 337-339.

65. **Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F., and Peterhansel, C.** (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. Nat Biotechnol *25*, 593-599.

66. **Kelly, G., and Latzko, E.** (1976). Inhibition of spinach-leaf phosphofructokinase by 2-phosphoglycollate. FEBS Lett *68*, 55-58.

67. **Khersonsky, O., and Tawfik, D.S.** (2010). Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem *79*, 471-505.

68. **Kille, S., Acevedo-Rocha, C.G., Parra, L.P., Zhang, Z.G., Opperman, D.J., Reetz, M.T., and Acevedo, J.P.** (2013). Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. ACS Synth Biol *2*, 83-92.

69. **Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H.** (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (A Complete Set of *E. coli* K-12 ORF Archive): Unique Resources for Biological Research. DNA Res *12*, 291-299.

70. **Knowles, J.R.** (1989). The mechanism of biotin-dependent enzymes. Annu Rev Biochem *58*, 195-221.

71. **Krakow, G., Barkulis, S.S., and Hayashi, J.A.** (1961). Glyoxylic acid carboligase: an enzyme present in glycolate-grown *Escherichia coli*. J Bacteriol *81*, 509.

72. **Kroeger, J.K., Zarzycki, J., and Fuchs, G.** (2011). A spectrophotometric assay for measuring acetyl–coenzyme A carboxylase. Anal Biochem *411*, 100-105.

73. **Leister, D.** (2019). Thawing out frozen metabolic accidents. BMC Biol *17*, 8.

74. **Li, W.J., Jayakody, L.N., Franden, M.A., Wehrmann, M., Daun, T., Hauer, B., Blank, L.M., Beckham, G.T., Klebensberger, J., and Wierckx, N.** (2019). Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440. Environ Microbiol *21*, 3669–3682.

75. **Li, X., Grigorieff, N., and Cheng, Y.** (2010). GPU-enabled FREALIGN: accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors. J Struct Biol *172*, 407-412.

76. **Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., and McCoy, A.J.** (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861-877.

77. **Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A., and Hanson, M.R.** (2014). A faster Rubisco with potential to increase photosynthesis in crops. Nature *513*, 547-550.

78. **Lindenkamp, N., Schürmann, M., and Steinbüchel, A.** (2013). A propionate CoA-transferase of *Ralstonia eutropha* H16 with broad substrate specificity catalyzing the CoA thioester formation of various carboxylic acids. Appl Microbiol Biotechnol *97*, 7699-7709.

79. **Liu, C., Wang, Q., Xian, M., Ding, Y., and Zhao, G.** (2013). Dissection of malonyl-coenzyme A reductase of *Chloroflexus aurantiacus* results in enzyme activity improvement. PloS one *8*, e75554.

80. **Liu, D., Ramya, R.C.S., and Mueller-Cajar, O.** (2017). Surveying the expanding prokaryotic Rubisco multiverse. FEMS Microbiol Lett *364*, fnx156.

81. **Long, B.M., Hee, W.Y., Sharwood, R.E., Rae, B.D., Kaines, S., Lim, Y.-L., Nguyen, N.D., Massey, B., Bala, S., and von Caemmerer, S.** (2018). Carboxysome encapsulation of the CO_2-fixing enzyme Rubisco in tobacco chloroplasts. Nat Commun *9*, 1-14.

82. **Lorimer, G.H., Badger, M.R., and Andrews, T.J.** (1976). The activation of ribulose-1,5-bisphosphate carboxylase by carbon dioxide and magnesium ions. Equilibria, kinetics, a suggested mechanism, and physiological implications. Biochemistry *15*, 529-536.

83. **Maier, A., Fahnenstich, H., Von Caemmerer, S., Engqvist, M., Weber, A., Flügge, U.-I., and Maurino, V.** (2012). Transgenic introduction of a glycolate oxidative cycle into *A. thaliana* chloroplasts leads to growth improvement. Front Plant Sci *3*.

84. **Martínez-García, E., and de Lorenzo, V.** (2012). Transposon-based and plasmid-based genetic tools for editing genomes of Gram-negative bacteria. In Synthetic Gene Networks (Springer), pp. 267-283.

85. **Martínez-García, E., and de Lorenzo, V.** (2011). Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. Environ Microbiol *13*, 2702-2716.

86. **Mastronarde, D.N.** (2005). Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol *152*, 36-51.

87. **Meyer, M., Schweiger, P., and Deppenmeier, U.** (2015). Succinic semialdehyde reductase Gox1801 from *Gluconobacter oxydans* in comparison to other succinic semialdehyde-reducing enzymes. Appl Microbiol Biotechnol *99*, 3929-3939.

88. **Miyazaki, K.** (2011). MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. Methods Enzymol *498*, 399-406.

89. **Moog, D., Schmitt, J., Senger, J., Zarzycki, J., Rexer, K.-H., Linne, U., Erb, T., and Maier, U.G.** (2019). Using a marine microalga as a chassis for polyethylene terephthalate (PET) degradation. Microbial cell factories *18*, 171.

90. **Mordhorst, S., Siegrist, J., Müller, M., Richter, M., and Andexer, J.N.** (2017). Catalytic Alkylation Using a Cyclic S-Adenosylmethionine Regeneration System. Angew Chem Int Ed *56*, 4037-4041.

91. **Mückschel, B., Simon, O., Klebensberger, J., Graf, N., Rosche, B., Altenbuchner, J., Pfannstiel, J., Huber, A., and Hauer, B.** (2012). Ethylene glycol metabolism by *Pseudomonas putida*. Appl Environ Microbiol *78*, 8531-8539.

92. **Mueller-Cajar, O.** (2017). The diverse AAA+ machines that repair inhibited Rubisco active sites. Front Mol Biosci *4*, 31.

93. **Mueller-Cajar, O., Morell, M., and Whitney, S.M.** (2007). Directed evolution of Rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. Biochemistry *46*, 14067-14074.

94. **Nakamura, Y., Kanakagiri, S., Van, K., He, W., and Spalding, M.H.** (2005). Disruption of the glycolate dehydrogenase gene in the high-CO_2-requiring mutant HCR89 of *Chlamydomonas reinhardtii*. Can J Bot *83*, 820-833.

95. **Nocek, B., Kochinyan, S., Proudfoot, M., Brown, G., Evdokimova, E., Osipiuk, J., Edwards, A.M., Savchenko, A., Joachimiak, A., and Yakunin, A.F.** (2008). Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria. Proc Natl Acad Sci USA *105*, 17730-17735.

96. **Nölke, G., Houdelet, M., Kreuzaler, F., Peterhänsel, C., and Schillberg, S.** (2014). The expression of a recombinant glycolate dehydrogenase polyprotein in potato (*Solanum tuberosum*) plastids strongly enhances photosynthesis and tuber yield. Plant Biotechnol J *12*, 734-742.

97. **Nowak, C., Beer, B., Pick, A., Roth, T., Lommes, P., and Sieber, V.** (2015). A water-forming NADH oxidase from *Lactobacillus pentosus* suitable for the regeneration of synthetic biomimetic cofactors. Front Microbiol *6*.

98. **Obexer, R., Godina, A., Garrabou, X., Mittl, P.R., Baker, D., Griffiths, A.D., and Hilvert, D.** (2017). Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. Nat Chem *9*, 50-56.

99. **Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R., and Parry, M.A.** (2016). Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. Plant J *85*, 148-160.

100. **Ort, D.R., Merchant, S.S., Alric, J., Barkan, A., Blankenship, R.E., Bock, R., Croce, R., Hanson, M.R., Hibberd, J.M., and Long, S.P.** (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. Proc Natl Acad Sci USA *112*, 8529-8536.

101. **Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., and Dasgupta, P.** (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 (IPCC).

102. **Peter, D.M., Vögeli, B., Cortina, N.S., and Erb, T.J.** (2016). A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters. Molecules *21*, 517.

103. **Peterhansel, C., and Maurino, V.G.** (2011). Photorespiration redesigned. Plant Physiol *155*, 49-55.

104. **Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E.** (2004). UCSF Chimera - a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

105. **Phillips, R., and Milo, R.** (2009). A feeling for the numbers in biology. Proc Natl Acad Sci USA *106*, 21465-21471.

106. **Post-Beittenmiller, D., Roughan, G., and Ohlrogge, J.B.** (1992). Regulation of plant fatty acid biosynthesis: Analysis of acyl-coenzyme A and acyl-acyl carrier protein substrate pools in spinach and pea chloroplasts. Plant Physiol *100*, 923-930.

107. **Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A.** (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290.

108. **Ranson, S.L., and Thomas, M.** (1960). Crassulacean acid metabolism. Annu Rev Plant Biol *11*, 81-110.

109. **Reger, A.S., Carney, J.M., and Gulick, A.M.** (2007). Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase. Biochemistry *46*, 6536-6546.

110. **Salvador, M., Abdulmutalib, U., Gonzalez, J., Kim, J., Smith, A.A., Faulon, J.-L., Wei, R., Zimmermann, W., and Jimenez, J.I.** (2019). Microbial genes for a circular and sustainable Bio-PET economy. Genes *10*, 373.

111. **Sambrook, J.F., and Russell, D.W.** (2001). Molecular Cloning: A Laboratory Manual, 3rd edn (Cold Spring Harbor Laboratory Press, New York).

112. **Savir, Y., Noor, E., Milo, R., and Tlusty, T.** (2010). Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. Proc Natl Acad Sci USA *107*, 3475-3480.

113. **Schada von Borzyskowski, L., Severi, F., Krüger, K., Hermann, L., Gilardet, A., Sippel, F., Pommerenke, B., Claus, P., Cortina, N.S., Glatter, T.*, et al.*** (2019). Marine Proteobacteria metabolise glycolate via the β-hydroxyaspartate cycle. Nature *575*, 500–504.

114. **Scheres, S.H.** (2012). A Bayesian view on cryo-EM structure determination. J Mol Biol *415*, 406-418.

115. **Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N.S., and Erb, T.J.** (2016). A synthetic pathway for the fixation of carbon dioxide *in vitro*. Science *354*, 900-904.

116. **Sedelnikova, O.V., Hughes, T.E., and Langdale, J.A.** (2018). Understanding the genetic basis of C_4 Kranz anatomy with a view to engineering C_3 crops. Annu Rev Genet *52*, 249-270.

117. **Selmer, T., Willanzheimer, A., and Hetzel, M.** (2002). Propionate CoA-transferase from *Clostridium propionicum*: Cloning of the gene and identification of glutamate 324 at the active site. Eur J Biochem *269*, 372-380.

118. **Seo, H., Kim, S., Son, H.F., Sagong, H.-Y., Joo, S., and Kim, K.-J.** (2019). Production of extracellular PETase from *Ideonella sakaiensis* using sec-dependent signal peptides in *E. coli*. Biochem Biophys Res Commun *508*, 250-255.

119. **Sharkey, T.D.** (1988). Estimating the rate of photorespiration in leaves. Physiol Plant *73*, 147-152.

120. **Shenoy, A.R., and Visweswariah, S.S.** (2003). Site-directed mutagenesis using a single mutagenic oligonucleotide and *Dpn*I digestion of template DNA. Anal Biochem *319*, 335-336.

121. **Shih, P.M., Occhialini, A., Cameron, J.C., Andralojc, P.J., Parry, M.A., and Kerfeld, C.A.** (2016). Biochemical characterization of predicted Precambrian RuBisCO. Nat Commun *7*, 10382.

122. **Shih, P.M., Zarzycki, J., Niyogi, K.K., and Kerfeld, C.A.** (2014). Introduction of a Synthetic CO_2-fixing Photorespiratory Bypass into a Cyanobacterium. J Biol Chem *289*, 9493-9500.

123. **Siegel, J.B., Smith, A.L., Poust, S., Wargacki, A.J., Bar-Even, A., Louw, C., Shen, B.W., Eiben, C.B., Tran, H.M., Noor, E., *et al.*** (2015). Computational protein design enables a novel one-carbon assimilation pathway. Proc Natl Acad Sci USA *112*, 3704-3709.

124. **Slack, C.R., and Hatch, M.D.** (1967). Comparative studies on the activity of carboxylases and other enzymes in relation to the new pathway of photosynthetic carbon dioxide fixation in tropical grasses. Biochem J *103*, 660.

125. **Son, H.F., Kim, S., Seo, H., Hong, J., Lee, D., Jin, K.S., Park, S., and Kim, K.J.** (2020). Structural insight into bi-functional malonyl-CoA reductase. Environ Microbiol *22*, 752-765.

126. **South, P.F., Cavanagh, A.P., Liu, H.W., and Ort, D.R.** (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. Science *363*.

127. **Sridhara, S., Wu, T.T., Chused, T.M., and Lin, E.C.C.** (1969). Ferrous-activated Nicotinamide Adenine Dinucleotide-linked Dehydrogenase from a Mutant of *Escherichia coli* Capable of Growth on 1,2-Propanediol. J Bacteriol *98*, 87-95.

128. **Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D., and Escalante-Semerena, J.C.** (2002). Sir2-dependent Activation of Acetyl-CoA Synthetase by Deacetylation of Active Lysine. Science *298*, 2390-2392.

129. **Starai, V.J., and Escalante-Semerena, J.C.** (2004). Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. J Mol Biol *340*, 1005-1012.

130. **Starai, V.J., Gardner, J.G., and Escalante-Semerena, J.C.** (2005). Residue Leu641 of acetyl-CoA synthetase (Acs) is critical for the acetylation of residue Lys609 by the protein acetyltransferase (Pat) enzyme of *Salmonella enterica*. J Biol Chem *280*, 26200-26205.

131. **Stec, B.** (2012). Structural mechanism of RuBisCO activation by carbamylation of the active site lysine. Proc Natl Acad Sci USA *109*, 18785-18790.

132. **Studer, R.A., Christin, P.-A., Williams, M.A., and Orengo, C.A.** (2014). Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. Proc Natl Acad Sci USA *111*, 2223-2228.

133. **Tabita, F.R., Hanson, T.E., Li, H., Satagopan, S., Singh, J., and Chan, S.** (2007). Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. Microbiol Mol Biol Rev *71*, 576-599.

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

135. **Tabor, S., and Richardson, C.C.** (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci USA *82*, 1074-1078.

136. **Tawfik, D.S., and Griffiths, A.D.** (1998). Man-made cell-like compartments for molecular evolution. Nat Biotechnol *16*, 652-656.

137. **Tcherkez, G.** (2013). Modelling the reaction mechanism of ribulose-1,5-bisphosphate carboxylase/oxygenase and consequences for kinetic parameters. Plant Cell Environ *36*, 1586-1596.

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

139. **Tong, L.** (2013). Structure and function of biotin-dependent carboxylases. Cell Mol Life Sci *70*, 863-891.

140. **Tong, L.** (2017). Striking diversity in holoenzyme architecture and extensive conformational variability in biotin-dependent carboxylases. In Advances in protein chemistry and structural biology (Elsevier), pp. 161-194.

141. **Toscano, M.D., Woycechowsky, K.J., and Hilvert, D.** (2007). Minimalist active-site redesign: teaching old enzymes new tricks. Angew Chem Int Ed *46*, 3212-3236.

142. **Trudeau, D.L., Edlich-Muth, C., Zarzycki, J., Scheffen, M., Goldsmith, M., Khersonsky, O., Avizemer, Z., Fleishman, S.J., Cotton, C.A.R., Erb, T.J., *et al.*** (2018). Design and in vitro realization of carbon-conserving photorespiration. Proc Natl Acad Sci USA *115*, E11455-E11464.

143. **Vagstad, A.L., Bumpus, S.B., Belecki, K., Kelleher, N.L., and Townsend, C.A.** (2012). Interrogation of global active site occupancy of a fungal iterative polyketide synthase reveals strategies for maintaining biosynthetic fidelity. J Am Chem Soc *134*, 6865-6877.

144. **Vamecq, J., and Poupaert, J.H.** (1990). Studies on the metabolism of glycolyl-CoA. Biochem Cell Biol *68*, 846-851.

145. **Villarreal, J.C., and Renner, S.S.** (2012). Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. Proc Natl Acad Sci USA *109*, 18873-18878.

146. **Walker, B.J., VanLoocke, A., Bernacchi, C.J., and Ort, D.R.** (2016). The Costs of Photorespiration to Food Production Now and in the Future. Annu Rev Plant Biol *67*, 107-129.

147. **Weber, A.P., and Bar-Even, A.** (2019). Update: Improving the Efficiency of Photosynthetic Carbon Reactions. Plant Physiol *179*, 803-812.

148. **Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., and McCoy, A.** (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr *67*, 235-242.

149. **Wurtzel, E.T., Vickers, C.E., Hanson, A.D., Millar, A.H., Cooper, M., Voss-Fels, K.P., Nikel, P.I., and Erb, T.J.** (2019). Revolutionizing agriculture with synthetic biology. Nat Plants *5*, 1207–1210.

150. **Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., and Oda, K.** (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). Science *351*, 1196-1199.

151. **Zarzycki, J., Brecht, V., Müller, M., and Fuchs, G.** (2009). Identifying the missing steps of the autotrophic 3-hydroxypropionate CO_2 fixation cycle in *Chloroflexus aurantiacus*. Proc Natl Acad Sci USA *106*, 21317-21322.

152. **Zarzycki, J., Sutter, M., Cortina, N.S., Erb, T.J., and Kerfeld, C.A.** (2017). *In Vitro* Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. Sci Rep *7*, 42757.

153. **Zelcbuch, L., Razo-Mejia, M., Herz, E., Yahav, S., Antonovsky, N., Kroytoro, H., Milo, R., and Bar-Even, A.** (2015). An *in vivo* metabolic approach for deciphering the product specificity of glycerate kinase proves that both *E. coli*'s glycerate kinases generate 2-phosphoglycerate. PLoS ONE *10*, e0122957.

154. **Zelitch, I., Schultes, N.P., Peterson, R.B., Brown, P., and Brutnell, T.P.** (2009). High glycolate oxidase activity is required for survival of maize in normal air. Plant Physiol *149*, 195-204.

155. **Zerihun, A., McKENZIE, B.A., and Morton, J.D.** (1998). Photosynthate costs associated with the utilization of different nitrogen–forms: influence on the carbon balance of plants and shoot–root biomass partitioning. New Phytol *138*, 1-11.

156. **Zhang, X., Zhang, B., Lin, J., and Wei, D.** (2015). Oxidation of ethylene glycol to glycolaldehyde using a highly selective alcohol dehydrogenase from *Gluconobacter oxydans*. J Mol Catal B Enzym *112*, 69-75.

157. **Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A.** (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332.

List of Abbreviations

2-PG	2-phosphoglycolate
2-PGA	2-phosphoglycerate
3-NPH	3-nitrophenylhydrazine
3OHP	3-hydroxypropionate bypass
3-PGA	3-phosphoglycerate
A. thaliana	*Arabidopsis thaliana*
A5P	arabinose-5-phosphate shunt
AbfT	4-OH-butyrate CoA transferase
ACC	acetyl-CoA carboxylase
ACS	acyl-CoA synthetase
ADP	adenosine diphosphate
AMP	adenosine monophosphate
Amp	ampicillin
app.	apparent
ATP	adenosine triphosphate
BC	biotin carboxylase domain
BCCP	biotin carboxyl carrier protein
BirA	biotin ligase
C. aminobutyricum	*Clostridium aminobutyricum*
C. aurantiacus	*Chloroflexus aurantiacus*
C. difficile	*Clostridioides difficile*
C. kluyveri	*Clostridium kluyveri*
C. propionicum	*Clostridium propionicum*
C. sativa	*Camelina sativa*
CAM	crassulacean acid metabolism
Cam	chloramphenicol
CBB cycle	Calvin-Benson-Bassham cycle
CCM	carbon concentrating mechanism
CDI	carbonyldiimidazole
CDS	correlated double sampling
CETCH cycle	crotonyl-coenzyme A (CoA)/ethylmalonyl-CoA/hydroxy-butyryl-CoA cycle
CoA	coenzyme A
CP	phosphocreatine
CPK	creatine phosphokinase
CPR	candidate phyla radiation
cryo-EM	crogenic electron microscopy

CT	carboxyltransferase
CTF	contrast transfer function
cv	column volumes
DAQ	data acquisition
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DH	dehydrogenase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dTTP	deoxythymidine triphosphate
E. coli	*Escherichia coli*
Ecm	ethylmalonyl-CoA mutase
EDC	1-(3-dimethyl-aminopropyl)-3-ethylcarbodiimide
EG	ethylene glycol
Eno	enolase
FAD	flavin adenine dinucleotide
FADS	fluorescence-activated droplet sorting
Fd_{ox}	oxidized ferredoxin
Fd_{red}	reduced ferredoxin
FSC	fourier shell correlation
FucO	(*L*)-lactaldehyde dehydrogenase
G6PDH	glucose-6-phosphate dehydrogenase
GA	glycolaldehyde
GAP	glyceraldehyde 3-phosphate
GCC	glycolyl-CoA carboxylase
Gcl	glyoxylate carboligase
gCoA	glycolyl-CoA
GCR	glycolyl-CoA reductase
GCS	glycolyl-CoA synthetase
GLB	golden lysogeny broth
GLC	glycerate bypass
GlcB	malate synthase
GlxK	glycerate kinase
Gox 0313	alcohol dehydrogenase
Gox1801	succinic semialdehyde/glyoxylate reductase
Gpm	phosphoglycerate mutase
Hbd	4-hydroxybutyryl-CoA dehydratase
Hbs	4-hydroxybutyryl-CoA synthetase

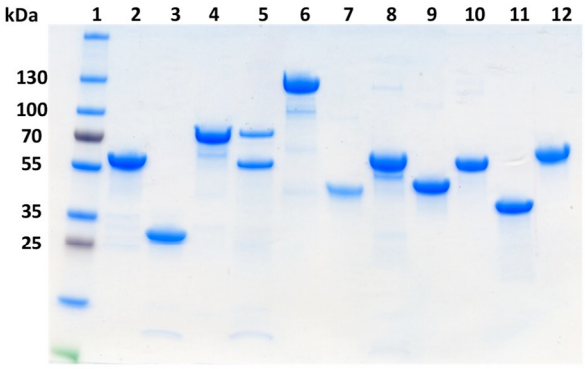
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC-MS	high performance liquid chromatography coupled mass spectrometry
IPTG	isopropyl β-*D*-1-thiogalactopyranoside
JGI	Joint Genome Institute
Kan	kanamycin
k_{cat}	turnover number
k_{cat}/K_M	catalytic efficiency
k_{cat}^c	carboxylation turnover number
k_{cat}^C/K_M^C	carboxylation catalytic efficiency
k_{cat}^O/K_M^O	oxygenation catalytic efficiency
K_M	Michaelis-Menten constant
L. pentosus	*Lactobacillus pentosus*
LB	lysogeny broth
LCC	long-chain acyl-CoA carboxylase
LDH	lactate dehydrogenase
M. extorquens	*Methylorubrum extorquens*
MatB	malonyl-CoA synthetase
Mcl	β-methylmalonyl-CoA mutase
Mcm	methylmalonyl-CoA mutase
Mco	methylsuccinyl-CoA oxidase
MCR	malonyl-CoA reductase
MDH	malate dehydrogenase
megaWHOP	megaprimer whole plasmid PCR
MOPS	3-(N-morpholino)propanesulfonic acid
MPI	Max Planck Institute
MRM	multiple reaction monitoring
Myo	myokinase
N. tabacum	*Nicotiana tabacum*
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
Nox	NADH oxidase
NPR	natural photorespiration
O. sativa	*Oryza sativa*
OA	oxaloacetate
OD	optical density
OX	glycolate oxidation pathway
P. dokdonensis	*Porphyrobacter dokdonensis*
P. putida	*Pseudomonas putida*
PBS	phosphate buffered saline

PCC	propionyl-CoA carboxylase
Pco	propionyl-CoA oxidase
pCoA	propionyl-CoA
PCR	polymerase chain reaction
PCS	propionyl-CoA synthetase
PCT	propionate CoA transferase
PDB	protein data bank
PDMS	polydimethylsiloxane
PduP	CoA-acylating propionaldehyde dehydrogenase
PEP	phosphoenolpyruvate
PEPC	phosphoenolpyruvate carboxylase
PET	polyethylene terephthalate
Pgp	phosphoglycolate phosphatase
P_i	inorganic phosphate
PK	pyruvate kinase
PMT	photomultiplier tube
PP	polyphosphate
PP_i	inorganic pyrophosphate
PPK	polyphosphate kinase
PTFE	polytetrafluoroethylene
QM/MM	quantum mechanics/molecular mechanics
R. eutropha	*Ralstonia eutropha*
R. pomeroyi	*Ruegeria pomeroyi*
RbcL	RuBisCO large subunit
RbcS	RuBisCO small subunit
RNAi	ribonucleic acid interference
RT	room temperature
Ru5P	ribulose 5-phosphate
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-1,5-bisphosphate
S. enterica	*Salmonella enterica*
SAD	single-wavelength anomalous scattering
$S_{C/O}$	CO_2/O_2 specificity factor
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SOEing	splicing by overlap extension
Spec	spectinomycin
SU	subunit
SucD	succinyl-CoA reductase
TaCo pathway	tartronyl-CoA pathway
TB	terrific broth

TCA cycle	tricarboxylic acid cycle
tCoA	tartronyl-CoA
TCR	tartronyl-CoA reductase
THF	tetrahydrofolate
Tris	tris(hydroxymethyl)aminomethane
TS1	target sequence upstream
TS2	target sequence downstream
U	unit (in µmol min^{-1})
UPLC-hrMS	ultra-high performance liquid chromatography coupled high resolution mass spectrometry
v_{max}	maximal specific enzyme activity (at saturating substrate concentration)
w/o	water-in-oil
wt	wildtype

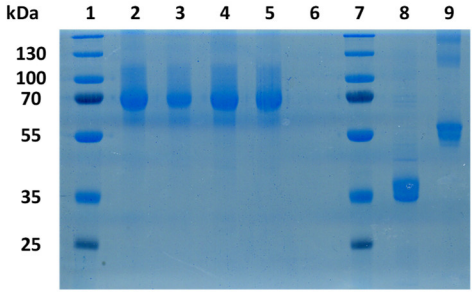
Supplementary Material

A



#	enzyme	organism	size [kDa]
1	Marker		
2	RuBisCO	*R. rubrum*	52.5
3	Pgp	*E. coli*	27.4
4	GCS	*Erythrobacter sp.*	73.0
5	GCC M4	*M. extorquens*	71.9 (α), 56.0 (β)
6	TCR	*C. aurantiacus*	133.1
7	GlxK	*E. coli*	41.1
8	Gpm	*E. coli*	58.5
9	Eno	*E. coli*	48.0
10	PK	*E. coli*	52.3
11	FucO	*E. coli*	43.0
12	PduP	*R. palustris*	55.5

B



#	enzyme	organism	size [kDa]
1	Marker		
2	ACS1	*Erythrobacter sp.*	73.0
3	GCS	*Erythrobacter sp.*	73.0
4	ACS1 VG	*Erythrobacter sp.*	73.0
5	ACS1 VS	*Erythrobacter sp.*	73.0
6	empty		
7	Marker		
8	Gox0313	*G. oxydans*	39.0
9	Nox	*L. pentosus*	51.9

Figure S1. Exemplary SDS gel analysis of proteins of the tartronyl-CoA pathway.
(**A**) Enzymes of the tartronyl-CoA pathway including malate read-out enzymes and ethylene glycol module enzymes. (**B**) Variants of *Erythrobacter* ACS1 and enzymes of the ethylene glycol module.

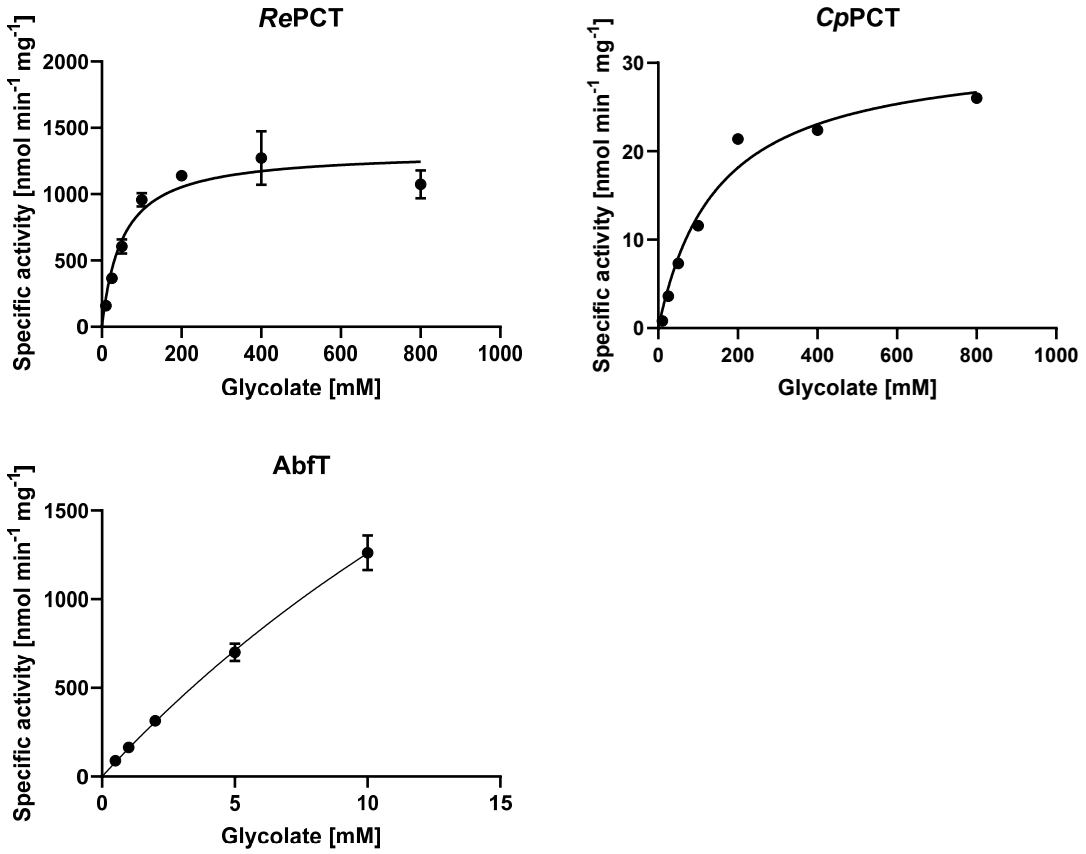


Figure S2. Michaelis-Menten kinetics of different acyl-CoA transferases.
*Re*PCT: propionyl-CoA transferase of *C. propionicum*, *Cp*PCT: propionyl-CoA transferase of *C. propionicum*, AbfT: 4-OH-butyryl-CoA transferase of *C. aminobutyricum*.

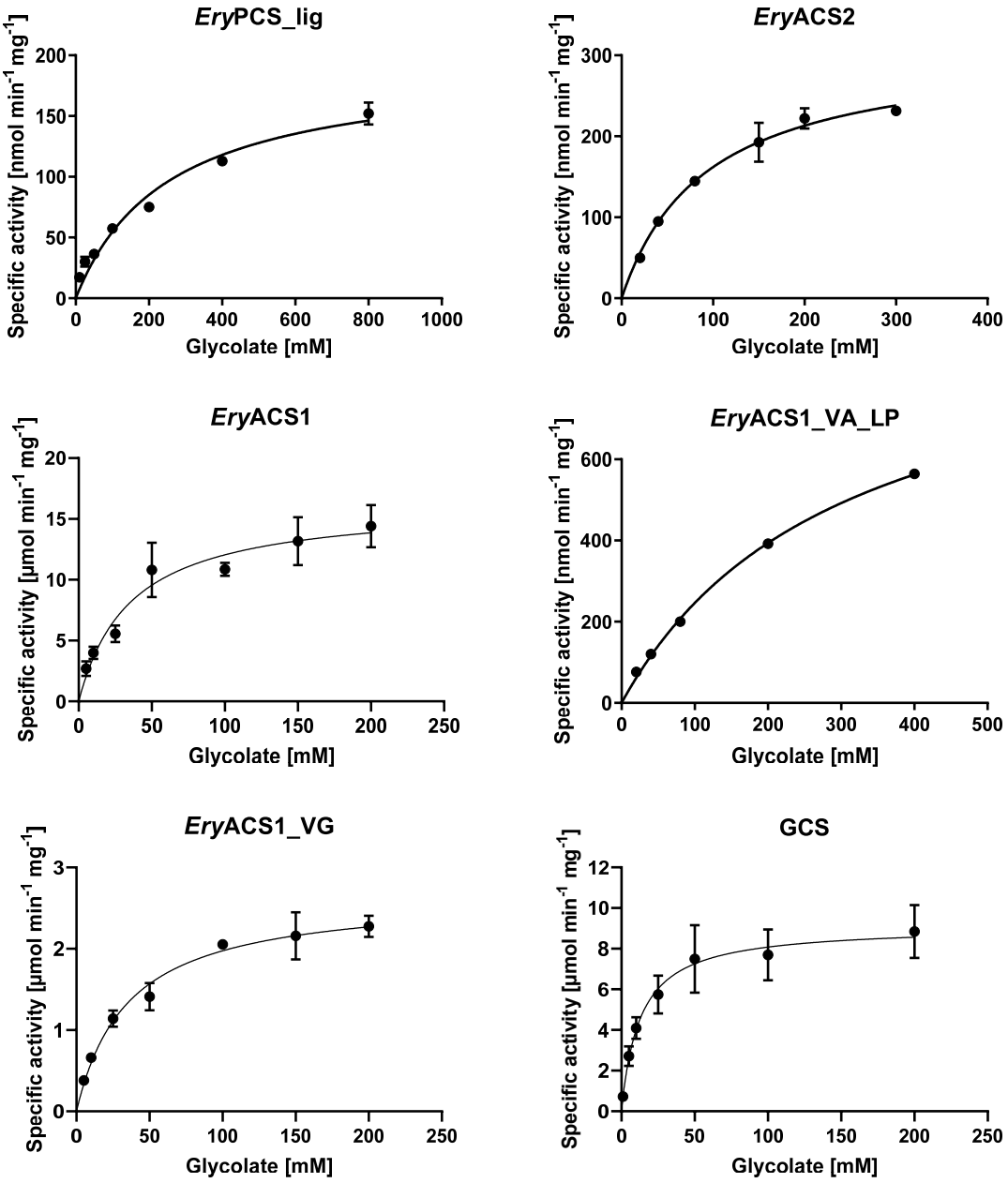


Figure S3. Michaelis-Menten kinetics of different acyl-CoA synthetases.
*Ery*PCS_lig: ligase domain of propionyl-CoA synthase from *Erythrobacter* sp. NAP1, *Ery*ACS2: acetyl-CoA synthetase 2 from *Erythrobacter* sp. NAP1, *Ery*ACS1: acetyl-CoA synthetase 1 from *Erythrobacter* sp. NAP1, *Ery*ACS1_VA_LP: acetyl-CoA synthetase 1 from *Erythrobacter* sp. NAP1 with two mutations (V379A and L641P), *Ery*ACS1_VG: acetyl-CoA synthetase 1 from *Erythrobacter* sp. NAP1 with V379G mutation, GCS: acetyl-CoA synthetase 1 from *Erythrobacter* sp. NAP1 with V379A mutation.

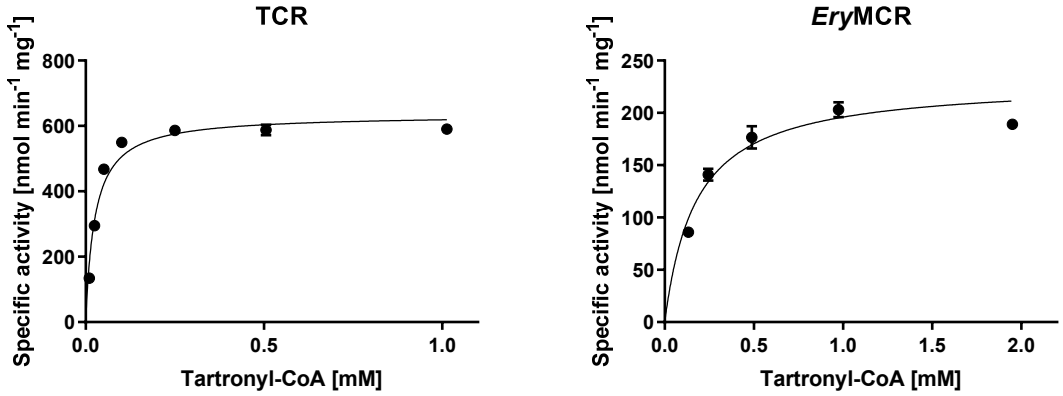


Figure S4. Michaelis-Menten kinetics of tartronyl-CoA reduction by two tested isoenzymes of malonyl-CoA reductases.
*Ca*MCR/TCR: malonyl-CoA reductase of *C. aurantiacus*. *Ery*MCR: malonyl-CoA reductase of *Erythrobacter sp*. NAP1.

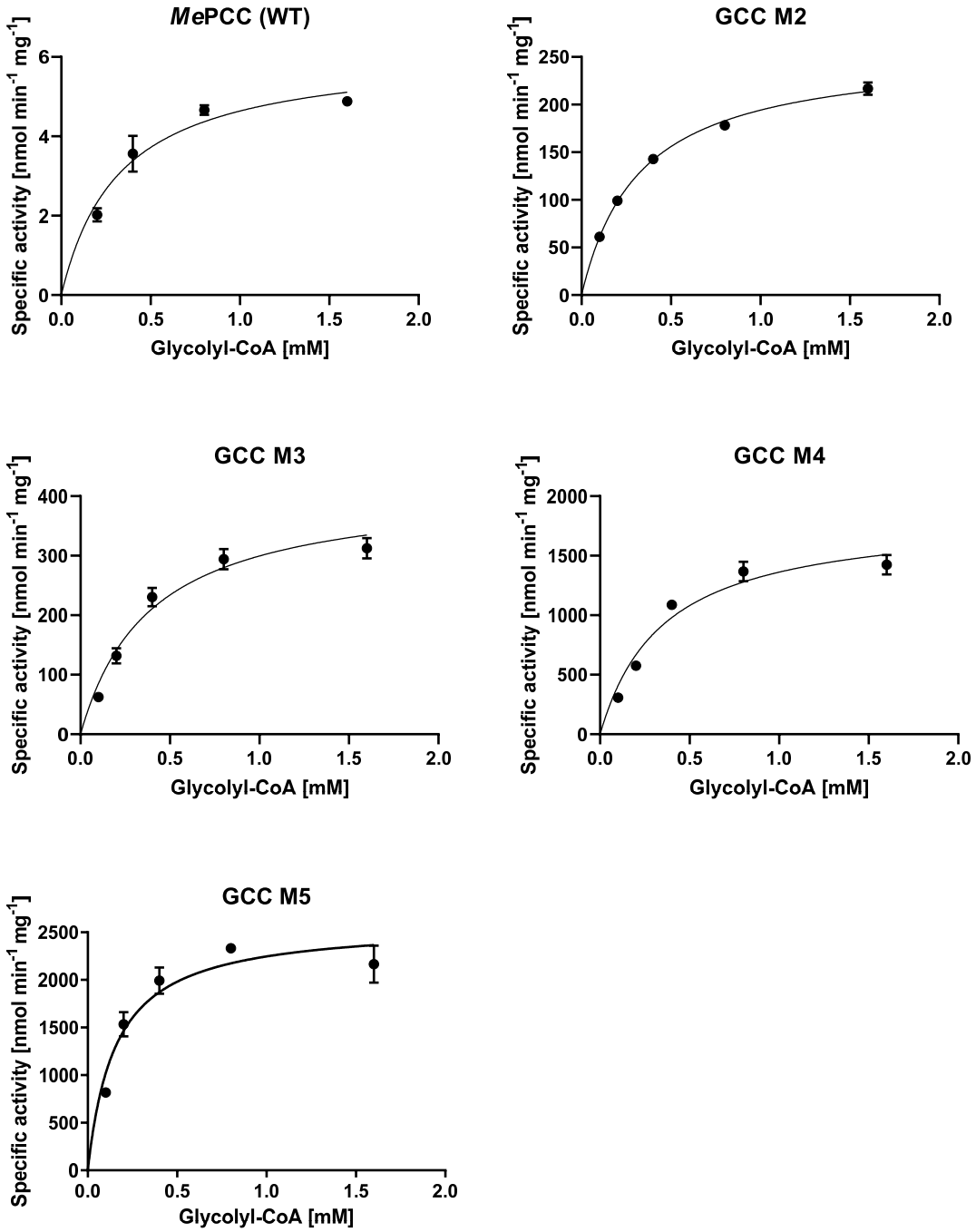


Figure S5. Michaelis-Menten kinetics of glycolyl-CoA carboxylases.
*Me*PCC (WT): propionyl-CoA carboxylase of *M. extorquens* AM1, GCC M2: *Me*PCC with Y143H and D407I mutations, GCC M3: *Me*PCC with L100S, Y143H and D407I mutations, GCC M4: *Me*PCC with L100S, Y143H, D407I and I450V mutations, GCC M5: *Me*PCC with L100S, Y143H, D407I, I450V and W502R mutations.

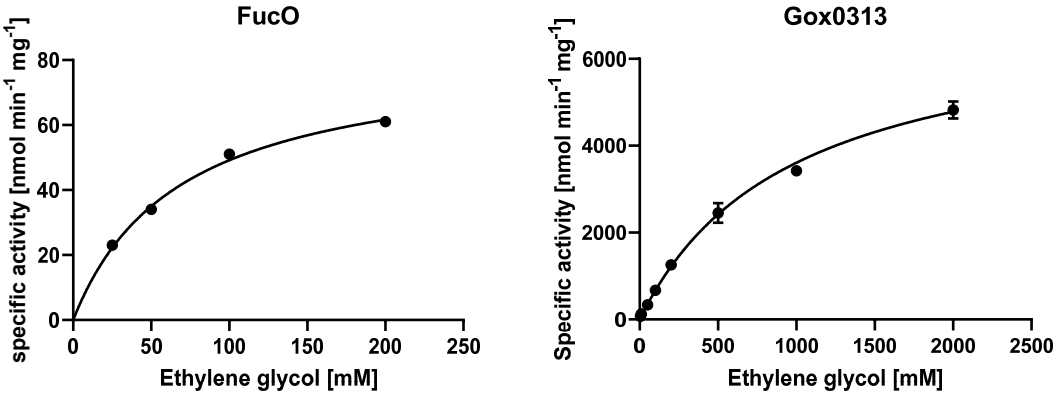


Figure S6. Michaelis-Menten kinetics of dehydrogenases FucO from *E. coli* and Gox0313 from *G. oxydans*.

Table S1. Enzymes and cofactors of different natural and synthetic photorespiratory bypasses.

	per reaction/turn				NPR	GLC	OX	A5P	3OHP	TaCo
2-PG molecules per 3-PGA regeneration					**2**	**2**	**1**	**1**	**1**	**1**
enzyme/process	**ATP**	**NAD(P)H**	**Fd_{red}**	**CO_2**						
additional turns of complete CBB cycle	-3	-2		+1	**2x**	**2x**	**3x**			
phosphoglycerate kinase	-1							**1x**		
glyceraldehyde 3-phosphate dehydrogenase		-1						**1x**		
phosphoribulokinase	-1							**1x**		
RuBisCO				+1				**1x**		
phosphoglycolate phosphatase										
glycolate oxidase										
catalase										
serine/glyoxylate AT										
glutamate/glyoxylate AT										
glutamine synthetase	-1				**1x**					
glutamine oxoglutarate AT			-2		**1x**					
serine hydroxymethyl-transferase										
glycine decarboxylase		+1		-1	**1x**					
hydroxypyruvate reductase		-1			**1x**					
glycerate kinase	-1				**1x**	**1x**				**1x**
glycolate dehydrogenase		+1				**2x**			**1x**	
glyoxylate carboligase				-1		**1x**				
tartronic semialdehyde reductase		-1				**1x**				
malate synthase										
malic enzyme		+1		-1			**1x**			
pyruvate dehydrogenase		+1		-1			**1x**			
glycolyl-CoA synthetase	-2							**1x**		**1x**
glycolyl-CoA reductase		-1						**1x**		
fructose-6-phosphate aldolase										
arabinose-5-phosphate isomerase										
MMC lyase										
mesaconyl-C1-CoA hydratase										
mesaconyl-CoA transferase										
mesaconyl-C4-CoA hydratase										
acetyl-CoA carboxylase	-1			+1					**1x**	
malonyl-CoA/tartronyl-CoA reductase		-2							**1x**	**1x**
propionyl-CoA synthase	-2	-1							**1x**	
pyruvate phosphate dikinase	-2								**1x**	
enolase										
phosphoglycerate mutase										
glycolyl-CoA carboxylase	-1			+1						**1x**
ATP					**-8**	**-7**	**-9**	**-4**	**-5**	**-4**
NAD(P)H					**-4**	**-3**	**-4**	**-2**	**-2**	**-2**
Fd_{red}					**-2**	**0**	**0**	**0**	**0**	**0**
CO_2					**+1**	**+1**	**+1**	**+1**	**+1**	**+1**

The stoichiometrically correct amounts of 2-phosphoglycolate required to regenerate one phosphoglycerate for each individual pathway were assumed. In case of AMP-forming enzymes (glycolyl-CoA synthetase, propionyl-CoA synthase and pyruvate phosphate dikinase), two ATP are taken into the calculations because of the requirement of an additional ATP for AMP regeneration by adenylate kinase. The physiological electron acceptor of glycolate dehydrogenase is not known. We assume the formation of one NAD(P)H here. NPR: Natural (canonical) photorespiration, GLC: glycerate bypass, OX: glycolate oxidation bypass, A5P: arabinose-5-phosphate bypass, 3OHP: 3-hydroxypropionate shunt, TaCo: tartronyl-CoA pathway, MMC lyase: malyl-CoA/β-methylmalyl-CoA/citramalyl-CoA lyase. Fd_{red}: reduced ferredoxin, AT: aminotransferase

Curriculum vitae

Die Seiten 159 – 160 (Curriculum vitae) enthalten persönliche Informationen. Sie sind daher nicht Bestandteil der Online-Veröffentlichung.

Acknowledgements

Die Seite 161 (Acknowledgements) enthält persönliche Informationen. Sie ist daher nicht Bestandteil der Online-Veröffentlichung.