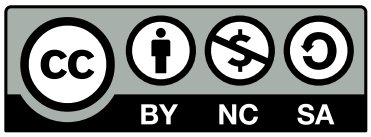
Engineering Enzymes and Pathways for Alternative CO_2 Fixation and Glyoxylate Assimilation

Dissertation

zur Erlangung des Grades eines
Doktor der Naturwissenschaften
(Dr. rer. nat.)
des Fachbereichs Biologie der Philipps-Universität
Marburg

Vorgelegt von
Pascal Pfister
aus Biedenkopf

Marburg, 2023



"IUCUNDI ACTI LABORES"

- Cicero

Contents

Publication list

The following articles were published during this doctorate;

★1. **Pascal Pfister**, Jan Zarzycki, and Tobias J. Erb. Structural basis for a cork-up mechanism of the intramolecular mesaconyl-CoA transferase. *Biochemistry*, 62(1):75–84, dec 2022

Contribution: Design and execution of all experiments, analysis of all data except of one structure (PDB 8APQ) and (co-)writing of the manuscript.

2. Helena Schulz-Mirbach, Alexandra Müller, Tong Wu, **Pascal Pfister**, Selçuk Aslan, Lennart Schada von Borzyskowski, Tobias J. Erb, Arren Bar-Even, and Steffen N. Lindner. On the flexibility of the cellular amination network in *E. coli*. *eLife*, 11, jul 2022

Contribution: Enzyme purification and kinetic characterisation of all enzymes and contribution to the manuscript

3. Maren Nattermann, Simon Burgener, **Pascal Pfister**, Alexander Chou, Luca Schulz, Seung Hwan Lee, Nicole Paczia, Jan Zarzycki, Ramon Gonzalez, and Tobias J. Erb. Engineering a highly efficient carboligase for synthetic one-carbon metabolism. *ACS Catalysis*, 11(9):5396–5404, apr 2021

Contribution: Collection and analysis of xray crystallography data (PDB 7AYG, 7B2E) and contribution to the manuscript

4. Marieke Scheffen, Daniel G. Marchal, Thomas Beneyton, Sandra K. Schuller, Melanie Klose, Christoph Diehl, Jessica Lehmann, **Pascal Pfister**, Martina Carrillo, Hai He, Selçuk Aslan, Niña S. Cortina, Peter Claus, Daniel Bollschweiler, Jean-Christophe Baret, Jan M. Schuller, Jan Zarzycki, Arren Bar-Even, and Tobias J. Erb. A new-to-nature carboxylation module to improve natural and synthetic CO_2 fixation. *Nature Catalysis*, 4(2):105–115, jan 2021

Contribution: Screening of Propionyl-CoA Carboxylases for activity on glycolyl-CoA, establishment of a biotin ligase expression strain and contribution to the manuscript

The following articles are submitted to peer review during the doctorate;

★ 5. **Pascal Pfister**, Christoph Diehl, Eric Hammarlund, Martina Carrillo, Tobias J. Erb. Enhancing substrate specificity of the succinyl-CoA reductase of *Clostridium kluyveri*. *Manuscript submitted*, feb 2023

Contribution: Discovery of SucD sidereaction with mesaconyl-C1-CoA, Design and execution analysis of all data (except kinetic characterisation of WT enzymes), Collection and analysis of xray crystallography data (PDB 8CEK, 8CEI, 8CEJ) and (co-)writing of the manuscript.

6. Maren Nattermann, Sebastian Wenk, **Pascal Pfister**, Nils Guntermann, Lennart Nickel, Charlotte Wallner, Jan Zarzycki, Nicole Paczia, Giancarlo Franciò, Walter Leitner, Arren Bar-Even and Tobias J. Erb. A new-to-nature cascade for phosphate-dependent formate reduction at the thermodynamic limit. *Manuscript in revision*, dec 2022

Contribution: Collection and analysis of xray crystallography data (PDB 8AFU, 8AFV) and contribution to the manuscript

7 Gabriele M. M. Stoffel, Matthias Tinzl, David Adrian Saez, Patrick D. Gerlinger, Rodrigo Recabarren, Carlos Castillo, Timothy Bradley, Hendrik Westedt, **Pascal Pfister**, Aharon Gomez, Nicole Paczia, Marc-Olivier Ebert, Esteban Vöhringer-Martinez, Tobias J. Erb. Development of the Biocatalytic Reductive Aldol Reaction. *Manuscript in revision*, nov 2022

Contribution: Collection and analysis of xray crystallography data (PDB 8BPP, 8BPQ) and contribution to the manuscript

Publications marked with ★, are featured as chapter in this thesis. The contributions of P.P. to the above mentioned publications or manuscripts are hereby assured.

Pascal Pfister Prof. Dr. Tobias J. Erb

Summary

Natural CO_2 fixation is mainly associated with the Calvin-Benson-Bassham (CBB) cycle found in many photoautotrophic organisms, e.g. cyanobacteria. The CBB cycle as well as its key enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) evolved in a atmosphere, that contained mainly CO_2 and merely any O_2. With emerging oxygenic photosynthesis and the oxygenation of the atmosphere, RuBisCO became increasingly inefficient. Its inefficiency to discriminate between both substrates, CO_2 and O_2, led to the evolution of carbon concentrating mechanisms (CCMs) and photorespiration. The latter is a metabolic route to remove the toxic side product of the oxygenase reaction, 2-phosphoglycolate (2PG) and recycle it into useable metabolites. During canonical photorespiration, at least one molecule of CO_2 would be released per two molecules of 2PG, reducing on biomass production at a notable margin. Among a variety of different approaches to mitigate this problem, examples for two of them will be discussed in this thesis. Synthetic photorespiration will be adressed via two chapters on the nature-inspired 3-hydroxypropionate (3OHP) bypass. Synthetic CO_2 fixiation will be features in one chapter about substrate selectivity in the new-to-nature crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle.

Photosynthetic organisms not always completely recycle photorespiratory 2PG, but also dephosphorylate and excrete glyoxylate to the surrounding medium. Other bacteria, like the thermophile *Chloroflexus aurantiacus* can feed on these acids and evolved a pathway, the 3OHP bi-cycle to metabolize them without the loss of CO_2. This inspired a synthetic photorespiration pathway, the 3OHP bypass. The first attempt

to introduce this pathway into the cyanobacterium *Synechococcus elongatus* were performed by Shih et al. [1].

Chapter 3 features the continued efforts to improve the 3OHP bypass in *S. elongatus*. A improved selection scheme, based on a carboxysome knockout strain and the pathway based detoxification of propionate were utilized to evolve a part of the 3OHP bypass in a turbidostat setup. The high CO_2 requiring strain improved its tolerance from 0.5% to 0.2% within 125 days.

Among the 3OHP bi-cycle enzymes are some catalysts with unique properties, like the intramolecular CoA transferase, mesaconyl-C1-C4-CoA CoA transferase (Mct). Chapter 4 is dedicated to a structural analysis on why this enzyme can be exclusively intramolecular. It has a narrow active site, that allows the CoA moiety of mesaconyl-CoA to blocks external acids from entering. A protein structure with trapped intermediates and kinetic analysis with external acids support this claim.

Additionally we investigated a promiscuous succinic semialdehyde dehydrogenase (SucD) that is featured in synthetic CO_2 fixation pathways, as described in chapter 2. SucD from *Clostridium kluyveri* is promiscuous to other CoA esters and especially active with mesaconyl-C1-CoA, another intermediate of the CETCH cycle. This side reaction will slowly drain mesaconyl-CoA from the pool of intermediates and lead to the accumulation of mesaconic semialdehyde. The specificity was addressed by solving the crystal structure of CkSucD and closing the active site by the substitution of an active site lysin to arginine. The mutation decreased site activity from 16% to 2%, but the overall efficiency decreased. In another SucD from *Clostridium difficile*, the same mutation had a comparable effect, changing the sidereaction from 12% to 2%, while conserving the overall efficiency. The designed enzyme is a wortwhile replacement for future iterations of the CETCH cycle.

Zusammenfassung

Der CBB-Zyklus ist der verbreiteste CO_2 Fixierungsweg, der in vielen photoautotrophen Organismen, z. B. Cyanobakterien, vorkommt. Der CBB-Zyklus sowie sein Schlüsselenzym RuBisCO entwickelten sich allerdings in einer Atmosphäre, die hauptsächlich CO_2 und nur wenig O_2 enthielt. Mit dem Aufkommen der oxygenen Photosynthese und der Sauerstoffanreicherung der Atmosphäre wurde RuBisCO zunehmend ineffizienter. Da dieses Enzym nur bedingt zwischen CO_2 und O_2 unterscheiden kann, entwickelten Organismen mit diesem Stoffwechsel sogenannte Kohlendioxid-Konzentrationsmechanismen (engl. CCMs) und Photorespiration. Letztere ist ein Stoffwechselweg zur Entfernung des toxischen Nebenprodukts der Oxygenase-Reaktion, 2-Phosphoglykolat (2PG), mit dem Ziel es als verwertbare Metaboliten zurück zu gewinnen. Bei der klassischen Photorespiration würde mindestens ein CO_2 für zwei Moleküle 2PG freigesetzt, was die Produktion von Biomasse deutlich reduziert. In dieser Arbeit werden zwei Beispiele für die verschiedenen Ansätze zur Lösung dieses Problems diskutiert. In zwei Kapiteln wird die synthetische Photorespiration anhand des von der Natur inspirierten 3-Hydroxypropionat (3OHP) Bypass (eng. 3OHP bypass) behandelt. Synthetische CO_2-Fixierung wird in einem Kapitel über die Substratselektivität eines Enzyms des kürzlich publizierten CETCH Zyklus behandelt. Photosynthetische Organismen recyceln photorespiratorisches 2PG nicht immer vollständig, sondern dephosphorylieren es auch zu Glyoxylat und geben es an die Umgebung ab. Andere Bakterien, wie das thermophile Bakterium *C. aurantiacus*, können diese Säuren aufnehmen und haben einen Stoffwechselweg, den 3OHP-Bi-Zyklus,

um sie ohne CO_2 Verlust zu verstoffwechseln. Diese Erkenntnis inspirierte die Entwicklung eines synthetischen, carboxylierenden Weg der Photorespiration an, die 3OHP Bypass. Der erste Versuch diesen Weg in das Cyanobakterium *S. elongatus* einzuführen, wurde von Shih et al. durchgeführt [1].

In Kapitel 3 werden die fortgesetzten Bemühungen zur Verbesserung des 3OHP Umweges in *S. elongatus* beschrieben. Es beschreibt ein verbessertes Selektionsschema, das auf einem Carboxysomen-Knockout-Stamm und der Gegenselektion von Propionat, als Wachstumshemmer beruht. Ein Teil des 3OHP Umweges konnte damit in einem Bioreaktor im Turbidostatmodus durch Selektion auf das schnellste Wachstum evolviert werden. Der Stamm mit initial hohem CO_2 Bedarf erhöhte seine Toleranz innerhalb von 125 Tagen von 0.5% auf 0.2 % CO_2.

Zu den Enzymen des 3OHP-Bi-Zyklus gehören Katalysatoren mit einzigartigen Eigenschaften, wie die intramolekulare CoA-Transferase, Mesaconyl-C1-C4-CoA CoA-Transferase (Mct). Das Kapitel 4 widmet sich der Strukturanalyse dieses Enzyms und der einzigartigen Selektivität auf den intramolekularen Transfer. Mct hat ein sterisch eingeschränktes aktives Zentrum, das durch den CoA-Teil von Mesaconyl-CoA für externe Säuren versperrt wird. Eine Proteinstruktur mit gebundenen Intermediaten und die Kinetik mit externen Säuren unterstützen diese Behauptung.

In Kapitel 2 wird die unspezifische Succinat-Semialdehyd-Dehydrogenase (SucD) die im hier ebenfalls beschriebenen synthetischen CO_2-Fixierungswegen CETCH vorkommt behandelt. Das Enzyme SucD aus *C. kluyveri* ist unspezifisch gegenüber anderen CoA-Estern, insbesondere gegenüber Mesaconyl-C1-CoA, einem weiteren Intermediat des CETCH Zyklus. Die Nebenreaktion entfernt Mesaconyl-C1-CoA aus dem Kreislauf und beeinträchtigt damit seine Funktion. Die Spezifizität wurde durch die Lösung der Kristallstruktur von CkSucD und die Schließung des aktiven Zentrums erhöht. Dazu wurde ein Lysin im aktiven Zentrum durch Arginin ersetzt. Die Mutation verringerte die Aktivität von 16% auf 2%, aber die Gesamteffizienz nahm ab. Bei einem Homolog aus *C. difficile* hatte diese Mutation eine vergleichbare Wirkung, indem sie

die Nebenreaktion von 12% auf 2% senkt ohne dabei die Effizienz mit Succiny-CoA zu beeinträchtigen. Das entworfene Enzym ist ein brauchbarer Ersatz für zukünftige Iterationen des CETCH-Zyklus.

Chapter 1

Introduction

1.1 Aim of this work

Anthropogenic climate change and hence the resulting climate crisis are one of the biggest challenges for our current and coming generations [2]. Artificial CO_2 fixation can contribute to the re-fixation of the emitted carbon from the atmosphere and use it for (i. e. drug producing) industrial applications [3, 4]. One synthetic CO_2 fixation pathway, the so called CETCH cycle may exceeds the efficiency of natural CO_2 fixation in the near future. To achieve this, a set of efficient catalysts is required to avoid bottlenecks and the generation of unwanted side products. An enzyme of this pathway, SucD has a notable sidereaction with mesaconyl-C1-CoA another intermediate of the CETCH cycle. We adressed this by solving the structure of SucD together with the mesaconyl-C1-CoA ligand and identified coordinating residues. These residues were targeted by site-directed-mutagensis and the resulting mutants were tested for improved specifity. We saw that a substitution on an active site lysine to arginine would occlude mesaconyl-C1-CoA from the active site in two related isoenzymes.
Besides whole new CO_2 fixation pathways also metabolic enhancements for natural photosynthesis and the CBB cycle could help to solve the problem of inefficient CO_2 fixation. The 3OHP bypass is a synthetic photorespiratory pathway that mittigates photorespiratory loss in carbon biomass. As the CETCH is also resulting in the formation of glyoxylate, photorespiratory pathways have the potential to increase the efficiency of synthetic CO_2 fixation even more [4, 5]. We prototyped the 3OHP bypass in the cyanobacterium *S. elongatus* PCC7942 based on a previous study [1]. We originally approached the prototyping with rational design, like genetic expression balancing, site directed mutagenesis of the bottleneck enzyme acetyl-CoA carboxylase (Acc), and also focused on a adaptive laboratory evolution (ALE) experiment to improve the pathway's integration into the host's metabolism.. We combined the selective pressure of photorespiratory carbon drain in a carboxysome deficient strain, with the toxicity of propionate, that can only be detoxified by the 3OHP bypass module, that we introduced. We operated a turbidostat for 125 days and observed an adaptation towards

lower concentrations of CO_2. However, further characterization has to be done, to verify the adaptations and mutations, that this population developed. The 3OHP bypass comprises enzymes with unique properties, like the intramolecular Mct. We explained the reaction mechanism of this CoA transferase by solving the structure, with reaction intermediates of all steps of catalysis. As mesaconyl-CoA enters the active site, the CoA moiety blocks the cavity entrance and the CoA tranfer is performed exclusively from and onto the original mesaconate moiety.

1.2 Metabolic Pathways for the Fixation of CO_2

1.2.1 Calvin-Benson-Bassham Cycle

The CBB cycle is a oxygen tolerant CO_2 assimilation pathway that is also known as the dark reaction of photosynthesis. It is the typical carbon assimilation cycle of oxygenic phototrophs, and therefore present in plants, algae and cyanobacteria [6]. Its wide distribution makes the CBB cycle the most important process for primary production on Earth [7,8].

Typically the reactions of the CBB cycle are categorized in 3 phases. The first is the actual carbon fixation step, which is catalyzed by the enzyme RuBisCO (Figure 1.1). This enzyme uses ribulose-1,5-bisphosphate and CO_2 as substrates and forms two molecules of 3-phosphoglycerate (3PG). The next phase is the reduction of 3PG to glyceraldehyde-3-phosphate (G3P), which consumes energy in form of adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH). In the following regeneration step G3P is converted back to ribulose-1,5-bisphosphate through the reductive pentose phosphate pathway to start the next cycle [6]. However, the energy equivalents in organism using the CBB cycle often, but not exclusively orginate from oxygenic photosynthesis [9,10].

Even though other carbon fixation pathways exist in nature, the lifestyle of oxygenic photosynthesis became historically relevant during the great oxygenation transi-

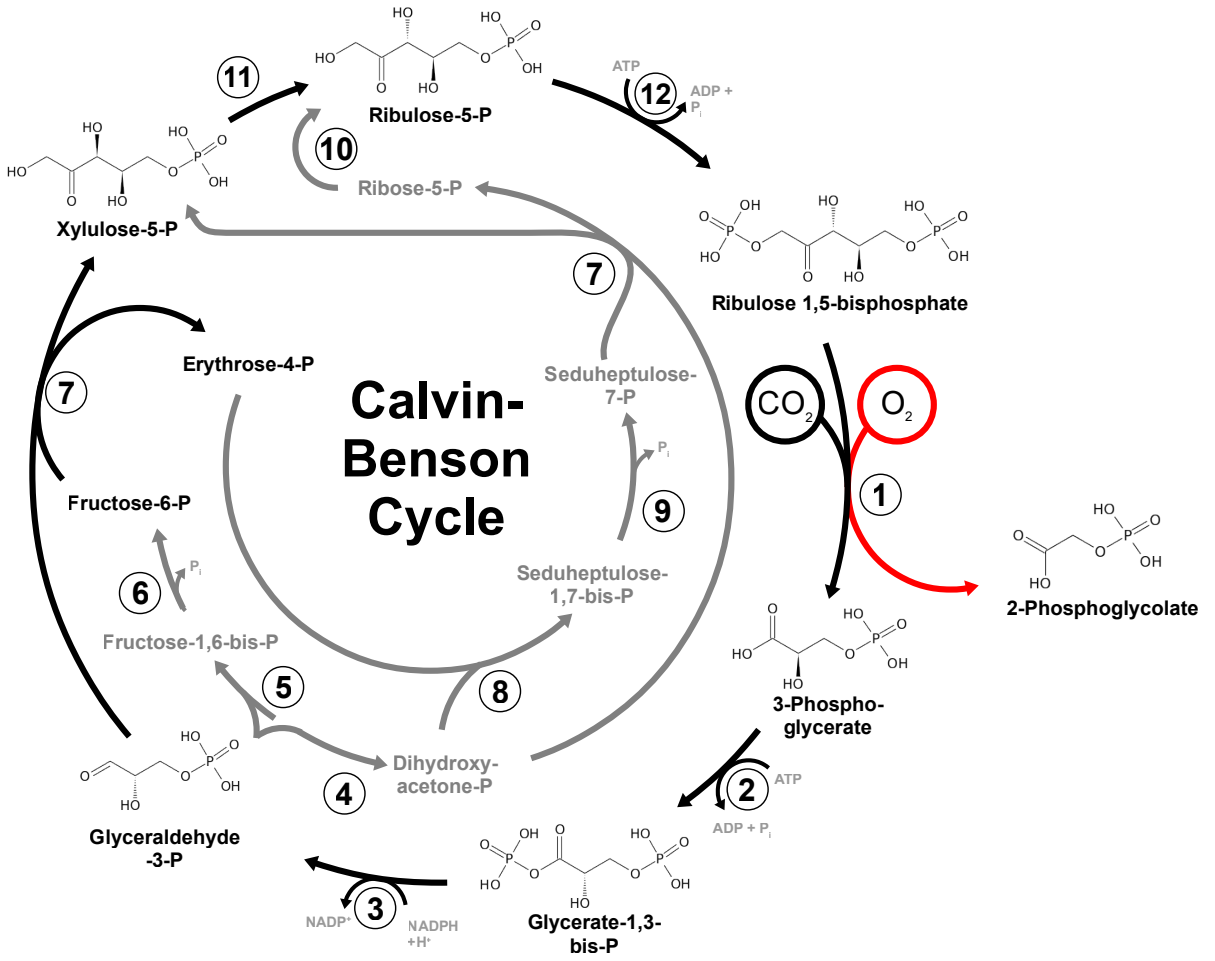


Figure 1.1: The Calvin-Benson-Bassham (reductive pentose phosphate) cycle. Ribulose-1,5-bisphosphate carboxylase/oxygenase(1), 3-phosphoglycerate kinase (2), glyceraldehyde-3-phosphate dehydrogenase (3), triose-phosphate isomerase (4), fructose-bisphosphate aldolase (5), fructose-bisphosphate phosphatase (6), transketolase (7), sedoheptulose-bisphosphate aldolase (8), sedoheptulose-bisphosphate phosphatase (9), ribose-phosphate isomerase (10), ribulose-phosphate epimerase (11) and phosphoribulokinase (12) [6, adapted])

tion (GOT) [6,8,11]. Prior to this event, which occurred approximately 2.4 billion years ago, cyanobacteria evolved the ability to use light energy to split water molecules and release oxygen gas as a byproduct [12]. This process transformed the atmosphere by producing large amounts of oxygen and reducing levels of carbon dioxide. This started an unprecedented transition in the history of life on Earth, as it resulted in the development of aerobic respiration and the evolution of complex life forms. As photosynthetic organisms began to produce oxygen through oxygenic photosynthe-

sis, the Earth's atmosphere changed from reducing to oxidizing. A consequence of the changed atmosphere, the temperature, sulfur and methane availability changed, causing a mass extinction of the prevalent oxygen sensitive life forms [11].

From an evolutionary perspective, the CBB cycle was a fast and effective metabolic invention. Under low O_2 concentrations in the atmosphere before and during the GOT the CBB cycle was performing efficiently. However with rising levels of oxygen, a momentous side reaction of RuBisCO became problematic. Instead of carboxylating ribulose-1,5-bisphosphate to form two molecules of 3PG, the substrate is oxidized with molecular oxygen leading to one molecule of 3PG and one molecule of 2PG [13,14].

Nowadays, with a high oxygen concentration in the ambient air RuBisCO has evolved at an interface between specificity and rate, to avoid the formation of 2PG while not sacrificing too much turnover. Evolutionary adaptions of RuBisCO on an enzymatic level towards higher affinity to CO_2 usually decreased turnover rate, raising the need to produce more and more of the enzyme. [15]. On a cellular level adaptations like CCMs had a positive effect by locally increasing the CO_2 concentration [16].

However, the main effort for an organism dealing with this side reaction lies in the removal of toxic 2PG by photorespiration (see Fig. 1.3). In canonical photorespiration, one molecule of CO_2 is released per 2 molecules of 2PG. The lost carbon has to be reassimilated throught the CBB cycle again. This strategy allows phototrophic organisms to regain most of the carbon from photorespiration, but the efficiency of the CBB is still lowered. Synthetic biology is trying to adress this with different approaches. There is a set of natural and synthetic CO_2 fixation pathways, that might replace the CBB in crop plants some day. A noteable metabolic strategy, among the six natural CO_2 fixation pathways is the 3OHP bi-cycle [6]. However, there are also synthetic CO_2 fixation pathways, like the CETCH cycle [3]. They harbor the chance for deliberate the matter of CO_2 fixation from evolutionary linearity, as they pair enzyme from different metabolic and ecological niches. Attempts to increase efficiency of the

CBB cycle have been undertakes as well. At this point, a set of different photorespiratory pathways are proposed and some even showed a positive impact on biomass production [17]. The aforementioned 3OHP bi-cycle also inspired a photrespiratory pathways, the 3OHP Bypass.

1.2.2 The 3-hydroxypropionate bi-cycle

The green non-sulfur bacterium *C. aurantiacus*, found in extreme environments like hot microbial spring mats can grow heterotrophically, mixotrophixally or autotrophically. *C. aurantiacus* grows on a variety of different organic acids, including on gly-

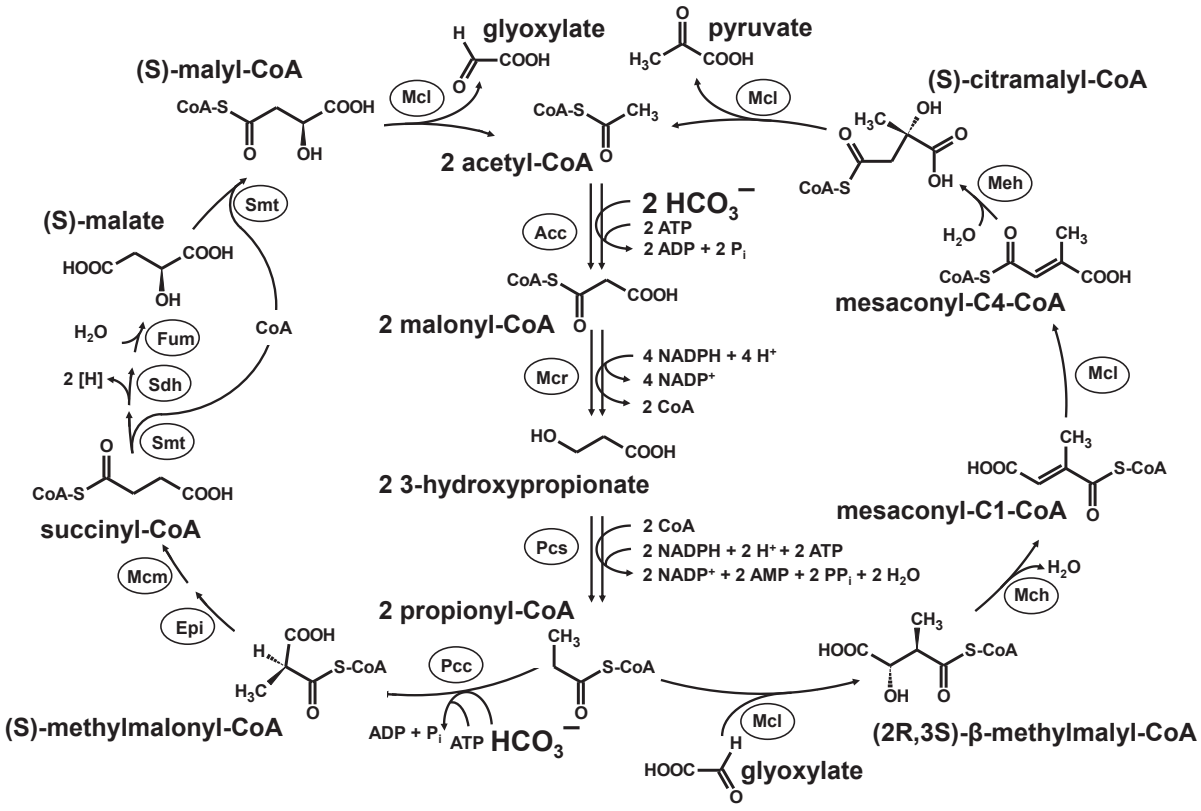


Figure 1.2: The 3OHP bi-cycle (adapted from [18]): acetyl-CoA carboxylase(Acc), malonyl-CoA reductase (Mcr), propionyl-CoA synthase (Pcs), propionyl-CoA carboxylase (Pcc), methylmalonyl-CoA epimerase (Epi), methylmalonyl-CoA mutase (Mcm), succinyl-CoA:(S)-malate-CoA transferase (Smt), succinate dehydrogenase (Sdh), fumarate hydratase (Fum), (S)-malyl-CoA/b-methylmalyl-CoA/(S)-citramalyl-CoA (MMC) lyase (Mcl), mesaconyl-C1-CoA hydratase (b-methylmalyl-CoA dehydratase) (Mch), mesaconyl-C1:C4-CoA CoA transferase (Mct), mesaconyl-C4-CoA hydratase (Meh).

colate and glyoxylate, but is also able to grow photoautotrophically under anoxic conditions. However, instead of the CBB cycle *C. aurantiacus* is able to assimilate CO_2 via the 3OHP bi-cycle (Figure 1.2). This recently discovered pathway consists of two part-cycles, that actually share a sequence of essential reactions. Both part-cycles share the carboxylation of acetyl-CoA and the reduction of the carboxylation product malonyl-CoA, which is reduced to 3OHP. The last shared reactions convert 3-hydroxypropionate to propionyl-CoA. In the first cycle propionyl-CoA is further carboxylated and transformed into (S)-malyl-CoA, which is then cleaved into acetyl-CoA and glyoxylate. Glyoxylate is condensed with another propionyl-CoA in the second cycle. In the end the second cycle yields acetyl-CoA and pyruvate (see Fig. 1.3). As the 3OHP bi-cycle is a unique pathway, these enzymes are of noteable scientific interest. propionyl-CoA synthase (Pcs) for example is a unique multidomain enzyme with a latent carboxylation capability [19,20]. Malonyl-CoA reductase (Mcr) is another complex enzyme with two reductase domains, that produce 3OHP from malonyl-CoA. This enzyme finds application in biotechnology [21,22]. Mesaconyl-C1-C4-CoA CoA transferase of *C. aurantiacus* also has an interesting reaction mechanism. It belongs to the CoA transferases of Family III, which have a well understood reaction mechanism for the transfer of a CoA moiety from a donor acyl-CoA, to an acceptor acid. However, Mct seems to exclusively transfer the CoA moiety on the same dicarboxylic acyl-group, that acted as CoA donor before, making it a genuine intramolecular CoA transferase (see 4). In the natural environment of *C. aurantiacus*, cyanobacteria excrete glycolate as a consequence of photorespiratory stress. *C. aurantiacus* utilizes this compound and oxydizes it to glyoxylate, which can be assimilated by the 3OHP bi-cycle. In case no external glycolate is available, the first partcycle assimilates CO_2 to produce glyoxylate. A glyoxylate dependent anabolism however, is an ideal basis to evolve a more efficient photorespiratory pathway. The 3OHP bi-cycle is therefore a good opportunity to bypass the quite deleterious photorespiration.

1.2.3 The CETCH Cycle as Synthetic CO_2 Fixation Pathway

The CETCH cycle is based on a combination of metabolic pathways found in bacteria [23]. It involves the use of 17 different enzymes, with crotonyl-CoA carboxylase/reductase (Ccr) as central carboxylase. These enzymes work together to convert CO_2 into organic compounds. It primarely produces glyoxlyate, that can be metabolized to complex molecules [3,4]. The CETCH cycle has shown promising results in in vitro experiments, where it has been shown to be more efficient than the Calvin cycle at fixing CO_2 [3]. However, it has not yet been implemented in living organisms, and more work needs to be done to optimize the pathway and to develop ways to introduce it into plants and other photosynthetic organisms [24].

To choose efficient catalysts is essential for developing a synthetic CO_2 fixation pathway like the CETCH cycle. In metabolic engineering, the choice of enzymes is critical to optimize the efficiency of metabolic pathways. By selecting enzymes with high catalytic efficiency and appropriate kinetics, the rate of the pathway can be increased, which in turn can improve the yield of the desired product.

In the CETCH cycle, the choice of enzymes is not less important for the optimization of CO_2 fixation. The enzymes used in the pathway need to have high catalytic activities, specificities for their substrates, and low levels of unwanted side reactions. In addition, the enzymes need to be able to function in the complex mixture of other enzymes and metabolites that are present in the pathway. Selecting enzymes with the desired characteristics, and by modifying the genes that encode these enzymes to optimize their expression and activity. Directed evolution or rational design can also be used to develop new enzymes with improved properties for the pathway. As we characterized the mechanistic properties of Mct of *C. aurantiacus*, we realized, that SucD, an important catalyst in the CETCH cycle, also accepts mesaconyl-C1-CoA and mesaconyl-C4-CoA as substrate. As mesaconyl-C1-CoA is a intermediate of the CETCH cycle, we sought to understand the molecular basics of this promiscuity and wanted to find a more specific enzyme for future CETCH cycle iterations (see

chapter 2). Lastly, the CETCH cycle primarely produces gloxylate, which leaves the same problem as photorespiration born 2PG, assimilation in a living organism would require an additional glyoxylate sink, that is ideally carbon neutral like the recently rediscovered β-hydroxyaspartate cycle (BHAC) [25]. An additional carboxylation, and a redirection into central carbon metabolites as it is done in the tartronyl-CoA (TaCo) pathway [5] or the 3OHP bypass, would allow an even more efficient carboxylation module.

1.2.4 The 3-hydroxypropionate Bypass

The 3OHP bypass is a synthetic pathway that has been envisioned to increase the carbon efficiency of photosynthesis and to mitigate the effects of photorespiration. The 3OHP bypass was inspired by the 3OHP bi-cycle [1, 18]. The 3OHP bi-cycle is an alternative CO_2 assimilation pathway found in *C. aurantiacus* that lives in extreme environments, such as hot springs [26]. This pathway is thought to have evolved as a way to cope with low levels of CO_2 and high temperatures, and it is more efficient than the CBB cycle at fixing carbon dioxide if CO_2 is scarce, as it makes use of the more readily available bicarbonate as the carboxylating carbon species.

Enzymes of the 3OHP bypass have been introduced into cyanobacteria to create a carbon-efficient photorespiratory pathway [1]. By using the 3OHP bypass, cyanobacteria and plants could bypass the energy-intensive steps of photorespiration and conserve more of the energy and carbon fixed during photosynthesis. While the 3OHP bypass has only conceptually been adressed [1], it has the potential to improve the efficiency of photosynthesis and to help plants grow more efficiently under stressful environmental conditions like drought or heat. The 3OHP bypass is an example of how synthetic biology and biotechnology can be used to develop new tools for improving efficient and sustainable agriculture.

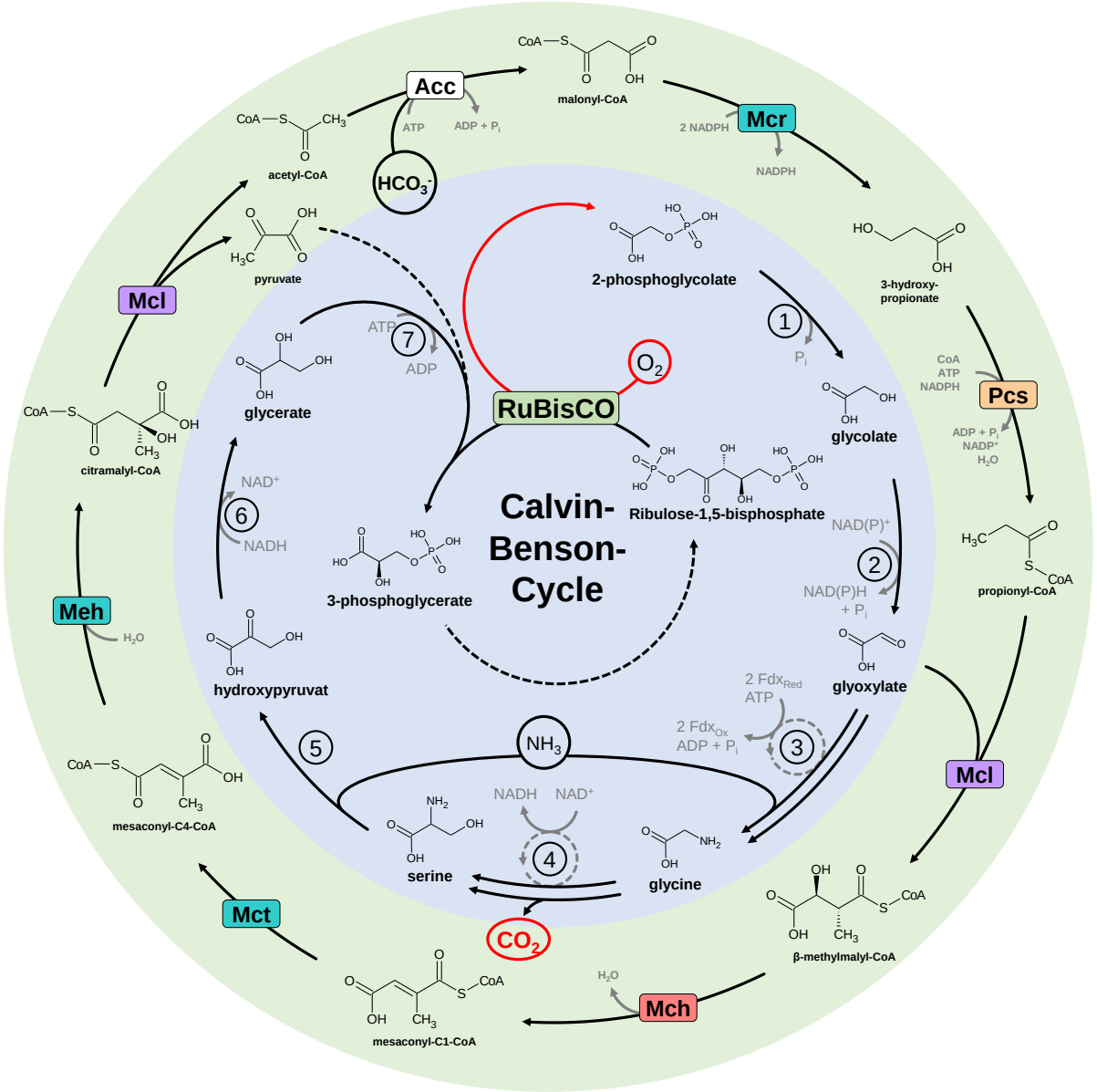


Figure 1.3: **Natural Photorespiration (inner circle) along with the 3OHP bypass (outer circle):** ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), 2-phosphoglycolate dephosphatase (1), glycolate dehydrogenase (2), glutamate-glyoxylate aminotransferase + glutamine synthetase + glutamine oxoglutarate aminotransferase (3), serine hydroxymethyltransferase (4), serine-glyoxylate transaminase + glycine decarboxylase (5), hydroxypyruvate reductase (6), D-glycerate 3-kinase (7). acetyl-CoA carboxylase (Acc), malonyl-CoA reductase (Mcr), propionyl-CoA synthase (Pcs), β-methylmalyl- CoA/(*S*)-citramalyl-CoA lyase (Mcl), mesaconyl-C1-CoA hydratase (Mch), mesaconyl-C1-C4-CoA CoA transferase (Mct), mesaconyl-C4-CoA hydratase (Meh). (adapted [1,6]).

1.3 The Cyanobacterium *S. elongatus*

1.3.1 *S. elongatus* as a host for synthetic photorespiration

S. elongatus PCC7942 is a unicellular cyanobacterium that has been extensively studied as a model organism for photosynthesis and carbon metabolism. It is a common model organism to investigate the molecular mechanisms of the CBB cycle, as well as other processes such as circadian rhythms [27] and light signaling [28]. *S. elongatus* PCC7942, like other cyanobacteria, has evolved CCMs to overcome the problem of photorespiration and improve the efficiency of CO_2 fixation. The key element of the cyanobacterial CCM is the carboxysome. The carboxysome is a proteinaceous microcompartment that contains a high concentration of RuBisCO and carbonic anhydrase [16]. By concentrating Rubisco and CO_2 within the carboxysome, the rate of CO_2 fixation is increased and the rate of photorespiration is reduced. This allows *S. elongatus* PCC7942 to use available carbon and energy resources more efficiently, and to adapt to changing environmental conditions.

In addition to carboxysomes, *S. elongatus* PCC7942 also uses other CCM components such as bicarbonate transporters [29], which help to concentrate CO_2 in the form of bicarbonate and increase the efficiency of photosynthesis. These CCM elements work together to optimize the efficiency of CO_2 fixation and minimize the effects of photorespiration. Cyanobacteria, and more specifically *S. elongatus* have a number of other adaptations that allow them to thrive as photosynthetic organisms. In addition to CCMs like carboxysomes, they also have unique storage and energy metabolism strategies, with the most relevant beeing glycogen. It is a primary storage compound for cyanobacteria. Glycogen is a sugar polymer that is metabolically close to ribulose-1,5-bisphosphate, the substrate for RuBisCO. This allows *S. elongatus* to refill CBB cycle intermediates directly in states of starvation [30].

Cyanobacteria are able to use light energy to generate ATP and reducing equivalents like NADPH through photosynthesis. This allows them to use energy-rich molecules

like glycogen or 3PG for anabolic processes like amino acid or cell wall biosynthesis. In contrast, heterotrophic organisms like *Escherichia coli* rely on catabolic processes like the TCA cycle to generate ATP and reducing equivalents trough respiration.
The TCA cycle in cyanobacteria is typically incomplete, with some of the intermediates being diverted to biosynthetic pathways rather than being fully oxidized for energy production. This allows the cyanobacteria to use the TCA cycle to generate precursors for amino acid biosynthesis and other anabolic processes, rather than solely for energy production.

1.3.2 The TCA Cycle in *S. elongatus* PCC 7942

The TCA cycle in most cyanobacteria is modified or incomplete. The incomplete version of the TCA cycle is referred to as horseshoe variant. Some cyanobacteria do not metabolize 2-oxoglutarate (2OG) to succinyl-CoA. Most marine cyanobacteria circumvent this via a decarboxylation to succinic semialdehyde and the subsequent conversion into succinate [32]. *S. elongatus* PCC 7942 is one of these strains lacking the 2OG decarboxylase and, therefore, produces succinate and 2OG in two seperate pools (Figure 1.4). Additionally, there is no evidence in the genome of PCC 7942 for a malate dehydrogenase [31]. The presence of malate and succinate in metabolomics analysis, however, indicate that an unidentified unspecific dehydrogenase is replacing the activity of malate dehydrogenase [31]. In reference to a potential succinyl-CoA shunt in this strain succinyl-CoA or more importantly the corresponding acid succinate would remain as a terminal compound that can be utilized for additional anabolic pathways.

1.3.3 *S. elongatus* has a Central Regulator (P_{II}) to Avoid Carbon Drainage

As cyanobacteria have a conservative interest on their organic molecules, the fluxes into different anabolic processes have to be finely regulated. In case of nutrient de-

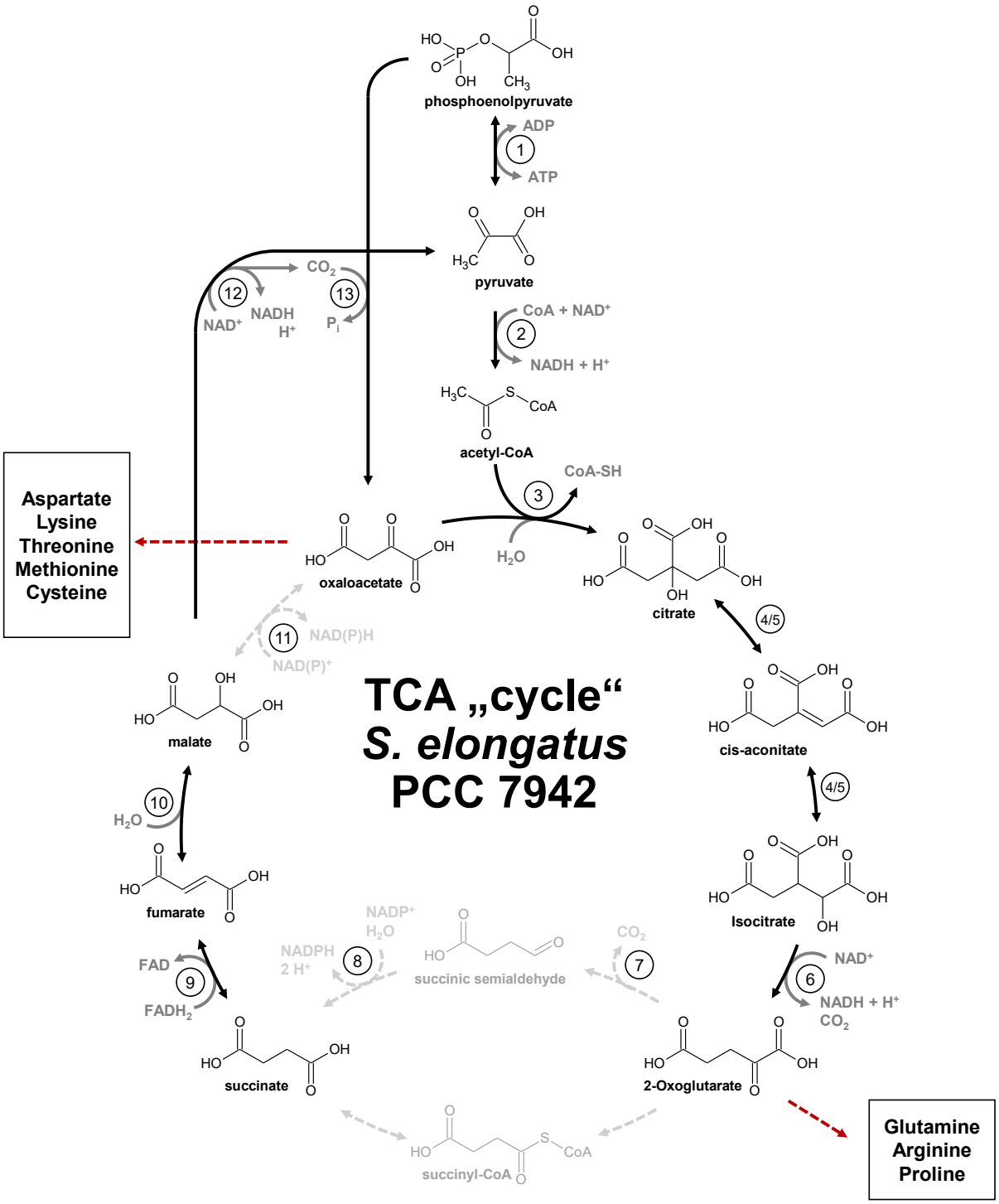


Figure 1.4: The TCA cycle in *S. elongatus* PCC 7942: pyruvate kinase(1), pyruvate dehydrogenase (2), 3. citrate synthase (3), aconitate hydratase 2 (4), 2-methylisocitrate dehydratase (5), isocitrate dehydrogenase (6), 2-oxoglutarate decarboxylase (7), aldehyde dehydrogenase (8); succinate dehydrogenase (9), fumarate hydratase (10), Unknown dehydrogenase (11), malate dehydrogenase (oxaloacetate-decarboxylating) (12), phosphoenolpyruvate carboxylase (13). Arrows shown in lightgray are not present in *S. elongatus* PCC 7942. Anabolic pathways for the production of amino acids are indicated in boxes with and red dashed lines [31–33].

ficiency like nitrogen starvation, the cells would avoid production of precursors for biosynthesis. For that matter, P_{II} is a central regulator that plays a critical role in the control of nitrogen metabolism in cyanobacteria [34]. It is encoded by the *glnB* gene and functions as a signaling protein that mediates signals from multiple nitrogen sources and modulates the activity of key enzymes involved in nitrogen metabolism. Specifically, P_{II} interacts with a number of target proteins to modulate their activity in response to changes in nitrogen availability. For example, P_{II} can bind to and regulate the activity of the enzyme N-acetylglutamate kinase, which is a key enzyme in the biosynthesis of arginine [34]. P_{II} can also interact with and modulate the activity of the transcription factor NtcA, which regulates the expression of many genes involved in nitrogen metabolism.

The activity of P_{II} is regulated by its binding to small molecules like ATP, ADP, and 2OG, which serve as indicators of the energy and nitrogen status of the cell. In response to changes in the availability of these small molecules, P_{II} can adopt different conformations that affect its interactions with target proteins and thus modulate their activity.

One of its targets, pirC [35], is an enzyme that regulates the drainage of metabolites from the sugar metabolism, which is important for maintaining carbon balance in the cell. The other target, acetyl-CoA carboxylase, is involved in the production of fatty acids, which are used as precursors for anabolism [36]. The 3OHP bypass also requires the activity of acetyl-CoA carboxylase.

Understanding the regulation of P_{II} could be important for optimizing the 3OHP bypass in cyanobacteria. By modulating P_{II}-Acc interaction, it may be possible to fine-tune the carbon and nitrogen balance in the cells and increase the efficiency of the 3OHP bypass. Additionally, the regulation of acetyl-CoA carboxylase by P_{II} could have implications for the production of other metabolites in cyanobacteria beyond the 3OHP bypass.

1.4 Previous Attempt to Implement the 3OHP Bypass into *S. elongatus*

The necessary genes from *C. aurantiacus* have already been introduced into PCC7942 for an improved photorespiration [1]. Initially, the gene encoding Acc was not integrated, as it already exists in all procaryotes. In this former study, the resulting strain showed no detectable growth improvement. Cell extract demonstrated activities for all required reactions, except the acetyl-CoA carboxylation, indicating, that this is the pathway bottleneck. We aimed to design a new improved iteration of the pathway and tried to solve problems that were identified [1,37,38].

The carboxysome shell proteins of *S. elongatus* PCC7942 are not essential under elevated CO_2 conditions. Knock-outs of the ccm operon resulted in a strain (ΔccmK-O) [39, 40], that was not able to produce carboxysomes anymore. Due to the lag of this CCM, the cells were not viable at ambient CO_2 concentrations. This strain could serve as host strain for artificial photorespiration, as it should grow better with an improved, more efficient photorespiration.

The authors of the original publication already speculated, that enzymes of a thermophilic organism may underperform at 30-37 °C, and that enzyme homologues from mesophilic host might be worthwhile to test and replace the ones, that do not perform better at lower temperatures. We identified β-methylmalyl-CoA/(S)-citramalyl-CoA lyase (Mcl) from *Candidatus Accumulibacter phosphatis*, mesaconyl-C1-CoA hydratase (Mch) from *Rhodobacter sphaeroides* and Pcs from *Erythrobacter spec.* as better variants [37].

Additionally, the regulatory elements of the genetic integration was fine tuned. Promoters that were known to work in *S. elongatus* PCC7942, the induceable promoter pLavS and the two constitutive promoters pPsbA1 and pCpt were used to express the gene for Mcl [37]. Mcl activities were measured in cell extracts, and the promoters were ranked by their relative values.

The identified bottleneck of the first iteration of the 3OHP bypass was the activity of Acc. Therefore, a variety of different other Accs were tested for *in-vitro* activity. Bacterial Accs usually consist of four subunits and their complex is notoriously instable during purification. However, we found one carboxylase, propionyl-CoA carboxylase (Pcc) from *Methylorubrum extorquens* that had a reasonable specific activity (s) of 100 mU/mg for acetyl-CoA. We performed site directed mutagenesis on this Pcc and generated a mutant Pcc_D407I, that showed elevated specific activities (s) for acetyl-CoA of 540 mU/mg. Next we integrated the designed Pcc_D407I and the Acc of *E. coli* as well as another copy of the Acc from *S. elongatus* under control of the best identified promoter into *S. elongatus* PCC7942 and verified integration. Surprisingly not even residual activities were measureable in these strains [37].

As we incubated lysates of these strains with additional (acyl-CoA carboxylase)biotin ligase (BirA), a ligase necessary for the maturation of Accs, we measured first residual acitivities of 7.8 ± 0.7 mU/mg for the Acc from *E. coli* in our cyanobacterial cell extract. Coproduction of the same BirA with Acc in *S. elongatus* was not leading to measureable activities [38].

We know now, that the biotin carboxyl carrier protein (BCCP) subunit of Accs and Pccs were highly regulated by the central regulator P_{II} in cyanobacteria [36]. Deletion of this central regulator is possible [41–43], but involves a co-deletion of more than one of its binding partners, and the resulting strain is not suitable for the burden of photorespiratory stress trough the deletion of its carboxysomes. It seems inevitable, to rather modify the BCCP subunit of Acc to break interaction with P_{II}, than vice versa. As the structure of the BCCP-P_{II} complex is undisclosed to date, we proposed an ALE experiment, to increase selective pressure on *S. elongatus* to unleash Acc from the controle of P_{II}, as will be discussed in chapter 3.

Chapter 2

Enhancing substrate specificity of *Clostridium* succinyl-CoA reductase for synthetic biology and biocatalysis

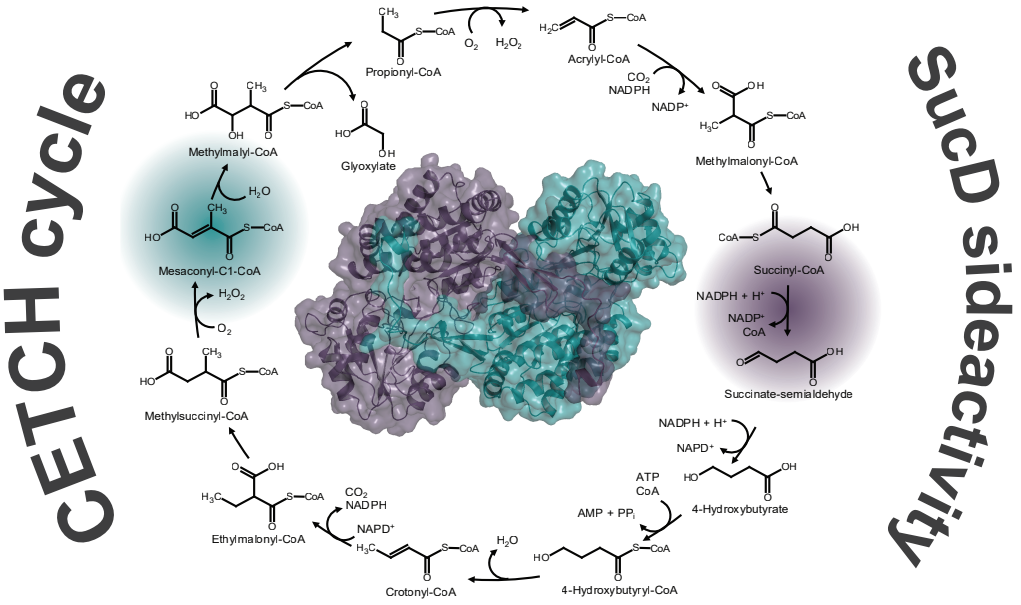
Pascal Pfister[1,*], Christoph Diehl[1,*], Eric Hammarlund[1], Martina Carrillo[1], Tobias J. Erb[1,2,°]

[1] Department of Biochemistry & Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch Str. 10, 35043 Marburg, Germany

[2] SYNMIKRO Center for Synthetic Microbiology, Karl-von-Frisch Str., 14, 35032 Marburg, Germany

* These two authors contributed equally to this work

° corresponding author: toerb@mpi-marburg.mpg.de



2.1 Abstract

Succinic semialdehyde dehydrogenase (SucD) is an acylating aldehyde reductase that catalyzes the NADPH-dependent reduction of succinyl-CoA to succinic semialdehyde and allows for co-assimilation of ethanol and succinate in *Clostridia*. The reaction sequence from succinate to crotonyl-CoA is of particular interest for several new-to-nature CO_2-fixation pathways, such as the CETCH cycle, in which SucD serves a key role. However, pathways like the CETCH cycle feature several CoA-ester intermediates, which could be potentially side-substrates for this enzyme. Here we show that the side reaction for most CETCH cycle metabolites is relatively small (<2%) with exception of mesaconyl-C1-CoA (16%), which represents a competing substrate in this pathway. We addressed this promiscuity by solving the crystal structure of a SucD of *C. kluyveri* in complex with $NADP^+$ and mesaconyl-C1-CoA. We further identified two residues (Lys70 and Ser243) that coordinate mesaconyl-C1-CoA at the active site. We targed those residues with site directed mutagenesis to improve succinyl-CoA over mesaconyl-C1-CoA reduction. The best resulting SucD variant, K70R, showed a

strongly reduced side for mesaconyl-C1-CoA, but also affected specific activity by a factor of 10. Transferring the same mutations into a SucD homologue from *C. difficile*, similarly decrease the side reaction of this enzyme from 12 to 2%, notably without changing specific activity of the enzyme. Overall, our structure-based engineering efforts provided a highly specific enzyme of interest for several applications in biocatalysis and synthetic biology.

2.2 Introduction

succinic semialdehyde dehydrogenase (SucD) is an acylating aldehyde reductase that catalyzes the NADPH dependent reduction of succinyl-CoA to succinic semialdehyde (SSA). In *C. kluyveri* and other *Clostridia* species, succinyl-CoA reductase operates in fatty acid fermentation, allowing for the co-assimilation of ethanol and succinate [44–49]. In these fermentations, succinate is first activated to succinyl-CoA, which is then reduced to SSA by SucD before being further converted into 4-hydroxybutyrate (Figure 2.1).

The reaction mechanism of SucD supposedly follows the canonical reaction mechanism of acylating aldehyde dehydrogenases [51]. In these enzymes, the respective acyl-CoA enters the active site, where the acyl-moiety is transferred onto an active site cysteine, leading to a covalent acyl-cysteine intermediate. The freed CoA moiety is protonated by a nearby histidine before exiting the active site, while the acyl-moiety is reduced to the aldehyde by NADPH.

Beyond its role in the fermentation of short-chain fatty acids, SucD has found increasingly attention for the realization of artificial CO_2 fixation pathways. The most prominent examples are the THETA [52]. and the CETCH cycle, which both have been established lately [3]. The CETCH cycle is a complex pathway for the assimilation of CO_2 [3]. It involves 17 different enzymes that were re-wired to form a new-to-nature pathway for the capture and conversion of CO_2. This pathway was reconstituted in

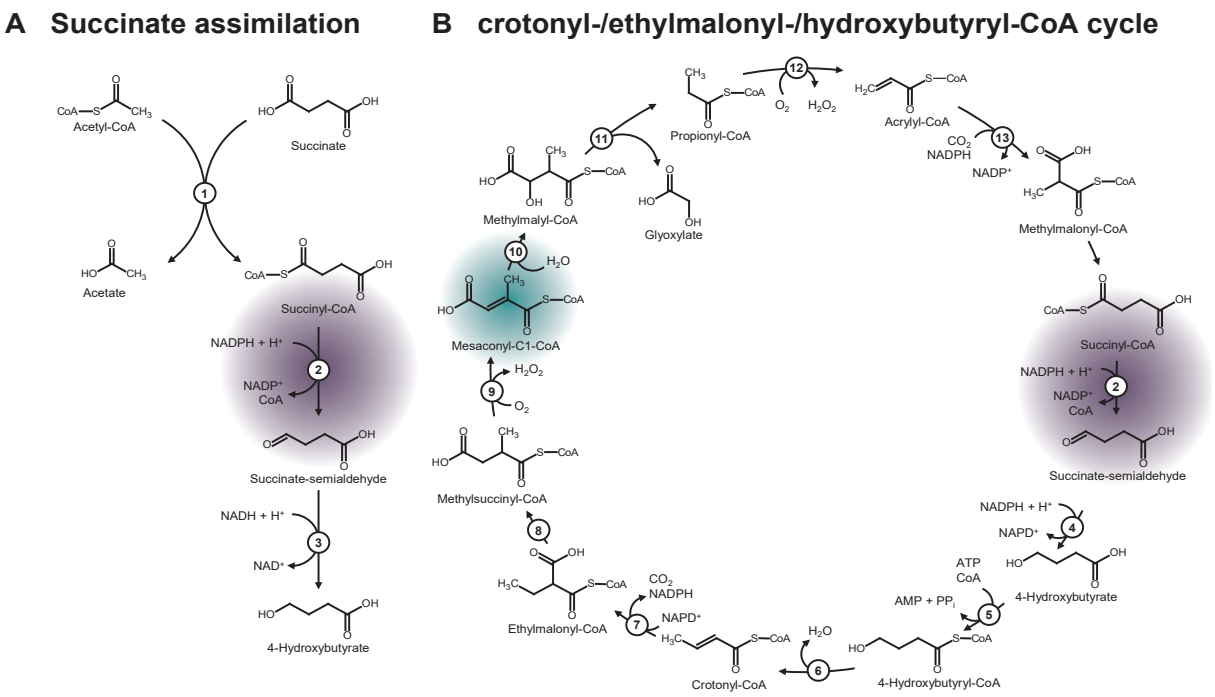


Figure 2.1: Succinate assimilation pathway of *C. kluyveri* [47, 50]. The enzymes, undefined CoA Transferase (1, Cat1 and Cat2), succinyl-CoA reductase (2, SucD) and 4-hydroxybutyrate dehydrogenase (3, 4Hbd) mediate the the conversion of succinate to 4-Hydroxybutyrate. The coassimilation steps of ethanol to balance reduction equivalents are not shown (A). The Reactions of the CETCH cycle (B) [3]; succinic semialdehyde reductase (4), 4-hydroxybutytyl-CoA synthetase (5), 4-hydroxybutytyl-CoA dehydratase (6), enoyl-CoA carboxylase/reductase (7 + 13) ethylmalonyl-CoA mutase (8),methylsuccinyl-CoA oxidase (9), mesaconyl-C1-CoA hydratase (10), β-methylmalyl-CoA lyase (11), propionyl-CoA oxidase (12). The reaction of SucD is highlighted in purple. The alternative substrate mesaconyl-C1-CoA is highlighted in teal.

vitro [3], optimized through rational and machine-learning approaches [24] and connected with downstream biosynthetic modules to produce different value-added compounds like polyketides and terpenes directly from CO_2 [53]. Similarly, the THETA cycle also involves 17 enzymes that together form new-to-nature CO_2-fixation pathway, which produces the central precursor acetyl-CoA as output molecule [52].

Key to the successful realization of new-to-nature pathways is the selection and/or design of suitable catalysts with exquisite substrate specificity and little promiscuity. Especially for the construction of complex new-to-nature networks that share structurally similar metabolites, such as the CETCH or THETA cycle, high substrate

specificity is essential to avoid draining of metabolites and accumulation of dead-end products, which ultimately lead to stalling of the metabolic network.
In this study, we assessed the substrate specificity of SucD from *C. kluyveri* (CkSucD) for its use in synthetic biology. We show that CkSucD selects against many CoA-esters, but has a significant side-activity with mesaconyl-C1-CoA, which is an important intermediate of both, the CETCH and the THETA cycle. We solved the crystal structure of SucD with mesaconyl-C1-CoA to identify amino acids conferring substrate binding at the active site. We further targeted these residues to improve substrate specificity of CkSucD and transferred the best mutation into the scaffold of SucD from *C. difficile* (CdSucD) to obtain a highly specific SucD variant.

2.3 Material and Methods

2.3.1 Plasmid generation

The plasmid containing His-tagged SucD of *C. kluyveri* [[3]] and *C. difficile* [5] was used as described in previous studies. Plasmids with point mutations were generated by side directed mutagenesis with a single primers depicted in Table 1 [54].

Table 1: Primers used in this study.

Primer	**Sequence**
PCC_E_seq1	CTTATGCGACTCCTGCATTAGG
pDuet_primus_rev	CGATTATGCGGCCGTGTACAATACG
SucD_Ck_K66R	GTTTATGAAGATAAAGTAGCTAGATGTCATTTGAAATCAGGAGC
SucD_Ck_K70R	GTTTATGAAGATAGAGTAGCTAAATGTCATTTGAAATCAGGAGC
SucD_Ck_K66R_K70R	GTTTATGAAGATAGAGTAGCTAGATGTCATTTGAAATCAGGAGC
SucD_Ck_S243N	CAATGGAATTATATGTAATTCAGAGCAATCAGTTATAGCTCCTGC
SucD_Ck_T112F	GCTACTACGCCTATATTTAATCCAGTGGTAACTC
SucD_Cd_K79R	GAAAAACAAGTCTAGGGCGAAGGTGATC

2.3.2 Synthesis of CoA esters

CoA ester synthesis was performed according to published protocols [55,56].

2.3.3 Gene expression and protein purification

Chemical competent E. coli BL21 DE3 cells were transformed with expression plasmids and selected for on LB agar plates using the respective antibiotics. For protein production, cells containing the plasmids were cultivated in 1 L salt-buffered TB medium at 100 r.p.m and a temperature of 37 °C until the culture reached an OD600 of 0.8. Afterwards, flasks were transferred to a shaker at 18 °C, induced with 0.25 mM IPTG and grown overnight. Cells were harvested for 10 min at 4 °C and 5,000 × g. The pellet was re-suspended in a 3-fold volume (3 mL per 1 g of cells) in lysis buffer (50 mM HEPES/KOH pH 7.8, 500 mM NaCl, 10% glycerol) and 10 μg/ml DNAse and 5 mM MgCl2 was added. Cells were lysed by ultra-sonication. The membrane fractions were removed by centrifugation at 18,000 × g and 4°C for 1 hour. The lysate was filtered through a 0.45 μm syringe filter before loading on a 1 ml HisTrap FF (pre equilibrated with lysis buffer) column attached to an Äkta Start (both from GE Healthcare, Freiburg, Germany). Unspecific bound proteins were washed off using lysis buffer with 75 mM imidazole. The bound protein was removed from the column using lysis buffer with 500 mM imidazole and collected in 1 ml fractions.

Batches for kinetic characterization were desalted using 2x5 ml HiTrap desalting columns (GE Healthcare, Freiburg, Germany) equilibrated with desalting buffer (50 mM HEPES/KOH pH 7.8, 200 mM NaCl, 10% glycerol). The fractions containing the protein of interest were pooled and concentrated using Amicon Ultra-4 centrifugal filters (Merck Millipore, Darmstadt, Germany). For storage, glycerol was added to a final concentration of 20% and aliquots were frozen in liquid nitrogen before storing them on -80°C.

Batches for crystallization were further purified using size exclusion chromatography.

Therefore a HiLoad 16/600 Superdex 200 pg attached to an Äkta Pure (both from GE Healthcare, Freiburg, Germany) was equilibrated with desalting buffer containing 20 mM TRIS HCl pH 7.8, 50 mM NaCl. The fractions containing the protein were pooled and concentrated in Amicon Ultra-15 centrifugal filters (Merck Millipore, Darmstadt, Germany), which were washed first to remove glycerol from the membranes.

2.3.4 Spectrophotometric assays

Reduction rates of CoA thioesters were determined by spectrophotometric monitoring of NADPH oxidation. A reaction mixture of 300 µL (200 mM HEPES pH 7.5, 400 µM NADPH, 400 nM CkSucD) was incubated at 30 °C for one minute. Varying amounts of CoA thioesters were added to the mixture to start the reaction. NADPH oxidation was monitored at a wavelength of 365 nm using an extinction coefficient of 3 300 M^{-1} cm^{-1}.

2.3.5 Crystallization and Structure Determination

The purified protein solution was spotted in different concentrations (5 and 10 mg/mL) on sitting-drop vapor-diffusion crystallization plates. 0.2 µL of each protein solution were mixed with 0.2 µL of crystallization condition. (For PDB 8CEI – 25% PEE propoxylate, 100 mM MES pH 6.5, 50 mM MgCl2; For PDB 8CEK – 45% PEE propoxylate, 100 mM MES pH 6.5, 400 mM KCl, 5 mM NADPH; For PDB 8CEJ – 100 mM magnesium acetate, 100 mM MOPS pH 7.5, 5 mM mesaconyl-C1-CoA, 12% PEG 8000). The drops equilibrated against 30 µL of protein free crystallization condition at 288 K. Crystals formed within 24 – 48 h. Cryoprotectant was added to the crystals to exceed a final concentration of 40% v/v (PEG 200 for PDB 8CEJ and PEE propoxylate for PDB 8CEI and 8CEK respectively). For crystals of PDB 8CEJ, mesaconyl-C1-CoA was added to 5 mM, additionally. All crystals were looped and frozen in liquid nitrogen. X-ray diffraction data were collected at the beamline P13 of the Deutsches

Elektronen-Synchrotron (DESY). The data sets were processed with the XDS software package [57]. The structures were solved by molecular replacement using a poly alanine search model of a propable aldehyde dehydrogenase from Listeria monocytogenes (PDB ID 3K9D). Molecular replacement was carried out using Phaser of the Phenix software package [58] and refined with Phenix.Refine. Additional modeling, manual refining and ligand fitting was done in COOT [59]. Final positional and B-factor refinements, as well as water-picking for the structure were performed using Phenix.Refine. The structure models were deposited at the PDB in Europe under PDB IDs 8CEJ, 8CEI and 8CEK. Data collection and refinement statistics are provided in Table 2.

Table 2: Data and Refinement Statistics for the Mct Crystal Structures. Statistics for the highest-resolution shell are in parentheses.

PDB ID	8CEI	8CEK	8CEJ
ligands	-	NADPH	mesaconyl-C1- CoA, mesaconate
resolution range (Å)	39.5 - 2.2 (2.3 - 2.2)	29.7 - 2.1 (2.2 - 2.1)	24.6 - 2.1 (2.2 - 2.1)
space group	P 1 21 1	I 2 2 2	I 2 2 2
unit cell dimensions			
a, b, c (Å)	86.2 89.3 137.3	140.0 190.8 190.9	141.3 189.7 189.5
α, β, γ (deg)	90.0 104.6 90.0	90.0 90.0 90.0	90.0 90.0 90.0
total reflections	200863 (19650)	276277 (27303)	294931 (29294)
unique reflections	102055 (10025)	138152 (13652)	147477 (14647)
multiplicity	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)
completeness (%)	99.19 (98.07)	99.85 (99.89)	99.83 (99.90)
mean I/σ(I)	7.28 (2.57)	9.74 (1.28)	10.86 (3.98)
R-merge	0.04911 (0.2416)	0.04709 (0.5583)	0.03242 (0.1402)
R-pim	0.04911 (0.2416)	0.04709 (0.5583)	0.03242 (0.1402)
CC1/2	0.997 (0.949)	0.999 (0.718)	0.998 (0.955)
reflections used in refinement	101689 (10008)	138025 (13637)	147363 (14634)
R-work	0.2424 (0.3033)	0.1980 (0.2902)	0.2949 (0.3344)
R-free	0.2670 (0.3073)	0.2221 (0.3261)	0.3095 (0.3281)

Table 2: Data and Refinement Statistics for the Mct Crystal Structures. Statistics for the highest-resolution shell are in parentheses.

PDB ID	8CEI	8CEK	8CEJ
number of non- hydrogen atoms	14600	14888	15096
macromolecules	13576	13630	13666
ligands	0	192	42
solvent	1024	1066	1388
protein residues	1788	1795	1794
RMS(bonds)	0.002	0.004	0.004
RMS(angles)	0.40	0.58	0.43
Ramachandran			
favored (%)	98.31	98.15	97.08
allowed (%)	1.52	1.62	2.59
outliers (%)	0.17	0.22	0.34
Rotamer outliers (%)	0.83	0.62	0.48
Clashscore	2.71	2.41	2.00
Average B-factor	33.59	47.07	29.29
macromolecules	33.40	46.70	29.05
ligands	36.12	67.05	27.29
solvent	1	48.11	31.76
Twin fraction (law)	-	-	0.4 (-h, -l, -k)

2.3.6 Structural Modelling of CdSucD and CkSucD Mutants

A structure model of CdSucD was generated using the software package SWISS-MODEL (www. swissmodel.expasy.org) [60–62], by providing PDB 8CEJ and the respective sequence files.

2.4 Results

2.4.1 SucD from *C. kluyveri* is promiscuous with mesaconyl-C1-CoA

To investigate the substrate specificity of SucD, we determined the activity of CkSucD with its native substrate, succinyl-CoA, and different alternative CoA esters (Table 3). CkSucD displayed a k_{cat}/K_M of 4.1 × 105 $M^{-1}s^{-1}$ for succinyl-CoA and 2% or less catalytic efficiency for most other CoA esters. One notable exception was mesaconyl-C1-CoA, which showed a catalytic efficiency of 6.7 × 104 $M^{-1}s^{-1}$, corresponding to 16% of its native reaction with succinyl-CoA. Overall, this data indicated that the use of CkSucD could be problematic with pathways featuring mesaconyl-C1-CoA as metabolite, such as the CETCH or THETA cycle.

Table 3: Kinetic parameters of SucD variants. "±" indicates SE. Fits of Michaelis Menten kinetic are shown in Figure 2.4 and Figure 2.7.

Enzyme	Substrate	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [M^{-1} s^{-1}]	
CkSucD	Succinyl-CoA	7.8 ± 0.43	22.6 ± 4.9	3.5×10^5	} 16%
	Mesaconyl-C1-CoA	1.6 ± 0.07	28.5 ± 4.7	5.6×10^4	
	Mesaconyl-C4-CoA	0.2 ± 0.02	25.7 ± 6.8	8.2×10^3	
	Propionyl-CoA	0.1 ± 0.04	145.2 ± 144	6.9×10^2	
	Crotonyl-CoA	0.3 ± 0.03	332.6 ± 54.8	9.0×10^2	
	4OH-Butyryl-CoA	0.2 ± 0.01	48.9 ± 10.8	4.1×10^3	
CkSucD K70R	Succinyl-CoA	0.9 ± 0.03	19.9 ± 2.3	4.7×10^4	} 2%
	Mesaconyl-C1-CoA	0.1 ± 0.01	86.1 ± 15.1	1.1×10^3	
CdSucD	Succinyl-CoA	28.7 ± 4.1	213.8 ± 49.4	1.3×10^5	} 12%
	Mesaconyl-C1-CoA	0.65 ± 0.08	38.3 ± 14.9	1.6×10^4	
CdSucD K79R	Succinyl-CoA	7.3 ± 0.51	38.6 ± 7.4	1.9×10^5	} 2%
	Mesaconyl-C1-CoA	0.5 ± 0.02	144.3 ± 12.3	3.7×10^3	

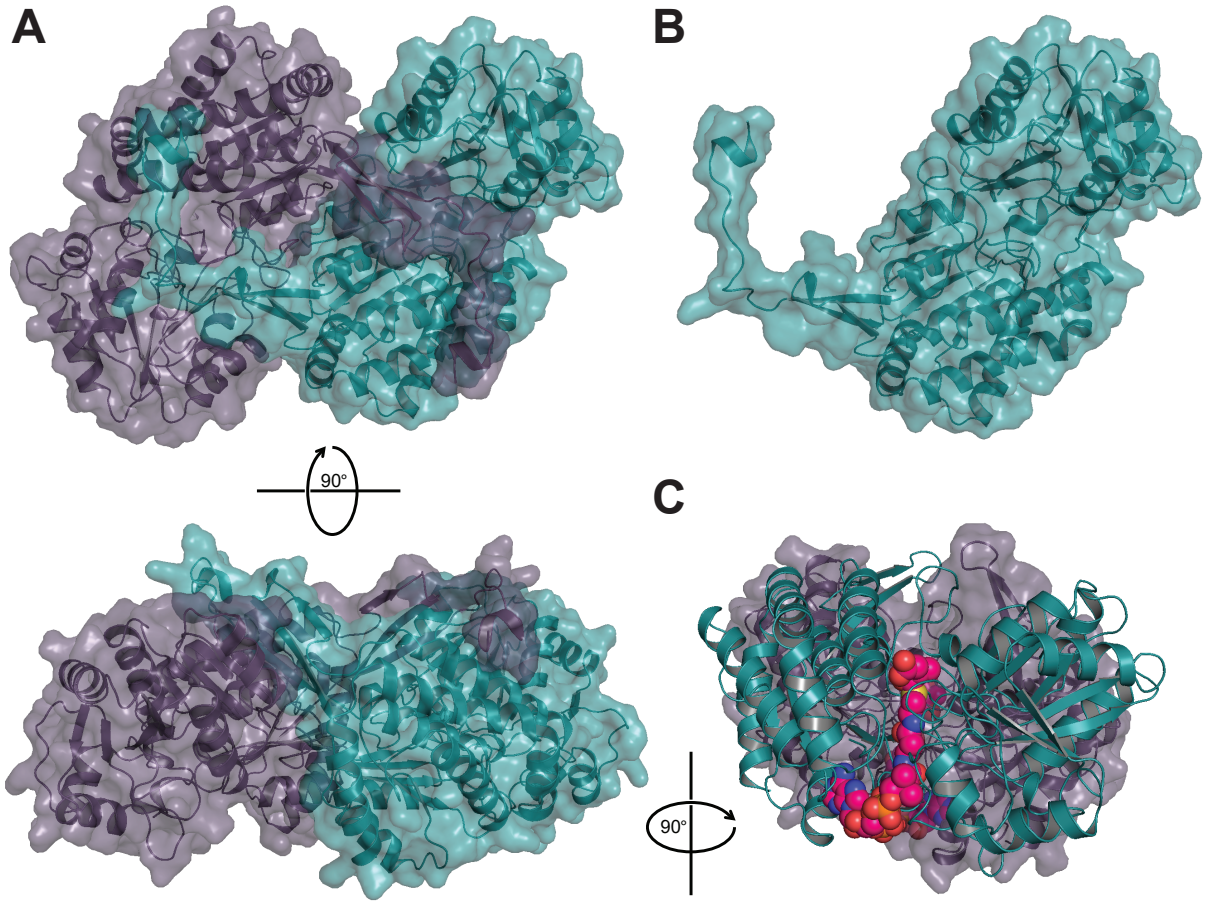


Figure 2.2: Structure of CkSucD with bound ligands; CkSucD (pdb 8CEI) forms a homodimer (A) with an overlapping C-terminal loop (B). CoA ester substrates cross the whole subunit (pdb 8CEJ) to reach towards the active site. Mesaconyl-C1-CoA is shown in pink (C).

2.4.2 Crystal Structure of CkSucD Identifies Molecular Basis for Mesaconyl-C1-CoA Binding

To understand the basis of substrate specificity in CkSucD, we solved the crystal structure of the enzyme in complex with mesaconyl-C1-CoA. CkSucD forms a homodimeric complex (Figure 2.2A). Each monomer (Figure 2.2B) has an extended C-terminal loop that reaches into the second subunit within the complex. CoA ester ligands are coordinated within a tubular cavity that reaches to the surface bound active site on the other side of the monomer (Fig 2C). The reaction mechanism of CkSucD is likely analogous to that of propionaldehyde dehydrogenase (PduP) [51]. An active site cysteine (Cys242) plays a key role by forming a covalent bond with

the acyl moiety of the CoA ester and releasing the CoA moiety (Figure 2.3B). Active site residues His364 and Thr365 assist in proton donation to the released CoA (Figure 2.3C). In our structure, we observed mesaconyl-C1-CoA coordinated at the active site (Figure 2.3A) with a occupancy of 75% (pdb 8CEJ,C). However, in the rest of the subunits, the occupancy could be rather assigned to a mesaconyl-cysteine (pdb 8CEJ, A, B, D), indicating that we also trapped the covalent intermediate of the reaction mechanism in part of our crystals. The secondary structure of CkSucD shares high similarity with and PduP from Rhodopseudomonas palustris (pdb 5JFN, Figure 2.6A, RMSD 1.031 over 302 residues) and is – at 32% identities – the closest protein structure with a trapped intermediate [7, 22]. The active site of propionaldehyde dehydrogenases is identical in respect to the mechanistically relevant residues, but differs in the active site surroundings (Figure 2.6B). Compared to PduP, residues that restrict the active site pocket and coordinate the acyl moiety of propionyl-CoA (Leu158, Leu483 and Val331) are absent in CkSucD (aligned residues in CkSucD are Lys70, Thr395 and Ser243), which creates a pocket that is differently shaped and more spacious [4, 7] (Figure 2.6A). Mesaconyl-C1-CoA is coordinated through Ser243, which is located next to the catalytically active His242 and forms a hydrogen bond to the carboxy group of mesaconyl-C1-CoA. The terminal carboxyl group of mesaconyl-C1-CoA is further coordinated through a hydrogen bond to Lys70 (Figure 2.3C).

2.4.3 Active Site Mutagenesis to Increase Selectivity of CkSucD Against Mesaconyl-C1-CoA

Based on our structures, we identified several residues that we targeted to increase selectivity of CkSucD against mesaconyl-C1-CoA. To discriminate sterically and electronically against the methyl group of mesaconyl-C1-CoA, we replaced Lys70 by a bulkier, positively charged arginine. For the same reason, we also exchanged Lys66, which is located on the same α-helix as Lys70, by an arginine. We also mutated Ser243 to an asparagine to allow for hydrogen bonding to the mesaconyl-C1-CoA car-

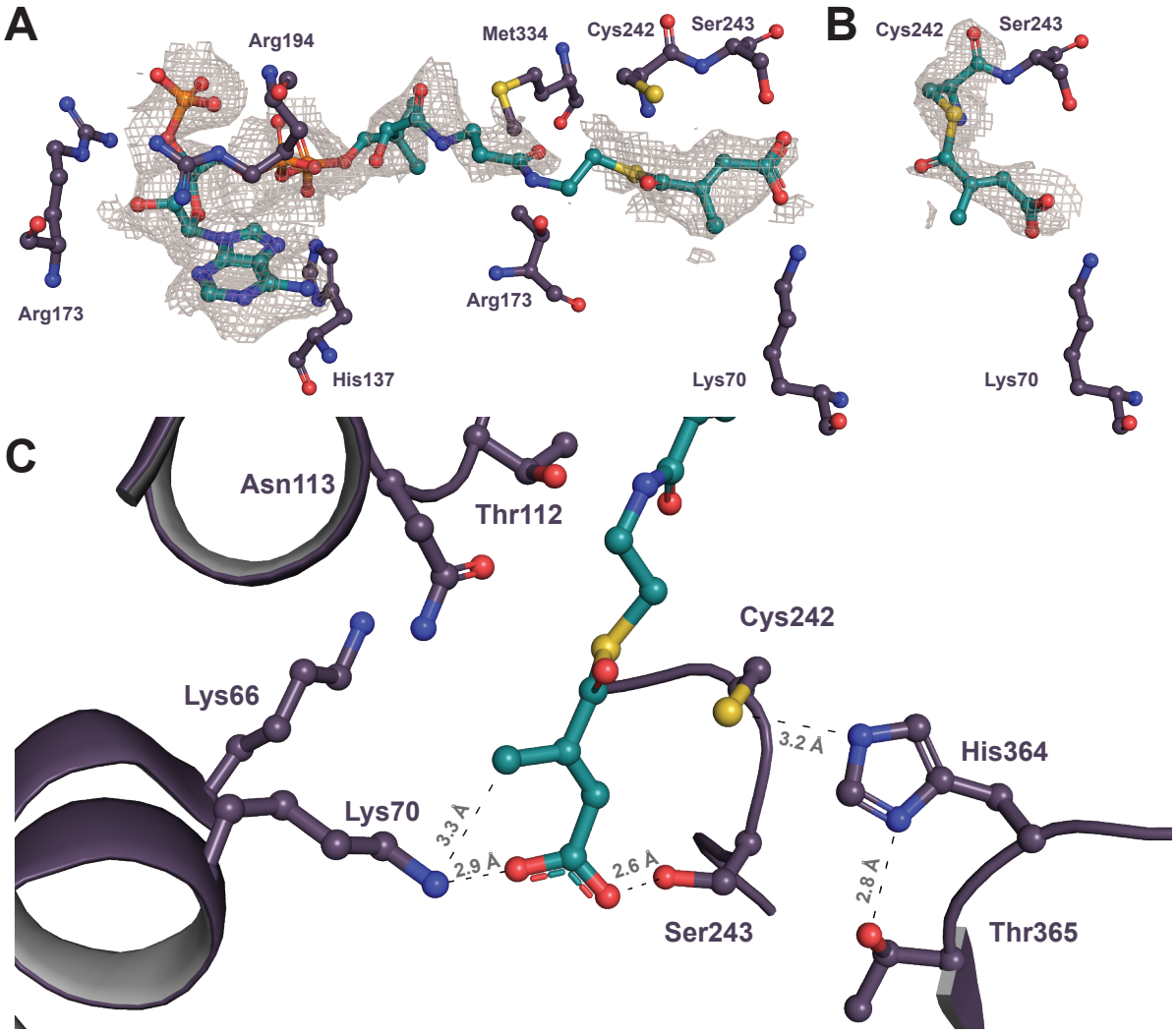


Figure 2.3: The active site of CkSucD. The acyl moiety of mesaconyl-C1-CoA is coordinated by Lys70 and Ser243 (A). Cys242 is forming a mesaconyl-cystein intermediate (B). Thr365 and His364 donate a proton to form the product (C). Peptide residues are presented in violet, mesaconylated intermediates are presented in teal.

boxy group, while increasing steric constraints against the methyl group. Finally, we introduced a phenylalanine at the position of peripheral Thr112, which coordinates the amide group of the cysteamine in the CoA moiety and controls access to the active site, with the idea to restrict mesaconyl-CoA accommodation. Most single mutations were soluble, but showed only residual or non-detectable activity with succinyl-CoA (Table 4). Only variants K70R and K66R showed relevant turnover rates, albeit at one tenth and one twentieth of wildtype activity, respectively. Double variants K70R_K66R and K70R_S243N did result in insoluble of non-active protein, which left us with the

K70R mutant, as this variant also had shown some improved specificity (8% relative activity of mesaconyl-C1-CoA to succinyl-CoA reduction) compared to wildtype (16%) and the K66R variant (34%) in our screen.

Table 4: Specific activities ($\underline{s}$) of different CkSucD mutants. "±" indicates SE and "n.d." not detectable.

		Specific activity [s^{-1}]		
mutation	solubility	succinyl-CoA	mesaconyl-C1-CoA	Side activity [%]
wildtype	+	7.8 ± 0.43	1.6 ± 0.07	20.5
K66R	+	0.48 ± 0.02	0.164 ± 0.01	34.2
K70R	+	0.84 ± 0.014	0.072 ± 0.01	8.6
S243N	+	n.d.	n.d.	-
T112F	+	0.10 ± 0.001	0.03 ± 0.003	30
K66R_K70R	-	-	-	-
K70R_S243N	+	n.d.	n.d.	-

2.4.4 CkSucD K70R Shows Increased Selectivity, Albeit at Ten-fold Reduced Catalytic Efficiency

In the following, we characterized CkSucD wildtype and CkSucD K70R in more detail (Table 3). The catalytic activity of active site variant CkSucD_K70R for succinyl-CoA was reduced by one order of magnitude (from ∼4 ×105 to ∼5 ×104 M^{-1} s^{-1}). However, the catalytic efficiency for mesaconyl-C1-CoA reduction had dropped more than fifty-fold (from ∼6 ×104 to ∼1 ×103 M^{-1} s^{-1}). This was caused by an 3-fold increased apparent K_M values of for mesaconyl-C1-CoA (from ∼30 to ∼90 μM), while the drop in specific activity contributed by roughly a factor of two. Together, these factors decreased the relative catalytic efficiency of K70R with mesaconyl-C1-CoA from 16% to 2%; yet, this specificity increase came at a ten-fold decreased catalytic efficiency for the original substrate.

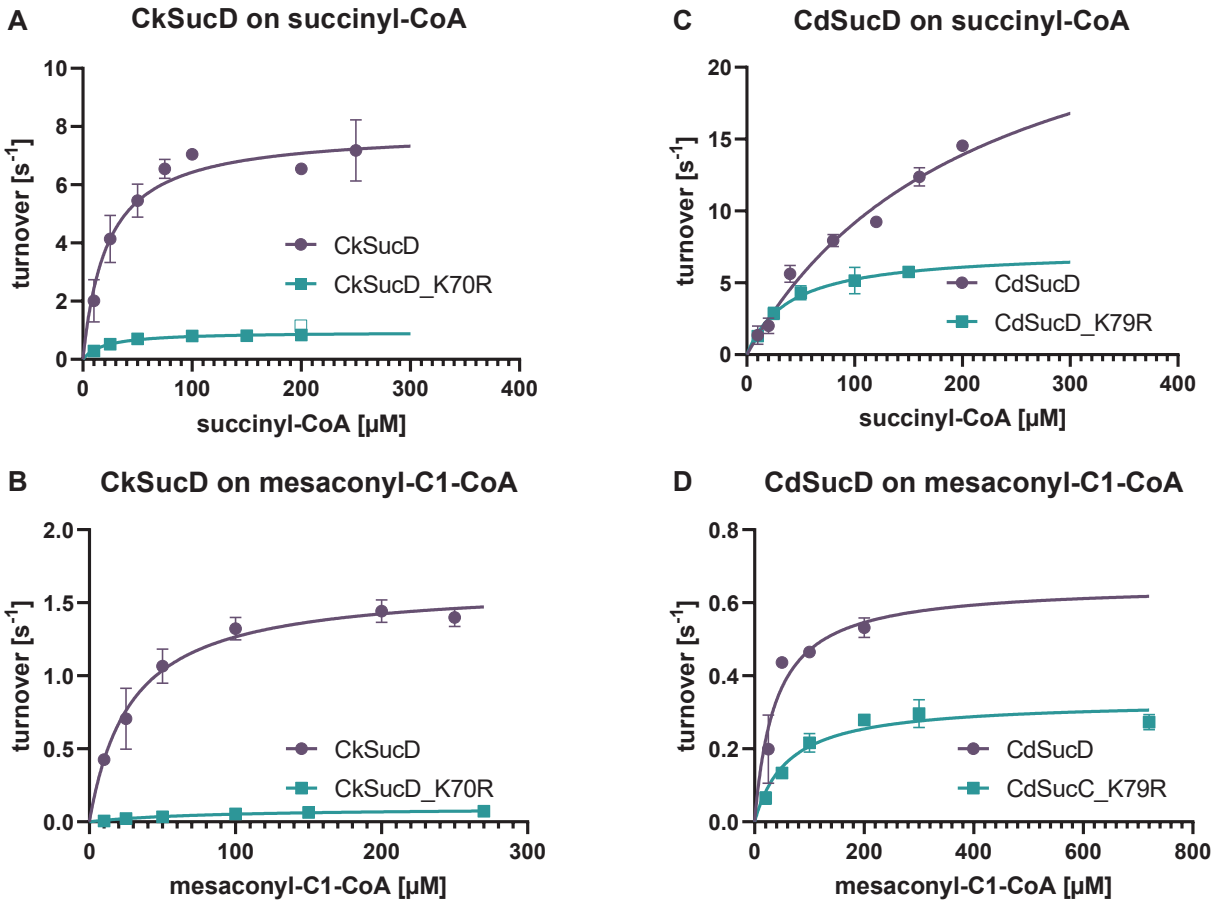


Figure 2.4: Michaelis Menten kinetics of sucD variants. Activities of CkSucD and the K70R variant on succinyl-CoA and mesaconyl-C1-CoA (A+B). Activities of CdSucD and the K79R variant on both substrates (C+D).

2.4.5 *C. difficile* SucD KI70R Shows Increased Selectivity At a High Catalytic Efficiency

We also tested the effects of the K70R mutation could be transferred to other homologs. To that end, we introduced the equivalent substitution (K79R) into SucD from *C. difficile* (CdSucD). Notably, this mutation did not negatively affect the catalytic efficiency of the reaction with succinyl-CoA (catalytic efficiency was actually slightly increased), while the catalytic efficiency with mesaconyl-C1-CoA dropped more than fourfold. This was mainly caused by a 3-fold increased apparent K_M. for mesaconyl-C1-CoA. Overall, the K70R mutation reduced the relative catalytic efficiency with mesaconyl-C1-CoA to about 2%, while the catalytic efficiency for the native substrate

remained virtually unchanged.

2.5 Discussion

In this study, we investigated the substrate specificity of SucD, an essential enzyme in ethanol-succinate fermentation and a key enzyme in several new-to-nature CO_2 fixation pathways that were developed recently. We show that the enzyme is an efficient succinyl-CoA reductase with a k_{cat} of 7.8 ± 0.43 s^{-1}, but also possesses a significant side activity with mesaconyl-CoA at 16% specific activity. To understand the molecular basis of this promiscuity, we solved the crystal structure without ligands (pdb 8CEI), as well as in the NADPH- (pdb 8CEK) and mesaconyl-C1 CoA bound state (pdb 8CEJ, which also includes a catalytically trapped mesaconate-cysteine intermediate). Our structures at a resolution between 2.1 and 2.2 Å, (re-)confirm catalytically active residues and residues necessary for the coordination and binding of mesaconyl-C1-CoA. Ser243 and Lys70, which we subsequently targeted for site-directed mutagenesis, coordinate the distal carboxy group of mesaconyl-C1-CoA. To engineer substrate specificity of CkSucD, we created different active site mutants, of which K70R decreased the relative catalytic efficiency with mesaconyl-C1-CoA from 16 to 2%, yet, this mutation also decreased the overall catalytic efficiency for succinyl-CoA by ten-fold. When transferring this mutation into the closely related homolog CdSucD, relative catalytic efficiency dropped again to 2%, notably, however, without affecting catalytic activity with succinyl-CoA, yielding a highly specific, yet highly active enzyme. Note that this 2% side reactivity represents a upper limit, as the selectivity might even increase, especially in situations, in which the enzyme faces low concentrations of the respective CoA esters, as the apparent K_M for mesaconyl-C1-CoA (∼145 mM) is considerably higher than for succinyl-CoA (∼ 40 mM).

Overall, the CdSucD K70R variant created in this study is a highly specific succinyl-CoA reductase for future use in the construction and operation of new-to-nature path-

ways, such as the CETCH or the THETA cycle, as well as other metabolic networks featuring mesaconyl-C1-CoA as a metabolite.

2.6 Acknowledgements

The authors thank Dr. Jan Zarzycki for great stimulating scientific discussions. The authors thank DESY (Hamburg, Germany) for the use of the beamline P13, and especially Johanna Hakanpää and Isabel Bento for assistance with beamline operation.

2.6.1 Author contributions

P.P., C.D. and T.J.E. conceived the project. P.P., M.C., C.D. and E.H. performed kinetic characterization. C.D., E.H. and P.P generated the mutant library. C.D. and P.P. generated protein crystals. P.P. analyzed structural x-ray data. The manuscript was written by P.P., C.D. and T.J.E. with contribution from all other authors.

2.6.2 Accession Codes

PDB-IDs for CkSucD are 8CEI (without ligands), 8CEK (with NADPH) and 8CEJ (with mesaconyl-C1-CoA).
Uniprot IDs are P38947 for CkSucD and A0A031WJ42 for CdSucD.

2.7 Supporting Information

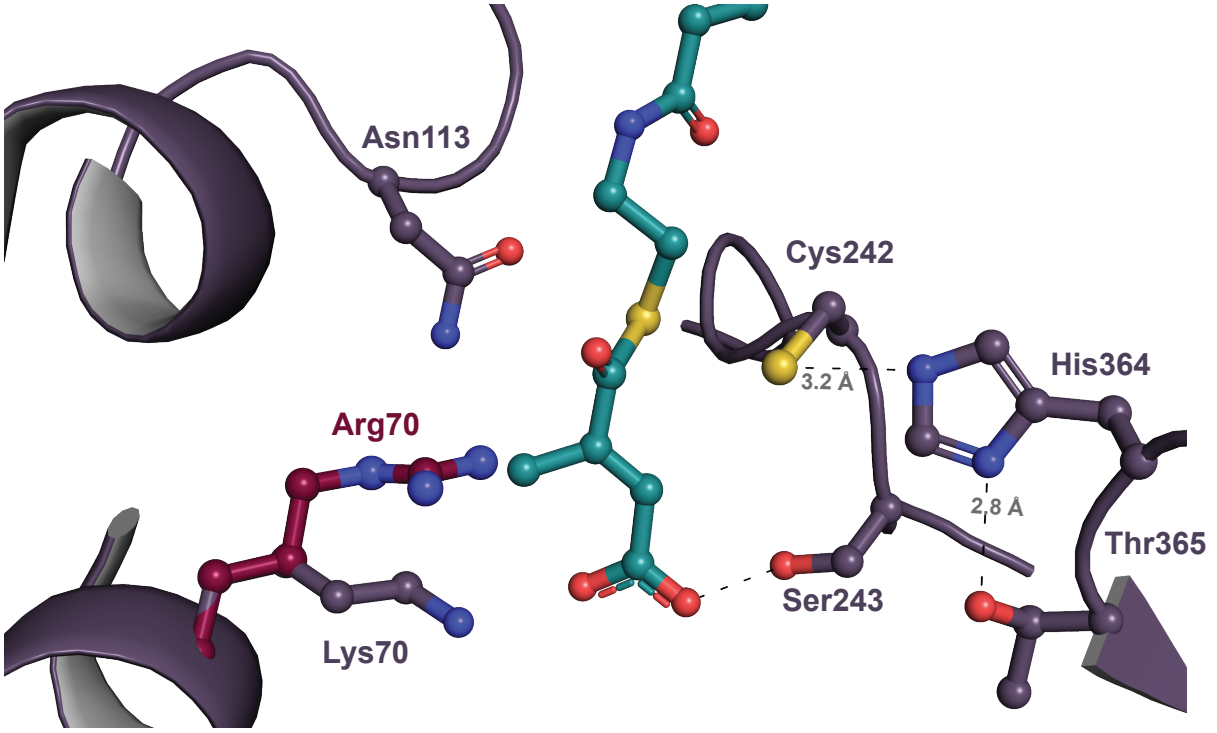


Figure 2.5: CkSucD active site with K70R mutation. Peptide residues are presented in violet, mesaconylated intermediates are presented in teal. K70R is colored in red.

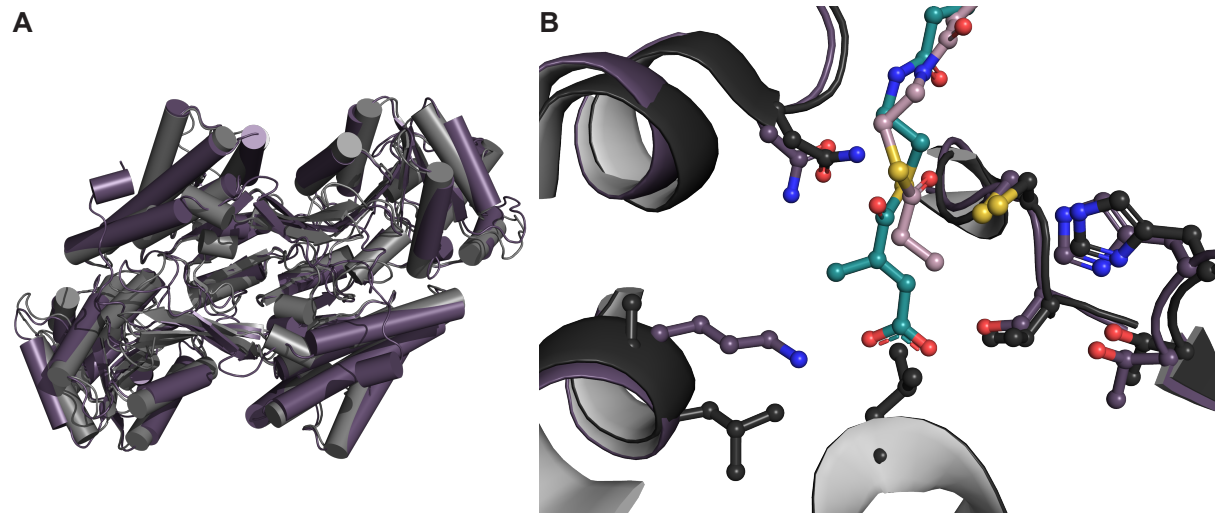


Figure 2.6: Comparison of PduP and CkSucD. Secondary structure comparison of CkSucD (violet) and PduP (gray). RMSD value is 1.05 over 302 residues (A). Active site of PduP has different spatial elements. Mesaconyl-C1-CoA (teal) as coordinated in CkSucD would be occluded in PduP by Leu483 (B).

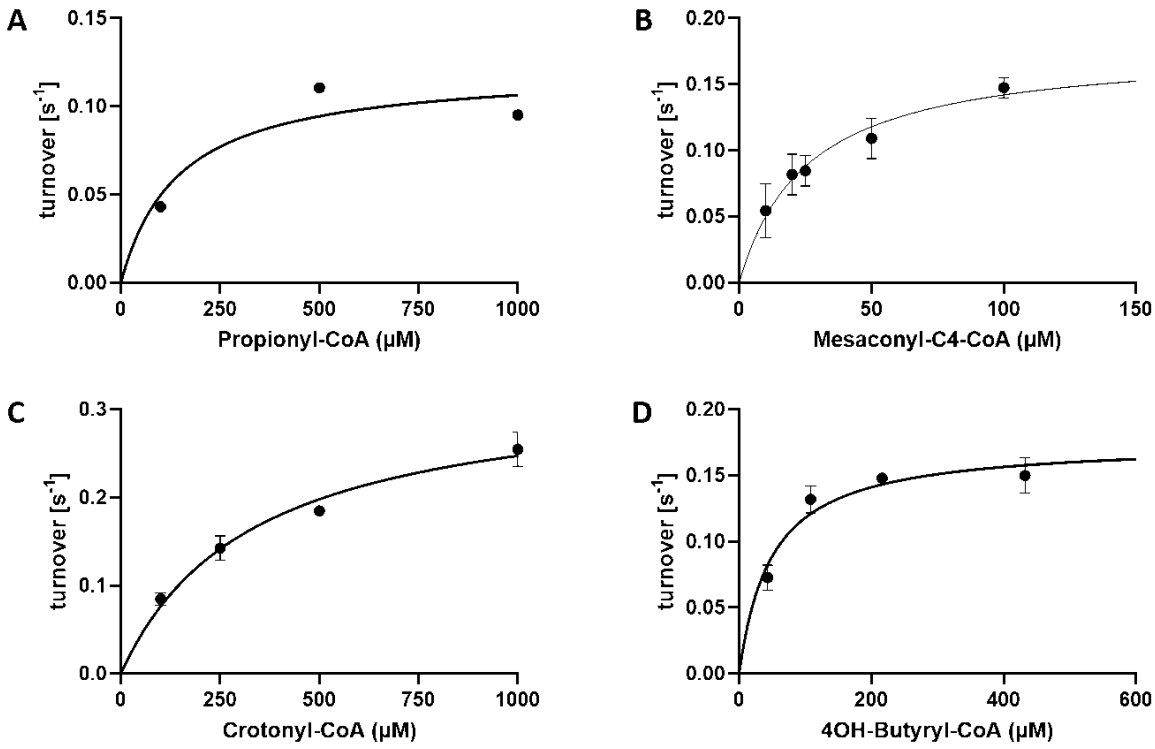


Figure 2.7: Michaelis Menten kinetics of CkSucD for different acyl-CoAs: Propionyl-CoA (A), Mesaconyl-C4-CoA (B), Crotonyl-CoA (C), 4OH-Butyryl-CoA.

Chapter 3

Efficient Propionate mediated Photorespiration in *S. elongatus* PCC 7942

3.1 Abstract

RuBisCO as the key enzyme of the CBB cycle is the most abundant enzyme on earth [7,8]. Its efficiency, however, is held back by its unfavorable oxygenase sidereaction with dioxygen. The resulting 2PG needs a whole metabolic machinery (photorespiration), that even releases CO_2 and ammonium to be recycled back into usable compounds. Synthetic pathways to alleviate this photorespiratory burden are focused on avoiding decarboxylation or the release of ammonia, CO_2 or even both [25]. It has been proven beneficial, if the decarboxylation is relocated into the chloroplast so its level rise in proximity of RuBisCO [17]. The 3OHP bypass is a nature inspired synthetic photorespiratory pathway, that incorporates an additional carboxylation step per oxygenation by RuBisCO [1,18]. So far its potential to avoid photorespiratory carbon drain has been shown theoretically. Here, we introduced the first reaction steps of the bypass into the cyanobacterium *S. elongatus* PCC 7942. We see a growth benefit by feeding propionate, the substrate photorespiratory gloxylate is coassimiliated through the introduced pathway. The resulting strain was confronted with continiously increasing photorespiratory burden for 125 days to evolve and adjust to this selective condition. The slowly increasing growth rate shows, that the 3OHP bypass has potential to replace natural photorespiration for the better.

3.2 Introduction

Due to the promiscous nature of the RuBisCO reaction, all organisms fixing CO_2 via the CBB cycle have a way to remove toxic 2PG from its metabolome [13–15]. The recycling of 2PG via the C_2-cycle (see Figure 1.3, blue circle) into CBB cycle precursors is more energy and carbon sustainable than excretion or the complete oxidation. However, this is still requiring additional energy equivalents and releases one ammonium and CO_2 per two molecules of 2PG. The liberated carbon and ammonium have to be re-assimilated in additional steps.

It was proposed that photorespiration takes place in more than 20% of the RuBisCO

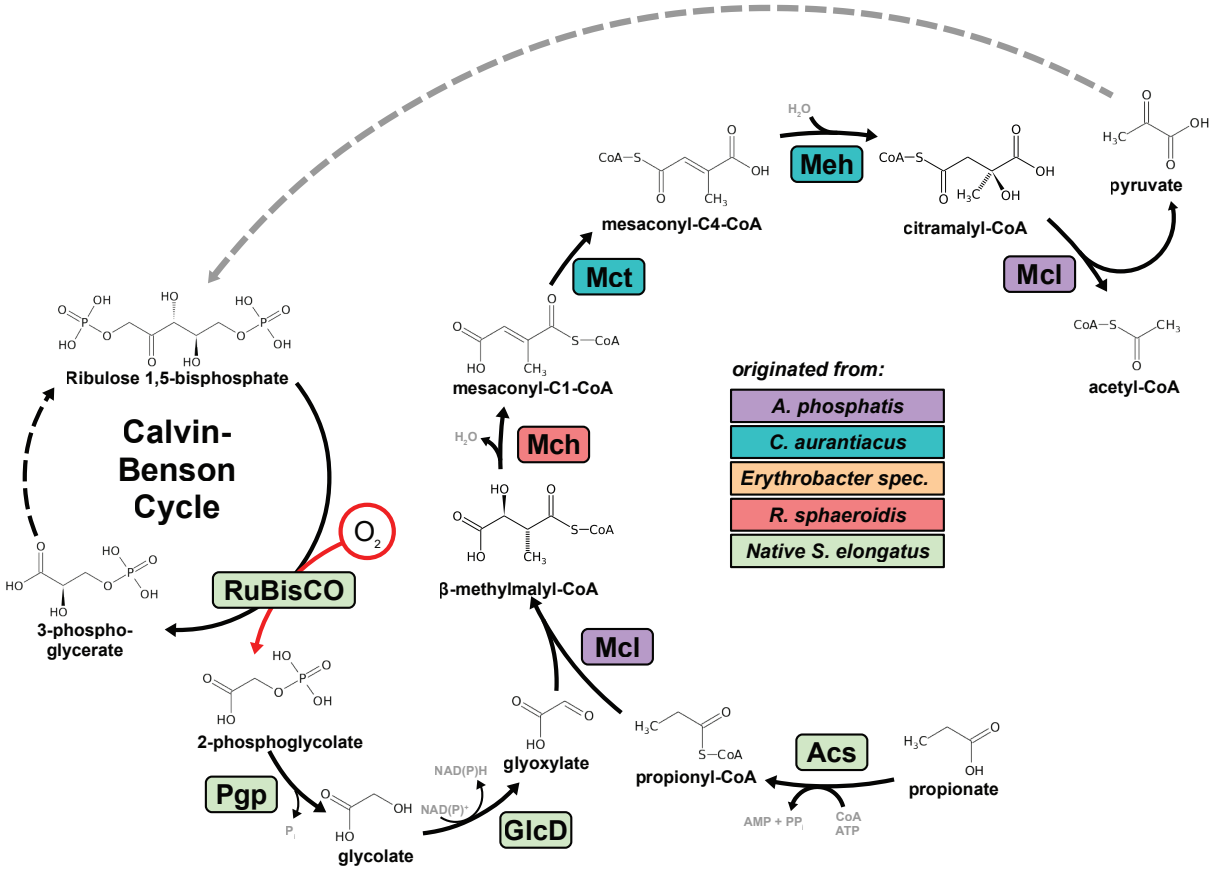


Figure 3.1: **Mcl-route of the 3OHP bypass for the assimilation of propionate and photorespiratory glyoxylate.** Acetyl-CoA synthase (Acs), β-methylmalyl-CoA/(S)-citramalyl-CoA lyase (Mcl), mesaconyl-C1-CoA hydratase (Mch), mesaconyl-C1-C4-CoA CoA transferase (Mct), mesaconyl-C4-CoA hydratase (Meh). The colors of the enzymes represents the organism. Green enzymes are native to *S. elongatus* PCC7942.

reactions in C_3 plants [63], setting back their biomass production by a notable margin. Probably due to the gradual manner of oxygen accumulation in our atmosphere, evolutionary adaptation favored the exclusion of oxygen from the reaction centers of RuBisCO with so called CCMs, like carboxysomes in cyanobacteria, pyrenoids in the chloroplast or the evolution of C_4 plants [16,29,64].

Recent discoveries show, that a more efficient strategy of the assimilation of glyoxlate derived from 2PG takes place in organisms that are living in close proximity with cyanobacteria and algae [25,26]. One of them is the 3OHP bi-cycle [18], present in the thermophilic green non-sulfur bacterium *C. aurantiacus*. It was first discovered in hot

microbial spring mats, where organic acids like glycolate or glyoxylate are excreted by cyanobacteria [65,66]. Although *C. aurantiacus* can assimilate CO_2 to gloxylate in the first part of the 3OHP bi-cycle, it preferably grows mixotrophically on small organic acids like glyoxylate entering metabolism via the second part of the bi-cycle (see Fig. 1.2) [26,67]. The 3OHP bi-cycle assimilates glyoxylate more efficiently, as it produces pyruvate from one molecule of glyoxylate by incorporation of an additional carboxylation step. That makes this pathway a favorable carbon positive photorespiratory pathway.

The hypothesis that *C. aurantiacus* is feeding on cyanobacterial excretions [26,65,66] implied the metabolic advantages of a cyanobacterium, that could utilze the second part of the 3OHP bicycle for the efficient removal of photorespiratory glyoxylate with an unimpaired biomass production due to an additional carboxylation reaction. The combination of those metabolic pathways was already attempted in a previous study [1]. There all the required genes from *C. aurantiacus* were introduced into the single cell freshwater cyanobacterium *S. elongatus* PCC7942 and identfied the native activity of Acc as a bottleneck for this pathway by cell extract activity assays. As bacteria should have a functioning Acc as it is an essential enzyme for fatty acid biosynthesis, absence of detectable Acc activity remained undisclosed.

The reaction of Acc consists of two catalytic reaction. A BCCP domain with a biotin covalently linked to an active site lysine is carboxylated by the biotin carboxylase (BC) domain. This step requires ATP. The carboxybiotin serves as carbon donor in carboxyl transferase (CT) domain, where the carboxygroup is transfered onto an acyl-CoA ester.

Acc is a complex multidomain enzyme, that undergoes post-translational modification as well as metabolic regulation. A dedicated enzyme, BirA, attaches the biotin covalently to the active site lysine. We hypothesized, that the production of a heterologous Acc should be supported with the corresponding BirA from the originating organism. Even if the native BirA activity was promiscious enough, the coexpression

of an additional copy should ensure proper biotinylation.

We tested the co-integration of both, other copies of Acc genes, as well as the gene encoding BirA. However, Acc activity was neither detectable with supplemented biotin nor without.

Cyanobacteria accumulate glycogen as storage for sugar molecules [30]. Even though some cyanobacteria produce polyhydroxyalkanoates (PHA) from acetyl-CoA, they would also heavily rely on glycogen. The metabolic route from glycogen to produce the essential CBB cycle intermediate ribulose-1,5-bisphosphate (RuBP) is shorter and less energy intensive than gluconeogenesis from polyhydroxybutyrate (PHB). As the tricarboxylic acid (TCA) cycle is rather an anabolic, than a katabolic pathway for cyanobacteria, the drainage of valuable sugar compounds into the lower glycolysis or even the TCA cycle is highly disadvantageous. A central regulator, P_{II} is responsible for many signalling cascades guarding these precious metabolites [34, 35]. It also interacts with the BCCP domain of Acc [36].

Cyanobacteria have an incomplete TCA cycle, with a few exceptions, that compensate the missing enzymes with alternative reactions (see Figure 1.1). However, flux through these alternative TCA reactions is usually very low and also rather serve anaplerotic functions [32]. Furthermore, the ability to feed on external substrates like fatty acids or sugars is a rare trait among cyanobacteria. In fact organic acids even inhibit growth. Acrylate, propionate and 3OHP are described as growth inhibitors for *S. elongatus* PCC 7942 [68]. The enzyme Acs of this strain has promiscuity for those other acids, and activates them to the respective acyl-CoA ester. This drains the CoA pool as *S. elongatus* has no way of metabolizing them. Even though exposure to these acids reportedly leads to mutations withing Acs that make it more specific [68], the toxicity can only completely be circumvented if those acyl-CoA esters can be metabolized. The 3OHP bypass has propionyl-CoA and 3OHP-CoA as intermediates and would allow the detoxification of propionate as 2PG is getting removed.

We propose that a *S. elongatus* strain without carboxysomes, expressing the necessary

genes from the 3OHP bypass is detoxifying propionate via this metabolic path. They were able to adapt to lower concentrations of CO_2 than a ΔccmK-O strain, as they metabolize gloxylate in a carbon positive manner.

3.3 Materials and Methods

3.3.1 CoA ester Synthesis and Purification

To measure enzyme activities and enzyme specificities acetyl-CoA and propionyl-CoA were synthesized according to an established protocol [56]. 80 mg of CoA were solved in 4 ml sodium bicarbonate buffer (10 mg/ml), then 20 µl acetic anhydride or 30 µl propionic anhydride were slowly added to the buffer on ice (100 µmol CoA : 200 µmol propionic/acetic anhydride). Complete conversion was tested by Ellmans reagent (50 mM MOPS pH 7.5, 10 mM DTNB, 10 mM EDTA). The reaction was adjusted to pH 3.0 with formic acid to stabilize the CoA thioester products. CoA esters were purified by high Performance Liquid Chromatography (HPLC) (buffer, 25 mM ammonium-formate pH 4.2; eluent, methanol). After purification, samples were frouen in liquid nitrogen, lyophilized and stored dry at $-20\,°C$

3.3.2 Media Preparation

BG-11 medium

Medium to cultivate *S. elongatus* PCC 7942 contained HEPES buffer and sodiumnitrate [69,70]. Five Stocksolutions were prepared and autoclaved indepedently (1 [100x] - 0.28 mM $MgNa_2$ EDTA, 3 mM ferric ammonium citrate, 3.12 mM citric acid, 1 mM $CaCl_2$; 2 [100x] - 30 mM $MgSO_4$; 3 [100x] - 33 mM K_2HPO_4; 4 [50x] - 1 M HEPES/KOH pH 8.0; 5 [1000x] - 46.3 mM H_3BO_3, 9.15 mM $MnCl_2$, 0.77 mM $ZnSO_4$, 0.32 mM $CuSO_4$, 1.61 mM Na_2MoO_4, 0.17 mM $CoCl_2$). All stocksolutions were mixed in appropriate amounts and 1.5 g/L sodium nitrate were added and the complete medium was

autoclaved again.

For plates 2x BG-11 was prepared as well as 2.4% Agar and both solutions were mixed in equal parts after autoclaving. The culture was incubated at 30 °C without light for 16 hours.

3.3.3 Culture Handling

Cultivation of *S. elongatus*

To sustain phototrophic growth, cultures were kept in light incubators(Minitron, Infors HT, Basel) at 30 °C with a constant CO_2 concentration of 3.0% and an light intensity of 100 μE m^{-2} s^{-1}. Liquid cultures of *S. elongatus* were grown in erlenmeyer-flasks with BG-11 medium or on BG-11 plate. After reaching stationary phase, cultures were diluted (1:1000) in fresh medium.

Plasmids

The plasmid pJZ148 created in a former study [38] was used to integrate the genes encoding Mcl from *A. phosphatis*, Mch from *R. sphaeroides* as well as Mct and Meh from *C. aurantiacus* as well as a spectinomycin resistance gene into the neutral site 1 of *S. elongatus* PCC7942.

Genome Integration

2 mL of a fresh culture of *S. elongatus* grown to an OD_{730} between 1 and 2 were centrifuged at 17,000 × g at room temperature. After discarding the supernatant, the cells were resuspended in 400 μL of BG-11 Medium. 500 to 1,000 ng of plasmid DNA were add to the cells. The culture was streaked on a sterile filter on BG-11 plates. After 24 hours of incubation at 30 °C, the filter was transfered to a BG-11 plate with 10 μg/mL spectinomycin.

3.3.4 Enzymatic Assays

For the test of enzyme activities in cell extracts, cells were lysed by sonication. After centrifugation for 1 h at 4 °C and 100 000 × g, the supernatant was filtered (0.45 µm pore size) and stored on ice. To run the assay 20 µL extract were used in a reaction mixture of 300µL (200 mM HEPES pH 7.2, 20 µL lysate, 1 mM propionyl-CoA, 3 mM glyoxylate, 3 mM $MgCl_2$). 10 µL samples were taken at 0, 20 60 and 120 minutes and quenched directly with 40 µL formic acid(>99%). Denatured protein was removed by centrifugation at 17 000 × g and 4 °C for 20 minutes and supernatant was used for HPLC analysis. The Chromatography was done with ammonium formate buffer (pH 4.2) using a reversed-phase C18 column (Gemini-NX 10u; 100 x 21.2 mm) and signal was detected at 260 nm. The compounds were eluted with a gradient from 5 to 95% of methanol over the time cause of 10 minutes.

3.3.5 Turbidostat Cultivation

Turbidostat Parameters

The turbidostat cultivation was conducted in a DasGIP Photobioreactor (Eppendorf, Jülich, Germany) at a working volume of 0.3 L at 37°C. The turbidostat was filled with BG-11, with 5 mM of propionate. The same medium was used for the Inflow medium. On days 7, 10, 20, 29, 45, and 107 of the evolution experiment, the CO_2 level was set to 0.4, 0.3, 0.2, 0.4, 0.3, 0.2 % in air, respectively. The turbidostat was constantly flushed with 60 sL/h. The targeted OD value was set from 1.5 initially to 1.5, 0.9 and 1.5 again on day 24, 30 and 42, respectively to compensate for stalled growth with fresh medium. The dilution rate was estimated by the derivation of the absolute feed pump volume as provided by the system. From day 103-106 a pump head overdrive due to empty feed bottles caused wrongly reported pump values, so the estimated dilution rate from day 100 - 110 were not considered accurate. After refilling the medium reservoir, the turbidostat was functional as before.

Calculating substrate concentration and growth rate

In a turbidostat experiment, the dilution rate (D) is defined by the culture volume (V) and the flow rate (F);

$$D = \frac{F}{V}$$

D equals the specific growth rate μ, which can be used to calculate the doubling time T_d;

$$T_d = \frac{ln(2)}{\mu}$$

3.4 Results

3.4.1 Integrating a Functional 3OHP Partcycle into *S. elongatus* PCC 7942

To create a strain, that was able to coassimilate propionate and glyoxylate effectively, the genes encoding Mcl from *A. phosphatis*, Mch from *R. sphaeroides* as well as Mct and Meh from *C. aurantiacus* were integrated into the wildtype (WT:M) and a carboxysome knockout strain (ΔK-O:M), that suffers from increased photorespiration rates. The integration success was assessed by colony PCR. To verify the function of the integrated genes beyond their presence in the genome, HPLC based enzyme assay were performed. Cell extracts of the ΔccmK-O strain and ΔccmK-O with the Mcl-route integration were tested for the conversion of propionyl-CoA and glyoxylate to acetyl-CoA and pyruvate.

The resulting chromatograms indicated the consumption of propionyl-CoA and the formation of acetyl-CoA in extracts of the strain containing the 3OHP bypass genes (Figure 3.2A). The chromatography also showed several peaks for different other CoA esters, that represent the intermediates (β-methylmalyl-CoA, mesaconyl-C1-CoA, mesaconyl-C4-CoA and citramalyl-CoA) of the reaction sequence. As Mcl is multifunctional and also produces (S)-malyl-CoA from acetyl-CoA and glyoxylate

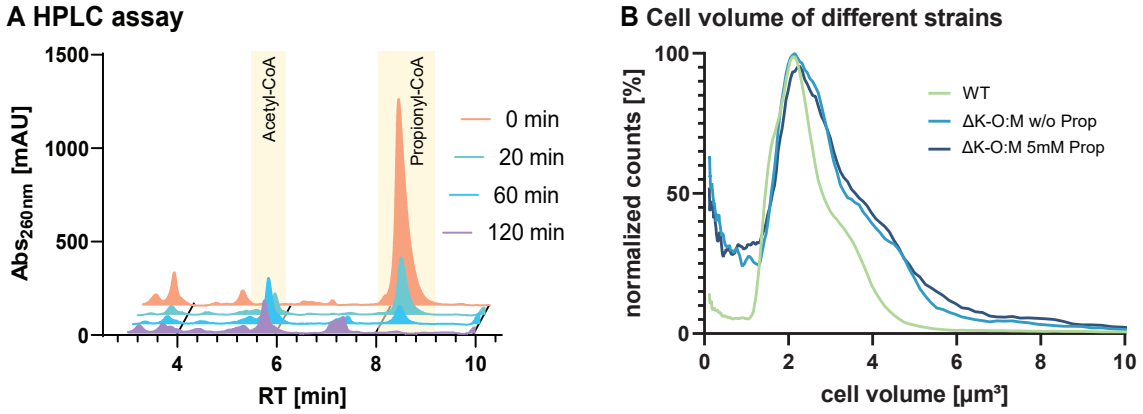


Figure 3.2: **Characterization of strain using Mcl-route.** HPLC assays of cell extract verifies reaction from propionyl-CoA to acetyl-CoA (A). Coulter Counter experiment determined cell volumes. ΔK-O:M has slightly elongated cells. Effect of propionate does not reflect in morphology (B)

this intermediate also occured over time. However, this sidereaction is of only hypothetical importance, as it is completely reversible and Mcl does not favor the synthesis of malyl-CoA. In living cells of *S. elongatus*, both acetyl-CoA and pyruvate would be incorporated into biomass by gluconeogenesis or amino acid biosynthesis.

The morphology of the resulting strain was investigated by fluorescence microscopy. As chlorphyll is a fluorophore, the viable cells were identified by fluroescent signal . The carboxysome lacking strain, that also contained the Mcl-route seem to have a slightly elongated morphology. This is a typical stress sign for procaryotes, and indicates a direct impairment even though the cells were cultivated in relaxing conditions. As the cells were grown in 5 mM of propionate, the cells did appeared even more elongated, but statistical analysis using a coulter counter experiment showed that cell sizes remained stable with 5 mM of propionate (Figure 3.2B).

3.4.2 Propionate Detoxification via the 3OHP bypass

The metabolism of some cyanobacteria, including *S. elongatus* PCC 7942 is known to be inhibited by propionate [68]. Sideactivity of Acs activates propionate to propionyl-CoA. As it can not be further metabolized by these strains, it depletes the free CoA pool. The strain inheriting the first reaction steps of the bypass, was able to metabolize

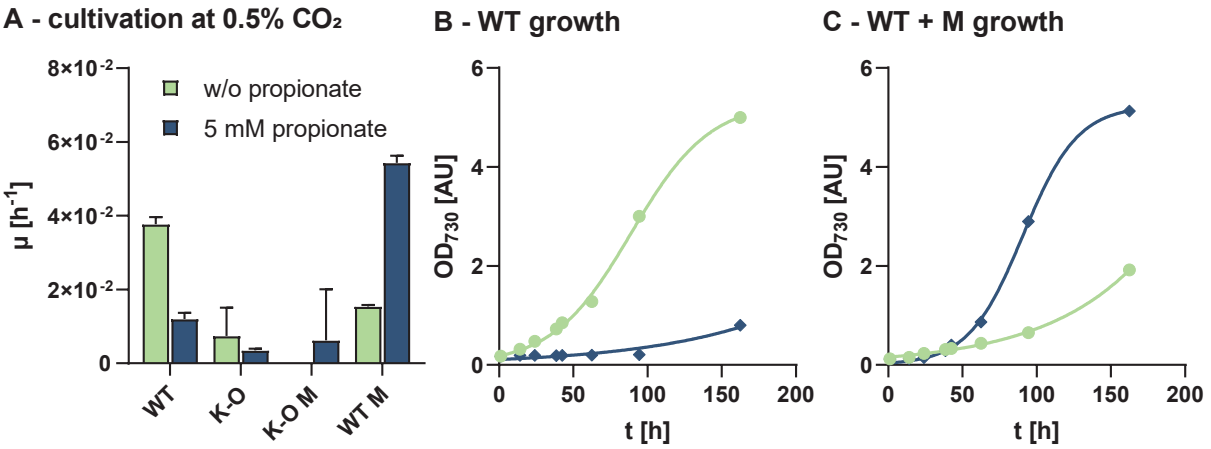


Figure 3.3: **Growth of different strains in presence of propionate.** Growthrate of different strains at 0.5% of CO_2 (A). Wildtype grows better in absence of propionate (B). WT:M strain grows better in presence of propionate (B). Error bars indicate SE.

propionyl-CoA while also removing the toxic photorespiratory glycolate/glyoxylate. Different strains of *S. elongatus* with this pathway were grown in presence of propionate (Figure 3.3) under 1.0% CO_2. Strains not expressing the bypass genes were more negatively affected by the presence of propionate. Wildtype growth rate decreased from 3.7 × 10^{-2} to 1.2 × 10^{-2}, wheras ΔK-O decreased from 7.4 × 10^{-3} to 3.4 × 10^{-3}. The strains containing the integration were positively affected by propionate. ΔK-O:M only grew in presence of propionate, and the growthrate of WT:M even increased by 3-fold. However, these results also indicate, that the production of 3OHP cycle enzymes came with a defined growth disadvantage. WT as well as the Δccm strain both grew notably worse with the genes of the 3OHP bypass, if no propionate was added (Figure 3.3). The burden of producing four additional enzymes might cause this growth deficit.

Growth experiments indicated that the beneficial effect of propionate feeding was limited to a low concentration of propionate. The growthrates decreased for concentrations above 5 mM (Figure 3.4). A Monod plot showed a theoretical μ_{max} of 1.1 × 10^{-2} h^{-1} at 2 mM. Even though photomixotrophic growth with cofeeding does not follow Monod kinetics, propionate served as additional feedstock. Growth was not affected when CO_2 levels were changed between 0.7% and 0.5%. Propionate and gly-

oxylate are metabolized to pyruvate and acetyl-CoA. As these molecules are central carbon metabolites this served mainly anabolic functions in *S. elongatus* PCC7942. The toxicity of propionate and glyoxylate seemed to overrule the growth benefits of the additionally produced acetyl-CoA and pyruvate at these concentrations.

These results are vital for the design and establishment of continious cultivation under selective conditions. The initial CO_2 tolerance of the designed ΔK-O:M strain was determined around 0.5%. Its ideal propionate concentration for growth was observed between 1 and 5 mM. With this information, a continuous ALE experiment was designed

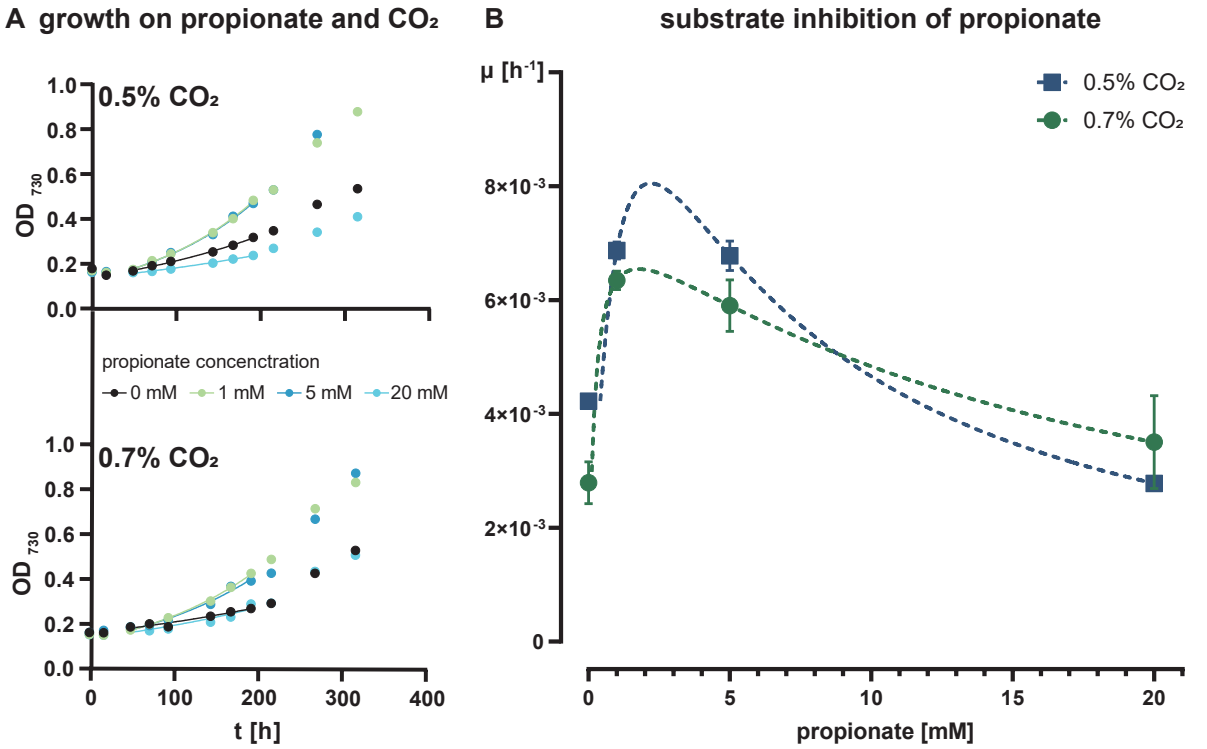


Figure 3.4: **Growth of ΔK-O:M strain on propionate.** The ΔK-O:M strains grow in different concentrations of CO_2 and propionate. Lines indicate area of logarithmic growth (A). A plot of this growth reveals the detoxification effect of propionate is beneficial in concentrations between 1 mM and 5 mM (B). Error bars indicate SE.

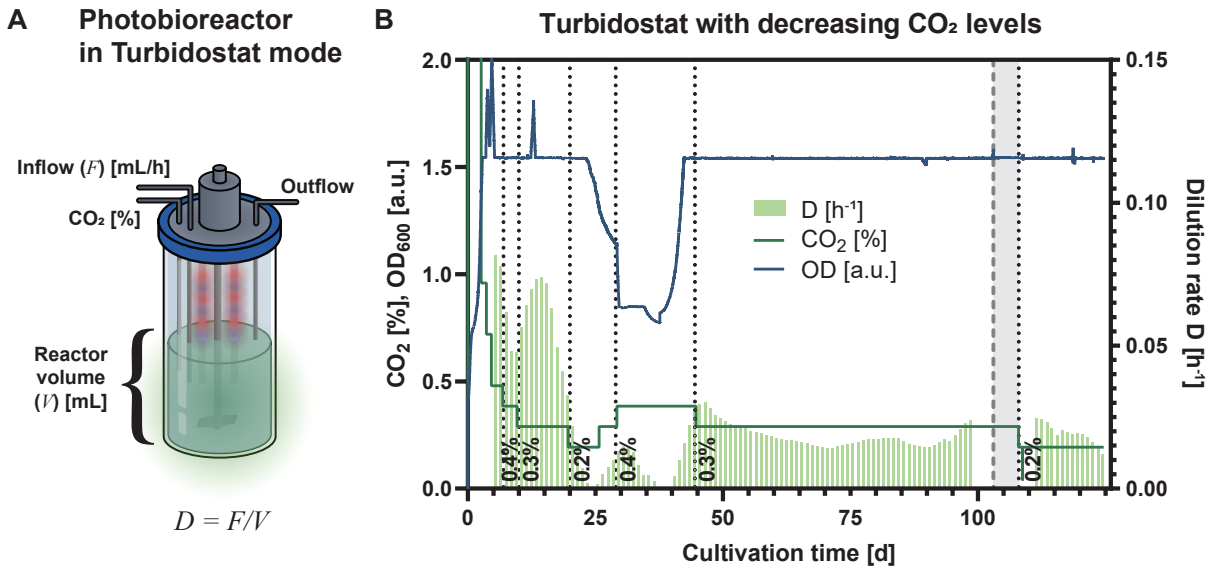


Figure 3.5: **Gradual evolution in a turbidostat continous cultivation setup.** Long time evolution of ΔK-O:M strain in a Turbidostat system. Total Pump volume indicates the dilution rate (D). Dilution/growth rate per day is presented in green bars (A). Dilution rate defined growth rate of the Turbidostat population. Strains were gradually exposed to less CO_2, while recovering in growth rate (B).

3.4.3 3OHP Bypass Strain Adapts to Lowered CO_2 Levels

To adapt the strain slowly to lowered concentrations of CO_2 while co-feeding of propionate, ΔK-O:M was cultivated in a continuous turbidostat photobioreactor for 125 days. This allowed fine control of growth parameters and adaptive responses to changes in growth behaviour. The CO_2 levels were initially set to 3.0%, and then lowered to 1.0, 0.7, 0.5, 0.4, 0.3 and ultimately to 0.2% (Figure 3.5). However, lowering the CO_2 level to 0.2% after the first 20 days, resulted in growth stagnation, and the OD even decreased from OD 1.5 to OD 1.2 (after day 25). Growth was recovered by an increase of the CO_2 level, and lowering the OD setpoint to OD 0.9 at day 30 to replenish propionate enriched medium. After the steady state was re-established, the OD setpoint was set back to 1.5 at day 45, and as soon as this reached a steady state, the cells were cultivated for 65 days at CO_2 levels of 0.3 %. After 55 days of incubation, the cells started to grow faster with a local optimum of 0.025 h^{-1}. As this indicated that the culture adapted to the low CO_2 levels, it was decreased to 0.2% at day 108,

without the stagnation or drop in OD that had been observed earlier at day 25. Even though the estimated growthrate was decreasing over the course of 15 days, the longer adaptation on 0.3 %, allowed cells to be viable on 0.2 %. After 125 days of cultivation, the reactor was paused for maintenance. The final cell population was transfered to flasks and will be further characterized for mutations and exact growthrates. A strain of this population may then be used to complete the photorespiratory cycle bypass and/or undergow annother round of ALE.

3.5 Discussion

The functional integration of the 3OHP bypass has been held back by the endogenous regulation of the Acc. Sufficient Acc activity for a photorespiratory sink has not yet been measured in *S. elongatus*. However, the 3OHP bypass has the theoretical potential to improve carbon fixation and thus result in better growth and potentially crop yields. [5].

The ΔK-O:M strain that transformed with the 3OHP bypass genes resulted in detectable enzyme activity for the conversion of propionyl-CoA to acetyl-CoA, indicating production and functionality of the enzymes. Judged by morphology change (Figure 3.2), the strain seems to be affected by the burden of lacking the carboxysome and additionally expressing the genes for the Mcl-route. The strain does not seem to change morphology as propionate is introduced. This ties along with the detectable growth benefit that was observed in all strain, when the Mcl-route was combined with propionate supplementation (Figure 3.3+3.4).

We show that the resulting strain was not able to grow at ambient CO_2 levels, but in 0.5% of CO_2 in presence of 5 mM propionate. The toxicity of propionate was not completely avoided, as concentrations of more than 5 mM resulted in reduced growth. Based on these findings, we started a continious turbidostat evolution experiment, by

decreasing the level of CO_2, while co-feeding propionate. We found, that the cells were able to adapt to the decreased levels of CO_2. After a growing for 65 days on 0.3% of CO_2, the cells were able to even sustain growth at 0.2% CO_2, as well. At this point, the resulting population most likely also changed their tolerance towards external propionate again.

As these results were obtained realtively recent, the population from this experiment could not be carefully characterized yet. However, we speculate that genetic changes in genes of central metabolism, inorganic carbon transporters, and most hopefully also in the integrated parts of the genome have occurred. We therefore propose, that the population will be carefully tested for growth behaviour in flasks to reproduce or even exceed the tolerance towards decreasing concentrations of CO_2. Additionally, the propionate tolerance should be determined again, by growing the same strain at 0.2% CO_2, with varying concentrations of propionate. And to see what caused the new tolerance of the population, whole genome sequencing should be performed in order to identify genomic mutations. The next steps clearly will be the reproduction of the here shown set of experiments with the integration of the remaining genes for the completion of the pathway and the cultivation of the resulting strain in presence of propionate or even 3OHP at ambient CO_2 levels while slowly removing the amount of externally fed substrate.

Our results so far are encouraging and show that our ALE strategy enabled growth at CO_2 concentrations that were not achievable with the parental ΔK-O:M strain. Future ALE experiments with the whole bypass will hopefully be able to force the re-purposing of Acc for the photorespiratory bypass in a similar fashion.

Besides the realization of the 3OHP Bypass in Cyanobacteria, we investigated on an enzyme with unique properties. The Mct from *C. aurantiacus* performs an intramolecular CoA transfer. Chapter 4 is dedicated to its structural analysis and the description of its reaction mechanism

Chapter 4

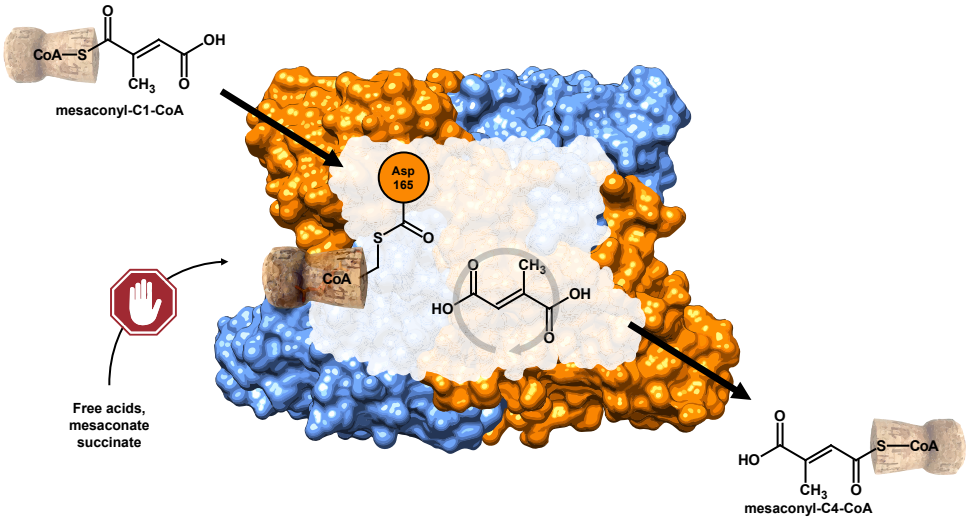
Structural Basis for a Cork-Up Mechanism of the Intra-Molecular Mesaconyl-CoA Transferase

Pascal Pfister[1], Jan Zarzycki[1], and Tobias J. Erb[1,2,*]

[1] Department of Biochemistry & Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch Str. 10, 35043 Marburg, Germany

[2] SYNMIKRO Center for Synthetic Microbiology, Karl-von-Frisch Str., 14, 35032 Marburg, Germany

* corresponding author: toerb@mpi-marburg.mpg.de



4.1 Abstract

Mct is one of the key enzymes of the 3OHP bi-cycle for autotrophic CO_2 fixation. Mct is a family III/Frc family CoA transferase that catalyzes an unprecedented intra-molecular CoA transfer from the C1-carboxyl group to the C4-carboxyl group of mesaconate at catalytic efficiencies $>10^6$ M^{-1} s^{-1}. Here, we show that the reaction of Mct proceeds without any significant release of free CoA or the transfer to external acceptor acids. Mct catalyzes intra-molecular CoA transfers at catalytic efficiencies that are at least more than 6 orders of magnitude higher compared to inter-molecular CoA transfers, demonstrating that the enzyme exhibits exquisite control over its reaction. To understand the molecular basis of the intra-molecular CoA transfer in Mct, we solved crystal structures of the enzyme from *C. aurantiacus* in its apo form, as well as in complex with mesaconyl-CoA and several covalently enzyme-bound intermediates of CoA and mesaconate at the catalytically active residue Asp165. Based on these structures, we propose a reaction mechanism for Mct that is similar to inter-molecular family III/Frc family CoA transferases. However, in contrast to the latter that undergo opening and closing cycles during the reaction to exchange substrates, the central cavity of Mct remains sealed ("corked-up") by the CoA moiety, strongly

favoring the intra-molecular CoA transfer between the C1 and the C4 position of mesaconate.

4.2 Introduction

The thermophilic green non-sulfur bacterium *C. aurantiacus* uses the 3OHP bi-cycle for autotrophic CO_2 fixation. [18,26,71] In the first part of the 3OHP bi-cycle, CO_2 is captured via two biotin-dependent carboxylases yielding glyoxylate as the primary CO_2-fixation product. In the second part of the 3OHP bi-cycle, glyoxylate is condensed with propionyl-CoA into (2R,3S)-β-methylmalyl-CoA. [72] Methylmalyl-CoA is rearranged and converted into acetyl-CoA and pyruvate, the final CO_2-fixation product. [18] The rearrangement sequence of the second cycle starts via dehydration of methylmalyl-CoA into mesaconyl-C1-CoA (2-methylfumaryl- CoA). [73] The CoA moiety of mesaconyl-C1-CoA is then transferred from the C1- to the C4-carboxyl group by Mct, resulting in mesaconyl-C4-CoA (3-methylfumaryl-CoA). Mesaconyl-C4-CoA is further converted into (3S)-citramalyl-CoA, which is ultimately cleaved into acetyl-CoA and pyruvate (Fig. 4.1).

The Mct reaction is a key reaction in the 3OHP bi-cycle. It conserves the energy-rich CoA-ester bond during C1-/C4-transfer without the release of mesaconate or the transfer of CoA onto other acceptors, which would result in a loss of intermediates and require additional ATP for (re-)activation of the free mesaconate. Overall, this makes the intra-molecular C1-/C4-CoA transfer by Mct an elegant and energetically highly efficient solution.

CoA transferases have been traditionally categorized into three different families, although recent phylogenetic analysis indicates that the evolutionary history of family I and II CoA transferases is more complex and that CoA transferases fall into six different monophyletic groups [74] (see Table 6 for Pfams). "Family II" members (i.e., members of the CitF and MdcA families) are enzyme complexes that naturally use

acyl-carrier proteins during catalysis but are also able to accept CoA esters as substrates in vitro [75–77]. In contrast, "family I" members (i.e., members of the Cat1, OXCT1, and Gct families) and family III members (i.e., members of the Frc family) are lone-standing enzymes that typically catalyze the inter-molecular CoA transfer between a CoA donor and an acceptor acid in a similar fashion [78–81]. The initial step in these enzymatic reactions is the nucleophilic attack of an active site, glutamate ("family I" members) or aspartate (family III/Frc family members), on the donor CoA ester, resulting in an acyl-enzyme anhydride and free $CoAS^-$. The $CoAS^-$ subsequently attacks the acyl-enzyme anhydride, releasing the donor acid and yielding a γ-glutamyl-bound ("family I") or β-aspartyl-bound (family III/Frc family) enzyme-CoA thioester intermediate. The acceptor acid attacks the enzyme-CoA thioester to release $CoAS^-$ and forms another acyl-enzyme anhydride. In the last step, this anhydride is re-attacked by the $CoAS^-$, releasing the new CoA thioester [77,78,80,82,83].

The catalytic mechanisms of "family I" and family III/Frc family CoA transferases follow similar principles. However, while "family I" transferases use a classical ping-pong mechanism [79–81], family III/Frc family enzymes show a modified mechanism, in which access of small acceptor acids to the active site may be gated either through a flexible glycine loop [78,84–88] or even larger domain movements as observed for crotonobetainyl-CoA:carnitine CoA transferase (CaiB) [84]. The glycine-rich loop presumably opens and closes during the catalytic cycle to allow access of the acceptor acid upon formation of the β-aspartyl-CoA intermediate with the donor acid still present at the active site. After CoA transfer, the newly formed acceptor acid-CoA thioester and the then free donor acid are released. Crystallographic evidence for these enzyme-bound intermediate states was presented for the formyl-CoA transferase (Frc) of Oxalobacter formigenes [78].

While Mct falls within canonical family III/Frc family CoA transferases, the enzyme catalyzes an unprecedented intra-molecular CoA transfer, in which the acceptor acid (i.e., the second carboxylic group of mesaconate) is already part of the CoA donor.

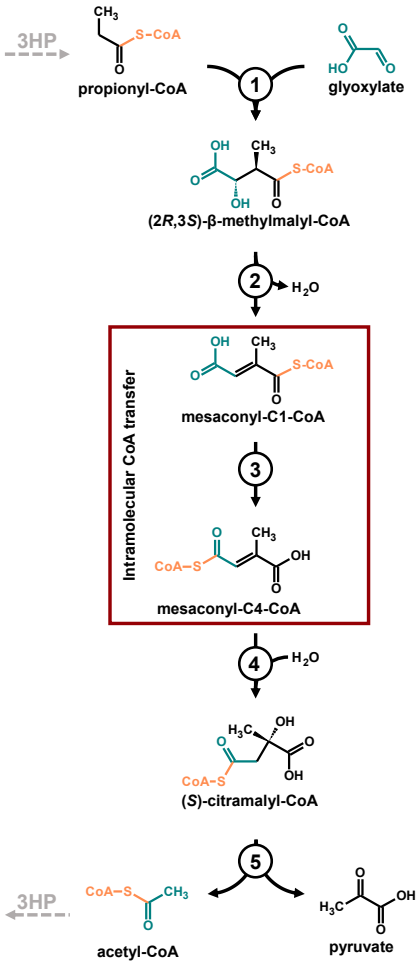


Figure 4.1: **Reaction sequence of the 3OHP bi-cycle involving Mct [18]:** (S)-malyl-CoA/(2R,3S)-β-methylmalyl-CoA/(3S)-citramalyl-CoA lyase (1 and 5), mesaconyl-C1-CoA hydratase (2), mesaconyl-C1:C4-CoA CoA transferase (3), and mesaconyl-C4-CoA hydratase (4). Metabolic connection to the 3OHP bi-cycle is indicated in dashed lines. The atoms originating from glyoxylate are colored in teal. The CoA moiety is colored in orange.

Since there is no need to introduce an additional substrate during the catalytic cycle, it has been speculated that the active site stays fully closed during catalysis [18, 89]. This hypothesis is consistent with the observation that small inactivating molecules that could react with the acyl-enzyme anhydride intermediate, such as hydroxylamine or borohydride, had little or even no effect on Mct activity [18]. However, this also means that mesaconate would need to re-orient within the active site of Mct to enable CoA transfer from C1 to C4. Because of these proposed major differences to the catalytic cycle of inter-molecular CoA transferases, the mechanism of intra-molecular

CoA transfer by Mct remained elusive.

Recently, the structure of the Mct homologue from *Roseiflexus castenholzii* (PDB 7XKG) was reported in its apo form [89]. This structure showed that a flexible glycine-rich loop that supposedly gates catalysis in some other inter-molecular family III/Frc family CoA transferases [78,87] is absent in Mct, indicating that the reaction may proceed differently in the intra-molecular CoA transferases. Based on the structure of the apoenzyme, molecular dynamics simulations with mesaconyl-C1- and C4-CoA were performed [89] and a mechanism for the intra-molecular CoA transfer of Mct was proposed, which differed from the canonical family III/Frc family CoA transferases. Notably, a direct, water-assisted attack of the free $CoAS^-$ onto the free carboxy group of mesaconate has been postulated [89]. However, this mechanism seems biochemically infeasible and support for this mechanism is lacking. Here, we sought to further biochemically and structurally characterize Mct from *C. aurantiacus* to better understand the molecular basis of catalysis. We show that Mct is virtually an exclusive intra-molecular CoA transferase and provide atomic-resolution crystal structures of the enzyme with different bound intermediates. Based on this data, we propose a mechanism for Mct that is similar to those of inter-molecular family III/Frc family transferases with the enzyme's active site being "corked-up" by the substrate's CoA moiety. This active site sealing likely favors the intra-molecular CoA transfer over inter-molecular CoA transfer in Mct by several orders of magnitude, resulting in a highly selective enzyme.

4.3 Materials and Methods

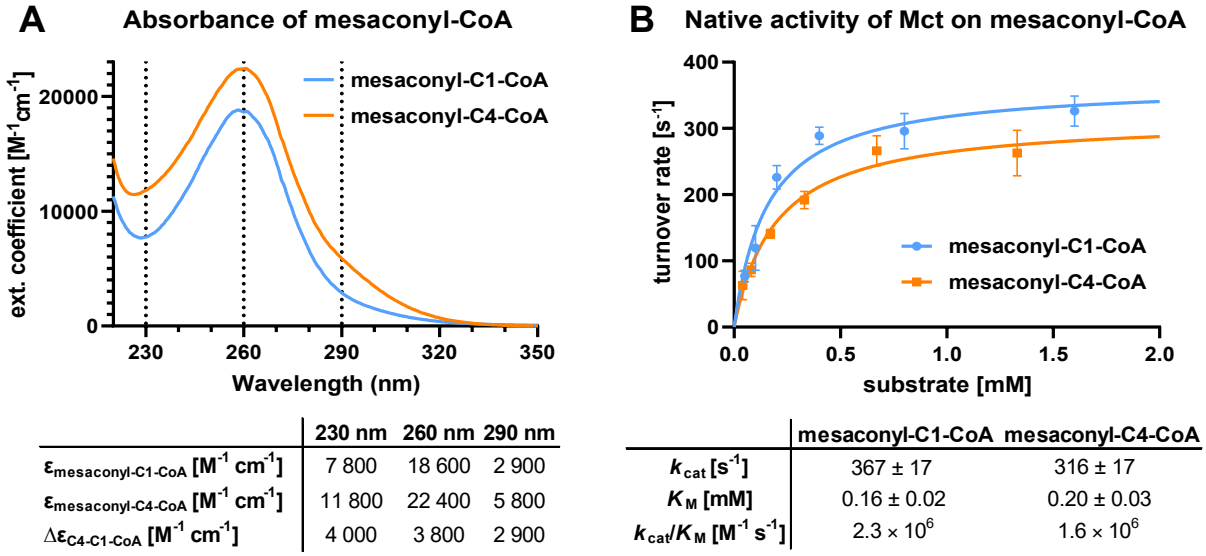
4.3.1 Synthesis of CoA Thioesters

Synthesis of Mesaconyl-C1- and Mesaconyl-C4-CoA

First, 0.5 M mesaconic acid (116 mg) was dissolved in 2 mL of diethylether on ice. Then, 80 µL water-free pyridine and 94 µL of ice-cold ethyl chloroformate were added under constant stirring. After 15 min, the supernatant containing mesaconic anhydride was slowly added to a CoA solution (2.5 mM CoA, 25 mM $NaHCO_3$). After 30 min of constant stirring on ice, pH was adjusted to pH of 3.0 with HCl. [90] Free CoA, mesaconyl-C1-CoA, and mesaconyl-C4-CoA were separated by HPLC (Agilent 1260 Infinity HPLC) with a Gemini 10 µm NX-C18 110 Å column (Phenomenex) in a gradient from 14 to 50% methanol in buffer (25 mM NH_4COOH/HCOOH, pH 4.2) over 10 min at a flow rate of 25 mL/min. Mesaconyl-C1- and C4-CoA could be differentiated by their UV spectra (Figure 4.2) and retention times. The retention times for CoA, mesaconyl-C1-CoA, and mesaconyl-C4-CoA were 1.9, 3.9, and 4.8 min, respectively. Peak fractions were pooled, frozen in liquid nitrogen, and subsequently lyophilized. The resulting powder was stored at $-20\,°C$ and solved in ddH2O before use. Purity was confirmed by HPLC-MS. Both CoA thioesters were >99% pure. They did not show any cross-contamination with the respective other mesaconyl-CoA derivative (Figure 4.7) or with free CoA, as judged by Ellman's reagent.

Synthesis of Other CoA Esters

All other CoA thioesters were synthesized and purified according to previously established protocols [56,91].



	230 nm	260 nm	290 nm
$\varepsilon_{\text{mesaconyl-C1-CoA}}$ [M^{-1} cm^{-1}]	7 800	18 600	2 900
$\varepsilon_{\text{mesaconyl-C4-CoA}}$ [M^{-1} cm^{-1}]	11 800	22 400	5 800
$\Delta\varepsilon_{\text{C4-C1-CoA}}$ [M^{-1} cm^{-1}]	4 000	3 800	2 900

	mesaconyl-C1-CoA	mesaconyl-C4-CoA
k_{cat} [s^{-1}]	367 ± 17	316 ± 17
K_M [mM]	0.16 ± 0.02	0.20 ± 0.03
k_{cat}/K_M [M^{-1} s^{-1}]	2.3×10^6	1.6×10^6

Figure 4.2: **Spectrophotometric data for mesaconyl-CoA thioesters and the CoA transferase reaction.** (A) UV spectra of mesaconyl-C1-CoA (blue) and mesaconyl-C4-CoA (orange). Extinction coefficients at 230, 260, and 290 nm for each substrate are given below The difference in absorbance at 290 nm ($\Delta\varepsilon_{4\text{-C1-CoA}}$) was used for photometric activity assays. (B) Michaelis–Menten plot for Mct activity with mesaconyl-C1-CoA (blue) and mesaconyl-C4-CoA (orange), respectively. Kinetic values for both substrates are given in the bottom table. SD values are indicated.

4.3.2 Gene Expression and Protein Purification

The expression plasmid pMCTCa_JZ05 encoding a His-tagged Mct from *C. aurantiacus* [18] was used for protein production. The plasmid was transformed into *E. coli* BL21 DE3, grown in 2 L Terrific Broth [92] for 24 h at 25 °C without induction. The cells were harvested at 4 °C and 8000g and resuspended in a threefold volume (3 mL per 1 g of cells) of loading buffer (50 mM MOPS/KOH pH 7.8, 150 mM NaCl, 75 mM imidazole). The cells were lysed using an LM10 microfluidizer (H10Z chamber, Microfluidics, Westwood, MA) at 18 000 psi. The lysate was heat-precipitated at 70.0 °C for 20 min and kept on ice for downstream purification. The unwanted denatured proteins were removed by centrifugation at 4 °C and 100 000g for 1 h. The cell extract was filtered (0.4 μm syringe filter), and Mct was purified by nickel affinity chromatography (elution buffer 50 mM MOPS/KOH, pH 7.8, 150 mM NaCl, 500 mM imidazole) using a 1 mL HisTrap FF column (Cytiva, Freiburg, Germany). Afterward, the elu-

ate was desalted in low-salt buffer (50 mM MOPS/KOH, pH 7.8, 50 mM NaCl) and further purified by anion exchange (Q-HP 16/10 column, Cytiva, Freiburg, Germany) using a gradient with high-salt buffer (50 mM MOPS/KOH, pH 7.8, 300 mM NaCl) over 20 min. The enzyme eluted between NaCl concentrations of 100 and 150 mM. The purity of the enzyme was checked by SDS-PAGE at each purification step, and Mct was concentrated by centrifugal filters (Amicon by Merck, Darmstadt, Germany) with a 30 kDa cutoff. Protein concentrations were determined using a NanoDrop (Thermo-Fisher Scientific) and applying a calculated molar extinction coefficient of 49 000 M^{-1} cm^{-1} coefficient at 280 nm.

4.3.3 Determination of the Extinction Coefficient of Mesaconyl-CoA Derivatives

UV spectra (220 - 350 nm) of both HPLC-purified mesaconyl-CoA forms were recorded via spectrophotometer (Cary 60, Agilent). CoA thioester concentrations were measured by depletion in a coupled NADPH-dependent assay [93] by reduction via succinate-semialdehyde dehydrogenase (SucD, EC:1.2.1.76) spectrophotometrically [3] (Cary 60, Agilent) in a 1 cm quartz cuvette (300 µL assay volume; 200 mM HEPES, pH 7.5, 70 nM SucD, 0.7 mM NADPH and about 0.25 mM mesaconyl-CoA) at 365 nm and 37 °C.

4.3.4 Determination of Enzymatic Activity

Spectrophotometric Assay

To examine the activity of the intra-molecular CoA transfer of Mct, a spectrophotometric assay was used. Mesaconyl-C4-CoA has a higher extinction coefficient at 290 nm ($\varepsilon_{290\,nm}$ = 5 800 M^{-1} cm^{-1}) than mesaconyl-C1-CoA ($\varepsilon_{290\,nm}$ = 2 900 M^{-1} cm^{-1}). Therefore, the conversion of mesaconyl-C1-CoA was measured by the increase in absorbance at 290 nm ($\Delta\varepsilon_{290\,nm}$ = 2 900 M^{-1} cm^{-1}). To conduct the measurements, 150 µL

assay volume (200 mM HEPES/KOH, $pH_{25\,°C}$ 8.0, 22 nM Mct, and varying concentrations of mesaconyl-CoA) was incubated at 55 °C, and change in absorbance at 290 nm was monitored over time in a 3 mm quartz cuvette. The reaction was started with the substrate (ranging from 50 to 1600 μM for mesaconyl-C1-CoA and 40 to 1300 μM for mesaconyl-C4-CoA).

HPLC-MS-Based Assay

To test for alternative CoA acceptors, Mct was preincubated in reaction buffer (200 mM HEPES, pH 7.5) and supplemented with 20 mM of the corresponding carboxylic acid. After 5 min of preincubation, the reaction was started by the addition of 1 mM mesaconyl- C1-CoA. Samples were taken after 0 and 20 min and stopped on ice by the addition of HCl to a final concentration of 100 mM. The precipitated enzyme was removed by centrifugation (4 °C and 17 000 × g), and the supernatants were analyzed by HPLC-MS and for the presence of alternative CoA thioesters. To evaluate and quantify the kinetics of succinate as acceptor acids, an enzyme assay was performed (55 μL; 200 mM HEPES/KOH, $pH_{25\,°C}$ 8.0, 6 μM Mct, 1 mM mesaconyl- C4-CoA, and varying concentrations of succinate). The reaction was started with the addition of succinate and incubated for 20 min at 55 °C. At 0, 1, and 20 min, a sample of 5 μL was taken and quenched in 45 μL of formic acid. The precipitated enzyme was removed by centrifugation (4 °C and 17 000 × g), and the supernatants were analyzed by HPLC-MS for the presence of succinyl-CoA.

Determination of CoA thioesters was performed using a HiRes-LC-MS. The chromatographic separation was performed on a Thermo Scientific Vanquish HPLC system using a Kinetex Evo C18 column (150 × 2.1 mm2, 100 A, 1.7 μm, Phenomenex) equipped with a 20 × 2.1 mm2 guard column of similar specificity at a constant eluent flow rate of 0.25 mL/min and a column temperature of 25 °C with eluent A being 50 mM ammonium formate at a pH of 8.1 water and eluent B being MeOH (Honeywell). The injection volume was 1 μL. The elution profile consisted of the following steps

and linear gradients: 0-2 min constant at 0% B; 2-10 min from 0 to 80% B; 10-12 min constant at 80% B; 12-12.1 min from 80 to 0% B; and 12.1-15 min constant at 0% B. A Thermo Scientific ID-X Orbitrap mass spectrometer was used in positive mode with an electrospray ionization source and the following conditions: ESI spray voltage 3500 V, sheath gas at 50 arbitrary units, auxiliary gas at 10 arbitrary units, sweep gas at 1 arbitrary unit, ion transfer tube temperature at 300 °C, and vaporizer temperature at 350 °C. Detection was performed in full-scan mode using the orbitrap mass analyzer at a mass resolution of 240 000 in the mass range 800-900 (m/z). Extracted ion chromatograms of the $[M^+ H]^+$ forms were integrated using Tracefinder software (Thermo Scientific). Absolute concentrations for succinyl-CoA were calculated based on an external calibration curve.

4.3.5 Crystallization of Mct X-ray Structure Determination

The purified protein solution was spotted in different concentrations (3, 6, and 8 mg/mL) on sitting-drop vapor-diffusion crystallization plates. First, 0.2 µL of each protein solution was mixed with 0.2 µL of crystallization condition. The drops were equilibrated against 30 µL of protein-free crystallization condition at 288 K. The resulting crystals of condition A (200 mM sodium chloride, 100 mM sodium potassium phosphate, pH 6.2, and 50% v/v poly(ethylene glycol) 200) appeared after 5 days. In wells containing condition B (35% 2-methyl-2,4-pentanediol and 100 mM sodium/potassium phosphate, pH 6.2) crystals appeared after 2 days and grew until the 5^{th} day of incubation. The crystals in condition A were directly snap-frozen in liquid nitrogen, whereas the crystals of condition B were transferred into a drop containing higher concentrations of cryoprotectant and a mixture of both forms of mesaconyl-CoA (40% MPD, sodium/ potassium phosphate, pH 6.2, 5 mM mesaconyl-CoA) for 2 min and were then frozen in liquid nitrogen. X-ray diffraction data were collected at the beamline ID29 of the European Synchrotron Radiation Facility (ESRF) and the beamline P14 of the Deutsches Elektronen-Synchrotron (DESY). The data sets

were processed with the XDS software package [57]. The structures were solved by molecular replacement using a polyalanine search model of the formyl-CoA:oxalate CoA transferase from Acetobacter aceti (PDB ID 3UBM) [94]. Molecular replacement was carried out using Phaser of the Phenix software package [58] and refined with Phenix.Refine. Additional modeling, manual refining, and ligand fitting were done in COOT [59]. Final positional and B-factor refinements, as well as water-picking for the structure, were performed using Phenix.Refine. The Mct structure models were deposited at the PDB in Europe under PDB IDs 8APR and 8APQ. Data collection and refinement statistics are provided in Table 5.

Table 5: Data and Refinement Statistics for the Mct Crystal Structures. Statistics for the highest-resolution shell are in parentheses.

crystal	Mct - apo form	Mct with bound substrates
beamline	ESRF ID29, Grenoble, France	DESY P14, Hamburg, Germany
PDB ID	8APR	8APQ
ligands	3 Cl^-	mesaconyl-C1-CoA, mesaconate, CoA
wavelength	0.96862	0.97660
resolution range (Å)	29.2-2.1 (2.2-2.1)	29.7-2.5 (2.6-2.5)
space group	C121	$P3_221$
unit cell dimensions		
a, b, c (Å)	172.6, 103.5, 95.3	193.8, 193.8, 252.0
α, β, γ (deg)	90.0, 119.6, 90.0	90.0, 90.0, 120.0
total reflections	576 260 (51 434)	1 110 749 (111 537)
unique reflections	84 090 (8166)	189 902 (18 751)
multiplicity	6.9 (6.3)	5.8 (5.9)
completeness (%)	99.52 (96.95)	99.68 (99.38)
mean I/σ(I)	13.47 (2.28)	13.57 (2.68)
Rmerge	0.1011 (0.8707)	0.08911 (0.7038)
Rpim	0.04167 (0.3762)	0.0401 (0.3118)
CC1/2	0.998 (0.864)	0.998 (0.804)
reflections used in refinement	84 021 (8139)	189 871 (18750)
Rwork	0.1808 (0.2690)	0.1850 (0.2386)
Rfree	0.2125 (0.3305)	0.1995 (0.2702)

Table 5: Data and Refinement Statistics for the Mct Crystal Structures. Statistics for the highest-resolution shell are in parentheses.

crystal	Mct - apo form	Mct with bound substrates
number of non-hydrogen atoms	9886	20187
macromolecules	9354	18795
ligands	3	337
solvent	529	1055
protein residues	1212	2435
RMS (bonds)	0.007	0.002
RMS (angles)	0.86	0.49
Ramachandran		
favored (%)	98.09	97.36
allowed (%)	1.66	2.39
outliers (%)	0.25	0.25
rotamer outliers (%)	0.52	0.10
clashscore	1.60	1.07
average B-factor	41.76	50.62
macromolecules	41.62	50.30
ligands	35.78	69.70
solvent	44.32	50.19

4.4 Results

4.4.1 Mct Is a Highly Efficient Intra-Molecular Mesaconyl-CoA Transferase

For the spectrophotometric kinetic characterization of Mct, we first synthesized and purified mesaconyl-C1-CoA and mesaconyl-C4-CoA. We revisited the UV spectra of both CoA thioesters to determine their exact extinction coefficients at 230, 260, and 290 nm. While the overall spectra of both compounds were similar, the C1 and C4 species showed distinct differences. Compared to mesaconyl-C1-CoA, the spectrum

of mesaconyl-C4-CoA resembled more those of other α,β-unsaturated CoA esters like crotonyl- or acrylyl-CoA, exhibiting a higher overall extinction coefficient at 260 nm and a more pronounced shoulder in the region between 280 and 340 nm (Figure 4.2A). We then used the difference in the extinction coefficients at 290 nm ($\Delta\varepsilon_{290\,nm}$ = 2900 M^{-1} cm^{-1}) to determine the catalytic properties of Mct from *C. aurantiacus* at the organism's optimum growth temperature of 55 °C with mesaconyl-C1-CoA and mesaconyl-C4-CoA in a continuous photometric assay. Starting with either of the substrates, the enzyme showed remarkably high V_{max} values of 495 and 430 μmol min^{-1} mg^{-1} for mesaconyl-C1- and C4-CoA, corresponding to k_{cat} values of 370 and 320, respectively (Figure 4.2B). The K_M values for both CoA esters were 0.16 and 0.2 mM, resulting in catalytic efficiencies (k_{cat}/K_M) of 2.3 × 106 and 1.6 × 106 M^{-1} s^{-1} for mesaconyl-C1-CoA and mesaconyl-C4-CoA, respectively (Figure 4.2B). These kinetic parameters are in line with previously published values [18] while also considering the revised extinction coefficients.

4.4.2 Mct Strongly Discriminates against Other Substrates

Next, we wanted to assess Mct's ability to use succinate as an alternative dicarboxylic acid-CoA acceptor when externally provided during catalysis with mesaconyl-CoA. We detected only a negligible side activity (i.e., formation of succinyl-CoA) with an extremely low catalytic efficiency for the CoA transfer onto succinate (k_{cat}/K_M = 0.49 M^{-1} s^{-1}), which is more than 6 orders of magnitude lower compared to the interconversion of the two different mesaconyl-CoA thioesters (see Figure 4.3A). This strong selectivity against free succinate was accompanied by a very high apparent K_M for this alternative substrate (>25 mM).

Having identified a very low, but detectable activity with succinate, we sought to test other central carbon metabolites as potential acceptor acids and several alternative acyl-CoA thioesters as potential CoA donors. To that end, we preincubated different carboxylic acids (mesaconate, succinate, malate, crotonate, and acetate) individually

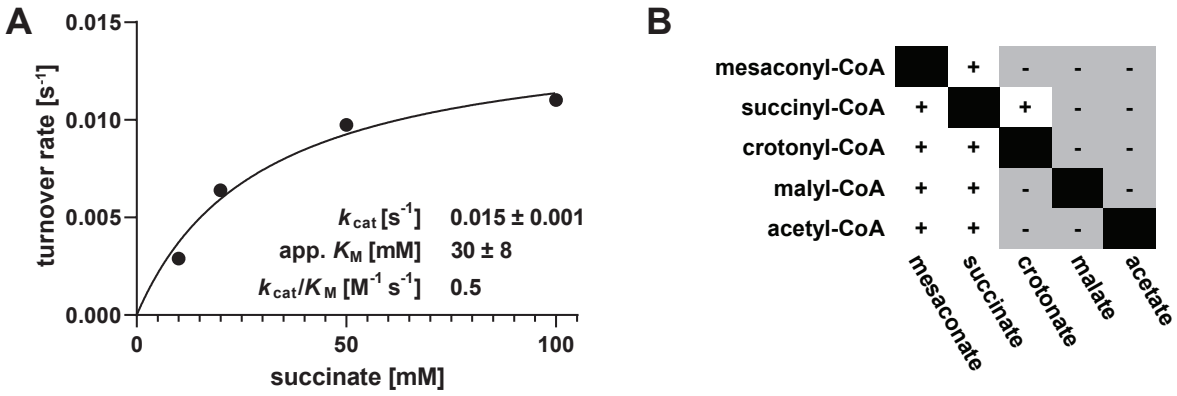


Figure 4.3: **Testing externally provided acids and CoA esters for inter-molecular CoA transfer.** (A) Kinetic parameters for mesaconyl-CoA:succinate transfer show very poor catalytic efficiency for succinate as an alternative CoA acceptor. The SD is indicated. (B) Testing different potential CoA donors and CoA acceptors shows that only succinate and mesaconate serve as CoA acceptors, whereas all tested CoA esters serve a as CoA donor for succinate and mesaconate. "+" indicates that CoA was transferred onto the respective acid, as confirmed by HPLC–MS. "–" indicates that the formation of a corresponding CoA thioester could not be detected by HPLC–MS.

at concentrations of 20 mM for 5 min with Mct, before the reaction was started with 1 mM of either mesaconyl-, succinyl-, crotonyl-, or acetyl- CoA. In these assays, Mct also accepted crotonate as an alternative acceptor acid, when succinyl-CoA was provided as a CoA donor (Figure 4.3B). However, activity with crotonate as a CoA acceptor was comparable to or even lower than for succinate and several orders of magnitude lower than the intra-molecular reaction with mesaconyl-CoA alone. This demonstrated that Mct is able to efficiently discriminate against other carboxylic acids during catalysis. When testing mesaconate as a CoA acceptor with different alternative CoA donors, we found that all of the tested CoA esters could in general serve as substrates (Figure 4.3B). However, mesaconyl-CoA formation only occurred when mesaconate was provided in nonphysiologically high concentrations (20 mM). These results are in line with previous data that reported a lack of detectable radioactive products when either 14C-labeled mesaconyl-CoA or 14C-labeled mesaconate was used with their respective unlabeled counterparts [18]. We, therefore, reason that these (side) reactions are likely irrelevant under physiological conditions. Taken together, our data show that Mct can neither serve as a mesaconyl-CoA:carboxylic acid-CoA transferase nor pos-

sess a significant activity as acyl-CoA:mesaconate CoA transferase and therefore is an authentic intra-molecular CoA transferase.

4.4.3 Crystal Structure Reveals Snapshots of the Catalytic Cycle

Next, we became interested in understanding the structural determinants underlying substrate discrimination in Mct. We first solved the crystal structure of Mct from C.aurantiacus in its apo form without substrates at 2.1 Å resolution. Similar to the recently solved crystal structure of the homologue from Roseiflexus [89] and the other family III/Frc family CoA transferases [78, 84, 86, 94], Mct of *C. aurantiacus* is an intertwined homodimer (Figure 4.4) [85], where the polypeptide chains are threaded through a hole in the neighboring subunit (Figure 4.4B), respectively. A Rossman fold is formed between the C- and the N-termini of the enzyme. Residues Leu8 to Ala195 of the N-terminus form the essential part of the Rossman fold motif, followed by a loop that completely wraps around the adjacent subunit of the Mct dimer. This loop ends in a structure on the opposite side of the Rossman fold harboring three antiparallel β-strands and five short α-helices. Another loop reaches back to the described N-terminal structure, in which residues following Thr398 complete the Rossman fold.

We also solved another structure of Mct under different crystallization conditions and with substrate soaking at 2.5 Å resolution. Under these conditions, we detected three dimers of Mct in the asymmetric unit (ASU). The structure of the soaked crystal showed additional electron densities at the six active sites representing different states of bound substrates and/or reaction intermediates. The active sites are located in cavities that are formed directly at the dimerization interfaces between the two subunits. They are located adjacent to the Rossman fold of each monomer and harbor the catalytically active Asp165 residue, which itself is part of the last helix of the Rossman fold. Although the electron densities at the active sites were slightly ambiguous, representing somewhat mixed states, we were able to model mesaconyl-C1-CoA (Figure 4.5), as well as Asp165-mesaconate anhydride intermediates with free CoA, and a β-

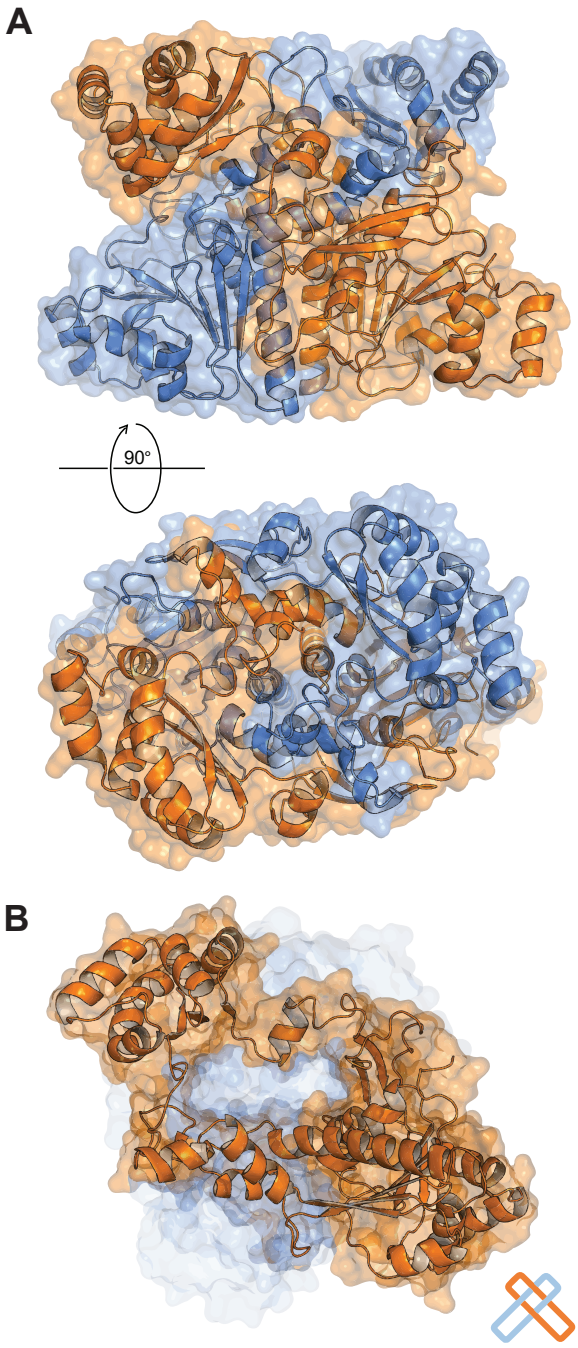


Figure 4.4: **Active site of Mct.** (A) Two subunits shown in orange and blue form an intertwined dimer depicted in cartoon and surface representations. (B) Family III CoA transferases form an interlocked dimer. Shown in surface representation are Mct subunit A in orange and subunit B in transparent blue. The polypeptide chains are interlocked and each is threaded through a hole in the neighboring subunit, as represented by the pictogram in the lower right corner.

aspartyl-CoA intermediate with free mesaconate into the different active sites present in the ASU, respectively (Figure 4.6).

In the active site with bound mesaconyl-C1-CoA, the mesaconyl moiety rests in

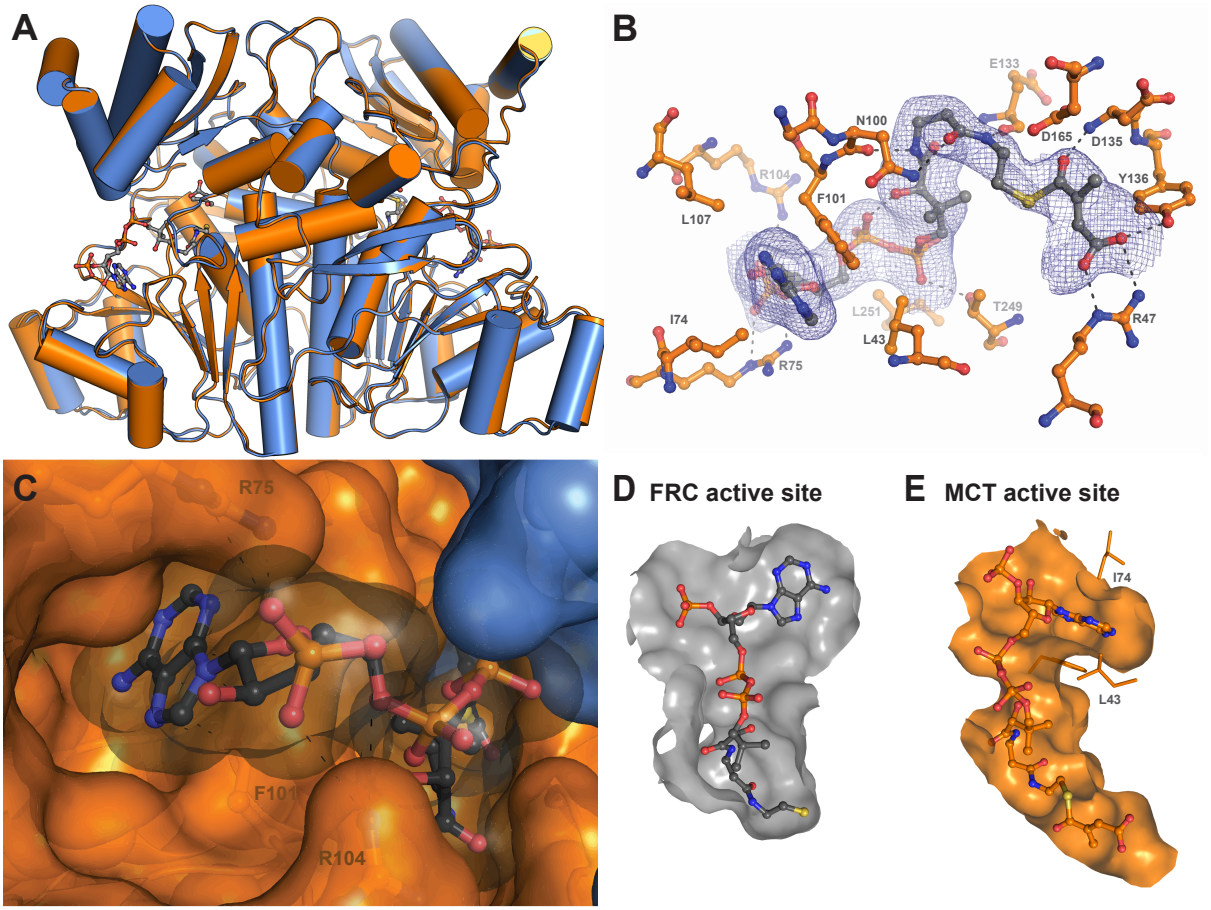


Figure 4.5: **Structure of Mct.** (A) Overlay of the apo form (blue) of Mct and the ligand-bound structure (orange) with an RMSD of 0.267 Å between 675 Cα-pairs. Bound substrates/intermediates are shown (gray) in ball and stick representation. (B) The mesh represents a simulated annealing omit map (Fo–Fc) at 2.0 σ, showing mesaconyl-C1-CoA bound to the active site. Polar interactions with surrounding residues are shown with dashed lines. (C) A transparent surface representation depicts the adenosyl moiety of CoA strongly coordinated at the mouth of the substrate tunnel by F101, R75, and R104, resulting in a cork-like sealing of the active site. (D) Shown is a slice through the active site cavity of the inter-molecular formyl-CoA transferase of *O. formigenes* (Frc) in complex with CoA (PDB 1P5R) and (E) the active site cavity of the intra-molecular Mct of *C. aurantiacus* in complex with mesaconyl-C1-CoA (gray) for comparison. Here L43 constricts the mouth of the active site cavity.

close proximity to the catalytic Asp165. The terminal carboxy group of mesaconyl-CoA is coordinated by Arg47 and Tyr136 (see Figure 4.5B). Notably, Arg47 also oc-

cupies the corresponding space of the flexible glycine-rich loop that is found in some inter-molecular CoA transferases [78,85], preventing conformational changes, such as active-site opening or closing in Mct.

The phosphopantetheine arm of CoA is well coordinated along the active site tunnel of Mct, and the carbonyl-oxygen of the thioester bond engages in a hydrogen bridge with the peptide nitrogen of Asp135. The amide nitrogen, the amide oxygen of the β-alanine, and the cysteamine moiety of CoA are coordinated by the backbone oxygen of Glu133 and the side chain of Asn100, respectively. Arg75 and Arg104 coordinate with the phosphate of the adenosyl group. The adenine ring itself is wedged in between Phe101 and Ile74, engaging in a staggered π-stack with the phenylalanine (Figure 4.5B). Notably, the adenosyl group of CoA adopts a different, perpendicular ("kinked") orientation to what is found in the other family III/Frc family enzymes (Figure 4.5D,E) [78,84–86]. In addition to this difference in adenine binding, Mct also harbors Leu43, which narrows the entrance to the active site of Mct substantially compared to inter-molecular CoA transferases (Figure 4.5D,E). A leucine or isoleucine residue in this position is conserved in all CoA transferases that catalyze intra-molecular CoA transfer, i.e., Mct from *C. aurantiacus*, *R. castenholzii* [89], Candidatus Accumulibacter phosphatis [1], and the γ1-endosymbiont of the gutless worm *Olavius algarvensis* [95]. Overall, the tight binding of CoA along the substrate tunnel together with kinking of the adenine prevents trapped molecules from escaping and other molecules from entering the active site (Figure 4.5B-E), effectively sealing the catalytic site in a "cork-like" fashion.

Importantly, the CoA moiety also plugs those active sites, in which mesaconate is covalently bound to Asp165, indicating that the CoA moiety does not exchange during catalysis, which is consistent with our biochemical observations. This is also supported by previous experiments that concluded through radioactive labeling that no external mesaconate was involved in the reaction mechanism of Mct [18]. Notably, electron densities in different active sites in the ASU allowed us to place the

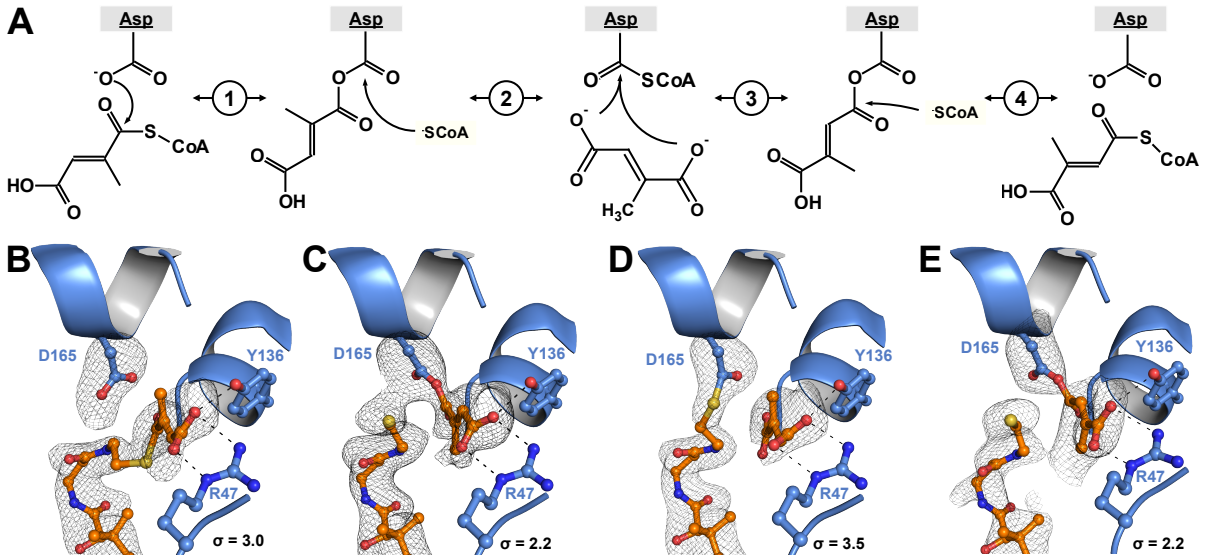


Figure 4.6: **Proposed reaction mechanism of Mct.** (A) An aspartate residue attacks the thioester bond of mesaconyl-C1-CoA (1). The free CoA attacks the mesaconyl-C1-aspartate and liberates mesaconate from the anhydride bond with Asp165 (2). Mesaconate flips around in the active site. The C4-carboxyl group attacks the aforeformed β-aspartyl-CoA (3). Finally, free CoA attacks the mesaconyl-C4-anhydride bond yielding mesaconyl-C4-CoA (4). (B–E) Crystallographically identified intermediates of the Mct reaction (PDB 8APQ). Simulated annealing omit maps (Fo–Fc) are shown as black mesh and the respective σ values are given. (B) Mesaconyl-C1-CoA was observed in chain A, (C) mesaconyl-C1-aspartate and CoA in chain E, (D) β-aspartyl-CoA and mesaconate in chain D, and (E) mesaconyl-C4-aspartate and CoA in chain F.

Asp165-mesaconate anhydride in the C1- as well as the C4-bound orientation (Figure 4.6). We did not observe any electron density that would accommodate an additional mesaconate molecule in any of the active sites. Altogether, these structures suggest that the intra-molecular transfer follows a similar mechanism as canonical inter-molecular, family III transferases that work with two distinct substrates - except that the substrate is not exchanged and may passively re-orient itself during catalysis. A small pocket around Arg47 appears large enough for mesaconate to change orientation randomly. Supporting this hypothesis, in one active site, we actually observed clear electron density for a β-aspartyl-CoA intermediate at Asp165 and a free mesaconate molecule in the aforementioned pocket (Figure 4.6D).

In summary, our structure with bound reaction intermediate states provides additional evidence and an explanation of how Mct catalyzes the intra-molecular CoA transfer favoring it over an inter-molecular transfer. Note that we did not observe any significant conformational changes between the apo form and the substrate-bound form of Mct (Figure 4.5A). This steric constraint of the apoenzyme together with the tight binding of CoA may effectively prevent access to the active site (Figure 4.5B-E), explaining how Mct is able to exclude other CoA acceptor carboxylic acids during the interconversion of the two forms of mesaconyl-CoA. Based on our crystal structures with different substrate-bound states (Figure 4.6B-E), Mct follows the canonical reaction mechanism proposed for the other family III CoA transferases (Figure 4.6A) with neither conformational changes nor exchanges of substrates taking place during the reaction.

4.5 Discussion

Here, we biochemically and structurally characterized Mct, an unusual family III/Frc family CoA transferase that catalyzes an intra-molecular CoA transfer. Our structure with covalently enzyme-bound intermediates provides evidence that the enzyme

follows the mechanism for inter-molecular family III/Frc family CoA transferases as proposed by Berthold et al. [78] Based on our data, we suggest that upon mesaconyl-CoA entering the active site, Asp165 attacks the thioester bond, forming a mesaconyl-C1-aspartate anhydride and free CoA. The Asp165-bound mesaconate is displaced by an attack of the free CoA, resulting in a β-aspartyl-CoA, and releasing mesaconate into the active site cavity, where it can freely rotate within an extended pocket close to the catalytically active aspartate residue. At this step, any of the two carboxyl groups of mesaconate can attack the aspartyl-CoA, yielding either mesaconyl-C1-CoA or mesaconyl-C4-CoA.

The proposed reaction mechanism alone, however, does not explain why the reaction is specific for an intra-molecular transfer and how CoA transfer to other acceptor acids is prevented. The tight coordination of the CoA moiety effectively closes the active site and leads to an enclosed, "corked-up" reaction chamber, excluding diffusion of molecules in or out of the active site. Additionally, we did not observe any significant conformational changes in our two crystal structures, which could allow mesaconate to leave the active site or other acceptor acids to enter. While it could be in principle possible that an alternative acceptor acid may become trapped in the active site before mesaconyl-CoA or another CoA donor threads into the active site tunnel, this seems to be an unlikely event, as our assays with alternative acceptor acids showed that inter-molecular transfer is extremely rare and only takes place at very high, nonphysiologically relevant concentrations of these acids. Such a trapped acceptor molecule could interfere with the re-orientation of the mesaconate released from mesaconyl-CoA, rather resulting in the re-formation of the mesaconyl-CoA than of an alternative CoA thioester. Interestingly, not all tested potential acceptor acids could serve as a substrate. In particular, acetate that should be small enough to occupy the active site cavity was not used by Mct. On the other hand, succinate was accepted in the presence of varying CoA donors. Yet, Mct showed only poor catalytic efficiency (at least 6 orders of magnitude lower than for the intra-molecular CoA transfer) with succinate

as the CoA acceptor. Preventing the diffusion of substrates in and out of the active site is likely the reason why the Mct reaction proceeds so fast in comparison to the intermolecular transfers catalyzed by the other family III CoA transferases [78,88,96–100]. In conclusion, our data provide detailed molecular insights into the structural and mechanistic differences between intra- and inter-molecular family III CoA transferases, explaining how "corking up" the active site with the CoA substrate allows Mct to achieve excellent selectivity toward its native substrates, efficiently preventing unwanted side reactions.

4.6 Author contribution

P.P. and J.Z. conceived the work. P.P. designed and performed experiments and analyzed the data together with J.Z. and T.J.E. P.P. purified and crystallized Mct. P.P. and J.Z. solved and refined the crystal structures. P.P., J.Z., and T.J.E. wrote the manuscript.

4.7 Acknowledgments

The authors thank Nicole Paczia and Peter Claus for HPLC-MS analytics. The authors acknowledge support from the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and Didier Nurizzo for assistance at beamline ID29. The authors thank Marieke Scheffen for collecting X-ray data and Guillaume Pompidor for assistance at beamline P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). The authors thank Christoph Diehl for providing Ssr and purified succinyl-CoA. This work was funded by the Max Planck Society.

4.8 Supporting Information

4.8.1 Supporting Results

Reliability of mesaconyl-CoA separation

Stored samples of mesaconyl-CoA were tested for their purity via HPLC-MS as described in the methods section. The sample of pooled mesaconyl-C1-CoA contained less than 0.3% of mesaconyl-C4-CoA judged by relative ion count (Fig. 4.7). The sample of mesaconyl-C4-CoA contained no detectable contamination of mesaconyl-C1-CoA judged by relative ion count

Mesaconyl-CoA hydrolysis

Mesaconyl-CoA derivatives were incubated in previous described assay matrix (300 µL containing 500 µM CoA, 200 mM HEPES, $pH_{55\,°C}$ 7.5) at 55 °C for 180 hours. Mesaconyl-C1-CoA was additionally incubated in presence of 3 µg Mct. At different time points, CoA hydrolysis was monitored by measuring the remaining absorbance at 290 nm. Samples were checked for alternative byproducts via HPLC-MS. The only detectable products were free CoA or dimerized CoA. The hydrolysis half-lives (Figure 4.7B) were longer than 30 h and 14 h for mesaconyl-C1-CoA and mesaconyl-C4-CoA, respectively. When mesaconyl-C1-CoA was incubated with Mct, the hydrolysis half-life was similar to that of mesaconyl-C4-CoA (about 12 h), within the margin of error. Here, Mct produced mesaconyl-C4-CoA, which then hydrolyzes faster than mesaconyl-C1-CoA alone. The uniform hydrolysis of mesaconyl-CoA in the presence of Mct with similar rates as mesaconyl-C4-CoA alone suggests that Mct remained active throughout the whole experiment. Judging from these results, the effect of mesaconyl-CoA instability on our enzymatic assays was negligible.

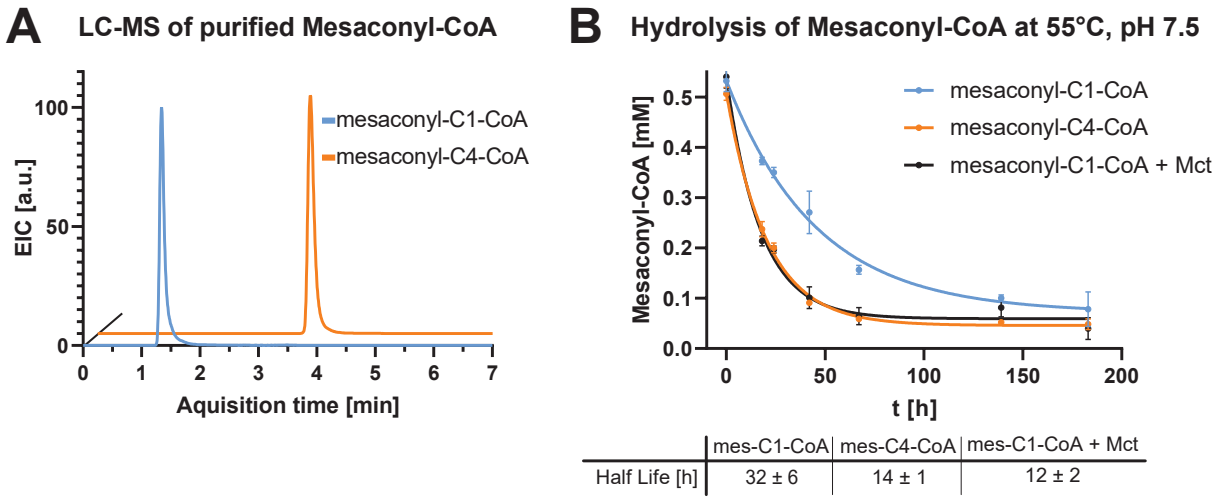


Figure 4.7: **Purity and stability of mesaconyl-CoAs.** (A) HPLC-MS EIC chromatograms of purified mesaconyl-C1-CoA (blue) and mesaconyl-C4-CoA (orange). (B) Mesaconyl-CoA hydrolysis over time. Mesaconyl-C1-CoA alone (blue) has a 2-fold higher half-life then mesaconyl-C4-CoA (orange) and mesaconyl-C1-CoA incubated with Mct (black) at 55 °C pH 7.5. "±" indicates the 95% confidence interval.

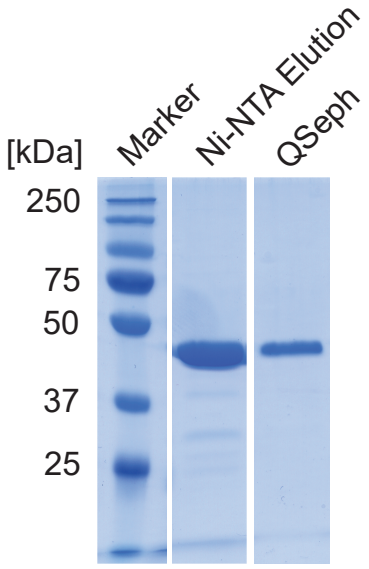


Figure 4.8: **SDS-PAGE results of purification steps of Mct.** Precision Plus Protein Dual Color (Bio-rad, USA) was used as marker (left lane) for size determination. Enriched Mct (10 µg) after Ni-NTA affinity purification (middle lane) already shows as dominant band at 45 kDa. After the anion exchange (right lane, 6 µg), the only visible band left is at around 45 kDa. Mct has a calculated mass of 44.8 kDa per monomer.

Enzyme purity

Protein purity was assessed by SDS-PAGE. Stored samples of the enzyme after Ni-NTA affinity chromatography, as well as after anion exchange chromatography were

boiled in 4× SDS-loading buffer (0.2 M Tris-HCl, 0.4 M DTT, 277 mM 8.0% (w/v) SDS, 6 mM Bromphenol blue, 4.3 M glycerol) for 5 minutes at 100 °C. After staining the gel for 10 minutes with GelCode™ Blue Safe staining solution (Thermo Fisher, USA) it was destained with ddH2O. After Ni-NTA affinity chromatography, a band corresponding to the size of Mct was already dominant on the gel. A few bands not corresponding to Mct could be removed by anion exchange chromatography using Q-sepharose.

4.8.2 Supporting Tables

CoA-transferase family*	Canonical nomenclature	Pfam
Cat1	Family I	13336, 02550
OXCT1	Family I	01144
Gct	Family I	01144
CitF	Family II	04223
McdA	Family II	16957
Frc	Family III	02515

Table 6: Nomenclature and PFAM IDs for the different CoA transferase families. *Proposed classification according to Hackmann, 2022.

Chapter 5

Summarizing Discussion

5.1 Realizing the Photorespiratory 3OHP Bypass by Strain Evolution

In a recent study, Shih et al. [1] introduced genetic constructs for the 3OHP Bypass into the cyanobacterium *S. elongatus*. They integrated the genes, originating from *C. aurantiacus* into the wild type strain *S. elongatus* PCC7942. Catalytic activities for all enzymes but the Acc were detectable. The resulting strain showed no clear growth benefit or disadvantage. This suggested that the Chloroflexus enzymes were not interfering with the host's metabolism too much and it should be possible to establish the pathway in this cyanobacterium. We used the carboxysome knockout, generated by Cameron et al. [40]. We also identified mesophilic alternatives for the original *Chloroflexus* enzymes. The replaced isoenzymes for Mcl, Mch, Pcs were found in *A. phosphatis*, *R. sphaeroides* and *E. sp.* NAP1, respectively [37]. For Acc we found a set of 3 candidates; the four subunit Acc of *E. coli* and *S. elongatus* and the two subunit Pcc with a D407I substitution, that lowered specificity to propionyl-CoA and increased activity with acetyl-CoA. These candidates were tested in *E. coli* cell extracts in case of Accs and as purified enzyme in case of Pcc. All of these candidates showed promising activity, but unfortunately after we introduced the corresponding genes into *S. elongatus*, we were unable to detect any increase in Acc activity in cell extracts. We know that Acc in cyanobacteria is tightly regulated by a central regulatory protein, P_{II} [36], that is involved in quasi-essential regulation cascades. Nitrogen metabolism, carbon drain from upper glycolysis and most noteable for our aim, fatty acid biosynthesis. Lipid body accumulation in a P_{II} knockout from Hauf, et al. [36], suggests that indeed the interaction is required for the down-regulation of fatty acid production, of which Acc catalyzes the first committed step. P_{II} also interacts with PirC [35], that is directly affecting actyl-CoA levels, so the unleashed activity of Acc alone, might not be responsible for the accumulation of lipids in this mutant strain. However, its contribution to the phenotype is indisputable.

P_{II} cannot not be deleted, as it is to involved in cellular functions and the resulting strain would be too unstable to sustain a full synthetic pathway. A more subtle approach would be the modification of BCCP. Due to lack of structural data, the molecular basis for the P_{II}-BCCP interaction remain elusive for now, impeding a rational approach to change the binding of the two proteins. Therefore we devised an an approach, that could result in increased Acc activity, exploring a broad range of possible mutations and adaptations.

In our first ALE iteration we only focused on a subsection of the bypass that did not need to rely on Acc activity and would allow for a relatively short generation time, which would be crucial to accumulate enough beneficial mutations. Therefore, we made use of the described toxicity of propionate in *S. elongatus* [68]. We suspected, that propionate would be co-metabolized with 2PG in a strain with the 3OHP Bypass, and Acc activity was not required to generate additional central carbon metabolites, like pyruvate and acetyl-CoA in this case. After integration of a construct containing only genes encoding Mcl, Meh, Mct and Mch, activities of all enzymes were verified by HPLC assays in cell extracts of the generated strain (Figure 3.2). We tested effects of propionate on a strain with the first required steps of the bypass. And in fact, a distinct growth benefit was detectable, when the transformants were grown with propionate (Figure 3.3). Even though the strain was able to use propionate to detoxify some of the emerging 2PG, it did not grow in CO_2 levels below 0.5% (Figure 3.4).

The strain was cultivated in a continuous turbidostat bioreactor to adapt to even lower CO_2 levels, thus further increasing photorespiratory pressure to use the bypass. The feedstock contained standard minimal growth medium (BG-11) and 5 mM sodium propionate As soon as a steady state was reached, the CO_2 level was decreased to 0.4%, 0.3%, and ultimately 0.2% within the matter of 120 days. At the end, the cells maintained growth at 0.2% of CO_2, which is a noteable improvement as the parent strain (ΔK-O) below 0.5%. To complete our experiment, the resulting population will be characterized by sequencing, and growth characterisation at different propionate

concentrations. Even though a final evaluation remains to be performed, we consider our selection strain a suitable vehicle to use propionate mediated photorespiration, that could ultimately enable us to realize the complete 3OHP Bypass in *S. elongatus*, and to disrupt canonical photorespiration pathways in this host as well. This might also allow *S. elongatus* to even grow at lower CO_2 concentrations in the absence of carboxysomes. The establishment of the whole 3OHP bypass may be again done in a stepwise fashion, first only adding the gene encoding Pcs and feeding 3OHP instead of propionate, ensuring Pcs is sufficiently active before adding the next parts of the cycle, i.e. Mcr and Acc. These experiments are planned for the near future and will rely on the presented findings and the here obtained selection strain.

5.2 The Role of Mct for the 3-hydroxypropionate Bypass

A heat precipitation step was crucial for the purification and activity of Mct from *C. aurantiacus*. In purification batches, in which the heat precipitation was performed at 60 °C instead of 70 °C, the turnover rate was found to be reduced from 370 s^{-1} to 100 s^{-1}. As Mct is usually produced at around 55 °C in optimally growing *Chloroflexi*, it might not fold correctly at 30-37 °C, temperatures at which our cyanobacterial strains are usually growing. We hypothesize, that misfolded protein correctly folds at 70 °C, or denatures, leaving only correctly folded enzyme in solution. In approaches to introduce the 3OHP Bypass into phototrophic organisms, it will virtually exclusively be used in mesophilic conditions. The drop in turnover rate, regardless of its cause, should be accounted for in experimental design.

Nonetheless, Mct's unique mechanism for the intramolecular CoA-transfer of mesaconyl-CoA derivatives lead us to investigate. Other CoA transferases described to perform intramolecular transfer, however, we found that these enzymes usually also accept external acids, like succinate [78, 81]. As presented (see 4.4.2), CaMct is a genuine intramolecular CoA transferase, that does not accept readily external mesaconate. The

structural data we present, strongly suggests, that Mct has a more narrow active site cavity, that is plugged by the CoA moiety itself. That way no external acid can enter or leave the active site during catalysis.
We also propose, that the reaction mechanism is canonical to CoA transferases of Family III, and that a previously proposed mechanism was not correct [89]. We support our claim by resolving all intermediate steps of this reaction mechanism in our crystal structure.
The implication of this is that Mct is indeed an enyzme with a to date unique catalytic mechanism, that transfers the CoA moiety from one of the carboxyl groups to the other in the same mesaconate molecule.

5.3 CETCH Cycle Reactions Inspired an Acc Independent Photorespiratory Cycle

Beyond the scope of the 3OHP bypass for photorespiration, there are other strategies aiming at increasing biological CO_2 fixation. One of these strategies is to replace the natural CBB cycle with a new-to-nature artificial pathways like the CETCH cycle. These pathways are typically described as modules. The CETCH cycle is a CO_2 fixation module that generates gloxylate as primary end product. The TaCo pathway [5] or the 3OHP Bypass are modules for the carbon positive assimilation of glyoxylate into central carbon metabolites. Modularity indicates a a plug-and-play system, with an interface that both modules can communicate over, in the case of metabolism an intermediate like glyoxylate. That this modularity is possible with the CETCH cycle for *in vitro* reconstituted pathways has recently been shown [4]. Also the combination of the CETCH cycle with the TaCo pathway has been validated [5]. The combination with the 3OHP Bypass would also be possible, yet requires a cetrain amount of careful precautions.
As described in chapter 2 for SucD, the CETCH cycle enzyme has crossreactive side

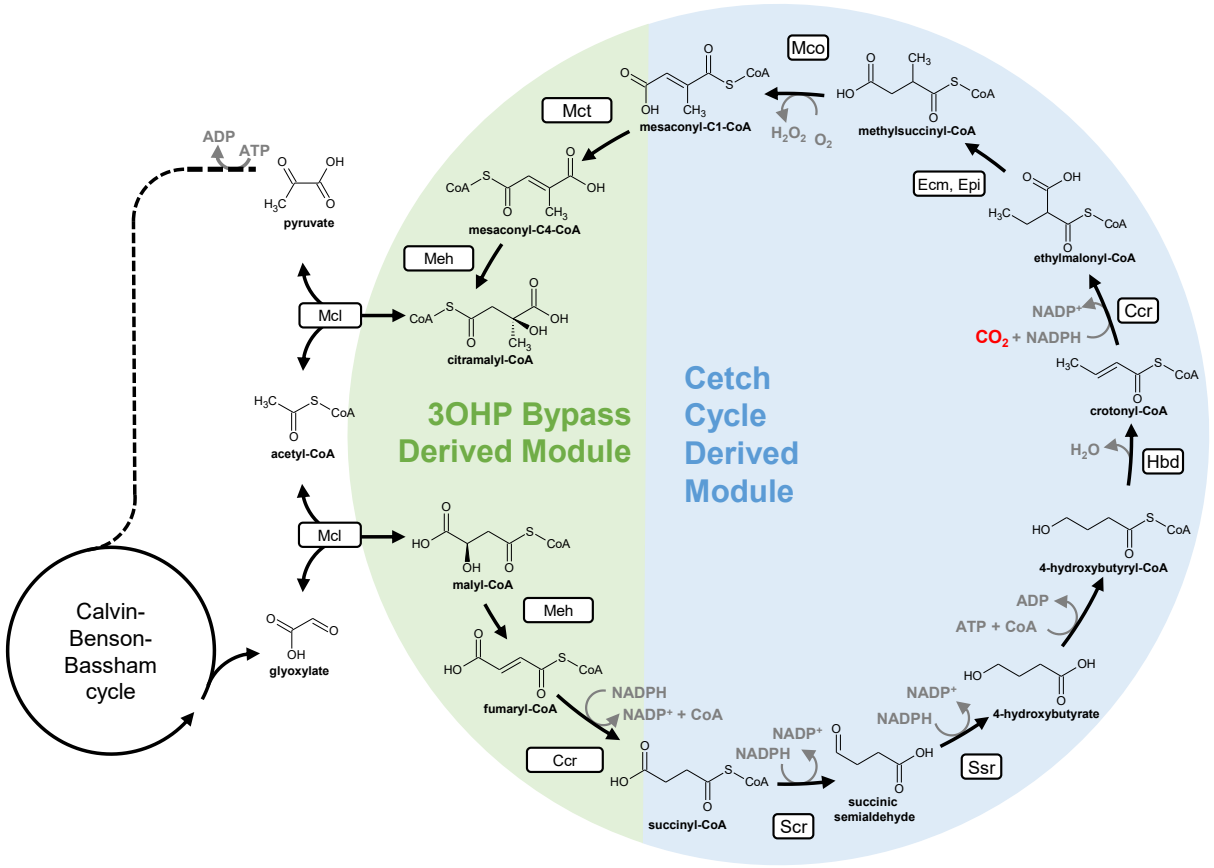


Figure 5.1: **Combination of 3OHP bypass and the CETCH cycle via the fumaryl-CoA (FuCo) shunt.** β-methylmalyl-CoA/(S)-citramalyl-CoA lyase (Mcl), mesaconyl-C4-CoA hydratase (Meh), mesaconyl-C1-C4-CoA CoA transferase (Mct), succinyl-CoA reductase (ssr), succinic semialdehyde reductase (ssr), 4-hydroxybutyryl-CoA synthase (Hbs), 4-hydroxybutyryl-CoA dehydratase, (Hbd), crotonyl-CoA carboxylase/reductase (Ccr), ethylmalonyl-CoA mutase and epimerase (Ecm, Epi), methylsuccinyl-CoA oxidase (Mco).

activities. From the initial design, the CETCH cycle already features Mcl, Mch and their intermediates, like the 3OHP bypass. If the CETCH cycle were to be extended by the 3OHP bypass to assimilate glyoxylate into pyruvate and adding another CO_2 fixing step, Mct would affect pools of mesaconyl-C1-CoA, that could lead to draining of CETCH intermediates. The in the following proposed FuCo shunt (Figure 5.1) could serve as a anapleirotic pathway mitigating this problem [101]. It is based on the reactions catalyzed by Mcl, Meh and Ccr to convert acetyl-CoA and gloxylate to malyl-CoA, fumaryl-CoA and ultimately succinyl-CoA. As the enzymes for these reaction are present in both pathways, it would not require any additionaly integration.

The anapleirotic process would occur as a self correcting pathway without additional efforts.

The FuCo shunt could therefore serve as a photorespiration pathway that uses the reactions of the CETCH cycle to detoxify 2PG to succinyl-CoA, and the arrive at mesaconyl-C1-CoA. Mct and Mcl would subsequently lead to the production of pyruvate as final carboxylation product and acetyl-CoA to re-enter the FuCo shunt again. This pathway would be another way to overcome the problem of tight Acc regulation in *S. elongatus*. Even though this pathways involves 11 instead of just 7 exogenic enzymes, the benefits would be, that Ccr activity was found to be easily detectable and thus viable to use in cyanobacteria [101].

5.4 Combining the CETCH Cycle and 3OHP Bypass as a Whole New Carbon Fixation Module

The CETCH cycle is designed to produce glyoxylate as direct carboxylation product. The 3OHP Bypass is designed to assimilate glyoxylate to central carbon metabolites, while perfoming an additional carboxylation. Combining both pathways for an ultimately efficient carbon assimilation strategy seems obvious. As discussed before, this coincides with a shared reactions and intermediates, that could be alleviated by anapleirotic pathways. Even though glyoxylate seems the connecting intermediate for such a hybrid pathway, it might be possible to avoid glyoxylate and hence the shared reactions completely. Reactions of Mcl and Mch are starting from mesaconyl-C1-CoA in the CETCH cycle, while yielding mesaconyl-C1-CoA in the 3OHP Bypass. Excluding the shared reactions would form a glyoxylate free CO_2-fixation cycle (Figure 5.2), that produces pyruvate from CO_2 via 2 enoyl-CoA carboxylase/reductase (Ecr) reactions and the carboxylation trough Acc. This pathways would involve even two reaction steps less than the 3OHP bi-cycle, and has interfaces with central metabolism on many levels. First of all, pyruvate could be converted to succinyl-CoA, with yet

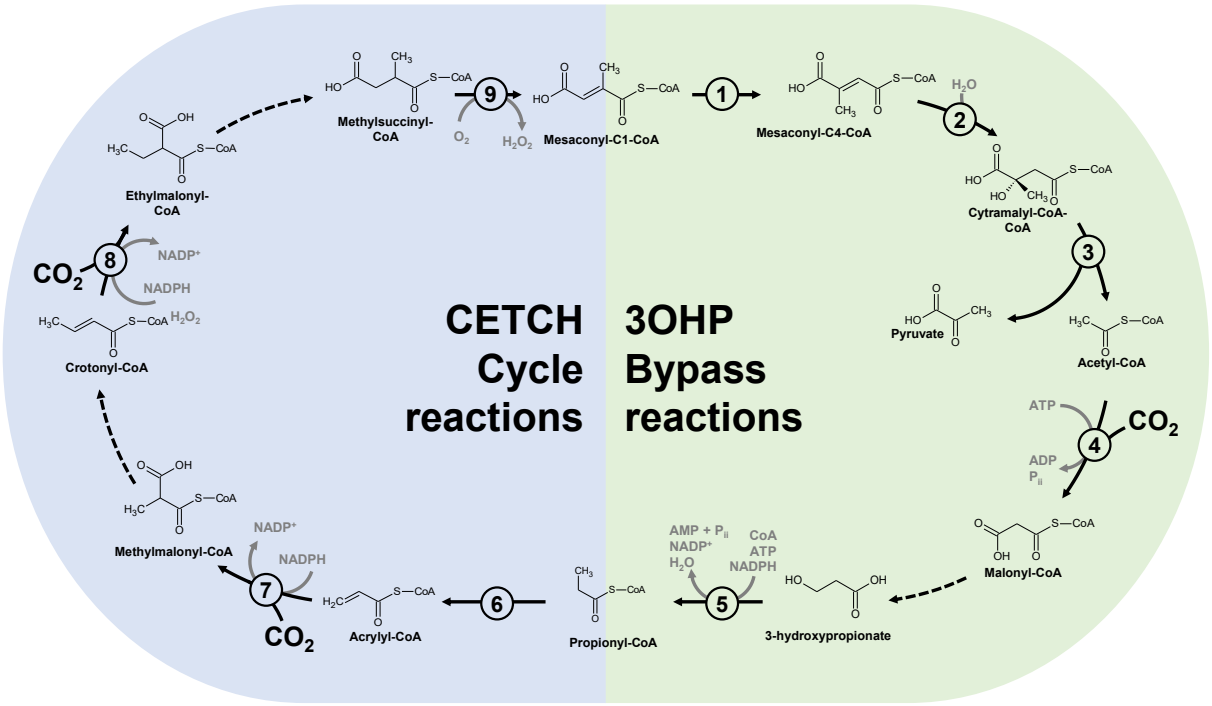


Figure 5.2: **Combination of 3OHP bypass and the CETCH cycle.** mesaconyl-C1-C4-CoA CoA transferase (1), mesaconyl-C4-CoA hydratase (2), β-methylmalyl-CoA/(S)-citramalyl-CoA lyase (3), acetyl-CoA carboxylase (4), propionyl-CoA synthase (5), propionyl-CoA oxidase (6), enoyl-CoA carboxylase/reductase (7,8), methylsuccinyl-CoA oxidase (9). CETCH reactions are highlighted blue, 3OHP Bypass reactions are highlighted green. Dashed lines represents several reaction steps, that were shown before.

another carboxylation by Phosphoenolpyruvate carboxylase (PEPC) or via pyruvate carboxylase. This autoregeneration would lead to a very high pathway stability, the insensitivity against metabolites beeing removed from the active pool trough sidereactions like SucD forming aldehydes from CoA esters. It was discussed that cyanobacteria have a conservative glycogen storage system to quickly replenish CBB cycle intermediates. Bacteria relying on PHB as storing compount instead of glycogen could benefit from such an CO_2-fixation metabolism, as acetyl-CoA, the initial monomer of PHB could enter the pathway. This new CETCH-3OHP cycle does not seem to be without merit and underscores the modularity of natural as well as artificial metabolism.

5.5 Structural Biology as a Method to Extend Enzyme Solution Space in Synthetic Metabolism

The here presented findings have a strong focus on structural biology, and enzymatic reaction mechanism. Synthetic pathways are heavily relying on effective enzymatic catalysts. Structural information can provide an educated idea of a biochemical property. Even though it is possible to use kinetic data, directed mutagenesis, high throughput screening and selection schemes alone to improve enzymatic properties, the information of structural data goes beyond the implications for a certain application. Structural data is shared with the whole scientific community, helps to understand evolution and can be used to train AI based prediction tools to provide structural and catalytic information for enzymes from non-model organisms.

Mct is a quite efficient enzyme catalyst. Its activity for mesaconyl-C1-CoA and mesaconyl-C4-CoA is 2.3×10^6 and 1.6×10^6, respectively. There are no notable sidereactions and if so, it would also catalyze the reaction reversibly. For CkSucD the catalytic efficiency is lower, with 4.1×10^5 for succinyl-CoA and 6.7×10^4 for mesaconyl-C1-CoA. CkSucD is a efficient catalyst as well, but its notably high promiscuity to mesaconyl-C1-CoA affect the synthetic pathways it is used in.

In chapter 2, the enzyme structure and active site residues of SucDs are discussed. Its promiscuity for mesaconyl-C1-CoA was described in crystal structure and follow up characterizations of the wild type and mutant versions. Based on these findings, a small set of residue changes were found to lead to a noteable increase in speficity of SucD. A change of Lys70 for CkSucD and Lys79 for CdSucD to an arginine residue resulted in an improvement from 16 to 2% for CkSucD and 12 to 2% for CdSucD. The new SucD variants that were generated might improve a future iteration of the CETCH cycle by a notable margin, but the impact of the enzyme structures (PDB 8CEK, 8CEI, 8CEJ) are beyond this rather narrow field of use. The structures are the first ones of clostridial SucDs and among they relative aldehyde dehydrogenases

(<30% Identities) it is the only one next to PduP [51], that was crystalized with ligands. In scope of the CETCH cycle, the obvious goal is to generate a more specific SucD variant. In the scope of new-to-nature pathways, there might be a requirement to also reduce other dicarboxylic acids, or their corresponding CoA esters, with larger, bulkier side chains. Our data suggest, that the respective lysine could also be replaced with a smaller residue like asparagine or alanine, to make it more efficient for mesaconyl-C1-CoA or other bulkier acyl-CoAs.

In any case, this work once more highlights how pathway design and implementation goes hand in hand with careful biochemical and structural characterizations of enzymes, sometimes requiring to employ different strategies to optimize enzyme function and pathway integration.

Bibliography

[1] P. M. Shih, J. Zarzycki, K. K. Niyogi, and C. A. Kerfeld. Introduction of a synthetic CO_2-fixing photorespiratory bypass into a cyanobacterium. *J. Biol. Chem.*, 289:9493, 2014.

[2] IPCC. *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Summary for Policymakers. Cambridge University Press, Cambridge, UK and New York, USA, 2022.

[3] T. Schwander, L. Schada von Borzyskowski, S. Burgener, N. S. Cortina, and T. J. Erb. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science*, 354:900, 2016.

[4] C. Diehl, P. D. Gerlinger, N. Paczia, and T. J. Erb. Synthetic anaplerotic modules for the direct synthesis of complex molecules from CO_2. *Nat Chem Biol*, 19(2):168–175, Feb 2023.

[5] M. Scheffen, D. G. Marchal, T. Beneyton, S. K. Schuller, M. Klose, C. Diehl, J. Lehmann, P. Pfister, M. Carrillo, H. He, S. Aslan, N. S. Cortina, P. Claus, D. Bollschweiler, J.-C. Baret, J. M. Schuller, J. Zarzycki, A. Bar-Even, and T. J. Erb. A new-to-nature carboxylation module to improve natural and synthetic CO_2 fixation. *Nature Catalysis*, 4(2):105–115, January 2021.

[6] I. A. Berg. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Appl. Environ. Microbiol.*, 77(6):1925–1936, Mar 2011.

[7] J. A. Raven. Rubisco: still the most abundant protein of Earth? *New Phytol*, 198(1):1–3, Apr 2013.

[8] A. Bar-Even, E. Noor, and R. Milo. A survey of carbon fixation pathways through a quantitative lens. *J Exp Bot*, 63(6):2325–2342, Mar 2012.

[9] A. Pohlmann, W. F. Fricke, F. Reinecke, B. Kusian, H. Liesegang, R. Cramm, T. Eitinger, C. Ewering, M. Pötter, E. Schwartz, A. Strittmatter, I. Voss, G. Gottschalk, A. Steinbüchel, B. Friedrich, and Botho Bowien. Genome sequence of the bioplastic-producing "knallgas" bacterium *Ralstonia eutropha* h16. *Nature Biotechnology*, 24:1257–62, 11 2006.

[10] S. H. Kopf and D. K. Newman. Photomixotrophic growth of *Rhodobacter capsulatus* SB1003 on ferrous iron. *Geobiology*, 10(3):216–222, May 2012.

[11] H. D. Holland. The oxygenation of the atmosphere and oceans. *Philos Trans R Soc Lond B Biol Sci*, 361(1470):903–915, Jun 2006.

[12] N. Nelson and A. Ben-Shem. The complex architecture of oxygenic photosynthesis. *Nat Rev Mol Cell Biol*, 5(12):971–982, Dec 2004.

[13] O. Warburg. *Über die Geschwindigkeit der photochemischen Kohlensäurezersetzung in lebenden Zellen*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1928.

[14] C. H. Foyer, A. J. Bloom, G. Queval, and G. Noctor. Photorespiratory metabolism: genes, mutants, energetics, and redox signaling. *Annu Rev Plant Biol*, 60:455–484, 2009.

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

[16] A. Turmo, C. R. Gonzalez-Esquer, and C. A. Kerfeld. Carboxysomes: metabolic modules for CO2 fixation. *FEMS Microbiol Lett*, 364(18), Oct 2017.

[17] P. F. South, A. P. Cavanagh, H. W. Liu, and D. R. Ort. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, 363(6422):eaat9077, Jan 2019.

[18] J. Zarzycki, V. Brecht, M. Müller, and G. Fuchs. Identifying the missing steps of the autotrophic 3-hydroxypropionate CO_2 fixation cycle in *Chloroflexus aurantiacus*. *Proc. Natl. Acad. Sci. U.S.A.*, 106:21317, 2009.

[19] I. Bernhardsgrütter, B. Vögeli, T. Wagner, D. M. Peter, N. S. Cortina, J. Kahnt, G. Bange, S. Engilberge, E. Girard, F. Riobé, O. Maury, S. Shima, J. Zarzycki, and T. J. Erb. The multicatalytic compartment of propionyl-CoA synthase sequesters a toxic metabolite. *Nat Chem Biol*, 14(12):1127–1132, Dec 2018.

[20] I. Bernhardsgrütter, K. Schell, D. M. Peter, F. Borjian, D. A. Saez, E. Vöhringer-Martinez, and T. J. Erb. Awakening the sleeping carboxylase function of enzymes: Engineering the natural CO_2-binding potential of reductases. *J Am Chem Soc*, 141(25):9778–9782, Jun 2019.

[21] Q. Wang, C. Liu, M. Xian, Y. Zhang, and G. Zhao. Biosynthetic pathway for poly(3-hydroxypropionate) in recombinant *Escherichia coli*. *J Microbiol*, 50(4):693–697, Aug 2012.

[22] C. Rathnasingh, S. M. Raj, Y. Lee, C. Catherine, S. Ashok, and S. Park. Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant *Escherichia coli* strains. *J Biotechnol*, 157(4):633–640, Feb 2012.

[23] T. J. Erb, G. Fuchs, and B. E. Alber. (2S)-Methylsuccinyl-CoA dehydrogenase closes the ethylmalonyl-CoA pathway for acetyl-CoA assimilation. *Mol Microbiol*, 73(6):992–1008, Sep 2009.

[24] A. Pandi, C. Diehl, A. Yazdizadeh Kharrazi, S. A. Scholz, E. Bobkova, L. Faure, M. Nattermann, D. Adam, N. Chapin, Y. Foroughijabbari, C. Moritz, N. Paczia,

N. S. Cortina, J. L. Faulon, and T. J. Erb. A versatile active learning workflow for optimization of genetic and metabolic networks. *Nat Commun*, 13(1):3876, Jul 2022.

[25] L. Schada von Borzyskowski, F. Severi, K. Krüger, L. Hermann, A. Gilardet, F. Sippel, B. Pommerenke, P. Claus, N. S. Cortina, T. Glatter, S. Zauner, J. Zarzycki, B. M. Fuchs, E. Bremer, U. G. Maier, R. I. Amann, and T. J. Erb. Marine proteobacteria metabolize glycolate via the β-hydroxyaspartate cycle. *Nature*, 575(7783):500–504, 2019.

[26] J. Zarzycki and G. Fuchs. Coassimilation of organic substrates via the autotrophic 3-hydroxypropionate bi-cycle in *Chloroflexus aurantiacus*. *Appl. Environ. Microbiol.*, 77:6181, 2011.

[27] A. Taton, C. Erikson, Y. Yang, B. E. Rubin, S. A. Rifkin, J. W. Golden, and S. S. Golden. The circadian clock and darkness control natural competence in cyanobacteria. *Nat Commun*, 11(1):1688, Apr 2020.

[28] N. F. Tsinoremas, M. R. Schaefer, and S. S. Golden. Blue and red light reversibly control *psbA* expression in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J Biol Chem*, 269(23):16143–16147, Jun 1994.

[29] T. Omata, G. D. Price, M. R. Badger, M. Okamura, S. Gohta, and T. Ogawa. Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proc Natl Acad Sci U S A*, 96(23):13571–13576, Nov 1999.

[30] A. Makowka, L. Nichelmann, D. Schulze, K. Spengler, C. Wittmann, K. Forchhammer, and K. Gutekunst. Glycolytic Shunts Replenish the Calvin-Benson-Bassham Cycle as Anaplerotic Reactions in Cyanobacteria. *Mol Plant*, 13(3):471–482, Mar 2020.

[31] D. Schwarz, A. Nodop, J. Hüge, S. Purfürst, K. Forchhammer, K. P. Michel, H. Bauwe, J. Kopka, and M. Hagemann. Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the Cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant Physiol.*, 155(4):1640–1655, Apr 2011.

[32] S. Zhang and D. A. Bryant. The tricarboxylic acid cycle in cyanobacteria. *Science*, 334(6062):1551–1553, Dec 2011.

[33] J. T. Broddrick, B. E. Rubin, D. G. Welkie, N. Du, N. Mih, S. Diamond, J. J. Lee, S. S. Golden, and B. O. Palsson. Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proc. Natl. Acad. Sci. U.S.A.*, 113(51):E8344–E8353, 12 2016.

[34] K. Forchhammer, K. A. Selim, and L. F. Huergo. New views on P_{II} signaling: from nitrogen sensing to global metabolic control. *Trends Microbiol*, 30(8):722–735, Aug 2022.

[35] T. Orthwein, J. Scholl, P. t, S. Lucius, M. Koch, B. Macek, M. Hagemann, and K. Forchhammer. The novel P_{II}-interactor PirC identifies phosphoglycerate mutase as key control point of carbon storage metabolism in cyanobacteria. *Proc Natl Acad Sci U S A*, 118(6), Feb 2021.

[36] W. Hauf, K. Schmid, E. C. Gerhardt, L. F. Huergo, and K. Forchhammer. Interaction of the Nitrogen Regulatory Protein GlnB (P_{II}) with Biotin Carboxyl Carrier Protein (BCCP) Controls Acetyl-CoA Levels in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Front Microbiol*, 7:1700, 2016.

[37] P. Pfister. Optimizing the synthetic photorespiratory 3-hydroxypropionate bypass. B.Sc. thesis, Philipps-Universität Marburg, Bibliothek des Fachbereichs Biologie, Karl-von-Frisch-Str. 8, 35037 Marburg, Deutschland, Sep 2016.

[38] P. Pfister. The Photorespiratory 3-Hydroxypropionate Bypass in *Synechococcus elongatus* PCC 7942. M.Sc. thesis, Philipps-Universität Marburg, Bibliothek des

Fachbereichs Biologie, Karl-von-Frisch-Str. 8, 35037 Marburg, Deutschland, Aug 2018.

[39] B. D. Rae, B. M. Long, M. R. Badger, and G. D. Price. -carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS One*, 7(8):e43871, 2012.

[40] J. C. Cameron, S. C. Wilson, S. L. Bernstein, and C. A. Kerfeld. Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell*, 155(5):1131–1140, Nov 2013.

[41] J. Espinosa, M. A. Castells, K. B. Laichoubi, and A. Contreras. Mutations at pipX suppress lethality of P_{II}-deficient mutants of *Synechococcus elongatus* PCC 7942. *J. Bacteriol.*, 191(15):4863–4869, Aug 2009.

[42] K. Forchhammer and N. Tandeau de Marsac. Functional analysis of the phosphoprotein P_{II} (glnB gene product) in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J Bacteriol*, 177(8):2033–2040, Apr 1995.

[43] A. Forcada-Nadal, J. L. Llácer, A. Contreras, C. Marco-Marín, and V. Rubio. The P_{II}-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. *Front Mol Biosci*, 5:91, 2018.

[44] B. Söhling and G. Gottschalk. Molecular analysis of the anaerobic succinate degradation pathway in *Clostridium kluyveri*. *J Bacteriol*, 178(3):871–80, 1996. Söhling, B Gottschalk, G eng Research Support, Non-U.S. Gov't 1996/02/01 J Bacteriol. 1996 Feb;178(3):871-80. doi: 10.1128/jb.178.3.871-880.1996.

[45] P. Hillmer and G. Gottschalk. Particulate nature of enzymes involved in the fermentation of ethanol and acetate by *Clostridium kluyveri*. *FEBS Lett*, 21(3):351–354, 1972.

[46] V. K. Madan, P. Hillmer, and G. Gottschalk. Purification and properties of nadp-dependent l(+)-3-hydroxybutyryl-CoA dehydrogenase from *Clostridium kluyveri*. *Eur J Biochem*, 32(1):51–6, 1973.

[47] B. Söhling and G. Gottschalk. Purification and characterization of a Coenzyme-A-dependent succinate-semialdehyde dehydrogenase from *Clostridium kluyveri*. *Eur. J. Biochem.*, 212(1):121–127, Feb 1993.

[48] U. Scherf, B. Söhling, G. Gottschalk, D. Linder, and W. Buckel. Succinate-ethanol fermentation in *Clostridium kluyveri*: purification and characterisation of 4-hydroxybutyryl-CoA dehydratase/vinylacetyl-CoA delta 3-delta 2-isomerase. *Arch Microbiol*, 161(3):239–45, 1994.

[49] S. Gencic and D. A. Grahame. Diverse Energy-Conserving Pathways in Clostridium difficile: Growth in the Absence of Amino Acid Stickland Acceptors and the Role of the Wood-Ljungdahl Pathway. *J Bacteriol*, 202(20), Sep 2020.

[50] H. Seedorf, W. F. Fricke, B. Veith, H. Bruggemann, H. Liesegang, A. Strittmatter, M. Miethke, W. Buckel, J. Hinderberger, F. Li, C. Hagemeier, R. K. Thauer, and G. Gottschalk. The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. *Proc Natl Acad Sci U S A*, 105(6):2128–33, 2008.

[51] J. Zarzycki, M. Sutter, N. S. Cortina, T. J. Erb, and C. A. Kerfeld. In Vitro Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. *Sci Rep*, 7:42757, 02 2017.

[52] S Luo, C. Diehl, H. He, Y. Bae, M. Klose, P. Claus, N. S. Cortina, C. A. Fernandez, R. McLean, A. A. Ramírez Rojas, D. Schindler, N. Paczia, and T.J. Erb. Construction and modular implementation of the synthetic theta cycle for CO_2-fixation in vitro and in vivo. in press, Jan 2023.

[53] S. Sundaram, C. Diehl, N. S. Cortina, J. Bamberger, N. Paczia, and T. J. Erb. A modular in vitro platform for the production of terpenes and polyketides from CO_2. *Angew Chem Int Ed Engl*, 60(30):16420–16425, 2021.

[54] A. R. Shenoy and S. S. Visweswariah. Site-directed mutagenesis using a single mutagenic oligonucleotide and Dpni digestion of template DNA. *Anal Biochem*, 319(2):335–6, 2003.

[55] P. Pfister, J. Zarzycki, and T. J. Erb. Structural basis for a cork-up mechanism of the intra-molecular mesaconyl-CoA transferase. *Biochemistry*, 62(1):75–84, 2023. doi: 10.1021/acs.biochem.2c00532.

[56] D. M. Peter, B. Vogeli, N. S. Cortina, and T. J. Erb. A chemo-enzymatic road map to the synthesis of CoA esters. *Molecules*, 21:517, 2016.

[57] W. Kabsch. Xds. *Acta Crystallogr., Sect. D*, 66:125, 2010.

[58] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart. Phenix: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D*, 66:213, 2010.

[59] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of coot. *Acta Crystallogr., Sect. D*, 66:486, 2010.

[60] N. Guex, M. C. Peitsch, and T. Schwede. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*, 30 Suppl 1:S162–173, Jun 2009.

[61] G. Studer, C. Rempfer, A. M. Waterhouse, R. Gumienny, J. Haas, and T. Schwede. QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics*, 36(6):1765–1771, Mar 2020.

[62] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*, 46(W1):W296–W303, Jul 2018.

[63] F. A. Busch. Current methods for estimating the rate of photorespiration in leaves. *Plant Biol (Stuttg)*, 15(4):648–655, Jul 2013.

[64] A. Bräutigam and U. Gowik. Photorespiration connects C_3 and C_4 photosynthesis. *J. Exp. Bot.*, 67(10):2953–2962, May 2016.

[65] A. L. Ruff-Roberts, J. G. Kuenen, D. M. Ward, and D. M. Ward. Distribution of cultivated and uncultivated cyanobacteria and *Chloroflexus*-like bacteria in hot spring microbial mats. *Appl Environ Microbiol*, 60(2):697–704, Feb 1994.

[66] R. W. Castenholz and B. K. Pierson. Ecology of thermophilic anoxygenic phototrophs. *Anoxygenic photosynthetic bacteria*, pages 87–103, 1995.

[67] M. M. Bateson and D. M. Ward. Photoexcretion and fate of glycolate in a hot spring cyanobacterial mat. *Appl Environ Microbiol*, 54(7):1738–1743, Jul 1988.

[68] M. B. Begemann, E. K. Zess, E. M. Walters, E. F. Schmitt, A. L. Markley, and B. F. Pfleger. An organic acid based counter selection system for cyanobacteria. *PLoS One*, 8(10):e76594, 2013.

[69] R.A. Andersen. *Algal Culturing Techniques*. Academic Press. "Elsevier/Academic Press", 2005.

[70] M. M. Allen. Simple Conditions for Growth of Unicellular Blue-green Algae on Plates. *J. Phycol.*, 4(1):1–4, Mar 1968.

[71] G. Strauss and G. Fuchs. Enzymes of a novel autotrophic CO_2 fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle. *Eur. J. Biochem.*, 215:633, 1993.

[72] S. Herter, A. Busch, and G. Fuchs. L-malyl-Coenzyme A lyase/beta-methylmalyl-Coenzyme A lyase from *Chloroflexus aurantiacus*, a bifunctional enzyme involved in autotrophic CO_2 fixation. *J. Bacteriol.*, 184:5999, 2002.

[73] J. Zarzycki, A. Schlichting, N. Strychalsky, M. Müller, B. E. Alber, and G. Fuchs. Mesaconyl-Coenzyme A hydratase, a new enzyme of two central carbon metabolic pathways in bacteria. *J. Bacteriol.*, 190:1366, 2008.

[74] T. J. Hackmann. Redefining the Coenzyme A transferase superfamily with a large set of manually annotated proteins. *Protein Sci.*, 31:864, 2022.

[75] W. Buckel and A. Bobi. The enzyme complex citramalate lyase from *Clostridium tetanomorphum*. *Eur. J. Biochem.*, 64:255, 1976.

[76] P. Dimroth, W. Buckel, R. Loyal, and H. Eggerer. Isolation and function of the subunits of citramalate lyase and formation of hybrids with the subunits of citrate lyase. *Eur. J. Biochem.*, 80:469, 1977.

[77] W. Buckel, U. Dorn, and R. Semmler. Glutaconate CoA-transferase from *Acidaminococcus fermentans*. *Eur. J. Biochem.*, 118:315, 2005.

[78] C. L. Berthold, C. G. Toyota, N. G. Richards, and Y. Lindqvist. Reinvestigation of the catalytic mechanism of formyl-CoA transferase, a class III CoA-transferase. *J. Biol. Chem.*, 283:6519, 2008.

[79] F. Solomon and W. P. Jencks. Identification of an enzyme-gamma-glutamyl Coenzyme A intermediate from Coenzyme A transferase. *J. Biol. Chem.*, 244:1079, 1969.

[80] T. Selmer and W. Buckel. Oxygen exchange between acetate and the catalytic glutamate residue in glutaconate CoA-transferase from acidaminococcus fermentans. implications for the mechanism of CoA-ester hydrolysis. *J. Biol. Chem.*, 274:20772, 1999.

[81] J. Heider. A new family of CoA-transferases. *FEBS Lett.*, 509:345, 2001.

[82] U. Jacob, M. Mack, T. Clausen, R. Huber, W. Buckel, and A. Messerschmidt. Glutaconate CoA-transferase from *Acidaminococcus fermentans*: the crystal structure reveals homology with other CoA-transferases. *Structure*, 5:415, 1997.

[83] M. Mack and W. Buckel. Conversion of glutaconate CoA-transferase from acidaminococcus fermentans into an acyl-CoA hydrolase by site-directed mutagenesis. *FEBS Lett.*, 405:209, 1997.

[84] E. S. Rangarajan, Y. Li, P. Iannuzzi, M. Cygler, and A. Matte. Crystal structure of *Escherichia coli* crotonobetainyl-CoA: carnitine CoA-transferase (caib) and its complexes with CoA and carnitinyl-CoA. *Biochemistry*, 44:5728, 2005.

[85] S. Ricagno, S. Jonsson, N. Richards, and Y. Lindqvist. Formyl-CoA transferase encloses the CoA binding site at the interface of an interlocked dimer. *EMBO J.*, 22:3210, 2003.

[86] A. Gruez, V. Roig-Zamboni, C. Valencia, V. Campanacci, and C. Cambillau. The crystal structure of the *Escherichia coli* yfdw gene product reveals a new fold of two interlaced rings identifying a wide family of CoA transferases. *J. Biol. Chem.*, 278:34582, 2003.

[87] S. Jonsson, S. Ricagno, Y. Lindqvist, and N. G. Richards. Kinetic and mechanistic characterization of the formyl-CoA transferase from oxalobacter formigenes. *J. Biol. Chem.*, 279:36003, 2004.

[88] K. Schühle, J. Nies, and J. Heider. An indoleacetate-CoA ligase and a phenylsuccinyl-CoA transferase involved in anaerobic metabolism of auxin. *Environ. Microbiol.*, 18:3120, 2016.

[89] Z. Min, X. Zhang, W. Wu, Y. Xin, M. Liu, K. Wang, X. Zhang, Y. He, C. Fan, Z. Wang, and X. Xu. Crystal structure of an intramolecular mesaconyl-

Coenzyme A transferase from the 3-hydroxypropionic acid cycle of *Roseiflexus castenholzii*. *Front. Microbiol.*, 13:923367, 2022.

[90] E. R. Stadtman. Preparation and assay of acyl Coenzyme-a and other thiol esters - use of hydroxylamine. *Methods Enzymol.*, 3:931, 1957.

[91] B. Vögeli, K. Geyer, P. D. Gerlinger, S. Benkstein, N. S. Cortina, and T. J. Erb. Combining promiscuous acyl-CoA oxidase and enoyl-CoA carboxylase/reductases for atypical polyketide extender unit biosynthesis. *Cell Chem. Biol.*, 25:833, 2018.

[92] J. Sambrook and D. W. Russell. *Molecular Cloning: A Laboratory Manual*, volume 1. COLD SPRING HARBOR LABORATORY PRESS CSHLP, 2001.

[93] R. M. C. Dawson. *Data for Biochemical Research*, volume XII. Oxford Science Publication, 1986.

[94] E. A. Mullins, C. M. Starks, J. A. Francois, L. Sael, D. Kihara, and T. J. Kappock. Formyl-Coenzyme A (CoA):oxalate CoA-transferase from the acidophile *Acetobacter aceti* has a distinctive electrostatic surface and inherent acid stability. *Protein Sci.*, 21:686, 2012.

[95] M. Kleiner, C. Wentrup, C. Lott, H. Teeling, S. Wetzel, J. Young, Y. J. Chang, M. Shah, N. C. VerBerkmoes, J. Zarzycki, G. Fuchs, S. Markert, K. Hempel, B. Voigt, D. Becher, M. Liebeke, M. Lalk, D. Albrecht, M. Hecker, T. Schweder, and N. Dubilier. Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proc. Natl. Acad. Sci. U.S.A.*, 109:E1173, 2012.

[96] T. Elssner, C. Engemann, K. Baumgart, and H. P. Kleber. Involvement of Coenzyme A esters and two new enzymes, an enoyl-CoA hydratase and a CoA-transferase, in the hydration of crotonobetaine to l-carnitine by *Escherichia coli*. *Biochemistry*, 40:11140, 2001.

[97] F. Borjian, U. Johnsen, P. Schönheit, and I. A. Berg. Succinyl-CoA:mesaconate CoA-transferase and mesaconyl-CoA hydratase, enzymes of the methylaspartate cycle in *Haloarcula hispanica*. *Front. Microbiol.*, 8:1683, 2017.

[98] J. Kim, D. Darley, T. Selmer, and W. Buckel. Characterization of (r)-2-hydroxyisocaproate dehydrogenase and a family III Coenzyme A transferase involved in reduction of l-leucine to isocaproate by *Clostridium difficile*. *Appl. Environ. Microbiol.*, 72:6062, 2006.

[99] M. Schürmann, B. Hirsch, J. H. Wubbeler, N. Stoveken, and A. Steinbüchel. Succinyl-CoA:3-sulfinopropionate CoA-transferase from variovorax paradoxus strain tbea6, a novel member of the class III Coenzyme A (CoA)-transferase family. *J. Bacteriol.*, 195:3761, 2013.

[100] S. Friedmann, A. Steindorf, B. E. Alber, and G. Fuchs. 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, 2006.

[101] M. W. Burgis. Roadmap for cyanobacteria engineering: Integration of a Fumaryl-CoA shunt in *Synechococcus elongatus* PCC 7942. B.Sc. thesis, Philipps-Universität Marburg, Bibliothek des Fachbereichs Biologie, Karl-von-Frisch-Str. 8, 35037 Marburg, Deutschland, May 2021.

List of Figures

List of Tables

Abbreviations

2PG 2-phosphoglycolate

2OG 2-oxoglutarate

3OHP 3-hydroxypropionate

3PG 3-phosphoglycerate

Acc acetyl-CoA carboxylase

Acs acetyl-CoA synthase

ALE adaptive laboratory evolution

ATP adenosine triphosphate

BirA (acyl-CoA carboxylase)biotin ligase

BC biotin carboxylase

BCCP biotin carboxyl carrier protein

BHAC β-hydroxyaspartate cycle

CBB Calvin-Benson-Bassham

CCM carbon concentrating mechanism

Ccr crotonyl-CoA carboxylase/reductase

CETCH crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA

CT carboxyl transferase

Ecr enoyl-CoA carboxylase/reductase

FuCo fumaryl-CoA

G3P glyceraldehyde-3-phosphate

GOT great oxygenation transition

HPLC high Performance Liquid Chromatography

Mch mesaconyl-C1-CoA hydratase

Mcl β-methylmalyl-CoA/(S)-citramalyl-CoA lyase

Mct mesaconyl-C1-C4-CoA CoA transferase

Mcr malonyl-CoA reductase

Meh mesaconyl-C4-CoA hydratase

NADPH nicotinamide adenine dinucleotide phosphate

Pcc propionyl-CoA carboxylase

Pcs propionyl-CoA synthase

PduP propionaldehyde dehydrogenase

PHB polyhydroxybutyrate

PEPC Phosphoenolpyruvate carboxylase

PHA polyhydroxyalkanoates

RuBP ribulose-1,5-bisphosphate

RuBisCO ribulose-1,5-bisphosphate carboxylase/oxygenase

SucD succinic semialdehyde dehydrogenase

SSA succinic semialdehyde

TaCo tartronyl-CoA

TCA tricarboxylic acid

Chapter 6

Appendix

6.1 Acknowledgments

I want to express my sincere gratitude to all the people that made this accomplishment possible. My whole doctorate, I had the feeling that everyone around me was driven with the extraordinary curiosity and passion for science, that motivated and inspired me all the way.
First and foremost, I want to thank my supervisor, Tobias Erb. Even though I saw you on so many stages, I will never forget your excitement and engagement for science as you were substitute lecturer in our bachelor lesson on anaplerosis. You taught me that scientists better be humble, and value the importance of every single insight.
My thanks also go to my instructor and mentor Jan Zarzycki. You raised me from the first plasmid preps to everything I have done scientifically thereafter. All this time, you had an open ear, and helped me with even the tiniest problems. I simply couldn't have done it without you.
I thank the scientific role models that played an important role in my studies. Lars-Oliver Essen and Anke Becker, as my thesis advisory committee, but also Michael Bölker and Johann Heider. I am very happy that my examination committee consists of scientist, that I know and appreciate from the early stages of my scientific career.
I had a great set of collaborators. Chris, Maren, Simon, Gabo and Matthias, you provided me your precious enzymes that I had nothing better to do with than to precipitate them and call for more. Thank you for your endurance in our lasting crystallography projects. Marieke, Daniel and Helena, I thank you for all the inspiring discussions about enzymes, evolution and statistics.
I thank Michael, Vanessa and Eric. I hope I taught you as much, as I could learn trough you. I hope you are missing our morning briefings as much as I do.
My thanks also go to all the great discussion partners I had in our day starting coffee

routine. Luca, Alberto, Nicole, Philipp, René and all the rest of the Erbly birds.
I want to thank all the great people working with cyanobacteria for sharing their methods, contacts, ideas and friendly small talk, Moritz and Khaled.
Mein Dank geht auch an meine Familie; Besonders an Valeria und Gerda, aber auch an meine Eltern und Geschwister. Eure Liebe gibt mir das sichere Gefühl, das mich stets aufgefangen hat.

6.2 Validation Reports

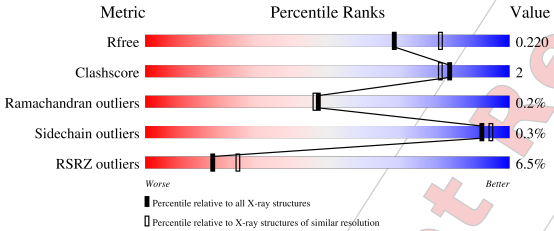
As depositions for PDB 8CEK, 8CEJ and 8CEJ are not publicly available yet, the respective reports are attached hereafter;

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	453	7% 94% 5%
1	B	453	7% 94% 6%
1	C	453	5% 92% 6%
1	D	453	7% 94% 5%

Validation Pipeline (wwPDB-VP) : 2.32.1

Full wwPDB X-ray Structure Validation Report (i)

Feb 14, 2023 – 05:04 pm GMT

PDB ID : 8CEK
Title : Succinyl-CoA Reductase from Clostridium kluyveri (SucD) with NADPH
Deposited on : 2023-02-02
Resolution : 2.15 Å(reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.32.1
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

2 Entry composition

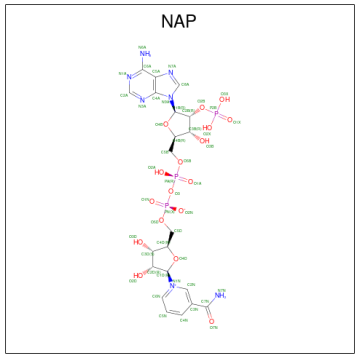
There are 3 unique types of molecules in this entry. The entry contains 14888 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Succinate-semialdehyde dehydrogenase (acetylating).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	1	0
			3416	2163	577	661	15			
1	B	449	Total	C	N	O	S	0	0	0
			3407	2157	575	660	15			
1	C	448	Total	C	N	O	S	0	0	0
			3400	2152	574	659	15			
1	D	449	Total	C	N	O	S	0	0	0
			3407	2157	575	660	15			

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 3 is water.

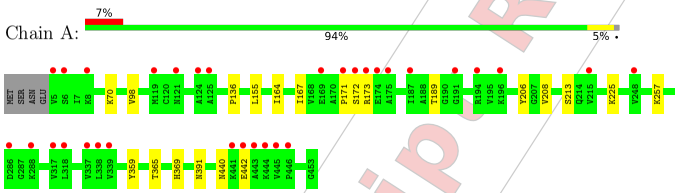
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	271	Total	O	0	0
			271	271		
3	B	283	Total	O	0	0
			283	283		
3	C	211	Total	O	0	0
			211	211		
3	D	301	Total	O	0	0
			301	301		

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

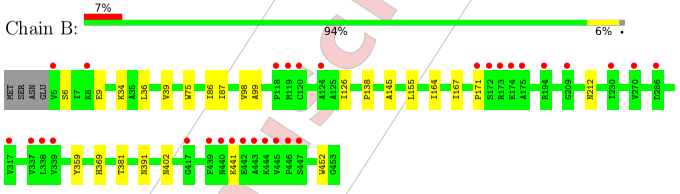
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

Chain A:



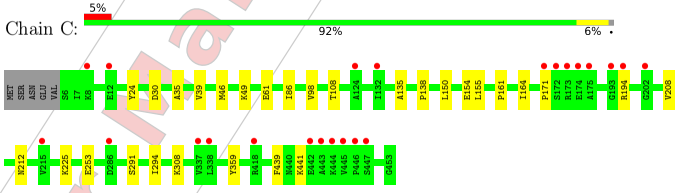
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

Chain B:



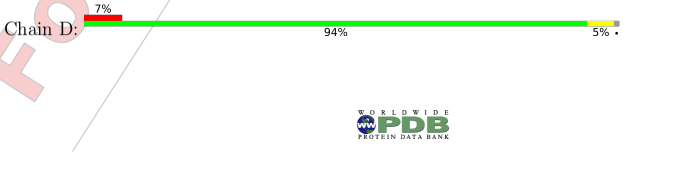
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

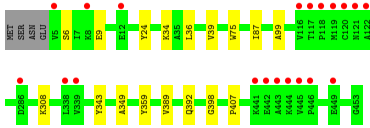
Chain C:



- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	140.05Å 190.83Å 190.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.66 – 2.15 29.66 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.66-2.15) 99.9 (29.66-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.53 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.198 , 0.222 0.198 , 0.220	Depositor DCC
R_{free} test set	2004 reflections (1.45%)	wwPDB-VP
Wilson B-factor (Å^2)	38.6	Xtriage
Anisotropy	0.422	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 44.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.047 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	14888	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.25 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.4257e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAP

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.28	0/3473	0.47	0/4696
1	B	0.28	0/3464	0.46	0/4685
1	C	0.28	0/3457	0.47	0/4675
1	D	0.29	0/3464	0.46	0/4685
All	All	0.28	0/13858	0.47	0/18741

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3416	0	3467	18	0
1	B	3407	0	3455	13	0
1	C	3400	0	3446	20	0
1	D	3407	0	3455	14	0
2	A	48	0	25	2	0
2	B	48	0	25	2	0
2	C	48	0	25	4	0
2	D	48	0	25	2	0
3	A	271	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	283	0	0	2	0
3	C	211	0	0	6	0
3	D	301	0	0	2	0
All	All	14888	0	13923	67	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (67) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:225:LYS:NZ	3:A:602:HOH:O	2.22	0.73
1:C:225:LYS:NZ	3:C:606:HOH:O	2.26	0.68
1:A:70:LYS:NZ	3:A:605:HOH:O	2.27	0.65
1:A:189:THR:HA	1:A:208:VAL:CG2	2.27	0.65
1:A:189:THR:HA	1:A:208:VAL:HG22	1.80	0.63
1:A:136:PRO:HG3	1:A:167:ILE:HD11	1.82	0.61
1:D:36:LEU:O	1:D:39:VAL:HG12	2.02	0.60
2:B:501:NAP:H8A	2:B:501:NAP:H51A	1.87	0.57
1:A:172:SER:OG	1:A:173:ARG:N	2.33	0.57
2:D:501:NAP:O2A	3:D:602:HOH:O	2.18	0.55
1:C:86:ILE:HG12	1:C:98:VAL:HG22	1.90	0.54
1:D:138:PRO:HG3	1:D:171:PRO:HG3	1.89	0.54
1:B:155:LEU:HD13	1:B:164:ILE:HD11	1.90	0.53
1:B:452:TRP:O	3:B:601:HOH:O	2.19	0.52
1:D:173:ARG:NH2	2:D:501:NAP:O2X	2.42	0.52
1:C:439:PHE:HE2	1:C:441:LYS:HD2	1.75	0.52
1:C:155:LEU:HD13	1:C:164:ILE:HD11	1.92	0.51
1:D:87:ILE:HD13	1:D:99:ALA:HB2	1.94	0.50
1:B:87:ILE:HD13	1:B:99:ALA:HB2	1.94	0.49
1:C:138:PRO:HG3	1:C:171:PRO:HG3	1.94	0.49
1:C:291:SER:HA	1:C:294:ILE:HG13	1.95	0.49
1:B:6:SER:HB3	1:B:9:GLU:HG3	1.94	0.49
1:C:24:TYR:OH	1:C:161:PRO:HD3	2.14	0.48
1:A:208:VAL:HB	2:A:501:NAP:N7N	2.29	0.48
1:A:208:VAL:O	1:A:208:VAL:HG23	2.14	0.47
1:D:6:SER:HB3	1:D:9:GLU:HG3	1.95	0.47
1:C:30:ASP:OD2	3:C:603:HOH:O	2.20	0.47
1:D:308:LYS:NZ	3:D:608:HOH:O	2.40	0.46
1:D:24:TYR:OH	1:D:161:PRO:HD3	2.15	0.46
1:C:39:VAL:CG2	1:C:154:GLU:HG3	2.46	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:440:ASN:OD1	1:A:442:GLU:HG3	2.17	0.45
1:B:34:LYS:HB2	1:B:75:TRP:CH2	2.52	0.45
1:B:86:ILE:HG12	1:B:98:VAL:HG22	1.97	0.45
1:A:172:SER:HG	1:A:173:ARG:H	1.63	0.45
1:A:206:TYR:HB3	3:A:690:HOH:O	2.17	0.45
1:C:308:LYS:HG2	1:D:264:GLU:OE2	2.17	0.45
1:A:189:THR:HA	1:A:208:VAL:HG21	1.98	0.44
1:D:173:ARG:HA	1:D:176:ALA:HB3	1.99	0.44
1:A:257:LYS:HB3	1:A:257:LYS:HE2	1.69	0.44
1:C:108:THR:HG22	1:C:135:ALA:HB3	2.00	0.44
1:A:155:LEU:HD13	1:A:164:ILE:HD11	2.00	0.43
1:B:138:PRO:HG3	1:B:171:PRO:HG3	1.99	0.43
1:C:194:ARG:HH11	2:C:501:NAP:C8A	2.31	0.43
1:B:145:ALA:HA	1:B:167:ILE:HD13	2.01	0.43
2:B:501:NAP:H51A	2:B:501:NAP:C8A	2.47	0.43
1:C:39:VAL:HG22	1:C:154:GLU:HG3	2.00	0.43
1:D:34:LYS:HB2	1:D:75:TRP:CH2	2.54	0.43
1:A:172:SER:HB2	2:A:501:NAP:N6A	2.34	0.42
1:D:212:ASN:HA	1:D:246:GLN:HG3	2.00	0.42
1:A:98:VAL:HG11	1:B:381:THR:HA	2.00	0.42
1:B:402:ASN:HB2	3:B:611:HOH:O	2.19	0.42
1:D:343:TYR:CG	1:D:349:ALA:HB2	2.54	0.42
1:D:398:GLY:O	1:D:407:PRO:HA	2.19	0.42
2:C:501:NAP:H52A	3:C:615:HOH:O	2.20	0.42
1:C:35:ALA:O	1:C:39:VAL:HG23	2.19	0.42
1:C:253:GLU:HG2	3:C:729:HOH:O	2.20	0.42
1:B:36:LEU:HD11	1:B:126:ILE:HD12	2.01	0.42
1:B:36:LEU:O	1:B:39:VAL:HG12	2.20	0.42
1:A:369:HIS:HA	1:A:391:ASN:OD1	2.20	0.42
1:A:213:SER:HA	1:A:365:THR:O	2.20	0.41
1:C:194:ARG:HH12	2:C:501:NAP:P2B	2.43	0.41
1:C:49:LYS:HG3	3:C:748:HOH:O	2.20	0.41
1:C:46:MET:HE1	1:C:150:LEU:HD11	2.02	0.41
1:B:369:HIS:HA	1:B:391:ASN:OD1	2.20	0.41
1:C:61:GLU:OE1	3:C:604:HOH:O	2.22	0.40
1:D:389:VAL:CG1	1:D:392:GLN:HG3	2.51	0.40
1:C:208:VAL:HB	2:C:501:NAP:N7N	2.35	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	448/453 (99%)	439 (98%)	8 (2%)	1 (0%)	47	46
1	B	447/453 (99%)	440 (98%)	6 (1%)	1 (0%)	47	46
1	C	446/453 (98%)	439 (98%)	6 (1%)	1 (0%)	47	46
1	D	447/453 (99%)	437 (98%)	9 (2%)	1 (0%)	47	46
All	All	1788/1812 (99%)	1755 (98%)	29 (2%)	4 (0%)	47	46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	PRO
1	C	212	ASN
1	B	212	ASN
1	D	212	ASN

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	363/366 (99%)	362 (100%)	1 (0%)	92	95
1	B	362/366 (99%)	360 (99%)	2 (1%)	86	90
1	C	361/366 (99%)	360 (100%)	1 (0%)	92	95
1	D	362/366 (99%)	361 (100%)	1 (0%)	92	95
All	All	1448/1464 (99%)	1443 (100%)	5 (0%)	92	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	359	TYR
1	B	359	TYR
1	B	441	LYS
1	C	359	TYR
1	D	359	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAP	D	501	-	45,52,52	2.21	6 (13%)	56,80,80	1.65	16 (28%)
2	NAP	C	501	-	45,52,52	2.15	7 (15%)	56,80,80	1.62	14 (25%)
2	NAP	B	501	-	45,52,52	2.18	8 (17%)	56,80,80	1.61	13 (23%)
2	NAP	A	501	-	45,52,52	2.17	6 (13%)	56,80,80	1.64	16 (28%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAP	D	501	-	-	7/31/67/67	0/5/5/5
2	NAP	C	501	-	-	9/31/67/67	0/5/5/5
2	NAP	B	501	-	-	5/31/67/67	0/5/5/5
2	NAP	A	501	-	-	5/31/67/67	0/5/5/5

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	501	NAP	P2B-O2B	11.05	1.80	1.59
2	A	501	NAP	P2B-O2B	10.92	1.79	1.59
2	B	501	NAP	P2B-O2B	10.62	1.79	1.59
2	C	501	NAP	P2B-O2B	10.55	1.79	1.59
2	A	501	NAP	PN-O5D	5.09	1.79	1.59
2	B	501	NAP	PN-O5D	5.07	1.79	1.59
2	C	501	NAP	PN-O5D	4.87	1.79	1.59
2	D	501	NAP	PN-O5D	4.87	1.79	1.59
2	D	501	NAP	C7N-N7N	2.86	1.38	1.33
2	B	501	NAP	C7N-N7N	2.82	1.38	1.33
2	C	501	NAP	C7N-N7N	2.79	1.38	1.33
2	A	501	NAP	C7N-N7N	2.78	1.38	1.33
2	D	501	NAP	C2N-N1N	2.61	1.38	1.35
2	B	501	NAP	C2A-N1A	2.54	1.38	1.33
2	C	501	NAP	C2A-N1A	2.48	1.38	1.33
2	A	501	NAP	C2A-N1A	2.42	1.38	1.33
2	A	501	NAP	C2N-N1N	2.35	1.37	1.35
2	A	501	NAP	O2B-C2B	-2.31	1.35	1.44
2	D	501	NAP	C2A-N1A	2.31	1.38	1.33
2	C	501	NAP	O2B-C2B	-2.30	1.35	1.44
2	B	501	NAP	O2B-C2B	-2.29	1.35	1.44
2	C	501	NAP	C2N-N1N	2.26	1.37	1.35
2	D	501	NAP	O2B-C2B	-2.23	1.36	1.44
2	B	501	NAP	C2D-C1D	2.21	1.57	1.53
2	C	501	NAP	C4A-N3A	2.10	1.38	1.35
2	B	501	NAP	C2N-N1N	2.01	1.37	1.35
2	B	501	NAP	C4A-N3A	2.01	1.38	1.35

All (59) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	501	NAP	PN-O3-PA	-5.60	113.59	132.83
2	A	501	NAP	PN-O3-PA	-5.57	113.72	132.83
2	B	501	NAP	PN-O3-PA	-5.56	113.75	132.83
2	C	501	NAP	PN-O3-PA	-5.54	113.82	132.83
2	A	501	NAP	O3X-P2B-O2X	3.14	119.62	107.64
2	C	501	NAP	O3X-P2B-O2X	3.08	119.40	107.64
2	D	501	NAP	O3X-P2B-O2X	3.04	119.27	107.64
2	B	501	NAP	O3X-P2B-O2X	3.02	119.17	107.64
2	B	501	NAP	O2B-P2B-O1X	-2.91	98.14	109.39
2	C	501	NAP	O2B-P2B-O1X	-2.86	98.36	109.39
2	B	501	NAP	PA-O5B-C5B	-2.83	105.08	121.68
2	D	501	NAP	O2B-P2B-O1X	-2.82	98.52	109.39
2	A	501	NAP	O2B-P2B-O1X	-2.81	98.54	109.39
2	B	501	NAP	O7N-C7N-C3N	2.79	122.97	119.63
2	D	501	NAP	O2A-PA-O1A	2.78	126.00	112.24
2	B	501	NAP	O2A-PA-O1A	2.78	125.97	112.24
2	A	501	NAP	PA-O5B-C5B	-2.76	105.49	121.68
2	A	501	NAP	O2A-PA-O1A	2.74	125.80	112.24
2	C	501	NAP	PA-O5B-C5B	-2.74	105.64	121.68
2	C	501	NAP	O2A-PA-O1A	2.73	125.76	112.24
2	D	501	NAP	PA-O5B-C5B	-2.72	105.71	121.68
2	C	501	NAP	O7N-C7N-C3N	2.68	122.83	119.63
2	B	501	NAP	PN-O5D-C5D	-2.66	106.06	121.68
2	A	501	NAP	O7N-C7N-C3N	2.64	122.80	119.63
2	C	501	NAP	PN-O5D-C5D	-2.64	106.22	121.68
2	D	501	NAP	O7N-C7N-C3N	2.60	122.74	119.63
2	D	501	NAP	PN-O5D-C5D	-2.58	106.56	121.68
2	A	501	NAP	PN-O5D-C5D	-2.48	107.15	121.68
2	C	501	NAP	O4B-C4B-C3B	2.43	109.93	105.11
2	D	501	NAP	O5D-PN-O1N	-2.40	99.71	109.07
2	D	501	NAP	O2N-PN-O1N	2.38	124.00	112.24
2	A	501	NAP	O5D-PN-O1N	-2.36	99.83	109.07
2	C	501	NAP	O5D-PN-O1N	-2.36	99.85	109.07
2	B	501	NAP	O5D-PN-O1N	-2.35	99.90	109.07
2	B	501	NAP	O2N-PN-O1N	2.33	123.76	112.24
2	C	501	NAP	O2N-PN-O1N	2.33	123.75	112.24
2	A	501	NAP	O2N-PN-O1N	2.32	123.73	112.24
2	D	501	NAP	C5B-C4B-C3B	-2.31	106.54	115.18
2	C	501	NAP	C5B-C4B-C3B	-2.27	106.69	115.18
2	D	501	NAP	O2N-PN-O5D	-2.26	97.24	107.75
2	C	501	NAP	O2N-PN-O5D	-2.25	97.32	107.75
2	B	501	NAP	O2N-PN-O5D	-2.23	97.41	107.75
2	A	501	NAP	O2N-PN-O5D	-2.18	97.62	107.75

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	NAP	C5B-C4B-C3B	-2.18	107.02	115.18
2	D	501	NAP	O7N-C7N-N7N	-2.17	119.49	122.58
2	A	501	NAP	C3B-C2B-C1B	-2.17	98.80	102.89
2	A	501	NAP	O7N-C7N-N7N	-2.17	119.50	122.58
2	B	501	NAP	O7N-C7N-N7N	-2.14	119.53	122.58
2	C	501	NAP	O7N-C7N-N7N	-2.14	119.54	122.58
2	D	501	NAP	O4B-C4B-C3B	2.10	109.27	105.11
2	D	501	NAP	C2A-N1A-C6A	-2.10	115.17	118.75
2	B	501	NAP	O2X-P2B-O1X	2.09	118.85	110.68
2	D	501	NAP	C3B-C2B-C1B	-2.08	98.98	102.89
2	A	501	NAP	O4B-C4B-C3B	2.07	109.21	105.11
2	A	501	NAP	C2A-N1A-C6A	-2.07	115.21	118.75
2	D	501	NAP	O2X-P2B-O1X	2.06	118.75	110.68
2	C	501	NAP	O2X-P2B-O1X	2.06	118.75	110.68
2	B	501	NAP	C5B-C4B-C3B	-2.03	107.58	115.18
2	A	501	NAP	O2X-P2B-O1X	2.03	118.61	110.68

There are no chirality outliers.

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	501	NAP	C5B-O5B-PA-O1A
2	D	501	NAP	C5B-O5B-PA-O3
2	B	501	NAP	O4B-C4B-C5B-O5B
2	C	501	NAP	C3D-C4D-C5D-O5D
2	D	501	NAP	C3D-C4D-C5D-O5D
2	C	501	NAP	O4B-C4B-C5B-O5B
2	C	501	NAP	O4D-C4D-C5D-O5D
2	A	501	NAP	PA-O3-PN-O5D
2	B	501	NAP	PA-O3-PN-O5D
2	C	501	NAP	PA-O3-PN-O5D
2	D	501	NAP	PA-O3-PN-O5D
2	B	501	NAP	C2B-O2B-P2B-O1X
2	D	501	NAP	C2B-O2B-P2B-O1X
2	D	501	NAP	O4D-C4D-C5D-O5D
2	D	501	NAP	C5B-O5B-PA-O2A
2	A	501	NAP	C3D-C4D-C5D-O5D
2	C	501	NAP	C3B-C4B-C5B-O5B
2	C	501	NAP	C3B-C2B-O2B-P2B
2	C	501	NAP	C1B-C2B-O2B-P2B
2	A	501	NAP	C2B-O2B-P2B-O1X
2	B	501	NAP	C5D-O5D-PN-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	501	NAP	C2B-O2B-P2B-O2X
2	A	501	NAP	O4D-C4D-C5D-O5D
2	C	501	NAP	PN-O3-PA-O1A
2	A	501	NAP	C5D-O5D-PN-O1N
2	B	501	NAP	C5D-O5D-PN-O1N

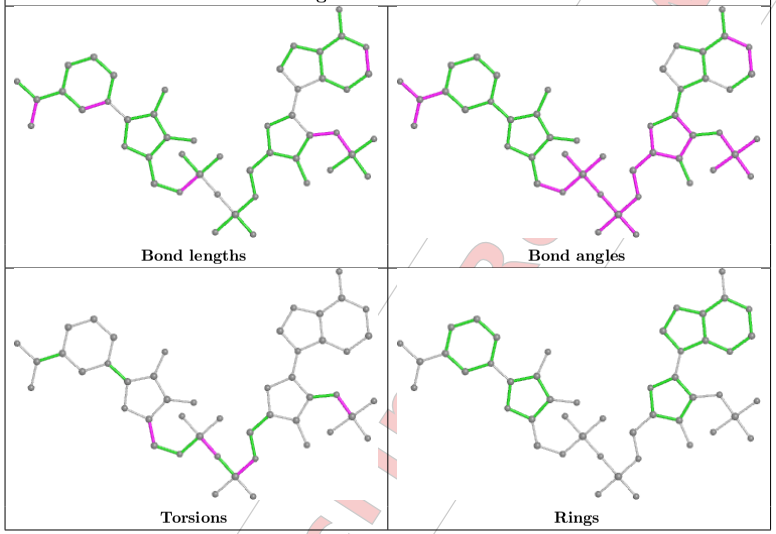
There are no ring outliers.

4 monomers are involved in 10 short contacts:

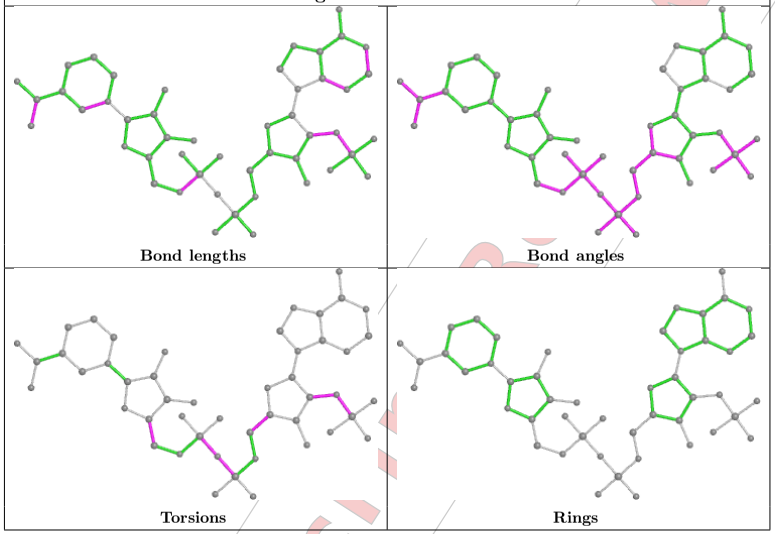
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	501	NAP	2	0
2	C	501	NAP	4	0
2	B	501	NAP	2	0
2	A	501	NAP	2	0

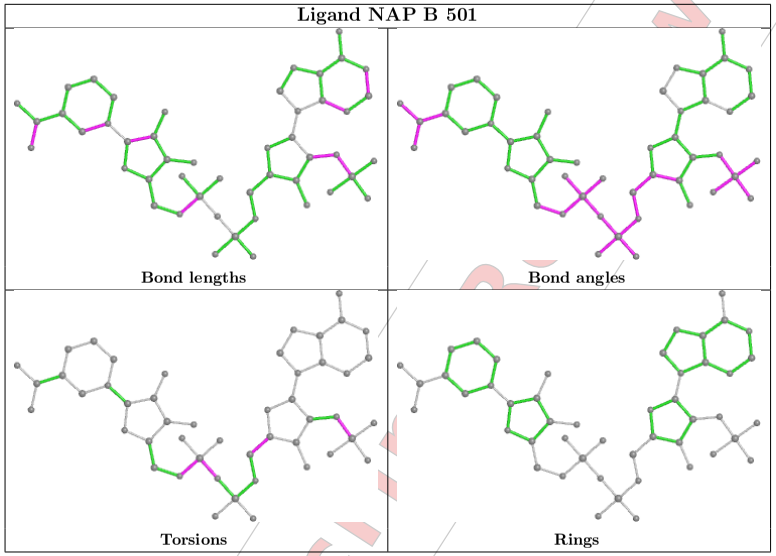
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

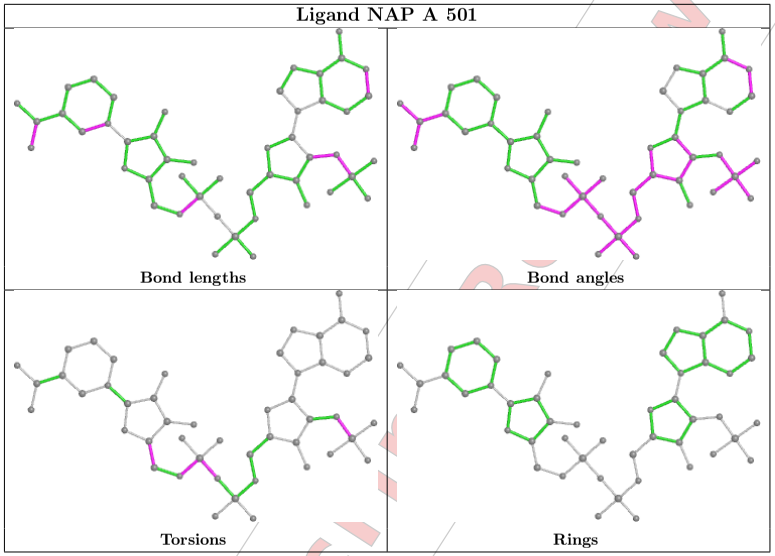
Ligand NAP D 501



Ligand NAP C 501







5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å^2)	Q<0.9
1	A	449/453 (99%)	0.23	32 (7%)	16	22	32, 43, 70, 112	0
1	B	449/453 (99%)	0.23	30 (6%)	17	24	27, 42, 72, 112	0
1	C	448/453 (98%)	0.24	23 (5%)	28	36	29, 47, 77, 124	0
1	D	449/453 (99%)	0.17	31 (6%)	16	23	27, 41, 71, 121	0
All	All	1795/1812 (99%)	0.22	116 (6%)	18	25	27, 43, 74, 124	0

All (116) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	445	VAL	7.1
1	B	445	VAL	6.4
1	C	174	GLU	6.3
1	C	443	ALA	5.7
1	B	171	PRO	5.7
1	D	175	ALA	5.6
1	C	445	VAL	5.5
1	D	445	VAL	5.4
1	C	175	ALA	5.3
1	B	446	PRO	5.2
1	A	194	ARG	5.2
1	B	175	ALA	5.2
1	C	444	LYS	5.2
1	C	172	SER	4.9
1	B	441	LYS	4.8
1	C	171	PRO	4.8
1	B	444	LYS	4.8
1	D	174	GLU	4.7
1	A	173	ARG	4.7
1	A	171	PRO	4.6
1	B	5	VAL	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	171	PRO	4.4
1	C	442	GLU	4.0
1	D	443	ALA	4.0
1	D	173	ARG	3.9
1	B	442	GLU	3.9
1	D	444	LYS	3.8
1	A	174	GLU	3.7
1	C	446	PRO	3.7
1	B	174	GLU	3.7
1	D	442	GLU	3.5
1	A	5	VAL	3.5
1	A	339	VAL	3.5
1	A	443	ALA	3.4
1	B	439	PHE	3.4
1	A	317	VAL	3.3
1	B	8	LYS	3.3
1	C	132	ILE	3.3
1	D	441	LYS	3.2
1	C	8	LYS	3.2
1	A	6	SER	3.2
1	D	194	ARG	3.2
1	A	338	LEU	3.2
1	C	194	ARG	3.2
1	A	215	VAL	3.2
1	C	173	ARG	3.1
1	A	444	LYS	3.1
1	A	196	LYS	3.0
1	D	122	ALA	3.0
1	A	187	ILE	3.0
1	B	443	ALA	3.0
1	C	418	ARG	2.9
1	D	286	ASP	2.9
1	D	172	SER	2.9
1	B	124	ALA	2.9
1	D	120	CYS	2.9
1	D	338	LEU	2.9
1	D	446	PRO	2.8
1	A	446	PRO	2.8
1	C	12	GLU	2.8
1	A	8	LYS	2.7
1	B	172	SER	2.7
1	A	442	GLU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	118	PRO	2.7
1	A	125	ALA	2.7
1	C	193	GLY	2.7
1	A	248	VAL	2.7
1	B	317	VAL	2.7
1	D	132	ILE	2.7
1	B	286	ASP	2.6
1	A	318	LEU	2.6
1	B	447	SER	2.5
1	B	173	ARG	2.5
1	D	124	ALA	2.5
1	B	194	ARG	2.5
1	C	286	ASP	2.5
1	C	124	ALA	2.5
1	D	8	LYS	2.5
1	A	286	ASP	2.5
1	B	338	LEU	2.5
1	C	215	VAL	2.5
1	B	120	CYS	2.5
1	A	175	ALA	2.4
1	B	118	PRO	2.4
1	A	337	VAL	2.4
1	B	417	GLY	2.4
1	A	172	SER	2.4
1	A	441	LYS	2.4
1	C	337	VAL	2.4
1	D	187	ILE	2.4
1	B	337	VAL	2.3
1	D	247	SER	2.3
1	D	117	THR	2.3
1	A	191	GLY	2.3
1	C	338	LEU	2.2
1	D	119	MET	2.2
1	B	230	ILE	2.2
1	B	209	GLY	2.2
1	D	116	VAL	2.2
1	A	121	ASN	2.2
1	A	288[A]	LYS	2.1
1	A	124	ALA	2.1
1	D	12	GLU	2.1
1	B	119	MET	2.1
1	B	440	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	5	VAL	2.1
1	B	339	VAL	2.1
1	D	121	ASN	2.1
1	D	339	VAL	2.1
1	D	177	LYS	2.1
1	C	447	SER	2.0
1	A	119	MET	2.0
1	A	169	GLU	2.0
1	C	202	GLY	2.0
1	D	449	GLU	2.0
1	B	270	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

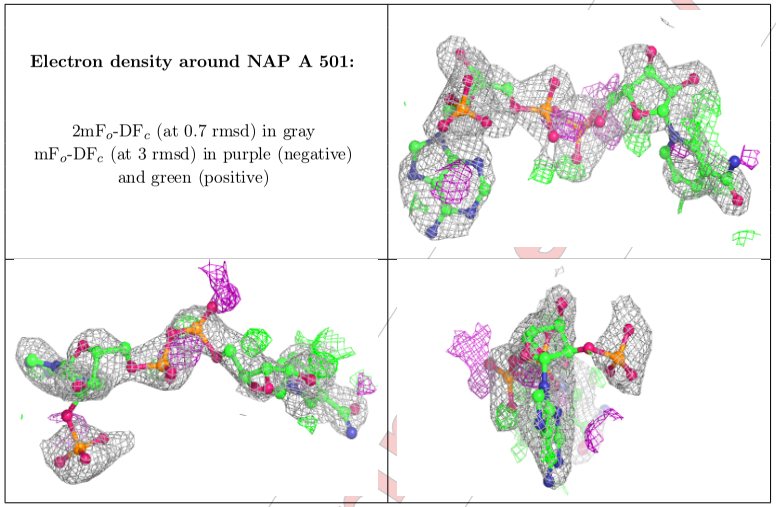
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAP	A	501	48/48	0.63	0.32	52,71,84,85	48
2	NAP	D	501	48/48	0.64	0.30	40,64,75,78	48
2	NAP	C	501	48/48	0.65	0.32	50,68,86,91	48
2	NAP	B	501	48/48	0.65	0.31	51,66,79,81	48

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

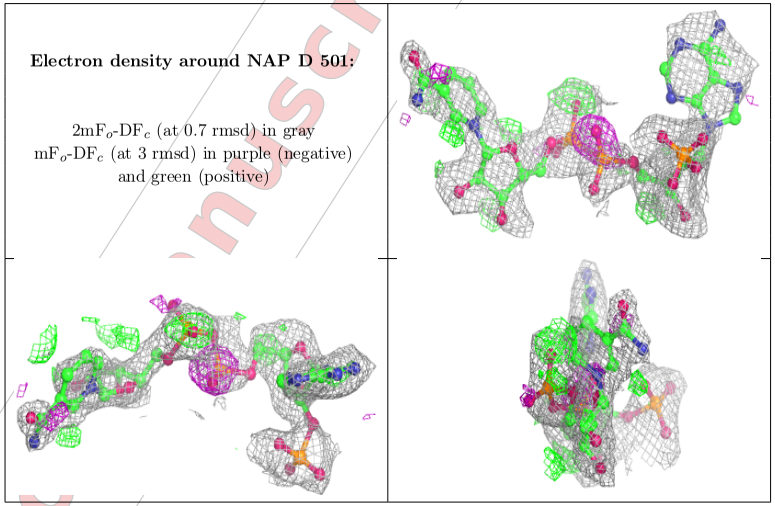
Electron density around NAP A 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



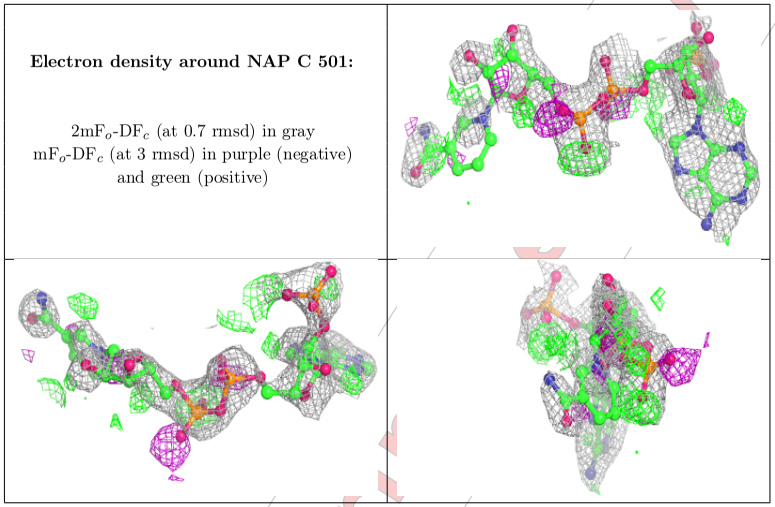
Electron density around NAP D 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



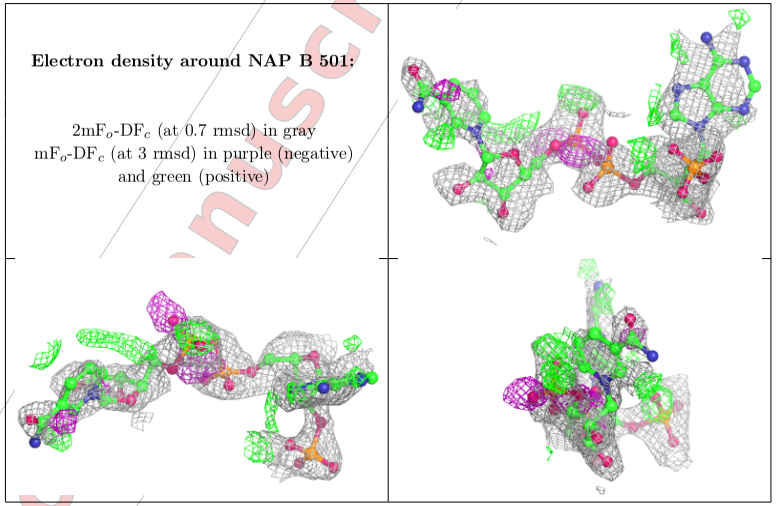
Electron density around NAP C 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around NAP B 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

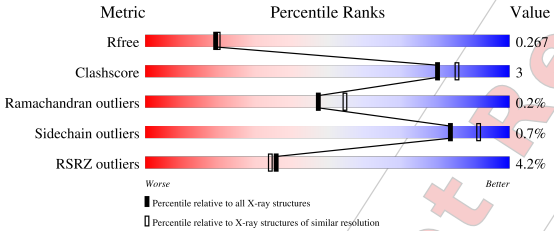
There are no such residues in this entry.

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	453	5% 91% 8%
1	B	453	3% 92% 7%
1	C	453	4% 90% 9%
1	D	453	5% 91% 7%

WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 3, 2023 – 05:09 pm GMT

PDB ID : 8CEI
Title : Succinyl-CoA Reductase from Clostridium kluyveri (SucD)
Deposited on : 2023-02-02
Resolution : 2.20 Å(reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.32.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.32.1

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14600 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Succinate-semialdehyde dehydrogenase (acetylating).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	447	Total	C	N	O	S	0	0	0
			3394	2149	573	657	15			
1	B	447	Total	C	N	O	S	0	0	0
			3394	2149	573	657	15			
1	C	447	Total	C	N	O	S	0	0	0
			3394	2149	573	657	15			
1	D	447	Total	C	N	O	S	0	0	0
			3394	2149	573	657	15			

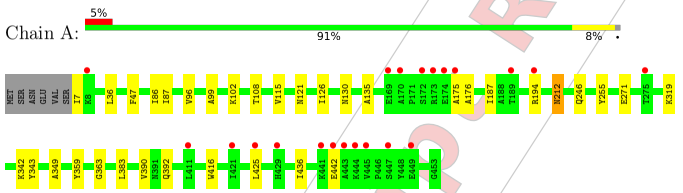
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	227	Total	O	0	0
			227	227		
2	B	259	Total	O	0	0
			259	259		
2	C	281	Total	O	0	0
			281	281		
2	D	257	Total	O	0	0
			257	257		

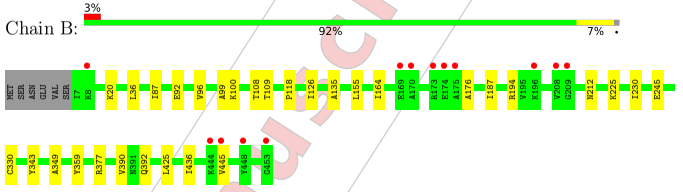
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

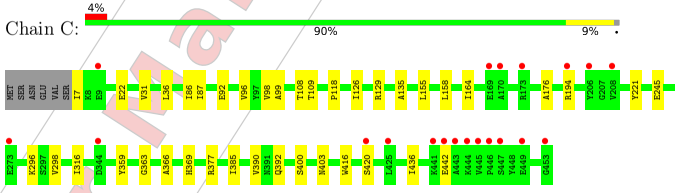
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)



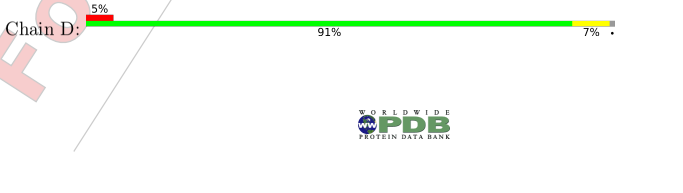
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)



- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)



- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

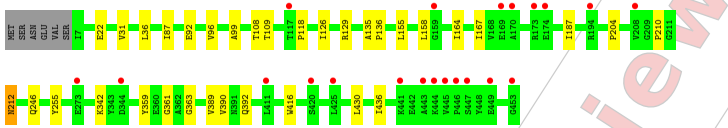


4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	86.23Å 89.34Å 137.32Å 90.00° 104.64° 90.00°	Depositor
Resolution (Å)	39.47 – 2.20 38.77 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.5 (39.47-2.20) 98.1 (38.77-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.03 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.242 , 0.267 0.243 , 0.267	Depositor DCC
R_{free} test set	1969 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 34.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.58$, $< L^2 > = 0.43$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14600	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 58.47 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0362e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.24	0/3451	0.43	0/4667
1	B	0.24	0/3451	0.43	0/4667
1	C	0.24	0/3451	0.43	0/4667
1	D	0.24	0/3451	0.43	0/4667
All	All	0.24	0/13804	0.43	0/18668

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3394	0	3441	18	0
1	B	3394	0	3441	19	0
1	C	3394	0	3441	22	0
1	D	3394	0	3441	18	0
2	A	227	0	0	1	0
2	B	259	0	0	4	0
2	C	281	0	0	3	0
2	D	257	0	0	0	0
All	All	14600	0	13764	74	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (74) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:330:CYS:SG	2:B:501:HOH:O	2.43	0.76
1:C:36:LEU:HD11	1:C:126:ILE:HD12	1.79	0.64
1:D:36:LEU:HD11	1:D:126:ILE:HD12	1.79	0.64
1:B:36:LEU:HD11	1:B:126:ILE:HD12	1.86	0.57
1:C:296:LYS:NZ	2:C:506:HOH:O	2.38	0.57
1:C:108:THR:HG21	1:C:194:ARG:HH21	1.70	0.56
1:D:87:ILE:HD13	1:D:99:ALA:HB2	1.87	0.56
1:C:22:GLU:HA	1:C:129:ARG:HD2	1.88	0.56
1:D:96:VAL:HB	1:D:436:ILE:HB	1.88	0.54
1:A:36:LEU:HD11	1:A:126:ILE:HD12	1.89	0.54
1:B:225:LYS:NZ	2:B:517:HOH:O	2.40	0.54
1:C:87:ILE:HD13	1:C:99:ALA:HB2	1.90	0.53
1:A:271:GLU:HG3	1:A:319:LYS:HD2	1.91	0.52
1:B:20:LYS:NZ	2:B:522:HOH:O	2.43	0.51
1:A:383:LEU:O	1:B:100:LYS:HE2	2.11	0.51
1:B:194:ARG:NH1	2:B:527:HOH:O	2.45	0.50
1:C:176:ALA:HB1	1:C:194:ARG:HH22	1.76	0.49
1:B:109:THR:HG21	1:B:118:PRO:HG3	1.94	0.49
1:A:86:ILE:HD13	1:B:377:ARG:HH11	1.78	0.49
1:A:108:THR:HG22	1:A:135:ALA:HB3	1.95	0.49
1:C:7:ILE:N	2:C:516:HOH:O	2.46	0.48
1:C:363:GLY:HA3	1:C:416:TRP:HB2	1.94	0.48
1:B:108:THR:HG21	1:B:194:ARG:HH21	1.79	0.48
1:D:155:LEU:HD13	1:D:164:ILE:HD11	1.95	0.47
1:C:109:THR:HG21	1:C:118:PRO:HG3	1.96	0.47
1:A:87:ILE:HD13	1:A:99:ALA:HB2	1.97	0.47
1:A:96:VAL:HB	1:A:436:ILE:HB	1.96	0.47
1:C:31:VAL:HG12	1:C:158:LEU:HD13	1.97	0.47
1:B:87:ILE:HD13	1:B:99:ALA:HB2	1.96	0.47
1:A:343:TYR:CG	1:A:349:ALA:HB2	2.50	0.46
1:C:86:ILE:HG12	1:C:98:VAL:HG22	1.98	0.46
1:C:366:ALA:HB3	1:C:385:ILE:HD13	1.98	0.46
1:D:390:VAL:O	1:D:392:GLN:HG2	2.15	0.45
1:A:187:ILE:HD13	1:A:425:LEU:HD21	1.99	0.45
1:B:176:ALA:HB1	1:B:194:ARG:HH22	1.82	0.45
1:C:108:THR:HG22	1:C:135:ALA:HB3	1.98	0.45
1:D:31:VAL:HG12	1:D:158:LEU:HD13	2.00	0.44
1:D:389:VAL:CG1	1:D:392:GLN:HG3	2.48	0.44
1:B:108:THR:HG22	1:B:135:ALA:HB3	1.99	0.44
1:D:187:ILE:HD11	1:D:430:LEU:HD21	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:96:VAL:HB	1:C:436:ILE:HB	1.98	0.44
1:C:155:LEU:HD13	1:C:164:ILE:HD11	2.00	0.44
1:B:230:ILE:HG23	1:B:245:GLU:HG2	1.99	0.44
1:D:255:TYR:CD1	1:D:342:LYS:HG2	2.53	0.44
1:B:96:VAL:HB	1:B:436:ILE:HB	1.99	0.44
1:A:102:LYS:N	1:A:130:ASN:OD1	2.47	0.44
1:A:176:ALA:HB1	1:A:194:ARG:HH22	1.83	0.44
1:A:121:ASN:ND2	2:A:517:HOH:O	2.48	0.43
1:D:22:GLU:HA	1:D:129:ARG:HD2	2.01	0.43
1:D:363:GLY:HA3	1:D:416:TRP:HB2	2.00	0.43
1:D:389:VAL:HG12	1:D:392:GLN:HG3	2.01	0.43
1:C:298:VAL:HG13	1:C:316:ILE:HD11	2.00	0.43
1:B:155:LEU:HD13	1:B:164:ILE:HD11	2.00	0.43
1:B:187:ILE:HD13	1:B:425:LEU:HD21	2.01	0.42
1:B:390:VAL:O	1:B:392:GLN:HG2	2.19	0.42
1:C:377:ARG:NH1	2:C:520:HOH:O	2.49	0.42
1:C:245:GLU:OE1	1:C:245:GLU:N	2.52	0.42
1:B:343:TYR:CG	1:B:349:ALA:HB2	2.55	0.42
1:C:390:VAL:O	1:C:392:GLN:HG2	2.19	0.42
1:A:7:ILE:HD11	1:A:175:ALA:HB1	2.02	0.41
1:A:47:PHE:HB3	1:A:115:VAL:HG21	2.02	0.41
1:A:212:ASN:HA	1:A:246:GLN:HG3	2.02	0.41
1:D:108:THR:HG22	1:D:135:ALA:HB3	2.01	0.41
1:A:390:VAL:O	1:A:392:GLN:HG2	2.19	0.41
1:B:245:GLU:OE1	1:B:245:GLU:N	2.46	0.41
1:C:400:SER:HB3	1:C:403:ASN:ND2	2.36	0.41
1:D:109:THR:HG21	1:D:118:PRO:HG3	2.02	0.41
1:D:212:ASN:HA	1:D:246:GLN:HG3	2.02	0.41
1:A:255:TYR:CD1	1:A:342:LYS:HG2	2.56	0.41
1:C:221:TYR:CE2	1:C:369:HIS:HB3	2.56	0.41
1:A:363:GLY:HA3	1:A:416:TRP:HB2	2.03	0.41
1:C:420:SER:HA	1:D:204:PRO:HB3	2.03	0.41
1:D:136:PRO:HG3	1:D:167:ILE:HD11	2.03	0.41
1:D:210:PRO:O	1:D:361:GLY:HA2	2.20	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	445/453 (98%)	437 (98%)	7 (2%)	1 (0%)	47	55
1	B	445/453 (98%)	437 (98%)	7 (2%)	1 (0%)	47	55
1	C	445/453 (98%)	437 (98%)	8 (2%)	0	100	100
1	D	445/453 (98%)	439 (99%)	5 (1%)	1 (0%)	47	55
All	All	1780/1812 (98%)	1750 (98%)	27 (2%)	3 (0%)	47	55

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	212	ASN
1	D	212	ASN
1	B	212	ASN

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	360/366 (98%)	358 (99%)	2 (1%)	86	93
1	B	360/366 (98%)	357 (99%)	3 (1%)	81	90
1	C	360/366 (98%)	357 (99%)	3 (1%)	81	90
1	D	360/366 (98%)	358 (99%)	2 (1%)	86	93
All	All	1440/1464 (98%)	1430 (99%)	10 (1%)	84	91

All (10) residues with a non-rotameric sidechain are listed below:

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	447/453 (98%)	0.38	21 (4%)	31	30	22, 32, 53, 103	0
1	B	447/453 (98%)	0.31	13 (2%)	51	49	21, 30, 46, 112	0
1	C	447/453 (98%)	0.42	20 (4%)	33	32	21, 30, 52, 102	0
1	D	447/453 (98%)	0.31	21 (4%)	31	30	20, 31, 50, 120	0
All	All	1788/1812 (98%)	0.36	75 (4%)	36	34	20, 30, 52, 120	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	170	ALA	7.4
1	B	170	ALA	6.4
1	D	169	GLU	5.5
1	A	441	LYS	5.4
1	D	170	ALA	5.3
1	B	173	ARG	5.1
1	C	173	ARG	4.9
1	D	173	ARG	4.6
1	A	173	ARG	4.5
1	C	441	LYS	4.3
1	A	175	ALA	4.2
1	B	174	GLU	4.2
1	D	441	LYS	4.1
1	D	344	ASP	3.8
1	B	175	ALA	3.7
1	B	445	VAL	3.7
1	C	449	GLU	3.6
1	A	443	ALA	3.6
1	C	170	ALA	3.5
1	A	172	SER	3.4
1	C	443	ALA	3.4

Continued on next page...

Mol	Chain	Res	Type
1	A	359	TYR
1	A	442	GLU
1	B	92	GLU
1	B	359	TYR
1	B	445	VAL
1	C	92	GLU
1	C	359	TYR
1	C	442	GLU
1	D	92	GLU
1	D	359	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	444	LYS	3.4
1	B	444	LYS	3.2
1	A	275	THR	3.1
1	D	446	PRO	3.0
1	D	443	ALA	2.9
1	C	208	VAL	2.9
1	B	169	GLU	2.8
1	A	445	VAL	2.7
1	A	169	GLU	2.7
1	A	421	ILE	2.7
1	A	444	LYS	2.7
1	D	447	SER	2.6
1	C	169	GLU	2.6
1	C	206	TYR	2.6
1	C	344	ASP	2.6
1	D	194	ARG	2.6
1	D	449	GLU	2.6
1	C	447	SER	2.6
1	A	174	GLU	2.6
1	A	411	LEU	2.5
1	C	425	LEU	2.5
1	D	444	LYS	2.4
1	C	453	GLY	2.4
1	C	445	VAL	2.4
1	A	442	GLU	2.4
1	C	446	PRO	2.3
1	D	208	VAL	2.3
1	A	425	LEU	2.3
1	B	448	TYR	2.3
1	D	411	LEU	2.3
1	D	453	GLY	2.3
1	B	208	VAL	2.3
1	C	442	GLU	2.2
1	C	9	GLU	2.2
1	D	273	GLU	2.2
1	B	209	GLY	2.2
1	A	194	ARG	2.2
1	B	8	LYS	2.2
1	B	453	GLY	2.2
1	D	445	VAL	2.1
1	C	273	GLU	2.1
1	D	159	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	194	ARG	2.1
1	D	420	SER	2.1
1	A	449	GLU	2.1
1	C	420	SER	2.1
1	D	174	GLU	2.0
1	A	189	THR	2.0
1	A	447	SER	2.0
1	B	196	LYS	2.0
1	D	425	LEU	2.0
1	A	429	HIS	2.0
1	A	8	LYS	2.0
1	D	117	THR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

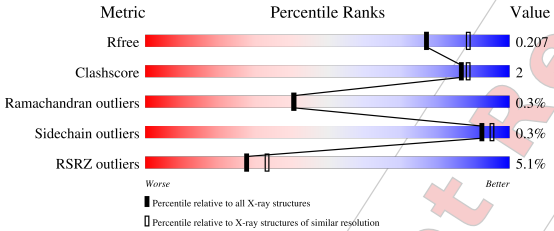
There are no such residues in this entry.

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	453	5% 94% 5%
1	B	453	5% 92% 7%
1	D	453	5% 94% 5%
2	C	454	5% 93% 6%

Validation Pipeline (wwPDB-VP) : 2.32.1

Full wwPDB X-ray Structure Validation Report (i)

Feb 7, 2023 – 06:49 pm GMT

PDB ID : 8CEJ
Title : Succinyl-CoA Reductase from Clostridium kluyveri (SucD) with Mesaconyl-C1-CoA
Deposited on : 2023-02-02
Resolution : 2.10 Å(reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.32.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15096 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Succinate-semialdehyde dehydrogenase (acetylating).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	0	0
			3415	2162	575	663	15			
1	B	449	Total	C	N	O	S	0	0	0
			3415	2162	575	663	15			
1	D	449	Total	C	N	O	S	0	0	0
			3415	2162	575	663	15			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	242	UL3	CYS	conflict	UNP P38947
B	242	UL3	CYS	conflict	UNP P38947
D	242	UL3	CYS	conflict	UNP P38947

- Molecule 2 is a protein called Succinate-semialdehyde dehydrogenase (acetylating).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	450	Total	C	N	O	P	S	0	0	0
			3463	2183	582	679	3	16			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	501	OA9	-	expression tag	UNP P38947

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	299	Total	O	0	0
			299	299		
3	B	372	Total	O	0	0
			372	372		

Continued on next page...

Continued from previous page...

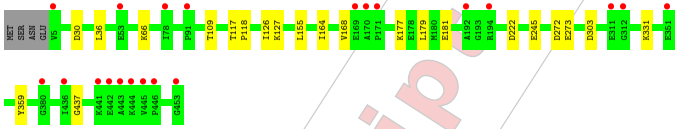
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	342	Total	O	0	0
			342	342		
3	D	375	Total	O	0	0
			375	375		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

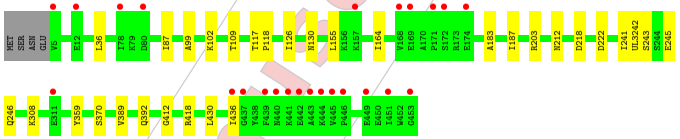
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

Chain A:



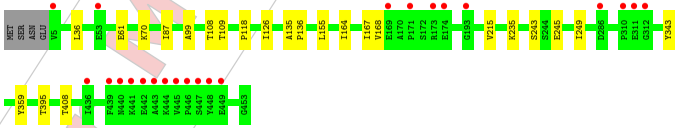
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

Chain B:



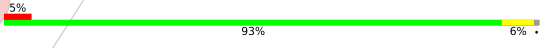
- Molecule 1: Succinate-semialdehyde dehydrogenase (acetylating)

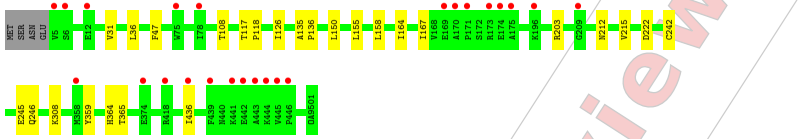
Chain D:



- Molecule 2: Succinate-semialdehyde dehydrogenase (acetylating)

Chain C:





4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	141.30Å 189.66Å 189.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.89 – 2.10 24.89 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.89-2.10) 99.7 (24.89-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	306.98 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.186 , 0.210 0.187 , 0.207	Depositor DCC
R_{free} test set	1990 reflections (1.35%)	wwPDB-VP
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 28.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.417 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15096	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.07 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3481e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: OA9, UL3

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.24	0/3457	0.43	0/4674
1	B	0.26	0/3457	0.43	0/4674
1	D	0.25	0/3457	0.43	0/4674
2	C	0.25	0/3464	0.43	0/4685
All	All	0.25	0/13835	0.43	0/18707

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3415	0	3450	15	0
1	B	3415	0	3450	17	1
1	D	3415	0	3450	12	0
2	C	3463	0	3455	13	1
3	A	299	0	0	7	1
3	B	372	0	0	2	3
3	C	342	0	0	2	3
3	D	375	0	0	1	1
All	All	15096	0	13805	57	6

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (57) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:36:LEU:HD11	2:C:126:ILE:HD12	1.80	0.62
1:D:36:LEU:HD11	1:D:126:ILE:HD12	1.81	0.61
1:A:36:LEU:HD11	1:A:126:ILE:HD12	1.83	0.61
1:A:66:LYS:NZ	3:A:505:HOH:O	2.34	0.60
2:C:365:THR:O	3:C:601:HOH:O	2.17	0.58
1:A:303:ASP:OD2	3:A:501:HOH:O	2.18	0.56
2:C:245:GLU:OE1	2:C:245:GLU:N	2.38	0.55
1:D:87:ILE:HD13	1:D:99:ALA:HB2	1.88	0.54
1:D:109:THR:HG21	1:D:118:PRO:HG3	1.88	0.54
1:A:109:THR:HG21	1:A:118:PRO:HG3	1.88	0.54
1:B:389:VAL:HG12	1:B:392:GLN:HG3	1.90	0.53
1:D:243:SER:OG	1:D:395:THR:HA	2.08	0.52
1:B:155:LEU:HD13	1:B:164:ILE:HD11	1.90	0.52
1:B:308:LYS:NZ	3:B:516:HOH:O	2.42	0.52
1:B:36:LEU:HD11	1:B:126:ILE:HD12	1.92	0.50
1:A:245:GLU:OE1	1:A:245:GLU:N	2.38	0.50
1:A:177:LYS:NZ	3:A:518:HOH:O	2.44	0.50
2:C:308:LYS:NZ	3:C:625:HOH:O	2.42	0.49
1:B:183:ALA:O	1:B:203:ARG:NH1	2.40	0.49
1:B:218:ASP:HB2	1:B:370:SER:HB2	1.93	0.49
1:D:235:LYS:NZ	3:D:532:HOH:O	2.44	0.48
2:C:155:LEU:HD13	2:C:164:ILE:HD11	1.95	0.48
1:B:245:GLU:OE1	1:B:245:GLU:N	2.37	0.48
2:C:136:PRO:HG3	2:C:167:ILE:HD11	1.96	0.47
2:C:242:CYS:SG	2:C:364:HIS:NE2	2.87	0.47
2:C:215:VAL:HG21	2:C:245:GLU:HG3	1.96	0.47
1:D:61:GLU:OE1	1:D:61:GLU:N	2.45	0.47
1:D:249:ILE:HG22	1:D:343:TYR:HD2	1.79	0.47
1:A:155:LEU:HD13	1:A:164:ILE:HD11	1.96	0.47
1:D:70:LYS:NZ	1:D:408:THR:O	2.47	0.46
1:D:155:LEU:HD13	1:D:164:ILE:HD11	1.98	0.46
2:C:47:PHE:CZ	2:C:150:LEU:HD12	2.50	0.46
1:A:331:LYS:NZ	3:A:531:HOH:O	2.48	0.46
1:D:108:THR:HG22	1:D:135:ALA:HB3	1.98	0.45
1:B:187:ILE:HD11	1:B:430:LEU:HD21	1.99	0.45
1:B:241:ILE:HG22	1:B:243:SER:H	1.81	0.45
1:D:215:VAL:HG21	1:D:245:GLU:CG	2.46	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:177:LYS:NZ	1:A:181:GLU:OE2	2.46	0.45
1:B:87:ILE:HD13	1:B:99:ALA:HB2	1.99	0.44
1:B:102:LYS:N	1:B:130:ASN:OD1	2.50	0.44
1:A:117:THR:HB	1:A:118:PRO:HD3	1.98	0.44
1:B:412:GLY:N	3:B:532:HOH:O	2.50	0.44
1:D:136:PRO:HG3	1:D:167:ILE:HD11	1.99	0.44
1:A:30:ASP:OD2	3:A:502:HOH:O	2.21	0.44
1:B:241:ILE:HG23	1:B:242:UL3:C11	2.48	0.44
1:A:30:ASP:OD1	1:A:127:LYS:NZ	2.45	0.43
1:A:273:GLU:OE2	3:A:503:HOH:O	2.21	0.43
1:A:168:VAL:HG21	1:A:179:LEU:HD22	2.00	0.42
2:C:31:VAL:HG12	2:C:158:LEU:HD13	2.01	0.42
1:B:212:ASN:HA	1:B:246:GLN:HG3	2.02	0.42
2:C:117:THR:HB	2:C:118:PRO:HD3	2.02	0.42
2:C:212:ASN:HA	2:C:246:GLN:HG3	2.02	0.42
2:C:108:THR:HG22	2:C:135:ALA:HB3	2.01	0.41
1:A:272:ASP:OD1	3:A:504:HOH:O	2.22	0.41
1:B:117:THR:HB	1:B:118:PRO:HD3	2.03	0.41
1:B:109:THR:HG21	1:B:118:PRO:HG3	2.03	0.41
1:B:109:THR:HG22	1:B:117:THR:HG22	2.02	0.40

All (6) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:418:ARG:NH1	2:C:203:ARG:NH1[7_444]	2.05	0.15
3:B:853:HOH:O	3:C:865:HOH:O[7_444]	2.05	0.15
3:D:751:HOH:O	3:D:751:HOH:O[3_555]	2.06	0.14
3:B:714:HOH:O	3:C:679:HOH:O[7_444]	2.12	0.08
3:A:637:HOH:O	3:A:680:HOH:O[3_554]	2.14	0.06
3:B:710:HOH:O	3:C:840:HOH:O[7_444]	2.18	0.02

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	446/453 (98%)	433 (97%)	11 (2%)	2 (0%)	34	32
1	B	446/453 (98%)	431 (97%)	13 (3%)	2 (0%)	34	32
1	D	446/453 (98%)	434 (97%)	12 (3%)	0	100	100
2	C	447/454 (98%)	435 (97%)	10 (2%)	2 (0%)	34	32
All	All	1785/1813 (98%)	1733 (97%)	46 (3%)	6 (0%)	41	41

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	437	GLY
1	A	222	ASP
1	B	222	ASP
2	C	222	ASP
2	C	436	ILE
1	B	436	ILE

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	361/365 (99%)	360 (100%)	1 (0%)	92	95
1	B	361/365 (99%)	360 (100%)	1 (0%)	92	95
1	D	361/365 (99%)	359 (99%)	2 (1%)	86	90
2	C	362/366 (99%)	361 (100%)	1 (0%)	92	95
All	All	1445/1461 (99%)	1440 (100%)	5 (0%)	92	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	359	TYR
1	B	359	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	359	TYR
1	D	168	VAL
1	D	359	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	448/453 (98%)	0.67	21 (4%)	31	37	19, 27, 46, 150	0
1	B	448/453 (98%)	0.60	24 (5%)	25	31	18, 24, 47, 187	0
1	D	448/453 (98%)	0.64	23 (5%)	28	33	19, 25, 46, 192	0
2	C	449/454 (98%)	0.63	24 (5%)	26	32	18, 24, 49, 188	0
All	All	1793/1813 (98%)	0.64	92 (5%)	28	33	18, 25, 49, 192	0

All (92) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	443	ALA	14.2
1	D	442	GLU	13.4
1	A	443	ALA	8.6
1	B	443	ALA	8.1
1	B	442	GLU	8.1
2	C	442	GLU	7.6
1	A	444	LYS	7.1
1	D	441	LYS	6.8
1	A	446	PRO	6.8
1	D	312	GLY	6.7
2	C	444	LYS	6.6
2	C	443	ALA	6.4
1	B	441	LYS	5.6
1	B	444	LYS	5.0
1	A	442	GLU	4.7
1	B	446	PRO	4.6
2	C	441	LYS	4.3
1	B	453	GLY	4.2
1	A	445	VAL	4.2
1	B	445	VAL	4.0
1	D	169	GLU	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	311	GLU	3.9
2	C	436	ILE	3.9
1	D	446	PRO	3.8
1	D	445	VAL	3.7
1	A	436	ILE	3.7
1	D	449	GLU	3.7
2	C	439	PHE	3.7
1	A	169	GLU	3.6
1	A	441	LYS	3.5
1	B	174	GLU	3.5
1	B	449	GLU	3.4
1	B	172	SER	3.4
1	D	173	ARG	3.4
1	B	439	PHE	3.4
2	C	171	PRO	3.2
1	B	440	ASN	3.1
1	A	453	GLY	3.0
1	D	447	SER	3.0
1	B	169	GLU	3.0
1	D	444	LYS	3.0
1	D	440	ASN	2.9
2	C	6	SER	2.8
1	D	5	VAL	2.7
1	A	170	ALA	2.7
1	B	5	VAL	2.7
1	B	168	VAL	2.7
1	B	171	PRO	2.7
2	C	12	GLU	2.7
1	D	310	PRO	2.6
1	A	311	GLU	2.6
1	A	312	GLY	2.6
2	C	78	ILE	2.6
1	A	171	PRO	2.6
1	D	53	GLU	2.5
2	C	169	GLU	2.5
1	D	174	GLU	2.5
2	C	170	ALA	2.5
2	C	75	TRP	2.4
1	A	53	GLU	2.4
1	B	437	GLY	2.4
1	D	286	ASP	2.4
2	C	445	VAL	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	351	GLU	2.4
2	C	358	MET	2.4
2	C	5	VAL	2.4
2	C	175	ALA	2.4
1	A	78	ILE	2.4
1	B	436	ILE	2.4
2	C	196	LYS	2.3
1	D	171	PRO	2.3
1	A	5	VAL	2.3
2	C	446	PRO	2.3
1	D	448	TYR	2.3
2	C	374	GLU	2.3
1	A	194	ARG	2.2
1	B	157	LYS	2.2
1	B	451	ILE	2.2
1	D	436	ILE	2.2
2	C	418	ARG	2.2
1	D	439	PHE	2.1
1	A	380	GLY	2.1
1	D	193	GLY	2.1
1	B	78	ILE	2.1
2	C	209	GLY	2.1
1	A	91	PRO	2.1
1	B	80	ASP	2.1
2	C	173	ARG	2.1
1	D	311	GLU	2.1
1	A	192	ALA	2.1
2	C	174	GLU	2.1
1	B	12	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.

Die vorliegende Dissertation wurde von September 2018 bis Februar 2023 am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg unter Leitung von Prof. Dr. Tobias J. Erb angefertigt.

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

Erstgutachter: Prof. Dr. Tobias J. Erb
Zweitgutachter: Prof. Dr. Lars O. Essen

Tag der Disputation: _____

Einverständniserklärung

Ich erkläre mich hiermit einverstanden, dass die vorliegende Arbeit

Engineering Enzymes and Pathways for Alternative CO_2 Fixation and Glyoxylate Assimilation

in Bibliotheken allgemein zugänglich gemacht wird. Dazu gehört, dass sie

- von der Bibliothek der Einrichtung, in der ich meine Arbeit angefertigt habe, zur Benutzung in ihren Räumen bereit gehalten wird;
- in konventiellen un maschinenlesbaren Katalogen, Verzeichnissen und Datenbanken verzeichnet wird;
- der UB für lokale Nutzung und für Fernleihe zur Verfügung steht;
- im Rahmen der urheberrechtlichen Bestimmungen für Kopierzwecke genutzt werden kann.

Marburg, June 21, 2023

Pascal Pfister (Autor) Tobias Erb (Betreuender Hochschullehrer)

Eidesstattliche Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel "Engineering Enzymes and Pathways for Alternative CO_2 Fixation and Glyoxylate Assimilation" 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, June 21, 2023
Pascal Pfister