# The design and realization of synthetic pathways for the fixation of carbon dioxide *in vitro*


Dissertation

zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

Dem Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Thomas Schwander

aus Luzern, Schweiz

Marburg/Lahn, Deutschland, 2017

Die Untersuchungen zur vorliegenden Arbeit wurden von November 2014 bis November 2017 unter der Betreuung von Herrn Dr. Tobias Jürgen Erb in Marburg am Max-Planck-Institut für terrestrische Mikrobiologie in der Abteilung „Biochemistry and Synthetic Metabolism“ durchgeführt.

Vom Fachbereich Biologie

der Philipps-Universität Marburg als Dissertation

angenommen am: 09.01.2018

Erstgutachter: Dr. Tobias Erb

Zweitgutachter: Prof. Dr. Wolfgang Buckel

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Lars-Oliver Essen

Prof. Dr. Hans-Ulrich Mösch

Prof. Dr. Johann Heider

Tag der mündlichen Prüfung: 23.01.2018

# Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „**The design and realization of synthetic pathways for the fixation of carbon dioxide *in vitro***“ 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. November 2017

Thomas Schwander

# Publications

Part of this work was published in the following articles:

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

**T. Schwander**, R. McLean, J. Zarzycki, T.J. Erb. Structural basis for substrate specificity of methylsuccinyl-CoA dehydrogenase: an unusual member of the Acyl-CoA dehydrogenase family (submitted)

R. McLean, **T. Schwander**, N.S. Cortina, T.J. Erb. HOPAC cycle, another synthetic pathway for the fixation of carbon dioxide *in vitro*. (in preparation)

Other Publications:

M. Konneke, D.M. Schuber, P.C. Brown, M. Hügler, S. Standfest, **T. Schwander**, L. Schada von Borzyskowski, T.J. Erb, D.A. Stahl, I.A. Berg. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for $CO_2$ fixation. *Proc Natl Acad Sci* **111**: 8239-8244.

**T. Schwander**, T.J. Erb. (2016). Do it your (path)way – synthetische Wege zur $CO_2$-Fixierung. BIOspektrum **22**: 590-592

S. Burgener, **T. Schwander**, T.J. Erb. Molecular basis for converting methylsuccinyl-CoA dehydrogenase into an oxidase (draft manuscript)

*"Man is a little germ that lives on an unimportant rock ball that revolves about an insignificant star on the outer edges of one of the smaller galaxies. But on the other hand if you think about that for a few minutes… I am absolutely amazed to discover myself on this rock ball rotating around a spherical fire. It's a very odd situation! And the more I look at things, I cannot get rid of the feeling that existence is quite weird."*

*"And then more so, when this so-called insignificant little creature has inside his skull a neurological contraption that is able to center itself in the midst of these incredible expanses of galaxies and start measuring the whole thing. That is quite extraordinary!"*

Alan Watts (1915 – 1973), Philosopher

# Contents

# Summary

The fixation of inorganic carbon and the conversion to organic molecules is a prerequisite for life and the foundation of the carbon cycle on Earth. Since the industrial revolution, this carbon cycle has become imbalanced and consequently the atmospheric carbon dioxide ($CO_2$) concentration is increasing and is a major cause of global warming. On the other hand, atmospheric $CO_2$ can also be considered an important carbon feedstock of the future. Human society has not yet come up with a viable solution to convert this inorganic $CO_2$ back into reduced carbon compounds and is still relying on natural $CO_2$ fixation. In contrast, nature has evolved multiple solutions to reduce $CO_2$ and incorporate it into organic molecules. The involved pathways differ in their cofactor requirements and are often limited to anoxic conditions. The most common oxygen tolerant $CO_2$ fixation pathway is the Calvin-Benson-Bassham (CBB) cycle found in plants, algae and many prokaryotes. Its central carboxylase Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) is a fairly slow catalyst, which also performs a side reaction with oxygen. This leads to an energetically wasteful process called photorespiration that results in the loss of already fixed carbon. Many attempts have been made to improve the carboxylation reaction of RubisCO directly or to redesign the photorespiration pathway to a more energy efficient process. The emerging field of synthetic biology offers an alternative approach by designing novel pathways for the fixation of $CO_2$. This approach additionally includes the design and engineering of enzymes enabling molecular conversions not found in nature. Although, several such artificial pathways have been designed, none of them have been realized so far. This reveals an existing gap between the design and the realization and implementation of such a synthetic $CO_2$ fixation pathway.

In this work we designed several synthetic oxygen-tolerant $CO_2$ fixation pathways in a bottom-up approach, by freely combining feasible enzyme reactions. The pathways were designed around an efficient central carboxylase from the family of enoyl-CoA carboxylases/reductases. Some members of this family belong to the most efficient carboxylases known so far, do not accept oxygen as a substrate and only require the ubiquitous cofactor NADPH.

The theoretical analysis of thermodynamic and energetic properties of the designed pathways for $CO_2$ fixation also showed that they are comparable or even more energy efficient than naturally occurring oxygen-tolerant $CO_2$-fixing pathways. We were able to realize two of these cycles *in vitro* and investigated their efficiencies for the fixation of inorganic $CO_2$ into organic molecules. The results narrowed the gap between theoretical design and *in vivo* transplantation. The following results were achieved:

1. First, the **C**rotonyl-CoA/**ET**hylmalonyl-**C**oA/**H**ydroxybutyryl-CoA (CETCH) cycle was established *in vitro* and its $CO_2$ fixation efficiency was increased in several rounds of optimization, including the principle of metabolic proofreading to recycle undesired side products. The current version CETCH 5.4 is a reaction network of 17 enzymes originating from nine different organisms of all three domains of life, including three rationally engineered enzymes. It is energized by ATP- and NADPH-regeneration modules and converts $CO_2$ into organic molecules at a rate of 5 nmol $CO_2$ $min^{-1}$ $mg^{-1}$ protein. A stepwise incorporation of $^{13}CO_2$ into the cycle intermediates confirmed a continuous operation of the cycle for multiple rounds of conversion. The CETCH cycle expands the solution space beyond the six naturally evolved $CO_2$ fixation pathways by a seventh, man-made alternative that is thermodynamically more efficient than the CBB cycle of plants.

2. Second, the **H**ydr**O**xy**P**ropionyl-CoA/**A**crylyl-**C**oA (HOPAC) cycle was established *in vitro* and its turnover number was increased over several rounds of optimization. The current version

HOPAC 4.1 is a reaction network of 15 enzymes originating from eight different organisms (including bacteria and archaea). We also applied rational active site engineering to improve enzymes towards desired reactions. A stepwise incorporation of $^{13}CO_2$ into the cycle intermediates confirms a continuous operation for multiple rounds of conversion. The HOPAC cycle is the second synthetic $CO_2$ fixation cycle established *in vitro*, but still requires more optimization to reach the level of the CETCH cycle. Nevertheless, it increases the available options for future applications and allows the parallel optimization of both pathways (HOPAC and CETCH) and their comparison.

3. The methylsuccinyl-CoA dehydrogenase (Mcd) from *R. sphaeroides* was a crucial component for the construction of the CETCH and the HOPAC cycle. This enzyme belongs to the family of flavin dependent acyl-CoA dehydrogenases, which catalyze the α-,β-carbon bond desaturation of acyl-CoA thioesters. The electrons from the oxidation reactions are naturally transferred to the membrane bound electron transport chain via electron transfer flavoproteins (ETF). In the CETCH cycle, we engineered the enzyme towards using molecular oxygen as a direct electron acceptor and in the HOPAC cycle the side reactivity was exploited to convert succinyl-CoA to fumaryl-CoA. In turn, to elucidate the enzyme's substrate specificity and the engineered oxidase activity, we crystallized the protein and obtained a high resolution structure. It is the first protein structure of an acyl-CoA dehydrogenase converting an α-methyl branched γ-carboxylic acid CoA thioester. The enzyme also possesses an unusual N-terminal ~170 residue extension and we were able to assign a tertiary structure and function to this domain.

In summary, this study laid the foundation for the development of artificial pathways for the fixation of $CO_2$. We demonstrated that it is possible to convert inorganic $CO_2$ into organic molecules via two synthetic pathways *in vitro*. These results narrow the gap between theoretical design of synthetic $CO_2$ fixation pathways and their application *in vivo*. Additionally, we solved the crystal structure of methylsuccinyl-CoA dehydrogenase, revealing a novel fold for a structural domain that is unique among the well described family of acyl-CoA dehydrogenases.

# Zusammenfassung

Die Fixierung von anorganischem Kohlenstoff und dessen Umwandlung in organische Moleküle bildet die Grundlage des Kohlenstoffkreislaufs und ist notwendig für das gesamte Leben auf der Erde. Seit der industriellen Revolution ist dieser Kreislauf jedoch gestört und als Konsequenz messen wir einen stetigen Anstieg der Konzentration an Kohlenstoffdioxid ($CO_2$) in der Atmosphäre. Dieser Anstieg ist unteranderem einer der Hauptgründe für die Klimaerwärmung. Jedoch kann dieses $CO_2$ auch als eine wichtige Kohlenstoffquelle der Zukunft betrachtet werden. Unsere Gesellschaft hat allerdings bis heute keinen optimalen Prozess gefunden, um dieses anorganische Molekül zurück zu gewinnen und ist weiterhin auf natürlich Prozesse angewiesen. Die Natur hat mehrere Stoffwechselwege evolviert, um $CO_2$ zu reduzieren und wieder in organische Moleküle einzubauen. Viele dieser Stoffwechselwege sind jedoch nur unter anaeroben Bedingungen funktionsfähig. Der meist verbreitete sauerstofftolerante Stoffwechselweg zur Fixierung von $CO_2$ ist der Calvin-Benson-Bassham (CBB) Zyklus, der in Pflanzen, Algen und Mikroorganismen vorkommt. Die zentrale Carboxylase Ribulose-1,5-Bisphosphat-Carboxylase/Oxygenase (RubisCO) ist jedoch ein vergleichbar langsamer Katalysator und besitzt eine intrinsische Seitenreaktion mit Sauerstoff. Diese Seitenreaktion führt zu einem energetisch ungünstigen Prozess, welcher Photorespiration genannt wird. Photorespiration führt unteranderem zu dem Verlust von fixiertem $CO_2$ und reduziert damit die Gesamteffizienz der Photosynthese. Viele Versuche wurden bereits unternommen, die RubisCO selbst oder den Prozess der Photorespiration zu verbessern. Die aufkommende Disziplin der synthetischen Biologie bietet einen neuen Ansatz, in dem komplett neue Stoffwechselwege zur Fixierung von $CO_2$ entworfen werden. Dieser Ansatz beinhaltet auch das konstruieren (*engineering)* und entwickeln von Enzymen, um neue Reaktionen zu ermöglichen. Solche künstliche Stoffwechselwege wurden schon in früheren Untersuchungen entworfen, aber keiner konnte bisher verwirklicht werden. Dies veranschaulicht eine bestehende Lücke zwischen dem Entwerfen von synthetischer $CO_2$ Fixierung und dessen tatsächlicher Umsetzen und Anwenden.

Ziel dieser Arbeit war es mehrere sauerstofftolerante und synthetische Stoffwechselwege zur $CO_2$ Fixierung zu entwickeln, indem enzymatische Reaktionen kombiniert wurden. Das zentrale Enzym dieser künstlichen Prozesse stammt aus der Familie der Enoyl-CoA-Carboxylasen/Reduktasen, da diese zu den effizientesten Carboxylasen zählen, die bisher bekannt sind. Zudem besitzen sie keine Seitenreaktion mit Sauerstoff und benötigen für die Katalyse nur den ubiquitären Cofaktor NADPH.

Die theoretische Analyse dieser neuen künstlichen Stoffwechselwege zeigte eine vergleichbare oder sogar bessere Energieeffizienz gegenüber natürlichen sauerstofftoleranten Prozessen zur $CO_2$ Fixierung. Von diesen theoretischen Konstrukten, konnten wir zwei Stoffwechselwege *in vitro* realisieren und deren $CO_2$ Fixierungseffizienz analysieren. Diese Resultate verkleinern die bestehende Lücke zwischen theoretischem Entwerfen und der Etablierung künstlicher $CO_2$ Fixierung *in vivo.* Die folgenden Resultate wurden erreicht:

1. Erstens wurde der **C**rotonyl-CoA/**ET**hylmalonyl-**C**oA/**H**ydroxybutyryl-CoA (CETCH) Zyklus *in vitro* realisiert und die $CO_2$ Fixierungseffizienz wurde in mehreren Schritten optimiert. Dies beinhaltete auch das Prinzip des *metabolic proofreading*, um Nebenprodukte zu recyceln. Die aktuelle Version CETCH 5.4 ist ein Reaktionsnetzwerk, bestehend aus 17 Enzymen von neun verschiedenen Organismen aus allen Domänen des Lebens. Zudem wurden drei Enzyme *engineered*, um bestimmte Reaktion zu katalysieren. Der Stoffwechselweg wird durch ATP- und NADPH-Regenerationssysteme angetrieben und wandelt 5 nmol $CO_2$ $min^{-1}$ $mg^{-1}$ Protein um. Die schrittweise Inkorporation von $^{13}CO_2$ in die Intermediate zeigte, dass der Zyklus sich für mehrere Runden drehte. Der CETCH Zyklus erweitert die sechs natürlich evolvierten

Lösungen zur Fixierung von $CO_2$ um eine Siebente, von Menschenhand geschaffene. Zudem ist dieser Zyklus thermodynamisch sogar energieeffizienter als der CBB Zyklus von Pflanzen.

2. Zweitens wurde der **H**ydr**O**xy**P**ropionyl-CoA/**A**crylyl-**C**oA (HOPAC) Zyklus *in vitro* realisiert und die $CO_2$ Fixierungseffizienz wurde schrittweise optimiert. Die aktuelle Version HOPAC 4.1 ist ein Reaktionsnetzwerk bestehend aus 15 Enzymen von acht verschiedenen Organismen (Bakterien und Archaeen). Unteranderem wurde für den HOPAC Zyklus das rationale Entwickeln (*engineering*) von Enzymen benutzt, um die Katalyse von gewünschten Reaktionen zu ermöglichen. Die schrittweise Inkorporation von $^{13}CO_2$ in die Intermediate bestätigte auch für diesen Zyklus, dass mehrere Reaktionsrunden aufrechterhalten werden konnten. Der HOPAC Zyklus ist der zweite künstliche Stoffwechselweg zur $CO_2$ Fixierung, der *in vitro* etabliert wurde, jedoch bedarf dieser noch weiterer Optimierungen im Vergleicht zum CETCH Zyklus. Der HOPAC Zyklus erweitert die Verfügbarkeit von möglichen Optionen für zukünftige Anwendungen und erlaubt die parallele Weiterentwicklung beider Zyklen (CETCH und HOPAC) und deren Vergleich.

3. Die Methylsuccinyl-CoA-Dehydrogenase (Mcd) von *R. sphaeroides* war eine wichtige Komponente für die Funktionsfähigkeit des CETCH und HOPAC Zyklus. Dieses Enzym gehört zur Familie der flavin abhängigen Acyl-CoA-Dehydrogenasen, welche die α-,β-Kohlenstoffbindung von Acyl-CoA-Thioestern desaturieren. Die Elektronen aus der katalysierten Oxidationsreaktion werden im Organismus via einem elektronen-transferierenden Flavoprotein (ETF) zur Elektronentransportkette transportiert. Im CETCH Zyklus wurde dieses Enzym zu einer Oxidase *engineered*, um Sauerstoff als direkten Elektronenakzeptor zu verwenden und im HOPAC Zyklus wurde die intrinsische Seitenreaktion mit Succinyl-CoA ausgenutzt. Um die Substratspezifität und die eingebaute Oxidaseaktivität besser zu verstehen, wurde dieses Protein kristallisiert und eine hochauflösende Struktur gelöst. Es ist die erste Struktur einer Acyl-CoA-Dehydrogenase, die einen α-Methyl verzweigten γ-Carbonsäure CoA Thioester akzeptiert. Dieses Enzym enthält im Gegensatz zu anderen Acyl-CoA-Dehydrogenasen eine N-terminale ~170 Aminosäure lange Erweiterung und wir konnten dieser Domäne eine tertiäre Struktur und Funktion zuordnen.

Zusammengefasst legt diese Arbeit den Grundstein für die Entwicklung von künstlichen Stoffwechselwegen zur Fixierung von $CO_2$. Wir konnten *in vitro* zeigen, dass es möglich ist anorganisches $CO_2$ mit Hilfe von zwei synthetischen Stoffwechselwegen in organische Moleküle zu überführen. Die Resultate dieser Arbeit schmälern die Lücke zwischen theoretischem Entwerfen von künstlichen Prozessen für die $CO_2$ Fixierung und deren Umsetzung *in vivo.* Zusätzlich wurde eine neue Struktur in der bereits gut erforschten Familie der Acyl-CoA-Dehydrogenasen gelöst.

# 1. Introduction

## 1.1. Carbon dioxide in the atmosphere

In recent years, atmospheric carbon dioxide ($CO_2$) has become a main focus of interest, because it is the primary anthropogenic greenhouse gas released by the human economical thrive. Burning fossil fuels releases the $CO_2$ that was sequestered by nature over several geological eras millions of years ago. Since the industrial revolution, the concentration of $CO_2$ rose from 280 ppm to 400 ppm today and the annual anthropogenic release of ~36 Gt of $CO_2$ is expected to rise further (**Fig. 1**). The accumulation of $CO_2$ in the atmosphere correlates undeniably with the experienced rise of the global temperature (*1*). Human society and economy has a very high demand for energy and fossil fuels will still be the main source for years until more ecological technologies are developed and viable. Hence, fighting global warming will be one of the major and most difficult tasks in the future.

Whereas the $CO_2$ in the atmosphere poses a threat to our society, it is also considered as an important carbon feedstock for the future. The sustainable capture and utilization of $CO_2$ will be crucial to realize a green economy that is independent from petrochemical sources. Yet, compared to the anthropogenic release of $CO_2$, our capabilities to recapture or reuse atmospheric $CO_2$ are still limited. Chemistry still struggles with developing catalysts and processes that allow the conversion of $CO_2$ under mild conditions into a multi-carbon, value-added product in a carbon-neutral, let alone carbon-beneficial, sustainable way (*2-4*). In contrast, natural autotrophic carbon fixation annually transforms gigatons of inorganic $CO_2$ into biomass. More than 90% is converted via the Calvin-Benson-Bassham (CBB) cycle by plants, algae and microorganisms. The remainder is converted through alternative $CO_2$ fixation pathways (*5, 6*). Despite the existence of diverse natural solutions for carbon fixation, the biotechnological application of these enzymes and pathways is limited and proved challenging so far (*7, 8*). Additionally, biological $CO_2$ fixation mainly delivers biomass and not a dedicated product. In turn, the obtained biomass needs further processing to yield a useful product (*e.g.* biofuels) (*9*).

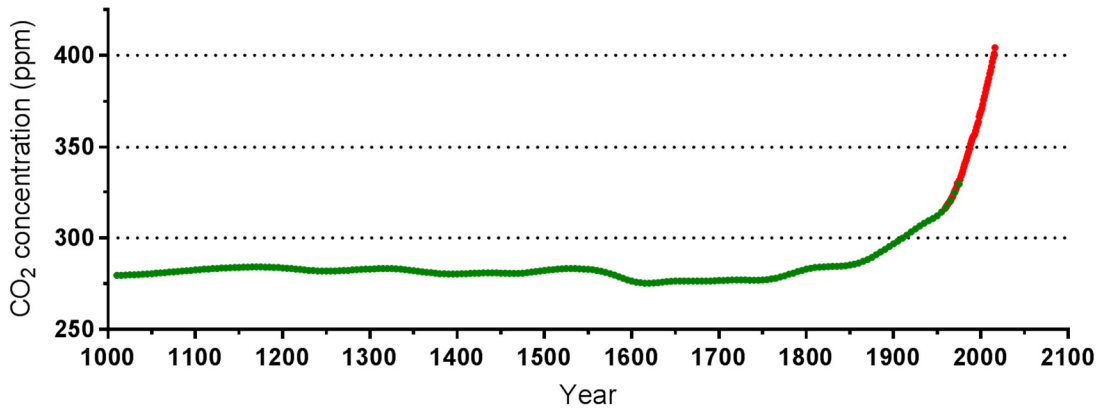



**Fig. 1 Annual mean of $CO_2$-concentration in the atmosphere.** In green are the annual mean measured in the Law Dome Ice Core (*10*) and in red are the data measured on Mauna Loa, Hawaii by the National Oceanic & Atmospheric Administration (NOAA)/ Earth System Research Laboratory (ERSL). www.esrl.noaa.gov/gmd/ccgg/trends/

## 1.2. Diversity of natural carbon fixation pathways

The fixation of inorganic carbon and the conversion to organic molecules is a prerequisite for life and a vital part in the carbon cycle on our planet. The most prominent $CO_2$ fixation pathway in nature, is the CBB cycle, which is responsible for the autotrophic growth of plants, algae, cyanobacteria and many other bacterial species. Although this pathway seems to be a late invention during evolution compared to some other $CO_2$ fixation pathways (*11, 12*), it was very successful and able to spread in nature. This pathway was first discovered in the late 1940s and is described in a series of publications by the group of Melvin Calvin (*13-15*). The ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) is the key enzyme of the CBB cycle and catalyzes the carboxylation of ribulose 1,5-bisphosphate and produces two molecules of 3-phosphoglycerate (**Fig. 2**). This intermediate is phosphorylated to 1,3-bisphosphoglycerate before it is reduced to glyceraldehyde 3-phosphate. Whereas one glyceraldehyde 3-phosphate is used for biosynthesis, five of these $C_3$ molecules are recycled via the pentose phosphate pathway to yield three $C_5$ compounds (ribulose 5-phsophate). Subsequent phosphorylation of the ribulose 5-phosphate by phosphoribulokinase (PRK) regenerates the carboxylation substrate. Consequently, three cycles are required to produce one molecule of glyceraldehyde 3-phosphate and regenerate the $CO_2$ accepting intermediate. The vast success of the CBB cycle in evolution can be partially accounted for by the fact that all its enzymes are oxygen tolerant (*5*). However, in the presence of molecular oxygen RubisCO also exhibits an oxygenase activity, which produces the toxic side product 2-phosphoglycerate (*16*). Photorespiration is a wasteful detoxifying process in organisms to prevent the accumulation of the side product and recycle it to 3-phosphoglycerate. Since this pathway involves a decarboxylation step, $CO_2$ fixation efficiency of the organism is reduced (*17, 18*).

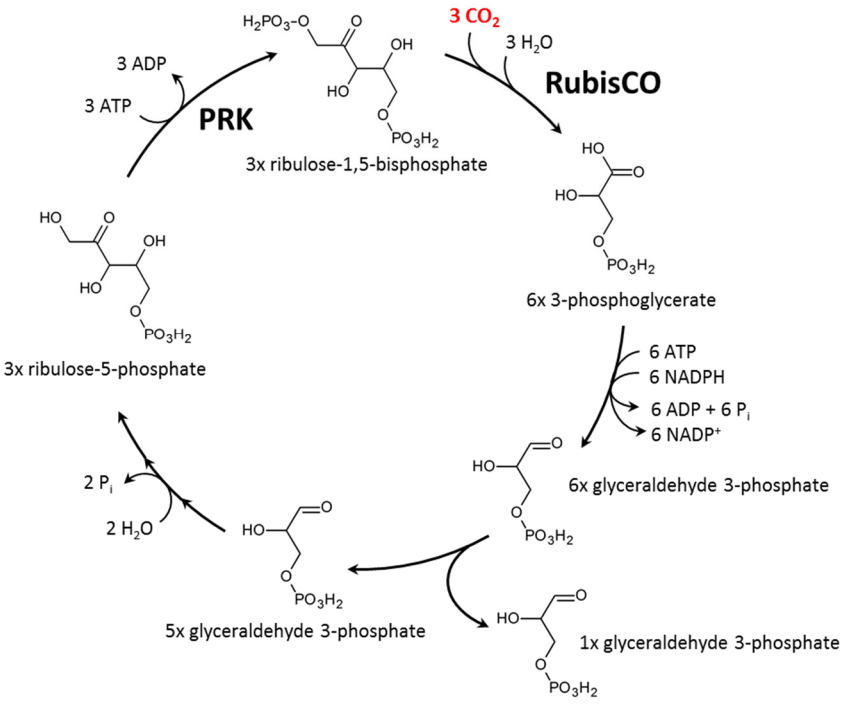



**Fig. 2 Simplified scheme of the CBB cycle.** Note, that not all individual enzymatic conversions are shown. The two key enzymes are ribulose-5-phosphate kinase (phosphoribulokinase, PRK) and ribulose-1,5,-bisphosphate carboxylase/oxygenase (RubisCO). Glyceraldehyde-3-phosphate is withdrawn and my enter gluconeogenesis or glycolysis for further conversion.

After the discovery of the CBB cycle, it was assumed to be the only carbon fixation pathway, which fit with the central biological dogma of biochemical unity of life at that time (*19*). But this dogma only stood for two decades because several alternative $CO_2$ fixation pathways have been discovered to date. These pathways are adapted to the environmental conditions given by a certain habitat to allow efficient fixation of inorganic $CO_2$ (*5*).

Already in the 1960s, the reductive citric acid (rTCA) cycle was discovered as an alternative $CO_2$ fixation pathway (*20*). The name already implies, that the pathway reverses the reactions of the oxidative citric acid cycle (Krebs cycle) and forms acetyl-coenzyme A (CoA) from two molecules of $CO_2$ (**Fig. 3**). Subsequently, acetyl-CoA is reductively carboxylated to pyruvate by a ferredoxin-dependent pyruvate synthase. Three enzymes are characteristic for this cycle and enable reversed flux through the pathway. Succinate dehydrogenase is replaced by fumarate reductase, the nicotinamide dinucleotide ($NAD^+$)-dependent 2-oxoglutarate dehydrogenase is replaced by 2-oxoglutarate:ferredoxin oxidoreductase (*21*) and the citrate synthase is replaced by adenosine triphosphate (ATP) citrate lyase (*22*). These enzymes support the reversion of the Krebs cycle, but the overall conversion from succinate to acetyl-CoA is still endergonic (ΔG°´ +27 kJ) (*6*). Although, the rTCA cycle is quite energy efficient, it needs reduced ferredoxin to drive the system thermodynamically. Conclusively, the pathway is restricted to anaerobic organisms.

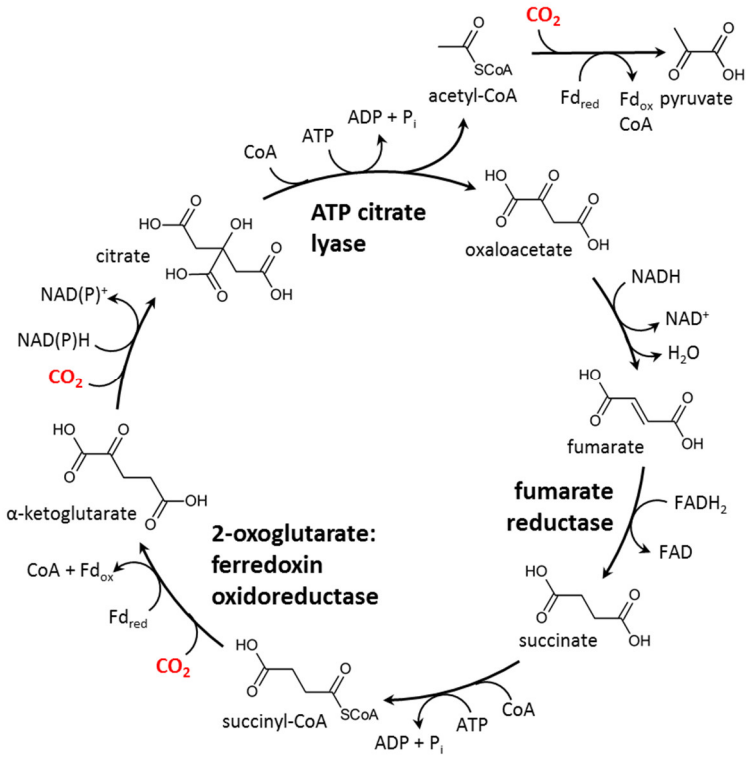



**Fig. 3 Simplified scheme of the rTCA cycle adapted from (*6*).** Note, that not all individual enzymatic conversions are shown. The key enzymes are ATP-citrate lyase, 2-oxoglutarate:ferredoxin oxidoreductase and fumarate reductase. The carboxylation reactions are catalyzed by isocitrate dehydrogenase, 2-oxoglutarate:ferredoxin oxidoreductase and pyruvate:ferredoxin oxidoreductase (pyruvate synthase).

Moreover, the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) is found in acetogenic bacteria and methanogenic archaea, which live close to the thermodynamic limit, and requires the least input of ATP among all known $CO_2$-fixing pathways. It is not only used for the fixation of $CO_2$, but also for the conservation of energy (*23-25*). The genomic reconstruction of LUCA (last universal common ancestor) predicts that this pathway already operated at the beginning of life (*26*). In this pathway, acetyl-CoA is produced as the primary assimilation product from two molecules of $CO_2$ (**Fig.**

**4**). The first molecule of $CO_2$ is reduced to carbon monoxide and binds to the nickel containing Fe-S cluster of the key enzyme, CO dehydrogenase/acetyl-CoA synthase. In acetogens, the second $CO_2$ molecule is reduced to formate and coupled to the cofactor tetrahydrofolate ($H_4F$). Subsequently, the formyl-$H_4F$ is further reduced to methyl-$H_4F$ and the methyl group is also transferred to the CO dehydrogenase/acetyl-CoA synthase, and acetyl-CoA is formed (*27*). Subsequently, acetyl-CoA can be carboxylated by pyruvate synthase to form pyruvate. Despite its high energy efficiency, this pathway is only functional under strictly anaerobic conditions, as it utilizes highly oxygen-sensitive metal co-factors and reduced ferredoxin. Therefore, it is only found in a limited number of anaerobic organisms.

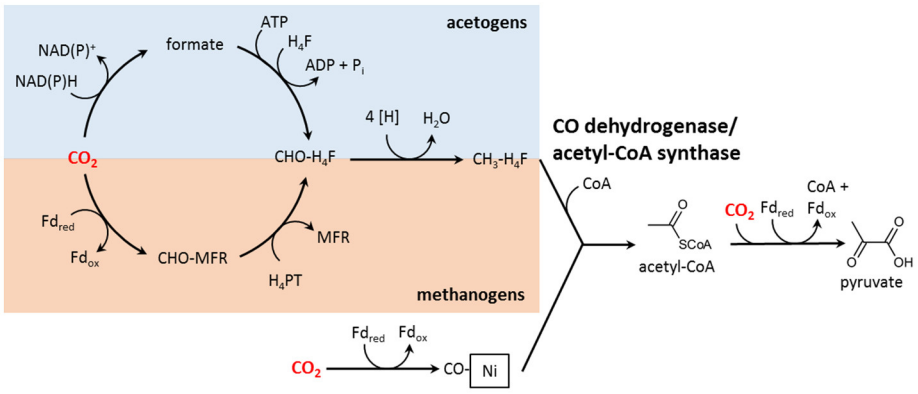



**Fig. 4 Simplified scheme of the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) adapted from (*6*).** One $CO_2$ is reduced to a methyl group bound on a carrier molecule. This reaction sequence differs in acetogens and methanogens (indicated by a blue or red box), because different reductants and coenzymes are applied. A second $CO_2$ molecule is reduced to a nickel bound carbon monoxide in the active site of CO dehydrogenase/acetyl-CoA-synthase. This enzyme also combines the reduced methyl group with CO to form acetyl-CoA. Acetyl-CoA can be further reductively carboxylated to pyruvate. The $CO_2$ fixing enzymes are formate dehydrogenase, formyl-methanofuran dehydrogenase, CO dehydrogenase/acetyl-CoA-synthase, pyruvate:ferredoxin oxidoreductase. MFR: methanofuran; $H_4F$: tetrahydrofolate; $H_4PT$: tetrahydropterin ($H_4F$ is present in acetogens and $H_4PT$ in methanogens).

The following three $CO_2$ fixation pathways have been elucidated in the laboratory of Georg Fuchs and are often termed Fuchs cycles I, II and III.

The first hints towards the 3-hydroxypropionate (3HP) bi-cycle were already published in the late 1980s (*28*). Yet, it took another 20 years to elucidate the full reaction sequence of the pathway. Two metabolic cycles are involved for the fixation of $CO_2$ (**Fig. 5**). In the first cycle acetyl-CoA is carboxylated to malonyl-CoA by acetyl-CoA carboxylase. The reduction of malonyl-CoA to propionyl-CoA is catalyzed by a bi-functional malonyl-CoA reductase (*29*) and a tri-functional propionyl-CoA synthase (*30*). Subsequently, propionyl-CoA is carboxylated to methylmalonyl-CoA by propionyl-CoA carboxylase (Pcc) and isomerized to succinyl-CoA. The oxidation of succinyl-CoA to malyl-CoA proceeds via the free carboxylic acids. A CoA transfer from succinyl-CoA to malate by succinyl-CoA:(*S*)-malate-CoA transferase leads to the formation (*S*)-malyl-CoA. The latter is then cleaved into acetyl-CoA and glyoxylate. In the second cycle, acetyl-CoA is again carboxylated followed by reduction to propionyl-CoA (identical to the first cycle), however, propionyl-CoA is now condensed with glyoxylate (the product of the first cycle) to form β-methylmalyl-CoA. After a series of $C_5$-transformations, (*S*)-citramalyl-CoA is formed and subsequently cleaved into acetyl-CoA and pyruvate. All condensation and cleavage reactions are catalyzed by one tri-functional malyl-CoA/β-methylmalyl-CoA/citramalyl-CoA lyase (*31-33*). Interestingly the whole pathway involves 19 reactions, but requires only 13 enzymes to catalyze them, none of which are oxygen sensitive. Considering also the more efficient carboxylases

than RubisCO, it is interesting, that this pathway was less successful during evolution and appears to be restricted to *Chloroflexaceae*.

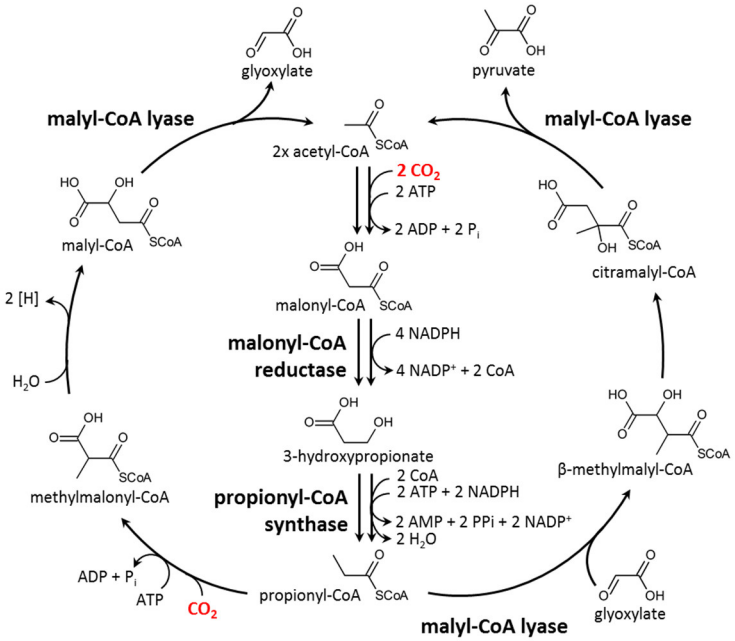



**Fig. 5 Simplified scheme of the 3HP bi-cycle adapted from (*6*).** Note, that not all individual enzymatic conversions are shown. The key enzymes are the bifunctional malonyl-CoA reductase, trifunctional propionyl-CoA synthase and the malyl-CoA/β-methylmalyl-CoA/citramalyl-CoA lyase. The carboxylation reactions are catalyzed by acetyl-CoA carboxylase and propionyl-CoA carboxylase. In the first cycle (left) one molecule of glyoxylate is produced, which is used in the second cycle (right) to produce pyruvate.

The 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle for $CO_2$ fixation operates in autotrophic thermoacidophilic *Crenarchaeota* and mesophilic *Thaumarchaeota (5, 34, 35)*. The first part of the cycle is identical to the 3HP bi-cycle and converts acetyl-CoA with two carboxylation steps to succinyl-CoA via propionyl-CoA (**Fig. 6**). However, the key enzymes for the reduction of malonyl-CoA to propionyl-CoA are not homologous in archaea and *Chloroflexaceae*, suggesting an independent and convergent evolution of the two pathways (*33, 35, 36*). The second part of the cycle converts succinyl-CoA to 4-hydroxybutyryl-CoA via 4-hydroxybutyrate. The key enzyme 4-hydroxybutyryl-CoA dehydratase catalyzes the dehydration to crotonyl-CoA, which is eventually oxidized to acetoacetyl-CoA. Subsequent cleavage of acetoacetyl-CoA yields two molecules of acetyl-CoA, closing the cycle. All enzymes are oxygen tolerant and the pathway can be found in aerobic as well as anaerobic organisms. Although the key enzyme (4-hydroxybutyryl-CoA dehydratase), is a radical enzyme and homologs in clostridia are susceptible to inactivation by molecular oxygen, the archaeal enzymes seem to be sufficiently oxygen insensitive (*34*).

The last natural $CO_2$ fixation cycle discovered is the dicarboxylate/4-hydroxybutyrate (DC/4HB) cycle and operates in archaea as well (*37*). Similar to the pathways described before, acetyl-CoA is also carboxylated twice to yield succinyl-CoA (**Fig. 6**). However, a different reaction sequence with two different carboxylases is used. Unlike before, acetyl-CoA is first reductively carboxylated by pyruvate synthase, to produce pyruvate. Subsequently, pyruvate is activated to phosphoenol pyruvate (PEP) and further carboxylated to oxaloacetate by PEP carboxylase. After the conversion to succinyl-CoA, two molecules of acetyl-CoA are produced via 4-hydroxybutyrate as described for the 3HP/4HB cycle. Since

the enzymes and electron carriers of the DC/4HB cycle are susceptible to oxygen, this pathway is restricted to anaerobic archaea.

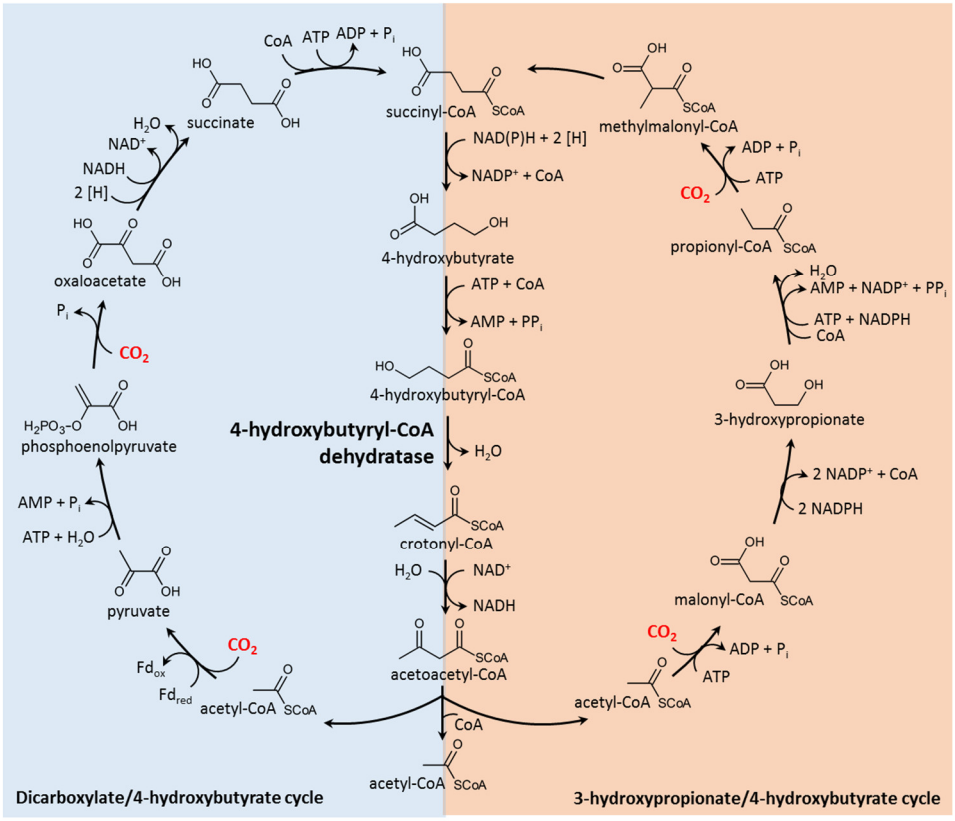



**Fig. 6 Simplified scheme of the 3HP/4HB cycle (red) and the DC/4HB cycle (blue) adapted from (*6*).** Note, that not all individual enzymatic conversions are shown. The key enzyme is the 4-hydroxybutyryl-CoA dehydratase. While the 3HP/4HB cycle depends on biotin dependent acyl-CoA carboxylases, the DC/4HB cycle catalyzes the carboxylation reactions via pyruvate:ferredoxin oxidoreductase and the phosphoenolpyruvate carboxylase. Both cycles convert succinyl-CoA to two acetyl-CoA via 4-hydroxybutyrate.

Conclusively, many different strategies to fix $CO_2$ for autotrophic growth have evolved in nature, of which six have been discovered so far. Each pathway evolved in a specific environment, thus representing an adaptation to a certain ecological niche. The occurrence of a particular pathway in an organism is highly dependent on the physiological conditions, such as energy requirements and type of reducing equivalents used for the reduction of the $CO_2$. Further determinants are their requirements for special cofactors, metals and especially their tolerance towards molecular oxygen.

**Table 1 Pathways for autotrophic carbon fixation**, adapted from (*36*)

| Pathway | ATP equivalents for synthesis of one pyruvate | Reductants for synthesis of on pyruvate (10 [H]) | $CO_2$-fixing enzymes | Key enzymes | Oxygen tolerant |
|---|---|---|---|---|---|
| **Calvin-Benson-Bassham cycle** | 7 | 5 NAD(P)H | RubisCO ($CO_2$) | RubisCO, phosphoribulo-kinase | Yes |
| **Reductive citric acid cycle** | 2-3 | 2 NAD(P)H, 1 NADH or an unknown donor of fumarate reductase, 2 ferredoxin | 2-Oxoglutarate synthase ($CO_2$), isocitrate dehydrogenase ($CO_2$) or 2-oxoglutarate carboxylase ($HCO_3^-$), pyruvate synthase ($CO_2$) | 2-Oxoglutarate synthase, ATP citrate lyase | No |
| **Reductive acetyl-CoA pathway** | ~1 | NAD(P)H, Ferredoxin, F420-$H_2$ in Euryarchaeota | CO dehydrogenase/Acetyl-CoA synthase ($CO_2$), formate dehydrogenase ($CO_2$), Pyruvate synthase ($CO_2$) | CO dehydrogenase/ Acetyl-CoA synthase and enzymes reducing $CO_2$ to methyl-$H_4$F | No |
| **3-Hydroxypropionate bi-cycle** | 7 | NAD(P)H, but 1 quinone is reduced by succinate dehydrogenase | Acetyl-CoA and propionyl-CoA carboxylase ($HCO_3^-$) | Malonyl-CoA reductase, propionyl-CoA synthase, malyl-CoA lyase | Yes |
| **3-Hydroxypropionate/4-hydroxybutyrate cycle** | 9 | NAD(P)H, but 1 quinone is reduced by succinate dehydrogenase | Acetyl-CoA and propionyl-CoA carboxylase ($HCO_3^-$) | Acetyl-CoA/propionyl-CoA carboxylase, enzymes reducing malonyl-CoA to propionyl-CoA, 4-hydroxybutyryl-CoA dehydratase | Yes |
| **Dicarboxylate/4-hydroxybutyrate cycle** | 5 | 2-3 ferredoxin 1-2 NAD(P)H, 1 unknown donor of fumarate reductase | Pyruvate synthase ($CO_2$), PEP carboxylase ($HCO_3^-$) | 4-Hydroxybutyryl-CoA dehydratase | No |

## 1.3. Carboxylases in nature

Carboxylases are the key enzymes of all $CO_2$-fixing pathways and are among the most important enzymes in the biosphere. Since they catalyze the incorporation of inorganic $CO_2$ into organic molecules, they are crucial for the carbon cycle and supply all life on earth with reduced carbon compounds. Although, all carboxylases transfer a $CO_2$ molecule to an organic substrate, they differ in the underlying mechanism, as well as co-substrates and cofactors used (*38*). Carboxylases are not only function in autotropic pathways, where the carbon fixation reaction delivers the initial precursor for all biomass formed in a cell (as discussed above); they serve a broad range of physiologically relevant functions (*39*).

(i) Carboxylases can be found in assimilatory pathways, where they are important for the functionalization of inert substrates such as phenol, acetophenone, epoxypropane or acetone (*40-42*) as an alternative to oxygen-dependent activation in anaerobic organisms. Additionally, by-fixation of $CO_2$ can occur in assimilatory pathways, where substrates are converted to central metabolites by carboxylation. For example, the shortest odd-numbered acyl-CoA ester, propionyl-CoA ($C_3$), cannot be metabolized by classical β-oxidation and is therefore carboxylated to (2*S*)-methylmalonyl-CoA for assimilation via succinyl-CoA ($C_4$) (*43*). Furthermore, the assimilation of acetate, even-numbered fatty acids, alkanes, alcohols, esters and the bacterial storage polymer polyhydroxybutyrate (PHA) occurs with an initial conversion to acetyl-CoA. In turn, acetyl-CoA ($C_2$) is transformed into metabolites of greater complexity ($C_3$ and $C_4$) via carboxylation reactions. Anaerobic organisms are able to directly carboxylate acetyl-CoA to pyruvate via pyruvate:ferredoxin oxidoreductase, the same enzyme found in the rTCA cycle, DC/4HB cycle and reductive acetyl-CoA pathway (*44*). On the other hand, aerobic organisms are known to assimilate acetyl-CoA via the glyoxylate cycle, methylaspartate pathway and the ethylmalonyl-CoA pathway (*45-47*). Interestingly, the methylaspartate pathway comprises the carboxylation of propionyl-CoA, but a preceding decarboxylation step results in an overall neutral carbon yield. In contrast, the ethylmalonyl-CoA pathway applies two carboxylation reactions to convert three molecules of acetyl-CoA into two $C_4$ molecules (succinyl-CoA and malate), resulting in a positive carbon yield within this pathway. Note that one of these carboxylation reactions is performed by a nicotinamide dinucleotide phosphate (NADPH) dependent enoyl-CoA carboxylase, crotonyl-CoA carboxylase/reductase (Ccr).

(ii) Carboxylases can also be found in other anaplerotic pathways, which are essential to maintain the function of central metabolism. The most important anaplerotic reaction is the replenishing of the TCA cycle intermediate, oxaloacetate, from pyruvate (*48*). The conversion involves either the carboxylation by PEP carboxylase or the biotin-dependent pyruvate carboxylase. Notably, the ethylmalonyl-CoA and methylaspartate pathway are also anaplerotic pathways and there is a functional overlap between assimilatory and anaplerotic carboxylases in some organisms.

(iii) In biosynthetic pathways, specific building blocks for the cells are generated. Carboxylases are involved in the essential biosynthesis of fatty acids, by providing activated acyl-CoA esters (*49*). Acetyl-CoA is carboxylated to malonyl-CoA by a biotin-dependent acetyl-CoA carboxylase and subsequent decarboxylation yields a reactive enolate that acts as a nucleophile for the extension of the growing carbon chain in a Claisen condensation-like reaction. Furthermore, carboxylases are providing substrates for the phylogenetically related polyketide synthases. The most common extender units methylmalonyl-CoA and ethylmalonyl-CoA are formed by a biotin-dependent Pcc or NADPH-dependent Ccr (*50*). Homologues of Ccr form the class of carboxylating enoyl-thioester reductase (ECR) are known to accept a broad range of substrates, providing a diverse spectrum of polyketide extender units (*51, 52*).

(iv) Finally, carboxylases are important for balancing the redox state in a cell by using $CO_2$ as electron acceptor. For example, the CBB cycle with RubisCO is absolutely essential in purple non-sulfur bacteria during anaerobic photoheterotrophic growth and acts as electron sink (*53, 54*). But not only the CBB cycle has been found to be important in redox balancing, also the ethylmalonyl-CoA pathway that uses reductive carboxylation serves as an electron sink in *R. sphaeroides* under photoheterotrophic conditions (*55*).

Conclusively, nature evolved many ways to catalyze the fixation of $CO_2$ into organic matter not only for autotrophic growth, but for many diverse physiologically important processes. $CO_2$ is readily available in the atmosphere as gas or dissolved in water as bicarbonate ($HCO_3^-$) and is effectively used as simple $C_1$ extension in biosynthetic pathways. While nature has found many mechanisms to reduce and incorporate $CO_2$ under mild conditions, our technology and chemistry is still lacking behind in making efficient use of this molecule. The increasing concentration of carbon dioxide in the atmosphere requires new methods to cope with this greenhouse gas and nature provides a multitude of solutions for this problem. Understanding, applying and improving naturally evolved $CO_2$ fixation in organic chemistry, biotechnology and synthetic biology is therefore an important scientific task.

## 1.4. Improving natural carbon fixation

In the past, several possibilities to improve and apply biological $CO_2$ fixation mechanisms have been explored. Initial attempts focused on the key enzyme, RubisCO, since the $CO_2$ fixation rate of the CBB cycle is restricted by low activity of the carboxylase (1 –10 $s^{-1}$) (*56*). In a cell, sufficient flux is achieved by high expression levels of RubisCO, which can account for up to 50% of soluble protein (*57*). Moreover, RubisCO accepts molecular oxygen as substrate, which leads to photorespiration and in turn to the loss of ammonia ($NH_3$) and carbon dioxide (see above). This oxygenation reaction of the enzyme occurs under ambient conditions approximately every fifth turn over. Therefore, RubisCO was subjected to directed evolution or engineering towards higher reaction rates with increased $CO_2/O_2$ specificity (*58-60*). These attempts were met with limited success, because it became apparent that there is an inherent trade-off between activity and specificity. RubisCOs with a higher turnover rate show even less specificity towards $CO_2$, resulting in a higher oxygenation rate (*61, 62*). In nature, certain organisms evolved different $CO_2$-concentration mechanisms (CCM) to optimize the operational conditions for RubisCO and decrease the oxygenation side reaction (*63*). Accordingly, different attempts focused on these CCMs, by either mimicking the C4-plants mechanism for $CO_2$ concentration (*64*) or expressing carboxysome proteins for the generation of a heterologous $CO_2$-fixing organelle (*65*). Further strategies aim to decrease the deleterious side effects of the oxygenation reaction, by improving/replacing the photorespiration pathway (*56*). A recently published study even proposed a synthetic photorespiration route based on the 3HP bi-cycle. Compared to the canonical pathway, this synthetic solution prevents not only the loss of ammonia, but results in the additional fixation of $CO_2$ via different carboxylases (*66*). Another attempt involves the transplantation of natural $CO_2$ fixation pathways into heterologous host organisms. Nevertheless, the introduction of the 3HP bi-cycle into *E. coli* (*67*) or the introduction of the natural 3HP/4HB cycle into *P. furiosus (68)* were less successful. The most successful transplantation so far involves the heterologous operation of the CBB cycle in *E. coli*, which required the introduction of three genes. (*69*). This genetically modified organsim was able to drive the heterologous CBB cycle with reducing equivalents and ATP derived from the oxidation of pyruvate in the TCA cycle. In this strain, a substantial amount of biomass is still produced from pyruvate, but sugar phosphates are exclusively generated from $CO_2$ via the implemented CBB cycle. Initially, no growth of the so-called hemi-autotrophic organism could be observed and it was decided to evolve the strain under constant selective pressure. Xylose enabled the continuous operation of the heterologous CBB cycle and the population evolved towards growth on pyruvate and $CO_2$ after about

50 days. Interestingly, genome sequencing revealed mutations of enzymes at branching points of the CBB cycle and the authors concluded that fine-tuning of the carbon flux between the carbon fixation cycle and biosynthetic pathways was crucial. They also determined that almost 35% of cellular carbon originated from carbon dioxide. Despite this semi-autotrophic growth and the fact that the organism applies the "inefficient" CBB cycle, this study is a major step towards transplanting $CO_2$ fixation pathways into heterotrophic organisms. This study also lays a foundation for future attempts to introduce more efficient or artificial $CO_2$ fixation pathways into heterologous hosts.

In contrast to these efforts, the emerging field of synthetic biology provides an even more radical approach, by creating completely artificial $CO_2$-fixing pathways. Since nature offers many different solutions for the fixation of inorganic $CO_2$, completely novel routes can be designed around efficient carboxylases. Previous efforts have been made to understand and describe the thermodynamics and kinetics of central carbon metabolism and also the stability of autocatalytic cycles (*70-72*). These findings provide the foundation for the design of novel pathways. The free combination of different enzymatic reactions, including a more efficient carboxylase than RubisCO, leads to kinetically and thermodynamically favored pathways. Several such synthetic routes for $CO_2$ fixation have already been theoretically and computationally considered, by using the entire repertoire of approximately 5000 metabolic enzymes described so far (*73*). In this study, the authors followed a constraint-based modeling approach to design thermodynamically feasible pathways based on the efficient carbon fixation enzyme PEP carboxylase. These designed synthetic pathways were calculated to be approximately 2-3 times faster than the CBB cycle. It is a very intriguing idea to design artificial $CO_2$-fixing pathways with superior thermodynamics and kinetics, compared to the naturally evolved solutions. Nevertheless, the strong focus on the computational analysis and the *in silico* design was not followed up by the assembly of a synthetic carbon fixation route *in vitro* or *in vivo* so far.

We decided to take a slightly different approach to design artificial $CO_2$ fixation pathways. We did not restrict the design on naturally occurring reactions, but more on biochemically feasible conversions (*i.e.* falling into one of the standard E.C. reaction classes) and reactions, which require simple cofactors, such as CoA, NAD(P)H, ATP, or flavin adenosine dinucleotide (FAD). Conclusively, this design principle includes not yet discovered or unprecedented reactions and permits a more extended solution space. Nevertheless, these yet unknown reactions could be achieved by enzyme engineering and evolution. Comparable to the previously described approach, we designed our pathways around an efficient central carboxylase. In this case, we decided to base our synthetic $CO_2$ fixation pathway on the carboxylation reaction catalyzed by CoA-dependent carboxylases and in particular enoyl-CoA carboxylase/reductases (ECRs), since they show favorable catalytic properties in comparison to other carboxylases.

As aforementioned, none of such designed artificial $CO_2$ fixation pathways have been realized so far, emphasizing an existing gap between theoretical design and experimental realization in synthetic biology. A common problem is the attempt to assemble a synthetic pathway directly into a living organism. Often, there is a lack of understanding the complex interplay of the artificial pathway with the host organism that leads to undesired effects, such as side-reactions and toxicity. Therefore, the realization of synthetic pathways requires a different strategy, which allows to understand, test and optimize it in a cell free environment (*74-76*). A prime example is the research field of synthetic biochemistry, which focuses on biologically produced compounds in cell free systems via synthetic pathways (*77, 78*). This approach offers the advantage that only the production of the desired compound is important without considering the survivability of an organism and could lead to higher rates and yields than in production strains (*77, 78*). Yet, the stability of such a synthetic system is highly dependent on the regeneration and balancing of energy and redox equivalents (*79, 80*). These achievements prove that the *in vitro* approach is a viable platform for the development of synthetic

metabolic pathways. Therefore, we aim to narrow the existing gap in the field of synthetic $CO_2$ fixation, by first realizing artificial pathways *in vitro.* This allows the prior testing, optimization and investigation before an *in vivo* application is attempted.

## 1.5. Aim of this study

A general aim of this study was to design efficient artificial pathways for the fixation of $CO_2$ based on the highly efficient carboxylation reaction of enoyl-CoA carboxylase/reductase (*52*). Following the approach of synthetic biochemistry, we focused on the experimental proof of these artificial pathways *in vitro*.

### 1.5.1. Design of synthetic CO2 fixation

In the beginning several possible pathways were designed by using a principle called metabolic retrosynthesis. After considering different carboxylases, our designs were constructed around the highly efficient carboxylation reaction performed by the class of ECR. Given a certain ECR carboxylation reaction as starting point, molecular conversions from the standard reaction classes defined by the Enzyme Commission (including other carboxylation reactions) were used to regenerate the starting substrate. Notably, the designs were not restricted to known enzymes, but rather to reactions that are biochemically feasible. Additionally, the fixed $CO_2$ was channeled into a primary assimilation product via a dedicated output reaction. With this design principle, several prospective cycles were obtained and further analyzed for thermodynamic and energetic properties. The most promising candidates were selected for their realization, testing and optimization in a defined cell free environment.

### 1.5.2. Realization of synthetic $CO_2$ fixation

Furthermore, this study focused not only on the theoretical design, but approached the realization and optimization of the artificial $CO_2$ fixation cycles in a reductionist approach by assembling the artificial pathway from its principle components *in vitro*. Two promising candidates from the design phase, the **C**rotonyl-CoA/**ET**hylmalyonyl-**C**oA/**H**ydroxybutyryl-CoA (CETCH) cycle and the **H**ydr**o**xy**p**ropionyl-CoA/**A**crylyl-**C**oA (HOPAC) cycle, were realized in a cell free environment with purified enzymes. Both pathways were improved stepwise in their $CO_2$ fixation efficiency within the boundaries of the *in vitro* system.

# 2. Theoretical design and analysis of synthetic $CO_2$ fixation cycles

## 2.1. Metabolic retrosynthesis

A known bottleneck in natural $CO_2$ fixation is the efficiency of the carboxylating enzyme in a given pathway (*34, 81, 82*). To identify a suitable carboxylase as the core reaction for the synthetic pathways, we first compared the different biochemical and kinetic properties of all known major classes of carboxylating enzymes (*83*) (**Table 2** and **Fig. 7**).

**Table 2 Comparison of different carboxylase classes**.

| Carboxylase class | Oxygen sensitivity | Catalytic efficiency $k_{cat}$ $K_M^{-1}$ ($M^{-1}$ $s^{-1}$) [a] | $v_{max}$ (U $mg^{-1}$) |
|---|---|---|---|
| **Enoyl-CoA reductases/ carboxylases (ECRs)** | No | $7.0 \pm 5.1 \cdot 10^5$ [b],[d] | 130 [e] |
| **Acetyl-CoA carboxylase** | No | $4.9 \pm 3.7 \cdot 10^5$ [c] | 18 [e] |
| **RubisCO** | No, but side reaction with oxygen | $2.9 \pm 1.1 \cdot 10^5$ [c] | 3.5 [e] |
| **Propionyl-CoA carboxylase** | No | $1.6 \pm 1.6 \cdot 10^5$ [c] | 30 [e] |
| **PEP carboxykinase** | No | $1.2 \pm 1.1 \cdot 10^5$ [c] | 39 [e] |
| **2-oxoglutarate carboxylase** | No | $0.6 \pm 0.7 \cdot 10^5$ [c] | 15 [e] |
| **Pyruvate carboxylase** | No | $0.4 \pm 0.4 \cdot 10^5$ [c] | 30 [e] |
| **Methylcrotonyl-CoA carboxylase** | No | $0.3 \pm 0.3 \cdot 10^5$ [c] | 6 [e] |
| **2-Keto acid synthases** | Yes | -- | -- |

[a] $k_{cat}/K_M$–values are given in respect to the carboxylation substrate of each carboxylase, for individual data, see **Fig. 7** below. [b] from Peter et al. 2015 (*52*) and (*84*). [c] from the BRENDA database Chang et al. 2015 (*85*); http://www.brenda-enzymes.org/). Taken into account were only entries that provided $k_{cat}$ and $K_M$ measurements for the same enzyme in the same study. [d] Note that ECRs include enzymes from primary and secondary metabolism. Only enzymes from primary metabolism were evaluated. [e] from Bar-Even A. et al. 2010 (*73*). Specific activity are given for saturating $CO_2/HCO_3^-$ concentrations.

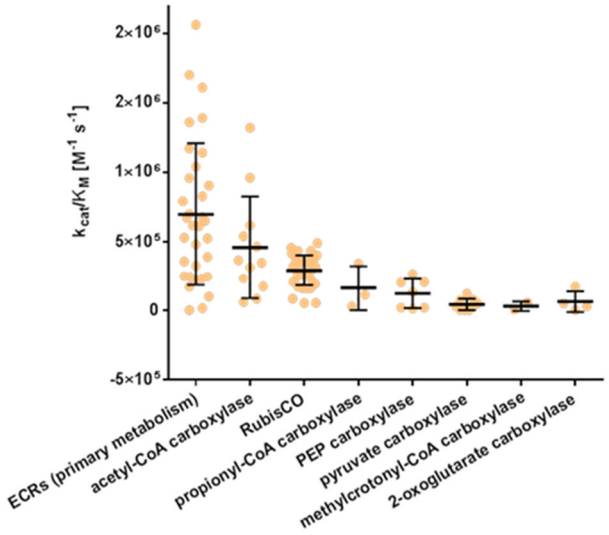



**Fig. 7 Catalytic efficiencies of individual enzymes within a given class of carboxylases different classes of carboxylases**. Each dot represents the parameter of an individual enzyme according to the sources used in **Table 2**.

From this analysis, we decided to design our synthetic $CO_2$ fixation pathways around the carboxylation reaction catalyzed by CoA-dependent carboxylases and in particular enoyl-CoA carboxylase/reductases (ECRs), since they show favorable catalytic properties in comparison to other carboxylases. ECRs are a recently discovered class of carboxylases, which are important in primary central metabolism but can also be found in secondary metabolism (*47*). Notably, this class of enzymes has not been discovered in any pathway for autotrophic growth so far. Compared to RubisCO and other carboxylases, ECRs span a broad substrate spectrum (*52*), which enables the design of multiple synthetic pathways starting from different carboxylation substrates. This increases the possibilities to design an efficient $CO_2$ fixation cycle. Additionally, these enzymes are preferable, because they are oxygen-insensitive, do not accept oxygen as substrate, require only the ubiquitous cofactor NADPH for the reduction, and catalyze the fixation of $CO_2$ with high catalytic efficiency (**Table 2** and **Fig. 7**) (*38*).

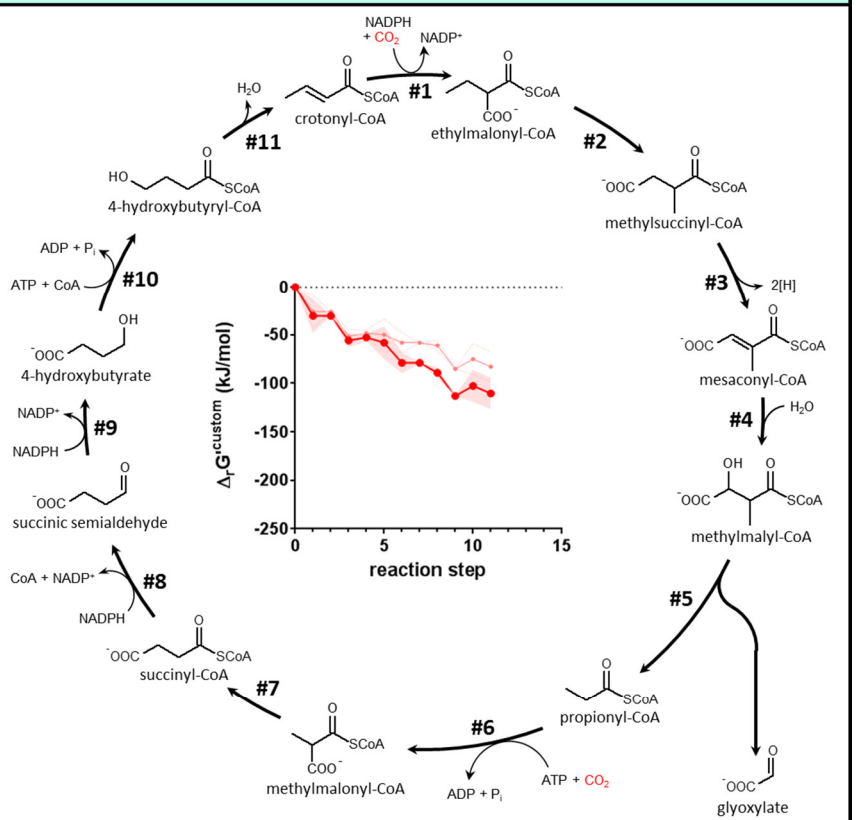
According to the initial analysis of the central carboxylation reaction, we designed several theoretical $CO_2$ fixation pathways, in a process we termed metabolic retrosynthesis. Comparable to chemical retrosynthesis, where a reaction sequence for the formation of a product is planned beforehand, we assembled a theoretical metabolic pathway from possible enzymatic catalyzed reactions applying the following constrains. Starting from a given central ECR reaction, the initial substrate is regenerated in the pathway to allow continuous cycling. In contrast to earlier approaches, we did not restrict our design to known enzymes (*73, 86*), but rather considered all reactions that seemed biochemically feasible (*i.e.* fall into one of the standard E.C. reaction classes) and required simple cofactors, such as CoA, NAD(P)H, ATP, or flavin adenosine dinucleotide (FAD). Additionally, the fixed $CO_2$ is channeled into a product via a dedicated output reaction. Following this metabolic retrosynthesis approach, we obtained several prospective cycles, of which the **C**rotonyl-CoA/**ET**hylmalonyl-**C**oA/ 4-**H**ydroxybutyryl-CoA (**CETCH**), **H**ydr**O**xy**P**ropionyl-CoA/**A**crylyl-**C**oA (**HOPAC**), **C**rotonyl-CoA/ **HY**droxyethylmalonyl-CoA/**ME**thymalonyl-CoA (**CHYME**), **CR**otonyl-CoA/**ACE**toacetate (**CRACE**), **FU**maryl-CoA/**ME**thylmalonyl-CoA/**S**uccinyl-CoA (**FUMES**), **H**ydroxycrotonyl-CoA/**IT**aconyl-CoA/ **ME**thylmalonyl-CoA (**HITME**) and **HY**droxycrotonyl-CoA/**P**ropionyl-CoA/**A**cetyl-CoA (**HYPA**) cycle were considered further (**Fig. 8**).

## 2.2. Analysis of thermodynamic and enzymatic properties

Next, we examined the theoretical drafts for certain properties. First, we evaluated the thermodynamic feasibility of these theoretical cycles, by calculating the Gibbs free energy profiles under standard and biologically relevant conditions (**Fig. 8** and chapter 8.1). This verified that the reaction sequence is driven in the direction of $CO_2$ fixation and no endergonic barrier is incorporated in the system. Although, some reaction in the pathways require the input of energy, they are coupled to exergonic steps to preserve the thermodynamic driving force. Another important factor is the availability of enzymes to catalyze the intendent reactions. Therefore, we searched bioinformatic databases for enzymes that are able sustain our drafts. For the CETCH and the HOPAC cycle we were able to identify potential enzyme candidates for each reaction. With the approach of synthetic biology we can choose from the vast diversity of enzymes in nature and considering further characterization of homologs, even more suitable enzymes can be obtained for the operation of the pathway. On the contrary, for some of the designed $CO_2$ fixation cycles we could not identify a candidate for each reaction. Nevertheless, the field of enzyme design and engineering offers great potential to fill these gaps, by obtaining novel biologically catalyzed reactions (*87*).

## A CETCH

*Crotonyl-CoA / ETthylmalonyl-CoA / 4-Hydroxybutyryl-CoA*




**Overall reaction:**

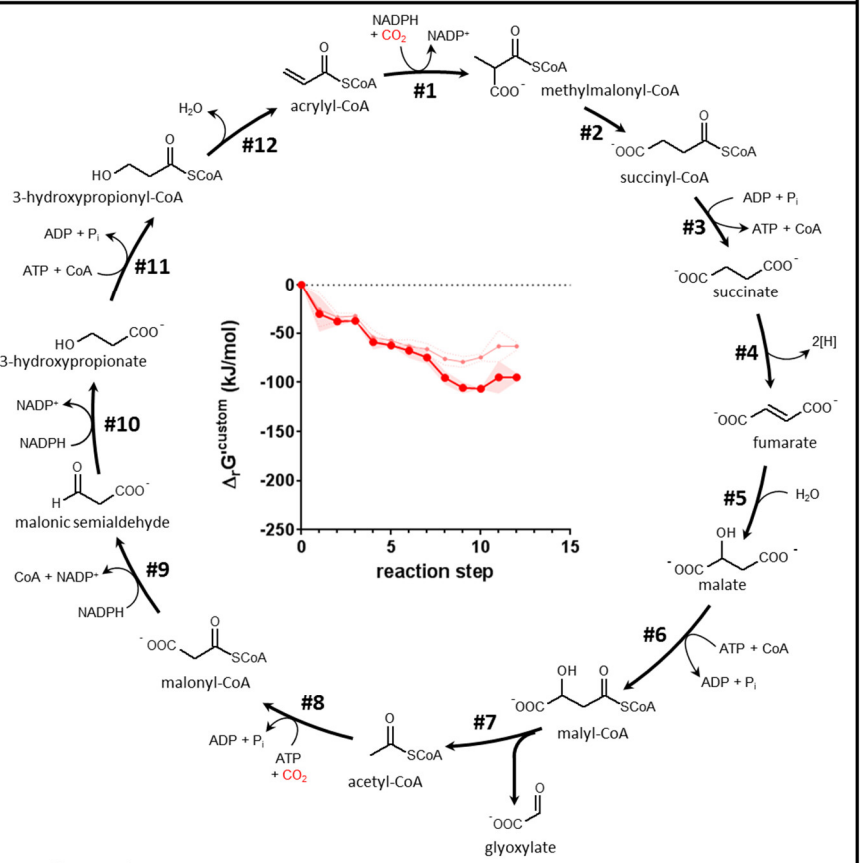
$2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 |
|---|---|---|---|---|---|---|---|---|---|---|

## B HOPAC

*HydrOxyPropionyl-CoA / ACrylyl-CoA*




**Overall reaction:**

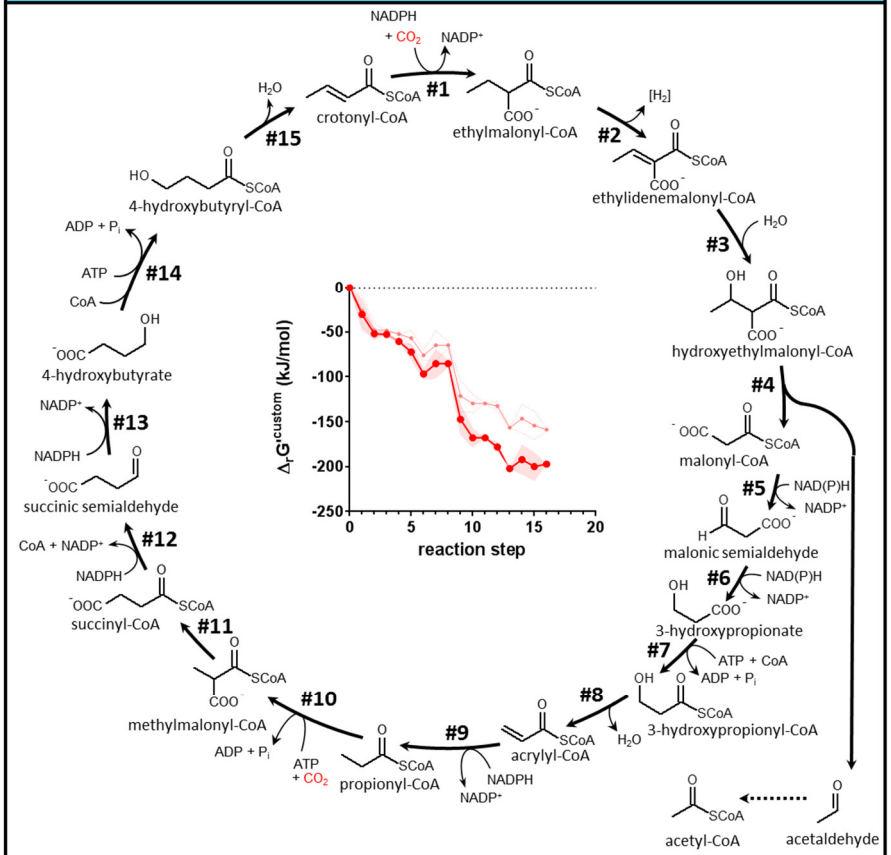
$2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 | #12 |
|---|---|---|---|---|---|---|---|---|---|---|---|

## C CHYME

*Crotonyl-CoA / HYdroxyethylmalonyl-CoA / MEthylmalonyl-CoA*




**Overall reaction:**

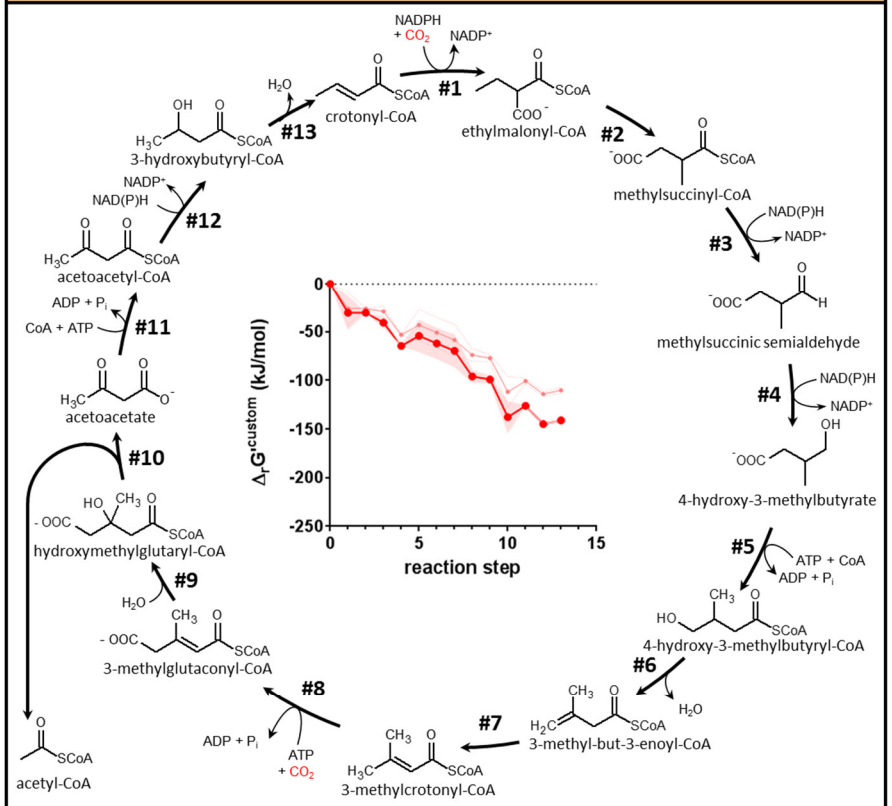
$2\ CO_2 + 6\ NAD(P)H + 3\ ATP + FAD \rightarrow acetaldehyde + 6\ NAD(P)^+ + 3\ ADP + 3\ P_i + FADH_2$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 | #12 | #13 | #14 | #15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

## D CRACE

*CRotonyl-CoA / ACEtoacetate*




**Overall reaction:**

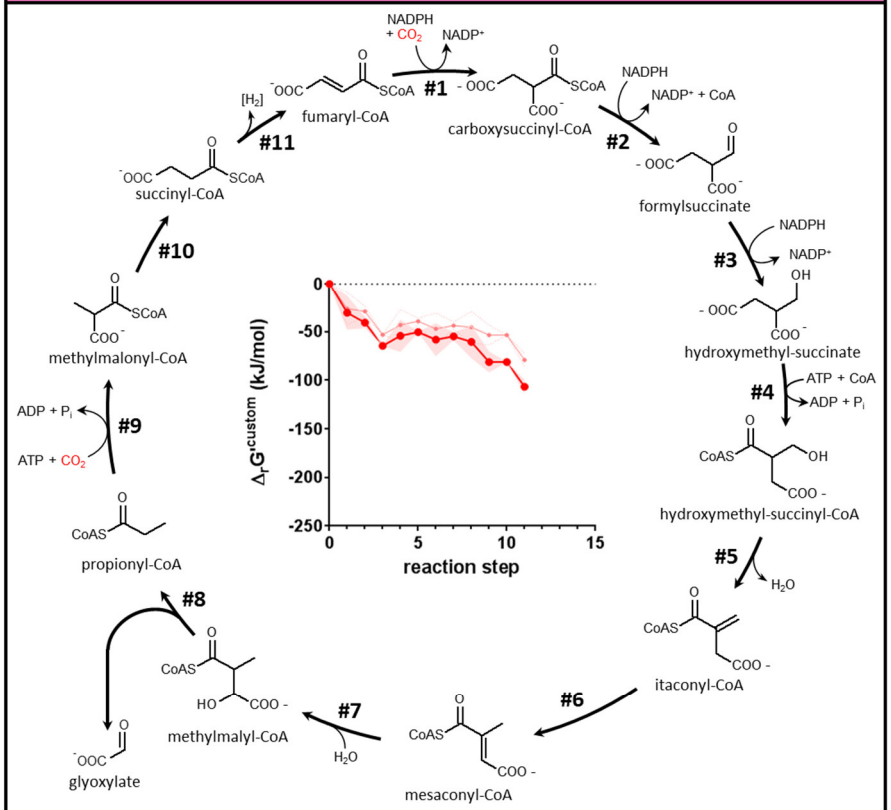
$2\ CO_2 + 4\ NAD(P)H + 3\ ATP \rightarrow acetyl\text{-}CoA + 4\ NAD(P)^+ + 3\ ADP + 3\ P_i$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 | #12 | #13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ✓ | ✓ | | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

## E FUMES

*FU*maryl-CoA / *ME*thylmalyl-CoA / *S*uccinyl-CoA




**Overall reaction:**

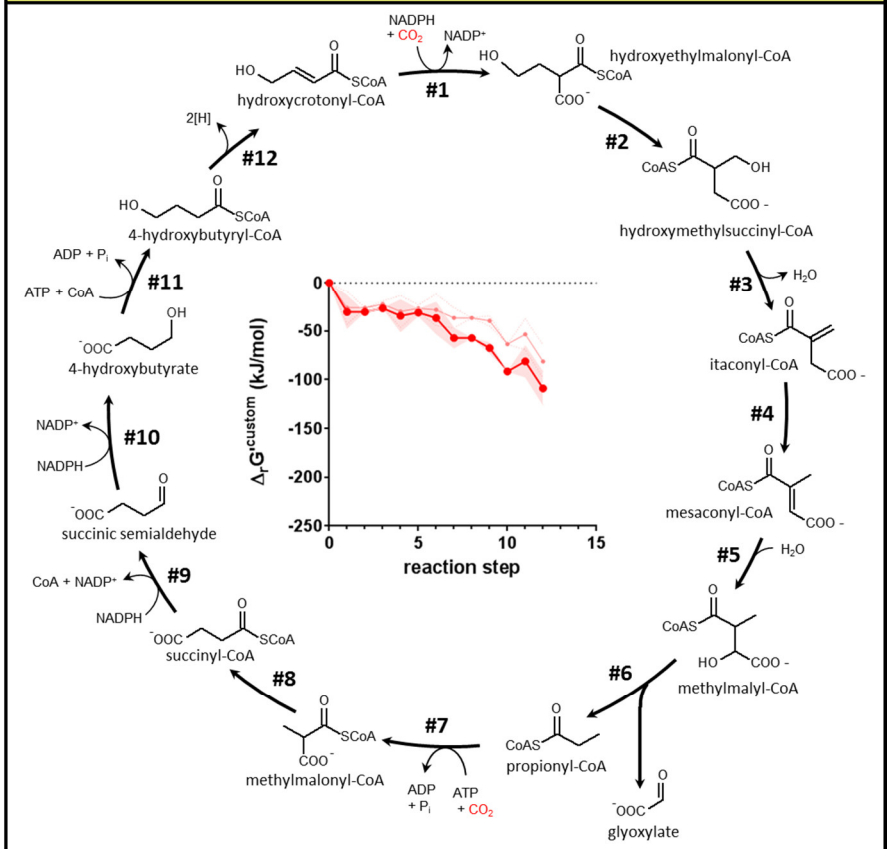
$2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | #5 | #6 | **#7** | **#8** | **#9** | **#10** | #11 |
|---|---|---|---|---|---|---|---|---|---|---|

## F HITME

*H*ydroxycrotonyl-CoA / *IT*aconyl-CoA / *ME*thylmalonyl-CoA




**Overall reaction:**

$2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$

**Potential enzyme candidates available:**

| #1 | #2 | #3 | #4 | **#5** | **#6** | **#7** | **#8** | **#9** | **#10** | **#11** | #12 |
|---|---|---|---|---|---|---|---|---|---|---|---|

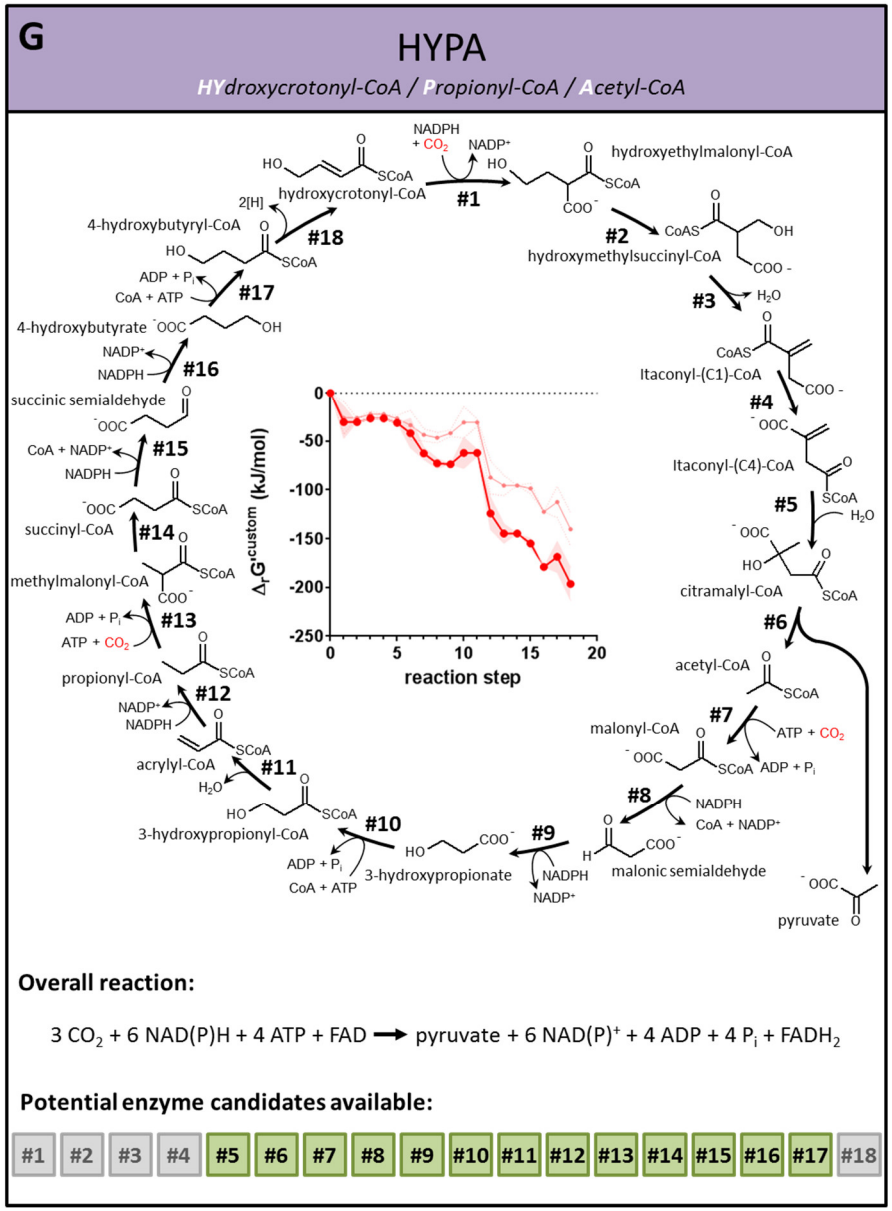



**Fig. 8 Topology and thermodynamic profiles of seven selected synthetic $CO_2$ fixation cycles.**

Depicted are seven promising synthetic $CO_2$ fixation cycles that were conceived by metabolic retrosynthesis. (**A**) CETCH cycle, (**B**) HOPAC cycle, (**C**) CHYME cycle, (**D**) CRACE cycle, (**E**) FUMES cycle, (**F**) HITME cycle, and (**G**) HYPA cycle.

The theoretical cycles were drafted based on a given ECR reaction (#1 in each reaction scheme). Note that the stereochemistry of individual reactions (*i.e.*, the requirement for epimerases and racemases) was not taken into account at the design stage, but only during the realization phase.

The Gibbs free energy profile of these theoretical cycles is shown at the center of the respective pathway. The Gibbs free energy profiles were based on calculations and estimations using the eQuilibrator software (*88*) (http://equilibrator.weizmann.ac.il/) or other sources (indicated in the tables in chapter 8.1). Shown in dark red are the summarized Gibbs free energies along the individual reactions for $\Delta_rG'$ at pH 7.8, ionic strength I = 0.2, and assuming the following metabolite concentrations: substrates (CoA ester, acids, aldehydes) = 0.2 mM; products (CoA ester, acids, aldehydes) = 0.2 mM; free CoA = 0.5 mM; Quinone / Quinol (2[H]) = 0.1 mM; ATP = 1.2 mM; ADP = 0.6 mM; AMP = 0.2 mM; NADPH = 3.6 mM; $NADP^+$ = 0.4 mM. Shown in light red is the summarized $\Delta_rG'^m$ (assuming a concentration of 1 mM for each metabolite) as comparison.

The overall reaction summarized from the individual reaction steps is given below each cycle.

An evaluation of enzyme candidates available to realize the theoretical cycles is also given. Reactions for which potential candidates could be identified are highlighted in green boxes and reactions for which obvious candidate enzymes were missing according to bioinformatic analysis and database queries are shown in grey boxes.

## 2.3. Analysis of energetic properties

We assessed the energetic cost of the synthetic drafts and natural $CO_2$ fixation pathway, by estimating the presumable ATP- and NAD(P)H-consumption of each cycle for the conversion of $CO_2$ into its primary fixation product. First, we derived the overall reaction stoichiometry of each cycle and added the number of NAD(P)H, ATP and $FADH_2$ required or generated during the conversion of $CO_2$ into the respective primary $CO_2$ fixation product.

**CETCH cycle**: $2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$
**HOPAC cycle**: $2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$
**CHYME cycle**: $2\ CO_2 + 6\ NAD(P)H + 3\ ATP + FAD \rightarrow acetaldehyde + 6\ NAD(P)^+ + 3\ ADP + 3\ P_i + FADH_2$
**CRACE cycle**: $2\ CO_2 + 4\ NAD(P)H + 3\ ATP \rightarrow acetyl\text{-}CoA + 4\ NAD(P)^+ + 3\ ADP + 3\ P_i$
**FUMES cycle**: $2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$
**HITME cycle**: $2\ CO_2 + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow glyoxylate + 3\ NAD(P)^+ + 2\ ADP + 2\ P_i + FADH_2$
**HYPA cycle**: $3\ CO_2 + 6\ NAD(P)H + 4\ ATP + FAD \rightarrow pyruvate + 6\ NAD(P)^+ + 4\ ADP + 4\ P_i + FADH_2$
**CBB cycle**: $3\ CO_2 + 6\ NAD(P)H + 9\ ATP \rightarrow 3\text{-}phosphoglyceraldehyde + 6\ NAD(P)^+ + 9\ ADP + 9\ P_i$
**3HP bi-cycle**: $3\ HCO_3^- + 6\ NAD(P)H + 5\ ATP + FAD \rightarrow pyruvate + 6\ NAD(P)^+ + 3\ ADP + 3\ P_i + 2\ AMP + 4\ PP_i + FADH_2$
**3HP/4HB cycle**: $2\ HCO_3^- + 4\ NAD(P)H + 4\ ATP \rightarrow acetyl\text{-}CoA + 4\ NAD(P)^+ + 4\ ADP + 2\ P_i$

To allow for comparison between the synthetic cycles and their naturally existing aero tolerant counterparts, we then normalized all pathways against a common multi-carbon product, either phosphoglycerate or acetyl-CoA and added the additional NAD(P)H, ATP and $FADH_2$ required or generated during these conversions. The results are listed in **Table 3**.

The CETCH, HOPAC, FUMES and HITME cycles that generate glyoxylate as primary $CO_2$ fixation product were normalized against phosphoglycerate according to the following scheme *(89)*:

**CETCH, HOPAC, FUMES, HITME cycles: 2 glyoxylate to phosphoglycerate**

(1) $2\ glyoxylate \rightarrow tartronic\ semialdehyde + CO_2$ (glyoxylate carboligase)
(2) $tartronic\ semialdeyhde + NADH \rightarrow glycerate + NAD^+$ (tartronic semialdehyde reductase)
(3) $glycerate + ATP \rightarrow 3\text{-}phosphoglycerate + ADP + P_i$ (glycerate kinase)

The CHYME cycle that generates acetaldehyde as primary $CO_2$ fixation product, was normalized against acetyl-CoA according to the following scheme:

**CHYME cycle: acetaldehyde to acetyl-CoA**

(1) $acetaldehyde + CoA + NADP^+ \rightarrow acetyl\text{-}CoA + NADPH$ (acylating dehydrogenase)

The HYPA cycle that generates pyruvate as primary $CO_2$ fixation product, was normalized against acetyl-CoA according to the following scheme:

**HYPA cycle: pyruvate to acetyl-CoA**

(1) $pyruvate + CoA + NADP^+ \rightarrow acetyl\text{-}CoA + CO_2 + NADPH$ (pyruvate dehydrogenase)

The CBB cycle that generates 3-phosphoglyceraldehyde as primary $CO_2$ fixation product was normalized against acetyl-CoA and phosphoglycerate according to the following schemes (*34*):

**CBB cycle: phosphoglyceraldehyde to phosphoglycerate:**

(1) $phosphoglyceraldehyde + P_i + NAD^+ \rightarrow 1,3\text{-}bisphosphoglycerate + NADH$ (glyceraldehyde-3-P dehydrogenase)
(2) $1,3\text{-}bisphosphoglycerate + ADP \rightarrow ATP + 3\text{-}phosphoglycerate$ (phosphoglycerate kinase)

**CBB cycle phosphoglyceraldehyde to acetyl-CoA:**

| | |
|---|---|
| (1) *phosphoglyceraldehyde + $P_i$ + $NAD^+$ → 1,3-bisphosphogylcerate + NADH* | (glyceraldehyde-3-P dehydrogenase) |
| (2) *1,3-bisphosphoglycerate + ADP → 3-phosphoglycerate+ ATP* | (phosphoglycerate kinase) |
| (3) *3-phosphoglycerate → 2-phosphoglycerate* | (phosphoglycerate mutase) |
| (4) *2-phosphoglycerate + $H_2O$ → PEP* | (enolase) |
| (5) *PEP + ADP → pyruvate + ATP* | (pyruvate kinase) |
| (6) *pyruvate + CoA + $NADP^+$ → acetyl-CoA + $CO_2$ + NADPH* | (pyruvate dehydrogenase) |

The 3HP bi-cycle that generates pyruvate as primary $CO_2$ fixation product was normalized against acetyl-CoA and phosphoglycerate according to the following schemes *(34).*

**3HP bi-cycle: pyruvate to phosphoglycerate:**

| | |
|---|---|
| (1) *pyruvate + ATP → PEP + AMP + $PP_i$* | (pyruvate:phosphate dikinase) |
| (2) *PEP → 2-phosphoglycerate + $H_2O$* | (enolase) |

**3HP bi-cycle: pyruvate to acetyl-CoA**

| | |
|---|---|
| (1) *pyruvate + CoA + $NADP^+$ → acetyl-CoA + $CO_2$ + NADPH* | (pyruvate dehydrogenase) |

The 3HP/4HB cycle that generates acetyl-CoA as primary $CO_2$ fixation product was normalized against phosphoglycerate according to the following scheme (*34*):

**3HP/4HB cycle: acetyl-CoA to phosphoglycerate:**

| | |
|---|---|
| (1) *acetyl-CoA + $CO_2$ + ATP → malonyl-CoA + ADP + $P_i$* | (acetyl-CoA carboxylase) |
| (2) *malonyl-CoA + 2 NADPH → 3-hydroxypropionate + CoA + 2 $NADP^+$* | (malonyl-CoA reductase) |
| (3) *3-HP + CoA + NADPH + ATP → propionyl-CoA + $NADP^+$ + ADP + $P_i$ + $H_2O$* | (propionyl-CoA synthase) |
| (4) *propionyl-CoA + $CO_2$ + ATP → methylmalonyl-CoA + ADP + $P_i$* | (propionyl-CoA carboxylase) |
| (5) *methylmalonyl-CoA → succinyl-CoA* | (methylmal.-CoA mutase) |
| (6) *succinyl-CoA + GDP + Pi → succinate + CoA + GTP* | (succinyl-CoA synthase) |
| (7) *succinate + $FAD^+$ → fumarate + $FADH_2$* | (succinate dehydrogenase) |
| (8) *fumarate + $H_2O$ → malate* | (fumarase) |
| (9) *malate + $NAD^+$ → oxaloacetate + NADH* | (malate dehydrogenase) |
| (10) *oxaloacetate + ATP → PEP + $CO_2$ + ADP* | (PEP carboxykinase) |
| (11) *PEP → 2-phosphoglycerate + $H_2O$* | (enolase) |

To calculate the ATP equivalents required per $CO_2$-molecule converted into 3-phosphoglycerate or acetyl-CoA, we converted all NAD(P)H and $FADH_2$ into ATP equivalents, assuming a P/O ratio of 2.5 NAD(P)H and 1.5 for $FADH_2$ (*90*). Additionally, all AMP-forming reactions were taken into account with two ATP equivalents (**Table 3**). To account for different P/O ratios for NAD(P)H reported in the literature (*90*), we also calculated the ATP equivalents required per $CO_2$-molecule converted into 3-phosphoglycerate and acetyl-CoA for a changing P/O ratio in comparison to natural oxygen tolerant pathway (especially the CBB cycle) (**Fig. 9**).

Furthermore, the photons required per $CO_2$-molecule converted into 3-phosphoglycerate in a theoretical photosynthetic process was calculated. First, we accounted for the amount of photons in a linear electron flow process to generate the number of ferredoxins and NAD(P)H molecules required, assuming a coupled stoichiometry of 4 photons per 1 NAD(P)H and 1.28 ATP generated *(91)*. If $FADH_2$ was formed through a given pathway, the energy of $FADH_2$ is conserved on the level of the plastoquinone pool by assuming a stoichiometry of 2 photons (and no ATP) to generate 1 NADPH from $FADH_2$. If the number of ATP molecules generated through the coupled linear electron flow was not sufficient, the extra ATP required was calculated from a cyclic electron flow process with a stoichiometry of 2 photons per 1.71 ATP generated *(91)*. The additional number of ATP-molecules left from the overall process was also determined (**Table 3** and **Fig. 10**).

**Table 3 Comparison of synthetic $CO_2$ fixation cycles with known naturally evolved, aero-tolerant $CO_2$ fixation pathways.**

Normalized to 3-phosphoglycerate as fixation product

| Pathway | Primary $CO_2$-fixation product | Normalized $CO_2$ fixation product | $\Sigma_{NAD(P)H}$ | $\Sigma_{ATP}$ | $\Sigma_{FADH2}$ | ATP equiv. per $CO_2$ | Photons per $CO_2$ (additional ATP molecules left) [a] |
|---|---|---|---|---|---|---|---|
| **CETCH cycle** | glyoxylate | 3-phosphoglycerate | -7 | -5 | 2 | -6.5 | 24 (+ 1.4 ATP) |
| **HOPAC cycle** | glyoxylate | 3-phosphoglycerate | -7 | -5 | 2 | -6.5 | 24 (+ 1.4 ATP) |
| **FUMES cycle** | glyoxylate | 3-phosphoglycerate | -7 | -5 | 2 | -6.5 | 24 (+ 1.4 ATP) |
| **HITME cycle** | glyoxylate | 3-phosphoglycerate | -7 | -5 | 2 | -6.5 | 24 (+ 1.4 ATP) |
| **CBB cycle (without $O_2$)** | 3P-glyceraldehyde | 3-phosphoglycerate | -5 | -8 | 0 | -6.8 | 22 (+0.1 ATP) |
| **CBB cycle (with $O_2$)** [b] | 3P-glyceraldehyde | 3-phosphoglycerate | -7 | -12 | 0 | -9.9 | 33 (+0.2 ATP) |
| **3HP bi-cycle** | pyruvate | 3-phosphoglycerate | -6 | -9 | 1 | -7.5 | 26 (+0.8 ATP) |
| **3HP/4HB cycle** [c] | acetyl-CoA | 3-phosphoglycerate | -6 | -7 | 1 | -6.8 | 23 (+0.3 ATP) |
| **CETCH 5.4 (as realized)** | glyoxylate | 3-phosphoglycerate | -9 | -3 | 4 | -6.5 [d] /-8.5[e] | 28 (+3.4 ATP) |
| **HOPAC 4.1 (as realized)** | glyoxylate | 3-phosphoglycerate | -7 | -5 | 2 | -6.5 | 24 (+ 1.4 ATP) |

Normalized to acetyl-CoA as fixation product

| Pathway | Primary $CO_2$-fixation product | Normalized $CO_2$ fixation product | $\Sigma_{NAD(P)H}$ | $\Sigma_{ATP}$ | $\Sigma_{FADH2}$ | ATP equiv. per $CO_2$ | Photons per $CO_2$ (additional ATP molecules left) [a] |
|---|---|---|---|---|---|---|---|
| **CHYME cycle** | acetaldehyde | acetyl-CoA | -5 | -3 | 1 | -7.0 | 18 (+ 2.1 ATP). |
| **CRACE cycle** | acetyl-CoA | acetyl-CoA | -4 | -3 | 0 | -6.5 | 16 (+ 2.1 ATP) |
| **HYPA cycle** | pyruvate | acetyl-CoA | -5 | -4 | 1 | -7.5 | 18 (+ 1.1 ATP) |
| **CBB cycle (without $O_2$)** | 3P-glyceraldehyde | acetyl-CoA | -4 | -7 | 0 | -8.5 | 19 (+ 0.7 ATP) |
| **CBB cycle (with $O_2$)** [b] | 3P-glyceraldehyde | acetyl-CoA | -6 | -11 | 0 | -13 | 30 (+ 0.8 ATP) |
| **3HP bi-cycle** | pyruvate | acetyl-CoA | -5 | -7 | 1 | -9.0 | 21 (+ 0.7 ATP) |
| **3HP/4HB cycle** [c] | acetyl-CoA | acetyl-CoA | -4 | -4 | 0 | -7.0 | 16 (+ 1.1 ATP) |

To calculate the ATP equivalents required per $CO_2$-molecule converted into 3-phosphoglycerate or acetyl-CoA, we converted all NAD(P)H and $FADH_2$ into ATP equivalents, assuming a P/O ratio of 2.5 NAD(P)H, 1.5 for $FADH_2$.and 3 for ferredoxin. Additionally, every AMP-forming reaction was accounted for two ATP equivalents. [a] These calculations consider the conservation of electrons from additional $FADH_2$ generated through a given pathway. [b] Assuming a photorespiration rate of 20% and additional -0.4 ferredoxin per phosphoglycerate. [c] Calculation for the more energy efficient 3HP/4HB cycle from Thaumarchaeota. [d] Calculations for CETCH 5.4, where the electrons from the two oxidase reactions could be conserved. [e] This calculation does not consider the conservation of additional $FADH_2$ created in CETCH 5.4.

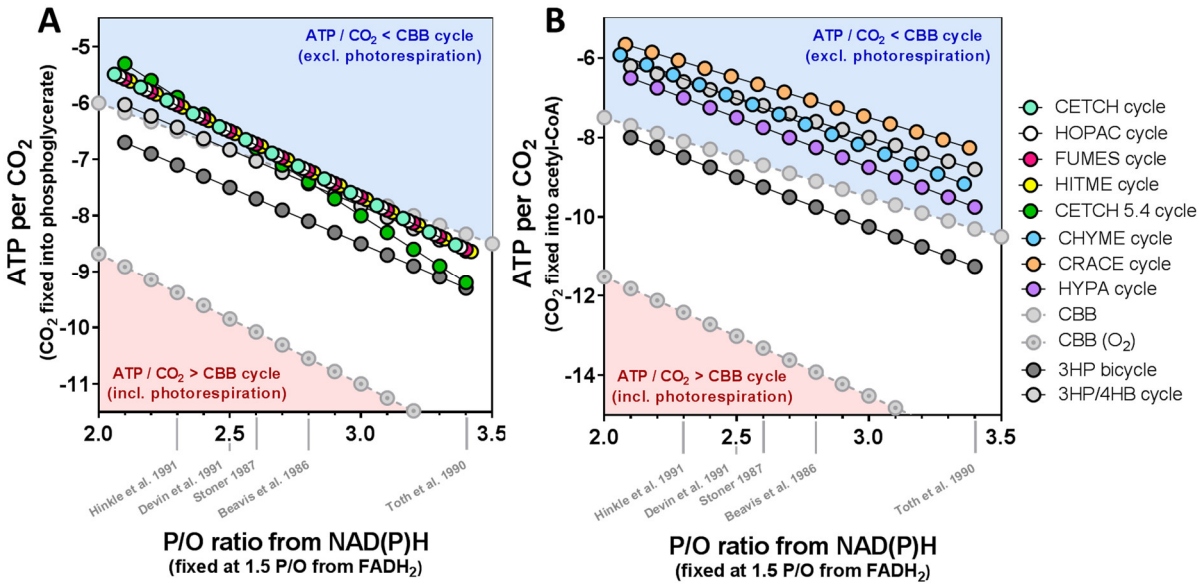



**Fig. 9 Comparison of ATP equivalents required per $CO_2$ converted into phosphoglycerate or acetyl-CoA in respect to a changing P/O ratio.** Comparison of the ATP equivalents required per $CO_2$ fixed (y-axis) in respect to different P/O ratios (x-axis), with some experimentally reported P/O ratios indicated on the x-axis. The efficiency of the CBB cycle was calculate with and without photorespiration indicated by the blue and red area respectively. (A) ATP equivalents per $CO_2$ fixed into phosphoglycerate for theoretical CETCH, HOPAC , HITME, and FUMES cycle drafts, as well as the experimentally realized CETCH 5.4 cycle compared to natural oxygen tolerant carbon fixation cycles. (B) ATP equivalents per $CO_2$ fixed into acetyl-CoA for theoretical CHYME, CRACE, and HYPA cycle drafts compared to natural oxygen tolerant carbon fixation cycles.

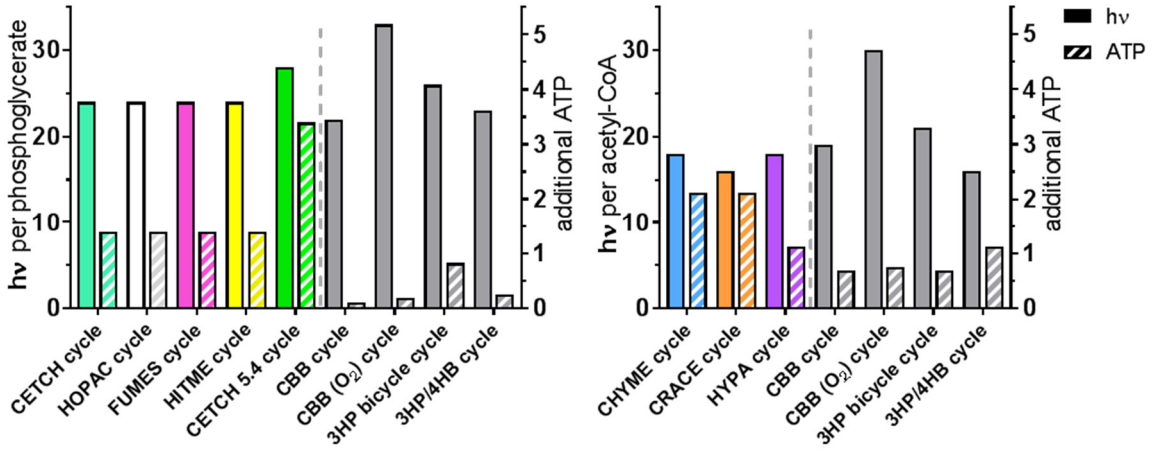



**Fig. 10 Photons required per phosphoglycerate or acetyl-CoA molecule formed from $CO_2$ in a photosynthetically-coupled process**. Shown are the total number of photons required for the formation of phosphoglycerate (A) and acetyl-CoA (B) for the theoretical cycle drafts, as well as the experimentally realized CETCH 5.4 cycle in comparison to natural oxygen tolerant $CO_2$ fixation pathways (left y-axis and filled bars). In case of all synthetic cycles, linear electron flow alone would be sufficient to generate all NADPH and ATP required. Through the fixed stoichiometry of 1:1.28 per 4 photons in linear electron flow, ATP is produced in excess compared to NADPH for these cycles. In case of the natural cycles, linear electron flow alone is not sufficient to produce all ATP required. The additional ATP needed is generated through cyclic electron flow (at a stoichiometry of 1.71 ATP per 2 photons). ATP produced in excess after all photons are taken into account for linear and cyclic electron flow are indicated in the patterned bars (right y-axis).

In summary, the energetic calculations and analysis showed, that the synthetic pathways are in general comparable or even more energy efficient than natural occurring oxygen tolerant $CO_2$ fixation pathways. Especially, if the cycles are compared to the CBB cycle, all pathways are superior, due to the fact, that RubisCO incorporates molecular oxygen and energy is lost in photorespiration. For example, the CETCH, HOPAC, FUMES and HITME cycle require 1/3 ATP less per $CO_2$ per phosphoglycerate formed from $CO_2$ compared to the CBB under oxygenic conditions. This also corresponds to at least nine photons less per phosphoglycerate in a theoretical photosynthetic process (**Table 3**). Conclusively, these results suggested that our synthetic cycles could provide very promising solutions to a more efficient C-C bond formation under aerobic conditions that compete with (and eventually surpass) any known naturally existing solution.

In the first phase of this study, we designed seven promising synthetic pathways for the fixation of $CO_2$. As a proof of principle, we strongly focused on realizing some of these candidates *in vitro.* This intermediate step between a direct *in vivo* transplantation, allows the analysis and optimization of $CO_2$ fixation capabilities in a controlled cell-free environment. Additionally, this will serve as a proof that the metabolic retrosynthesis approach is a viable option for the design of functional and novel $CO_2$ fixation pathways. From the seven promising pathway designs, the CETCH and HOPAC cycle were considered further for the *in vitro* realization. For both pathways, the intended enzymatic reactions can already be found in bioinformatics databases. As aforementioned, the missing reactions could be obtained by enzyme engineering or directed evolution. This would require the initial development of novel catalysts. The CETCH and HOPAC cycle did not require initial enzyme engineering and we could focus on the assembly of the pathway *in vitro.*

# 3. Realization of the CETCH cycle

From the theoretical drafts, we first aimed at realizing one of these synthetic cycles. To that end we searched bioinformatics databases for enzymes that could sustain the theoretical CETCH cycle (**Fig. 11**). We were able to identify one or more potential enzyme candidates for each reaction of the CETCH cycle. These candidates were expressed and characterized to select a final set of 12 enzymes, each with a specific activity >1 U $mg^{-1}$ protein (at 250 µM substrate concentration), that could catalyze the proposed reaction sequence of the CETCH cycle (**Table 4**).

**Table 4 List of enzymes considered, tested, characterized and used to realize the CETCH cycle.**

| Rxn | Abb. | Enzyme full name | Name | Organism | $v_{max}$ (U $mg^{-1}$) | $K_M$ (mM) | Source of kinetic data |
|---|---|---|---|---|---|---|---|
| **1** | Pco | propionyl-CoA oxidase | Pco | *A. thaliana* | 12 | 0.044 | This work |
| **1** | Acx4 | acyl-CoA oxidase | Acx4 | *A. thaliana* | 39 | 0.013 | This work |
| **1→2** | Pcc | propionyl-CoA carboxylase | PccAB | *M. extorquens* | 24 | 0.1. | This work |
| **2+8** | Ccr | crotonyl-CoA carboxylase/reductase | Ccr | *M. extorquens* | 110 | 0.17 | This work |
| **2+8** | Ccr | crotonyl-CoA carboxylase/reductase | CcrR152A | *M. extorquens* | 85 | 0.82 | This work |
| **2+8** | Ccr | crotonyl-CoA carboxylase/reductase | Ccr | *K. setae* | 21 | 0.01 | Peter (*84*) |
| **2+8** | Ccr | crotonyl-CoA carboxylase/reductase | Ccr | *L. borgpetersenii* | 75 | 0.07 | Peter (*84*) |
| **3+9** | Epi | emC/mmC epimerase | Epi | *R. sphaeroides* | 440 | 0.08 | Erb et al. 2008 (*92*) |
| **3** | Mcm | methylmalonyl-CoA mutase | Mcm | *R. sphaeroides* | 20 | 0.14 | Erb et al. 2009 (*92*) |
| **4** | Scr | succinyl-CoA reductase | SucD | *C. kluyveri* | 29 | 0.003 | Söhling et al. 1993 (*93*) |
| **5** | Ssr | succinic semialdehyde reductase | AKR7a2 | *H. sapiens* | 3.9 | 0.013 | This work |
| **5** | Ssr | sulfolactaldehyde reductase | YihU | *E. coli* | 2 | 26.7 | This work |
| **6** | Hbs | 4-hydroxybutyryl-CoA synthetase | Nmar0206 | *N. maritimus* | 2 | 0.19 | Könneke et al. 2014 (*34*) |
| **6** | Hbs | 3-hydroxypropionyl-CoA synthetase | Nmar1309 | *N. maritimus* | 3 | 28 | Könneke et al. 2014 (*34*) |
| **6** | Hbs | medium-chain fatty acid CoA ligase | AlkK | *P. putida* | 4 | 69.7 | This work |
| **7** | Hbd | 4-hydroxybutyryl-CoA dehydratase | Nmar0207 | *N. maritimus* | 26 | 0.06 | Könneke et al. 2014 (*34*) |
| **7** | Hbd | 4-hydroxybutyryl-CoA dehydratase | AbfD | *C.aminobutyricum* | 19 | 0.06 | Könneke et al., 2014 (*34*) |
| **9** | Ecm | ethylmalonyl-CoA mutase | Ecm | *R. sphaeroides* | 7 | 0.06 | Erb et al. 2008 (*92*) |
| **10** | Mco | methylsuccinyl-CoA oxidase | Mco | *R. sphaeroides.* | 0.1 | 0.03 | This work |
| **10** | Mcd | methylsuccinyl-CoA dehydrogenase | Mcd | *R. sphaeroides* | 6 | 0.04 | Erb et al. 2009 (*94*) |
| **11** | Mch | mesaconyl-CoA hydratase | Mch | *R. sphaeroides* | 1500 | n.d. | Zarzycki et al. 2008 (*95*) |
| **12** | Mcl | β-methylmalyl-CoA lyase | Mcl1 | *R. sphaeroides* | 5 | 0.01 | Erb et al. 2010 (*96*) |
| **13** | Mas | malate synthase | GlcB | *E. coli* | 36 | 0.021 | Anstrom et al. 2003 (*97*) |
| **14** | Mct | malyl-CoA thioesterase | Mcl2 | *R. sphaeroides* | 200 | 0.09 | Erb et al. 2010 (*96*) |
| **15** | Fdh | formate dehydrogenase (D221A) | Fdh | *M. vaccae* | 1.4 | 0.37 | Hoelsch et al. 2013 (*98*) |
| **16** | Pkk | polyphosphate kinase | Pkk2 | *S. meliloti* | 12 | 0.032 | Nocek et al. 2008 (*99*) |
| **17** | Kat | katalase | KatE | *E. coli* | 11740 | 86.5 | Sevinc et al. 1995 (*100*) |
| **18** | Gcl | Glyoxylate carboligase [a] | Gcl | *E. coli* | 18 | 0.9 | Kaplun et al. 2008 (*101*) |
| **19** | Gar | Tartronate semialdehyde reductase[a] | GarR | *E. coli* | 410 [b] | 0.2 [b] | Gotto et al. 1961 (*102*) |
| **20** | Glk | Glycerate kinase [a] | GlxK | *E. coli* | 960 | 138 | This work |

The reaction numbering is according to (**Fig. 17, Fig. 18, Fig. 19, Fig. 21** and **Fig. 22**). Whereas reactions from 1 to 12 represent the core sequence of the CETCH cycle, reactions from 13 to 20 are additional enzymes introduced over the process of optimization and for the conversion of the primary substrate. Highlighted in green are those enzymes used in the final version of CETCH 5.4. [a] These enzymes were tested as proof-of-principle that they could convert glyoxylate into phosphoglycerate, demonstrating a potential connection with central carbon metabolism and allowing for the thermodynamic comparison of CETCH with the CBB. [b] Values for *Pseudomonas ovalis* enzymes

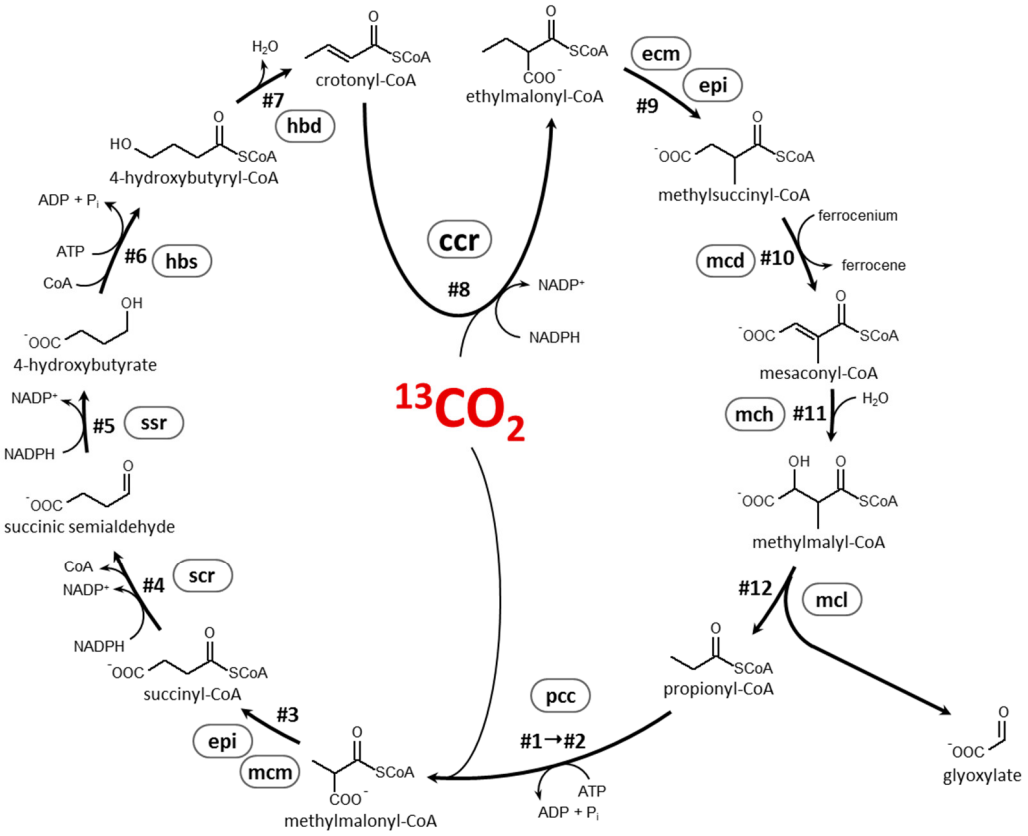



**Fig. 11 Topology of the CETCHcycle from the initial design phase.** Enzyme names are abbreviated according to **Table 4**.

## 3.1. CETCH 1.0 (discontinuous/stepwise reconstitution)

To demonstrate feasibility of the CETCH cycle, we first reconstituted its central $CO_2$ fixation reaction sequence stepwise. We provided cofactors and the chemically synthesized intermediate propionyl-CoA in a buffer containing $NaHCO_3$ as a source of $CO_2$. To this mixture, Pcc was added first, followed by the subsequent and stepwise addition of the other eleven enzymes required for the assembly of the pathway. HPLC-MS analysis demonstrated the stepwise transformation of propionyl-CoA into key intermediates of the pathway (**Fig. 12**), indicating that this first version of the CETCH cycle (CETCH 1.0) was in principle functional.

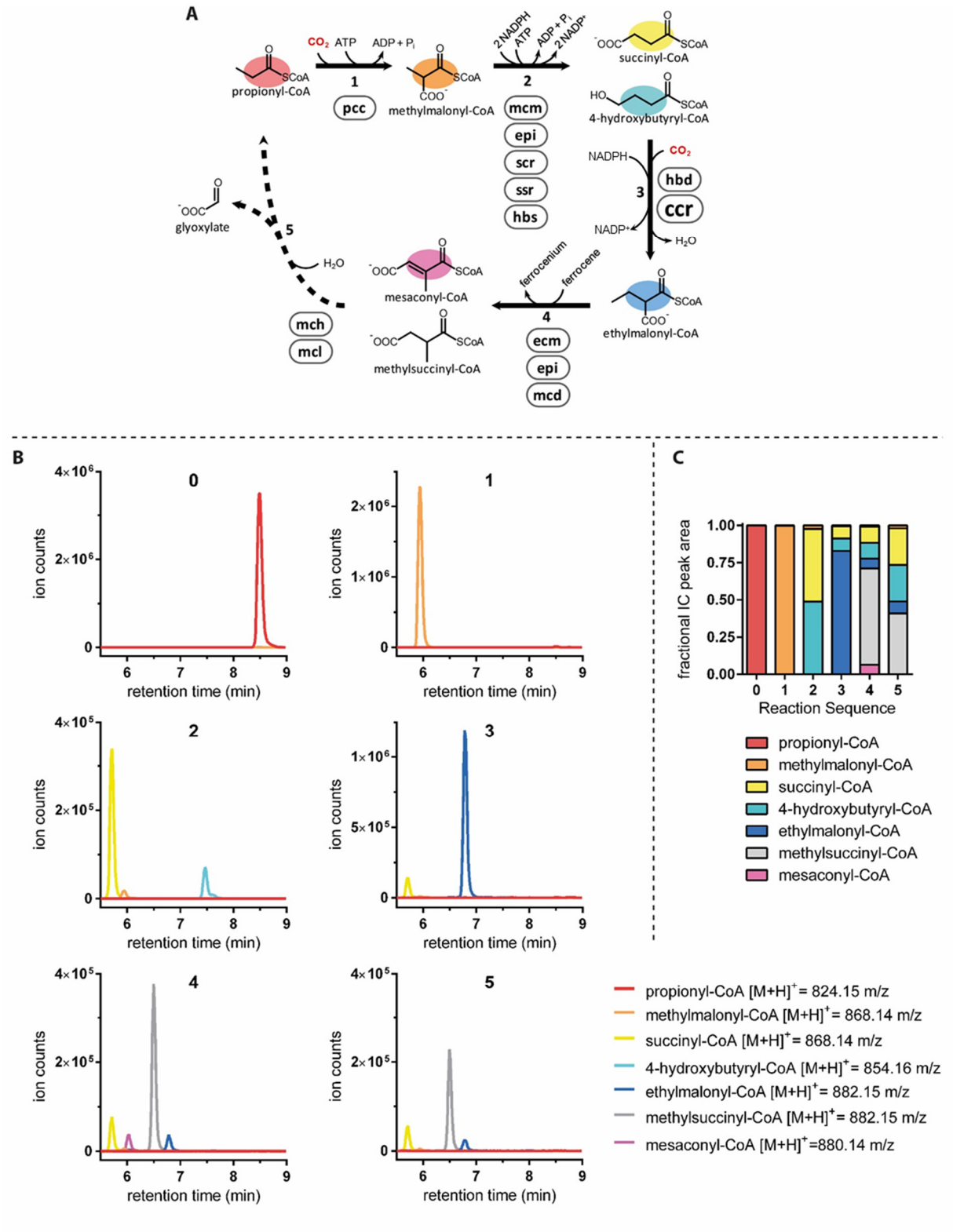



**Fig. 12 Stepwise reconstitution of the CETCH cycle.** (A) Reconstitution scheme of the CETCH cycle. Enzyme names are abbreviated according to **Table 4**. The CETCH cycle was reconstituted in five individual steps through the subsequent addition of one or multiple enzymes, as indicated for each step. (B) HPLC-MS analysis of the reaction mixture from step 0 (before addition of Pcc) to step 5 (after all enzymes were added). Shown are the extracted high resolution-MS traces of individual intermediates, confirming their stepwise formation. Color coding of intermediates is according to panel (A). (C) Fractional peak areas of these key intermediates (in total ion counts) from step 0 to 5, coloring according to panel (A).

## 3.2. CETCH 1.0 (continuous cycling is blocked on the level of Mcd)

After we had demonstrated the individual reactions of the CETCH cycle (CETCH 1.0) in a discontinuous fashion and stepwise, we aimed at establishing a continuous operating cycle that could fix $CO_2$ in multiple turns. Additionally, by using $^{13}C$ labeled bicarbonate ($NaH^{13}CO_3$), incorporation into the intermediates occurs according to the numbers of turns and enables the analysis of the continuous operation of the synthetic cycle (**Fig. 13**). In a first attempt, we provided all enzymes, cofactors (NADPH, ATP and CoA) and the artificial electron acceptor ferrocenium (100 µM) from the beginning and started the cycle by addition of propionyl-CoA (**Fig. 14A**). However, under these continuous operation conditions, CETCH 1.0 was not functional, as demonstrated by intermediate detection and $^{13}C$-labeling on HPLC-MS. The conversion of CoA-ester intermediates stopped at the level of methylsuccinyl-CoA, indicating that the methylsuccinyl-CoA dehydrogenase (Mcd) reaction was rate limiting (Rxn #10, **Fig. 14B-C**). Mcd belongs to the family of FAD-dependent acyl-CoA dehydrogenases (ACD), which are naturally coupled via electron transfer flavoproteins to the membrane bound ubiquinone pool (via electron transfer flavoprotein:ubiquinone oxidoreductase) and eventually to the respiratory chain. In the CETCH 1.0 we added the artificial electron acceptor ferrocenium together with the enzyme to replace this complicated electron transfer cascade. However, under continuous operation the use of ferrocenium with Mcd was not possible. The maximum ferrocenium concentration in the assay was limited to 0.1 mM, causing protein precipitation at higher concentrations. Moreover, ferrocenium has an unfavorable redox behavior under the assay conditions (*i.e.* it reacts spontaneously with NADPH, free CoA and has a slow re-oxidation rate (*103*)) preventing the effective cycling of CETCH 1.0.

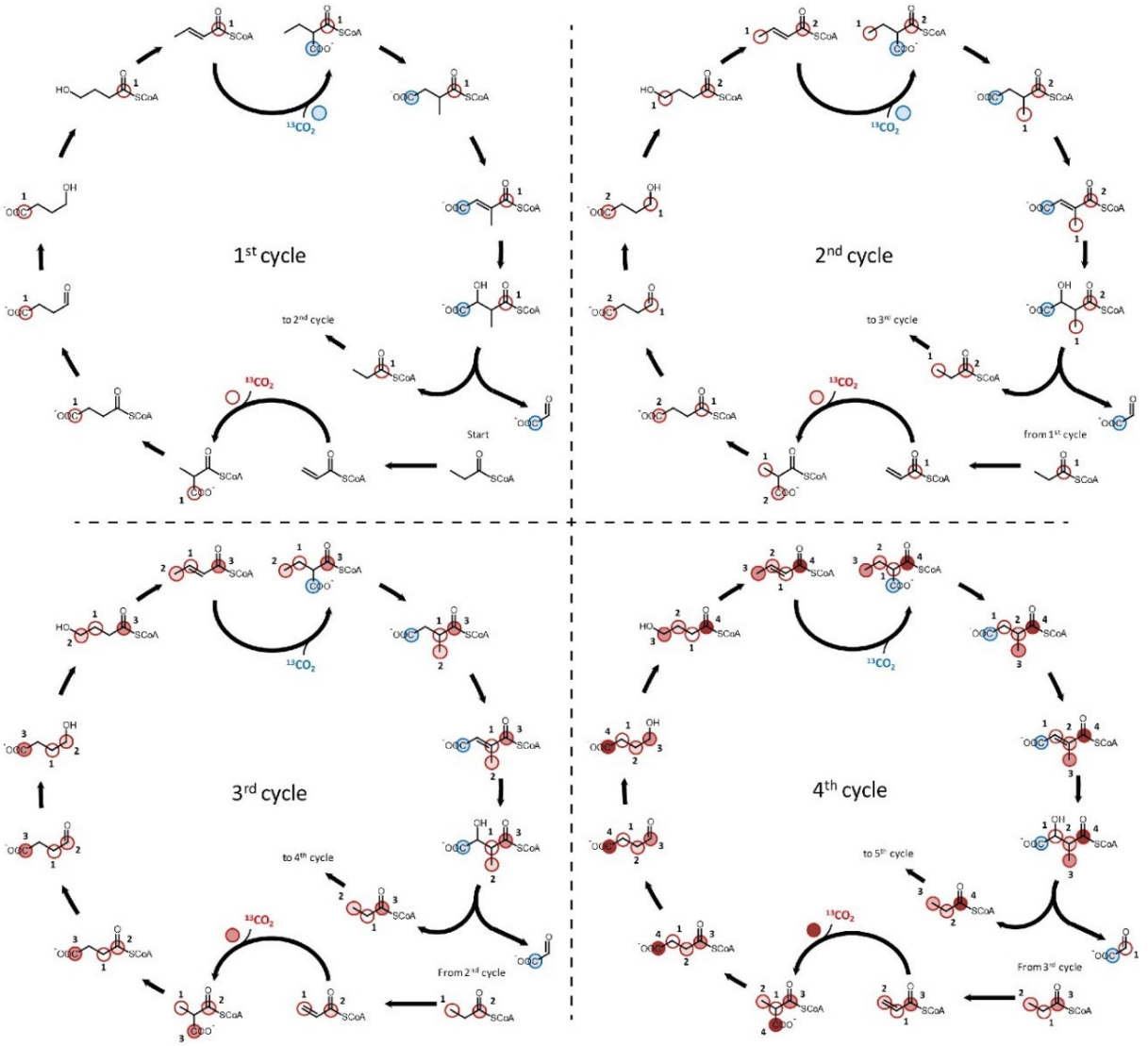



**Fig. 13 Expected $^{13}$C-label incorporation of the CETCH cycle from $^{13}CO_2$.** Shown is the expected labeling pattern for each turn of the CETCH cycle in its final version 5.4 (but is essentially the same for all other versions). The cycle features two $CO_2$-incorporation steps per turn shown in blue and red dots. The carbon incorporated into the $C_4$-acceptor crotonyl-CoA is shown in blue. This carbon ends up in the output molecule glyoxylate during each turn. The carbon incorporated into the $C_3$-acceptor propionyl-CoA (CETCH 1.0 to 4.0) or acrylyl-CoA (CETCH 5.0 to 5.4) is shown in red. This label stays with the acceptor molecule during each turn of the cycle (indicated by a darker red color). Note that the CETCH cycle includes carbon skeleton-rearranging steps catalyzed by mutases, resulting in a complex labeling pattern.

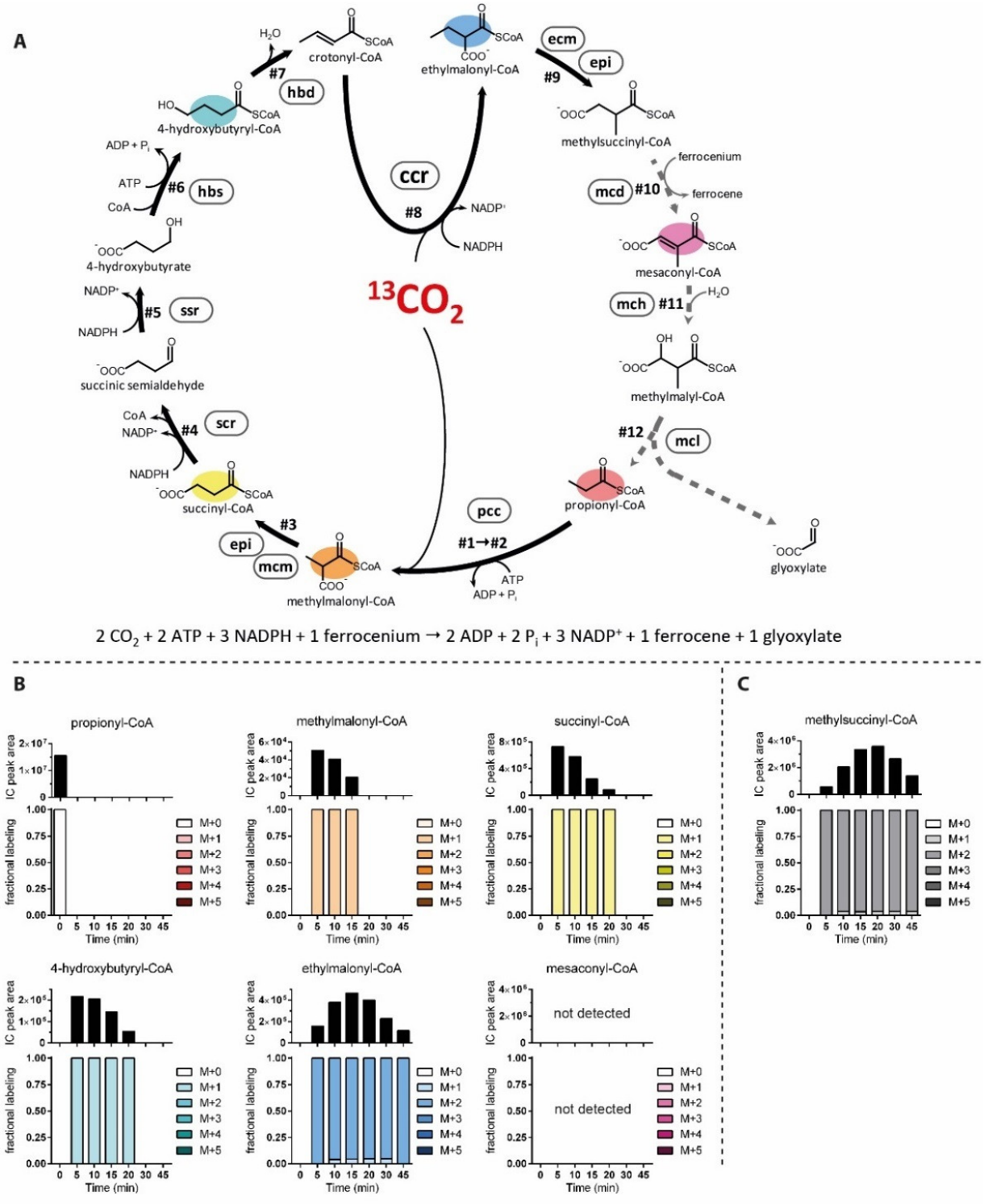



**Fig. 14 CETCH 1.0 (continuous cycling is blocked on the level of Mcd).** (A) Topology of the CETCH cycle version 1.0, operated in continuous mode. Enzyme names are abbreviated according to **Table 4**. (B) Dynamics of key intermediates of CETCH 1.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labeling pattern of six different CoA esters with color-coding according to (**A**). Fractional labeling showed that the cycle was not turning and that CoA esters accumulated at the level of methylsuccinyl-CoA, indicating that it was stalled at the level of Mcd (Rxn #10). (C) Level and fractional labeling pattern of methylsuccinyl-CoA.

## 3.3. CETCH 2.0 (engineering of Mcd to Mco allows multiple turning)

In CETCH 2.0 we aimed at overcoming the limitation caused by Mcd by engineering the enzyme to directly use molecular oxygen as electron acceptor. Modeling of Mcd onto the structure of short-chain acyl-CoA dehydrogenase from *Homo sapiens* (38% amino acid identity, 68% coverage, PDB ID: 2VIG) indicated a negatively charged, narrow pocket opposite of the redox active N5 of the isoalloxazine ring of the FAD cofactor. The same FAD pocket appeared to be wider in acyl-CoA oxidases (*e.g.*, short chain acyl-CoA oxidase 4 from *A. thaliana*, PDB ID: 2IX5) that are able to reoxidize the FAD cofactor with oxygen (**Fig. 16A**). A sequence alignment, comparing dehydrogenases with oxidases, revealed a conserved threonine (Thr317 in Mcd), which is replaced by a glycine in oxidases. The structural model showed, that the Thr315 is in proximity to the N5 of the isoalloxazine ring of the FAD cofactor (**Fig. 15**). We assumed that this residue could modulate the electrochemical potential of the FAD and might also prevent the active binding of molecular oxygen. Introduction of a Thr317Gly mutation into Mcd from *R. sphaeroides* already supported initial oxidase activity (**Fig. 16** and **Table 5**). Next, at the position of Glu377 in Mcd, the oxidases poses a conserved asparagine, which is also close to the FAD cofactor in the active site and might be involved in the oxidase activity (**Fig. 15** and **Fig. 16A**). Introduction of the single Glu377Asn mutation showed very low oxidase activity and in combination with the Thr317Gly mutation (T317G, E377N), oxidase activity was lower than of the single Thr317Gly variant (**Fig. 16** and **Table 5**). Additionally, we investigated the Trp315 of Mcd, since it is positioned in the active site tunnel and mutations might contribute to the solvation of the active site and change the accessibility for molecular oxygen (**Fig. 16A**). A single mutation of Trp315Ala catalyzed no reaction with oxygen, but a Trp315Phe mutation showed very low oxidase activity (**Fig. 16B**). The combination of the Trp315Phe with the Thr317Gly mutation (W315F, T317G) increased the oxidase activity, compared to the single Thr317Gly variant. Interestingly, the additional Glu377Asn mutation (W315F, T317G, E377N) increased the reaction rate with oxygen further, even though it showed a negative effect in the double variant (T317G, E377N) (**Fig. 16C** and **Table 5**). Conclusively, three point mutations (T317G, E377N, W315F) allowed us to convert methylsuccinyl-CoA dehydrogenase into a functional methylsuccinyl-CoA oxidase (Mco) that catalyzed the oxidation of methylsuccinyl-CoA with a $v_{max}$ of 97.3 ± 5.4 mU $mg^{-1}$ and at an apparent $K_M$ of 27.4 ± 5.0 µM.

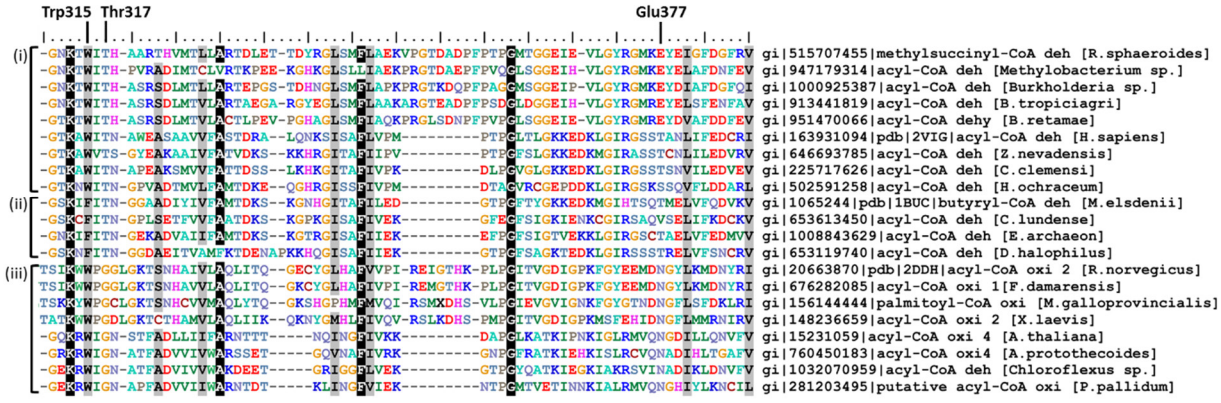



**Fig. 15 Sequence alignment of different ACDs and oxidases.** (i) Represents ACD from different organisms. (ii) Represents homologs from the butyryl-CoA dehydrogenases from the anaerobic living M. elsdenii, which is known to have an intrinsic oxidase activity (*104*). (iii) Represents different acyl-CoA oxidases. Trp315 (from Mcd) is conserved in dehydrogenases as well as oxidases. This residue is positioned in the active site tunnel and replacing it by a phenylalanine (as found in anaerobic dehydrogenases) might enable solvation of the active site. Thr315 (from Mcd) is a conserved active site residue amongst dehydrogenases and is in proximity to the N5 of the FAD cofactor. This residue is replaced by a glycine in oxidases. Glu377 is close to the active site and the FAD cofactor, while some dehydrogenases poses a glutamate, the oxidases show a conserved asparagine.

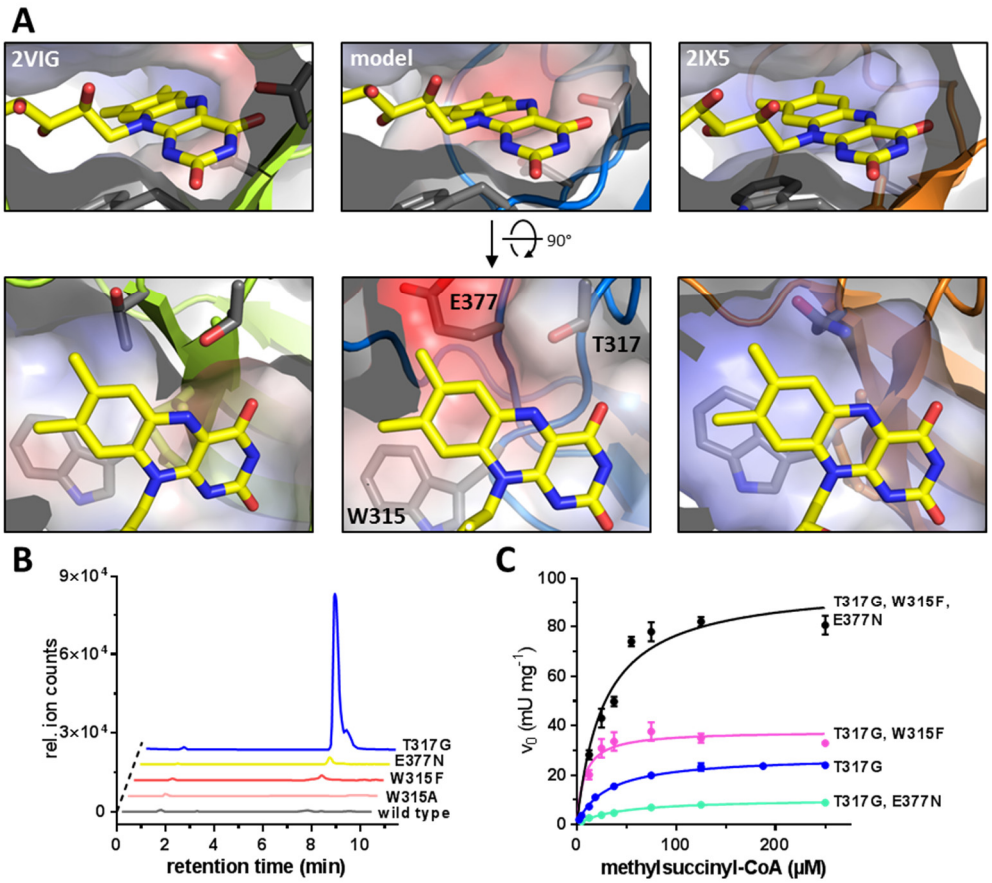



**Fig. 16 Structure guided engineering of Mcd into Mco.** (A) Active site comparison of the human short-chain acyl-CoA dehydrogenase (PDB ID: 2VIG, green backbone) with a model of Mcd from *Rhodobacter sphaeroides* (blue backbone, modeled by the SWISS-MODEL server with 2VIG as template) and the short-chain acyl-CoA oxidase 4 from *Arabidopsis thaliana* (PDB ID: 2IX5, orange backbone). The three residues that were targeted to introduce oxidase activity into Mcd are highlighted. (B) HPLC-MS based analysis of wild-type and different single active-site variants for oxidation of methylsuccinyl-CoA into mesaconyl-CoA with molecular oxygen as electron acceptor. The screen identified three substitutions (W315F, T317G and E377N) that were combined into double and triple variants. (C) Michaelis-Menten graphs of single, double and triple variants characterized in more detail. The triple variant showed the best kinetic parameters (see also **Table 5**).

**Table 5 Kinetic parameters for the Mcd variants.**

| Mutation | e⁻ acceptor | $v_{max}$ (mU mg$^{-1}$) | apparent $K_M$ (µM) |
|---|---|---|---|
| **wt** [a] | ferrocenium | 6000 | 20 |
| **wt** | $O_2$ | No activity [b] | |
| **Thr317Gly** | $O_2$ | 27.7 ± 0.7 | 30.0 ± 2.5 |
| **Trp315Ala** | $O_2$ | No activity [b] | |
| **Trp315Phe** | $O_2$ | Low activity [b] | |
| **Glu377Asn** | $O_2$ | Low activity [b] | |
| **Thr317Gly, Glu377Asn** | $O_2$ | 10.4 ± 0.3 | 39.9 ± 3.3 |
| **Thr317Gly, Trp315Phe** | $O_2$ | 37.8 ± 2.0 | 7.6 ± 2.5 |
| **Thr317Gly, Glu377Asn, Trp315Phe** | $O_2$ | **97.3 ± 5.4** | **27.4 ± 5.0** |

[a] values were taken from Erb et al. 2009 (*94*). [b] Kinetic parameters were not determined, due to low product formation rate observed in the HPLC based analysis (**Fig. 16B**)

When we replaced Mcd by the engineered oxidase (Mco) in CETCH 2.0, fractional labeling of CoA-esters showed that the $^{13}CO_2$ is stepwise incorporated and the majority of propionyl-CoA was labelled once after 5 minutes and twice after 30 minutes. This indicates, that the majority of propionyl-CoA was already turn over twice in the pathway after 30 min. Although the total amount of intermediates in the cycle decreased over time, these experiments demonstrated that CETCH 2.0 could in principle turn several times (**Fig. 17**).

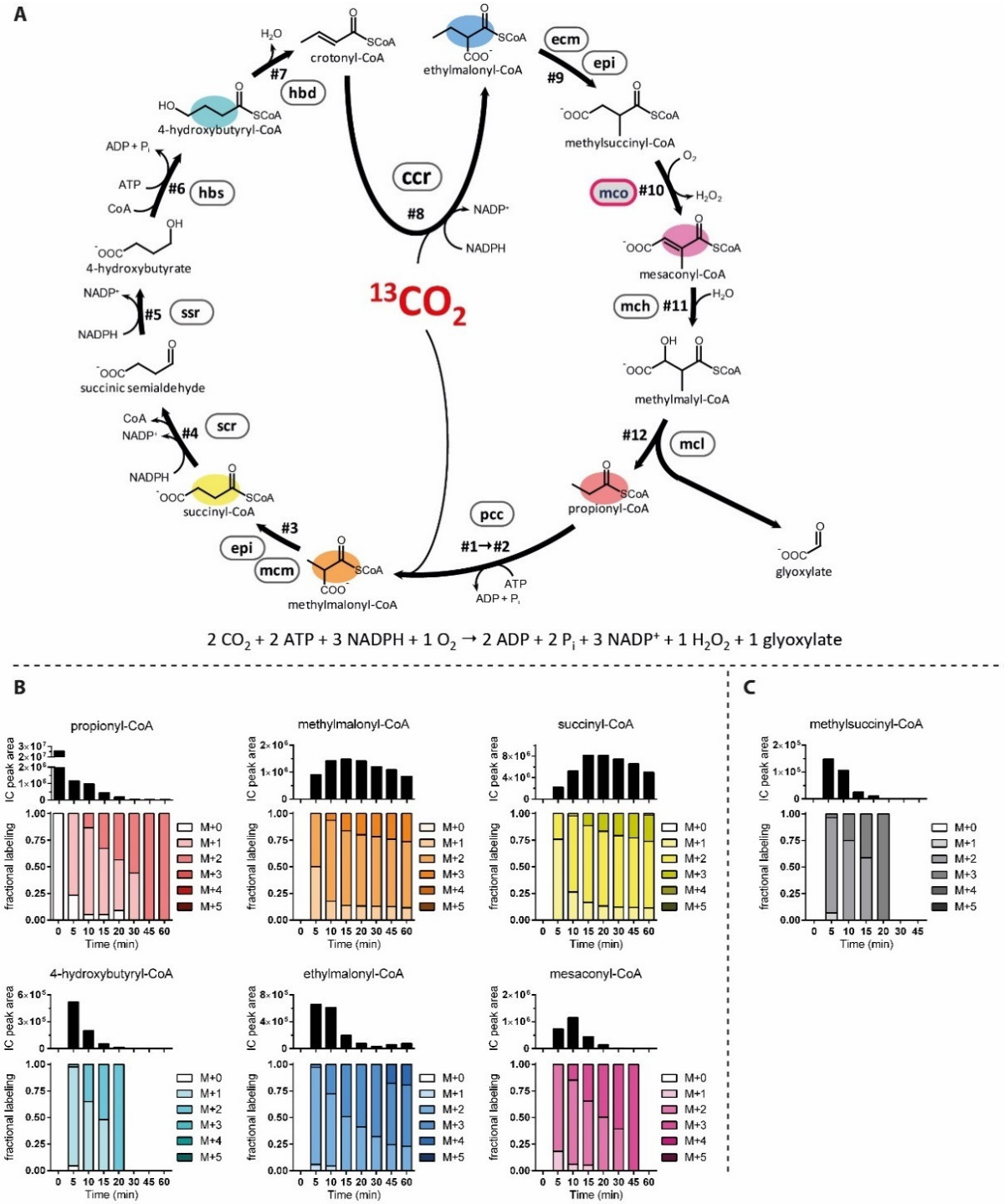



**Fig. 17 CETCH 2.0 (replacing Mcd by Mco allows multiple turning).** (A) Topology of the CETCH cycle version 2.0. Enzyme names are abbreviated according to **Table 4**. Mcd was replaced by Mco (boxed in red) (B) Dynamics of key intermediates of CETCH 2.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labeling pattern of six different CoA ester with color-coding according to (A). Replacing Mcd by Mco allowed to operate the cycle beyond rxn #10 and for multiple turns. (**C**) Level and fractional labeling pattern of methylsuccinyl-CoA (compared to CETCH 1.0, **Fig. 14**).

## 3.4. CETCH 3.0 (read-out module and NADPH regeneration)

For CETCH 3.0 we implemented a read-out module to directly quantify $CO_2$ fixation (**Fig. 18A**). Starting from propionyl-CoA, the CETCH cycle converts two molecules of $CO_2$ per turn into one molecule of glyoxylate. To measure glyoxylate formation we used a malate synthase (Mas) from *Escherichia coli*, which condenses glyoxylate with externally added acetyl-CoA to yield malate. Additionally, we introduced a NADPH regeneration module to continuously replenish the NADPH-pool (**Fig. 18A**). To that end, we added an engineered formate dehydrogenase from *Mycobacterium vaccae* (*98*), which could directly operate in the formate-based buffer system. HPLC-MS quantification of the malate produced from the CETCH 3.0 cycle, yielded 0.3 fixed $CO_2$-molecules per propionyl-CoA acceptor molecule over 90 min (**Fig. 22C**). CoA ester analysis showed that ethylmalonyl-CoA, methylmalonyl-CoA, malonyl-CoA and malyl-CoA accumulated over time in CETCH 3.0 (**Fig. 18B-C**). The accumulation of the two alkylmalonyl-CoA ester suggested that the coenzyme $B_{12}$-dependent mutase reactions in the cycle were affected. Buildup of malyl-CoA, which is not a direct intermediate of the CETCH cycle, indicated a side reaction of the acetyl-CoA pool with glyoxylate, presumably catalyzed by β-methylmalyl-CoA lyase (Mcl), which is known to possess this activity intrinsically (*96*). This side reaction leads to a loss of malate production, since the primary assimilation product (glyoxylate) is converted into a different output molecule. The accumulation of malonyl-CoA is presumably caused because propionyl-CoA carboxylase (Pcc) is able to carboxylate acetyl-CoA with a low activity (**Fig. 18**).

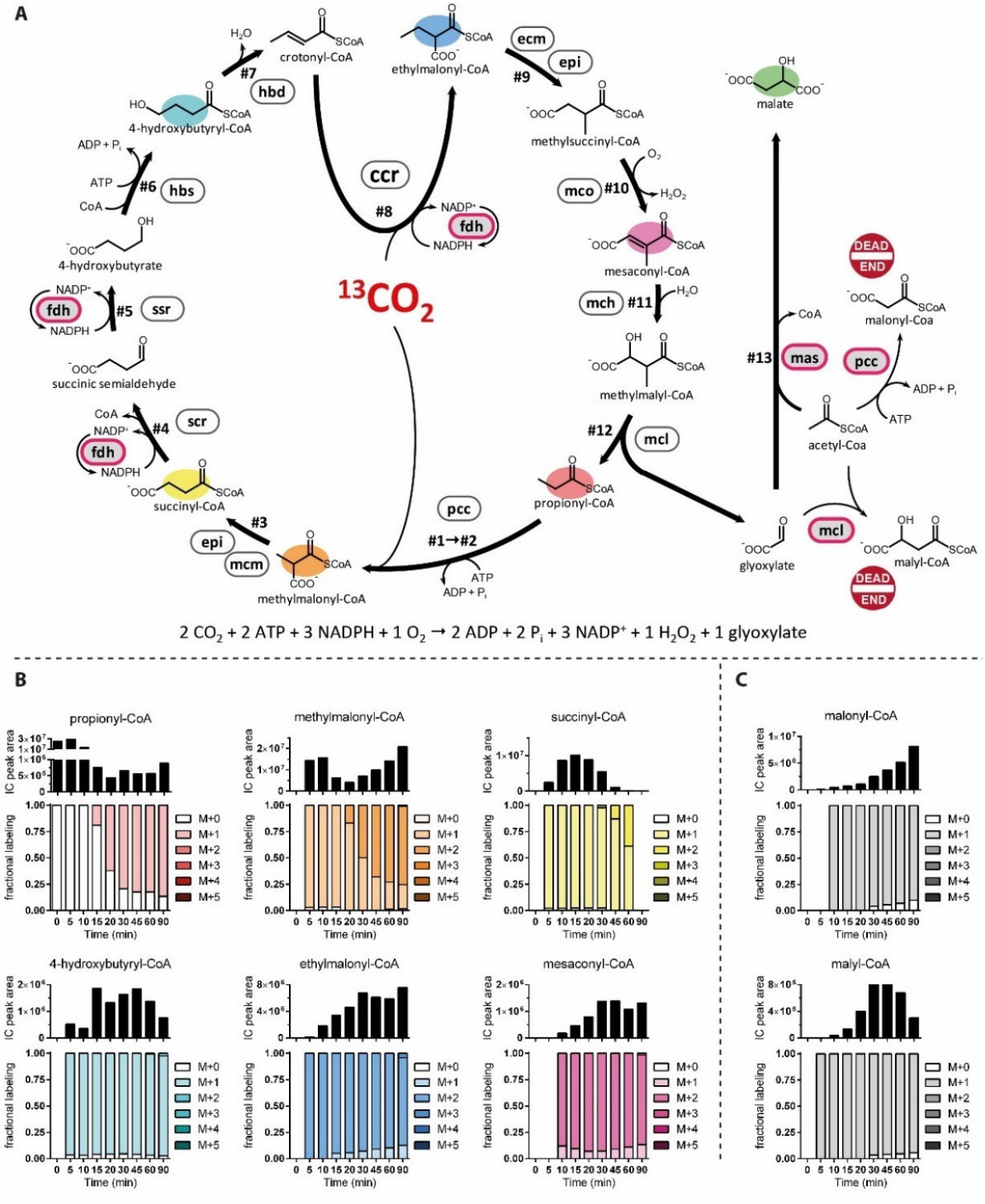



**Fig. 18 CETCH 3.0 (read-out module and NADPH regeneration).** (A) Topology of the CETCH cycle version 3.0. Enzyme names are abbreviated according to **Table 4**. The newly added enzymes Fdh (NADPH-regeneration) and Mas (read-out module) converting the primary $CO_2$ fixation product glyoxylate into malate are boxed in red. (B) Dynamics of key intermediates of CETCH 3.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA esters with color-coding according to (A). (C) Level and fractional labeling pattern of malyl-CoA and malonyl-CoA, which are dead-end metabolites that accumulated over time.

## 3.5. CETCH 4.0 (proof reading & protection from oxidative damage)

In CETCH 4.0, we took care of the unwanted malyl-CoA formation observed in CETCH 3.0. We added malyl-CoA thioesterase (Mct) from *Rhodobacter sphaeroides* that catalyzes the specific hydrolysis of malyl-CoA into malate and CoA (**Fig. 19A**). We reasoned that Mct would act as a metabolic proof reading-enzyme that frees the CoA bound malate deriving from the unwanted side reactivity of Mcl. We also addressed the mutase reactions that were apparently limiting in CETCH 3.0 by adding 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid and catalase (Kat) from *E. coli* to protect $B_{12}$, as well as other coenzymes, cofactors, intermediates and proteins from oxidative damage and radical oxygen species (such as $H_2O_2$, produced by Mco). With these changes, we could increase efficiency of CETCH 4.0 by a factor of five to 1.5 fixed $CO_2$ equivalents per acceptor molecule within 30 min (**Fig. 22C**), after which $CO_2$ fixation leveled off. CoA ester analysis still showed an accumulation of malonyl-CoA, which was accompanied by an unexpected high ATP-hydrolysis and acetyl-CoA consumption rate (**Fig. 19C**). We suspected that malonyl-CoA formation was caused by Pcc, which is known to accept acetyl-CoA besides propionyl-CoA (*105*). This side activity of Pcc apparently depleted the acetyl-CoA pool within short time and caused the unproductive hydrolysis of ATP, preventing a more efficient cycling of CETCH 4.0.

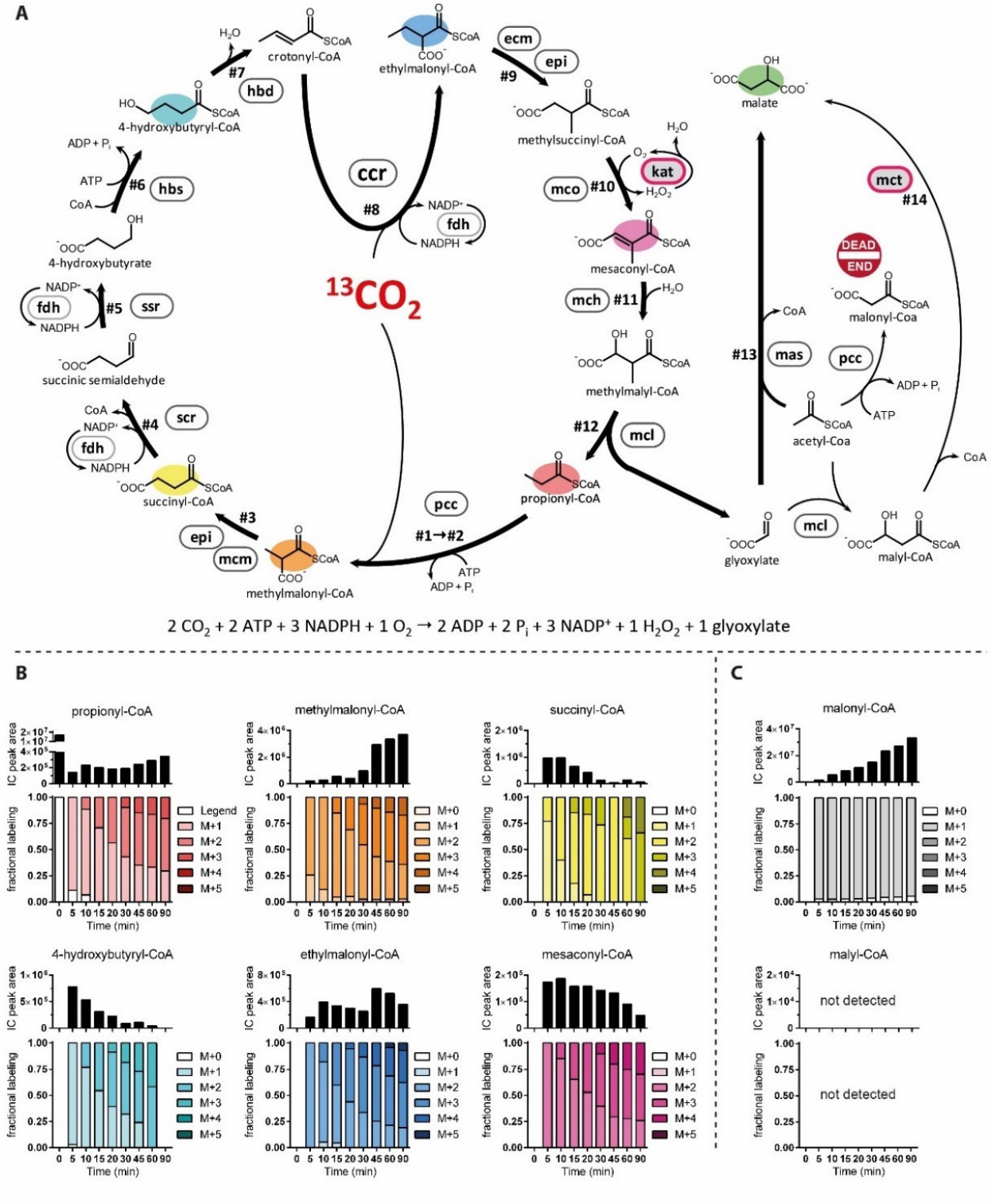



**Fig. 19 CETCH 4.0 (proof reading by Mct prevents malyl-CoA build-up)** (A) Topology of the CETCH cycle version 4.0. Enzyme names are abbreviated according to **Table 4**. The newly added enzymes Kat ($H_2O_2$ protection) and Mct (malyl-CoA proof-reading) are boxed in red. (B) Dynamics of key intermediates of CETCH 4.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA ester with color-coding according to (A). (C) Level and fractional labeling pattern of malyl-CoA and malonyl-CoA. While build-up of the malyl-CoA was efficiently prevented through application of Mct (compared to CETCH 3.0, **Fig. 18**), malonyl-CoA still accumulated over time.

## 3.6. CETCH 5.0 (reaction sequence redesign & addition of ATP regeneration)

For CETCH 5.0, we concentrated on eliminating the unwanted side production of malonyl-CoA, as well as maintaining a stable ATP-pool. The side reaction of Pcc with acetyl-CoA resulted not only in the production of malonyl-CoA, but more importantly lead to an undesired hydrolysis of ATP. First, we redesigned the core sequence of the CETCH cycle to become independent of the promiscuous Pcc. Realizing that Ccr also accepts acrylyl-CoA as non-native substrate we drafted an alternative reaction sequence that would allow us to form methylmalonyl-CoA from propionyl-CoA via reductive carboxylation of acrylyl-CoA (**Fig. 21A**).

So far, no enzyme is known that would specifically oxidize propionyl-CoA into acrylyl-CoA. We thus screened different oxidases for an activity with propionyl-CoA and identified a short-chain acyl-CoA oxidase 4 (Acx4) from *Arabidopsis thaliana* that catalyzed the oxidation of propionyl-CoA with a catalytic efficiency of $k_{cat}\,K_M^{-1}$ = 2.4·10$^6$ M$^{-1}$ s$^{-1}$ ($k_{cat}$ = 30.8 ± 1 s$^{-1}$, $K_M$ = 13 ± 1 µM). Unfortunately, the enzyme also oxidized 4-hydroxybutyryl-CoA, another intermediate of the CETCH cycle, with a similar catalytic efficiency of $k_{cat}\,K_M^{-1}$ = 6.9· 10$^5$ M$^{-1}$ s$^{-1}$ ($k_{cat}$ = 9.7 ± 0.2 s$^{-1}$; $K_M$ = 14 ± 1 µM). To decrease the unwanted side-reactivity of Acx4 with 4-hydroxybutyryl-CoA, we engineered the enzyme towards higher substrate specificity. The structure of Acx4 revealed an active site pocket, which is large enough to accept different short-chain acyl-CoA esters (*106*) (**Fig. 20A**). We argued, that mutations, decreasing the overall size of this active site pocket, would be sufficient to enable the enzyme to discriminate the short propionyl-CoA ($C_3$) from the longer 4-hydroxybutyryl-CoA ($C_4$ with an additional hydroxyl group). Therefore we targeted several residues lining the cavity. Eventually, a Thr134Leu mutation in Acx4 allowed us to lower the activity of the enzyme with 4-hydroxybutyryl-CoA by two orders of magnitude to a $k_{cat}\,K_M^{-1}$ = 1.3 · 10$^3$ M$^{-1}$ s$^{-1}$ ($k_{cat}$ = 0.3 ± 0.02 s$^{-1}$; $K_M$ = 239 ± 44 µM) while the reaction of the engineered enzyme with propionyl-CoA was only affected by a factor of ten at a $k_{cat}\,K_M^{-1}$ = 2.2 · 10$^5$ M$^{-1}$ s$^{-1}$ ($k_{cat}$ = 9.8 ± 0.4 s$^{-1}$; $K_M$ = 44 ± 5 µM). In summary, our engineered, propionyl-CoA oxidase (Pco) showed a strongly improved discrimination factor of 180 between the desired oxidation of propionyl-CoA and the unwanted side-reaction with 4-hydroxybutyryl-CoA (**Fig. 20** and **Table 6**). Note that Pco additionally serves a proof reading function in the CETCH cycle. Ccr possess a very low reduction reaction besides its reductive carboxylation reaction, which could lead to a loss of intermediates from the cycle over time (through the reduction of acrylyl-CoA to propionyl-CoA). Pco could work against this loss by recycling propionyl-CoA to acrylyl-CoA.

In assays with the Pco/Ccr combination instead of Pcc, formation of malonyl-CoA was not detectable anymore and ATP hydrolysis rate was strongly decreased, indicating the successful implementation of the Pco/Ccr bypass (**Fig. 21C**). However, we realized that the reaction of 4-hydroxybutyryl-CoA synthetase (Hbs) operates close to the thermodynamic equilibrium ($\Delta_rG'$ = 10 kJ mol$^{-1}$) and thus is very sensitive to changes in the ATP and ADP + $P_i$ concentrations. We reasoned that it might not be sufficient to decrease ATP hydrolysis, but to actively maintain a high ATP concentration in the assay to keep the thermodynamic driving force intact. We tested different ATP-regeneration/addition systems with the CETCH core sequence. Of the systems tested, a polyphosphate transferase from *Sinorhizobium meliloti (99)* worked most reliably. In summary, the Pco/Ccr combination together with the ATP-regeneration system increased the efficiency of the cycle to 4.3 $CO_2$-equivalents per acceptor molecule over a running time of 90 min (**Fig. 22C**).

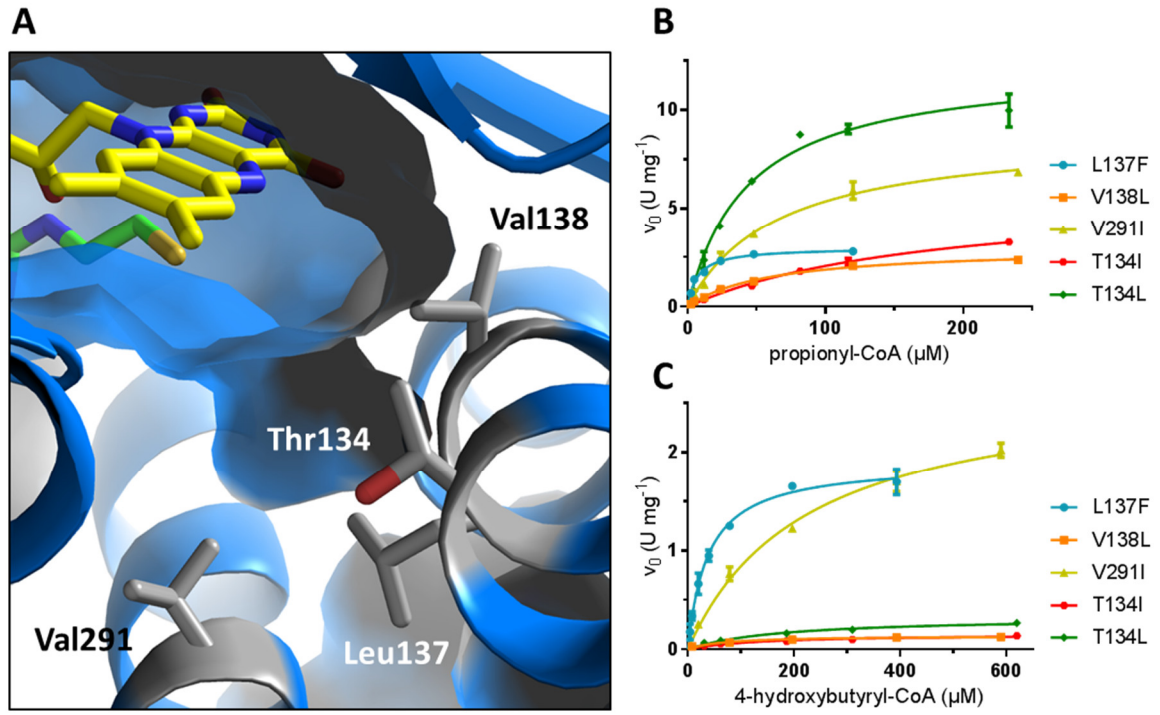



**Fig. 20 Engineering of Acx4 towards a specific propionyl-CoA oxidase (Pco).** The short-chain oxidase 4 (ACX4) from *A. thaliana* was engineered towards desired substrate (propionyl-CoA) specificity, because we detected activity with 4-hydroxybutyryl-CoA. (A) The active site structure (PDB ID: 2IX5) shows a large substrate binding pocket, which is surrounded by the residues Thr134, Thr137, Val138 and Val291 (*106*). These residues were mutated to more bulky ones with the intention to close the active site and enable the discrimination of the larger 4-hydroxybutyryl-CoA ($C_4$ with additional hydroxyl group) from the shorter propionyl-CoA ($C_3$) (B) Michaelis-Menten kinetics of all characterized variants with propionyl-CoA as substrate. (C) Michaelis-Menten kinetics of all characterized variants with 4-hydroxybutyryl-CoA as substrate.

**Table 6 Kinetic data of all Acx4 variants analyzed**.

| | **propiony-CoA** | | **4-hydroxybutyryl-CoA** | | |
|---|---|---|---|---|---|
| | $v_{max}$ (U mg$^{-1}$) | app. $K_M$ (µM) | $v_{max}$ (U mg$^{-1}$) | app. $K_M$ (µM) | $k_{cat}$ $K_M^{-1}$ (propionyl-CoA) / $k_{cat}$ $K_M^{-1}$ (4-hydroxybutyryl-CoA) |
| **wt** | **38.9 ± 1.0** | **12.5 ± 1.3** | **12.2 ± 0.2** | **13.6 ± 1.2** | **4** |
| **L137F** | 3.0 ± 0.1 | 7.0 ± 0.7 | 1.90 ± 0.05 | 38.0 ± 3.4 | 9 |
| **V138L** | 3.0 ± 0.1 | 59.2 ± 4.7 | 0.14 ±0.01 | 63 ± 20 | 23 |
| **V138I** | n.d. | | n.d. | | |
| **V291I** | 8.9 ± 0.3 | 65.4 ± 6.0 | 2.73 ± 0.14 | 225 ± 29 | 12 |
| **V291L** | n.d. | | n.d. | | |
| **V291F** | n.d. | | n.d. | | |
| **T134I** | 5.1 ± 0.3 | 148 ± 16 | 0.18 ± 0.01 | 182 ± 38 | 34 |
| **T134L** | **12.3 ± 0.5** | **43.7 ± 5.3** | **0.38 ± 0.03** | **239 ± 45** | **180** |

n.d.: no detectable activity

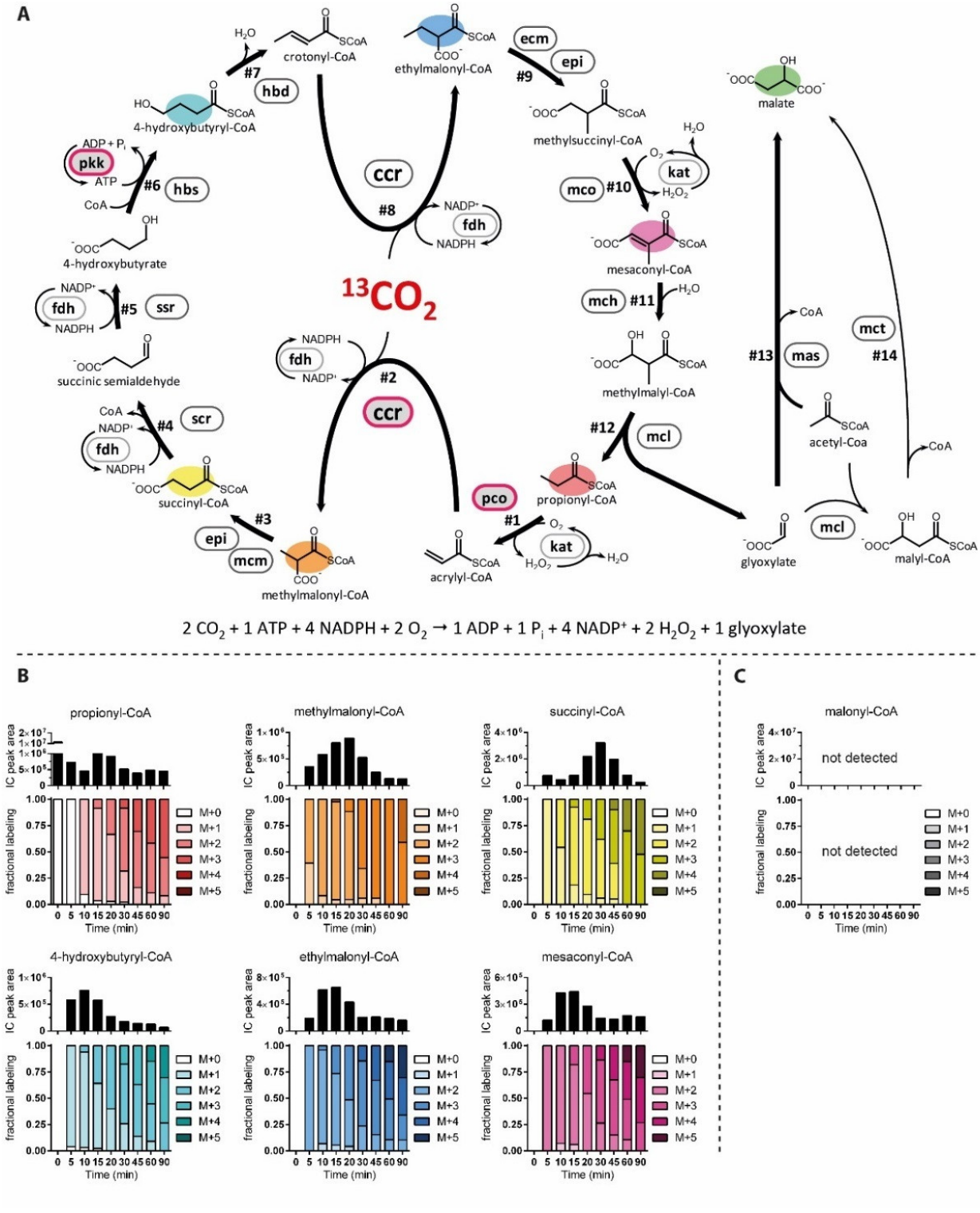



**Fig. 21 CETCH 5.0 (reaction sequence redesign & ATP regeneration).** (A) Topology of the CETCH cycle version 5.0. Enzyme names are abbreviated according to **Table 4**. The newly added enzymes, Pkk (ATP regeneration) and the Pco/Ccr combination (replacing promiscuous Pcc) are boxed in red. (B) Dynamics of key intermediates of CETCH 55.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA esters with color-coding according to (A). (C) Level and fractional labeling pattern of malonyl-CoA. The build-up of the dead-end metabolite malonyl-CoA (compare to CETCH 4.0, **Fig. 19**) was effectively prevented through these changes.

## 3.7. CETCH 5.1-5.4 (further improvements).

In further versions of the CETCH cycle (CETCH 5.1-5.4), we tested the influence of minor variations of the assay onto $CO_2$ fixation efficiency. In CETCH 5.1, we tested the influence of different Ccr homologs onto $CO_2$ fixation efficiency. We substituted a mutated Ccr used so far (*Methylobacterium extorquens;* Arg152Ala; $k_{cat}$ = 70 $s^{-1}$; $K_{M,crotonyl\text{-}CoA}$= 820 µM) by the wt enzyme from *M. extorquens* ($k_{cat}$ = 90 $s^{-1}$; $K_{M,crotonyl\text{-}CoA}$= 180 µM), a Ccr from *Kitasatospora setae* ($k_{cat}$ = 17 $s^{-1}$ ; $K_{M,crotonyl\text{-}CoA}$= 11 µM), as well as a Ccr from *Leptospira borgpetersenii* ($k_{cat}$= 57 $s^{-1}$ ; $K_{M,crotonyl\text{-}CoA}$= 65 µM). The use of either of these alternative Ccrs improved $CO_2$ fixation efficiency to about 5.1 fixed $CO_2$-equivalents and we used the *M. extorquens* wt enzyme in the following. In CETCH 5.2, we added 10 mM calcium to the assay, reasoning that calcium precipitates inorganic phosphate that accumulated over time and might interfere with the Hbs reaction, which operates close to the thermodynamic equilibrium (see above). However, addition of calcium did not increase $CO_2$ fixation efficiency. In CETCH 5.3, we exchanged the tris(hydroxymethyl)aminomethane (Tris) buffer used to store the purified proteins with a MOPS-based buffer system. We intended to reduce the final concentration of Tris in the assay, because we speculated that the Tris could react with the succinic semialdehyde intermediate (trapping the semialdehyde as Tris-succinic imine adduct). However, this alternation did not change $CO_2$ fixation efficiency. In CETCH 5.4, we exchanged the ammonium-formate reaction buffer used so far with a 100 mM MOPS KCl buffer (pH 7.5), which eventually increased efficiency of the cycle to a maximum of 5.4 fixed $CO_2$-equivalents per acceptor molecule over 90 min (**Fig. 22**).

## 3.8. CETCH 5.4 (additional features and limitations)

So far, all versions of the CETCH cycle were tested in buffer containing 50 mM $HCO_3^-$. $HCO_3^-$ was equilibrated with carbonic anhydrase to provide dissolved $CO_2$ in the assay. To investigate the ability of CETCH 5.4 to directly fix atmospheric $CO_2$, we quantified malate production rate of the cycle in presence and absence of $HCO_3^-$ in the buffer. Omitting $HCO_3^-$ from the buffer reduced the maximal malate production rate to approximately 73% (70% and 75%, n=2) compared to the positive control (*i.e.*, buffer with 50 mM $HCO_3^-$). This experiment showed that CETCH 5.4 is able to fix atmospheric $CO_2$ but suggests that under the chosen experimental conditions transfer of atmospheric $CO_2$ into solution becomes limiting. For CETCH 5.4, we also reinvestigated the requirement for catalase, coenzyme $B_{12}$ and ascorbic acid as protecting additives. We tested the influence of the three factors on malate production by running the cycle in the absence of each of these compounds. While omitting coenzyme $B_{12}$ and ascorbic acid did not show significant differences, the absence of catalase lowered malate production by about 55% (57% and 54%, n =2), indicating that an active catalase is important for an operational CETCH cycle. Finally, we turned our attention to the fact that malate production in CETCH 5.4 started to plateau around 90 min, reaching a final concentration of about 500-600 µM malate, questioning what factor limits conversion of $CO_2$ into malate. When we added a fresh batch of propionyl-CoA or ATP to a plateaued CETCH cycle, malate production by the cycle could not be restarted. Likewise, when we added the four enzymes that were suspected to be the most instable in the CETCH enzyme mix (*i.e.*, succinyl-CoA reductase, Scr; methylmalonyl-CoA mutase, Mcm; ethylmalonyl-CoA mutase, Ecm and 4-hydroxybutyryl-CoA synthetase, Hbs), to the enzyme assay at 15, 30, 60, and 90 min, production of malate could not be increased. Finally, and most interestingly, when we started a cycle in the presence of 500 µM malate, the cycle did not accumulate significant amounts of additional malate, while in the control experiment (without malate addition from the beginning), the cycle was operating normally. These experiments suggested that it is neither enzyme activity nor substrate supply that limits $CO_2$ fixation in version 5.4 of the CETCH cycle, but that it is rather product inhibition by the formed malate.

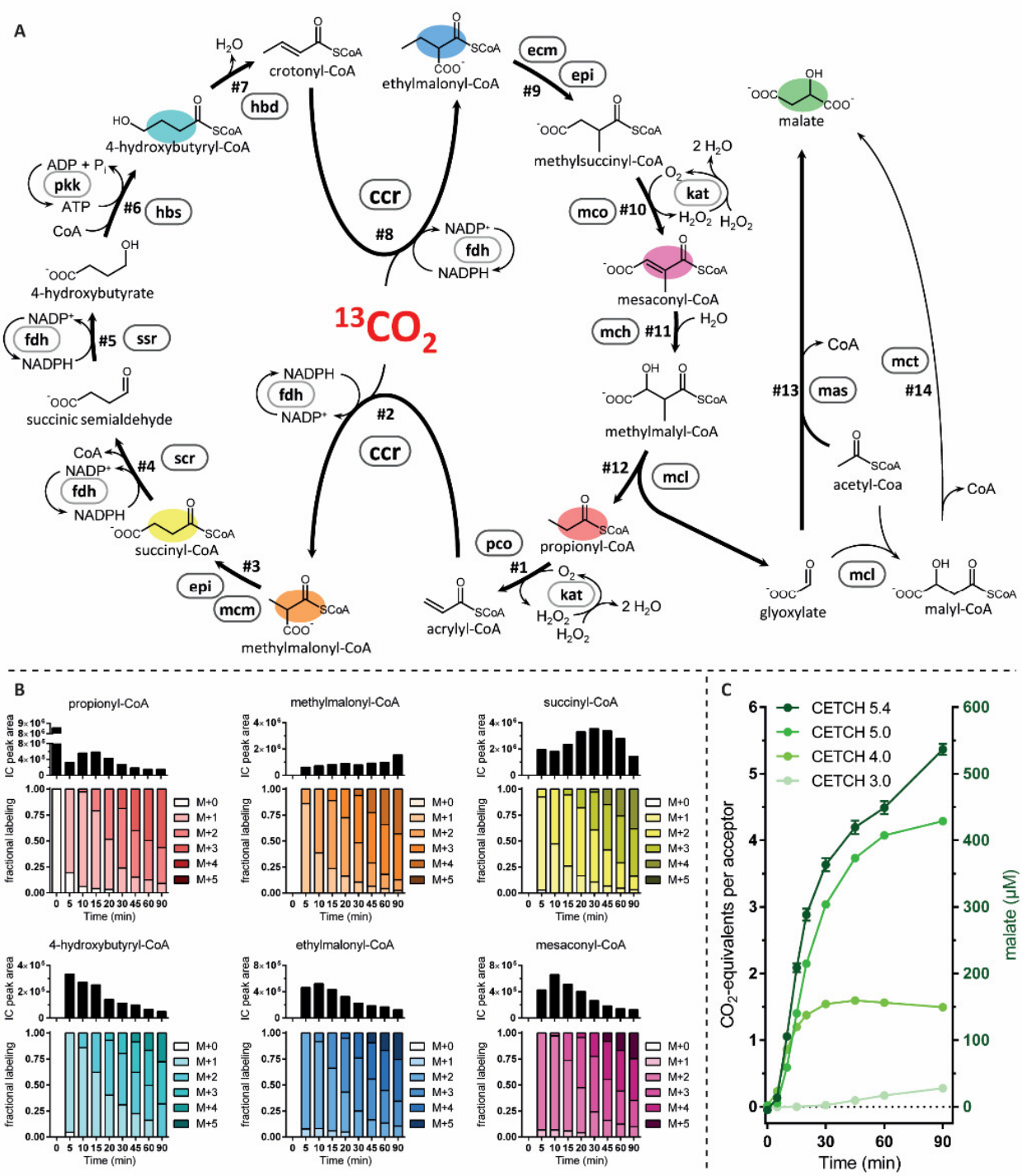



**Fig. 22 CETCH 5.4 (further improvements).** (A) Topology of the CETCH cycle (version 5.4), including proofreading and cofactor regenerating enzymes. Enzyme names are abbreviated according to **Table 4**. (B) Dynamics of key intermediates of CETCH 5.4 over 90 minutes. Shown are the levels of six different intermediates, as well as their fractional labeling patterns from the incorporation of $^{13}CO_2$ for each turn of the cycle. Intermediate coloring according to panel A. (C) Left x-axis: $CO_2$ fixation efficiency of the CETCH cycle over the course of its optimization. $CO_2$ fixation efficiency is defined as the $CO_2$-equivalents fixed per acceptor molecule in the cycle (*i.e.* starting amount of propionyl-CoA). Right x-axis: Absolute malate concentration formed over the course of 90 minutes in CETCH 5.4. The final assay (0.52 ml) contained 2.3 mg ml$^{-1}$ protein of cycle core enzymes plus 0.8 mg ml$^{-1}$ auxiliary enzymes and produced 540 µM malate over 90 min, which corresponds to 1080 µM fixed $CO_2$.

## 3.9. Conclusions

Over the course of the optimization, $CO_2$ fixation efficiency in the CETCH cycle improved by almost a factor of 20 until version 5.4 (**Fig. 22**). In its current form, CETCH 5.4 represents an *in vitro* enzyme network that is able to fix $CO_2$ at a rate of 5 $nmol^{-1}$ $mg^{-1}$ of core cycle protein. This is notably comparable to the very scarce literature on attempts to measure the complete CBB cycle *in vitro* (*107*).

The synthetic pathway comprises reaction sequences, which are also part of natural occurring pathways (*i.e*. ethylmalonyl-CoA pathway, 3HP/4HB cycle and anaerobic succinate degradation pathway). Nevertheless, with the approach of synthetic biology, we considered the vast diversity of enzymes in nature, which might not be found within these natural pathways.

In total, the 13 core reactions together with the readout module, auxiliary proofreading and cofactor regeneration processes are catalyzed by 17 enzymes from nine different organisms and all three domains of life, including bacteria, archaea, plants and humans. This also includes the engineering of three enzymes towards a desired reaction.

The combination of such diverse enzymes into a synthetic pathway poses the challenge that these enzymes face metabolites to which they were never exposed in their native metabolic context, causing undesired side-reactions. Inspired by biology we pursued different strategies to overcome this problem. We replaced the original reaction sequence to circumvent problematic reactions, applied enzyme engineering to minimize side reactions of promiscuous enzymes, and introduced metabolic proofreading enzymes to correct for the formation of side products. The first two strategies have been applied before in biotechnology and synthetic biology to improve productivity (*108, 109*). In contrast, the concept of metabolic proofreading has only been considered in few cases for synthetic pathway design (*77*), although it exists apparently also in naturally evolved pathways (*110*). We believe that this design principle should be considered more systematically for metabolic engineering in the future.

The redesign of the reaction sequence for the CETCH cycle from version 5.0 on, solely depends on the reductive carboxylation reaction of Ccr. This reallocation of the efficient carboxylase as key reaction in a $CO_2$-fixing pathway goes beyond simply improving or reshuffling naturally existing autotrophic $CO_2$ fixation enzymes and pathways. The first successful in vitro reconstitution of a synthetic enzymatic network for the conversion of $CO_2$ into organic products *in vitro* opens the door for several future applications (further discussed in chapter 6)

# 4. Realization of the HOPAC cycle

The HOPAC cycle is the second feasible candidate pathway resulting from our metabolic retrosynthesis approach. From the theoretical draft, we aimed at realizing this synthetic cycle *in vitro*. To that end we searched the available databases for enzymes that could sustain the theoretical HOPAC cycle (**Fig. 23**). We were able to identify one or more potential enzyme candidates for each reaction of the HOPAC cycle. These candidates were expressed and characterized to select a final set of 10 enzymes that were able to catalyze the proposed reaction sequence of the HOPAC cycle (**Table 7**).

**Table 7 List of enzymes considered, tested, characterized and used to realize the HOPAC cycle.**

| Rxn | Abb. | Enzyme full name | Name | Organism | $v_{max}$ (U mg$^{-1}$) | $K_M$ (mM) | Source of kinetic data |
|---|---|---|---|---|---|---|---|
| **1** | mPc | propionyl-CoA carboxylase | Pcc βD407I | *M. extorquens* | 1.1 | 0.9 | This work |
| **1** | Pcc | propionyl-CoA carboxylase | Pcc | *M. extorquens* | 0.1 | 1.3 | This work |
| **2** | Mcr | malonyl-CoA reductase | Mcr | *C. aurantiacus* | 10[c] | 0.03 | Huegler et al. (2002) (*29*) |
| **3** | Hps | 3-hydroxypropionyl-CoA synthetase | Nmar1309 | *N. maritimus* | 0.6 | 1.2 | Könneke et al. 2014 (*34*) |
| **3+4** | tPcs | propionyl-CoA synthase (truncated) | tPcs | *Erythrobacter* | 1.3[d] | 0.2[d] | Bernhardsgrütter et al. (2017) (*111*) |
| **3+4** | tPcs | propionyl-CoA synthase (truncated) | tPcs | *C. aurantiacus* | 4[c] [d] | 0.02[d] | Alber et. al. (2002) (*30*) |
| **4** | Ech | (*R*)-specific enoyl-CoA hydratase | PhaJ | *P. aeruginosa* | 6200[e] | 0.03[e] | Fukui et al. (1989) (*112*) |
| **5** | Ccr | crotonyl-CoA carboxylase/reductase | Ccr | *M. extorquens* | 16[f] | 0.5[f] | Erb et. al. (2009) (*38*) |
| **6** | Mcm | methylmalonyl-CoA mutase | Mcm | *R. sphaeroides* | 20 | 0.14 | Erb et al. (2009) (*92*) |
| **7** | Mcd | methylsuccinyl-CoA dehydrogenase | Mcd | *R. sphaeroides* | 0.3[g] | 0.14[g] | This work |
| **7** | Gdh | Glutaryl-CoA dehydrogenase | Gdh | *D. multivorans* | no activity [g] | | This work |
| **8** | Mch | mesaconyl-CoA hydratase | Mch | *R. sphaeroides* | 1500[h] | n.d. | Zarzycki et al. (2008) (*95*) |
| **8** | Hpd | 3-hydroxypropionyl-CoA dehydratase | Msed2001 | *M. sedula* | 150[i] | 0.06[i] | Teufel et. al. (2009) (*113*) |
| **9** | Mcl | malyl-CoA lyase | Mcl1 | *R. sphaeroides* | 5 | 0.01 | Erb et al. 2010 (*96*) |
| **10+13** | Smt | succinyl-CoA:malyl-CoA transferase | SmtAB | *C. aurantiacus* | 7.5[c] | 0.5 | Friedmann et al. (2006) (*114*) |
| **11** | Sdh | succinate dehydrogenase | SdhA | *E. coli* | n.d. | n.d. | |
| **12** | Fuh | fumarate hydratase | FumA | *E. coli* | 1900 | 0.5 | Lussenburg et. al. (2013) (*115*) |
| **14** | Etf | electron transfer flavoprotein | EtfAB | *R. sphaeroides* | n.d. | n.d. | |
| **15** | Fdh | formate dehydrogenase (D221A) | Fdh | *M. vaccae* | 1.4 | 0.37 | Hoelsch et al. (2013) (*98*) |
| **16** | Pkk | polyphosphate kinase | Pkk2 | *S. meliloti* | 12 | 0.032 | Nocek et al. (2008) (*99*) |
| **17** | Kat | katalase | KatE | *E. coli* | 11740 | 86.5 | Sevinc et al. (1995) (*100*) |
| **18** | Sod | superoxide dismutase | SodB | *E. coli* | n.d. | n.d. | |
| **15** | Gcl | glyoxylate carboligase [a] | Gcl | *E. coli* | 18 | 0.9 | Kaplun et al. (2008) (*101*) |
| **16** | Gar | tartronate semialdehyde reductase[a] | GarR | *E. coli* | 410[b] | 0.2[b] | Gotto et al. (1961) (*102*) |
| **17** | Glk | glycerate kinase [a] | GlxK | *E. coli* | 960 | 138 | This work |

The reaction numbering is according to (**Fig. 23, Fig. 29, Fig. 32, Fig. 33, Fig. 34** and **Fig. 35**). The reactions from 1 to 9 represent the core sequence of the HOPAC cycle. Highlighted in green are those enzymes used in the final version of HOPAC 4.1. n.d. not determined. [a] These enzymes would convert glyoxylate into phosphoglycerate, as a potential connection of the HOPAC cycle with central carbon metabolism. [b] Values for *Pseudomonas ovalis* enzymes. [c] Values determined at 55°C. [d] Values for the overall reaction of all three domains to propionyl-CoA. The reaction rate for the truncated (only synthetase and hydratase) of *C. aurantiacus* is 0.5 U mg$^{-1}$ at 45°C (*30*). [e] Values for crotonyl-CoA from *A. caviae*. [f] Values from *R. sphaeroides.* [g] Values for succinyl-CoA. [h] Values for mesaconyl-CoA. [i] Values determined at 65°C.

## 4.1. Initial considerations

To find suitable enzymes for all steps in the HOPAC cycle, we investigated certain reactions in more detail prior to the assembly of the whole cycle. The studies of these candidate enzymes are summarized in the following chapters (chapter 4.1.1 to 4.1.4)

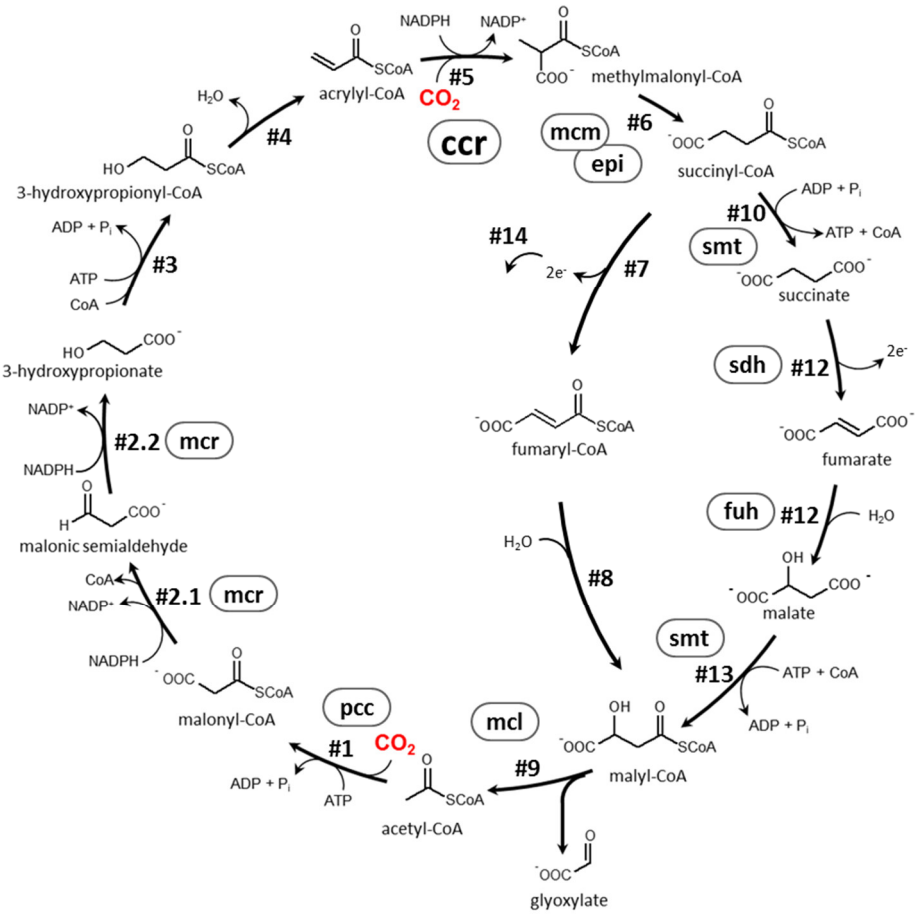



**Fig. 23 Topology of the HOPAC cycle from the initial design phase.** Enzyme names are abbreviated according to **Table 7.** The enzymes for reactions #3, #4, #7, #8 and #14 will be discussed in chapters 4.1.1 to 4.1.4. The reaction sequence includes an alternative route bypassing the membrane bound succinate dehydrogenase (Sdh, reaction #12). Instead, a direct oxidation of succinyl-CoA to fumaryl-CoA, followed by a hydration to malyl-CoA is shown (reactions #7 and #8).

### 4.1.1. Formation of 3-hydroxypropionyl-CoA

The ATP dependent formation of 3-hydroxypropionyl-CoA from 3-hydroxypropionate and CoA (**Fig. 23** reaction #3) is limited by the availability of suitable enzyme candidates. This reaction is found in autotrophic $CO_2$ fixation pathways of archaea (e.g. *Metallosphaera sedula* or *Nitrosopumilus maritimus*) or bacteria (*Chloroflexus aurantiacus*) (*30, 34, 95*). The propionyl-CoA synthase (Pcs) from *C. aurantiacus* consists of three domains (synthetase, dehydratase and reductase) converting 3-hydroxypropionate and CoA into propionyl-CoA with 3-hydroxypropionyl-CoA and acrylyl-CoA as intermediates. The Michaelis-Menten kinetics parameters for the synthetase domain were previously determined ($K_m$ for 3-hydroxypropionate of 15 µM and a $v_{max}$ of 4 U $mg^{-1}$ at 55°C) (*30*). This enzyme is of interest, because it was previously shown that the cleavage of the reductase domain creates a shortened Pcs producing and releasing acrylyl-CoA, thus combining two of the reactions in the HOPAC cycle (**Fig. 23** reaction #3 and #4) (*30, 116*). Recently, a homolog of the Pcs was found in *Erythrobacter*

*sp.* NAP1. The enzyme was characterized with favorable kinetic parameters compared to the Pcs from *C. aurantiacus* ($K_m$ for 3-hydroxypropionate of 200 µM and a $v_{max}$ of 1.3 U $mg^{-1}$ at 30°C) and the protein structure was solved (*111*). The information from the crystal structure and the previously described truncation site (from *C. aurantiacus*) were used to remove the reductase domain of the Pcs from *Erythrobacter*. Nevertheless, all these truncated enzymes (stop codon after Glu1428 or Val1214) were inactive. We then tried to cleave off the dehydratase domain as well, to obtain an AMP-forming 3-hydroxypropionyl-CoA synthetase (stop codon after Asn881 or Ser900). Again, these further truncated enzymes showed no activity.

Eventually, we decided to use the ADP-forming 3-hydroxypropionyl-CoA synthetase (Nmar1309) from *N. maritimus* for the development of the HOPAC cycle. This enzyme was described to perform under mesophilic conditions with an apparent $K_m$ for 3-hydroxypropionate of 1.2 ± 0.2 mM and a $v_{max}$ of 0.59 ± 0.03 U $mg^{-1}$(*34*). Additionally, the application of an ADP-forming synthetase reduces the ATP equivalents required for the reaction sequence, however, rendering the synthetase reaction reversible.

### 4.1.2. Oxidation of succinyl-CoA

The original design considered a reaction sequence including succinate as an intermediate, which is oxidized by the succinate dehydrogenase to fumarate (**Fig. 23** reaction #12). This enzyme is naturally part of the tricarboxylic acid (TCA) cycle and is comprised of four different subunits (SdhA, SdhB, SdhC, SdhD). SdhA contains the catalytic active site, whereas SdhB is responsible for the electron transfer and together with SdhC and SdhD for the reduction of ubiquinone. Moreover, the SdhC and SdhD are membrane bound (*117*). This complexity of the enzyme is not suitable for a cell-free development and testing of the HOPAC cycle. Nevertheless, the catalytic subunit (SdhA) was express and purified from *E. coli.* Although it was obtained in a soluble form, no activity with different electron acceptors could be observed.

Therefore, a new reaction sequence was designed to circumvent the membrane bound SdhA reaction. This solution involves the direct oxidation of succinyl-CoA to fumaryl-CoA, which is then subsequently hydrated to malyl-CoA (**Fig. 29** reactions #7 and #8). The oxidation reaction of succinyl-CoA is not naturally catalyzed by any known enzyme so far. Initially, we tested the non-decarboxylating glutaryl-CoA dehydrogenase from *Desulfococcous multivorans*, which catalyzes the oxidation of glutaryl-CoA ($C_5$ carboxylic acid CoA thioester) to glutaconyl-CoA (*118*). Unfortunately, this enzyme showed no activity with succinyl-CoA ($C_4$ carboxylic acid CoA thioester). Therefore, the best candidate for this reaction step was a (2*S*)-methylsuccinyl-CoA dehydrogenase (Mcd) found in alphaproteobacteria (*e.g. R. sphaeroides*, *P. denitrificans* or *M. extorquens*) and actinobacteria (*e.g. Streptomyces spp., Frankia spp.*). This enzyme specifically catalyzes the FAD dependent oxidation of methylsuccinyl-CoA to mesaconyl-CoA and was already used as an engineered version (methylsuccinyl-CoA oxidase, Mco) in the CETCH cycle (**Fig. 22** reaction #10). Even though the enzyme is highly specific towards methylsuccinyl-CoA, it also accepts succinyl-CoA with approximately 0.5% of the relative activity for the natural substrate (*94*). Mcd evolved towards a high specificity for the additional methyl group of methylsuccinyl-CoA to prevent the side reaction with the central carbon metabolism intermediate succinyl-CoA. We therefore focused on rational site-directed mutagenesis of the Mcd from *R. sphaeroides* (RsMcd) to change the substrate specificity and obtain a potential candidate enzyme for the oxidation of succinyl-CoA *in vitro*. The enzyme has a 22-34% identity to other acyl-CoA dehydrogenases (ACDs), if only theC-termianl ~390 amino acids are considered, because Mcd contains an unusual N-terminal extension of ~170 amino acids. Moreover, homology modeling of RsMcd onto the structure of the short-chain acyl-CoA dehydrogenase from *Homo sapiens* (38% amino acid identity,

68% coverage, PDB ID: 2VIG) showed no conclusive mechanism for substrate binding (especially for the carboxylic acid of succinyl-CoA) and did not allow for rational design of the active site. Therefore, we attempted to obtain a crystallize structure for the Mcd. We were able to solve the structure of a close homolog from *P. denitrificans* (PdMcd; 88% identity in the C-terminal catalytic domain to RsMcd) at a resolution of 1.37 Å. Although, PdMcd was crystallized in the presence of mesaconyl-CoA only FAD was found to be bound in the active site; theobserved weak electron density did not allow the modelling of a CoA ester in the active site (see chapter 5).

Nevertheless, the solved structure allowed rational active site engineering on RsMcd, to increase the desired side activity towards succinyl-CoA as substrate. The difference between the natural and the desired substrate is the methyl branch at the C2 position. Although the missing methyl group in succinyl-CoA poses less steric hindrance, the wild type enzyme is able to efficiently distinguish between succinyl-CoA and (2*S*)-methylsuccinyl-CoA. An alanine at position 282 (285 in PdMcd) allows for the accommodation of the C2-methyl group of modeled fitted methylsuccinyl-CoA molecule (**Fig. 24A**). In comparison, the isobutyryl-CoA dehydrogenase from *H. sapiens* (PDB ID: 2RX0) also forms a cavity between Tyr136 and Ile103 to accommodate the branched C2-methyl group (**Fig. 24B**). This enzyme shows again reduced or no activity for the unbranched acyl-CoA esters (propionyl-CoA and butyryl-CoA respectively). It is assumed, that the wider cavity allows unbranched substrates to adapt a non-productive conformation (*119*). Closing this cavity with larger residues at this position could force the non-branched substrate to bind in a more productive conformation. In comparison, the glutaryl-CoA dehydrogenase from *D. multivorans* (PDB ID: 3MPI) catalyzes the oxidation of the unbranched glutaryl-CoA ($C_5$ carboxylic acid) and the cavity at the C2 position is occupied by the more steric Phe126. Therefore, we mutated the Ala282 of the RsMcd (Ala285 in PdMcd) to bulkier non-polar residues (Phe, Leu, Ileu and Val). Only the Ala282Val variant showed slightly improved product formation rates (**Fig. 25**), whereas all other variants were inactive (see also chapter 5).

At this point, we decided that further engineering of Mcd would be required, to obtain a more efficient enzyme for the oxidation of succinyl-CoA. In turn, we applied the wild type enzyme from *R. sphaeroides* for the initial realization of the HOPAC cycle, despite its low activity. This enzyme could be replaced in the future by an advanced engineered and more active succinyl-CoA dehydrogenase.

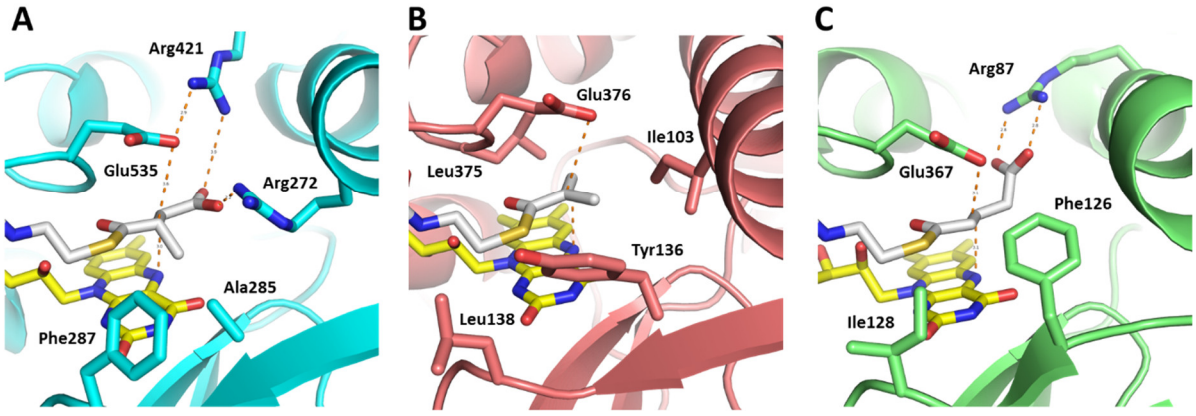



**Fig. 24 Active site structure of different acyl-CoA dehydrogenases.** (A) Active site of methylsuccinyl-CoA dehydrogenase from *P. denitrificans*. The (2*S*)-methylsuccinyl-CoA is modeled into the structure according to glutaryl-CoA in panel (C). The C2 methyl group is positioned in the cavity formed by Ala285 (between Phe287 and Arg272). (B) Active site of isobutyryl-CoA dehydrogenase from *H. sapiens* (PDB ID: 2RX0). The C2 methyl group of the product is positioned in the cavity between Tyr136 and Ile103. (C) Active site of glutaryl-CoA dehydrogenase (PDB ID: 3MPI). The enzyme catalyzes the oxidation of glutaryl-CoA (unbranched $C_5$ carboxylic acid CoA ester, compared to branched C4 carboxylic acid CoA ester in Mcd) and at the C2 position of the substrate, the Phe126 prevents the formation of a cavity and the Arg87 forms a bidentate hydrogen bond to the carboxylic acid of glutaryl-CoA.

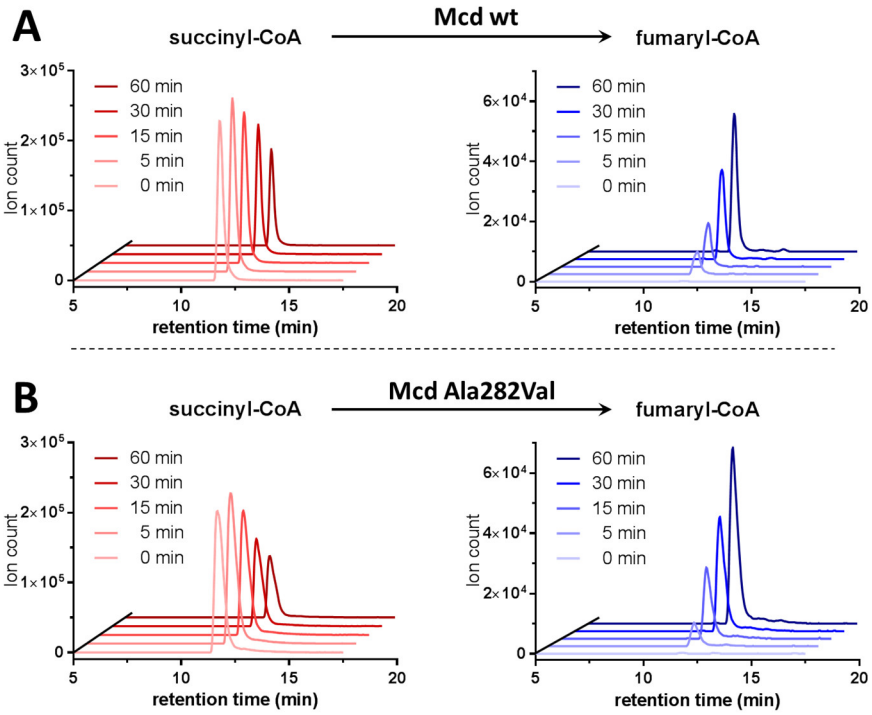



**Fig. 25 Fumaryl-CoA formation by either wild type or mutated Mcd (Ala282Val) from *R. sphaeroides***. The assay was performed with ferrocenium as electron acceptor (1.8 mM) and the enzyme concentration was 2 µM. (A) HPLC-MS analysis for the fumaryl-CoA formation of the wild type RsMcd. (B) HPLC-MS analysis of the Ala282Val variant of Mcd for the formation of fumaryl-CoA.

### 4.1.3. Electron acceptor for the oxidation of succinyl-CoA

Since the reaction catalyzed by Mcd is an oxidation, an electron acceptor was required for the *in vitro* conditions. As discussed for the CETCH cycle, artificial electron acceptors are not well suited for a complex *in vitro* system. In turn, Mcd was engineered into an oxidase (Mco) to accept molecular oxygen as terminal electron acceptor. We tested Mco (Mcd Trp315Phe, Thr317Gly and Glu377Asn) for the conversion of succinyl-CoA to fumaryl-CoA with oxygen as electron acceptor. We also included the mutation of Thr317Gly, which was crucial for implementing basal oxidase activity into the dehydrogenase. Although, the conversion of succinyl-CoA by Mco with oxygen was detected, the product formation over time was decreased to an insufficient level (**Fig. 26**).

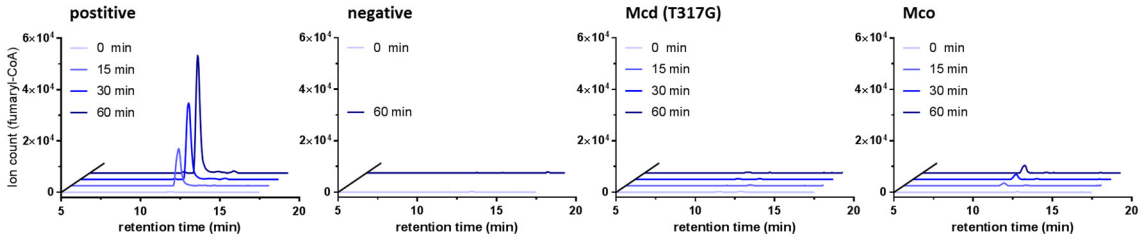



**Fig. 26 HPLC-MS chromatograms for the production of fumaryl-CoA (m/z=866) from succinyl-CoA**. For the positive control, ferrocenium hexafluorophosphate was used as electron acceptor, whereas none was used in the negative control. Mcd (Thr319Gly) contains a crucial mutation to implement a basal oxidase activity into Mcd. Mco is the engineered version of Mcd from the CETCH cycle and shows a reduced product formation rate compared to the wild type enzyme.

The natural electron acceptor of Mcd is an electron transfer flavoprotein (ETF), which shuttles the electrons from the oxidation reaction to the membrane-bound respiratory chain to conserve the energy. ETFs are soluble heterodimeric FAD-containing proteins and act as single electron acceptors for different acyl-CoA dehydrogenases (*120*). It has been shown, that ETFs interact naturally with molecular oxygen and are a source of superoxide and hydrogen peroxide in solution (*121*). Hence, ETF was a potential electron acceptor for the oxidation of succinyl-CoA by Mcd and could additionally be reoxidized by molecular oxygen. The produced superoxide and hydrogen peroxide could be removed by superoxide dismutase (Sod) and catalase (Kat). We were able to show, that ETF from *R. sphaeroides* was able to act as an electron acceptor for the oxidation of succinyl-CoA and the product formation was dependent on the ETF concentration (**Fig. 27A**). This indicated that the reoxidation of ETF by molecular oxygen was likely again a limiting factor in the overall reaction. It has been reported that the production rate for reactive oxygen species (ROS) by ETF was increased in clinically relevant variants, which implied a faster recycling of the electron acceptor. The first interesting variant comprised an aspartate to asparagine mutation, which affected the overall stability of the protein. The second was an arginine to cysteine mutation at the interface between the dehydrogenase and ETF, which probably destabilized the complex formation between ETF and the acyl-CoA dehydrogenase leaving ETF soluble but also exposed to dissolved oxygen (*121*). We introduced these two single mutations (βAsp125Asn or βArg188Cys) into ETF from *R. sphaeroides* and analyzed the product formation by Mcd (500 µM succinyl-CoA, 50 µM ETF, 5 µM Mcd) with 200 mM NaCl and 200 mM KCl salt concentration (to destabilize complex formation further). The βArg188Cys mutation strongly reduced the capability to accept electrons from the succinyl-CoA oxidation and no significant difference was observed for the βAsp125Asn mutation (**Fig. 27B**). Conclusively, further investigations into the reoxidation rate of ETF are required and alternative mutations might increase the reaction rate with molecular oxygen. Consequently, we applied the ETF wild type enzyme as an electron acceptor for the initial development of the HOPAC cycle.

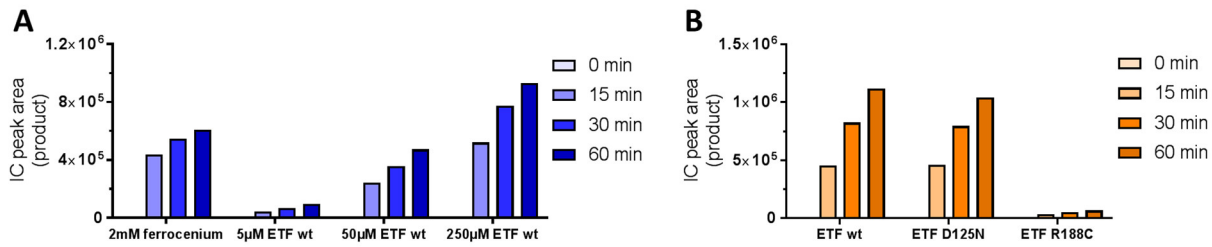



**Fig. 27 Assay of Mcd with succinyl-CoA as substrate and ETF as electron acceptor**. (A) The concentration of ETF was titrated (using 5µM Mcd and 500µM succinyl-CoA) and the product formation was observed via HPLC-MS. The product formation is dependent on the concentration of ETF in the assay. (B) Different mutations for ETF were tested for their effect on the production of fumaryl-CoA. Whereas, the Asp125 is an ETF internal mutation, Arg188 is found at the interface of ETF and dehydrogenase. The conditions applied in this assay were: 500 µM succinyl-CoA, 50 µM ETF, 5 µM Mcd, 200 mM NaCl and 200 mM KCl. Whereas the Arg188Cys mutation decreased product formation titer and rate, the Asp125Asn mutation had no effect under these conditions.

### 4.1.4. Fumaryl-CoA hydratase reaction (formation of malyl-CoA)

The reactions sequence of the HOPAC cycle (including the alternative route via fumaryl-CoA) contains the dehydration of 3-hydroxypropionyl-CoA and the hydration of fumaryl-CoA (**Fig. 23** reaction #4 and #8). These reversible reactions could possibly be catalyzed by a single enzyme, but the two substrates are very distinct requiring a promiscuous active site. Fumaryl-CoA is a compound, which is not common in any natural metabolism and we tested its conversion to malyl-CoA using mesaconyl-CoA hydratase (Mch) from *R. sphaeroides*. This enzyme naturally accepts the C2 branched mesaconyl-CoA (product of Mcd from methylsuccinyl-CoA). We discovered that Mch was able to accept the unbranched fumaryl-CoA and convert it to malyl-CoA (**Fig. 28A**). In turn, we tested Mch for the dehydration of 3-hydroxypropionyl-CoA to acrylyl-CoA, but the enzyme showed no activity towards this substrate. *Vice versa*, the (*R*)-specific enoyl-CoA hydratase (Ech) from *P. aeruginosa* was able to dehydrate 3-hydroxypropionyl-CoA but shows no activity with the non-natural fumaryl-CoA (**Fig. 28B**). Conclusively, our attempt to perform the two hydration/dehydration reactions in the HOPAC cycle by a single enzyme was not successful. Nevertheless, we discovered suitable candidates for both reactions.

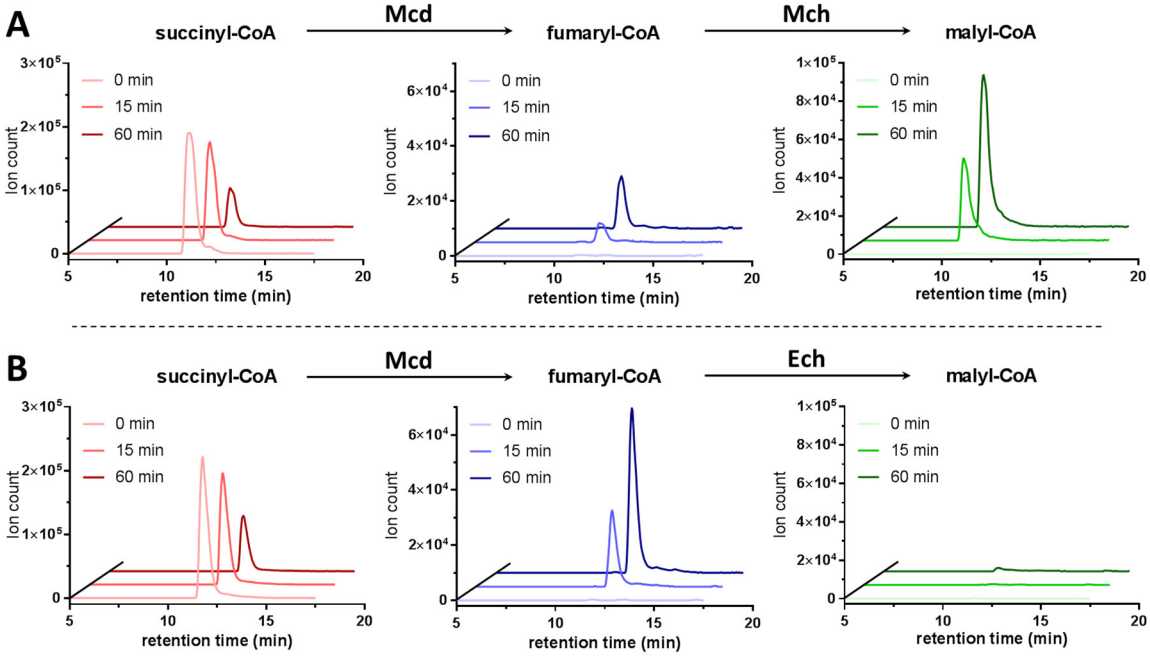



**Fig. 28 Assay of different hydratases in combination with the oxidation of succinyl-CoA.** (A) The mesaconyl-CoA hydratase (Mch) from *R. sphaeroides* is able to accept fumaryl-CoA as substrate and catalyzed the hydration to malyl-CoA. (B) On the contrary, the (*R*)-specific enoyl-CoA hydratase (Ech) from *P.* aeruginosa, which is used in the HOPAC cycle for the dehydration of 3-hydroxypropionyl-CoA shows no malyl-CoA formation from fumaryl-CoA.

## 4.2. HOPAC 1.0 (Pcc prevents cycling)

After optimizing individual steps of the HOPAC cycle, we combined these results and tested the complete reaction sequence in a cell free environment. In the first version of the HOPAC cycle (HOPAC 1.0) (**Fig. 29**), we provided all enzymes, cofactors (NADPH, ATP, CoA) at the beginning and started the assay by addition of 500µM acetyl-CoA. We used $^{13}C$ bicarbonate ($NaH^{13}CO_3$) as a source of labeled $^{13}CO_2$, because the incorporation into the intermediates allows to follow the number of turns of the cycle (**Fig. 30**). However, under these conditions we could not observe any conversion of the initially supplied acetyl-CoA nor the incorporation of $^{13}CO_2$ into intermediates by HPLC-MS. This indicated that the initial carboxylation of acetyl-CoA using Pcc from *M. extorquens* was a bottleneck in the reaction sequence.

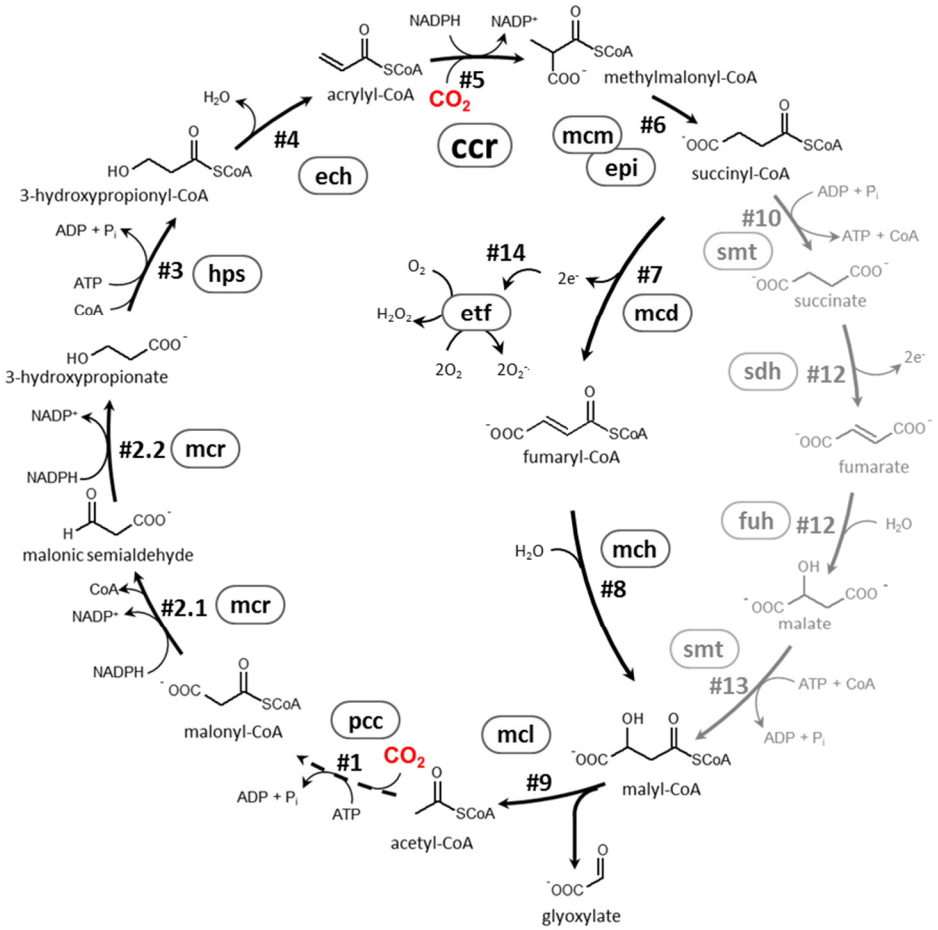



**Fig. 29 HOPAC 1.0 (Pcc prevents cycling).** Topology of the HOPAC 1.0 cycle. Enzyme names are abbreviated according to **Table 7.** The reaction sequence was altered from the initial design, bypassing the membrane bound succinate dehydrogenase (Sdh, reaction #12). Instead, a direct oxidation of succinyl-CoA to fumaryl-CoA (Mcd, reaction #7) followed by a hydration to malyl-CoA by mesaconyl-CoA hydratase (Mch, reaction #8) was applied. Pcc was insufficiently active with the provided starting substrate (acetyl-CoA) and prevented the operation of the $CO_2$ fixation cycle.

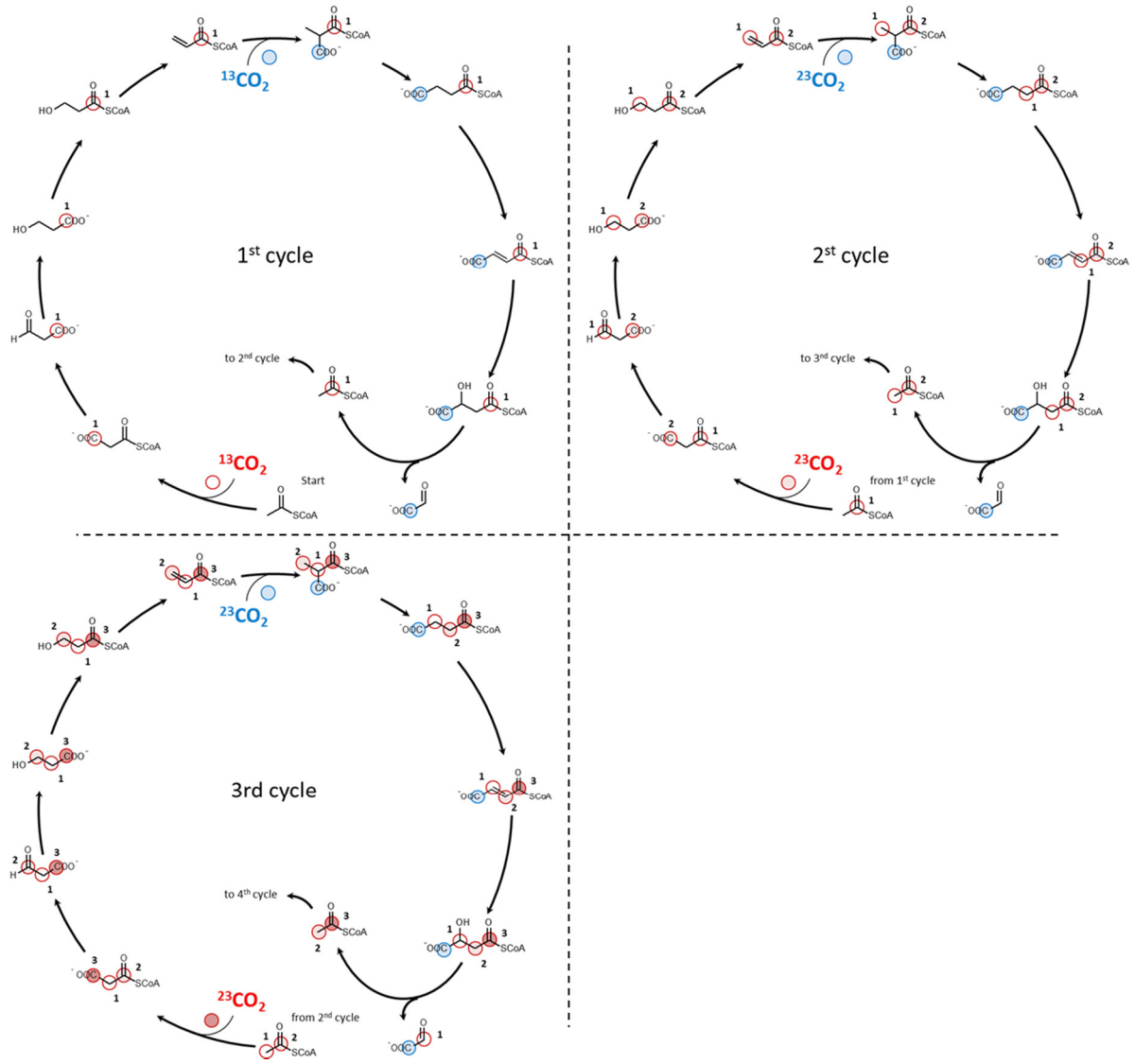



**Fig. 30 Expected $^{13}$C-label incorporation of the HOPAC cycle from $^{13}CO_2$.** Shown is the expected labeling pattern for each turn of the HOPAC cycle. The cycle features two $CO_2$-incorporation steps per turn shown in blue and red dots. The carbon incorporated into the $C_3$-acceptor acrylyl-CoA is shown in blue. This carbon ends up in the output molecule glyoxylate during each turn. The carbon incorporated into the $C_2$-acceptor acetyl-CoA is shown in red. This label stays with the acceptor molecule during each turn of the cycle (indicated by a darker red color). Note that the HOPAC cycle includes a carbon skeleton-rearranging step catalyzed by a mutase, resulting in a complex labeling pattern.

## 4.3. HOPAC 2.0 (Pcc variant mPc allows cycling)

In HOPAC 2.0 we aimed at overcoming the limitations caused by the insufficient activity of Pcc with acetyl-CoA, by analyzing naturally occurring acetyl-CoA carboxylases (Acc). Accs can be found in most living organisms (including archaea, bacteria, yeast, fungi, plants animals and humans) and are essential for the initial step of the fatty acid biosynthesis or the production of precursors for polyketide synthesis. The catalyzed two-step reaction is initiated by the biotin carboxylase (BC) followed by a transfer of the carboxylated biotin by the biotin carboxyl carrier protein (BCCP). The product is eventually formed by the carboxyl transferase (CT) domain, which consist of and α and β subunit. Acetyl-CoA carboxylases are well conserved among different organisms, but the organization of all different subunits is highly variable (**Fig. 31A**). Most eukaryotic Accs form a single, large (>200 kDa and >2000 amino acids) and multi-domain enzyme on a single polypeptide. Bacterial Accs consist of

multiple smaller subunits (up to 4 individually expressed genes). However, the holoenzyme complex is often unstable and the subunits tend to dissociate under various conditions (*122*). The archaeal Acc from *M. sedula* consist of three subunits and each domain (BC, BCCP and CT) forms an individual polypeptide. These proteins can be heterologously produced in *E. coli* and the holoenzyme is stable, but the specific activity at 30°C is only around 100 mU $mg^{-1}$ (*123*).

The Acc from *S. coelicolor* consists of only two subunits, because the BC and BCCP domains are combined to a single polypeptide. This enzyme requires an additional protein for the complex stability and to obtain the maximal activity of 490 mU $mg^{-1}$ with acetyl-CoA (*124*). Interestingly, this carboxylase shares the BC/BCCP subunit with a Pcc. In this case, the CT subunit from the Acc is exchanged with a propionyl-CoA specific CT (*125*). The main difference in the active site between these two CT domains is an aspartate to isoleucine substitution, determining the substrate specificity. Additionally, a single aspartate to isoleucine mutation (or *vice versa*) allows the inversion of substrate specificity within these two different CT domains (*126*). Accordingly, we decided to implement the same mutation at the homologous position in the CT domain of Pcc from *M. extorquens* (Asp407Ile). This enzyme has already been applied during the development of the CETCH cycle and can be heterologously produced in *E. coli*. Additionally, it consists of only two subunits (comparable to the Acc from *S. coelicolor*) and forms a stable holoenzyme complex. The mutation in this Pcc increased the catalytic efficiency for acetyl-CoA more than twelve times (wild type: 2.3 $10^2$ $s^{-1}$ $M^{-1}$; Asp407Ile: 2.7 $10^3$ $s^{-1}$ $M^{-1}$) to a $v_{max}$ of 1.15 ± 0.03 U $mg^{-1}$ and a $K_M$ for acetyl-CoA of 0.92 ± 0.06 mM (**Fig. 31B**).

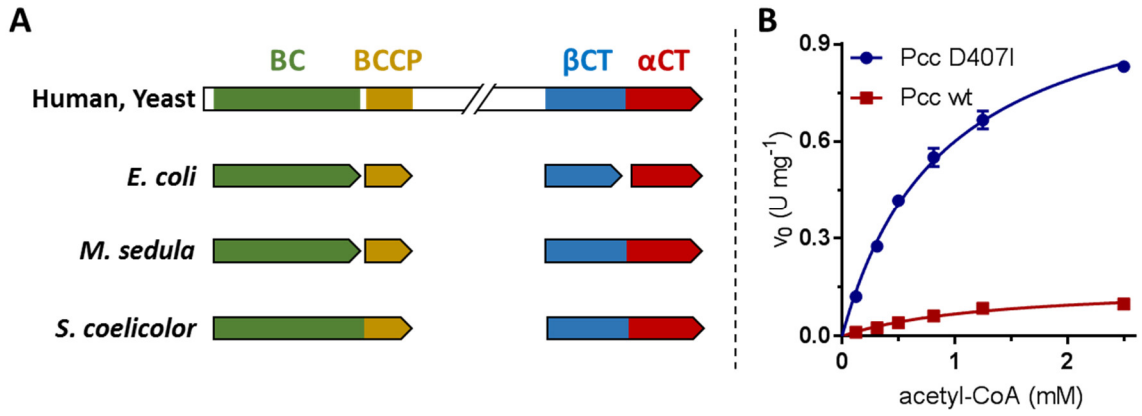



**Fig. 31 Comparison of different acetyl-CoA carboxylases.** (A) Domain organization of Accs from different organisms. Each arrow represents a single poly peptide chain. Adapted from (*122*). Accs consist of four domains: biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP) and the carboxyl transferase (comprising an α- and β-subunit). The Pcc from M. extorquens consists of two subunits, comparable to the Acc from *S. coelicolor*. (B) Michaelis-Menten kinetic of Pcc wt and the Asp407Ile variant from *M. extorquens* with acetyl-CoA as substrate. The variant shows an increased efficiency with acetyl-CoA.

Applying this Pcc variant (mPc) from *M. extorquens*, we tested the second version of the HOPAC (HOPAC 2.0) cycle and incorporation of $^{13}CO_2$ was observed after providing acetyl-CoA as starting substrate (250 µM). Fractional labeling of CoA-esters showed that the majority of acetyl-CoA was labelled once after 10 minutes and methylmalonyl-CoA is labeled up to 3 times indicating that some intermediates passed the cycle 1 and 1/2 times within 90 minutes. Although the total amount of intermediates in the cycle decreased over time, these experiments demonstrated that HOPAC 2.0 could in principle turn several times. (**Fig. 32**).

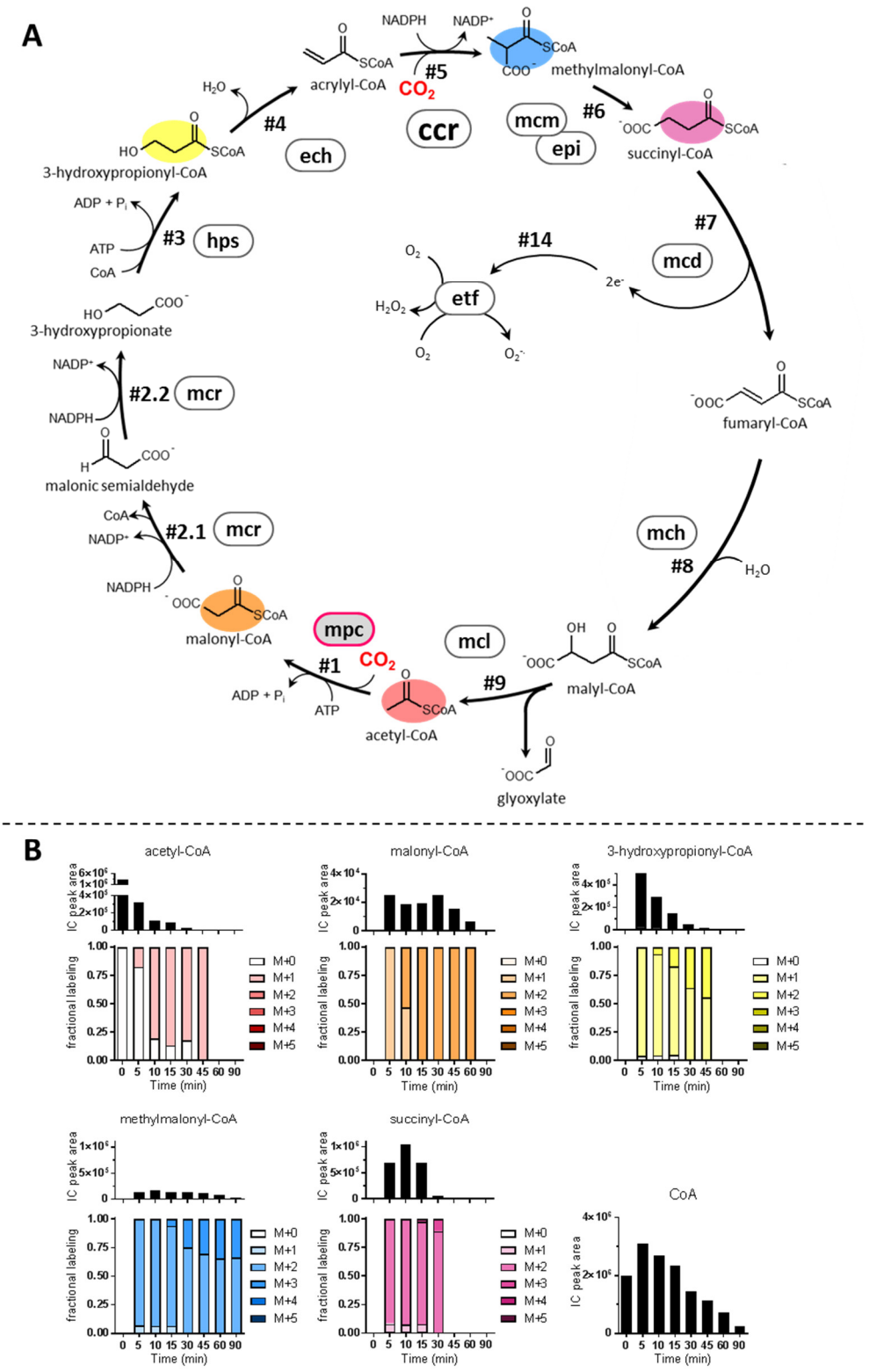



**Fig. 32 HOPAC 2.0 (Pcc variant mPc allows cycling)** (A) Topology of the HOPAC cycle version 2.0. Enzyme names are abbreviated according to **Table 7**. The Asp407Ile mutation in Pcc from *M. extorquens* (enzyme boxed in red) increases the activity towards acetyl-CoA sufficiently to allow for $^{13}CO_2$ incorporation by the pathway. (B) Dynamics of key intermediates of HOPAC 2.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA ester with color-coding according to (A).

## 4.4. HOPAC 3.0 (ATP and NADPH regeneration)

The recycling of cofactors (ATP and NADPH) was an important improvement in the CETCH cycle and was therefore also considered for the HOPAC cycle. Especially the reversible 3-hydroxypropionyl-CoA synthetase (Hps) reaction is highly sensitive to changes in the ATP and ADP + $P_i$ ratio, since the reaction operates close to the thermodynamic equilibrium ($\Delta_rG'$ = 11 kJ $mol^{-1}$). In HOPAC 3.0, we incorporated the regeneration systems involving polyphosphate kinase (Pkk) and formate dehydrogenase (Fdh) to provide a constant recycling of the ATP and NADPH cofactors.

In this version, the incorporation of $^{13}CO_2$ showed again that acetyl-CoA is labeled once after 10 min, but also incorporated a second label after 30 min. This indicates, that some intermediates passed the cycle already two times, in contrast to HOPAC 2.0 (**Fig. 33**). In HOPAC 3.0 we also observed an accumulation of methylmalonyl-CoA. The ATP-dependent reaction of Hps was not limiting anymore and the methylmalonyl-CoA mutase (Mcm) became a bottleneck in this version of the pathway. This indicates that Mcm was inactivated over time, since the radical based reaction of this enzyme is coenzyme $B_{12}$-dependent and might be susceptible to oxidative stress.

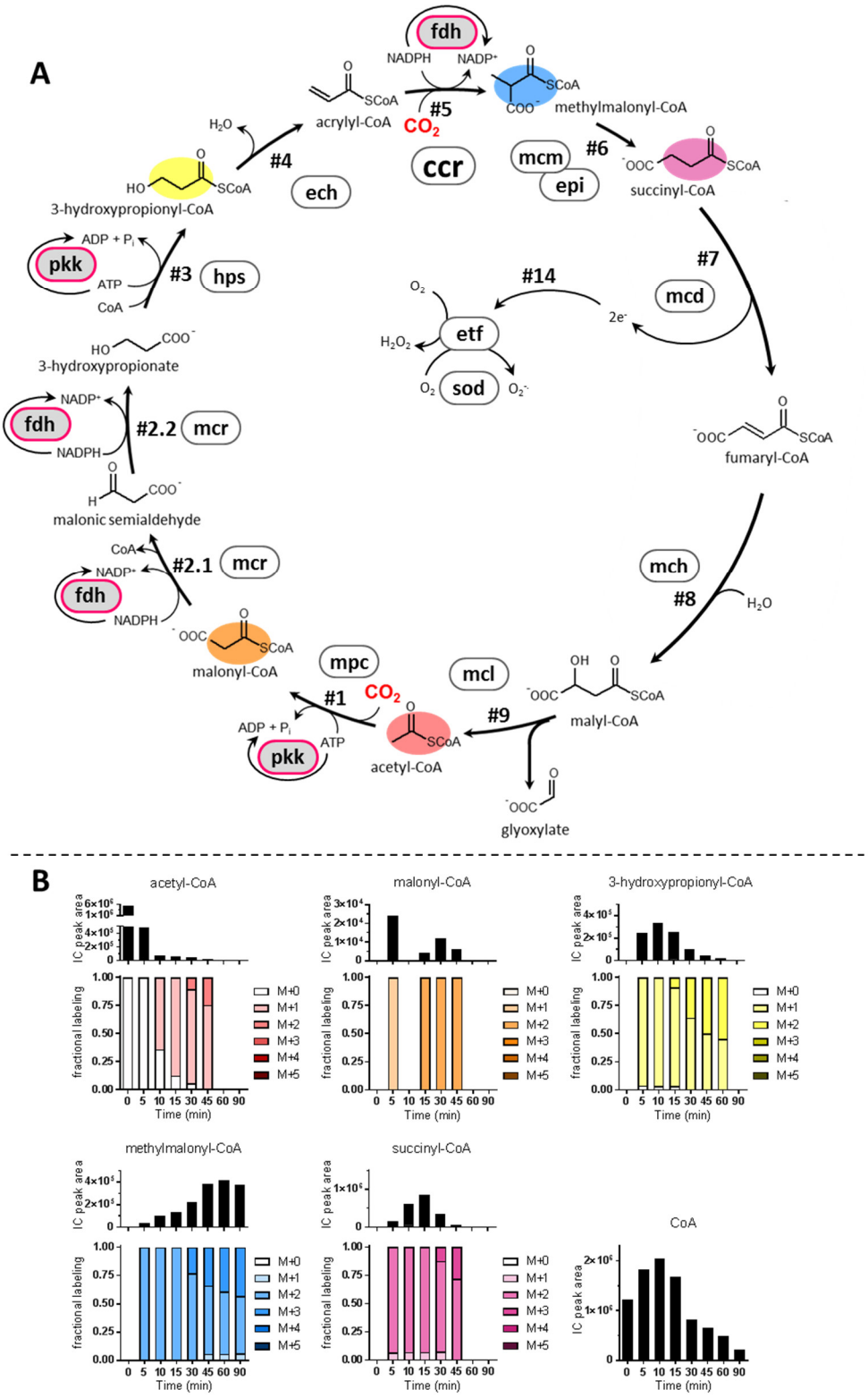



**Fig. 33 HOPAC 3.0 (ATP and NADPH regeneration)** (A) Topology of the HOPAC cycle version 3.0. Enzyme names are abbreviated according to **Table 7**. Addition of formate dehydrogenase (Fdh) and polyphosphate kinase (Pkk) (boxed in red) regenerate ATP and NADPH. (B) Dynamics of key intermediates of HOPAC 3.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA ester with color-coding according to (A). Compared to the HOPAC 2.0 cycling efficiency slightly increased, but there is an accumulation of methylmalonyl-CoA (blue) over time.

## 4.5. HOPAC 4.0 (Protection from reactive oxygen species)

From HOPAC 3.0 we suspected, that the cofactor of the $B_{12}$-dependent Mcm is susceptible to oxidative damage by ROS produced by ETF, which lead to an accumulation of methylmalonyl-CoA. Therefore, the next HOPAC version (HOPAC 4.0) included catalase (Kat) and superoxide dismutase (Sod) from *E. coli* to remove these ROS (**Fig. 34A**). These adjustments should protect the $B_{12}$ cofactor, as well as other coenzymes, cofactors, intermediates and proteins from oxidative damage. The assay was started by the addition of 250 µM of acetyl-CoA and the intermediates were analyzed by HPLC-MS. The $^{13}CO_2$ incorporation showed almost the same labeling pattern of the intermediates over time as observed in the previous version of the HOPAC cycle. In contrast, the accumulation of methylmalonyl-CoA was even further increased. This indicates that the ROS protection affected the overall performance of the cycle, but was insufficient to protect Mcm from inactivation. Coinciding with the accumulation of methylmalonyl-CoA after 30 min there was a depletion of succinyl-CoA, which underscores that Mcm was limiting the flux through the pathway. Nevertheless, the increased performance also resulted in the complete labeling of a small fraction of the intermediates after 90 min (*i.e*. 3-hydroxypropionyl-CoA and methylmalonyl-CoA) (**Fig. 34B**). This indicated that some intermediates passed the cycle already 2 and 1/2 times in this version. Additionally, we noticed the depletion of free CoA over time, which may lead to an insufficient performance of the 3-hydroxypropionyl-CoA synthetase. No side product formation could be observed so far that could explain the CoA depletion. The identification of such a side product would be important in order to devise a metabolic proofreading for its recycling. For example, a specific hydrolysis of an unwanted CoA derivative could replenish the CoA pool and help sustaining the performance of the cycle.

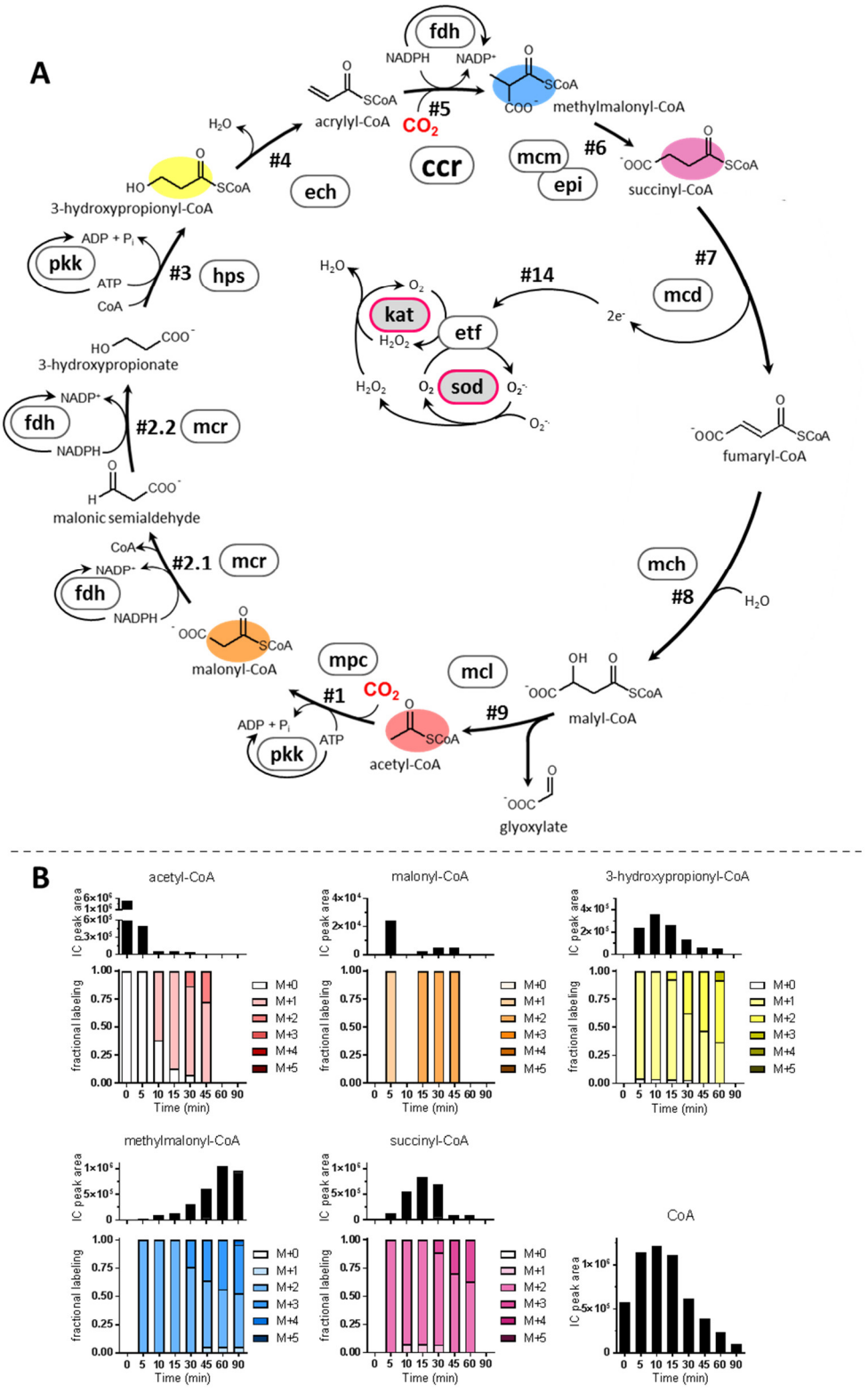



**Fig. 34 HOPAC 4.0 (Protection from reactive oxygen species)** (A) Topology of the HOPAC cycle version 4.0. Enzyme names are abbreviated according to **Table 7**. In this version, a superoxide dismutase (Sod) and catalase (Kat) (boxed in red) were added to protect the system from reactive oxygen species (B) Dynamics of key intermediates of HOPAC 4.0 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA ester with color-coding according to (A). Compared to the previous HOPAC 3.0 the labeling incorporation did not improve significantly and methylmalonyl-CoA accumulated even more. There is also an observable depletion of free CoA over time.

## 4.6. HOPAC 4.1 (further improvements)

According to the observed bottlenecks, we started the assay of HOPAC 4.1 again with 250 µM acetyl-CoA and supplied fresh Mcm and CoA after the first 20 min again. Addition of Mcm efficiently prevented the accumulation of methylmalonyl-CoA, which was observed in the previous version. The $^{13}CO_2$ incorporation improved significantly and all intermediates showed a fraction with the highest labeling possible after 90 min. This indicates, that some intermediates passed the cycle for at least three rounds of conversion at this point. The addition of CoA did not prevent its depletion in the end. The final concentrations was even lower than in HOPAC 4.0, which presumably resulted from the better overall performance in this version of the cycle (**Fig. 35B**). We identified an increase in the concentration of the oxidized dimeric form of CoA (CoAS-SCoA), which occurs simultaneously with the depletion of CoA. The oxidation of CoA implies an overall oxidative environment (**Fig. 35C**). Since the accumulation of methylmalonyl-CoA was prevented in this version, Mcd was able to perform more turn overs of the subsequent succinyl-CoA oxidation reaction. This reaction is responsible for the formation of ROS via ETF and might explain the increased oxidation of CoA. Additionally, the conversion of the free CoA pool into its oxidized form would eventually also stall the formation of 3-hydroxypropionyl-CoA by Hps. Moreover, oxidative stress could be detrimental for enzyme performance in general.

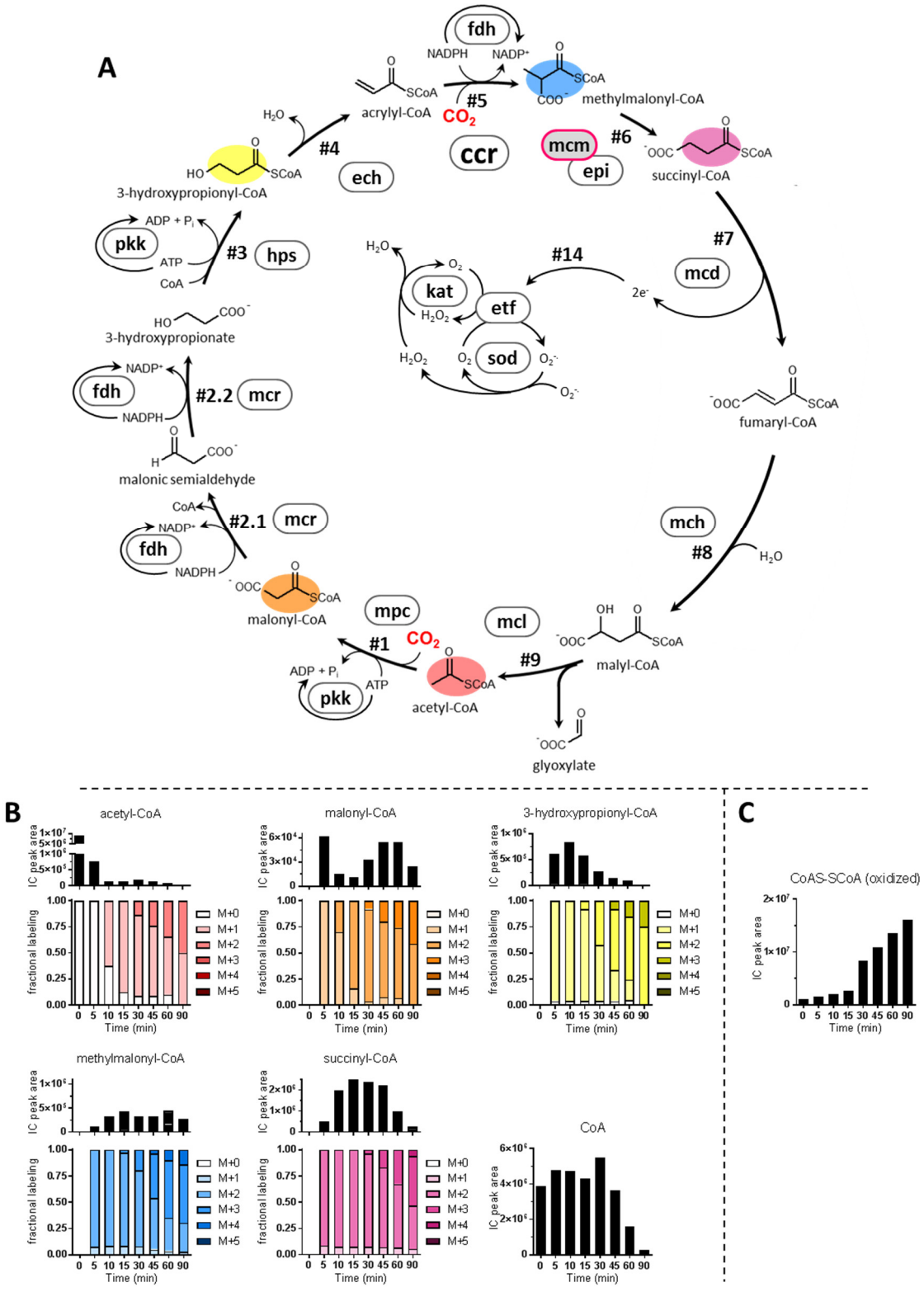



**Fig. 35 HOPAC 4.1 (further improvements)** (A) Topology of the HOPAC cycle version 4.1. Enzyme names are abbreviated according to **Table 7**. Methylmalonyl-CoA mutase (Mcm) (boxed in red) and CoA was added after 20 min. (B) Dynamics of key intermediates of HOPAC 4.1 over 90 minutes. Shown are the total levels (ion counts) and fractional labelling pattern of six different CoA ester with color-coding according to (A). The addition of Mcm prevented the accumulation of methylmalonyl-CoA compared to the previous version and increased the overall performance of the system. There is still an observable loss of free CoA over time. (C) The depletion of the free CoA pool is concurrent with the formation of the CoA dimer (CoAS-SCoA), which indicates an oxidative environment.

## 4.7. Conclusions

The HOPAC 4.1 cycle is the second synthetic pathway for the fixation of $CO_2$ *in vitro* and consists of 10 core reactions. The reaction sequence is catalyzed by 15 enzymes (including cofactor regeneration and ROS protection) originating from 8 different organisms (including bacteria and archaea) and includes the engineering of enzymes towards desired reactions. The HOPAC cycle is able to perform multiple rounds of conversion and incorporates $^{13}CO_2$ into organic molecules in a continuous fashion.

The reaction sequence, is comparable to the CETCH cycle. While the CETCH cycle converts an initial $C_3$ compound (propionyl-CoA), the HOPAC cycle converts an initial $C_2$ compound (acetyl-CoA) following a similar pattern of reaction types as the CETCH cycle. Hence, the energetic requirements (ATP equivalents per $CO_2$ fixed) are virtually the same as for the CETCH v2.0 cycle. Nevertheless, the intermediates produced, the catalyzing enzymes and the thermodynamic landscape is very different. Additionally, the reaction sequence closely resembles the first cycle of the 3-hydroxypropionate bi-cycle, which can be found in the thermophilic bacterium *Chloroflexus aurantiacus* (**Fig. 5**). The major difference in our design is that 3-hydroxypropionyl-CoA is not reduced with NADPH to propionyl-CoA for an ATP dependent carboxylation by Pcc. Instead, 3-hydroxypropionyl-CoA is first dehydrated to acrylyl-CoA, which is then reductively carboxylated by an ECR (**Fig. 35A**). This shortcut in the pathway reduces the energy requirement by one additional ATP. Moreover, the approach of synthetic biology allows us to “replace” the thermophilic enzymes with more suitable candidates for the desired mesophilic conditions. The application of a reversible ADP-forming acyl-CoA synthetase in the HOPAC cycle, also reduces the requirement for ATP, compared to the AMP-forming synthetases in the 3-hydroxypropionate bi-cycle.

For the realization of the HOPAC cycle, we first redesigned the original reaction sequence from the design phase to circumvent the problematic *in vitro* oxidation of succinate to fumarate by a synthetic route via an intermediate (fumaryl-CoA) not found in nature. Therefore, we exploited the side reactivity of Mcd and Mch for the direct conversion of succinyl-CoA to malyl-CoA. In this synthetic route we successfully applied ETF as electron acceptor for the oxidation reaction, because the engineered Mco from the CETCH cycle was not active with succinyl-CoA. Although ETF was reported to react with oxygen allowing the recycling of the electron acceptor, further investigations of the oxygen dependent reoxidation reaction are required. As we apply 50 µM of ETF in the HOPAC 4.1, the pool of ETF could be completely reduced leaving the reaction of Mcd dependent on the very slow reoxidation rate of ETF. Nevertheless, this approach actually represents an important initial step towards recycling electrons for energy conservation. The electrons from the acyl-CoA dehydrogenase reaction could be conserved via ETF for the production of ATP and NADPH on the level of ubiquinone or plastoquinone of isolated organelles or in a potential host organism.

We successfully applied enzyme engineering for the initial development of the HOPAC cycle, where the mutation of Pcc to accept acetyl-CoA as starting substrate was a crucial step for the operation of the cycle. We also investigated the engineering of Mcd toward accepting succinyl-CoA more efficiently, but these results were less successful in respect to the development of the HOPAC cycle. Additionally, the regeneration of cofactors (ATP and NADPH) as well as the protection from ROS increased the overall performance. Nevertheless, Mcm still seems to be the rate limiting step. We also observed the accumulation of the oxidized CoA dimer (CoAS-SCoA), which indicates an oxidative environment. This oxidative environment presumably arises from the reaction of Mcd coupled to ETF producing superoxide and hydrogen peroxide. This indicates that the protection from ROS via Sod and Kat might not be sufficient enough. Counteracting this redox state with a reducing agent like dithiotreitol (DTT) causes the problem of inhibiting the electron transfer efficiency from Mcd to ETF.

In a next step, the primary assimilation product (glyoxylate) could be converted into a secondary molecule to exactly quantify the fixation efficiency of the HOPAC cycle. In the CETCH cycle, glyoxylate was converted with acetyl-CoA into malate via the malate synthase. This read-out module is not applicable for the HOPAC cycle, since acetyl-CoA is a central intermediate of the cycle. Therefore a different read-out module is required (further discussed in chapter 6). The realization of the second synthetic pathway for the fixation of $CO_2$ *in vitro* increases the available options for future applications and allows for a parallel optimization of both pathways (HOPAC and CETCH) and their direct comparison (further discussed in chapter 6)

# *5.* Crystal structure of (2*S*)-methylsuccinyl-CoA dehydrogenase

The methylsuccinyl-Coenzyme A (CoA) dehydrogenase (MCD) is a member of the flavin dependent acyl-CoA dehydrogenase (ACD) family and catalyzes the oxidation of the α-branched (2*S*)-methylsuccinyl-CoA into the corresponding trans-2-enoyl-CoA thioester, mesaconyl-CoA. The enzyme plays a crucial role in the ethylmalonyl-CoA pathway for acetyl-CoA assimilation in many α-proteobacteria (*e.g*. *Rhodobacter sphaeroides, Methylobacterium extorquens and Paracoccus denitrificans*), actinobacteria (*e.g*. *Streptomyces spp., Frankia spp.*) and possibly spirochaetes (*e.g*. *Leptospira spp*.) (*94*). Other described members of the ACD family are short-chain acyl-CoA dehydrogenase (SCAD), medium-chain acyl-CoA dehydrogenase (MCAD), long-chain acyl-CoA dehydrogenase (LCAD), and very long-chain acyl-CoA dehydrogenase (VLCAD), which are involved in the β-oxidation of fatty acids. Additionally, isovaleryl-CoA dehydrogenase (IVD), isobutyryl-CoA dehydrogenase (IBD), short/branched-chain acyl-CoA dehydrogenase (SBCAD), and glutaryl-CoA dehydrogenase (GDH) are involved in the degradation of amino acids like leucine, valine, isoleucine, tryptophan, and lysine (*127-130*). A common feature to all of these enzymes is the presence of one non-covalently bound flavin adenosine dinucleotide (FAD) cofactor per active site. The α,β-desaturation of the CoA thioester is initiated by a proton abstraction from the α-carbon by a conserved catalytically active glutamate and a hydride transfer from the β-carbon to the N5 of the FAD cofactor (*131, 132*). The reduced co-factor is reoxidized by two sequential one-electron transfers to electron transfer flavoproteins (ETF), which in turn deliver the electrons to the membrane bound electron transport chain for energy conservation (*133*). For many members of the ACD family structures were solved by X-ray crystallography and almost all of them form homotetramers (dimers of dimers) consisting of four ~43 kDa subunits (*134*). An exception is the VLCAD, which forms a dimer of two ~67 kDa subunits (*135*). All enzymes in the ACD-family share an overall homologous fold, composed of an N-terminal α-helix domain, an intermediate β-barrel domain and a C-terminal α-helix domain. Again the VLCAD represents an exception as it comprises the central fold of ACDs, but also possesses a C-terminal ~180 amino acid extension forming an additional α-helix domain, responsible for dimer stability and the anchoring of the protein to the membrane (*135*).

The related family of acyl-CoA oxidases (ACO) are responsible for the β-oxidation of fatty acids in the peroxisomes of eukaryotes. ACOs also belong to the superfamily of ACDs and contain FAD as cofactor. Instead of transferring the electrons from the oxidation reaction to ETFs, however, they use molecular oxygen as a direct electron acceptor. Similar to VLCAD, these oxidases form homodimers that possess C-terminal extensions. Nevertheless, ACOs lack the transmembrane helix anchor and are thus soluble and not membrane bound. The catalytic domains of ACOs resemble to some extend the central fold of ACDs (*136*).

The different members of the ACD family exhibit substrate specificities towards distinct groups of CoA-thioesters, respectively, which are conveyed by defined active site architectures. The accommodation of all these different substrates within the same fold is therefore of particular interest. In SCAD (PDB ID: 1JQI) the active site is restricted by an isoleucine residue from helix G (*137*), whereas MCAD (PDB ID: 3MDE) possesses prolines in the same α-helix. This results in a change of the direction of the helix in MCAD, thereby widening the cavity to accommodate acyl-CoA thioesters of 4 to 16 carbons in length (*138, 139*). Bulkier residues in the active site of MCADs are substituted by glycines in VLCAD (PDB ID: 3B96), opening the pocket beyond these residues to allow the accommodation of acyl-CoA thioesters with up to 24 carbons (*135*). In the case of IVD (PDB ID: 1IVH) the catalytic glutamate residue is substituted by an alanine and another glutamate from the opposite α-helix G is assuming its space and function, resulting in a lateral expansion of the pocket and enable the accommodation of the β-branched substrate (*140*). Similarly, the active site cavity in IBD (PDB ID: 1RX0) is more narrow near

the top and more wide at the base to accept the α-branched substrate (*119*). In GDH (PDB ID: 3MPI) an arginine is important for binding of the C5 carboxylic acid moiety of the CoA thioester substrate (*118, 141*). Compared to all other acyl-CoA esters discussed above, (2*S*)-methylsuccinyl-CoA represents a substrate with mixed properties. It is short ($C_4$) but possesses a terminal carboxylic acid group, as well as an additional C2-methyl branch. To rationalize the function of MCD, it is important to decipher how substrate specificity is conferred in MCD and how a smaller substrate like succinyl-CoA, which lacks the methyl branch, is excluded from the active site of MCD.

Here we report the high resolution structure of MCD from *P. denitrificans* (PdMCD) that catalyzes the unprecedented α,β-desaturation of an α-methyl branched dicarboxylic acid CoA thioester. We used the structural information to rationalize the substrate specificity of MCD. This allowed us to engineer a MCD variant that exhibits improved catalytic efficiency with succinyl-CoA and a decreased efficiency with (2*S*)-methylsuccinyl-CoA, which is important for the development of the HOPAC cycle. For the CETCH cycle, an oxidase function was introduced into the MCD scaffold. The high resolution structure of MCD now provides a more in depth explanation for the engineered reactivity with oxygen.

## 5.1. Structure and overall topology

PdMCD forms a homodimer, consistent with our gel filtration analysis. The protein crystallized in the space group P $2_1$ $2_1$ $2_1$ with one dimer per asymmetric unit. The obtained well-ordered high resolution electron density map allowed to model almost the entire poly peptide chain (missing only 10 N-terminal residues) as well as the FAD cofactor in both active sites (**Table 8**).

**Table 8 Data collection and refinement statistics for PdMCD crystal structure**

| **PDB ID** | **6ES9** |
|---|---|
| **Ligands** | flavin adenine dinucleotide, Coenzyme A, sulfate |
| **Data collection** | |
| Wavelength | 0.979 |
| Space group | *P* $2_1$ $2_1$ $2_1$ |
| Cell dimensions | |
| *a b c* (Å) | 81.60 105.19 138.94 |
| *α β γ* (°) | 90.00 90.00 90.00 |
| Resolution range [Å] | 17.57 - 1.37 (1.42 - 1.37) |
| Number of observations | |
| total | 499135 (49543) |
| unique | 249876 (24803) |
| Redundancy | 2.0 (2.0) |
| Mean *I*/σ(*I*) | 15.25 (1.47) |
| $CC_{1/2}$ (%) | 99.9 (61.6) |
| Completeness (%) | 99.92 (99.97) |
| **Refinement** | |
| $R_{work}$ / $R_{free}$ | 0.1760 / 0.1847 |
| No. atoms | |
| protein | 8375 |
| ligands | 180 |
| solvent | 1389 |
| Mean B-factors | |
| protein | 20.52 |
| ligands | 24.86 |
| solvent | 31.31 |
| R.m.s deviations | |
| bond lengths [Å] | 0.016 |
| bond angles [º] | 1.41 |
| Ramachandran | |
| favored [%] | 98.53 |
| allowed [%] | 1.29 |
| outliers [%] | 0.18 |

Numbers in parentheses indicate values for highest resolution shell

The enzyme crystallized only in the presence of mesaconyl-CoA, the reaction product of MCDs. Although we could observe some additional electron density in the active sites, it was too ambiguous to model any ligand. Nevertheless, we observed electron densities at the surfaces of each monomer that allowed the modeling of the adenosyl phosphate parts of CoA or its derivative. The adenine ring is taking part in a cation-π stacking interactions between Arg81 and Trp476 (**Fig. 36** ). This interaction presumably stabilizes helix H, which is involved in a crystal contact, which may explain why the crystals only formed in the presence of mesaconyl-CoA. (**Fig. 36** and **Fig. 37**).

The core fold consists of ~390 amino acids and assumes the canonical central fold known from other tetrameric ACDs. Therefore, we followed the conventional secondary structure labeling of ACDs for the α-helical domain (helices A-F), the intermediary β-barrel domain (strands 1-7) and the C-terminal α-helical domain (helices G-K). The only difference to the central fold of other ACDs is a ~10 residue long loop region connecting β-strands 5 and 6 (**Fig. 37** and **Fig. 38A**). This loop is flanked by prolines (one N-terminal and three C-terminal) and contains a motif of three aspartates in a row.

## 5.2. N-terminal extension

In contrast to canonical tetrameric ACDs, PdMCD possesses an additional ~170 amino acids long N-terminal extension comprising helices A'-C', a small intermediate β-hairpin motive (strands 1' & 2'), followed by helices D'-E' (**Fig. 36** and **Fig. 37**). Superposing the monomer of PdMCD with the monomer of MCAD (PDB ID: 3MDE) shows, that the C-terminal ~390 residues structurally align well, but the N-terminus is organized in an additional domain. However, when superimposing the PdMCD monomer with the MCAD tetramer, it becomes apparent that the N-terminal ~170 residue extension of PdMCD structurally mimics the dimer-dimer interface of MCAD. The α-helices A'-E' of PdMCD align well with the α-helices G-K from the neighboring MCAD subunit (**Fig. 38B**). Exceptions are the small β-hairpin motive (not found in MCAD) and α-helix D' of PdMCD, which only loosely aligns with α-helix J from MCAD.

By emulating the dimer-dimer interface found in the tetrameric ACDs, the N-terminal extension of PdMCD is likely stabilizing the PdMCD dimer. Interestingly, all characterized ACDs exhibit a small cavity at the dimer–dimer interface. In contrast, the corresponding region in PdMCD is occupied by the unique β-hairpin structure. The hydrophilic residues in this motive form hydrogen bonds to the central fold of the opposite subunit lending additional stability to the dimer (**Fig. 36B** and **Fig. 38**). All residues from the N-terminal extension (helices A'-E') show ca. 20% sequence identity (32.5% Sequence similarity) to the structurally homologous residues from the central fold (helices G-K), indicating that the N-terminal extension originated from a partial gene duplication. The same is assumed for the C-terminal extension found in VLCAD that also mimics the dimer-dimer interface (*135*). Superposition of PdMCD with the structure of VLCAD (PDB ID: 3B96) demonstrates, that α-helices A'-C' of PdMCD align well with α-helices L-N of VLCAD. Whereas α-helix D' does not have a counterpart in VLCAD, α-helix E' from PdMCD lines up with the α-helix O from the other subunit of the VLCAD dimer (**Fig. 38**). Conclusively, in both PdMCD as well as VLCAD the stability of the dimer is increased by mimicking dimer-dimer interfaces of canonical tetrameric ACDs. Moreover, the dimeric ACOs also possess a comparable C-terminal extension and the PdMCD dimer also structurally aligns to some extent with ACOs (supplementary figure **SI Fig. 2**).

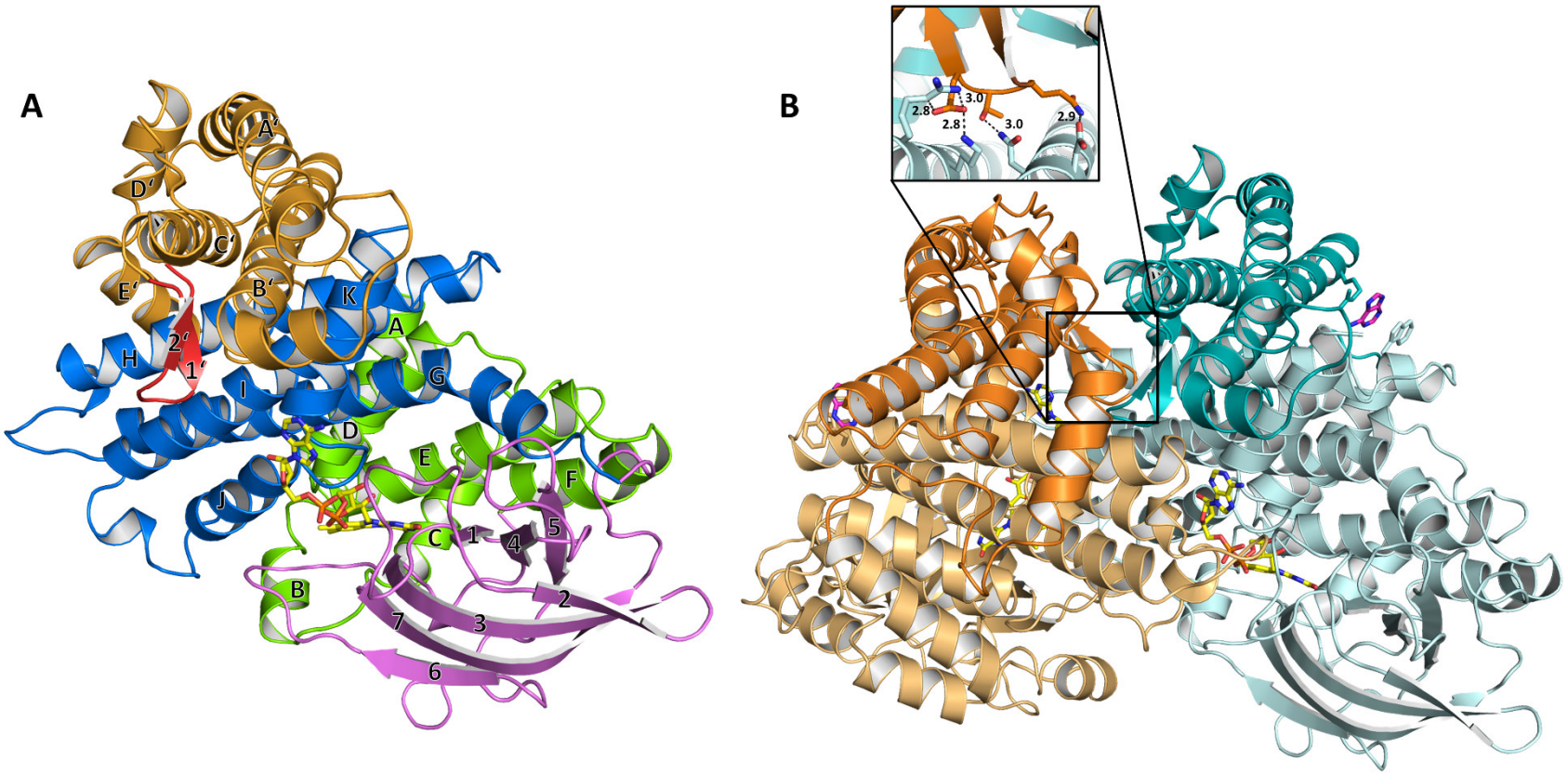



**Fig. 36 Overall structure of PdMCD.** (A) One monomeric subunit of PdMCD consists of five domains. The central fold found in other ACDs consists of an initial α-helical domain (green), an intermediary β-barrel domain (pink), and a C-terminal α-helical domain (blue). The N-terminal extension of PdMCD is shown in orange with the unique β-hairpin motif in red. The labeling of the secondary structure elements is in accordance to other ACDs. (B) PdMCD forms a homodimer and each subunit contains one FAD (yellow). The subunits are shown in orange and cyan with the N-terminal extension in darker shades, respectively. Adenosyl phosphate moieties were found to be bound on the surface of each subunit between Arg81 and Trp476. This interaction presumably stabilizes helix H, which is involved in a crystal contact. The dimeric form of PdMCD is strengthened by the β-hairpin motif in the N-terminal extension (highlighted in the box), which forms hydrogen bonds to the opposite subunit.

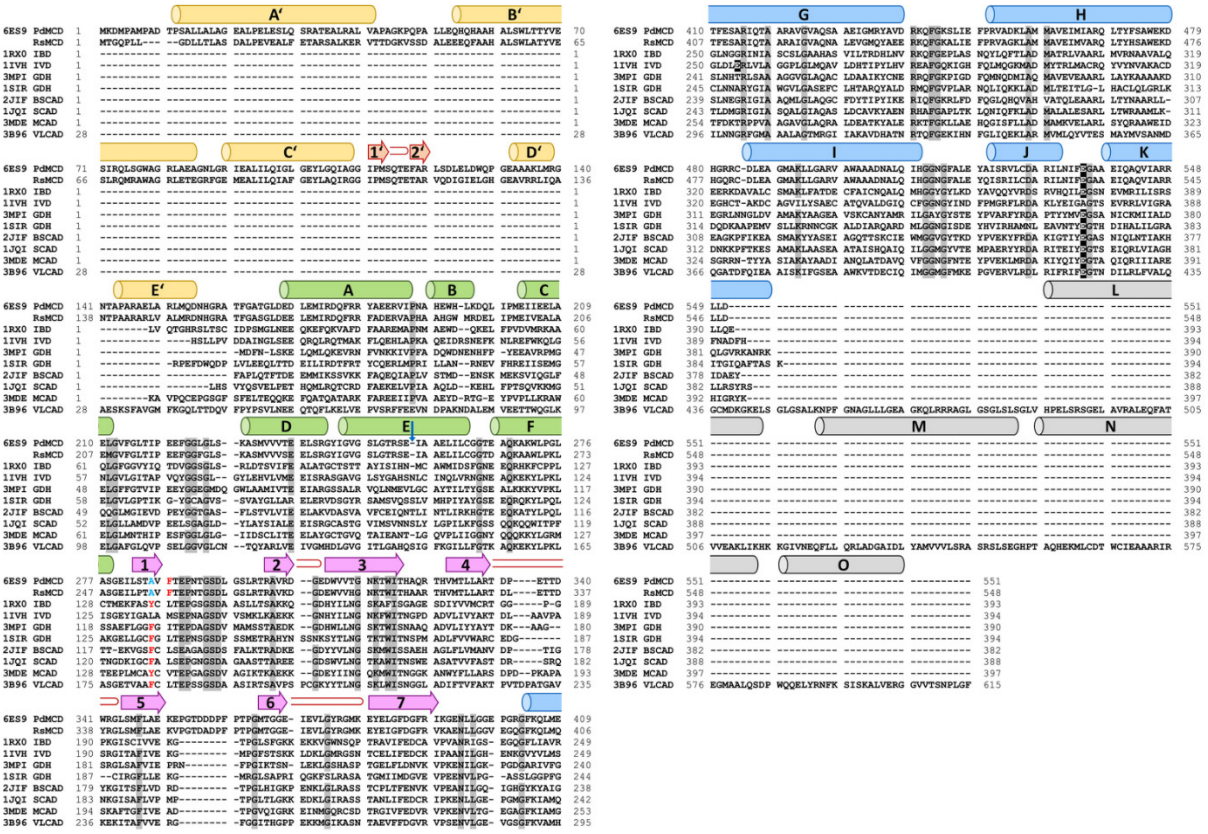


**Fig. 37 Sequence alignment of PdMCD, RsMCD and ACDs with known structures.** The sequences were aligned and the secondary structure elements are indicated (α-helices with barrels, β-sheets with arrows, and turns in red). The labeling of the secondary structure element follows the convention for canonical ACDs. The N-terminal domain of MCD is shown in orange. The central fold of all known ACDs is divided into two α-helical domains (green and blue) and an intermediary β-barrel domain (cyan). The secondary structure elements of the C-terminal extension of VLCAD is indicated in grey. Identical residues are shaded in gray and the active site glutamate is highlighted in black. The α-helix E shows a single amino acid deletion like in IVD and MCAD (indicated by a blue arrow). The Phe287 in PdMCD and Phe284 in RsMCD are shifted by two residues towards the C-terminus and the otherwise conserved phenylalanine or tyrosine at this position is replace with an alanine (indicated in red and blue letters).

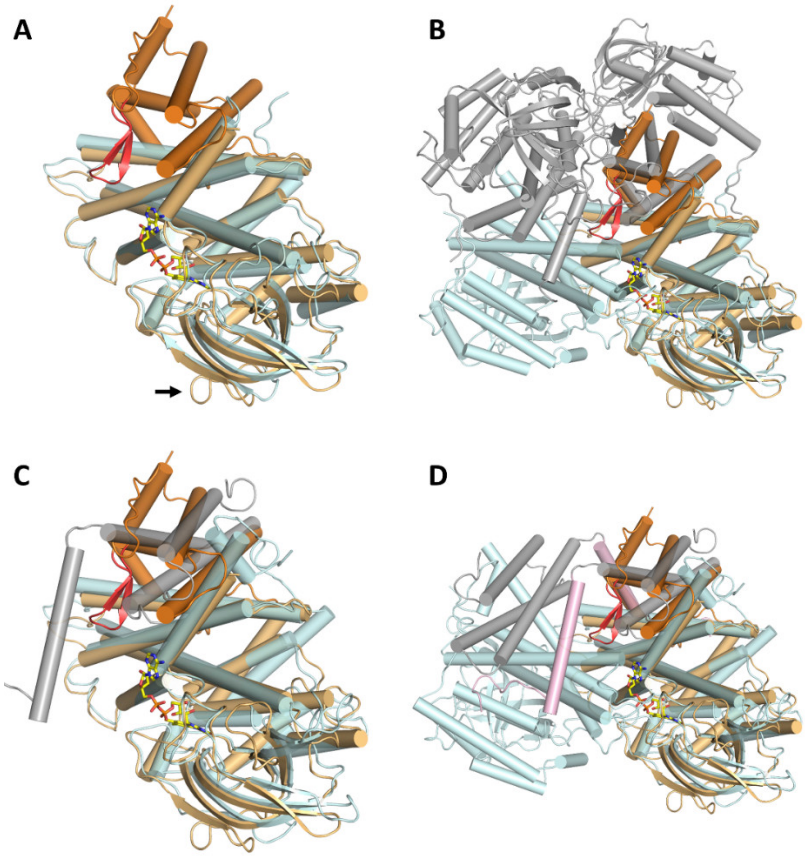



**Fig. 38 Superposition of PdMCD with other ACDs.** (A) Structural alignment of a PdMCD monomer in orange (N-terminal domain in dark orange) with the monomer of MCAD in cyan (PDB ID: 3MDE) (1.55 Å over 350 Cα-atoms per MCD monomer). The central fold of ACDs is conserved in the C-terminal domain of PdMCD. An extended loop region between β-strands 5 and 6 is indicated with an arrow. (B) Superposition of the PdMCD monomer with the tetramer (dimer of dimers in cyan and grey, respectively) of MCAD. The N-terminal extension of PdMCD aligns with the opposite monomer of MCAD and complements the dimeric interface. The unique β-hairpin motif (red) is located in a cavity at the interface, which can be found in all other ACDs known so far. (C) Superposition of a PdMCD monomer in orange (N-terminal domain in dark orange) with a monomer of VLCAD in cyan (PDB ID: 3B96) (1.85 Å over 313 Cα-atoms). VLCADs have a C-terminal extension (grey), which aligns with the N-terminal extension from PdMCD and also complements the dimeric interface. (D) Overlay of a PdMCD monomer with a VLCAD dimer (N-terminal domain: cyan; C-terminal domain: grey). The subunits of VLCAD dimers also interact via α-helix O (pink). This feature is lacking in PdMCD, but the neighboring subunits additionally interact via the β-hairpin motif (red).

## 5.3. The active site

The cofactor FAD binds to PdMCD in a drawn-out conformation as it has been observed in other ACDs as well (*118, 119, 135, 138*). Both subunits of the dimer contribute to the binding of FAD, which adopts a “butterfly like” conformation in respect to the isoalloxazine ring. This conformation has also been observed in other ACDs and was suggested to influence the redox potential of FAD for catalysis (*142*). This bending of the isoalloxazine ring is probably caused by hydrogen bonds to the polypeptide backbone (O2 with amine of Phe278, O4 with amine of Thr320, and N3 with carbonyl of Ala285) (**Fig. 39**) and stereoelectronic interaction of the isoalloxazine ring with aromatic residues (Phe287, Trp318, Phe515 and Phe534) (**Fig. 40A**).

Interestingly, the Phe287 in PdMCD is shifted two residues towards the C-terminus compared to a conserved tyrosine/phenylalanine in other ACDs (VLCAD: Phe174, MCAD: Tyr132, SCAD: Phe128 and IBD Tyr136) (**Fig. 37**). However, in a structural superposition, the aromatic ring of the shifted Phe287 is partially assuming the same space as the tyrosyl/phenyl rings in the other ACDs. The shift of this otherwise conserved residue might have functional implications as there is an alanine in the original

position in PdMCD, which provides more space at the base of the active site pocket (discussed below) (**Fig. 40C**). This is comparable to the structure of IBD (PDB ID: 1RX0). There, another conserved aromatic residue found in ACDs (PdMCD: Phe534, VLCAD: Phe421, MCAD: Tyr375 and SCAD: Tyr367) is replace with an isoleucine (Ile375 in IBD), which widens the cavity for substrate accommodation (*119*). Instead, the non-canonical Trp216 is oriented to partially assume the position of the tyrosyl/phenyl ring found in the other ACDs.

Another important residue for catalysis is a glutamate residue in position 535 in PdMCD, which is also conserved on other ACDs and was proposed to act as the catalytic base for the initial proton abstraction from the substrate. This conserved residue is positioned between α-helix J and K in canonical ACDs and in a homologues position in PdMCD (**Fig. 40E**). Only in IVD this glutamate is relocated to α-helix G, but it still occupies the same space in relation to the substrate.

A significant difference in the active site architecture of PdMCD, MCAD, and IBD is an amino acid deletion (**Fig. 37**) in α-helix E, which prevents a perturbation of this helix found in IVD, SCAD, VLCAD, and GDH. This results in a lateral widening of the active site to accommodate the sterically challenging substrates. Moreover, the same helix is shifted more than 2 Å to the back of the active site in PdMCD compared to IBD and MCAD, providing even more space for the terminal carboxylic acid group of methylsuccinyl-CoA, which likely interacts with at least two arginine residues (discussed below) (**Fig. 40**). Additionally, the loop between β-strand 6 and 7 (from Leu372 to Glu380) assumes a strikingly different conformation in PdMCD compared to other ACDs. The differing conformation of this loop is stabilized by various residue interactions (*i.e.* Tyr375 with the backbone carbonyl of Gln198, Lys379 with Glu279, and the backbone carbonyl of Tyr375 with Tyr243) and results in the positioning of Arg376 within the active site via a bidentate side-on interaction with Glu380 (**Fig. 40D**). Arg376 may contribute to substrate binding, which will be discussed next.

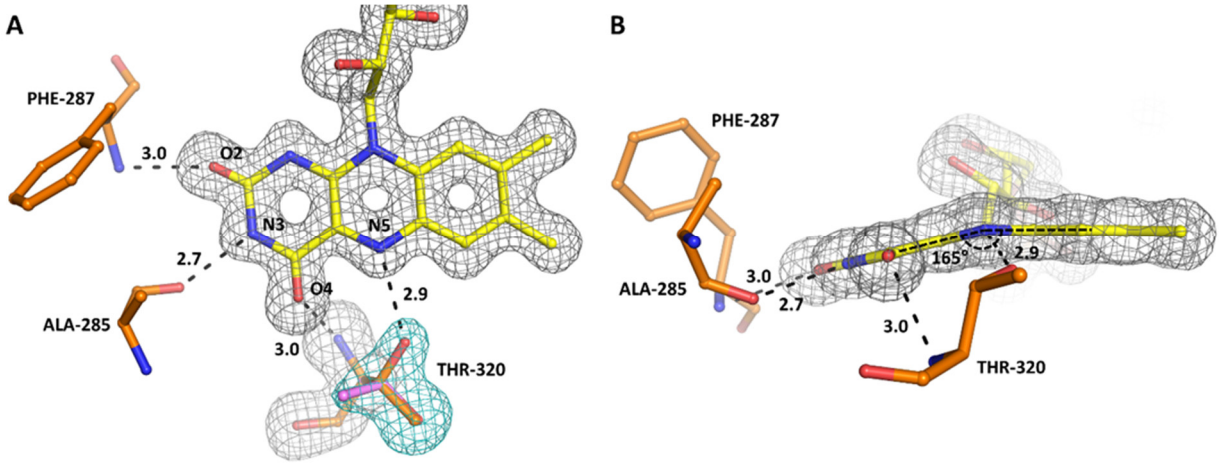



**Fig. 39 Superposition of $F_0$-$F_C$ electron density simulated annealing omit maps on a refined FAD and Thr320.** (A) Omit map at 3.0 σ for the FAD cofactor and Thr320 of PdMCD. The Thr320 forms a hydrogen bond to the N5 of the FAD. The electron density around this residue (cyan) indicates, that the threonine can assume a different rotamer conformation (purple) where the hydrogen bond is broken. The FAD cofactor forms additional hydrogen bonds with the polypeptide backbone. (B) The isolalloxazine ring of FAD assumes a distinct "butterfly-like" conformation (angle of 165°), which can be observed in other ACDs as well. This conformation is assumed to indicate a shift in the electron potential of the cofactor.

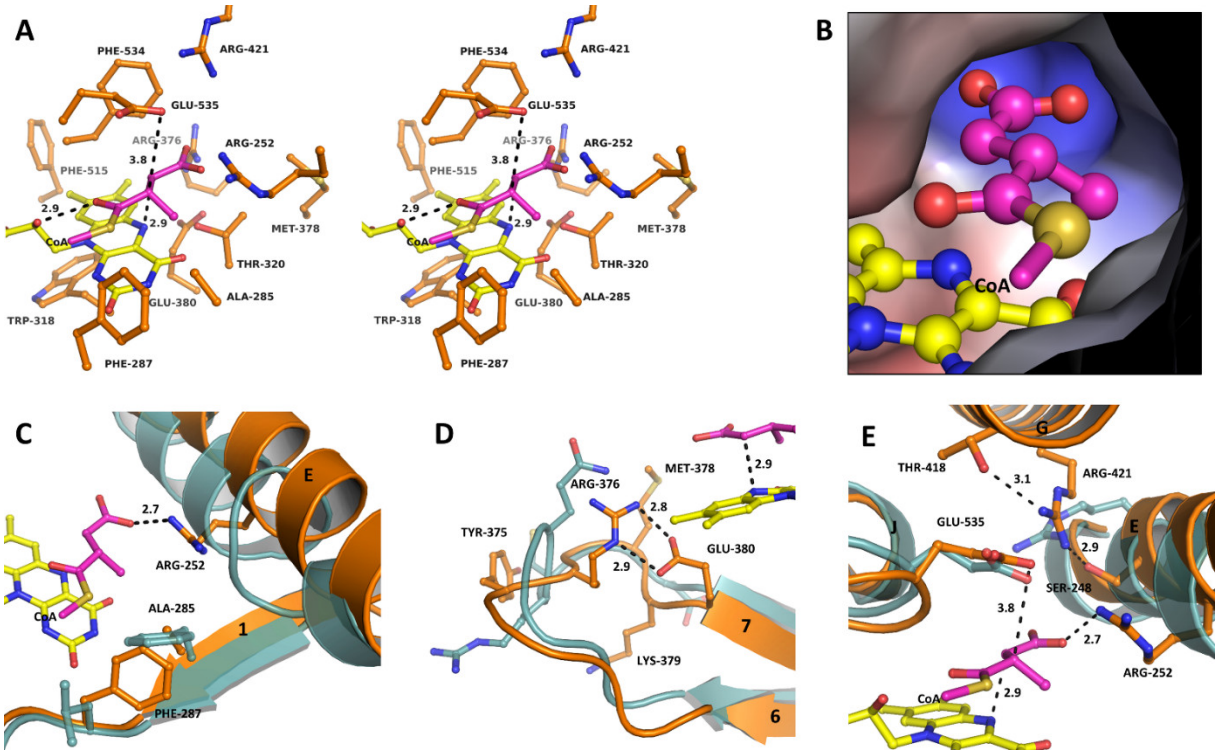



**Fig. 40 Active site architecture of PdMCD.** (A) Wall-eyed stereo view of the active site residues. The acyl moiety of (2*S*)-methylsuccinyl-CoA was fitted into the active site according to the conserved positioning of substrates found in other ACDs. The carbonyl group of the thioester bond forms a hydrogen bond to the 2'-hydroxyl of the FAD ribityl chain and the backbone amine of the catalytic Glu535. The C2 of the substrate is positioned below Glu535 for proton abstraction and the C3 is positioned above the N5 of FAD for the hydride transfer. The base of the active site cavity is lined by Arg376, Arg378 and Arg421 for binding of the carboxylic acid group of the substrate. (B) The interior surface of the active site shows a distinct pocket for the accommodation of the C2-methyl group and the base of the cavity is positively charged for the binding of the carboxylic acid of the substrate. (C) Overlay of PdMCD (orange) with VLCAD (PDB ID: 3B96) (cyan). The α-helix E of PdMCD is shifted away from the active site and a single residue deletion (as it can be found in IVD and MCAD) prevents a perturbation of this helix. A conserved phenylalanine/tyrosine residue in other ACDs is shifted in PdMCD. This shift allows to form the cavity to accommodate the C2-methyl branch of the substrate by Ala285. Moreover, Phe287 partially assumes an analogous position to the conserved aromatic residue in other ACDs. (D) Overlay of PdMCD (orange) with MCAD (PDB ID: 3MDE) (cyan) indicating a re-positioned loop between β-srands 6 and 7. This leads to the formation of a salt bridge between Glu380 and Arg376, which is thereby positioned within the active site. (D) Overlay of PdMCD (orange) with GDH (PDB ID: 3MPI) (cyan). The conserved catalytic glutamate is found in all ACDs (except IVD). In GDH a salt bridge to the substrate glutaryl-CoA is formed with an arginine from α-helix E. In case of PdMCD, Arg421 is in an analogous position but originates from α-helix G. Arg421 is held in place by Ser248 and Thr418, but the additonal Arg252 from α-helix E is in interaction distance to the fitted methylsuccinyl-CoA and may assume a different conformation upon substrate binding.

## 5.4. Substrate specificity

For ACDs the substrate specificity within the conserved fold of ACDs is of particular interest. MCD is highly specific for (2*S*)-methylsuccinyl-CoA without detectable activity towards (2*R*)-methylsuccinyl-CoA, butyryl-CoA and isobutyryl-CoA (*94*). Although, the enzyme accepts succinyl-CoA as a substrate, the rate of catalysis is ~200 fold lower than with (2*S*)-methylsuccinyl-CoA, suggesting that the C2-methyl group is crucial for the correct positioning of the substrate. Analogous to other ACDs, PdMCD shows a cleft on the surface that extends into the active site where the binding of the CoA thioester takes place. Although PdMCD was crystallized in the presence of the reaction product mesaconyl-CoA, only very weak electron density could be observed at the binding site, which did not allow the modeling of CoA or a derivative. However, it is well known from other ACDs that the thioester carbonyl oxygen forms hydrogen bonds with the 2'-hydroxyl group of the ribityl moiety of the riboflavin and the backbone amine hydrogen of the catalytic glutamate that acts as a base. Additionally, the C2 of the

substrate is required to be positioned directly below this glutamate for efficient proton abstraction, whereas C3 of the substrate needs to be positioned above the N5 of the isoalloxazine ring of FAD for the hydride transfer (*118, 119, 135, 138, 143*). Following these restrictions, we fitted (2*S*)-methylsuccinyl-CoA into the active site assuming the terminal carboxyl group to be planar relative to the carbon backbone of the molecule and parallel to the isoalloxazine ring of FAD. This would be also the case for the oxidized product (mesaconyl-CoA), in which the double bond would stabilized the planar conformation by delocalization. The modeled substrate fitted the active site surprisingly well (**Fig. 40B**), accommodating the C2-methyl branch as well as the carboxylic acid group without any significant clashing.

### 5.4.1. C2-methyl branch

The active site is laterally expanded forming a pocket lined by Ala285, Phe287, and Arg252, which provides space for the C2-methyl group (**Fig. 40A**). A similar pocket is not present in other ACDs where the alanine is substituted by conserved phenylalanine or tyrosine residues instead, which influence the electropotential of FAD (*142*). Interestingly, another phenylalanine can be found in PdMCD which is shifted towards the C-terminus (Phe287). As aforementioned this residue seems to partially assume a similar space as the conserved aromatic residue in the other ACDs (**Fig. 40C**). This observation underscores the role of Ala285 in substrate positioning and binding in MCD. The C2-methyl group of the fitted substrate protrudes into the cavity opened up by Ala285. The difference of the catalytic efficiencies between the branched and unbranched substrate might arise from a lower binding energy for succinyl-CoA lacking the C2-methyl group. The methyl group is likely also crucial for the precise positioning of the α-β-carbon bond of the substrate during catalysis. The absence of the methyl group allows succinyl-CoA to be more flexible and movable, which could affect its correct positioning and lead to unproductive binding states. A very similar observation has been made for the IBD, where propionyl-CoA is converted with only 5% of the relative catalytic rate compared to isobutyryl-CoA. The C2-metyhl branch hinders the free rotation around the α-carbon in the active site and wedges the substrate into position for efficient catalysis (*119*).

We investigated the role of different residues for substrate specificity in the homologous MCD of *Rhodobacter sphaeroides* (RsMCD), which we previously used for engineering efforts in the CETCH and HOPAC cycle. RsMCD exhibits 88% amino acid sequence identity (95% similarity) in catalytic core domain to that of PdMCD, with nearly identical catalytic properties (supplementary figures **SI Fig. 3** and **SI Fig. 4**). Interestingly, substitution of Ala282 in RsMCD (corresponding to Ala285 in PdMCD) into more sterically demanding residues like leucine, isoleucine, and phenylalanine yielded only inactive enzymes. These larger residues likely prevent accommodation of the C2 branch and also negatively impact the interaction of the carboxyl group with Arg249 (Arg 252 PdMCD) (discussed below). In contrast, mutation of the alanine into valine still resulted in some residual activity with (2*S*)-methylsuccinyl-CoA. The catalytic efficiency of this variant was decreased by 50% compared to the wild type, which was mostly due to a decreased turn over number. On the other hand, this variant exhibited an increased catalytic efficiency with unbranched succinyl-CoA due to a decrease in $K_M$ (**Table 9** and **Fig. 41A**). This indicates, that the mutation to valine prevents efficient binding of the branched methylsuccinyl-CoA, but facilitates productive binding of unbranched succinyl-CoA. We also investigated the role of the Phe284 in RsMCD (Phe287 in PdMCD), which is shifted towards the C-terminus by two positions compared to other ACDs and influences the redox potential of FAD. We additionally substituted the Phe284 with an alanine, valine or leucine within the Ala282Phe background of RsMCD, thus reversing the C-terminal shift. Nevertheless, these double mutants did not show any activity with succinyl-CoA. Interestingly, only the Ala282Phe/Phe284Ala swapping mutant was still able to convert (2*S*)-methylsuccinyl-CoA with less than 0.2% of relative catalytic efficiency (**Table 9** and **Fig. 41B**). Assuming the introduced Phe282 residue adopts a similar conformation as

observed in other ACDs (**Fig. 40**C) only a small lateral cavity would be formed, still partly accommodating the C2-methyl group of the substrate. However, it would likely result in a shift of the substrate and its α-β-carbon bond away from the optimal position for proton and hydride transfer. In contrast, the mutation into a bulkier valine or leucine in position 284 would prevent the Phe282 to adapt the canonical conformation and prevent the formation of the cavity, even resulting in a protrusion into the active site cavity. This likely leads to the exclusion of methylsuccinyl-CoA as well as succinyl-CoA.

**Table 9 Michaelis-Menten kinetic parameters for the wild type and mutant RsMCD with either (2*S*)-methylsuccinyl-CoA or succinyl-CoA as substrate.**

| | (2*S*)-methylsuccinyl-CoA | | | succinyl-CoA | | |
|---|---|---|---|---|---|---|
| | $v_{max}$ (U mg$^{-1}$) | app. $K_M$ (µM) | $k_{cat}/K_M$ ($s^{-1}$ $M^{-1}$) | $v_{max}$ (U mg$^{-1}$) | app. $K_M$ (µM) | $k_{cat}/K_M$ ($s^{-1}$ $M^{-1}$) |
| **RsMCD wt** | 79.6 ± 1.3 | 80 ± 5 | $1.0 \times 10^6$ | 0.31 ± 0.06 | 141 ± 52 | $2.3 \times 10^3$ |
| **RsMCD A282V** | 30.8 ± 0.6 | 64 ± 5 | $0.5 \times 10^6$ | 0.29 ± 0.04 | 57 ± 20 | $5.3 \times 10^3$ |
| **RsMCD A282L** | | n.d. | | | n.d. | |
| **RsMCD A282I** | | n.d. | | | n.d. | |
| **RsMCD A282F** | | n.d. | | | n.d. | |
| **RsMCD A282F/F284A** | 0.68 ± 0.03 | 439 ± 49 | $1.6 \times 10^3$ | | n.d. | |
| **RsMCD A282F/F284V** | | n.d. | | | n.d. | |
| **RsMCD A282F/F284L** | | n.d. | | | n.d. | |

n.d.: no detectable activity; Ala282 and Phe284 in RsMCD correspond to Ala285 and Phe287 in PdMCD, respectively.

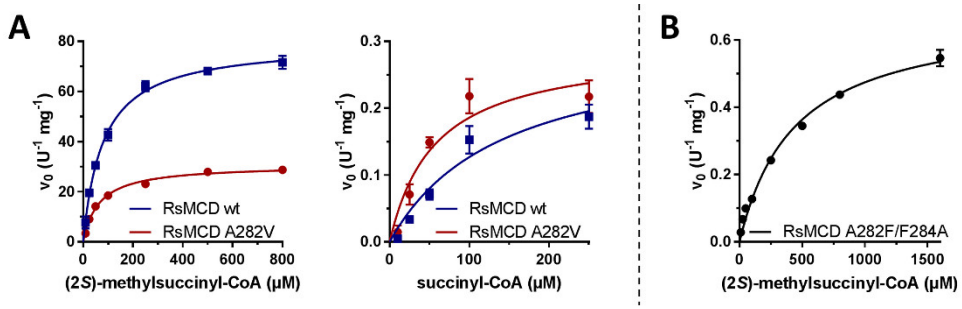



**Fig. 41 Kinetic properties of RsMCD wt, Ala282Val variant and the Ala282F/F284A double variants (correspond to Ala285 and Phe287 in PdMCD).** (A) Michaelis-Menten kinetic of RsMCD wt and Ala282Val variant with (2*S*)-methylsuccinyl-CoA and succinyl-CoA as substrate. The variant shows a decreased efficiency with (2*S*)-methylsuccinyl-CoA, but has a slightly increased efficiency with succinyl-CoA. (B) Michaelis-Menten kinetic of RsMCD Ala282Phe/Phe284Ala swapping mutant with (2*S*)-methylsuccinyl-CoA. This mutant only achieved 0.2% of relative catalytic efficiency with (2*S*)-methylsuccinyl-CoA was completely inactive with succinyl-CoA.

### 5.4.2. C4-carboxylic acid

It is striking that the inner surface at the base of the active site cavity is positively charge, which would be important for securing the carboxyl group of the substrate in the correct position (**Fig. 40B**). A similar observation can be made in the structure of GDH (PDB ID: 3MPI) where the Arg87 (from α-helix E) is responsible for the positive charge within the active site and forms a bidentate end-on interaction with the carboxyl group of glutaryl-CoA (*118*). In case of PdMCD Arg421 is located in an analogous position, but originates from α-helix G (**Fig. 40E**). Arg421 in PdMCD is at a distance of about 4 Å to the fitted carboxylic acid group of the substrate theoretically able to form a salt bridge. Furthermore, Arg376 from the aforementioned twisted loop between β-strands 6 and 7 limits the lower base of the

active site and likely contributes to substrate binding (**Fig. 40**). Whereas Arg376 and Arg421 are fixed by either Glu380 or Ser248 and Thr418, respectively, a third arginine at position 252 (from α-helix E) has no obvious interacting residue in the vicinity (**Fig. 40**). It is apparently only forming a water bridge to Ser283 and has a fairly close contact to the catalytic Glu535. The latter, however, is a very unlikely interaction partner, as this would increase the p$k_a$ of the catalytic base. It is therefore more probable that this arginine would undergo conformational changes upon substrate binding to approach the carboxylic acid group of methylsuccinyl-CoA to form mono- or bidentate side-on interactions. This would be comparable to the situation found in GDH, where Arg87 forms a bidentate interaction with the carboxyl group of glutaryl-CoA (*118*). Thus, we investigated the role of Arg252 of PdMCD by mutating it to alanine, glutamine and lysine. Each of these variants did not show any detectable activity with (2*S*)-methylsuccinyl-CoA or succinyl-CoA, demonstrating that this residue is essential for catalysis, likely by binding and correctly positioning the substrate.

## 5.5. Conclusions

The structure of PdMCD is unique among acyl-CoA dehydrogenases with an N-terminal extension to the common catalytic domain and a specific activity towards the complex substrate (2*S*)-methylsuccinyl-CoA. The active site architecture meats the substrates special demands by a basally positively charged active site to accommodate the carboxylic acid group of the substrate. We were able to demonstrate that the cavity formed by the Ala285 is required for the correct binding and positioning of the C2-branched substrate. The enzyme apparently evolved towards specifically preventing the oxidation of the unbranched succinyl-CoA, which is an essential intermediate in central carbon metabolism. The unspecific oxidation of succinyl-CoA to fumaryl-CoA would lead to a dead-end product. Furthermore, the hydrolysis of succinyl-CoA in the oxidative tricarboxylic acid cycle is used to generate an ATP equivalent and thus conserve energy. Our results show, how delicately the enzyme is evolved to distinguish between the presence and absence of a single methyl group. The exclusion of structurally similar but smaller substrates from an active site is not an easy task. The molecular basis for this ability probably arises from a lower binding energy for succinyl-CoA lacking the C2-methyl group. Therefore, the methyl group of (2*S*)-methylsuccinyl-CoA is likely crucial for the precise positioning of the α-β-carbon bond of the substrate during catalysis.

Yet, for different metabolic engineering and synthetic-biological applications, a soluble enzyme that catalyzes the α,β-carbon bond desaturation of succinyl-CoA to fumaryl-CoA is of interest. Partially closing the putative C2-binding pocket by replacing Ala282 in RsMCD with valine allowed us to change substrate specificity towards succinyl-CoA, albeit at reduced catalytic rate. This indicates that there is a potential to reverse the evolved substrate specificity of methylsuccinyl-CoA dehydrogenase.

Moreover, considering that the redox potential of the FAD cofactor in ACDs is precisely fine-tuned for the desaturation of carbon-carbon bonds of CoA thioesters and to transfer electrons onto an ETF, it is noteworthy, that we previously engineered the RsMCD into an oxidase (RsMCO) (see also chapter 3.3). RsMCO was able to accept molecular oxygen as electron acceptor instead of an ETF and catalyzed this reaction with approximately 1.5% of relative activity compared to the wild type dehydrogenase reaction (see chapter 3.3). We can now provide the structural basis for this engineered oxidase activity. A key mutation in the engineering RsMCO was to replace Thr317 (Thr320 in PdMCD) with glycine, which is conserved at this position in ACOs. The high resolution of the electron density map at the active site shows that this threonine in PdMCD can adopt different rotamer conformations, with one of them able to form a direct hydrogen bond to the N5 of the isoalloxazine ring of FAD (**Fig. 39**). Because this residue is conserved in the family of ACDs it might not only exclude molecular oxygen from the active site, but more importantly influence the electropotential of FAD by stabilizing the semiquinone state of the cofactor. This could permit the two single electron transfers to ETFs in

contrast to the two-electron transfer to molecular oxygen found in ACOs (*144*). Moreover, the engineered RsMCO also comprises a mutation of Trp315 (Trp 318 in PdMCD) to phenylalanine, which likely allows for a solvation of the active site and a modulation of the electropotential of FAD (*142, 143*). Additionally, Glu377 (Glu380 in PdMCD) was mutated to asparagine, which is not in direct contact to FAD. This residue is conserved among ACOs and probably stabilizes the formation of the superoxide anion intermediate (*106*).

In conclusion, the crystal structure of PdMCD presented in this study provides new insights into the molecular basis of reaction and substrate specificity in the ACD superfamily, and leads the way for engineering approaches to accommodate new electron acceptors and/or substrates at the active site of these highly abundant and relevant class of enzymes.

# 6. Discussion

## 6.1. Synthetic $CO_2$ fixation cycles

In this study we demonstrated the process from the theoretical design of synthetic $CO_2$ fixation cycles to the realization of two candidate pathways *in vitro*.

Initially, we designed promising pathways for the efficient fixation of $CO_2$, based on CoA-dependent carboxylases and in particular enoyl-CoA carboxylases/reductases. The metabolic retrosynthesis approach included biochemically feasible conversion and reactions that require simple cofactors, such as CoA, NAD(P)H, ATP, or flavin adenosine dinucleotide (FAD). The starting substrate and $CO_2$ acceptor molecule was regenerated via the cyclic pathway and the $CO_2$ fixation product (glyoxylate) was withdrawn from the cycle via a dedicated output reaction. Following these design principles, we design seven synthetic pathways for the fixation of $CO_2$. The theoretical analysis showed that these pathways are comparable, or even more energy efficient than naturally evolved oxygen-tolerant $CO_2$ fixation pathways. From these seven theoretical designs, we assembled two pathways (CETCH and HOPAC) from the principle building blocks *in vitro* and analyzed their capability to convert inorganic $CO_2$ into organic molecules.

The CETCH and HOPAC cycle were assembled with enzymes from various sources including all three domains of life. Moreover, we successfully applied rational enzyme engineering to optimize the performance of the cycles. The synthetic cycles operated in a defined cell-free environment and were optimized stepwise to perform multiple rounds of conversion. Both pathways comprised similar types of reactions, but the intermediates and enzymes involved were different. The realization of two different $CO_2$ fixation cycles allows for a future parallel optimization and comparison between the two systems. Whereas the HOPAC cycle is still in an earlier stage of development, the CETCH cycle was analyzed in more detail and optimized. An output module was established that condensed the primary fixation product glyoxylate with acetyl-CoA to form malate for the routine assay of the cycle. This module was not applicable for the HOPAC cycle, as acetyl-CoA is a key intermediate in this pathway. To compare the efficiency of $CO_2$ fixation in both cycles, it will be necessary to convert glyoxylate into the same product. A first option would be to use an aminotransferase for the conversion of glyoxylate with aspartate into glycine and oxaloacetate. However, this module would still require the addition of "external" aspartate as amino donor. Another module includes the conversion of the primary assimilation product into 2-phosphoglycerate via glyoxylate carboligase, tartronic semialdehyde reductase and glycerate kinase (*89*). This module is advantageous, because it does not require the addition of an "external" molecule. It condenses two molecules of glyoxylate ($C_2$) into one $C_3$ molecule. However, the condensation involves the release of a $CO_2$ molecule, which leads to an overall decrease of carbon fixation efficiency. We therefore want to also investigate a novel pathway that would conserve all the fixed carbon. In this pathway, two glyoxylate ($C_2$) would first be converted into β-hydroxyaspartate ($C_4$) and eventually to oxaloacetate (*145, 146*). Because the last two options would produce a central carbon metabolism intermediate from glyoxylate, they were also considered as a possible solutions to connect these synthetic $CO_2$ fixation cycles to the metabolism of a potential host organism in the future.

## 6.2. *In vivo* transplantation

Eventually, the *in vivo* transplantation of these synthetic $CO_2$ fixation cycles will close the existing gap between *in vitro* development and *in vivo* realization. The implementation of "synthetic autotrophy" into a phototrophic and/or heterotrophic organism would represent a milestone in synthetic biology comparable to a "metabolic heart transplantation" (*73*). It will require the remodeling and adaptation of the central carbon and energy metabolism of the host organism. Previous attempts focused on implementing natural $CO_2$ fixation pathways into *Escherichia coli (67, 69)* or *Methylobacterium extorquens (147).* This led to the incorporation of $CO_2$ into central metabolites of these organisms, but a self-sustained growth on $CO_2$ alone could not be achieved so far. Nevertheless, these studies already represent an important step towards realizing "synthetic autotrophy".

Interestingly, *M. extorquens* would be a likely host for the initial transplantation of a synthetic $CO_2$ fixation cycle, because of two reasons:

(i) This organism is able to grow on methanol as a sole source for energy and carbon. It has been shown that the energy metabolism (oxidation of methanol) can be separated from its assimilation into the central metabolism. This makes *M. extorquens* an ideal model organism. In fact, the bacterium has beed recently engineered towards the incorporation of $CO_2$ into biomass via an implemented CBB cycle (*147*). This will eventually lead to an artificial organoautorophic organism, once the metabolism is fully adapted to the CBB cycle and sustainable growth on $CO_2$ is achieved. Nevertheless, this organism still relies on the oxidation of methanol to $CO_2$ to gain energy, but the released $CO_2$ is partially assimilated via the implemented CBB cycle for biomass formation. Accordingly, the strain could be applied to host one of the synthetic $CO_2$ fixation cycles instead of the implemented CBB cycle.

(ii) *M. extorquens* already possesses the ethylmalonyl-CoA pathway (*148*), which covers already seven out of the eleven core reactions of the CETCH cycle (before version 5). Hence, "only" four enzymes (Scr, Ssr, Hbs and Hbd) would be required to complete the reaction sequence of the synthetic cycle in *M. extorquens*. The implementation of the CETCH cycle into *M. extorquens* would thus also result in an organoautotrophic organism that builds all biomass from $CO_2$, while deriving the required energy from the oxidation of methanol.

Therefore, the metabolic engineering of phototrophs is especially interesting, as they obtain the required energy from light. For example, it might be favorable to substitute the inefficient CBB cycle by an artificial and more efficient $CO_2$ fixation cycle in photoautotrophic organisms. The inherent reactivity of RubisCO with oxygen leads to the formation of 2-phosphoglycolate, which is subsequently detoxified in a process called photorespiration. This process involves the loss of already fixed $CO_2$ and is very energy demanding, thus reducing the efficiency of the CBB cycle. Furthermore, the inefficient RubisCO accounts for up to 50% of soluble protein (*57*) to achieve sufficient flux through the CBB cycle. Replacing the CBB cycle (with RubisCO as central carboxylase) with a more energy efficient process, would not only prevent photorespiration, but also decrease the cost for enzyme synthesis.

Conclusively, the *in vivo* transplantation of a synthetic $CO_2$ fixation pathway into a phototrophic, lithotrophic or heterotrophic organism, would sustains the pathway with energy and redox equivalents. Nevertheless, it would be a major challenge to connect the artificial pathway to the endogenous energy metabolism of the host. The evolved balance in the energy housekeeping would presumably be disturbed by the pathway and requires re-adaptation. Further considering the energy conservation *in vivo*, it is worth mentioning that we engineered the methylsuccinyl-CoA dehydrogenase towards an oxidase reactivity for the development of the pathways in a cell-free environment. In a host organism, the electrons from this oxidation reaction would be conserved for the production of ATP via the electron transport chain. Moreover, the synthetic route via fumaryl-CoA

in the *in vitro* HOPAC cycle could be substituted by the naturally occurring oxidation of succinate to fumarate in the TCA cycle.

However, the *in vivo* transplantation of a synthetic pathway will involve several difficulties. To achieve a more efficient autotrophic growth based on an artificial $CO_2$ fixation pathway, it has to become a main functional part of the central metabolism and the alternative pathways would need to be integrated within the native metabolism. Many problems could arise from the implementation of artificial pathways. For instance the adjustment of gene expression levels and enzyme biosynthesis, stability and activity of the recombinant enzymes, disadvantageous interactions of the synthetic pathway with the endogenous metabolic network, and incorrect regulations. However, the establishment of a dependency on the synthetic carbon fixation as a sole possibility for growth, could presumably enable natural selection to overcome these issues. A genetically engineered organism depending on a synthetic $CO_2$ fixation pathway for growth might not have a selective advantage in nature, but may be favorable from a biotechnological point of view (*73*).

Biotechnology relies strongly on heterotrophic organisms, which grow on carbon compounds, such as sugars or amino acids (*149-152*) to yield value added products. However, the utilization of $CO_2$ as a carbon feedstock in biotechnology, has lately gained significant interest. The implementation of artificial pathways for the fixation of $CO_2$ might be a favorable solution for the metabolic engineering of autotrophs that allow the direct production of value-added chemicals from $CO_2$.(*73, 153, 154*). These alternative pathways could be coupled to photosynthesis, reduced chemicals, hydrogen gas, or even electricity and provide a biotechnological and sustainable carbon capture and conversion strategy.

## 6.3. Synthetic biochemistry

Further improvements of the CETCH and HOPAC cycle in a cell-free environment could eventually also lead to a stable *in vitro* system for the formation of value added products from $CO_2$ in a synthetic biochemistry approach. The research field of synthetic biochemistry focuses on biologically produced compounds in cell free systems. A major challenge to current microbial production methods lies with the inherent restrictions by a living cell. The advantage of the cell-free approach is that only the production of the desired compound is important without considering the complexity and survivability of an organism. Previous studies have also shown that this could lead to higher rates and yields than in production strains (*77, 78*). In such an approach, the stability of a synthetic system is highly dependent on the regeneration and balancing of energy and redox equivalents (*79, 80*). These pathways start from a reduced compound and couple the energy gained from the *in situ* catabolic process to the anabolic pathway, which forms the product. Conclusively, the input molecule delivers the energy for the product formation. In our case, the synthetic cycles do not comprise a catabolic process and the input molecule is $CO_2$, which in turn requires the generation of energy and redox equivalents from a different source. Although the CETCH and HOPAC cycle, in their current versions, comprise regenerative systems for energy equivalents, they eventually depend on a limited pool of "external" input molecules (ATP is regenerated with Pkk from polyphosphate; NADPH is regenerated from formate oxidation by Fdh). In particular, the NADPH regeneration produces additional $CO_2$ in the process and is not an ideal solution for an efficient cell free process for $CO_2$ fixation. Therefore, we are interested in coupling the artificial $CO_2$ fixation pathway to isolated thylakoids from the chloroplasts of spinach or *Chlamydomonas reinhardtii* (*155, 156*). Thylakoids are able to provide the regeneration of ATP and NADPH from light via linear electron flow through photosystem II and photosystem I (**Fig. 42**). Additionally, there is a possibility to conserve the energy from the FAD dependent oxidation by the acyl-CoA dehydrogenase. In the HOPAC cycle the natural electron acceptor in the form of electron

transfer flavoprotein (ETF) was applied, but the energy is lost in the direct reduction of molecular oxygen. In an organism, the reduced ETF serves as an electron carrier, which can transfer the electrons to the ubiquinone pool in the membrane via the electron transfer flavoprotein: ubiquinone oxidoreductase (ETF-QO) (*157*). This system could be applied and engineered towards transferring the electrons from the oxidation reaction onto the plastoquinone pool of isolated thylakoids, where the energy could be conserved in the production of ATP and NADPH. Recently, we have started to experiment with the addition of isolated thylakoids to the CETCH 2.0 cycle, which enabled $CO_2$ fixation with light as energy source. These preliminary results, demonstrate the potential for the development of an efficient synthetic biochemistry process for the formation of value added products from $CO_2$ and light.

Towards that end, the primary fixation product of the CETCH 5.4 cycle (glyoxylate) could be condensed with acetyl-CoA to malate and subsequently converted into poly(L-malate) (PMA). PMA has received significant interest, as it can be applied as a pro-drug or in drug-delivery systems (*158*). It is a water soluble, biodegradable, biocompatible and bioabsorbable polymer without any toxicity (*159*). Yet, the biosynthetic pathway has not been fully elucidated and the polymer can be obtained from model production strains (*e.g. Penicillium cyclopium, Physarum polycephalum or Aureobasidium pullulans*) (*160, 161*). It would be interesting to investigate the PMA biosynthesis pathway and develop a synthetic biochemistry process to produce PMA from $CO_2$ and light. However, this requires also the sustainable production of acetyl-CoA, which might be achieved in the future by a different synthetic cycle that provides acetyl-CoA as primary fixation product (*i.e.* CHYME and CRACE cycle).

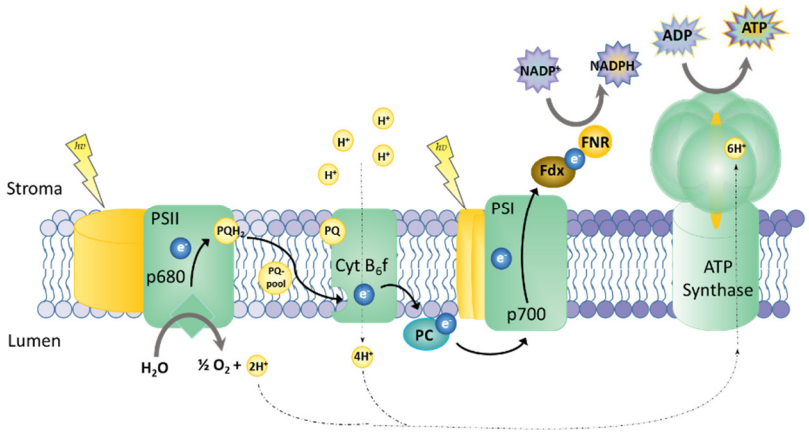



**Fig. 42 Scheme of the linear electron flow in the thylakoid membrane**. Photosystem II (PSII) generates oxygen from water and the electrons are excited by light. The high energy electrons are transferred via plastoquinones (PQ) to cytochrome $B_6f$ which generates a proton gradient across the membrane. The electrons are further transported via plastocyanin (PC) to the photosystem I (PSI) where the electrons are excited once more and are able to reduce $NADP^+$ via ferrodoxin (Fdx) and ferrodoxin NADP reductase (FNR). The proton gradient is used by the ATP synthase to phosphorylate ADP to ATP (adapted from Tarryn Miller).

## 6.4. System modeling

As aforementioned, cell-free reaction networks offer a range of advantages, including the absence of substrate or product toxicity, the absense of cell membranes and the cellular metabolism; leaving the operator with full control over the reaction networks. The latter includes the modification of the environment and enzyme composition for the optimization of the pathway and its reactions (*162*). Therefore, cell-free systems are used in a variety of applications (*e.g.* the synthesis of fine chemicals, fuels, polymers, hydrogen and even the generation of electricity). The operation of such a reaction network is highly attractive, since a broad spectrum of molecular conversions can be achieved with a high degree of selectivity. Furthermore, thermodynamically unfavored reactions can be accomplished by coupling them to exergonic steps in a network (*163*). Nevertheless, the increasing complexity of a cell-free reaction network leads often to a non-optimal behavior, since the interactions are difficult to predict and poorly understood. Therefore, a full system model, which is parameterized well enough might support the design and development of novel pathways. However, the parametrization is challenging due to nonlinear kinetics of enzymes, which are often accompanied by feedback, cooperative or allosteric regulation. Previous studies focused on the parametrization and optimization of linear pathways, which resulted in complex *in vitro* enzymatic reaction systems (*163, 164*). A full parametrization of the synthetic $CO_2$ fixation cycles would enable a model assisted optimization of the system.

In this study we rationally analyzed the reaction sequences for bottlenecks relying on the data from the individual experiments. With this approach, we were able to optimize the $CO_2$ fixation efficiency in stepwise fashion. However, for further optimizations, a model-based approach could be important. It could enable the optimization and testing of the $CO_2$ fixation efficiency *in silico* and a diversity of parameters can be assessed before testing and verification *in vitro*. The *in silico* model would allow to test the robustness of the whole system, as well as to fine-tune enzyme concentrations for optimal flux. Additionally, dynamic experiments could be performed, in which the perturbations of the system with different compounds is analyzed.

Since the $CO_2$ fixation pathway is circular, the stepwise incorporation of $^{13}CO_2$ enables the analysis of intermediates during the steady-state of the reaction sequence and results in a resolution of the cycle numbers. Theoretically speaking, the network can be analyzed according to a linear pathway, which repeats the same reactions multiple times.

Conclusively, the establishment of such a system model for a circular pathway would be highly interesting from a systems biology point of view. Additionally, the integration of the *in silico* analysis into the optimization process of synthetic $CO_2$ could prove invaluable. The *in vitro* platform for synthtetic $CO_2$ fixation cycles will advance the understanding of modulating a complex cyclic pathway *in silico* and vice versa. Finally, a well-parameterized model would certainly also support the future *in vivo* implementation.

## 6.5. Enzyme engineering

In the process of establishing the artificial $CO_2$ fixation cycles it became apparent, that enzyme engineering is a vital part for the development of synthetic pathways. The ongoing progress in synthetic pathway designs and eventually synthetic metabolisms could also drive the engineering of novel enzyme reactions and chemistry (*87*). In the course of this study, we engineered several enzymes towards a desired reaction. We used a relatively simple rational approach by evaluating sequences and active site architectures not relying on computational modeling approaches. The controlled

implementation of mutations into the enzymes already led to descent results without the need of screening large mutant libraries. Nevertheless, a further optimization of these initial obtained activities will certainly require more elaborate methods, like directed evolution of enzymes. Our rational approaches were only possible because enzyme kinetic and structural information wereavailable. For the engineering of Mcd, we initially used a homology model for the prediction of advantageous mutations. Although this was sufficient for the engineering of the oxidase (Mco) applied in the CETCH cycle, we required a more detailed crystal structure to understand the substrate specificity for its application in the HOPAC cycle. This led to the determination of the structure of the methylsuccinyl-CoA dehydrogenase and the investigation of its substrate specificity. Therefore, a novel structure was added to the well-characterized acyl-CoA dehydrogenase family. Conclusively, by building synthetic pathways, the requirement for novel reactions also drives the understanding of enzyme-catalyzed reactions and the knowledge about how to engineer and control them (*87*).

## 6.6. Closing remarks

Humans have become used to the production of $CO_2$ as a waste product following the economical and technological advances of our society. The price we pay is the increasing concentration of $CO_2$ in the atmosphere and the threat of global warming. On the other hand, the atmospheric $CO_2$ can also be considered as an important carbon feedstock of the future. Since the beginning of life, nature has evolved many different strategies to convert this inorganic molecule into organic biomass. Yet, the scientific and industrial progress has not found comparable solutions today and the application of $CO_2$-fixing enzymes and pathways in chemistry or biotechnology is still limited. Conclusively, we still rely on natural $CO_2$ fixation to convert the inorganic molecule into biomass. Many efforts have been made to improve this natural $CO_2$ fixation process, but with limited success. The field of synthetic biology offers the tempting possibility to circumvent the natural $CO_2$ fixation solutions and design optimized pathways in a bottom-up approach. Nature has evolved a vast diversity of enzymes to accomplish the catalysis of many chemical reactions and the continuous progress in biology enables us to harvest this potential. By freely combining different enzymatic reactions from various biological sources, completely novel $CO_2$ fixation routes can be constructed that are kinetically or thermodynamically more favorable compared to the naturally evolved $CO_2$ fixation pathways. Although several such synthetic routes have been theoretically considered, the implementation into a living organism is challenging. There is a lack of understanding the complex interplay between the different enzymes used in these synthetic networks, as well as the interference of these artificial pathways in the complex background of the host organism. Max Planck once said that "**Insight must precede application**" (*165*) and we considered that it is important to develop and understand such an artificial pathway first in a controlled environment. Eventually, the *in vitro* development of synthetic $CO_2$ fixation pathways narrows the gap between the theoretical design and the *in vivo* application and enables the isolated analysis and understanding of these reaction networks.

The results of this study have clearly proven, that it is possible to establish synthetic $CO_2$ fixation in a bottom up approach to investigate and optimize the network *in vitro*. We were able to go beyond the theoretical design and demonstrated two individual $CO_2$ fixation cycles, which are able to perform multiple rounds of conversion and incorporate inorganic carbon into organic molecules in a continuous fashion. The cycles rely on the efficient reductive carboxylation of enoyl-CoA esters. Although ECRs belong to the most efficient $CO_2$-fixing enzymes known to date, they were apparently not employed by nature for autotrophic $CO_2$ fixation during evolution, as far as we know. The reallocation of reductive carboxylation as key reaction for a synthetic autotrophic $CO_2$ fixation cycle thus goes beyond simply improving or reshuffling naturally existing autotrophic $CO_2$ fixation reactions and pathways. These

synthetic pathways expand the solution space of the six naturally evolved $CO_2$ fixation pathways by man-made alternatives that did not require the serendipity of evolution to bring together all components in space and time (*166*).

The here presented CETCH and HOPAC cycles are two examples for a bottom-up approach in synthetic biology. The successful assembly of complex systems from single biological components is part of the transition of biology from a descriptive discipline to a synthetic applied science. The successful reconstitution of a synthetic enzymatic network for the conversion of $CO_2$ into organic products that is superior to any chemical process or naturally existing $CO_2$-fixing solution opens the door for future applications. The *in vitro* pathways could be further improved in the future to become a stable synthetic biochemistry process as it has been shown for other artificial pathways (*77, 78*). This would enable the specific formation of value added products from $CO_2$ in a cell free environment. Moreover, the coupling of the artificial $CO_2$ fixation pathways to the energy production from light by thylakoids paves the way for a truly artificial and modular photosynthetic process. Finally, a future application includes the *in vivo* transplantation of these pathways into lithotrophic or photosynthetic organisms, breaking ground to improve $CO_2$ fixation (*167*). Considering different host organisms for the synthetic cycles will result in a different sources of energy for $CO_2$ fixation (*e.g.* lithotrophy, hydrogen or light). Moreover, there is the possibility to development artificial photosynthetic processes (*e.g.*, in combination with photovoltaics or artificial leaves) (*168*) or they might form the starting point for the design of a self-sustained, completely synthetic carbon metabolism in artificial or minimal cells (*169*).

# 7. Materials and Methods

## 7.1. Materials

Chemicals were obtained from Sigma-Aldrich (Munich, Germany) and CARL ROTH GmbH (Karlsruhe, Germany). Succinic semialdehyde was purchased from Santa Cruz Biotechnology Inc. (Dallas, USA) and $Na^{13}CO_3H$ was obtained from Cambridge Isotope Laboratories Inc. (Tewksbury, USA). Biochemicals and materials for cloning and protein expression were obtained from Thermo Fisher Scientific (St. Leon-Rot, Germany), New England Biolabs GmbH (Frankfurt am Main, Germany) and Macherey-Nagel GmbH (Düren, Germany). Carbonic anhydrase was bought from MP Biomedicals (Illkirch, France). Primers and synthesized genes were obtained from Eurofins MWG GmbH (Ebersberg, Germany) or Microsynth AG (Balgach, Switzerland). Materials and equipment for protein purification were obtained from GE Healthcare (Freiburg, Germany), BioRad (Munich, Germany) and Merck Millipore GmbH (Schwalbach, Germany).

## 7.2. Synthesis of CoA-thioesters

Crotonyl-CoA, succinyl-CoA and propionyl-CoA were synthesized from their respective anhydrides. Methylsuccinyl-CoA and Mesaconyl-CoA were synthesized via the mixed anhydride method starting from the free acids and obtaining a mixture of the isomeric 2-methyl and 3-methyl branched thioester (*170*). Ethylmalonyl-CoA was produced with 15 U crotonyl-CoA carboxylase from 6 mM crotonyl-CoA in 5 ml reaction buffer containing 60 mM ammonium carbonate pH 8, 50 mM $NaHCO_3$ and 7 mM NADPH. Methylmalonyl-CoA was produced with 300 U propionyl-CoA carboxylase from propionyl-CoA (20 mg, 24 mM) in 1 ml reaction buffer containing 100 mM ammonium carbonate pH 8, 250 mM $NaHCO_3$, 25 mM ATP, 25 mM NADH, 25 mM phosphoenolpyruvate, 10 mM $MgCl_2$ and 15 U of lactate dehydrogenase and pyruvate kinase each. 4-hydroxybutyryl-CoA was synthesized with Nmar0206 from 4-hydroxybutyrate, CoA and ATP as described in (*34*). All CoA-thioesters were purified using a HPLC (1260 Infinity, Agilent Technologies GmbH) with a Gemini® 10µm NX-C18 110 Å Column (Phenomenex, Aschaffenburg, Germany) as described before (*171*). The concentration of CoA-esters was quantified by determining the absorption at 260 nm ($\varepsilon$ = 22.4 $mM^{-1}$ $cm^{-1}$ for unsaturated and $\varepsilon$ = 16.4 $mM^{-1}$ $cm^{-1}$ for saturated CoA-thioesters). 4-hydroxybutyrate was synthesized from **γ-**Butyrolactone (*34*).

## 7.3. Bacterial strains and growth conditions

*E. coli* DH5α strain (Thermo Scientific™) was grown at 37°C in LB-medium and was used for cloning, site directed mutagenesis and DNA amplification. *E. coli* BL21 (DE3) (Thermo Scientific™) and *E. coli* Rosetta (DE3) pLys were used for protein expression and were grown in TB-medium. Antibiotics for selection purposes were used accordingly: 100 µg/ml ampicillin, 50 µg/ml kanamycin, 34 µg/ml chloramphenicol, 20 µg/ml Streptomycin and 50 µg/ml spectinomycin.

## 7.4. Plasmid List

**MM Table 1 List of all plasmids used in this study.**

| Vector | Relevant features | Source |
|---|---|---|
| **pET28a-PKK** | pET28a, His-SMc02148 (codon optimized Pkk2), kan$^r$ | Gift from Andexer, Jennifer |
| **Nmar0206** | pET16b, His-nmar0206, amp$^r$ | Könneke et al. 2014 (*34*) |
| **Nmar1309** | pET16b, His-nmar1309, amp$^r$ | Könneke et al. 2014 (*34*) |
| **p2BP1** | pLIC-SGC1, His-akr7a2, amp$^r$ | Gift from Opperman, Udo |
| **pTE701** | pET28b, His-phaJ, kan$^r$ | Gift from Vögeli, Bastian |
| **pBB541** | *E. coli* Chaperones GroEL, GroES, spectinomycin$^r$ | De Marco et al. 2007 (*172*) |
| **pEX-A-nmar0207_opt** | pEX-A, nmar0207, codon optimized for *E. coli, kan$^r$* | Eurofins MWG |
| **pEX-K4-acx4_opt** | pEX-K, acx4, codon optimized for *E. coli, kan$^r$* | Eurofins MWG |
| **pEXP5-CT/TOPO-GDH$_{Des}$** | pEXP5-CT/TOPO; His-gdh, amp$^r$ | Gift from Boll, Matthias |
| **pMCH_RS_JZ06** | pET16b, His-mch, amp$^r$ | Zarzycki et al. 2008 (*95*) |
| **pMCL1_RS_JZ03** | pET16b, His-mcl1, amp$^r$ | Erb et al. 2010 (*96*) |
| **pPCS_E_lig-hyd_JZ22** | pET16b, His-pcs (stop codon after V1214), amp$^r$ | Gift from Zarzycki, Jan |
| **pPCS_Lig_JZ26** | pET16b, His-pcs (stop codon after S900), amp$^r$ | Gift from Zarzycki, Jan |
| **pPCS_E_opt** | pSEVA471, His-pcs (codon optimized), streptomycin$^r$ | Gift from Grundling, Iria |
| **pJS12_Mcl2** | pET16b, His-mcl2, amp$^r$ | Erb et al. 2010 (*96*) |
| **pTrc-Mcr_Ca** | pTrc99A, Strep-mcr, amp$^r$ | Kroeger et al. 2011 (*173*) |
| **JW2943** | pCA24N, His-glcB, cam$^r$, ASKA collection | Kitagawa et al. 2005 (*174*) |
| **JW1721** | pCA24N, His-katE, cam$^r$, ASKA collection | Kitagawa et al. 2005 (*174*) |
| **JW3853** | pCA24N, His-yihU, cam$^r$, ASKA collection | Kitagawa et al. 2005 (*174*) |
| **JW0502** | pCA24N, His-glxK, cam$^r$, ASKA collection | Kitagawa et al. 2005 (*174*) |
| **pET21a(+)-fdh$_{D221A}$** | pET21, His-fdh (D221A), amp$^r$ | Hoelsch et al. 2013 (*98*) |
| **pTE22** | pET16b, His-mcd, amp$^r$ | Erb et al. 2009 (*94*) |
| **pTE33A** | pET16b, His-ecm, amp$^r$ | Erb et al. 2008 (*92*) |
| **pTE45** | pET16b, His-epi, amp$^r$ | Erb et al. 2008 (*92*) |
| **pTE71** | pET16b, His-ccr, amp$^r$ | Rosenthal et al. 2014 (*171*) |
| **pTE46** | pET16b, His-mcm, amp$^r$ | This work |
| **pTE380** | pCDFDuet-1, His-sucD, yihU, streptomycin$^r$ | This work |
| **pTE392** | pCDFDuet-1, His-etfA, etfB, streptomycin$^r$ | This work |
| **pTE393** | pRSET B, His-nmar0207 (codon optimized), amp$^r$ | This work |
| **pTE533** | pTE100, pccA, tet$^r$ | This work |
| **pTE534** | pET28b, pccB, kan$^r$ | This work |
| **pTE614** | pCDFDuet-1, pccA, streptomycin$^r$ | This work |
| **pTE615** | pCDFDuet-1, pccA,His-pccB, streptomycin$^r$ | This work |
| **pPccME_D407I** | pCDFDuet-1, pccA, His-pccB (D407I), streptomycin$^r$ | Gift from Zarzycki, Jan |
| **pTE801** | pET16b, His-mcd (T317G), amp$^r$ | This work |
| **pTE802** | pET16b, His-mcd (E377N), amp$^r$ | This work |
| **pTE803** | pET16b, His-mcd (Y378N), amp$^r$ | This work |
| **pTE804** | pET16b, His-mcd (T317G, E377N), amp$^r$ | This work |
| **pTE807** | pET16b, His-mcd (F531A), amp$^r$ | This work |
| **pTE808** | pET16b, His-mcd (F531K), amp$^r$ | This work |
| **pTE809** | pET16b, His-mcd (F531A), amp$^r$ | This work |
| **pTE810** | pET16b, His-mcd (W315F), amp$^r$ | This work |
| **pTE812** | pET16b, His-mcd (T317G, W315F), amp$^r$ | This work |
| **pTE813** | pET16b, His-mcd (T317G, W315F, E377N), amp$^r$ | This work |
| **pTE817** | pET16b, His-acx4_opt, amp$^r$ | This work |
| **pTE818** | pTE16b, His-acx4_opt (L137F), amp$^r$ | This work |
| **pTE819** | pET16b, His-acx4_opt (V138L), amp$^r$ | This work |
| **pTE820** | pET16b, His-acx4_opt ( V138I), amp$^r$ | This work |
| **pTE821** | pET16b, His-acx4_opt ( V291L), amp$^r$ | This work |
| **pTE822** | pET16b, His-acx4_opt (V291I), amp$^r$ | This work |

| | | |
|---|---|---|
| **pTE823** | pET16b, His-acx4_opt (V291F), amp$^r$ | This work |
| **pTE825** | pET16b, His-acx4_opt (T134L), amp$^r$ | This work |
| **pTE826** | pET16b, His-acx4_opt (T134I), amp$^r$ | This work |
| **pTE835** | pET16b, His-mcd (A282F), amp$^r$ | This work |
| **pTE836** | pET16b, His-mcd (A282L), amp$^r$ | This work |
| **pTE837** | pET16b, His-mcd (A282V), amp$^r$ | This work |
| **pTE843** | pET16b, His-mcd (A282I), amp$^r$ | This work |
| **pTE849** | pET28b, His-mcd (*P. denitrificans*), kan$^r$ | This work |
| **pTE857** | pSEVA471, His- pcs (stop codon after E1428), streptomycin$^r$ | This work |
| **pTE859** | pSEVA471, His-pcs (stop codon after N881), streptomycin$^r$ | This work |
| **pTE860** | pTE28b, His-mcd (*P. denitrificans*) (R252A), kan$^r$ | This work |
| **pTE861** | pTE16b, His-mcd (*P. denitrificans*) (R252K), kan$^r$ | This work |
| **pTE862** | pTE16b, His-mcd (*P. denitrificans*) (R252Q), kan$^r$ | This work |
| **pTE863** | pTE16b, His-etfA, etfB (D128N) streptomycin$^r$ | This work |
| **pTE864** | pTE16b, His-etfA, etfB (R191C) streptomycin$^r$ | This work |
| **pTE870** | pET16b, His-mcd (A282F, F284A), amp$^r$ | This work |
| **pTE871** | pET16b, His-mcd (A282F, F284V), amp$^r$ | This work |
| **pTE872** | pET16b, His-mcd (A282F, F284L), amp$^r$ | This work |

## 7.5. Cloning

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 (*175*). Plasmid isolation and PCR product purification was performed with kits from Macherey Nagel (Düren, Germany) according to the manufacturers protocols.

A his-tagged version of Mcm was produced by amplifying the gene encoding Mcm by PCR from chromosomal DNA of R. sphaeroides using the forward primer (5'-GGA AGA GCA TAT GAC CGA AGA TCT CGA TGC GTG GC -3') introducing an NdeI restriction site (underlined) at the initiation codon and the reverse primer (5'-GTC GGG CGG ATC CGG TCG GGT CAG C-3') introducing a HindIII site (underlined) after the stop codon. The PCR product was isolated and cloned into the expression vector pET16b (Invitrogen) resulting in pTE46.

The gene encoding for the succinyl-CoA reductase sucD was amplified by PCR from genomic DNA of *Clostridium kluyveri* strain DSM555 with Phusion High Fidelity DNA polymerase (Thermo Scientific™). The forward primer (5'-TAT ACC ATG GGC CAT CAT CAT CAT CAT CAT C-3') introduces an NcoI site (underlined), and the reverse primer (5'-TTT TGA ATT CCA GGA CTT ATC CCC ATA TTT CC-3') introduces an EcoRI site. The PCR product was isolated and cloned into the pCDFDuet-1 vector containing an N-terminal 6x His-Tag. Additionally, the gene for the 3-sulfolactaldehyde reductase yihU was amplified by PCR from with Phusion High Fidelity DNA polymerase (Thermo Scientific™) from the ASKA collection vector JW3853 by using the forward primer (5'-GAA TAC AAT TGG GCA GCA ATC GCG TT-3') introducing MunI site (underlined), and a reverse primer (5'-GCC GTC TCG AGT TAC ATT TTT ACT TTG-3') introducing XhoI site. The PCR product was isolated and also cloned into the pCDFDuet-1 containing the sucD gene resulting in the vector pTE380.

In order to clone the genes encoding for the propionyl-CoA carboxylase, the genes were amplified from genomic DNA of *M. extorquens* AM1 by PCR with Phusion High Fidelity DNA polymerase (Thermo Scientific™). For pccA, a forward primer (5'-GAC CGT GCA TAT GTT CGA TAA GAT CCT GAT TG-3') introducing the NdeI site (underlined) and a reverse primer (5'-GCT ACA TAA GCT TTC AGG CGA ATT CCA GGA TCA C-3') introducing the HindIII site (underlined) were used. The isolated PCR product was cloned into pTE100 (*89*) resulting in pTE533. For pccB, a forward primer (5'-GAC CGT GCA TAT GAA

GGA CAT CCT CGA GAA GC-3‘) introducing an NdeI site (underlined) and a reverse primer (5’-GAT ACA TGA ATT CTC AGA GCG GGA TGT TGT CGT G-3‘) introducing an EcoRI site (underlined) were used. The isolated PCR product was cloned into pET28b resulting in pTE534. In order to clone the genes into the pCDFDuet-1, a QuickChange® Site-Directed mutagenesis (Stratagene, La Jolla, USA) was performed on pTE533 to remove the NcoI site within the pccA gene. Therefore the forward (5’-GGT GCC ATC GCC GCA ATG GGC GAC AAG ATC-3’) and backward (5’-GAT CTT GTC GCC CAT TGC GGC GAT GGC ACC-3’) primer were used. Afterwards the pccA insert was amplified by PCR with Phusion High Fidelity DNA polymerase (Thermo Scientific™) using the forward primer (5’- CGG CTG CCA TAT GTT CGA TAA GAT CCT GAT TG-3‘) introducing NdeI site (underlined), and a reverse primer (5’-CAT GCG TGG TAC CTC AGG CGA ATT CCA GGA TC-3’) introducing KpnI site. The PCR product was isolated and cloned into the pCDFDuet-1 vector resulting in pTE614. The pccB gene was cut from pTE534with NcoI and SalI. The isolated insert was then cloned into pCDFDuet-1 containing the pccA gene, resulting in the vector pTE615.

The gene coding for Mcd from *P. denitrificans* was amplified using chromosomal DNA as template. Two oligonucleotides introducing restriction sites (underlined) were designed upstream (5’-ACA TGC ATA TGA AGG ACA TGC CCG CGA TG-3’; *Nde*I) and downstream (5’-ATT ATG CGG CCG CCT AGT CCA GTA GCC TGC GTG C-3’; *Not*I) of the gene coding of Mcd. PCR was performed with *Phusion® High-Fidelity DNA Polymerase* in GC-Buffer for 35 cycles, including denaturation for 60 s at 98°C, annealing for 30 s at 57°C and polymerization for 2 min at 72°C. The PCR product was cloned into the pET28b vector for expression resulting in plasmid pTE849.

The genes encoding for the electron transfer flavoprotein (etfA and etfB) were amplified from genomic DNA of *R. sphaeroides* by PCR with Phusion High Fidelity DNA polymerase (Thermo Scientific™). For etfA a forward primer introducing a BamHI site and a reverse primer introducing the HindIII site were used. The isolated PCR product was cloned into the pCDFDuet-1 vector. Subsequently, for the etfB gene a forward primer introducing an NdeI site and a reverse primer introducing the XhoI site were used and the isolated PCR product was also cloned into the pCDFDuet-1 vector already containing the etfA gene, resulting in the vector pTE392.

The gene encoding for the 4-hydroxybutyryl-CoA dehydratase Nmar0207 was synthesized by Eurofins MWG (Ebersberg, Germany). The sequence was optimized for the codon usage in *E. coli* (chapter 7.14). In order to clone the gene, a BamHI and an EcoRI site were introduced before and after the gene, respectively. The gene was cut from the vector with those restriction enzymes and the isolated insert was cloned into pRSET B (Thermo Fisher Scientific, St. Leon-Rot, Germany) resulting in pTE393.

The gene encoding for the short chain acyl-CoA oxidase 4 acx4 was synthesized by Eurofins MWG (Ebersberg, Germany). The sequence was optimized for the codon usage in *E. coli* (chapter 7.14). In order to clone the gene, an NdeI and an EcoRI site were introduced before and after the gene, respectively. The gene was cut from the vector with those restriction enzymes and the isolated insert was cloned into pET16b resulting in pTE817.

## 7.6. Site-directed mutagenesis

Site-directed mutagenesis was performed with the Quick Change® (Stratagene, La Jolla, USA) method and protocols. For the introduction of the mutations, 60 ng of template DNA and the corresponding primers (**MM Table 2**) were used.

**MM Table 2 List of Primers used for site directed mutagenesis**

| Gene | Organism | Template | Mutation | Primer forward 5' → 3' | Primer reverse 5' → 3' | Vector |
|---|---|---|---|---|---|---|
| **mcd** | *R. sphaeroides* | pTE22 | T317G | CAACAAGACCTGGATCGGCCATGCCGCGCG | GTGCGCGCGGCATGGCCGATCCAGGTCTTGTT | pTE801 |
| **mcd** | *R. sphaeroides* | pTE22 | E377N | CTACCGCGGCATGAAGAACTATGAGATCGG | CGAAGCCGATCTCATAGTTCTTCATGCCGCGGT | pTE802 |
| **mcd** | *R. sphaeroides* | pTE22 | Y378N | CGCGGCATGAAGGAAAACGAGATCGGCTT | GTCGAAGCCGATCTCGTTTTCCTTCATGCCGCG | pTE803 |
| **mcd** | *R. sphaeroides* | pTE801 | T317G, E377N | CTACCGCGGCATGAAGAACTATGAGATCGG | CGAAGCCGATCTCATAGTTCTTCATGCCGCGGT | pTE804 |
| **mcd** | *R. sphaeroides* | pTE22 | F531A | CATCCTCAACATCGCCGAGGGGGCGGCC | GGCCGCCCCCTCGGCGATGTTGAGGATG | pTE807 |
| **mcd** | *R. sphaeroides* | pTE22 | F531K | GCATCCTCAACATCAAAGAGGGGGCGGCC | CTCGGCCGCCCCCTCTTTGATGTTGAGGATGC | pTE808 |
| **mcd** | *R. sphaeroides* | pTE22 | W315A | GGCAACAAGACCGCGATCACCCATGCC | GGCATGGGTGATCGCGGTCTTGTTGCC | pTE809 |
| **mcd** | *R. sphaeroides* | pTE22 | W315F | GGCAACAAGACCTTCATCACCCATGCCGC | GCGGCATGGGTGATGAAGGTCTTGTTGCC | pTE810 |
| **mcd** | *R. sphaeroides* | pTE801 | T317G, W315F | GGCAACAAGACCTTCATCGGCCATGCCG | CGGCATGGCCGATGAAGGTCTTGTTGCC | pTE812 |
| **mcd** | *R. sphaeroides* | pTE812 | T317G, W315F, E377N | CTACCGCGGCATGAAGAACTATGAGATCGG | CGAAGCCGATCTCATAGTTCTTCATGCCGCGGT | pTE813 |
| **acx4_opt** | *A. thlaiana* | pTE817 | L137F | GTTGCTCAACGTTCATTTTCGTCCATTCAAG | CCCAGGCTTGAATGGACGAAAATGAACGTTGA | pTE818 |
| **acx4_opt** | *A. thlaiana* | pTE817 | V138L | GCTCAACGTTCATTCTGCTGCATTCAAGCCT | CATACCCAGGCTTGAATGCAGCAGAATGAACG | pTE819 |
| **acx4_opt** | *A. thlaiana* | pTE817 | V138I | CAACGTTCATTCTGATCCATTCAAGCCTGG | CCAGGCTTGAATGGATCAGAATGAACGTTG | pTE820 |
| **acx4_opt** | *A. thlaiana* | pTE817 | V291L | GTTAGCCGTGTCATGCTGGCCTGGCAACCT | GATAGGTTGCCAGGCCAGCATGACACGGCTAA | pTE821 |
| **acx4_opt** | *A. thlaiana* | pTE817 | V291I | CGTTAGCCGTGTCATGATTGCCTGGCAACC | GATAGGTTGCCAGGCAATCATGACACGGCTAA | pTE822 |
| **acx4_opt** | *A. thlaiana* | pTE817 | V291F | CGTTAGCCGTGTCATGTTTGCCTGGCAACCT | GATAGGTTGCCAGGCAAACATGACACGGCTAA | pTE823 |
| **acx4_opt** | *A. thlaiana* | pTE817 | T134L | GATGCCAGTTGCTCACTGTTCATTCTGGTCC | GGACCAGAATGAACAGTGAGCAACTGGCATC | pTE825 |
| **acx4_opt** | *A. thlaiana* | pTE817 | T134I | GATGCCAGTTGCTCAATCTTCATTCTGGTCC | GAATGGACCAGAATGAAGATTGAGCAACTGG | pTE826 |
| **mcd** | *R. sphaeroides* | pTE22 | A282F | GATCCTGCCGACCTTCGTCTTCACCGAG | CTCGGTGAAGACGAAGGTCGGCAGGATC | pTE835 |
| **mcd** | *R. sphaeroides* | pTE22 | A282L | GATCCTGCCGACCCTCGTCTTCACCGAG | CTCGGTGAAGACGAGGGTCGGCAGGATC | pTE836 |
| **mcd** | *R. sphaeroides* | pTE22 | A282V | GATCCTGCCGACCGTCGTCTTCACCGAG | CTCGGTGAAGACGACGGTCGGCAGGATC | pTE837 |
| **mcd** | *R. sphaeroides* | pTE22 | A282I | GATCCTGCCGACCATCGTCTTCACCGAG | CTCGGTGAAGACGATGGTCGGCAGGATC | pTE843 |
| **mcd** | *R. sphaeroides* | pTE835 | A282F, F284A | CCGACCTTCGTCGCCACCGAGCCGAAC | GTTCGGCTCGGTGGCGACGAAGGTCGG | pTE870 |
| **mcd** | *R. sphaeroides* | pTE835 | A282F, F284V | GCCGACCTTCGTCGTCACCGAGCCGAAC | GTTCGGCTCGGTGACGACGAAGGTCGGC | pTE871 |
| **mcd** | *R. sphaeroides* | pTE835 | A282F, F284L | GCCGACCTTCGTCCTCACCGAGCCGAAC | GTTCGGCTCGGTGAGGACGAAGGTCGGC | pTE872 |
| **mcd** | *P. denitrificans* | pTE849 | R252A | CTCGCTGGGCACCGCCAGCGAGATCGC | GCGATCTCGCTGGCGGTGCCCAGCGAG | pTE860 |
| **mcd** | *P. denitrificans* | pTE849 | R252K | CTCGCTGGGCACCAAAAGCGAGATCGCGG | CCGCGATCTCGCTTTTGGTGCCCAGCGAG | pTE861 |
| **mcd** | *P. denitrificans* | pTE849 | R252Q | CGCTGGGCACCCAGAGCGAGATCGC | GCGATCTCGCTCTGGGTGCCCAGCG | pTE862 |
| **pcs** | *Erythrobacter* | pPCS_E_opt | Stop codon after E1428 | CTGTTCACTACTCTGTAAATCCAAGCTGGC | GCCAGCTTGGATTTACAGAGTAGTGAACAG | pTE857 |
| **pcs** | *Erythrobacter* | pTE857 | Stop codon after N881 | CTTCTACCCTGCGTAACTGAGAGTCTCTGGA | GTTCGTCCAGAGACTCTCAGTTACGCAGGGTA | pTE859 |
| **etfB** | *R. sphaeroides* | pTE392 | D128N | GCAAGCAGGCCATCAACAATGACATGAACG | CGTTCATGTCATTGTTGATGGCCTGCTTGC | pTE863 |
| **etfB** | *R. sphaeroides* | pTE392 | R191C | GATGAACGAGCCGTGCTATGCCTCGC | GCGAGGCATAGCACGGCTCGTTCATC | pTE864 |

## 7.7. Heterologous enzyme production and purification

### 7.7.1. General protocol

All enzymes were purified according to the following protocol if not described otherwise. *E. coli* BL21 (DE3) were transformed with the expression plasmid (**MM Table 1**) and grown over night on LB-agar plates containing the antibiotic(s) necessary for selection. The grown colonies were used to inoculate 1 L TB medium for the expression culture. The cultures were grown at 37°C to an $OD_{600}$ of 1.5 and induced with 0.25 mM IPTG. The protein was produced over night at 21°C. The cells were harvested by centrifugation at 5000 rcf for 10 min and resuspended in lysis buffer (20 mM Tris-HCl pH 8.1, 500 mM NaCl, 10% Glycerol). Resuspended cells were stored at -20°C. Cells were lysed by ultrasonification and the lysate was cleared by ultracentrifugation at 50,000 rcf 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 (GE Healthcare, Freiburg, Germany) and unbound protein was removed with 20 ml of 20 mM Tris-HCl pH 8.1, 500 mM NaCl, 75 mM Imidazole. The protein was eluted in 20 mM Tris-HCl pH 8.1, 500 mM NaCl and 500 mM Imidazole. The protein was either desalted with a HiTrap 5 ml Desalting (GE Healthcare, Freiburg, Germany) column in 20 mM Tris-HCl pH 8.1, 200 mM NaCl or by a HiLoad 16/600 Superdex 200 pg (GE Healthcare, Freiburg, Germany) size exclusion column in 20 mM Tris-HCl pH 8.1, 150 mM NaCl. Elution fractions were concentrated with Amicon Ultra-4 centrifugal filters (Merck Millipore, Darmstadt, Germany). The concentration was determined on a Nanodrop 2000 (Thermo Scientific, St. Leon-Rot, Germany). The purified proteins were stored in 20% glycerol at -20°C up to several month if not mentioned elsewise. 2 mM $MgCl_2$ was added to enzymes containing magnesium.

### 7.7.2. Succinyl-CoA reductase (SucD) and succininc semialdehyde reductase (AKR7a2)

The general expression and purification protocol was applied. The enzymes were stored in 20% glycerol at -20°C for up to two weeks.

### 7.7.3. Ethylmalonyl-CoA mutase (Ecm) and methylmalonyl-CoA mutase (Mcm)

The expression and purification was conducted using the general protocol mentioned above. Additionally, these two enzymes were always purified by size exclusion FPLC as mentioned above. These proteins were stored with additional 2 mM of coenzyme $B_{12}$ and 20% glycerol at -20°C for up to several month

### 7.7.4. Acyl-CoA oxidases (Mco and Pco)

For expression, *E. coli* Rosetta (DE3) pLys was transformed with the expression plasmid (**MM Table 1**) and grown on LB-agar plates over night at 37°C. The grown colonies were used to inoculate 1 L TB-medium containing additionally 10 mg/L riboflavin. The culture was grown at 37°C to an $OD_{600}$ of 1.5 and induced with 0.25 mM IPTG. The protein was produced for 4h at 25°C and the cells were harvested as described above. For protein purification, the general protocol above was applied. Note that Pco can only be concentrated up to 1.2 mg $ml^{-1}$. The proteins were stored in 20% glycerol at -20°C for several weeks.

### 7.7.5. 4-hydroxybutyryl-CoA synthetase (Nmar0206) and 3-hydroxypropionyl-CoA synthetase (Nmar1309)

The general protocol as mentioned before was applied. Additionally, the BL21 (DE3) cells were also transformed with a plasmid (pBB541) for expression of the chaperone proteins GroES and GroEL from *E. coli* (*172*). All purification buffers contained 50 mM MOPS pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$ and 5 % v/v Glycerol. For the lysis, protease inhibitor (SIGMAFAST™ protease inhibitor (Sigma-Aldrich, Munich, Germany)) was added according to the provided protocol. Unspecific bound protein was washed with only 50mM of Imidazole. During concentration, the glycerol concentration was step wise increased to a final concentration of 20% and the protein was stored at -20°C for only up to three days.

### 7.7.6. 4-hydroxybutyryl-CoA dehydratase (Nmar0207)

The culture was grown in TB medium at 37°C aerobically to an $OD_{600}$ of 4. The culture was transferred to a sterile Schott Bottle and cooled down with closed lid to 21°C. 100 µM of $Fe(II)SO_4$, 100 µM Fe(III)citrate, 20 mM fumarate and 0.25 mM IPTG were added and the lid was closed firmly. The protein was produced under microaerobic conditions over night at 21°C. The following purification was performed aerobically as mentioned in the general protocol above, but all buffers contained 500 mM NaCl. The purified protein was stored in 20 % glycerol at -20°C for several weeks.

### 7.7.7. Electron transfer flavoprotein (EtfAB)

For expression, *E. coli* Rosetta (DE3) pLys was transformed with the expression plasmid (**MM Table 1**) and grown on LB-agar plates over night at 37°C. The grown colonies were used to inoculate 1 L TB-medium containing additionally 10 mg/L riboflavin. The culture was grown at 37°C to an $OD_{600}$ of 1.5 and induced with 0.25 mM IPTG. The protein was produced for overnight at 21°C and the cells were harvested as described above. For protein purification, the general protocol above was applied, but the unspecific bound protein was washed off with only 15 mM of imidazole. The protein was stored with 20% glycerol at -20°C for several weeks.

### 7.7.8. Malonyl-CoA reductase (Mcr)

*E. coli* BL21 (DE3) were transformed with the expression plasmid pTrc-Mcr_Ca and grown over night on LB-agar plates containing 100 µg/ml. The grown colonies were used to inoculate 1 L TB medium for the expression culture. The cultures were grown at 37°C to an $OD_{600}$ of 1.5 and induced with 0.25 mM IPTG. The protein was produced over night at 21°C. The cells were harvested by centrifugation at 5000 rcf for 10 min and resuspended in lysis buffer (50 mM MOPS pH 7.5, 200 mM KCl, 10% Glycerol). Resuspended cells were stored at -20°C. Cells were lysed by ultrasonification and the lysate was cleared by ultracentrifugation at 50,000 rcf 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, Freiburg, Germany) and unbound protein was removed with a buffer containing 50 mM MOPS pH 7.5, 200 mM KCl. The protein was eluted in 50 mM MOPS pH 7.5, 200 mM KCl with 2.5 mM desthiobiotin. The protein was desalted with a HiTrap 5 ml Desalting (GE Healthcare, Freiburg, Germany) column in 50 mM MOPS pH 7.5, 200 mM KCl. After the addition of 20% glycerol, the protein was stored at -20°C for several month.

## 7.8. Crystallization of Mcd from *P. denitrificans*

Competent *E. coli* Rosetta pLys (DE3) were transformed with the plasmid pTE849 and 1 liter cultures were grown at 37°C in TB medium with 50 µg kanamycin $ml^{-1}$ to an OD ~1 Expression was performed at 20°C over night after induction with 250 µM IPTG. The cells were harvested by centrifugation at 5000 rcf for 10 min and resuspended in 2 times the volume of lysis buffer (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 10% Glycerol). For the crystallization, the Mcd from *P. denitrificans* was purified as follows. Cells were lysed by ultrasonification and the lysate was cleared by ultracentrifugation at 50,000 rcf for 45 min at 4 °C and subsequently filtered through a 0.45 µm filter. The cleared lysate was loaded onto a 5 ml HisTrap FF column (GE Healthcare, Freiburg, Germany) and unbound protein was removed with 20 ml of 20 mM Tris-HCl pH 7.9, 500 mM NaCl, 75 mM Imidazole. The protein was eluted in 20 mM Tris-HCl pH 7.9, 500 mM NaCl and 500 mM Imidazole and subsequently desalted with a HiTrap 5 ml Desalting (GE Healthcare, Freiburg, Germany) column in 20 mM Tris-HCl pH 7.9, 250 mM NaCl. The His-Tag was removed by Thrombin (T4648, Sigma-Aldrich, Munich, Germany) mediated proteolysis (1 mg Thrombin per 10 mg Mcd) for 2 h at room temperature. The digested protein was passed through a 1 ml HisTrap FF column (GE Healthcare, Freiburg, Germany) and the flow through was collected. The digested protein was further purified via a HiLoad 16/600 Superdex 200 pg (GE Healthcare, Freiburg, Germany) size exclusion column in 20 mM Tris-HCl pH 7.9, 200 mM NaCl. Elution fractions were concentrated with Amicon Ultra-4 centrifugal filters (Merck Millipore, Darmstadt, Germany and the protein concentration was determined on a Nanodrop 2000 (Thermo Scientific, St. Leon-Rot, Germany).

Crystals were grown at 16°C using the sitting-drop vapor diffusion method. The Mcd protein adjusted to a concentration of 20 mg/ml in a buffer containing 20 mM Tris-HCl pH 7.9, 200 mM NaCl with a final concentration of 1.5 mM FAD and 3 mM mesaconyl-CoA. The protein solution was mixed in a 1:1 ratio with the crystallization buffer containing 30% PEG 5000 MME, 100 mM Tris pH8 and 200mM $LiSO_4$. Crystals were soaked in 30% glycerol before the being plunged into liquid nitrogen for freezing. X-ray diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France (beamline: ID29). The data was processed with XDS (*176*) and the CCP4 software package (*177*). The structure was solved using molecular replacement using Phaser-MR- and AutoBuild-programs of the Phenix software package (*178*). The structure of a putative acyl-CoA dehydrogenase (PDB ID: 4N5F) served as search model for the molecular replacement. Additional manual modeling and ligand fitting was performed using COOT (*179*). Because the FAD cofactor assumed a bend (butterfly like) conformation for the isoalloxazine ring the molecular restrains for the refinements had to be modified. The isoalloxazine ring was therefore separated into two planes along the N5/N10 axis of central

pyrazine ring. Additionally, the restrains for the bond angles over N5 and N10 were relaxed. Water-picking and further refinements were done by Phenix. The atomic coordinates (PDB ID: 6ES9) and structure factors have been deposited in the Protein Data Bank [http://www.pdb.org/].

## 7.9. Activity assays

### 7.9.1. Activity assay of sulfolactaldehyde reductase (YihU)

The activity was determined at 30°C in 150 µl of 100 mM MOPS pH 7.2 buffer containing 10 mM $MgCl_2$, 10 mM $KCl_2$, 1 mM $MnCl_2$ and 200 µM NADH and 7.8 µg purified YihU. Different concentrations of succinic semialdehyde (from 2 mM up to 100 mM) were used to start the reaction. The consumption of NADH was observed at 360 nm ($\Delta\varepsilon = 3.1\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany)

### 7.9.2. Activity assay of succinic semialdehyde reductase (AKR7a2)

The activity of AKR7a2 was measured in a phosphate buffer (100 mM, pH 7.4) containing 0.2 mM NADPH and 2 µg AKR7a2. The reaction was started by addition of succinic semialdehyde and the reaction was followed at 28 °C and 340 nm ($\Delta\varepsilon = 6.22\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-Vis spectrometer (Agilent Technologies GmbH, Waldbronn, Germany). Michaelis-Menten kinetic parameters were determined by varying the concentration of succinic semialdehyde from 2.4 µM to 480 µM.

### 7.9.3. Activity assay of medium-chain fatty acid CoA ligase (AlkK)

The activity of AlkK was measured in 450 µl of a 50 mM Tris-HCl buffer pH 7.8 containing 0.5 mM CoA, 1.6 mM ATP, 2 mM $MgCl_2$, 0.8 mM phosphoenolpyruvate, 0.5 mM NADH, 10 U myokinase , 6 U lactate dehydrogenase, 4 U pyruvate kinase and 7.2 µg purified AlkK. Different concentrations of 4-hydroxybutyrate were used to start the reaction. The consumption of NADH was observed at 30°C and 360 nm ($\Delta\varepsilon = 3.1\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany).

### 7.9.4. Activity assay of propionyl-CoA carboxylase (PccAB)

Carboxylation activity was measured in 200 µl of 100 mM HEPES pH 8.0 buffer containing 2 mM ATP, 5 mM $MgCl_2$, 100 mM $KHCO_3$, 0.25 mM NADPH, 0.65 µg purified propionyl-CoA carboxylase. The reaction was started with different concentrations of propionyl-CoA and the consumption of NADH was followed at 30 °C and 340 nm ($\Delta\varepsilon = 6.22\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany).

### 7.9.5. Activity assay of propionyl-CoA carboxylase D407I variant

Carboxylation activity was measured in 150 µl of 50 mM Tris pH 7.5 buffer containing 0.5 mM ATP, 8 mM $MgCl_2$, 40 mM $KHCO_3$, 0.2 mM NADH, 1 µl carbonic anhydrase (0.2 mg/ml), 24.6 µg Mcr and either 2.8 µg of Pcc (D407I) or 3.4 µg wt Pcc. The reaction was started with different concentrations of acetyl-CoA and the consumption of NADPH was followed at 30 °C and 340 nm ($\Delta\varepsilon = 6.22\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany).

### 7.9.6. Activity assay of crotonyl-CoA carboxylase/reductase

Carboxylation activity was measured in 100 µl of 100 mM phosphate buffer pH 8.0 containing 3 mM NADPH, 100 mM $NaHCO_3$ and 0.39 µg of Ccr. The reaction was started by addition of crotonyl-CoA and the consumption of NADPH was followed at 30°C and 340 nm ($\Delta\varepsilon = 6.22\ mM^{-1}\ cm^{-1}$) with a Cary 60 UV-

Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany). Michaelis-Menten kinetic parameters were determined by varying the concentration of crotonyl-CoA from 0.025 mM to 1.6 mM.

### 7.9.7. Activity assay of crotonyl-CoA carboxylase/reductase (R152A)

Carboxylation activity was measured in 100 µl of 77.5 mM Tris pH 8.0 containing 2.1 mM NADPH, 53.7 mM $NaHCO_3$ and 0.87 µg of Ccr (R152A). The reaction was started by addition of crotonyl-CoA and the consumption of NADPH was followed at 30°C and 340 nm ($\Delta\varepsilon$ = 6.22 $mM^{-1}$ $cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany). Michaelis-Menten kinetic parameters were determined by varying the concentration of crotonyl-CoA from 0.27 mM to 4.3 mM.

### 7.9.8. Activity assay of glycerate kinase (GlxK)

The activity of GlxK was measured in a coupled assay in potassium phosphate buffer (50 mM, pH 7.3) containing 10 mM $MgCl_2$, 0.5 mM ATP, 2.5 mM PEP, 0.2 mM NADH, 6 U lactate dehydrogenase, 4 U pyruvate kinase and 0.18 µg of GlxK. The reaction was started by addition of glycerate and consumption of NADH was followed at 30°C and 340 nm ($\Delta\varepsilon$ = 6.22 $mM^{-1}$ $cm^{-1}$) with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany). Michaelis-Menten kinetic parameters were determined by varying the concentration of glycerate from 15.625 µM to 1000 µM.

### 7.9.9. Activity assay of methylsuccinyl-CoA oxidase (Mco)

Oxidase activity screening of the enzyme variants was performed by product (mesaconyl-CoA) detection with HPLC-MS (1260 Infinity, Agilent Technologies GmbH)/MS (6130 quadrupole MS, Agilent Technologies GmbH) using the method described before *(41)*. The assay was prepared in 250 µl of 80 mM ammonium carbonate buffer containing 50 mM $NaHCO_3$ pH 7.5, 600 µM NADPH, 110 U crotonyl-CoA carboxylase/reductase (Ccr), 75 U ethylmalonyl-CoA mutase (Ecm), 2200 U epimerase (Epi) and 15 µg methylsuccinyl-CoA dehydrogenase (Mcd) variant. The reaction was started with 400 µM of crotonyl-CoA and samples were withdrawn at 5 min and 30 min. The reaction was quenched in a final concentration of 1% HCl. The Michaelis-Menten parameters were determined by a spectrophotometric assay on a Cary 60 UV-Vis spectrometer (Agilent Technologies GmbH, Waldbronn, Germany) by observing the formation of mesaconyl-CoA at 30°C and 290 nm ($\Delta\varepsilon$ = 3.1 $mM^{-1}$ $cm^{-1}$) starting with different concentration (from 2.5 – 250 µM) of (*S*)-methylsuccinyl-CoA. The reaction was performed in 100 µl of 37.5 mM phosphate buffer pH 7.5 containing 120 µg of Mcd variant.

### 7.9.10. Activity assay of methylsuccinyl-CoA dehydrogenase wt and variants

The activity of Mcd from *P. denitrificans* (PdMcd) and *R. sphaeroides* (RsMcd) was determined using the protocol based on the reduction of ferrocenium as electron acceptor, which has been described before (*94, 180*). The reaction was followed at 300 nm (adjusted $\Delta\varepsilon_{300}$=2.75 $mM^{-1}$ $cm^{-1}$) on a Carry 60 UV-Vis Spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany) and 30°C. The reaction mixture (0.1 ml) contained 50 mM potassium phosphate buffer (pH 7.5), 0.25 mM ferrocenium hexafluorophosphate. For the assay with succinyl-CoA, the assay contained 1.6 µM of the RsMcd wt or the RsMcd Ala282Val variant (corresponds to Ala285 in PdMcd). This assay was started by the addition of different concentrations of succinyl-CoA (10 µM – 800 µM). For the activity of the enzymes with (2*S*)-methylsuccinyl-CoA, the reaction contained additionally 7.5 µM of ethylmalonlyl-CoA mutase and 3 µM of epimerase. The following enzyme concentrations were used in this case: 0.008 µM RsMcd, 0.01µM RsMcd Ala282Val (corresponds to Ala285 in PdMcd), 1.9 µM RsMcd Ala282Phe/Phe284Ala (corresponds to Ala285/Phe287 PdMcd) or 0.007 µM PdMcd wt. The assay was started by addition of varying substrate (ethylmalonyl-CoA) concentrations (10 µM – 1600 µM).

### 7.9.11. Activity assay of short chain acyl-CoA oxidase (ACX4)

Activity of the oxidase variants was measured with a spectrophotometric assay on a Cary 60 UV-Vis spectrophotometer (Agilent Technologies GmbH, Waldbronn, Germany). Product formation from propionyl-CoA ($\Delta\varepsilon = 2.0\ mM^{-1}\ cm^{-1}$) and 4-hydroxybutyryl-CoA ($\Delta\varepsilon = 1.6\ mM^{-1}\ cm^{-1}$) was observed at 30°C and 290 nm. The reaction was performed in 150 µl phosphate buffer (45 mM, pH 7.5) and was started with various concentrations of propionyl-CoA or 4-hydroxybutyryl-CoA.

## 7.10. HPLC-MS based enzyme assays

### 7.10.1. Assay of Mco (Mcd W315F, T317G, E377N) and Mcd T317G) with succinyl-CoA

The assay of Mco and Mcd T317G was performed in a reaction mixture (0.25 ml) containing 100 mM (MOPS) (pH 7.5) and either contained 172.5 µg of Mco or 172.5 µg of Mcd T. The reaction was started with the addition of 750 µM succinyl-CoA. Samples (40 µl) were withdrawn at certain time points and the reaction was quenched in 4 µl of 40% formic acid.

### 7.10.2. Assay of Mcd with ETF (titration)

The assay of Mcd wit ETF was performed in a reaction mixture (0.25 ml) containing 50 mM (MOPS) (pH 7.5), 50 mM NaCl, 40.5 µg KatE, 303 µg SodB and 5 µM of Mcd wt enzyme. Four different reaction conditions were prepared with either 2 mM ferrocenium hexafluorophosphate, 5 µM ETF, 50µM ETF or 250 µM ETF as electron acceptor. The reaction was started with the addition of 500 µM succinyl-CoA. Samples (30 µl) were withdrawn at certain time points and the reaction was quenched in 3 µl of 40% formic acid and 1 µl of polyphosphate (500 mM).

### 7.10.3. Assay of ETF variants

The assay of the ETF variants was performed in a reaction mixture (0.25 ml) containing 50 mM (MOPS) (pH 7.5), 1.8 mM ferrocenium hexafluorophosphate, and 34 µg Mcd A282V variant. Two different reaction conditions were prepared with either 13.6 µg of Mch or 15.6 µg of PhaJ as hydratases. The reaction was started by the addition of 682 µM succinyl-CoA. Samples (40 µl) were withdrawn at certain time points and the reaction was quenched in 4 µl of 40% formic acid and 1 µl of polyphosphate (500 mM).

### 7.10.4. Assay of the potential fumaryl-CoA hydratases for the formation of malyl-CoA

The assay of the hydratases was performed in a reaction mixture (0.275 ml) containing 91 mM (MOPS) (pH 7.5), 250 mM NaCl, 250 mM KCl and 5µM Mcd wt. Three different reaction conditions were prepared with either 50 µM of ETF wt, 50 µM of ETF D125N or 50 µM of ETF R188C. The reaction was started by the addition of 500 µM succinyl-CoA. Samples (30 µl) were withdrawn at certain time points and the reaction was quenched in 3 µl of 40% formic acid and 1 µl of polyphosphate (500 mM).

### 7.10.5. Assay of Mcd and Mcd variants with succinyl-CoA and (2*S*)-methylsuccinyl-CoA

For the assay of the Arg252Ala/Gln/Lys variants of Mcd from *P. denitrificans* (PdMcd), single variants of Ala282Val/Leu/Ile/Phe of Mcd of *R. sphaeroides* (RsMcd) (corresponds to Ala285 in PdMcd) and the double variants of RsMcd (Ala282Phe with Phe284Ala/Val/Leu; corresponds to Ala285 and Phe287 in PdMcd) the reaction mixture (275 µl) contained 91 mM 3-morpholino-propanesulfonic acid (MOPS) at pH 7.5 and 1.8 mM ferrocenium hexafluorophosphate. The assay with succinyl-CoA contained 34µg of enzyme and was started with the addition of 680 µM succinyl-CoA. The assay with the natural substrate contained 1 µg Mcd variants, 110 µg ethylmalonyl-CoA mutase and 13 µg epimerase. To start the reaction, 680 µM ethylmalonyl-CoA was added. Samples (30 µl) were withdrawn at certain time

points and the reaction was quenched in 3 µl of 40% formic acid. The samples were analyzed by HPLC-MS described below.

## 7.11. CETCH cycle assays

### 7.11.1. Assay of CETCH 1.0 (discontinuous)

The discontinuous assay of the first CETCH cycle was performed in 250 µl of 180 mM ammonium formate buffer pH 7.5 containing 1 µl carbonic anhydrase (0.2 mg/ml), 1 mM ATP, 2 mM NADPH, 50 mM $NaHCO_3$, 0.5 mM CoA and 300 µM propionyl-CoA. All samples (35 µl each) that were withdrawn from the assay were quenched in formic acid with a final concentration of 2%. Between each addition of the following enzymes, the reaction mixture was incubated at 30°C for 10 min before taking a sample. First, a negative control sample was taken and the assay was started with the addition of 0.2 U PccAB. After sampling, 0.12 U Mcm, 2.2 U Epi, 0.62 U SucD, 0.02 U AKR7a2 and 0.14 U Nmar0206 were added and the assay was further incubated. Afterwards 0.14 U Nmar0207 and 2.1 U Ccr were added and the sample was withdrawn after 10 min. Subsequently, 0.07 U Ecm, 0.04 U Mcd and 100 µM ferrocenium hexafluorophosphate were added before further incubation. Eventually, 9.9 U Mch and 0.09 U Mcl1 were added and the reaction was incubated for 10 min at 30°C before withdrawing the last sample. The samples were analyzed for CoA-esters with the method described below.

### 7.11.2. Assay of CETCH 1.0 (continuous)

The continuous assay of the first CETCH cycle was performed in 600 µl buffer (124 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 5 mM ATP, 6.5 mM NADPH, 10 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaHCO_3$, 100 µM ferrocenium hexafluorophosphate and the enzymes with the corresponding amounts (**MM Table 3 and MM Table 4**). The reaction was started with the addition of 300 µM propionyl-CoA and incubated at 30°C. Samples (35 µl each) were withdrawn from the reaction mixture and quenched in formic acid with a final concentration of 2%. The samples were analyzed for CoA-esters with the method described below.

### 7.11.3. Assay of CETCH 2.0

The assay of CETCH 2.0 was performed in 400 µl buffer (125 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 5 mM ATP, 6.5 mM NADPH, 10 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaHCO_3$ and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). The reaction was started with the addition of 225 µM propionyl-CoA and incubated at 30°C. Samples (35 µl each) were withdrawn from the reaction mixture and quenched in formic acid with a final concentration of 2%. The samples were analyzed for CoA-esters with the method described below.

### 7.11.4. Assay of CETCH 3.0

The assay of CETCH 3.0 was performed in 550 µl buffer (118 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 5 mM ATP, 6.5 mM NADPH, 10 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 1 mM acetyl-CoA and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). The reaction was started with the addition of 250 µM propionyl-CoA and incubated at 30°C. Samples (40 µl each) were withdrawn from the reaction mixture and quenched in formic acid. After the 1 min and 2 min samples, 250 µM propionyl-CoA was additionally added. In general, all samples were quenched in formic acid with a final concentration of 2%. After the 30 min and 90 min samples, 1 mM of acetyl-CoA was added. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.5. Assay of CETCH 4.0

The assay of CETCH 4.0 was performed in 600 µl buffer (114 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 0.75 mM ATP, 4 mM NADPH, 2 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid, 1 mM acetyl-CoA and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). The reaction was started with the addition of 150 µM propionyl-CoA and incubated at 30°C. Samples (35 µl each) were withdrawn from the reaction mixture and quenched in formic acid to a final concentration of 2%. After the 30 min and 90 min samples, 1 mM ATP and acetyl-CoA each was added. The samples were analyzed for CoA-esters and malate with the method described below. Note that the buffer concentration from CETCH 1.0 to 4.0 slightly varied, due to the dilution of the assay to the desired volume from a 200 µM stock buffer. In the following assays, the buffer concentration was fixed to 100 mM.

### 7.11.6. Assay of CETCH 5.0

The assay of CETCH 5.0 was performed in 600 µl buffer (100 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 2 mM ATP, 4 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid, 1 mM acetyl-CoA, 20 mM polyphosphate and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). The reaction was started with the addition of 150 µM propionyl-CoA and incubated at 30°C. Samples (30 µl each) were withdrawn from the reaction mixture and quenched in formic acid to a final concentration of 2%. After the 5 min, 15 min, 30 min and 60 min samples additional enzymes were added each (Ecm: 18 mU, Mcm: 60 mU and Nmar0206: 72 mU). Additionally after the 60 min sample, 1 mM of acetyl-CoA was added. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.7. Assay of CETCH 5.1

The assay of CETCH 5.1 was performed in 520 µl buffer (100 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 2 mM ATP, 4 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid, 1 mM acetyl-CoA, 20 mM polyphosphate and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). To test the influence of different Ccrs, either 3.11 U/ml Ccr (*M. extorquens*), 1.0 U/ml Ccr (*K. setae*) or 1.5 U/ml Ccr (*L. borgpetersenii*) was added. The reaction was started with the addition of 150 µM propionyl-CoA and incubated at 30°C. Samples were withdrawn (30 µl each) from the reaction mixture and quenched in formic acid to a final concentration of 4%. After the 5 min, 15 min, 30 min and 60 min samples additional enzymes were added each (Ecm: 18 mU, Mcm: 60 mU and Nmar0206: 60 mU). Additionally after the 60 min sample, 1 mM of acetyl-CoA was added. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.8. Assay of CETCH 5.2

This assay was performed in the same way as CETCH 5.4 (see below). Additionally, the assay buffer contained 10 mM $CaCl_2$. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.9. Assay of CETCH 5.3

The purified enzymes for the CETCH 5.3 assay were pooled (double amounts of the assay; except Nmar0206, which was already stored in 25 mM MOPS pH 7.5 buffer) and the total concentration was determined by Nanodrop 2000 (Thermo Scientific, St. Leon-Rot, Germany). The storage buffer was exchanged to 25 mM MOPS pH 7.5 containing 250 mM NaCl and 10% Glycerol by using an Amicon Ultracel 10K centrifugal filter (Merck Millipore, Darmstadt, Germany) to wash the enzymes multiple

times. The enzyme solution was concentrated to the same concentration as before the buffer exchange. The assay was performed in 520 µl buffer (100 mM ammonium formate, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 2 mM ATP, 4 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid, 1 mM acetyl-CoA, 20 mM polyphosphate, 20 mM formate, the enzymes (exchanged buffer) and Nmar0206 with the corresponding amounts (**MM Table 3** and **MM Table 4**). Samples were withdrawn (30 µl each) from the reaction mixture and quenched in formic acid to a final concentration of 4%. After the 60 min sample, 1 mM of acetyl was added to the reaction mixture. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.10. Assay of CETCH 5.4

The assay of CETCH 5.4 was performed in 520 µl buffer (100 mM MOPS KCl, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 2 mM ATP, 4 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 5 mM ascorbic acid, 1 mM acetyl-CoA, 20 mM polyphosphate, 20 mM formate and the enzymes with the corresponding amounts (**MM Table 3** and **MM Table 4**). The reaction was started with the addition of 200 µM propionyl-CoA and incubated at 30°C. Samples were withdrawn (30 µl each) from the reaction mixture and quenched in formic acid to a final concentration of 4%. The samples were analyzed for CoA-esters and malate with the method described below.

### 7.11.11. Further assays of CETCH 5.4

For further investigation of the CETCH 5.4, the assay was performed using the protocol and conditions mentioned above for CETCH 5.4. To test the effect of the protecting agents for oxidative damage, either catalase, coenzyme $B_{12}$ or ascorbic acid was replaced by water. Further, the addition of 50 mM $NaH^{13}CO_2$ was omitted and replaced by water to check if carbon fixation can be achieved at ambient $CO_2$ concentration. To investigate why the production of malate ceases, either 5 mM ATP or 200 µM propionyl-CoA was added to the reaction after 60 min to restart the process. Additionally one assay contained 500 µM malate before the start of the assay. Samples were withdrawn (30 µl each) from the reaction mixture and quenched in formic acid to a final concentration of 4%. Malate was quantified using the method described below.

### 7.11.12. Lists of enzyme activities and amounts in the CETCH cycle

**MM Table 3 List of enzyme activities in CETCH assays.**

| Rxn Nr. | Abbr. | Tag | U $ml^{-1}$ in CETCH 1.0 | U $ml^{-1}$ in CETCH 2.0 | U $ml^{-1}$ in CETCH 3.0 | U $ml^{-1}$ in CETCH 4.0 | U $ml^{-1}$ in CETCH 5.0 |
|---|---|---|---|---|---|---|---|
| 1 | pco | Pco | - | - | - | - | 1.71 |
| 1→2 | pcc | PccAB | 2.41 | 2.30 | 2.41 | 1.02 | - |
| 2+8 | ccr | Ccr | - | - | - | - | - |
| 2+8 | ccr | Ccr (R152A) | 10.9 | 10.5 | 11.2 | 5.3 | 4.2 |
| 3+9 | epi | Epi | 3.67 | 5.50 | 2.00 | 3.67 | 3.67 |
| 3 | mcm | Mcm | 0.40 | 0.30 | 0.28 | 0.42 | 0.60 |
| 4 | scr | SucD | 2.91 | 2.88 | 2.16 | 3.60 | 2.17 |
| 5 | ssr | AKR7a2 | 0.10 | 0.11 | 0.10 | 0.10 | 0.10 |
| 6 | hbs | Nmar0206 | 0.61 | 0.61 | 0.58 | 0.90 | 0.80 |
| 7 | hbd | Nmar0207 | 0.48 | 0.42 | 0.28 | 0.52 | 0.39 |
| 9 | ecm | Ecm | 0.29 | 0.28 | 0.16 | 0.29 | 0.28 |
| 10 | mco | Mco | - | 0.14 | 0.13 | 0.13 | 0.13 |
| 10 | mcd | Mcd | 0.20 | - | - | - | - |
| 11 | mch | Mch | 16.50 | 24.75 | 20.18 | 16.50 | 16.50 |
| 12 | mcl | Mcl1 | 0.50 | 0.75 | 0.50 | 0.50 | 0.50 |
| 13 | mas | GlcB | - | - | 0.72 | 1.10 | 1.10 |
| 14 | mct | Mcl2 | - | - | - | 1.10 | 1.10 |
| - | fdh | Fdh | - | - | 0.76 | 0.75 | 0.93 |
| - | pkk | Pkk2 | - | - | - | - | 0.92 |
| - | kat | KatE | - | - | - | 1174.00 | 1761.00 |

| Rxn Nr. | Abbr. | Tag | U $ml^{-1}$ in CETCH 5.1 | U $ml^{-1}$ in CETCH 5.2 | U $ml^{-1}$ in CETCH 5.3 | U $ml^{-1}$ in CETCH 5.4 |
|---|---|---|---|---|---|---|
| 1 | pco | Pco | 1.73 | 1.73 | 1.73 | 1.73 |
| 1→2 | pcc | PccAB | - | - | - | - |
| 2+8 | ccr | Ccr | - | 3.11 | 3.11 | 3.11 |
| 2+8 | ccr | Ccr (R152A) | - | - | - | - |
| 3+9 | epi | Epi | 4.23 | 4.23 | 4.23 | 4.23 |
| 3 | mcm | Mcm | 0.58 | 0.58 | 0.58 | 0.58 |
| 4 | scr | SucD | 2.95 | 2.88 | 2.95 | 2.88 |
| 5 | ssr | AKR7a2 | 0.10 | 0.10 | 0.10 | 0.10 |
| 6 | hbs | Nmar0206 | 0.80 | 0.80 | 0.80 | 0.80 |
| 7 | hbd | Nmar0207 | 0.94 | 0.92 | 0.94 | 0.92 |
| 9 | ecm | Ecm | 0.28 | 0.28 | 0.28 | 0.28 |
| 10 | mco | Mco | 0.13 | 0.13 | 0.13 | 0.13 |
| 10 | mcd | Mcd | - | - | - | - |
| 11 | mch | Mch | 19.04 | 19.04 | 19.04 | 19.04 |
| 12 | mcl | Mcl1 | 0.50 | 0.50 | 0.50 | 0.50 |
| 13 | mas | GlcB | 1.01 | 1.01 | 1.01 | 1.01 |
| 14 | mct | Mcl2 | 0.76 | 0.76 | 0.76 | 0.76 |
| - | fdh | Fdh | 0.73 | 0.74 | 0.73 | 0.74 |
| - | pkk | Pkk2 | 0.74 | 0.74 | 0.74 | 0.74 |
| - | kat | KatE | 1523.9 | 1523.9 | 1523.9 | 1523.9 |

**MM Table 4 List of enzyme amounts in CETCH assays**

| Rxn Nr. | Abbr. | Tag | µg $ml^{-1}$in CETCH 1.0 | µg $ml^{-1}$in CETCH 2.0 | µg $ml^{-1}$in CETCH 3.0 | µg $ml^{-1}$in CETCH 4.0 | µg $ml^{-1}$in CETCH 5.0 |
|---|---|---|---|---|---|---|---|
| 1 | pco | Pco | - | - | - | - | 143 |
| 1→2 | pcc | PccAB | 101 | 99 | 100 | 43 | - |
| 2+8 | ccr | Ccr | - | - | - | - | - |
| 2+8 | ccr | Ccr (R152A) | 128 | 124 | 132 | 62 | 50 |
| 3+9 | epi | Epi | 8 | 13 | 5 | 8 | 8 |
| 3 | mcm | Mcm | 20 | 15 | 14 | 21 | 30 |
| 4 | scr | SucD | 100 | 99 | 75 | 124 | 75 |
| 5 | ssr | AKR7a2 | 27 | 29 | 25 | 26 | 25 |
| 6 | hbs | Nmar0206 | 303 | 306 | 291 | 450 | 400 |
| 7 | hbd | Nmar0207 | 18 | 16 | 11 | 20 | 15 |
| 9 | ecm | Ecm | 42 | 40 | 23 | 42 | 40 |
| 10 | mco | Mco | - | 1350 | 1309 | 1342 | 1333 |
| 10 | mcd | Mcd | 34 | - | - | - | - |
| 11 | mch | Mch | 11 | 17 | 13 | 11 | 11 |
| 12 | mcl | Mcl1 | 100 | 150 | 101 | 100 | 100 |
| 13 | mas | GlcB | - | - | 20 | 30 | 30 |
| 14 | mct | Mcl2 | - | - | 0 | 6 | 6 |
| - | fdh | Fdh | - | - | 545 | 533 | 667 |
| - | pkk | Pkk2 | - | - | - | - | 77 |
| - | kat | KatE | - | - | - | 100 | 150 |

| Rxn Nr. | Abbr. | Tag | µg $ml^{-1}$in CETCH 5.1 | µg $ml^{-1}$in CETCH 5.2 | µg $ml^{-1}$in CETCH 5.3 | µg $ml^{-1}$in CETCH 5.4 |
|---|---|---|---|---|---|---|
| 1 | pco | Pco | 144 | 144 | 144 | 144 |
| 1→2 | pcc | PccAB | - | - | - | - |
| 2+8 | ccr | Ccr | (see methods) | 29 | 29 | 29 |
| 2+8 | ccr | Ccr (R152A) | - | - | - | - |
| 3+9 | epi | Epi | 10 | 10 | 10 | 10 |
| 3 | mcm | Mcm | 29 | 29 | 29 | 29 |
| 4 | scr | SucD | 102 | 99 | 102 | 99 |
| 5 | ssr | AKR7a2 | 25 | 25 | 25 | 25 |
| 6 | hbs | Nmar0206 | 400 | 400 | 401 | 400 |
| 7 | hbd | Nmar0207 | 36 | 35 | 36 | 35 |
| 9 | ecm | Ecm | 40 | 40 | 40 | 40 |
| 10 | mco | Mco | 1279 | 1300 | 1314 | 1300 |
| 10 | mcd | Mcd | - | - | - | - |
| 11 | mch | Mch | 13 | 13 | 13 | 13 |
| 12 | mcl | Mcl1 | 99 | 99 | 99 | 99 |
| 13 | mas | GlcB | 28 | 28 | 28 | 28 |
| 14 | mct | Mcl2 | 4 | 4 | 4 | 4 |
| - | fdh | Fdh | 523 | 528 | 528 | 528 |
| - | pkk | Pkk2 | 62 | 62 | 62 | 62 |
| - | kat | KatE | 130 | 130 | 130 | 130 |

## 7.12. HOPAC cycle assays

### 7.12.1. Assay of HOPAC 1.0

The assay of HOPAC 1.0 was performed in 300 µl buffer (50 mM MOPS, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 3 mM ATP, 3 mM NADPH, 3 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$ and the enzymes with the corresponding amounts (**MM Table 5** and **MM Table 6**). The reaction was started with the addition of 500 µM acetyl-CoA and incubated at 30°C. Samples (30 µl each) were withdrawn from the reaction mixture and quenched in 3 µl of 40% formic acid. The samples were analyzed for CoA-esters with the method described below.

### 7.12.2. Assay of HOPAC 2.0

The assay of HOPAC 1.0 was performed in 300 µl buffer (50 mM MOPS, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 3 mM ATP, 3 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$ and the enzymes with the corresponding amounts (**MM Table 5** and **MM Table 6**). The reaction was started with the addition of 250 µM acetyl-CoA and incubated at 30°C. Samples (30 µl each) were withdrawn from the reaction mixture and quenched in 3 µl of 40% formic acid. The samples were analyzed for CoA-esters with the method described below.

### 7.12.3. Assay of HOPAC 3.0 and HOPAC 4.0

The assay of HOPAC 1.0 was performed in 300 µl buffer (50 mM MOPS, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 3 mM ATP, 3 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 20 mM formate and, 20 mM polyphosphate and the enzymes with the corresponding amounts (**MM Table 5** and **MM Table 6**). The reaction was started with the addition of 250 µM acetyl-CoA and incubated at 30°C. Samples (30 µl each) were withdrawn from the reaction mixture and quenched in 3 µl of 40% formic acid. The samples were analyzed for CoA-esters with the method described below.

### 7.12.4. Assay of HOPAC 4.1

The assay of HOPAC 1.0 was performed in 300 µl buffer (50 mM MOPS, pH 7.5) containing 1 µl carbonic anhydrase (0.2 mg/ml), 3 mM ATP, 3 mM NADPH, 5 mM $MgCl_2$, 0.5 mM CoA, 50 mM $NaH^{13}CO_3$, 0.1 mM coenzyme $B_{12}$, 20 mM formate and, 20 mM polyphosphate and the enzymes with the corresponding amounts (**MM Table 5** and **MM Table 6**). The reaction was started with the addition of 250 µM acetyl-CoA and incubated at 30°C. After 20 min 19 µg of Mcm, 0.5 µl CoA (100 mM) and 2 µl of coenzyme $B_{12}$ (5 mM) were added. Samples (30 µl each) were withdrawn from the reaction mixture and quenched in 3 µl of 40% formic acid. The samples were analyzed for CoA-esters with the method described below.

### 7.12.5. List of enzyme activities and amounts in the HOPAC cycle

**MM Table 5 List of enzyme activities in HOPAC assays**

| Rxn Nr. | Abbr. | Tag | U $ml^{-1}$ in HOPAC 1.0 | U $ml^{-1}$ in HOPAC 2.0 | U $ml^{-1}$ in HOPAC 3.0 | U $ml^{-1}$ in HOPAC 4.0 | U $ml^{-1}$ in HOPAC 4.1 |
|---|---|---|---|---|---|---|---|
| 1 | pcc | Pcc | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1 | mpc | Pcc D407I | 0.00 | 1.05 | 1.05 | 1.05 | 1.05 |
| 2 | mcr | Mcr | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 |
| 3 | hps | Nmar1309 | 0.30 | 0.60 | 0.60 | 0.60 | 0.61 |
| 4 | ech | PhaJ | 318.27 | 636.53 | 636.53 | 636.53 | 636.53 |
| 5 | ccr | Ccr | 1.61 | 6.16 | 6.16 | 6.16 | 6.16 |
| 6 | mcm | Mcm | 0.55 | 1.27 | 1.27 | 1.27 | 1.27 |
| 6 | epi | Epi | 7.92 | 7.92 | 7.92 | 7.92 | 7.92 |
| 7 | mcd | Mcd | 0.89 | 0.99 | 0.99 | 0.99 | 0.99 |
| 8 | mch | Mch_RS | 301.50 | 301.50 | 301.50 | 301.50 | 301.50 |
| 9 | mcl | Mcl1 | 0.51 | 0.53 | 0.53 | 0.53 | 0.53 |
| - | fdh | Fdh | - | - | 0.91 | 0.91 | 0.91 |
| - | pkk | Pkk2 | - | - | 0.98 | 0.98 | 0.98 |
| - | kat | KatE | - | - | - | 1267.92 | 1267.92 |
| - | sod | SodA | - | - | - | n.a | n.a |
| - | sod | SodB | - | - | - | n.a | n.a |
| 14 | etf | EtfAB | n.a. | n.a | n.a | n.a | n.a |

**MM Table 6 List of enzyme amounts in HOPAC assays**

| Rxn Nr. | Abbr. | Tag | µg $ml^{-1}$ in HOPAC 1.0 | µg $ml^{-1}$ in HOPAC 2.0 | µg $ml^{-1}$ in HOPAC 3.0 | µg $ml^{-1}$ in HOPAC 4.0 | µg $ml^{-1}$ in HOPAC 4.1 |
|---|---|---|---|---|---|---|---|
| 1 | pcc | Pcc | 249 | - | - | - | - |
| 1 | mpc | Pcc D407I | 0 | 950 | 950 | 950 | 950 |
| 2 | mcr | Mcr | 68 | 68 | 68 | 68 | 68 |
| 3 | hps | Nmar1309 | 500 | 997 | 997 | 997 | 1015 |
| 4 | ech | PhaJ | 51 | 103 | 103 | 103 | 103 |
| 5 | ccr | Ccr | 15 | 56 | 56 | 56 | 56 |
| 6 | mcm | Mcm | 28 | 63 | 63 | 63 | 63 |
| 6 | epi | Epi | 18 | 18 | 18 | 18 | 18 |
| 7 | mcd | Mcd | 2960 | 3290 | 3290 | 3290 | 3290 |
| 8 | mch | Mch_RS | 201 | 201 | 201 | 201 | 201 |
| 9 | mcl | Mcl1 | 102 | 106 | 106 | 106 | 106 |
| - | fdh | Fdh | - | - | 653 | 653 | 653 |
| - | pkk | Pkk2 | - | - | 81 | 81 | 81 |
| - | kat | KatE | - | - | - | 108 | 108 |
| - | sod | SodA | - | - | - | 52 | 52 |
| - | sod | SodB | - | - | - | 404 | 404 |
| 14 | etf | EtfAB | 330 | 2964 | 2964 | 2964 | 2964 |

## 7.13. UPLC-MS analysis

### 7.13.1. UPLC-high resolution MS of CoA esters

CoA esters were analyzed 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-18 column (50 mm x 2.1 mm, particle size 1.7 µm, Kinetex XB-C18, Phenomenex) using a mobile phase system comprised of 50 mM ammonium formate pH 8.1 and methanol. Chromatographic separation was carried out using the following gradient condition at a flow rate of 250 µl $min^{-1}$: 0 min 0% methanol; 1 min 0% methanol, 3 min 2.5% methanol; 9 min 23% methanol; 14 min 80 % methanol; 16 min 80% methanol. 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 200-1200 m/z.LC-MS data were analyzed using MassHunter Qualitative Analysis software (Agilent) and eMZed (*181*).

### 7.13.2. UPLC-high resolution MS of malic acid

The quenched samples were centrifuged for 10 min at 17,000 rcf, to remove precipitated proteins. The samples were diluted 1:4 in $H_2O$ for the injection into the HPLC-MS system. Standard curves for quantification (2 µM, 10 µM, 25 µM, 50 µM, 100 µM, 150 µM) were prepared. Malic acid was separated and detected using an Agilent 6550 iFunnel Q-TOF LC-MS system equipped with an electrospray ionization source set to negative ionization mode. The LC conditions were adapted from a previously reported method (*182*). The modified LC method is as follows: The analyte was separated on an aminopropyl column (100 mm x 2 mm, particle size 3 µm, 100 Å, Luna $NH_2$, Phenomenex) using a mobile phase system comprised of 95:5 20 mM ammonium acetate pH 9.3 (adjusted with ammonium hydroxide to a final concentration of approximately 10 mM) / acetonitrile (A) and acetonitrile (B). Chromatographic separation was carried out using the following gradient condition at a flow rate of 250 µl $min^{-1}$: 0 min 85% B; 7 min 0% B, 14 min 0% B; 15 min 85% B; 17.5 min 85% B. Column oven and autosampler temperature were maintained at 15°C. The ESI source was set to the following parameters: Capillary voltage was set at 3.5 kV and nitrogen gas was used as nebulizing (20 psig), drying (13 l/min, 225 C) and sheath gas (12 l/min, 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*. Autorecalibration was carried out using 113 *m/z* as reference mass. Subsequent peak integration of the malic acid isotopologues was performed using eMZed (*181*).

## 7.14. Optimized sequences

**Optimized sequence of 4-hydroxybutyryl-CoA dehydratase from *N. maritimus* (nmar0207_opt)**

5’-GGATCCATGGCCAATGTCCTTAAAACCATCCGTAGTGGGGACGACTACATTGAGTCACTGCGTGGTCGTGACCTGAAAGTCTATCTCT
TTGGCGAGTTGGTTAAGGAACCGGTTGATCATCCCATGATTCGCCCGTCGATTAACGCGGTAGCGGAAACCTACGACTTAGCTCTGCGTG
AAGAAGCGCTTGCGTCAGCCGATTCGTCAATCACTGGTCTGAAAGTAAATCGCTTCCTGCATATCGCTGAATCCGCCGAGGATCTGGTTC
TGCAAAACAAAATGCAGCGTAAACTGGGCCAGAATACCGGCACCTGTTTTCAGCGGTGTGTGGGCATGGATGCAATGAACTCGCTGCAT
AGTACTACCTTCGAAATCGATGAAAAGCACGGTACGGACTATCACAAACGGTTTCTGGAGTTCGTGAAGATGGTGCAGCAAGAGAATCT
GGTGATTGGCGGAGCAATGACCGATCCGAAAGGCGATCGCAGCAAAGGTCCCTCCGAACAAGATGATCCGGATCTGTTTACCCGCATTG
TGGATACAGACGAGAAAGGAGTGTACGTGAGTGGCGCGAAAGCACACCAGACGGGCTGTATTAACAGCCATTGGATTATCTTAATGCC
GACGATTCGTTTGACAGAAAGCGATAAAGATTGGGCGATTGTCGGTGCCATTCCAGCCGATGCGAAAGGCGTGACGTATATCTACGGTC
GTCAAAGCTGCGATACTCGCAGCATGGAGGAAGGGGATATCGATGATGGCAACGCGAAATTTGGTGGGCAGGAAGCGTTGATCATTCT
GGATCGCGTTTTCATCCCGTGGGACAAAGTGTTTATGCATGGCGAATACGAATTTGCCAGCATGCTGGTTGAACGCTTTACGTGCTATCA
TCGTCGCAGCTATGTGTGCAAAACGGGTCTGGGTGATGTACTGATCGGTGCAGCGGCTACCATTGCCGACTATAATGGCGTCCCTAAAG
TTTCGCACATCAAAGACAAAATTATCGAGATGACCCATCTCAACGAAACCATTTTCGCAGCGGGTATTGCATCCTCTCATCAAGGGCAGA
AAATGAAATCAGGAGTATATCTGAACGATGATATGCTCGCCCAGGTCTGCAAACACAACGTAACACGCTTCCCGTACGAAATCTCTCGCT
TAGCACAGGACATTGCTGGCGGCCTGGTCGTGACGTTGCCTTCTGAGAAGGACTTTCGTCACCCAGAAGCCGGTCCTTTGCTTAAGAAAT
ATCTGGCGGGTCGCAAAGGCGTGGACGTTGAAAACCGCATGCGCATTCTGCGGTTAATCGAAAATATGACACTTGGCCGTAATGCTGTT
GGATATCTGACCGAGAGTATGCATGGTGCTGGGTCGCCACAGGCACAACGCATTCAGATTCAGCGTCAGATGCAAGTTGGGTATAAGAA
GAATCTCGCGAAGAATTTAGCCGGCATTACTAACGATGTGGAAGAACCGAAAGAATCCAGCGAATACTTCAAACGCGTCTTTAAAACCA
AAGACTCTGTCCTGTAAGAATTC-3’

**Optimized sequence of short chain acyl-CoA oxidase 4 from *A. thaliana* (acx4_opt)**

5’-CATATGGCAGTGTTATCGTCAGCGGATCGTGCGTCTAACGAAAAGAAAGTGAAATCCAGCTATTTCGATCTTCCGCCTATGGAGATGT
CTGTGGCCTTTCCACAAGCAACACCAGCCAGCACTTTTCCACCGTGTACGAGCGATTACTATCACTTCAACGATCTGCTCACCCCGGAAGA
ACAGGCGATCCGCAAGAAAGTGCGCGAATGCATGGAAAAGGAAGTAGCTCCGATCATGACAGAGTACTGGGAGAAAGCCGAATTTCCG
TTTCACATTACCCCGAAACTGGGTGCAATGGGAGTGGCTGGTGGGAGCATCAAAGGCTATGGATGCCCCGGCTTGTCGATTACCGCGAA
TGCTATTGCCACCGCAGAGATTGCACGCGTTGATGCCAGTTGCTCAACGTTCATTCTGGTCCATTCAAGCCTGGGTATGCTGACCATCGCT
TTGTGTGGCTCTGAAGCGCAGAAAGAGAAATACTTACCGTCCTTGGCGCAACTGAACACTGTAGCGTGTTGGGCCCTGACGGAACCGGA
TAACGGCTCTGACGCGTCGGGTCTGGGTACCACAGCAACCAAAGTTGAAGGCGGTTGGAAAATCAATGGTCAGAAACGCTGGATTGGC
AATAGCACGTTTGCGGATCTGCTCATTATCTTTGCCCGTAATACGACGACTAACCAGATTAACGGCTTTATCGTGAAGAAGGATGCTCCA
GGCCTGAAAGCGACCAAGATTCCGAACAAAATCGGTCTGCGCATGGTTCAGAATGGCGACATTCTTCTGCAGAACGTGTTCGTTCCGGA
TGAGGATCGCTTGCCCGGCGTCAACTCCTTCCAAGACACCAGTAAAGTGTTAGCCGTTAGCCGTGTCATGGTTGCCTGGCAACCTATCGG
AATTTCGATGGGCATCTACGACATGTGTCATCGCTATCTGAAAGAACGGAAACAGTTTGGTGCTCCTTTAGCGGCCTTCCAGCTGAATCA
ACAGAAACTTGTGCAGATGCTGGGGAATGTACAAGCGATGTTTCTTATGGGCTGGCGTCTGTGCAAACTCTATGAAACGGGCCAGATGA
CTCCCGGACAGGCTTCGCTGGGGAAGGCATGGATTAGTTCCAAAGCCCGTGAAACCGCGTCATTAGGTCGTGAACTGCTGGGCGGCAAT
GGGATTCTGGCAGACTTTCTGGTCGCCAAAGCGTTCTGCGATTTGGAACCGATCTATACCTATGAGGGGACTTACGACATTAACACCCTC
GTAACCGGTCGCGAAGTCACGGGTATTGCGAGTTTCAAACCGGCAACACGCAGCCGGCTGTAAGAATTC -3’

# References


1. IPCC, "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: Geneva, Switzerland, 2014).
2. R. M. Cuellar-Franca and A. Azapagic, Carbon capture, storage and utilisation technologies: A critical analysis and comparison of their life cycle environmental impacts. *J. CO2 Util.* **9**, 82-102 (2015).
3. A. M. Appel *et al.*, Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO2 fixation. *Chem. Rev.* **113**, 6621-6658 (2013).
4. K. M. Yu *et al.*, Recent advances in CO2 capture and utilization. *ChemSusChem* **1**, 893-899 (2008).
5. I. A. Berg, Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925-1936 (2011).
6. G. Fuchs, Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631-658 (2011).
7. S. M. Glueck *et al.*, Biocatalytic carboxylation. *Chem. Soc. Rev.* **39**, 313-328 (2010).
8. L. Rosgaard *et al.*, Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants. *J. Biotechnol.* **162**, 134-147 (2012).
9. L. Brennan and P. Owende, Biofuels from microalgae-A review of technologies for production, processing, and extractions of biofuels and co-products. *Renew. Sust. Energ. Rev.* **14**, 557-577 (2010).
10. D. M. Etheridge *et al.*, Historical CO2 records from the Law Dome DE08, DE08-2, and DSS ice cores. In Trends: A Compendium of Data on Global Change. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A*, (1989).
11. G. Wachtershauser, Evolution of the 1st Metabolic Cycles. *Proc. Natl. Acad. Sci. U.S.A* **87**, 200-204 (1990).
12. F. R. Tabita *et al.*, Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* **71**, 576-599 (2007).
13. M. Calvin and A. A. Benson, The Path of Carbon in Photosynthesis. *Science* **107**, 476-480 (1948).
14. J. A. Bassham *et al.*, The Path of Carbon in Photosynthesis .8. The Role of Malic Acid. *J. Biol. Chem.* **185**, 781-787 (1950).
15. J. A. Bassham *et al.*, The Path of Carbon in Photosynthesis .21. The Cyclic Regeneration of Carbon Dioxide Acceptor. *J. Am. Chem. Soc.* **76**, 1760-1770 (1954).
16. G. Bowes *et al.*, Phosphoglycolate Production Catalyzed by Ribulose Diphosphate Carboxylase. *Biochem. Biophys. Res. Commun.* **45**, 716-722 (1971).
17. M. Eisenhut *et al.*, The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. *Plant Physiol.* **142**, 333-342 (2006).
18. M. Eisenhut *et al.*, The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. *Proc. Natl. Acad. Sci. U.S.A* **105**, 17199-17204 (2008).
19. J. Ormerod, 'Every dogma has its day'*: a personal look at carbon metabolism in photosynthetic bacteria. *Photosynthesis Res.* **76**, 135-143 (2003).
20. M. C. Evans *et al.*, A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. *Proc. Natl. Acad. Sci. U.S.A* **55**, 928-934 (1966).
21. U. Gehring and D. I. Arnon, Purification and properties of -ketoglutarate synthase from a photosynthetic bacterium. *J. Biol. Chem.* **247**, 6963-6969 (1972).
22. G. Antranikian *et al.*, Characterization of ATP citrate lyase from Chlorobium limicola. *J. Bacteriol.* **152**, 1284-1287 (1982).

23. L. G. Ljungdahl, The autotrophic pathway of acetate synthesis in acetogenic bacteria. *Annu. Rev. Microbiol.* **40**, 415-450 (1986).
24. S. W. Ragsdale and E. Pierce, Acetogenesis and the Wood-Ljungdahl pathway of CO2 fixation. *Biochim. Biophys. Acta* **1784**, 1873-1898 (2008).
25. R. K. Thauer *et al.*, Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579-591 (2008).
26. M. C. Weiss *et al.*, The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **1**, 16116 (2016).
27. S. W. Ragsdale, Enzymology of the wood-Ljungdahl pathway of acetogenesis. *Ann. N. Y. Acad. Sci.* **1125**, 129-136 (2008).
28. H. Holo, Chloroflexus-Aurantiacus Secretes 3-Hydroxypropionate, a Possible Intermediate in the Assimilation of CO2 and Acetate. *Arch. Microbiol.* **151**, 252-256 (1989).
29. M. Hugler *et al.*, Malonyl-coenzyme A reductase from Chloroflexus aurantiacus, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO2 fixation. *J. Bacteriol.* **184**, 2404-2410 (2002).
30. B. E. Alber and G. Fuchs, Propionyl-coenzyme A synthase from Chloroflexus aurantiacus, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO2 fixation. *J. Biol. Chem.* **277**, 12137-12143 (2002).
31. G. Strauss and G. Fuchs, Enzymes of a novel autotrophic CO2 fixation pathway in the phototrophic bacterium Chloroflexus aurantiacus, the 3-hydroxypropionate cycle. *Eur. J. Biochem.* **215**, 633-643 (1993).
32. S. Herter *et al.*, A bicyclic autotrophic CO2 fixation pathway in Chloroflexus aurantiacus. *J. Biol. Chem.* **277**, 20277-20283 (2002).
33. J. Zarzycki *et al.*, Identifying the missing steps of the autotrophic 3-hydroxypropionate CO2 fixation cycle in Chloroflexus aurantiacus. *Proc. Natl. Acad. Sci. U.S.A* **106**, 21317-21322 (2009).
34. M. Konneke *et al.*, Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc. Natl. Acad. Sci. U.S.A* **111**, 8239-8244 (2014).
35. I. A. Berg *et al.*, A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* **318**, 1782-1786 (2007).
36. I. A. Berg *et al.*, Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* **8**, 447-460 (2010).
37. H. Huber *et al.*, A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum Ignicoccus hospitalis. *Proc. Natl. Acad. Sci. U.S.A* **105**, 7851-7856 (2008).
38. T. J. Erb *et al.*, Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci. U.S.A* **106**, 8871-8876 (2009).
39. T. J. Erb, Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* **77**, 8466-8477 (2011).
40. S. A. Ensign *et al.*, New roles for CO2 in the microbial metabolism of aliphatic epoxides and ketones. *Arch. Microbiol.* **169**, 179-187 (1998).
41. B. Jobst *et al.*, ATP-dependent carboxylation of acetophenone by a novel type of carboxylase. *J. Bacteriol.* **192**, 1387-1394 (2010).
42. K. Schuhle and G. Fuchs, Phenylphosphate carboxylase: a new C-C lyase involved in anaerobic phenol metabolism in Thauera aromatica. *J. Bacteriol.* **186**, 4556-4567 (2004).
43. I. Olsen and J. M. Merrick, Identification of propionate as an endogenous CO2 acceptor in Rhodospirillum rubrum and properties of purified propionyl-coenzyme A carboxylase. *J. Bacteriol.* **95**, 1774-1778 (1968).
44. R. K. Thauer, Citric-acid cycle, 50 years on. Modifications and an alternative pathway in anaerobic bacteria. *Eur. J. Biochem.* **176**, 497-508 (1988).
45. H. L. Kornberg and H. A. Krebs, Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature* **179**, 988-991 (1957).
46. M. Khomyakova *et al.*, A methylaspartate cycle in haloarchaea. *Science* **331**, 334-337 (2011).

47. T. J. Erb *et al.*, Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. U.S.A* **104**, 10631-10636 (2007).
48. H. L. Kornberg, Anaplerotic Sequences in Microbial Metabolism. *Angew. Chem. Int. Ed. Engl.* **4**, 558-565 (1965).
49. J. E. Cronan and J. Thomas, Bacterial Fatty Acid Synthesis and Its Relationships with Polyketide Synthetic Pathways. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates* **459**, 395-433 (2009).
50. Y. A. Chan *et al.*, Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* **26**, 90-114 (2009).
51. M. C. Wilson and B. S. Moore, Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. *Nat. Prod. Rep.* **29**, 72-86 (2012).
52. D. M. Peter *et al.*, Screening and Engineering the Synthetic Potential of Carboxylating Reductases from Central Metabolism and Polyketide Biosynthesis. *Angew. Chem. Int. Ed. Engl.* **54**, 13457-13461 (2015).
53. D. L. Falcone and F. R. Tabita, Expression of endogenous and foreign ribulose 1,5-bisphosphate carboxylase-oxygenase (RubisCO) genes in a RubisCO deletion mutant of Rhodobacter sphaeroides. *J. Bacteriol.* **173**, 2099-2108 (1991).
54. J. B. McKinlay and C. S. Harwood, Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. *Proc. Natl. Acad. Sci. U.S.A* **107**, 11669-11675 (2010).
55. R. Laguna *et al.*, Acetate-dependent photoheterotrophic growth and the differential requirement for the Calvin-Benson-Bassham reductive pentose phosphate cycle in Rhodobacter sphaeroides and Rhodopseudomonas palustris. *Arch. Microbiol.* **193**, 151-154 (2011).
56. T. J. Erb and J. Zarzycki, Biochemical and synthetic biology approaches to improve photosynthetic CO2-fixation. *Curr. Opin. Chem. Biol.* **34**, 72-79 (2016).
57. R. J. Ellis, Most Abundant Protein in the World. *Trends Biochem. Sci.* **4**, 241-244 (1979).
58. M. R. Parikh *et al.*, Directed evolution of RuBisCO hypermorphs through genetic selection in engineered E.coli. *Protein Eng. Des. Sel.* **19**, 113-119 (2006).
59. D. N. Greene *et al.*, Artificially evolved Synechococcus PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochem. J.* **404**, 517-524 (2007).
60. N. E. Kreel and F. R. Tabita, Serine 363 of a Hydrophobic Region of Archaeal Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase from Archaeoglobus fulgidus and Thermococcus kodakaraensis Affects CO2/O2 Substrate Specificity and Oxygen Sensitivity. *PLoS One* **10**, e0138351 (2015).
61. G. G. Tcherkez *et al.*, Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. U.S.A* **103**, 7246-7251 (2006).
62. Y. Savir *et al.*, Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proc. Natl. Acad. Sci. U.S.A* **107**, 3475-3480 (2010).
63. J. A. Raven *et al.*, The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **363**, 2641-2650 (2008).
64. K. Kajala *et al.*, Strategies for engineering a two-celled C(4) photosynthetic pathway into rice. *J. Exp. Bot.* **62**, 3001-3010 (2011).
65. W. Bonacci *et al.*, Modularity of a carbon-fixing protein organelle. *Proc. Natl. Acad. Sci. U.S.A* **109**, 478-483 (2012).
66. P. M. Shih *et al.*, Introduction of a synthetic CO2-fixing photorespiratory bypass into a cyanobacterium. *J. Biol. Chem.* **289**, 9493-9500 (2014).
67. M. Mattozzi *et al.*, Expression of the sub-pathways of the Chloroflexus aurantiacus 3-hydroxypropionate carbon fixation bicycle in E. coli: Toward horizontal transfer of autotrophic growth. *Metab. Eng.* **16**, 130-139 (2013).

68. M. W. Keller *et al.*, Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proc. Natl. Acad. Sci. U.S.A* **110**, 5840-5845 (2013).
69. N. Antonovsky *et al.*, Sugar Synthesis from CO2 in Escherichia coli. *Cell* **166**, 115-125 (2016).
70. E. Noor *et al.*, Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput. Biol.* **10**, e1003483 (2014).
71. U. Barenholz *et al.*, Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. *Elife* **6**, (2017).
72. A. Bar-Even *et al.*, Thermodynamic constraints shape the structure of carbon fixation pathways. *Biochim. Biophys. Acta* **1817**, 1646-1659 (2012).
73. A. Bar-Even *et al.*, Design and analysis of synthetic carbon fixation pathways. *Proc. Natl. Acad. Sci. U.S.A* **107**, 8889-8894 (2010).
74. S. Billerbeck *et al.*, From understanding to designing enzymatic networks. *FEBS J.* **277**, 23-23 (2010).
75. J. R. Swartz *et al.*, Using cell-free biology to study systems biology. *Abstracts of Papers of the American Chemical Society* **227**, U255-U255 (2004).
76. J. A. Rollin *et al.*, New biotechnology paradigm: cell-free biosystems for biomanufacturing. *Green Chem.* **15**, 1708-1719 (2013).
77. P. H. Opgenorth *et al.*, A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nat. Chem. Biol.* **12**, 393-395 (2016).
78. T. P. Korman *et al.*, A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat. Commun.* **8**, 15526 (2017).
79. P. H. Opgenorth *et al.*, A synthetic biochemistry molecular purge valve module that maintains redox balance. *Nat. Commun.* **5**, 4113 (2014).
80. P. H. Opgenorth *et al.*, A molecular rheostat maintains ATP levels to drive a synthetic biochemistry system. *Nat. Chem. Biol.* **13**, 938-942 (2017).
81. S. P. Long *et al.*, Rising atmospheric carbon dioxide: plants FACE the future. *Annu. Rev. Plant Biol.* **55**, 591-628 (2004).
82. R. J. Spreitzer and M. E. Salvucci, Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.* **53**, 449-475 (2002).
83. L. Schada von Borzyskowski *et al.*, Evolutionary history and biotechnological future of carboxylases. *J. Biotechnol.* **168**, 243-251 (2013).
84. D. M. Peter, Substrate Promiscuity, Kinetics and Engineering of Enoyl-CoA Carboxylases/Reductases. *Doctoral Thesis, ETH Zurich, Switzerland*, (2016).
85. A. Chang *et al.*, BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Res.* **43**, D439-446 (2015).
86. M. Volpers *et al.*, Integrated In Silico Analysis of Pathway Designs for Synthetic Photo-Electro-Autotrophy. *PLoS One* **11**, e0157851 (2016).
87. T. J. Erb *et al.*, Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.* **37**, 56-62 (2017).
88. A. Flamholz *et al.*, eQuilibrator--the biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770-775 (2012).
89. L. Schada von Borzyskowski *et al.*, A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in Methylobacterium extorquens AM1 and other alphaproteobacteria. *ACS Synth. Biol.* **4**, 430-443 (2015).
90. P. C. Hinkle, P/O ratios of mitochondrial oxidative phosphorylation. *Biochim. Biophys. Acta* **1706**, 1-11 (2005).
91. D. M. Kramer and J. R. Evans, The importance of energy balance in improving photosynthetic productivity. *Plant Physiol.* **155**, 70-78 (2011).
92. T. J. Erb *et al.*, Ethylmalonyl-CoA mutase from Rhodobacter sphaeroides defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. *J. Biol. Chem.* **283**, 32283-32293 (2008).

93. B. Sohling and G. Gottschalk, Purification and characterization of a coenzyme-A-dependent succinate-semialdehyde dehydrogenase from Clostridium kluyveri. *Eur. J. Biochem.* **212**, 121-127 (1993).
94. T. J. Erb *et al.*, (2S)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol. Microbiol.* **73**, 992-1008 (2009).
95. J. Zarzycki *et al.*, Mesaconyl-coenzyme A hydratase, a new enzyme of two central carbon metabolic pathways in bacteria. *J. Bacteriol.* **190**, 1366-1374 (2008).
96. T. J. Erb *et al.*, The apparent malate synthase activity of Rhodobacter sphaeroides is due to two paralogous enzymes, (3S)-Malyl-coenzyme A (CoA)/{beta}-methylmalyl-CoA lyase and (3S)-Malyl-CoA thioesterase. *J. Bacteriol.* **192**, 1249-1258 (2010).
97. D. M. Anstrom *et al.*, Structure of the Escherichia coli malate synthase G:pyruvate:acetyl-coenzyme A abortive ternary complex at 1.95 A resolution. *Protein Sci.* **12**, 1822-1832 (2003).
98. K. Hoelsch *et al.*, Engineering of formate dehydrogenase: synergistic effect of mutations affecting cofactor specificity and chemical stability. *Appl. Microbiol. Biotechnol.* **97**, 2473-2481 (2013).
99. B. Nocek *et al.*, Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria. *Proc. Natl. Acad. Sci. U.S.A* **105**, 17730-17735 (2008).
100. M. S. Sevinc *et al.*, The cysteines of catalase HPII of Escherichia coli, including Cys438 which is blocked, do not have a catalytic role. *Eur. J. Biochem.* **230**, 127-132 (1995).
101. A. Kaplun *et al.*, Glyoxylate carboligase lacks the canonical active site glutamate of thiamine-dependent enzymes. *Nat. Chem. Biol.* **4**, 113-118 (2008).
102. A. M. Gotto and H. L. Kornberg, Crystalline tartronic semialdehyde reductase from Pseudomonas ovalis Chester. *Biochim. Biophys. Acta* **48**, 604-605 (1961).
103. B. W. Carlson and L. L. Miller, Oxidation of Nadh by Ferrocenium Salts - Rate-Limiting One-Electron Transfer. *J. Am. Chem. Soc.* **105**, 7453-7454 (1983).
104. S. Djordjevic *et al.*, 3-Dimensional Structure of Butyryl-Coa Dehydrogenase from Megasphaera-Elsdenii. *Biochemistry* **34**, 2163-2171 (1995).
105. C. S. Huang *et al.*, Crystal structure of the alpha(6)beta(6) holoenzyme of propionyl-coenzyme A carboxylase. *Nature* **466**, 1001-1005 (2010).
106. J. Mackenzie *et al.*, Controlling electron transfer in Acyl-CoA oxidases and dehydrogenases: a structural view. *J. Biol. Chem.* **281**, 31012-31020 (2006).
107. N. L. Gale and J. V. Beck, Evidence for the Calvin cycle and hexose monophosphate pathway in Thiobacillus ferrooxidans. *J. Bacteriol.* **94**, 1052-1059 (1967).
108. V. de Lorenzo *et al.*, Chemical reactivity drives spatiotemporal organisation of bacterial metabolism. *FEMS Microbiol. Rev.* **39**, 96-119 (2015).
109. C. Lerma-Ortiz *et al.*, 'Nothing of chemistry disappears in biology': the Top 30 damage-prone endogenous metabolites. *Biochem. Soc. Trans.* **44**, 961-971 (2016).
110. C. L. Linster *et al.*, Ethylmalonyl-CoA decarboxylase, a new enzyme involved in metabolite proofreading. *J. Biol. Chem.* **286**, 42992-43003 (2011).
111. I. Bernhardsgrütter *et al.*, *(to be published)*, (2017).
112. T. Fukui *et al.*, Expression and characterization of (R)-specific enoyl coenzyme A hydratase involved in polyhydroxyalkanoate biosynthesis by Aeromonas caviae. *J. Bacteriol.* **180**, 667-673 (1998).
113. R. Teufel *et al.*, 3-hydroxypropionyl-coenzyme A dehydratase and acryloyl-coenzyme A reductase, enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in the Sulfolobales. *J. Bacteriol.* **191**, 4572-4581 (2009).
114. S. Friedmann *et al.*, Properties of succinyl-coenzyme A:L-malate coenzyme A transferase and its role in the autotrophic 3-hydroxypropionate cycle of Chloroflexus aurantiacus. *J. Bacteriol.* **188**, 2646-2655 (2006).
115. B. M. van Vugt-Lussenburg *et al.*, Biochemical similarities and differences between the catalytic [4Fe-4S] cluster containing fumarases FumA and FumB from Escherichia coli. *PLoS One* **8**, e55549 (2013).

116. J. K. Rogers and G. M. Church, Genetically encoded sensors enable real-time observation of metabolite production. *Proc. Natl. Acad. Sci. U.S.A* **113**, 2388-2393 (2016).
117. J. Ruprecht *et al.*, Structure of Escherichia coli succinate:quinone oxidoreductase with an occupied and empty quinone-binding site. *J. Biol. Chem.* **284**, 29836-29846 (2009).
118. S. Wischgoll *et al.*, Structural basis for promoting and preventing decarboxylation in glutaryl-coenzyme a dehydrogenases. *Biochemistry* **49**, 5350-5357 (2010).
119. K. P. Battaile *et al.*, Structures of isobutyryl-CoA dehydrogenase and enzyme-product complex: comparison with isovaleryl- and short-chain acyl-CoA dehydrogenases. *J. Biol. Chem.* **279**, 16526-16534 (2004).
120. H. S. Toogood *et al.*, Dynamics driving function: new insights from electron transferring flavoproteins and partner complexes. *FEBS J.* **274**, 5481-5504 (2007).
121. J. V. Rodrigues and C. M. Gomes, Mechanism of superoxide and hydrogen peroxide generation by human electron-transfer flavoprotein and pathological variants. *Free Radic Biol Med* **53**, 12-19 (2012).
122. L. Tong, Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cell. Mol. Life Sci.* **62**, 1784-1803 (2005).
123. M. Hugler *et al.*, Characterization of acetyl-CoA/propionyl-CoA carboxylase in Metallosphaera sedula. Carboxylating enzyme in the 3-hydroxypropionate cycle for autotrophic carbon fixation. *Eur. J. Biochem.* **270**, 736-744 (2003).
124. L. Diacovich *et al.*, Kinetic and structural analysis of a new group of Acyl-CoA carboxylases found in Streptomyces coelicolor A3(2). *J. Biol. Chem.* **277**, 31228-31236 (2002).
125. E. Rodriguez and H. Gramajo, Genetic and biochemical characterization of the alpha and beta components of a propionyl-CoA carboxylase complex of Streptomyces coelicolor A3(2). *Microbiology* **145 ( Pt 11)**, 3109-3119 (1999).
126. L. Diacovich *et al.*, Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity. *Biochemistry* **43**, 14027-14036 (2004).
127. Y. Ikeda *et al.*, Separation and properties of five distinct acyl-CoA dehydrogenases from rat liver mitochondria. Identification of a new 2-methyl branched chain acyl-CoA dehydrogenase. *J. Biol. Chem.* **258**, 1066-1076 (1983).
128. K. Izai *et al.*, Novel fatty acid beta-oxidation enzymes in rat liver mitochondria. I. Purification and properties of very-long-chain acyl-coenzyme A dehydrogenase. *J. Biol. Chem.* **267**, 1027-1033 (1992).
129. T. V. Nguyen *et al.*, Identification of isobutyryl-CoA dehydrogenase and its deficiency in humans. *Mol. Genet. Metab.* **77**, 68-79 (2002).
130. A. Besrat *et al.*, Mammalian metabolism of glutaric acid. *J. Biol. Chem.* **244**, 1461-1467 (1969).
131. B. Pohl *et al.*, Studies on the reaction mechanism of general acyl-CoA dehydrogenase. Determination of selective isotope effects in the dehydrogenation of butyryl-CoA. *Eur. J. Biochem.* **160**, 109-115 (1986).
132. S. Ghisla *et al.*, Mechanistic studies with general acyl-CoA dehydrogenase and butyryl-CoA dehydrogenase: evidence for the transfer of the beta-hydrogen to the flavin N(5)-position as a hydride. *Biochemistry* **23**, 3154-3161 (1984).
133. F. L. Crane and H. Beinert, On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. II. The electron-transferring flavoprotein. *J. Biol. Chem.* **218**, 717-731 (1956).
134. Y. Ikeda *et al.*, Purification and characterization of short-chain, medium-chain, and long-chain acyl-CoA dehydrogenases from rat liver mitochondria. Isolation of the holo- and apoenzymes and conversion of the apoenzyme to the holoenzyme. *J. Biol. Chem.* **260**, 1311-1325 (1985).
135. R. P. McAndrew *et al.*, Structural basis for substrate fatty acyl chain specificity: crystal structure of human very-long-chain acyl-CoA dehydrogenase. *J. Biol. Chem.* **283**, 9435-9443 (2008).
136. Y. Nakajima *et al.*, Three-dimensional structure of the flavoenzyme acyl-CoA oxidase-II from rat liver, the peroxisomal counterpart of mitochondrial acyl-CoA dehydrogenase. *J Biochem* **131**, 365-374 (2002).

137. K. P. Battaile *et al.*, Crystal structure of rat short chain acyl-CoA dehydrogenase complexed with acetoacetyl-CoA: comparison with other acyl-CoA dehydrogenases. *J. Biol. Chem.* **277**, 12200-12207 (2002).
138. J. J. Kim *et al.*, Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U.S.A* **90**, 7523-7527 (1993).
139. F. L. Crane *et al.*, On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. I. The general fatty acyl coenzyme A dehydrogenase. *J. Biol. Chem.* **218**, 701-706 (1956).
140. K. A. Tiffany *et al.*, Structure of human isovaleryl-CoA dehydrogenase at 2.6 A resolution: structural basis for substrate specificity. *Biochemistry* **36**, 8455-8464 (1997).
141. Z. Fu *et al.*, Crystal structures of human glutaryl-CoA dehydrogenase with and without an alternate substrate: structural bases of dehydrogenation and decarboxylation reactions. *Biochemistry* **43**, 9674-9684 (2004).
142. S. Bhattacharyya *et al.*, Combined quantum mechanical and molecular mechanical simulations of one- and two-electron reduction potentials of flavin cofactor in water, medium-chain acyl-CoA dehydrogenase, and cholesterol oxidase. *J Phys Chem A* **111**, 5729-5742 (2007).
143. J. J. Kim and R. Miura, Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences. *Eur. J. Biochem.* **271**, 483-493 (2004).
144. B. Kuchler *et al.*, Biochemical characterization of a variant human medium-chain acyl-CoA dehydrogenase with a disease-associated mutation localized in the active site. *Biochem. J.* **337 ( Pt 2)**, 225-230 (1999).
145. H. L. Kornberg and J. G. Morris, Beta-Hydroxyaspartate pathway: a new route for biosyntheses from glyoxylate. *Nature* **197**, 456-457 (1963).
146. H. L. Kornberg and J. G. Morris, The Utilization of Glycollate by Micrococcus Denitrificans: The Beta-Hydroxyaspartate Pathway. *Biochem. J.* **95**, 577-586 (1965).
147. L. Schada von Borzyskowski *et al.*, An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. *(to be published),* (2017).
148. K. Schneider *et al.*, The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by Methylobacterium extorquens AM1 during growth on acetate. *J. Biol. Chem.* **287**, 757-766 (2012).
149. C. Dellomonaco *et al.*, Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**, 355-359 (2011).
150. C. J. Paddon *et al.*, High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-532 (2013).
151. M. Enquist-Newman *et al.*, Efficient ethanol production from brown macroalgae sugars by a synthetic yeast platform. *Nature* **505**, 239-243 (2014).
152. S. Galanie *et al.*, Complete biosynthesis of opioids in yeast. *Science* **349**, 1095-1100 (2015).
153. E. M. Nichols *et al.*, Hybrid bioinorganic approach to solar-to-chemical conversion. *Proc. Natl. Acad. Sci. U.S.A* **112**, 11461-11466 (2015).
154. C. Liu *et al.*, Water splitting-biosynthetic system with CO*2* reduction efficiencies exceeding photosynthesis. *Science* **352**, 1210-1213 (2016).
155. D. Seigneurin-Berny *et al.*, Purification of intact chloroplasts from Arabidopsis and spinach leaves by isopycnic centrifugation. *Curr. Protoc. Cell Biol.* **Chapter 3**, Unit 3 30 (2008).
156. I. Yacoby *et al.*, Photosynthetic electron partitioning between [FeFe]-hydrogenase and ferredoxin:NADP+-oxidoreductase (FNR) enzymes in vitro. *Proc. Natl. Acad. Sci. U.S.A* **108**, 9396-9401 (2011).
157. N. J. Watmough and F. E. Frerman, The electron transfer flavoprotein: ubiquinone oxidoreductases. *Biochim. Biophys. Acta* **1797**, 1910-1916 (2010).
158. S. Liu and A. Steinbuchel, Investigation of poly(beta-L-melic acid) production by strains of Aureobasidium pullulans. *Appl. Microbiol. Biotechnol.* **46**, 273-278 (1996).
159. E. Holler *et al.*, Biological and Biosynthetic-Properties of Poly-L-Malate. *FEMS Microbiol. Lett.* **103**, 109-118 (1992).

160. B. Willibald *et al.*, Is beta-poly(L-malate) synthesis catalysed by a combination of beta-L-malyl-AMP-ligase and beta-poly(L-malate) polymerase? *Eur. J. Biochem.* **265**, 1085-1090 (1999).
161. Y. K. Wang *et al.*, Enhanced production of Ca(2)(+)-polymalate (PMA) with high molecular mass by Aureobasidium pullulans var. pullulans MCW. *Microb. Cell. Fact.* **14**, 115 (2015).
162. J. K. Guterl and V. Sieber, Biosynthesis "debugged": Novel bioproduction strategies. *Eng. Life Sci.* **13**, 4-18 (2013).
163. C. Hold *et al.*, Forward design of a complex enzyme cascade reaction. *Nature Communications* **7**, (2016).
164. M. Bujara *et al.*, Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. *Nat. Chem. Biol.* **7**, 271-277 (2011).
165. A Portrait of the Max Planck Society. *https://www.mpg.de/short-portrait*, (2017).
166. F. Jacob, Evolution and tinkering. *Science* **196**, 1161-1166 (1977).
167. D. R. Ort *et al.*, Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci. U.S.A* **112**, 8529-8536 (2015).
168. R. K. Yadav *et al.*, A Photocatalyst-Enzyme Coupled Artificial Photosynthesis System for Solar Energy in Production of Formic Acid from CO2. *J. Am. Chem. Soc.* **134**, 11455-11461 (2012).
169. C. A. Hutchison *et al.*, Design and synthesis of a minimal bacterial genome. *Science* **351**, 1414-U1473 (2016).
170. D. M. Peter *et al.*, A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters. *Molecules* **21**, 517 (2016).
171. R. G. Rosenthal *et al.*, Direct evidence for a covalent ene adduct intermediate in NAD(P)H-dependent enzymes. *Nat. Chem. Biol.* **10**, 50-55 (2014).
172. A. de Marco, Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli. *Nat. Protoc.* **2**, 2632-2639 (2007).
173. J. K. Kroeger *et al.*, A spectrophotometric assay for measuring acetyl-coenzyme A carboxylase. *Anal. Biochem.* **411**, 100-105 (2011).
174. M. Kitagawa *et al.*, 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 (2005).
175. J. Sambrook and D. W. Russel, *Molecular Cloning, a Laboratory Manual*. (Cold Spring Harbor Laboratory Press, ed. 3, 2001).
176. W. Kabsch, Xds. *Acta. Crystallogr. D Biol. Crystallogr.* **66**, 125-132 (2010).
177. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta. Crystallogr. D Biol. Crystallogr.* **67**, 235-242 (2011).
178. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D Biol. Crystallogr.* **66**, 213-221 (2010).
179. P. Emsley *et al.*, Features and development of Coot. *Acta. Crystallogr. D Biol. Crystallogr.* **66**, 486-501 (2010).
180. T. C. Lehman *et al.*, An acyl-coenzyme A dehydrogenase assay utilizing the ferricenium ion. *Anal. Biochem.* **186**, 280-284 (1990).
181. P. Kiefer *et al.*, eMZed: an open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. *Bioinformatics* **29**, 963-964 (2013).
182. S. U. Bajad *et al.*, Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J. Chromatogr. A* **1125**, 76-88 (2006).

# 8. Supplementary information

## 8.1. Gibbs free energies use to calculate Gibbs free energy profiles

Gibbs free energy values of all reactions for natural and the designed synthetic $CO_2$ fixation pathways were calculated or taken from literature with their respective standard deviation if applicable (n.a = not applicable). $\Delta_rG'^m$: Gibbs free energy under standard conditions; $\Delta_rG'$ custom: Gibbs free energy under defined conditions (see **Fig. 8**).
[a]: calculated with eQuilibrator 2.0 (http://equilibrator.weizmann.ac.il) (*88*); [b]: taken from Fuchs et al. 2011 (*6*); [c]: estimated based on values given by eQuilibrator 2.0 for various short- and medium-chain CoA-ligases; [d]: estimated based on the value used for succinyl-CoA + NADPH ⇌ succinic semialdehyde + $NADP^+$; [e]: estimated based on the value used for succinic semialdehyde + NADPH ⇌ 4-hydroxybutyrate + $NADP^+$; [f]: estimated based on values given by eQuilibrator 2.0 for glutaryl-CoA + ubiquinone ⇌ glutaconyl-1-CoA + ubiquinol and succinate + ubiquinone ⇌ fumarate + ubiquinol; [g]: estimated based on values given by eQuilibrator 2.0 for acrylyl-CoA + $H_2O$ ⇌ 3-hydroxypropionyl-CoA, fumarate + $H_2O$ ⇌ malyl-CoA and mesaconyl-CoA + $H_2O$ ⇌ methylmalyl-CoA/citramalyl-CoA; [h]: estimated based on values given by eQuilibrator 2.0 for malyl-CoA ⇌ acetyl-CoA + glyoxylate, methylmalyl-CoA ⇌ propionyl-CoA + glyoxylate and citramalyl-CoA ⇌ acetyl-CoA + pyruvate; [i]: estimated based on the value used for 4-hydroxybutyrate + CoA + ATP ⇌ 4-hydroxybutyryl-CoA + ADP + Pi; [j]: estimated based on the value used for 4-hydroxybutyryl-CoA ⇌ crotonyl-CoA + $H_2O$; [k]: estimated based on the value given by eQuilibrator 2.0 for vinylacetyl-CoA ⇌ crotonyl-CoA; [l]: estimated based on the value used for crotonyl-CoA + NADPH + $CO_2$ ⇌ ethylmalonyl-CoA + $NADP^+$; [m]: estimated based on the value given by eQuilibrator 2.0 for 3-hydroxybutyryl-CoA ⇌ crotonyl-CoA + $H_2O$; [n]: estimated based on the value used for methylsuccinyl-CoA + ubiquinone ⇌ mesaconyl-CoA + ubiquinol; [o]: estimated based on the value given by eQuilibrator 2.0 for ethylmalonyl-CoA ⇌ methylsuccinyl-CoA; [p]: estimated based on the value given by eQuilibrator 2.0 for butyryl-CoA + ubiquinone ⇌ crotonyl-CoA + ubiquinol

### A) CETCH cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | crotonyl-CoA + NADPH + $CO_2$ | ethylmalonyl-CoA + $NADP^+$ | -26.1 ±17 | -30.4 ±17 | [a] |
| 2 | ethylmalonyl-CoA | methylsuccinyl-CoA | 0 ±0 | 0 ±0 | [a] |
| 3 | methylsuccinyl-CoA + UQ | mesaconyl-CoA + $UQH_2$ | -25.5 ±6.7 | -25.5 ±6.7 | [a] |
| 4 | mesaconyl-CoA + $H_2O$ | methylmalyl-CoA | 3.3 ±3.7 | 3.3 ±3.7 | [a] |
| 5 | methylmalyl-CoA | propionyl-CoA + glyoxylate | -1.6 ±16.4 | -5.6 ±16.4 | [a] |
| 6 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA | -8.2 ±10.3 | -20.7 ±10.3 | [a] |
| 7 | methylmalonyl-CoA | succinyl-CoA | 0 ±0 | 0 ±0 | [a] |
| 8 | succinyl-CoA + NADPH | succinic semialdehyde + CoA + $NADP^+$ | -3 ±5.4 | -10.2 ±5.4 | [a] |
| 9 | succinic semialdehyde + NADPH | 4-hydroxybutyrate + $NADP^+$ | -24 n.a. | -24 n.a. | [b] |
| 10 | 4-hydroxybutyrate + CoA + ATP | 4-hydroxybutyryl-CoA + ADP + $P_i$ | 10 ±16.1 | 10.4 ±16.1 | [c] |
| 11 | 4-hydroxybutyryl-CoA | crotonyl-CoA + $H_2O$ | -7.7 ±15.6 | -7.7 ±15.6 | [a] |

## B) HOPAC cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | acrylyl-CoA + NADPH + $CO_2$ | methylmalonyl-CoA + $NADP^+$ | -26 ±17.6 | -30.4 ±17.6 | [a] |
| 2 | methylmalonyl-CoA | succinyl-CoA | -7.6 ±4.1 | -7.6 ±4.1 | [a] |
| 3 | succinyl-CoA + ADP + $P_i$ | succinate + ATP + CoA | 1.1 ±2.6 | 0.6 ±2.6 | [a] |
| 4 | succinate + UQ | fumarate + $UQH_2$ | -21.6 ±7.1 | -21.6 ±7.1 | [a] |
| 5 | fumarate + $H_2O$ | malate | -3.4 ±0.6 | -3.4 ±0.6 | [a] |
| 6 | malate + ATP + CoA | malyl-CoA + ADP + $P_i$ | -5.8 ±7.1 | -5.4 ±7.1 | [a] |
| 7 | malyl-CoA | glyoxylate + acetyl-CoA | -3 ±5.8 | -6.9 ±5.8 | [a] |
| 8 | acetyl-CoA + ATP + $HCO_3^-$ | malonyl-CoA + ADP + $P_i$ | -9.9 ±6.5 | -20.7 ±6.5 | [a] |
| 9 | malonyl-CoA + NADPH | malonic semialdehyde + CoA + $NADP^+$ | -3 ±5.4 | -10.2 ±5.4 | [d] |
| 10 | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | 4.6 ±3.4 | -0.9 ±3.4 | [e] |
| 11 | 3-hydroxypropionate + ATP + CoA | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 ±16.1 | 11.6 ±16.1 | [a] |
| 12 | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 ±3.8 | 0 ±3.8 | [a] |

## C) CHYME cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | crotonyl-CoA + NADPH + $CO_2$ | ethylmalonyl-CoA + $NAPH^+$ | -26.1 ±17 | -30.4 ±17 | [a] |
| 2 | ethylmalonyl-CoA + UQ | ethylidenemalonyl-CoA + $UQH_2$ | -21.6 ±7.1 | -21.6 ±6.3 | [f] |
| 3 | ethylidenemalonyl-CoA + $H_2O$ | hydroxyethylmalonyl-CoA | -0.8 n.a. | -0.8 n.a. | [g] |
| 4 | hydroxyethylmalonyl-CoA | malonyl-CoA + acetaldehyde | -3.8 n.a. | -7.77 n.a. | [h] |
| 5 | malonyl-CoA + NADPH | malonic semialdehyde + $NADP^+$ + CoA | -4.6 ±8.8 | -11.8 ±8.8 | [a] |
| 6 | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | -19.1 ±5.8 | -24.5 ±5.8 | [a] |
| 7 | 3-hydroxypropionate + CoA + ATP | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 ±16.1 | 11.6 ±16.1 | [a] |
| 8 | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 ±3.8 | 0 ±3.8 | [a] |
| 9 | acrylyl-CoA + NADPH | propionyl-CoA + $NADP^+$ | -56.6 ±16.9 | -62 ±16.9 | [a] |
| 10 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA + ADP + $P_i$ | -8.2 ±10.3 | -20.7 ±10.3 | [a] |
| 11 | methylmalonyl-CoA | succinyl-CoA | 0 ±0 | 0 ±0 | [a] |
| 12 | succinyl-CoA + NADPH | succinic semialdehyde + CoA + $NADP^+$ | -3 ±5.4 | -10.2 ±5.4 | [a] |
| 13 | succinic semialdehyde + NADPH | 4-hydroxybutyrate + NADP+ | -24 n.a. | -24 n.a. | [b] |
| 14 | 4-hydroxybutyrate + CoA + ATP | 4-hydroxybutyryl-CoA + ADP + $P_i$ | 10 ±16.1 | 10.4 ±16.1 | [c] |
| 15 | 4-hydroxybutyryl-CoA | crotonyl-CoA + $H_2O$ | -7.7 ±15.6 | -7.7 ±15.6 | [a] |
| 16 | acetaldehyde + CoA + $NADP^+$ | acetyl-CoA + NADPH | -4.4 ±2.4 | 2.7 ±2.4 | [a] |

## D) CRACE cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | crotonyl-CoA + NADPH + $CO_2$ | ethylmalonyl-CoA + $NADP^+$ | -26.1 ±17 | -30.4 ±17 | [a] |
| 2 | ethylmalonyl-CoA | methylsuccinyl-CoA | 0 ±0 | 0 ±0 | [a] |
| 3 | methylsuccinyl-CoA + NADPH | methylsuccinic semialdehyde + $NADP^+$ | -3 ±5.4 | -10.2 ±5.4 | [d] |
| 4 | methylsuccinic semialdehyde + NADPH | 4-hydroxy-3-methylbutyrate + $NADP^+$ | -24 n.a. | -24 n.a. | [e] |
| 5 | 4-hydroxy-3-methylbutyrate + CoA + ATP | 4-hydroxy-3-methylbutyryl-CoA + ADP + $P_i$ | 10 ±16.1 | 10.4 ±16.1 | [i] |
| 6 | 4-hydroxy-3-methylbutyryl-CoA | 3-methyl-but-3-enoyl-CoA + $H_2O$ | -7.7 ±15.6 | -7.7 ±15.6 | [j] |
| 7 | 3-methyl-but-3-enoyl-CoA | 3-methylcrotonyl-CoA | -7.7 ±16.5 | -7.7 ±16.5 | [k] |
| 8 | 3-methylcrotonyl-CoA + $HCO_3^-$ + ATP | 3-methylglutaconyl-CoA + ADP + $P_i$ | -15.4 ±6.3 | -26.3 ±6.3 | [a] |
| 9 | 3-methylglutaconyl-CoA + $H_2O$ | hydroxymethylglutaryl-CoA | -3.1 ±3.7 | -3.1 ±3.7 | [a] |
| 10 | hydroxymethylglutaryl-CoA | acetyl-CoA + acetoacetate | -34.7 ±15.9 | -38.7 ±15.9 | [a] |
| 11 | acetoacetate + CoA + ATP | acetoacetyl-CoA + ADP + $P_i$ | 11 ±2.8 | 11.5 ±2.8 | [a] |
| 12 | acetoacetyl-CoA + NADPH | 3-hydroxybutyryl-CoA + $NADP^+$ | -13.2 ±2.5 | -18.7 ±2.5 | [a] |
| 13 | 3-hydroxybutyryl-CoA | crotonyl-CoA + $H_2O$ | 3.8 ±3.3 | 3.8 ±3.3 | [a] |

## E) FUMES cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | | Source |
|---|---|---|---|---|---|---|---|
| 1 | fumaryl-CoA + NADPH + $CO_2$ | carboxysuccinyl-CoA + $NADP^+$ | -26.1 | ±17 | -30.4 | ±17 | [l] |
| 2 | carboxysuccinyl-CoA + NADPH | formylsuccinate + CoA + $NADP^+$ | -3 | ±5.4 | -10.2 | ±5.4 | [g] |
| 3 | formylsuccinate + NADPH | hydroxymethylsuccinate + $NADP^+$ | -24 | n.a. | -24 | n.a. | [h] |
| 4 | hydroxymethylsuccinate + CoA + ATP | hydroxymethylsuccinyl-CoA + ADP + $P_i$ | 10 | ±16.1 | 10.4 | ±16.1 | [i] |
| 5 | hydroxymethylsuccinyl-CoA | itaconyl-CoA + $H_2O$ | 3.8 | ±3.3 | 3.8 | ±3.3 | [m] |
| 6 | itaconyl-CoA | mesaconyl-CoA | -7.7 | ±16.5 | -7.7 | ±16.5 | [k] |
| 7 | mesaconyl-CoA + $H_2O$ | methylmalyl-CoA | 3.3 | ±3.7 | 3.3 | ±3.7 | [a] |
| 8 | methylmalyl-CoA | propionyl-CoA + glyoxylate | -1.6 | ±16.4 | -5.6 | ±16.4 | [a] |
| 9 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA + ADP + $P_i$ | -8.2 | ±10.3 | -20.7 | ±10.3 | [a] |
| 10 | methylmalonyl-CoA | succinyl-CoA | 0 | ±0 | 0 | ±0 | [a] |
| 11 | succinyl-CoA + UQ | fumaryl-CoA + $UQH_2$ | -25.5 | ±6.7 | -25.5 | ±6.7 | [n] |

## F) HITME cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | | Source |
|---|---|---|---|---|---|---|---|
| 1 | hydroxycrotonyl-CoA + NADPH + $CO_2$ | hydroxyethylmalonyl-CoA + $NADP^+$ | -26.1 | ±17 | -30.4 | 17 | [l] |
| 2 | hydroxyethylmalonyl-CoA | hydroxymethylsuccinyl-CoA | 0 | ±0 | 0 | 0 | [o] |
| 3 | hydroxymethylsuccinyl-CoA | itaconyl-CoA + $H_2O$ | 3.8 | ±3.3 | 3.8 | 3.3 | [m] |
| 4 | itaconyl-CoA | mesaconyl-CoA | -7.7 | ±16.5 | -7.7 | 16.5 | [k] |
| 5 | mesaconyl-CoA + $H_2O$ | methylmalyl-CoA | 3.3 | ±3.7 | 3.3 | 3.7 | [a] |
| 6 | methylmalyl-CoA | propionyl-CoA + glyoxylate | -1.6 | ±16.4 | -5.6 | 16.4 | [a] |
| 7 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA | -8.2 | ±10.3 | -20.7 | 10.3 | [a] |
| 8 | methylmalonyl-CoA | succinyl-CoA | 0 | ±0 | 0 | 0 | [a] |
| 9 | succinyl-CoA + NADPH | succinic semialdehyde + CoA + $NADP^+$ | -3 | ±5.4 | -10.2 | 5.4 | [a] |
| 10 | succinic semialdehyde + NADPH | 4-hydroxybutyrate + $NADP^+$ | -24 | n.a. | -24 | n.a. | [b] |
| 11 | 4-hydroxybutyrate + CoA + ATP | 4-hydroxybutyryl-CoA + ADP + $P_i$ | 10 | ±16.1 | 10.4 | 16.1 | [c] |
| 12 | 4-hydroxybutyryl-CoA + UQ | hydroxycrotonyl-CoA + $UQH_2$ | -27.8 | ±17 | -27.8 | 17 | [p] |

## G) HYPA cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | | Source |
|---|---|---|---|---|---|---|---|
| 1 | hydroxycrotonyl-CoA + NADPH + $CO_2$ | hydroxyethylmalonyl-CoA + $NADP^+$ | -26.1 | ±17 | -30.4 | ±17 | [l] |
| 2 | hydroxyethylmalonyl-CoA | hydroxymethylsuccinyl-CoA | 0 | ±0 | 0 | ±0 | [o] |
| 3 | hydroxymethylsuccinyl-CoA | itaconyl-(C1)-CoA + H2O | 3.8 | ±3.3 | 3.8 | ±3.3 | [m] |
| 4 | itaconyl-(C1)-CoA | itaconyl-(C4)-CoA | 0 | ±0 | 0 | ±0 | [o] |
| 5 | itaconyl-(C4)-CoA + $H_2O$ | citramalyl-CoA | -4.5 | ±3.1 | -4.5 | ±3.1 | [a] |
| 6 | citramalyl-CoA | acetyl-CoA + pyruvate | -6.8 | ±15.3 | -10.8 | ±15.3 | [a] |
| 7 | acetyl-CoA + ATP + $HCO_3^-$ | malonyl-CoA + ADP + $P_i$ | -9.9 | ±6.5 | -20.7 | ±6.5 | [a] |
| 8 | malonyl-CoA + NADPH | malonic semialdehyde + $NADP^+$ + CoA | -3 | ±5.4 | -10.2 | ±5.4 | [a] |
| 9 | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | 4.6 | ±3.4 | -0.9 | ±3.4 | [a] |
| 10 | 3-hydroxypropionate + CoA + ATP | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 | ±16.1 | 11.6 | ±16.1 | [a] |
| 11 | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 | ±3.8 | 0 | ±3.8 | [a] |
| 12 | acrylyl-CoA + NADPH | propionyl-CoA + $NADP^+$ | -56.6 | ±16.9 | -62 | ±16.9 | [a] |
| 13 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA | -8.2 | ±10.3 | -20.7 | ±10.3 | [a] |
| 14 | methylmalonyl-CoA | succinyl-CoA | 0 | ±0 | 0 | ±0 | [a] |
| 15 | succinyl-CoA + NADPH | succinic semialdehyde + CoA + $NADP^+$ | -3 | ±5.4 | -10.2 | ±5.4 | [a] |
| 16 | succinic semialdehyde + NADPH | 4-hydroxybutyrate + $NADP^+$ | -24 | n.a. | -24 | n.a. | [b] |
| 17 | 4-hydroxybutyrate + CoA + ATP | 4-hydroxybutyryl-CoA + ADP + $P_i$ | 10 | ±16.1 | 10.4 | ±16.1 | [c] |
| 18 | 4-hydroxybutyryl-CoA + UQ | hydroxycrotonyl-CoA + $UQH_2$ | -27.8 | ±17 | -27.8 | ±17 | [p] |

## H) CBB cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | ribulose-1,5-bisphosphate + $CO_2$ + $H_2O$ | 2 3-phosphoglycerate | -37.8 ±7.5 | -37.3 ±7.5 | [a] |
| 2 | 3-phosphoglycerate + ATP | 1,3-bisphosphoglycerate + ADP | 18.7 ±0.9 | 17 ±0.9 | [a] |
| 3 | bisphosphoglycerate + NADPH | glyceraldehyde-3-phosphate + $NADP^+$ | -18.3 ±1 | -23.2 ±1 | [a] |
| 4 | glyceraldehyde-3-phosphate | dihydroxyacetone-phosphate | -5.5 ±1.1 | -5.5 ±1.1 | [a] |
| 5 | dihydroxyacetone-phosphate + glyceraldehyde- | fructose-1,6-bisphosphate | -2.4 ±4.4 | 1.6 ±4.4 | [a] |
| 6 | fructose-1,6-bisphosphate + $H_2O$ | fructose-6-phosphate + $P_i$ | -28.5 ±4.3 | -28 ±4.3 | [a] |
| 7 | fructose-6-phosphate + glyceraldehyde-3- | xylulose-5-phosphate + erythrose-4- | 10.2 ±3.8 | 10.2 ±3.8 | [a] |
| 8 | erythrose-4-phosphate + dihydroxyacetone- | seduheptulose-1,7-bisphosphate | 3.3 ±5.6 | 7.3 ±5.6 | [a] |
| 9 | seduheptulose-1,7-bisphosphate + $H_2O$ | seduheptulose-7-phosphate + $P_i$ | -25.1 ±2.4 | -24.6 ±2.4 | [a] |
| 10 | seduheptulose-7-phosphate +glyceraldehyde- | xylulose-5-phosphate + ribose-5-phosphate | -4.5 ±4.5 | -4.5 ±4.5 | [a] |
| 11 | ribose-5-phosphate | ribulose-5-phosphate | 2 ±1.5 | 2 ±1.5 | [a] |
| 12 | xylulose-5-phosphate | ribulose-5-phosphate | 3.4 ±2.3 | 3.4 ±2.3 | [a] |
| 13 | ribulose-5-phosphate + ATP | ribulose-1,5-bisphosphate + ADP | -24.2 ±4.5 | -25.9 ±4.5 | [a] |

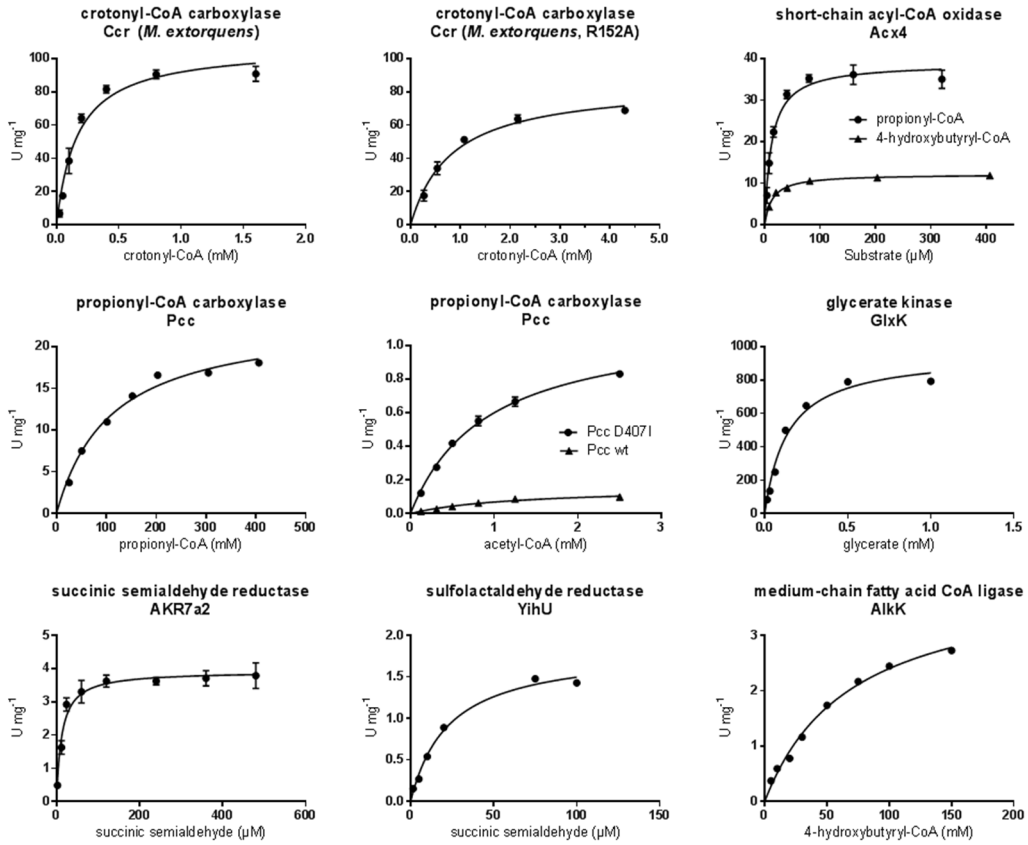
## I) 3HP bi-cycle

| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | Source |
|---|---|---|---|---|---|
| 1 | acetyl-CoA + ATP + $HCO_3^-$ | propionyl-CoA + ADP + $P_i$ | -9.9 6.5 | -20.7 6.5 | [a] |
| 2a | malonyl-CoA + NADPH | malonic semialdehyde + $NADP^+$ + CoA | -3 5.4 | -10.2 5.4 | [a] |
| 2b | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | 4.6 3.4 | -0.9 3.4 | [a] |
| 3a | 3-hydroxypropionate + CoA + ATP | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 16.1 | 11.6 16.1 | [a] |
| 3b | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 3.8 | 0 3.8 | [a] |
| 3c | acrylyl-CoA + NADPH | propionyl-CoA + $NADP^+$ | -56.6 16.9 | -62 16.9 | [a] |
| 4 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA + ADP + $P_i$ | -8.2 10.3 | -20.7 10.3 | [a] |
| 5+6 | methylmalonyl-CoA | succinyl-CoA | 0 0 | 0 0 | [a] |
| 7 | succinyl-CoA + malate | malyl-CoA + succinate | -4.8 7.4 | -4.8 7.4 | [a] |
| 8 | succinate + UQ | fumarate + $UQH_2$ | -21.6 7.1 | -21.6 7.1 | [a] |
| 9 | fumarate + $H_2O$ | malate | -3.4 0.6 | -3.4 0.6 | [a] |
| 7 | succinyl-CoA + malate | malyl-CoA + succinate | -4.8 7.4 | -4.8 7.4 | [a] |
| 10a | malyl-CoA | acetyl-CoA + glyoxylate | -3 5.8 | -6.9 5.8 | [a] |
| 1 | acetyl-CoA + ATP + $HCO_3^-$ | malonyl-CoA + ADP + $P_i$ | -9.9 6.5 | -20.7 6.5 | [a] |
| 2a | malonyl-CoA + NADPH | malonic semialdehyde + $NADP^+$ + CoA | -3 5.4 | -10.2 5.4 | [a] |
| 2b | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | 4.6 3.4 | -0.9 3.4 | [a] |
| 3a | 3-hydroxypropionate + CoA + ATP | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 16.1 | 11.6 16.1 | [a] |
| 3b | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 3.8 | 0 3.8 | [a] |
| 3c | acrylyl-CoA + NADPH | propionyl-CoA + $NADP^+$ | -56.6 16.9 | -62 16.9 | [a] |
| 10b | propionyl-CoA + glyoxylate | methylmalyl-CoA | 1.6 16.4 | 5.6 16.4 | [a] |
| 11 | methylmalyl-CoA | mesaconyl-CoA + $H_2O$ | -3.3 3.7 | -3.3 3.7 | [a] |
| 12 | mesaconyl-CoA (C1) | mesaconyl-CoA (C4) | 0 0 | 0 0 | [a] |
| 13 | mesaconyl-CoA + $H_2O$ | citramalyl-CoA | -3.1 3.7 | -3.1 3.7 | [a] |
| 10c | citramalyl-CoA | pyruvate + acetyl-CoA | -6.8 15.3 | -10.8 15.3 | [a] |

## J) 3HP/4HB cycle

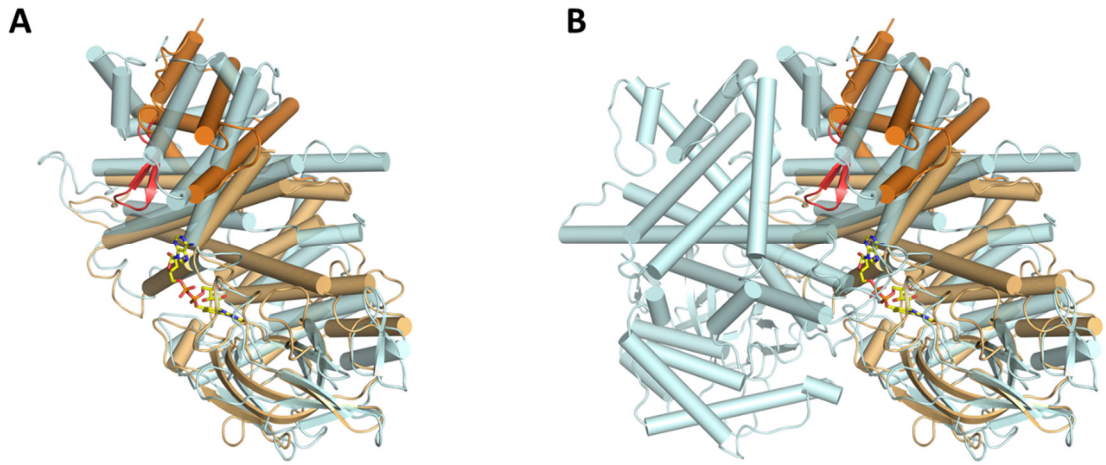
| Rxn | Substrates | Products | $\Delta_rG'^m$ (kJ mol$^{-1}$) | | $\Delta_rG'^{custom}$ (kJ mol$^{-1}$) | | Source |
|---|---|---|---|---|---|---|---|
| 1 | acetyl-CoA + ATP + $HCO_3^-$ | propionyl-CoA + ADP + $P_i$ | -9.9 | 6.5 | -20.7 | 6.5 | [a] |
| 2 | malonyl-CoA + NADPH | malonic semialdehyde + $NADP^+$ + CoA | -3 | 5.4 | -10.2 | 5.4 | [a] |
| 3 | malonic semialdehyde + NADPH | 3-hydroxypropionate + $NADP^+$ | 4.6 | 3.4 | -0.9 | 3.4 | [a] |
| 4 | 3-hydroxypropionate + CoA + ATP | 3-hydroxypropionyl-CoA + ADP + $P_i$ | 11.2 | 16.1 | 11.6 | 16.1 | [a] |
| 5 | 3-hydroxypropionyl-CoA | acrylyl-CoA + $H_2O$ | 0 | 3.8 | 0 | 3.8 | [a] |
| 6 | acrylyl-CoA + NADPH | propionyl-CoA + $NADP^+$ | -56.6 | 16.9 | -62 | 16.9 | [a] |
| 7 | propionyl-CoA + ATP + $HCO_3^-$ | methylmalonyl-CoA + ADP + $P_i$ | -8.2 | 10.3 | -20.7 | 10.3 | [a] |
| 8 | methylmalonyl-CoA | succinyl-CoA | 0 | 0 | 0 | 0 | [a] |
| 9 | succinyl-CoA + NADPH | succinic semialdehyde + CoA + $NADP^+$ | -3 | 5.4 | -10.2 | 5.4 | [a] |
| 10 | succinic semialdehyde + NADPH | 4-hydroxybutyrate + $NADP^+$ | -24 | n.a. | -24 | n.a. | [b] |
| 11 | 4-hydroxybutyrate + CoA + ATP | 4-hydroxybutyryl-CoA + ADP + $P_i$ | 10 | 16.1 | 10.4 | 16.1 | [c] |
| 12 | 4-hydroxybutyryl-CoA | crotonyl-CoA + $H_2O$ | -7.7 | 15.6 | -7.7 | 15.6 | [a] |
| 13 | crotonyl-CoA + $H_2O$ | 3-hydroxybutyryl-CoA | -3.8 | 3.3 | -3.8 | 3.3 | [a] |
| 14 | 3-hydroxybutyryl-CoA + $NADP^+$ | acetoacetyl-CoA + NADPH | 13.2 | 2.5 | 18.7 | 2.5 | [a] |
| 15 | acetoacetyl-CoA + CoA | 2 acetyl-CoA | -25.9 | 1.7 | -28.2 | 1.7 | [a] |

## 8.2. Michaelis-Menten Kinetics

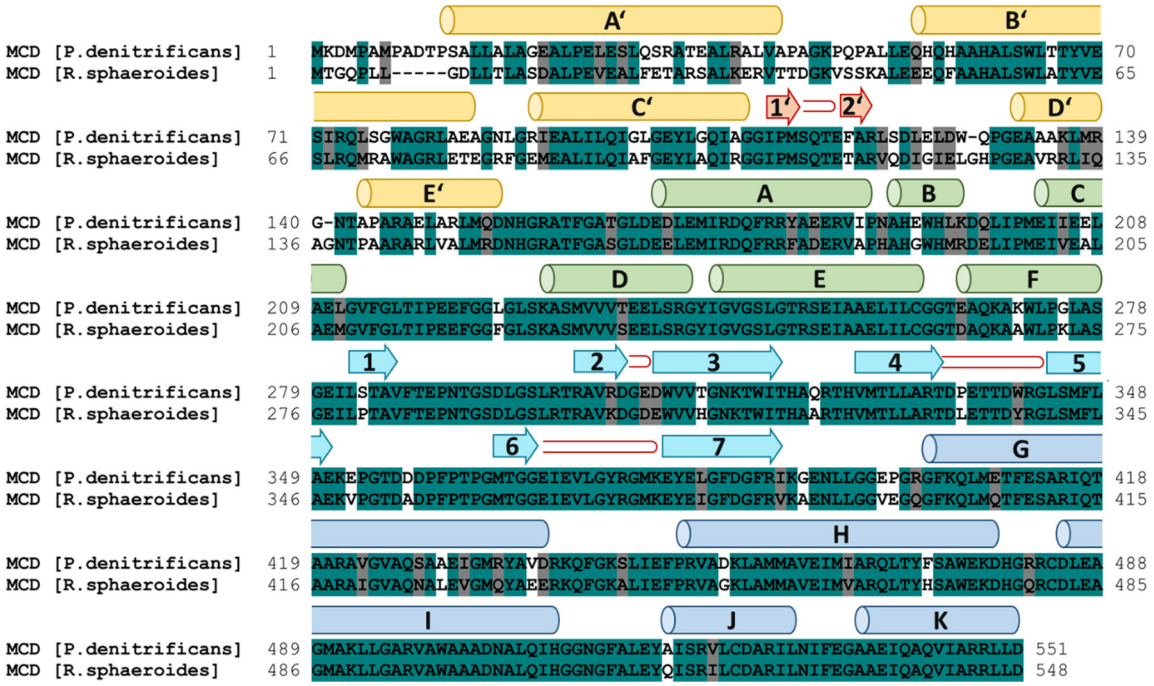



**SI Fig. 1 Additional Michaelis-Menten kinetics of enzymes characterized in this study (see Table 4 and Table 7 for values)**

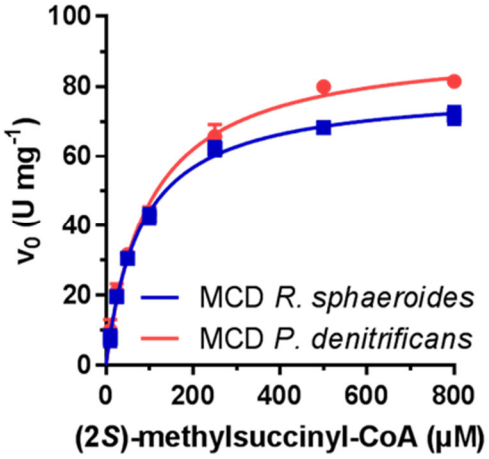
## 8.3. Crystal structure of MCD




**SI Fig. 2** (A) Superposition of a MCD monomer in orange (N-terminal domain in dark orange) with an ACO monomer (PDB ID: 1IS2) (4.82 Å over 591 Cα-atoms). The central fold is also conserved in ACOs and the structures aligns to some extent. (B) Overlay of a MCD monomer with the dimer of ACO in cyan.




**SI Fig. 3** The sequence alignment of the MCDs of *Paracoccus denitrificans* and *Rhodobacter sphaeroides* shows an overall identity of 77.8% and a similarity of 86.1%. The catalytic core domain (without the N-terminal extension) shows 88% identity and 95% similarity in the amino acid sequence.




| | (2*S*)-methylsuccinyl-CoA | | |
|---|---|---|---|
| | $V_{max}$ (U mg$^{-1}$) | app. $K_M$ (µM) | $k_{cat}/K_M$ ($s^{-1}$ $M^{-1}$) |
| MCD R.sphaeroides | 79.6 ± 1.3 | 80 ± 5 | $1.0 \times 10^6$ |
| MCD P.denitrificans | 92.6 ± 2.1 | 99 ± 8 | $1.0 \times 10^6$ |

**SI Fig. 4** Michaelis-Menten kinetics of the MCDs of *Rhodobacter sphaeroides* and *Paracoccus denitrificans* with (2*S*)-methylsuccinyl-CoA as substrate. Both enzymes show very similar catalytic properties.

# 9. Acknowledgments

# Curriculum vitae