Implementation of CO_2 fixation pathways into *Methylorubrum extorquens* AM1

Dissertation

kumulativ

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer.nat.)

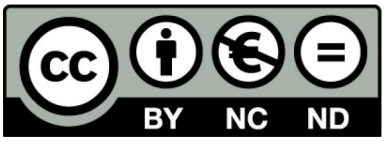
des Fachbereichs Biologie der Philipps-Universität Marburg

Vorgelegt von

Martina Carrillo Camacho

Aus Quito, Ecuador

Marburg, 2020



Die vorliegende Dissertation wurde von September 2015 bis April 2020 am Fachbereich Biologie unter Leitung von Prof. Dr. Tobias J. Erb angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am 08.07.2020.

Erstgutachter: Prof. Dr. Tobias J. Erb

Zweitgutachterin: Prof. Dr. Anke Becker

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Lennart Randau

Prof. Dr. Victor Sourjik

Tag der Disputation: 17.07.2020.

Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „**Implementation of CO_2 fixation pathways into *Methylorubrum extorquens* AM1**“ selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe. Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 23. April 2020

Martina Carrillo Camacho

“So let us then try to climb the mountain,

not by stepping on what is below us,

but to pull us up at what is above us,

for my part at the stars…”

M. C. Escher

A mathematician and artist.

TABLE OF CONTENTS

PART ONE: INITIAL REMARKS

SUMMARY

Growth is a prerequisite for life and metabolism plays a central role to achieve and sustain growth. Its function is to obtain energy needed for metabolic (inter)conversions as well as converting nutrients into cellular building blocks used to synthesize macromolecules and thereby the cell itself. The metabolic pathways present in an organism define and limit its growth capabilities. Nowadays, an organism's core metabolic network might be altered with relative ease through genetic engineering. The modification of biological systems to expand their growth potential was at the core of this thesis with an emphasis on the Alphaproteobacterium *Methylorubrum extorquens* AM1.

The following results were achieved:

1. Novel inducible promoters were designed for *M. extorquens* AM1 based on the LacI-*lacO* system that are tight before induction and exhibit good dynamic range. The expression level achieved with these promoters exceeds that of the standard P_{mxaF} promoter making them the strongest promoters described to date for this organism. New extrachromosomal elements termed 'mini-chromosomes' based on *repABC* cassettes were also identified, which maintain a stable unit copy number and faithful inheritance even in the absence of selection. These new DNA vehicles are compatible with each other and with the broad host plasmid used for *M. extorquens* AM1 so far. A conditionally unstable replicon, Mex-CM4, showed the behavior needed to establish CRISPR-Cas and similar techniques for which transient expression and fast extinction of the replicon are prerequisites.

2. The Calvin-Benson-Bassham (CBB) cycle was introduced into *M. extorquens* AM1 for heterologous CO_2 fixation and assimilation of methanol was disrupted to provide the cell with energy consistent with (chemo)organoautotrophic growth. The operation of a functional CBB cycle was confirmed with ^{13}C-tracer analysis in the engineered strain. Furthermore, a positive growth phenotype and increased cell viability was found dependent on a functional RubisCO.

3. *M. extorquens* AM1 was rationally engineered and further evolved by serial transfer to improve growth on the novel substrate ribose. Mutations throughout the evolutionary landscape were identified by whole genome (re)sequencing and further characterized genetically and biochemically. The CBB cycle was introduced into ribose evolved clones for simultaneous CO_2 fixation and acetate was provided for additional energy. These strains were evolved in semi-chemostat conditions to promote carbon fixation via the CBB cycle.

4. *In vivo* implementation of a synthetic CO_2 fixation cycle, the **c**rotonyl-CoA/**et**hylmalonyl-**C**oA/**h**ydroxybutyryl-CoA (**CETCH**) cycle, was explored

in *M. extorquens* AM1. Suitable candidate enzymes were identified and characterized biochemically for every step of the crotonyl-CoA regeneration module. These new candidates were combined to close the CETCH cycle and various *M. extorquens* AM1 strains were created to test growth-based selection schemes.

ZUSAMMENFASSUNG

Wachstum ist eine Voraussetzung für das Leben und der Stoffwechsel spielt eine zentrale Rolle, um Wachstum zu erreichen und zu erhalten. Seine Funktion besteht in der Konservierung von Energie, die für die metabolischen Reaktionen benötigt wird, sowie in der Umwandlung von Nährstoffen in zelluläre Bausteine, die zur Synthese von Makromolekülen und damit der Zelle selbst dienen. Die in einem Organismus vorhandenen Stoffwechselwege definieren und begrenzen daher seine Wachstumsfähigkeit. Heutzutage kann das zentrale Stoffwechselnetzwerk eines Organismus relativ einfach mithilfe der Gentechnik verändert werden. Die Modifikation biologischer Systeme zur Erweiterung ihres Wachstumspotentials stand im Mittelpunkt dieser Arbeit, wobei der Fokus auf das Alphaproteobakterium *Methylorubrum extorquens* AM1 gelegt wurde.

Die folgenden Ergebnisse wurden erzielt:

1. Es wurden neuartige induzierbare Promotoren für *M. extorquens* AM1 auf der Basis des LacI-*lacO*-Systems entwickelt, die vor der Induktion dicht sind und einen guten dynamischen Bereich aufweisen. Das mit diesen Promotoren erreichte Expressionsniveau übertrifft das des Standard-P_{mxaF}-Promotors, was sie zu den stärksten Promotoren macht, die bisher für diesen Organismus beschrieben wurden. Weiterhin wurden neue extrachromosomale Elemente identifiziert, die als "Mini-Chromosomen" bezeichnet werden und auf *repABC*-Kassetten basieren, die auch ohne Selektion eine stabile Einheitskopienzahl und eine sichere Vererbung erhalten. Diese neuen DNA-Vehikel sind untereinander und mit dem Broad-Host-(Range-)Plasmid, das bisher für *M. extorquens* AM1 verwendet wurde, kompatibel. Ein bedingt instabiles Replikon, Mex-CM4, zeigte das Verhalten, das zur Etablierung von CRISPR-Cas und ähnlichen Techniken erforderlich und für die eine transiente Expression und eine schnelle Auslöschung des Replikons Voraussetzung ist.
2. Der Calvin-Benson-Bassham (CBB)-Zyklus wurde in *M. extorquens* AM1 zur heterologen CO_2-Fixierung eingeführt. Dazu wurde die Kohlenstoffassimilierung aus Methanol unterbrochen, um die Zelle im Einklang mit dem (chemo)organotrophen Wachstum mit Energie zu versorgen. Der Betrieb eines funktionellen CBB-Zyklus wurde mit einer ^{13}C-Tracer-Analyse in dem manipulierten Stamm bestätigt. Darüber hinaus wurde ein positiver Wachstumsphänotyp und eine erhöhte Zellviabilität in Verbindung mit einer funktionellen RubisCO beobachtet.
3. *M. extorquens* AM1 wurde rational konstruiert und durch serielle Übertragung evolviert, um das Wachstum auf dem neuartigen Substrat Ribose zu verbessern. Mutationen während der gesamten Evolution wurden durch (Re-)Sequenzierung

des kompletten Genoms identifiziert und genetisch und biochemisch detailliert charakterisiert. Der CBB-Zyklus wurde in die Ribose-evolvierten Klone zur gleichzeitigen CO_2-Fixierung eingeführt, und Acetat lieferte zusätzliche Energie. Diese Stämme wurden unter semi-chemostatischen Bedingungen weiterevolviert, um die Kohlenstoff-Fixierung über den CBB-Zyklus zu fördern.

4. Die *In-vivo*-Implementierung eines synthetischen CO_2-Fixierungszyklus, des **C**rotonyl-CoA/**Et**hylmalonyl-**C**oA/**H**ydroxybutyryl-CoA (**CETCH**)-Zyklus, wurde in *M. extorquens* AM1 untersucht. Geeignete Enzymkandidaten für jeden Schritt des Crotonyl-CoA-Regenerationsmoduls wurden identifiziert und biochemisch charakterisiert. Diese neuen Kandidaten wurden miteinander kombiniert, um den CETCH-Zyklus zu schließen, und verschiedene *M. extorquens* AM1 Stämme wurden geschaffen, um wachstumsbasierte Selektionsstrategien zu testen.

PUBLICATIONS

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

Carrillo, M., Wagner, M., Petit, F., Dransfeld, A., Becker, A., and Erb, T. J. (2019) Design and control of extrachromosomal elements in *Methylorubrum extorquens* AM1. *ACS Synthetic Biology* **8**, 2451-2456.

Schada von Borzyskowski, L., Carrillo, M., Leupold, S., Glatter, T., Kiefer, P., Weishaupt, R., Heinemann, M., and Erb, T. J. (2018) An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. *Metab. Eng.* **47**, 423-433.

Carrillo, M., Peeck, L., Klose, M., Hervé, V., Brune, A., and Erb, T. J. Engineering and evolving a CO_2 fixing ribose metabolism in *Methylorubrum extorquens* AM1. (*draft manuscript*).

Carrillo, M., Schwander, T., Scheffen, M., Erb, T. J. Exploring *in vivo* implementation of synthetic carbon fixation cycles. (*draft manuscript*).

Publications that are not discussed in this thesis:

Döhlemann, J., Wagner, M., Happel, C., Carrillo, M., Sobetzko, P., Erb, T. J., Thanbichler, M., and Becker, A. (2017) A Family of Single Copy *repABC*-Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium *Sinorhizobium meliloti*. *ACS Synthetic Biology* **6**, 968-984.

Scheffen, M., Marchal, D. G., Beneyton, T., Schuller, S. K., Klose, M., Diehl, C., Lehmann, J., Pfister, P., Carrillo, M., Aslan, S., Cortina, N. S., Claus, P., Bollschweiller, D., Baret, J. C., Schuller, J. M., Zarzycki, J., Bar-Even, A., Erb, T. J. A new-to-nature carboxylation module improves natural and synthetic CO_2 fixation. (*submitted*).

INTRODUCTION

1.1 Metabolism

Theodor Schwann proposed the term 'metabolism' from the Greek μεταβολή (meaning change) and defined it as the complete set of chemical reactions in an organism. Its role in cells is to (i) obtain energy needed for metabolic (inter)conversions and (ii) convert nutrients to cellular building blocks used to synthesize macromolecules such as proteins, nucleic acids, and lipids. This view might be expanded to include the transport of nutrients or other compounds and the elimination of waste products. Because growth is a prerequisite for life and metabolism plays a central role to achieve and sustain growth, the mode in which an organism acquires nutrients (*e.g.* carbon and energy) is a defining characteristic. As such, organisms can be classified quite simply as being autotrophic or heterotrophic: the former can fix inorganic carbon into organic biomass, while the latter consumes previously fixed organic carbon. Schönheit *et al*. (2016) defined autotrophy when more than 50% of cellular carbon is the result of CO_2 fixation; otherwise, an organism is considered a heterotroph[1]. Similarly, prefixes are used to describe whether energy is harvested from light (photo-) or chemical molecules (chemo-). A further subclassification exists to describe whether an organic or inorganic chemical electron donor is used: (chemo)organotrophs and (chemo)litotrophs, respectively. The metabolic pathways present in an organism determine such distinctions, as well as the metabolic potential of an organism. With the onset of genetic engineering, an organism's core metabolic network might be altered with relative ease, giving rise to metabolic engineering and synthetic biology. The aims of these fields can be summarized as the manipulation and re-design of biological systems for useful purposes, which often entail the implementation of metabolic pathways into organisms to expand their potential. This objective was at the core of this thesis with an emphasis on the Alphaproteobacterium *Methylorubrum extorquens* AM1.

1.2 *Methylorubrum extorquens* AM1

The well-known model organism, *M. extorquens* AM1 was first isolated as an **a**irborne contaminant of **m**ethylamine by Peel and Quayle in 1961, who designated it *Pseudomonas* **AM1**[2]. Its genus was later renamed to *Methylobacterium* and recently reclassified as *Methylorubrum extorquens* AM1[3, 4]. It belongs to the non-taxonomic group of pink-pigmented facultative methylotrophic (PPFM) bacteria[5]. Sequencing revealed a multi-partite genome consisting of five replicons: the main chromosome, one megaplasmid, and three smaller plasmids[6]. As a strictly aerobic, heterotrophic organism, *M. extorquens* AM1 has a complete tricarboxylic acid (TCA) cycle to grow on a variety of organic substrates. Strikingly, C_6 and C_5 sugars are not utilized as sole carbon sources, even though the pentose phosphate pathway (PPP), as well as gluconeogenic and glycolytic genes, are present in the genome[5, 6]. As a representative of PPFM bacteria, *M. extorquens*

AM1 has been extensively studied for its ability to grow on reduced C_1 carbon sources such as methanol, formate, and methylamine.

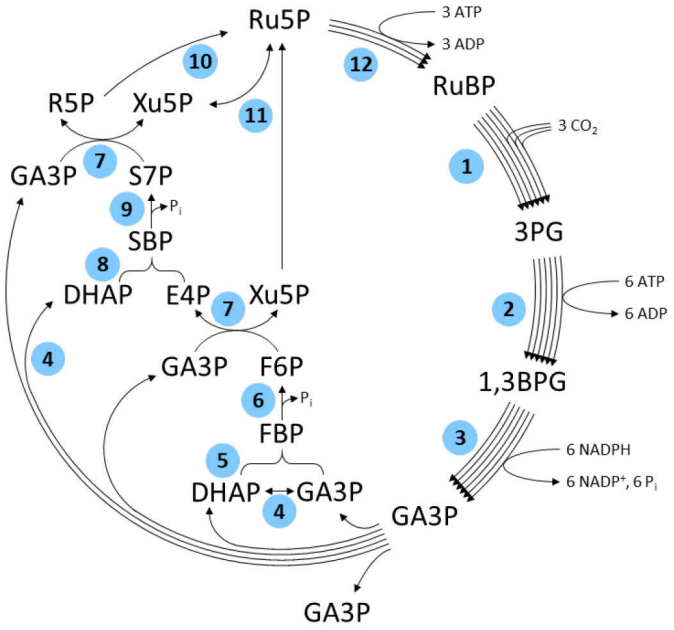
In this organism, C_1 carbon sources are assimilated heterotrophically via the serine cycle at the level of formate, while energy, in the form of reducing equivalents, is conserved by their complete oxidation to CO_2. A long-standing question surrounding the metabolic network of *M. extorquens* AM1 was the formation of glyoxylate – used as carbon backbone for C_1 assimilation and in anaplerosis during growth on C_2 carbon sources – since it lacks isocitrate lyase activity. This was solved by the elucidation of the ethylmalonyl-CoA pathway (EMCP) in which two carboxylation steps take place to yield a molecule of glyoxylate[7-9]. Being genetically tractable[10, 11], many genetic studies were performed to elucidate the complex metabolic network even before its genome was sequenced[6]. Indeed, several genetic tools are available, which have been summarized in Table 1[12-17]. Several bulk and value-added chemicals produced in *M. extorquens*, including the sesquiterpenoid α-humulene[18], 3-hydroxypropionate[19], and 1-butanol[20], are becoming increasingly relevant in a methanol- and formate-based bioeconomy[21]. All of these characteristics make *M. extorquens* AM1 an ideal organism for further study in this thesis.

1.3 Natural and synthetic autotrophy

Six autotrophic CO_2 fixation pathways have been described in nature so far[22]. Almost all of them are cycles in which CO_2 is fixed onto an existing carbon backbone that must be regenerated before a new round of CO_2 fixation can take place. An equilibrium between metabolites drained out of the cycle and incoming carbon is critical during autotrophic growth[23]. The reductive acetyl-CoA (Wood-Ljungdahl) pathway is an exception, since both carbons in acetyl-CoA originate from CO_2[24-26]. From the remaining autotrophic pathways, the Calvin-Benson-Bassham (CBB) cycle is the predominant CO_2 fixation cycle operating in plants, algae, cyanobacteria as well as many aerobic proteobacteria.

First characterized in the 1940s and 1950s, the CBB cycle entails the electrophilic addition of CO_2 to ribulose 1,5-bisphosphate (RuBP) yielding two molecules of 3-phosphoglycerate (3PG)[27-30]. This compound is reduced by gluconeogenic enzymes to glyceraldehyde 3-phosphate (GA3P), which can be used for cellular biosynthesis after three turns of the cycle. The regeneration of RuBP follows the interconversion of phosphorylated intermediates catalyzed by enzymes shared with the PPP to yield ribulose 5-phosphate (Ru5P), which is further phosphorylated to RuBP, closing the cycle (see **Fig. 1**). The key enzymes of this cycle are ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) and phosphoribulokinase (Prk). Sedoheptulose bisphosphatase (SBPase) can also be regarded as specific for this autotrophic cycle. The widespread nature of the CBB cycle is partly due to oxygen tolerance of all its enzymes, even if RubisCO accepts molecular oxygen as substrate resulting in the toxic product, 2-phosphoglycolate, which must be detoxified by the cell[31-33]. The large overlap with essential biosynthetic routes is another plausible

explanation to its success because only RubisCO and Prk are needed to close the PPP into an autotrophic CBB cycle. Finally, many organisms with a functional CBB cycle are only facultative autotrophs and will grow heterotrophically if organic carbon is available.



$$3\ CO_2 + 9\ ATP + 6\ NADPH \rightarrow 1\ GA3P + 9\ ADP + 8\ P_i + 6\ NADP^+$$

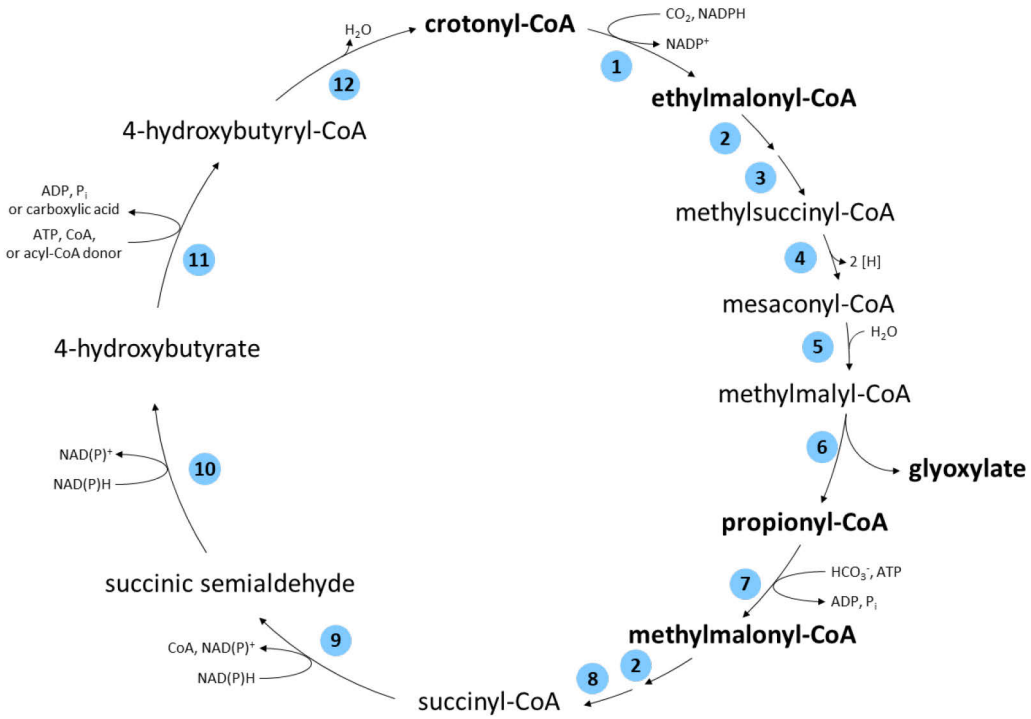
Fig. 1. The Calvin-Benson-Bassham (CBB) cycle. Enzymes: 1: ribulose 1,5-bisphosphate carboxylase/oxygenase; 2: 3-phosphoglycerate kinase; 3: glyceraldehyde 3-phosphate dehydrogenase; 4: triosephosphate isomerase; 5: fructose 1,6-bisphosphate aldolase; 6: fructose 1,6-bisphosphate phosphatase; 7: transketolase; 8: sedoheptulose 1,7-bisphosphate aldolase; 9: sedoheptulose 1,7-bisphosphate phosphatase; 10: ribose 5-phosphate isomerase; 11: ribulose-phosphate epimerase; and 12: phosphoribulokinase.

The alteration of an organism's metabolism by rational means is seen as the next logical step to go from understanding to rewiring a biological system. Much effort has been put into characterizing and broadening the spectrum of organic carbon sources an organism can grow on. While still immensely challenging, these approaches are often met with success as long as the uptake and catabolic pathways are well characterized. A more radical approach is to convert a heterotrophic organism into an autotrophic one. The first challenge lies in ensuring that the cell can still generate enough energy to sustain itself; followed by fine-tuning of CO_2 fixation *versus* metabolites drained out of the cycle for biomass generation.

Recent publications have engineered synthetic autotrophic growth through the CBB cycle in various model organisms: *Escherichia coli*[34, 35], *M. extorquens*[36] (as discussed in CHAPTER II and III of this thesis), and the yeast *Pichia pastoris*[37]. The first was a study where implementation of the CBB cycle into *Escherichia coli* yielded a so-called

"hemiautotrophic" growth[34]. The authors separated upper and lower metabolism by deleting phosphoglycerate mutase and fed pyruvate for energy and carbon for all biosynthetic routes downstream of 2-phosphoglycerate. Approximately 30% of cellular carbon originated from CO_2 fixation in upper metabolism, falling short of the 50% rule for autotrophic growth[1]. This was nevertheless a big step forward and required adaptive laboratory evolution (ALE) in a bioreactor, using xylose as a helper carbon source, before hemiautotrophic growth was possible. A follow up paper uncovered the mutations responsible for hemiautotrophic growth in *E. coli*[38]. In 2019, full (chemo)organoautotrophic growth of *E. coli* was achieved using the CBB cycle for carbon fixation and formate oxidation provided energy in the form of reducing equivalents[35]. In said study, ALE by feeding limiting amounts of xylose, as done previously[34], produced evolved clones that no longer relied on sugars and fixed all cellular carbon with RubisCO. The methylotrophic yeast *P. pastoris* naturally assimilates methanol by oxidizing it to formaldehyde, which is then assimilated via the xylulose monophosphate (XuMP) cycle, an alternative to the serine cycle of *M. extorquens*[39]. Thus, Gassler *et al.* (2019) blocked methanol assimilation via the XuMP cycle to exclusively generate energy from methanol and confined the heterologous CBB cycle to the peroxisomes for carbon fixation[37]. They obtained chemoorganoautotrophic growth on methanol and elevated CO_2 by expressing CBB enzymes. ALE was performed to improve the μ_{max} from 0.008 to 0.018 h^{-1}. Both studies constitute proof of principle that heterotrophic organisms can be converted into autotrophs through a combination of rational engineering and ALE.

Artificial CO_2 fixation cycles have been designed as alternate solutions to those found in nature. The objective is to overcome some limitations of natural CO_2 fixation cycles such as rate-limiting carboxylation steps, low energy efficiency, or sensitivity to oxygen[22]. In one study, several synthetic cycles were proposed around PEP carboxylase, chosen for its catalytic efficiency[40]. However, no cycle proposed in said study has been realized so far neither *in vitro* nor *in vivo*. More recently, seven artificial CO_2 fixation cycles designed around biochemically feasible conversions (*i.e.* fall into one of the E.C. standard classes) and requiring simple cofactors were proposed[41]. These cycles were all built around the family of enoyl-CoA carboxylases/reductases since they have good catalytic properties. Additional design criteria were favorable overall thermodynamics and low energetic costs. In said study, one of the cycles, the crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle, was realized *in vitro* using purified proteins and converted CO_2 into organic matter at a rate of 5 nmoles min^{-1} mg^{-1} protein. The CETCH cycle starts with the reductive carboxylation of crotonyl-CoA, followed by a series of carbon backbone rearrangements yielding methylmalyl-CoA. This compound is cleaved into propionyl-CoA and glyoxylate, the output compound of the cycle. Propionyl-CoA is carboxylated once more to methylmalonyl-CoA and rearranged to succinyl-CoA, which is used to regenerate crotonyl-CoA in four additional steps, thus closing the cycle (see **Fig. 2**). The implementation of the CETCH cycle *in vivo* into various organisms is ongoing in our laboratory and will be addressed for *M. extorquens* AM1 in CHAPTER IV of this thesis.



$CO_2 + HCO_3^- + 3\ NAD(P)H + 2\ ATP + FAD \rightarrow$ glyoxylate + $3\ NAD(P)^+$ + $2\ ADP + 2\ P_i + FADH_2$

Fig. 2. The CETCH cycle. Enzymes: 1: crotonyl-CoA carboxylase/reductase; 2: ethylmalonyl-CoA/ methylmalonyl-CoA epimerase; 3: ethylmalonyl-CoA mutase; 4: methylsuccinyl-CoA dehydrogenase; 5: mesaconyl-CoA hydratase, 6: β-methylmalyl-CoA lyase; 7: propionyl-CoA carboxylase, 8: methylmalonyl-CoA mutase; 9: succinyl-CoA reductase; 10: succinic semialdehyde reductase; 11: 4-hydroxybutyryl-CoA synthetase (a) or 4-hydroxybutyrate:CoA transferases (b); 12: 4-hydroxybutyryl-CoA dehydratase. Both carboxylation steps and the output molecule formed (glyoxylate) are highlighted in bold.

1.4 Inducible gene expression

Inducible promoters are fundamental tools in metabolic engineering as they bestow control over gene expression. Few inducible promoters have been described for *M. extorquens* AM1 in comparison to other model organisms, most of them utilizing the cumate inducible (CymR-*cmtO*) system from *Pseudomonas putida* F1[42] (see **Table 1**). Most CymR-*cmtO* resulting promoters are fairly leaky in *M. extorquens* and the highest expression is only 33% of the P_{mxaF} promoter[43-46]. It is important to mention that the CymR repressor is toxic to *M. extorquens*, making CymR controlled promoters unsuitable for gene expression. Presumably one of the most widely used inducible promoter is that of the *lac* operon from *E. coli* and derivatives thereof[47]. LacI-*lacO* inducible promoters are ideal for use in *M. extorquens* since no lactose repressor is found in the genome; making the system truly orthogonal. The P_{lac} promoter drives expression of *lacZ*, *lacY*, and *lacA* involved in lactose catabolism. Its repressor LacI binds to *lacO* operator sites within the P_{lac} promoter, thus blocking RNA polymerase and repressing transcription. Gene

expression is turned on by allolactose, which binds to the repressor leading to its dissociation from the operator site[48]. Nowadays, non-hydrolysable compounds such as isopropyl β-D-1-thiogalactopyranoside (IPTG) are used for induction with the same effect[48, 49]. Importantly, the LacI repressor is not toxic in contrast to CymR. In *M. extorquens*, the P_{lac} promoter conferred weak constitutive expression (without the LacI repressor) and was considered unsuitable for further work[12]. However, the authors did not consider altering the -33 or -10 regions to achieve stronger expression. In CHAPTER I of this thesis, we report the design of inducible promoters based on the LacI-*lacO* system for their use in *M. extorquens*[17]. Our promoters were broken down into the upstream, -33, interspacer, -10, spacer, and +1 regions. Specific sequences for each region were then put together resulting in inducible promoters with different properties. These novel promoters expand the genetic toolbox of *M. extorquens* AM1 and offer control over gene expression in a way that was not possible in the past.

1.5 Broad host range plasmids

Plasmids remain to this day extremely useful DNA vehicles for metabolic engineering. They occur naturally in different genera where they often carry genes which help organisms thrive under particular conditions. They are also key players during horizontal gene transfer for the acquisition of new traits among related bacteria. A plasmid's host range refers to the subset of microorganisms where it can replicate and be maintained effectively. Every plasmid has either a narrow or a broad host range (BHR). The latter were initially defined as those able to transfer between the *Enterobacteria* and *Pseudomonas* spp., although current understanding is that BHR plasmids can be maintained among members of different bacterial classes[50-52]. The most widely used BHR plasmids belong to the IncP or IncQ families. RK2-derivatives of the IncP family replicate well in *M. extorquens* AM1, while IncQ derivatives are used as suicide vectors for integration via homologous recombination. In a study aimed at identifying smaller and more accessible plasmids for this organism, only one was maintained without growth impairment[12]. This plasmid contains a spontaneous mutation within *traJ* as well as the *oriV*, *oriT*, *trfA* regions of RK2; refered in this thesis as *oriV-traJ'* origin of replication. All work in *M. extorquens* was carried out using a single shuttle plasmid carrying the *oriV-traJ'* origin (see **Table 1**). This has largely limited complex metabolic engineering in *M. extorquens* where pathways need to be divided into smaller modules, optimized individually, and combined later.

Table 1: Genetic tools available for *M. extorquens* AM1.

Genetic tool	Sub-category	Part Name	Description	Reference
Replicative plasmid	*oriV-traJ'* origin of replication	pDN19X, pCM80	RK2 derivative, spontaneous mutant.	Marx & Lidstrom (2001)[12]
Mini-chromosomes	*repABC* origin of replication	Mex-CM4, Mex-DM4, Nham-3, Mrad-JCM	Shuttle mini-chromosomes with *repABC* cassettes.	See CHAPTER I. Carrillo *et al*. (2019)[17]
Suicide plasmids		pAYC61		Chistoserdova *et al.* (1994)[10]
		pCM184, pCM351, pCM157	Cre-lox marker recycling.	Marx & Lidstrom (2002)[13]
		pCM433	SacB counterselection.	Marx (2008)[15]
Promoters		P_{lac}	Weak expression.	Marx & Lidstrom (2001)[12]
	Native promoters	P_{mxaF}	Strong expression.	Morris & Lidstrom (1992)[53]
		P_{fumC}, P_{coxB}, P_{tuf}	Weak to strong expression.	Schada von Borzyskowski *et al*. (2015)[16]
	Inducible promoters‡	$P_{mxaF\text{-}cmtO}$	Induced by cumate. *cymR* regulator encoded from a BHR plasmid or in the chromosome, respectively.	Choi *et al.* (2006)[44]; Chou & Marx (2012)[45]
		$P_{R/cmtO}$, $P_{R/tetO}$	Induced by cumate or anhydrotetracycline, respectively.	Chubiz *et al.* (2013)[46]
		$P_{META1p2148\text{-}cmtO}$	Induced by cumate.	Kaczmarczyk et al. (2013)
		$P_{A1/O4/O3}$, $P_{L/O4}$, $P_{L/O4/O3}$, $P_{L/O4/A1}$, $P_{A1/O5/O4}$, $P_{A1con/O5/O4}$, $P_{A1/O4}$, $P_{A1/O4s}$, $P_{A1/O4s\_GA}$, $P_{T5s/A1}$	LacI-*lacO* hybrid promoters. Induced by IPTG.	See CHAPTER I. Carrillo *et al*. (2019)[17]
Transposons	Mini Tn5 derivative	pCM639	Used for genetic screens. IsphoA/hah-Tc.	Marx *et al.* (2003)[14]
	Mariner transposon	pAlmar3	Used for genetic screens in PA1 strain.	Metzger *et al.* (2013)[54]
	Mini Tn7 derivative	pBRI170, pUX-BF13	MiniTn7 and helper plasmid. Insertion at *att*Tn7 site for ATCC 55366 strain.	Choi *et al.* (2006)[55]

Although BHR plasmids are almost ubiquitously used in genetic engineering, there are many drawbacks associated with them. First, most BHR plasmids have a medium to high copy number, which is not strictly controlled and varies due to stochastic partitioning of plasmids after cell division. This variability can lead to artifacts in long-term evolution approaches and cell-to-cell heterogeneity amongst producer cells. Secondly, the constant need for selection and associated costs or waste disposal concerns make stable strains where metabolic pathways are integrated more desirable in the long run. This extra step

is laborious and often 'good hits' using BHR plasmids fail to perform the same when in a chromosomal environment. Finally, the phenomenon of plasmid incompatibility is an important disadvantage of BHR plasmids as they rarely coexist in one cell and often segregate stochastically leading to cell-to-cell heterogeneity[56]. Therefore, we propose the use of so-called "mini-chromosomes" in *M. extorquens* by taking advantage of its natural multipartite genome architecture.

1.6 Multipartite genomes

The genome of most bacteria, which encodes their complete genetic information, is usually made up of a single circular chromosome. A multipartite genomic structure is found in approximately 10% of bacterial species, where a main chromosome is identified as well as smaller secondary replicon(s) (also termed second chromosomes, chromids, or megaplasmids)[57-59]. In both cases, DNA replication and partitioning of the nascent copy of the chromosome(s) to the daughter cell is a tightly controlled process[60, 61].

The formation and raison d'être of multipartite genomes has been debated over the years as summarized in various reviews[57, 59, 62]. Two hypotheses have been proposed to explain how essential secondary replicons came to be. According to the schism hypothesis, part of the main chromosome in a cell with a unified genome split off to form a secondary replicon leading to a multipartite genome. The plasmid hypothesis proposes instead that a cell acquired a large plasmid through horizontal gene transfer, which was domesticated into a stable secondary replicon by *e.g.* acquisition of essential core genes over time. In bacteria, most of the evidence supports the plasmid hypothesis as the DNA replication systems in secondary replicons are reminiscent of those found in smaller plasmids[57]. Whatever their origin, multipartite genomes exist in a number of bacteria with varied ecological niches.

As to their function, some have argued that stable secondary replicons allow bacteria with larger total genomes to replicate faster by having a smaller main chromosome and therefore double faster, as exemplified in some rhizobia[63]. Harrison *et al*. (2010) observed that total genome size of multipartite genomes averaged 5.73 ± 1.66 Mb (n=70) compared to 3.38 ± 1.81 Mb (n=729) in unified genomes[57]. Similar values were also found using larger datasets[59]. However, there are many species with multipartite genomes and relatively slow growth rates and *vice versa*. It might also be that multipartite genomes simply allowed fast acquisition of new traits, which were then assimilated over time as stable components, without affecting chromosomal architecture and/or disrupting important cellular processes[59]. The benefit of stable secondary replicons remains an open question.

1.6.1 Secondary replicons in Alphaproteobacteria

Large secondary replicons can be grouped into two types according to their replication mechanism: iteron or *repABC*-based. The second replicon, often called chromosome II, in members of the *Vibrio* species of Gammaproteobacteria is the best studied example of iteron-based replication systems[62, 64]. *repABC*-based replicons are found exclusively within Alphaproteobacteria with multipartite genomes where they are often referred to as 'megaplasmids', since they are generally ≥ 0.3 mega base pairs (Mb) in size[57, 59, 62]. They also generally have a copy number equal or similar to the main chromosome. The term 'chromid' was proposed to describe secondary replicons as a mixture between chromosome and plasmid under the condition that they carry essential genes no longer located in the main chromosome of said organism[57]. The term 'chromid' is not used in this thesis because (i) only a few megaplasmids (or secondary replicons) have been extensively studied, (ii) gene essentiality is hard to predict without experimental evidence, and perhaps more importantly, (iii) the presence or absence of essential genes is not relevant to our study. For the purpose of this thesis, we use the term 'megaplasmid' to describe large secondary replicons containing a *repABC*-based origin of replication with chromosomal-like properties.

The replication and partitioning of megaplasmids in Alphaproteobacteria is controlled by three genes organized together in an operon: *repA*, *repB*, and *repC*[65, 66]. RepC is responsible for replication, while RepA and RepB function during partitioning or segregation of the nascent megaplasmid to daughter cells. RepC proteins, analogous to DnaA, bind to the origin of replication for replication to start[67, 68]. In *repABC* cassettes, the replication origin is located within the *repC* coding sequence therefore RepC always functions in *cis*. Interestingly, RepC proteins have no significant homology to other replication initiator proteins and are only found in Alphaproteobacteria. In *Agrobacterium tumefaciens*, the amino-terminal domain of pTir10 RepC seems to be essential for DNA binding itself, while the carboxy-terminal domain discriminates between specific and non-specific binding[67]. Additionally, the last 39 amino acids of *Rhizobium etli* p42d RepC were linked to plasmid incompatibility[68]. However, it is not known whether these structural characteristics are common themes amongst RepC proteins of different *repABC* cassettes. RepA and RepB belong to the family of ParA and ParB partitioning proteins, respectively. They also function in very much the same way: RepA has ATPase activity while RepB binds to centromere-like sequences called *parS* sites[65, 69, 70]. A consensus *parS* sequence was identified after comparing multiple *repABC* operons: **GTT**NNNN**GC**NNNN**AAC**[65, 66, 71]. However, no consensus *parS* site(s) could be identified in some *repABC* operons, suggesting that these sites can be more divergent. Finally, although *parS* sites are essential for plasmid stability given that RepB binds to them for segregation, their number and location within *repABC* cassettes seems to vary widely[65, 66].

A stable unit copy number is achieved by the combined action of replication and partitioning systems. RepC levels are controlled at the transcriptional and post-transcriptional level. At the transcriptional level, RepA and RepB have been shown to perform autorepression on the entire operon by obstructing the RNA polymerase[72, 73]. A short, nontranslated counter-transcribed RNA (ctRNA) binds to the *repABC* mRNA post-transcriptionally. This ctRNA is typically encoded in the *repB*/*repC* intergenic region and binding is predicted to alter the secondary structure, thereby causing premature transcriptional termination, ribosomal binding site (RBS) sequestration and/or mRNA degradation[62, 65, 66]. As mentioned above, RepA, RepB, and *parS* sites are also extremely relevant to maintain a unit copy number. These components interact to ensure that one copy of the newly replicated megaplasmid is partitioned to each daughter cell. The reliable inheritance of *repABC* megaplasmids, including dispensable ones, shows the efficiency of this system and makes them very promising tools for use in synthetic biology.

Another feature of *repABC* megaplasmids is that they form incompatibility groups. Many Alphaproteobacteria, especially those belonging to the Rhizobiales, have more than one megaplasmid (or secondary replicon) each with a dedicated *repABC* operon for replication and partitioning[65]. Strikingly, *Rhizobium etli* CFN42 and *R. leguminosarum* 3841 have six *repABC*-type replicons each[74, 75]. This means that *repABC* replicons belonging to different compatibility groups can act independently without cross-talk. At the same time, two cassettes belonging to the same incompatibility group will not be tolerated. One way of avoiding cross-talk seems to be that RepA proteins show overall little homology to one another (less than 61%). The same is true for RepB, which are less than 51% identical. Divergent *parS* sites would also allow multiple RepB proteins in one bacterium since each will bind specifically to its cognate partner. In fact, *parS* sites provided in *trans* caused plasmid incompatibility with the parental *repABC* cassette[71, 76, 77]. This incompatibility is believed to be due to competition for the partitioning machinery between the two plasmids[66]. In addition, point mutations within the *parS* site reduced binding to RepB and eliminated plasmid incompatibility[71]. The RepC proteins can also play a role in plasmid incompatibility; however, their contribution is unclear since they only function in *cis*. In general, having multiple incompatibility groups means that while acquiring new *repABC* cassettes throughout evolution might have been challenging at first, once established they behave as independent chromosomal-like replicons. This opens the door for using them in metabolic engineering and synthetic biology.

1.6.2 *repABC* cassettes as artificial mini-chromosomes in synthetic biology

We believe *repABC*-based mini-chromosomes are very promising tools for Alphaproteobacteria. Döhlemann *et al*. (2017) were the first to introduce a family of heterologous *repABC*-based replicons into *Sinorhizobium meliloti* and related bacteria including *M. extorquens*[78]. These minimal cassettes are composed of the *repABC* operon and surrounding regions to include the necessary *parS* sites. When introduced into otherwise suicide vectors, they granted the replicons with the good properties of their

parental megaplasmids *i.e.*, unit copy number and stable inheritance in the absence of selection. As such, we have termed them mini-chromosomes. The drawbacks of BHR plasmids are overcome by using *repABC* cassettes for replication and faithful inheritance. At the same time, the addition of *E. coli* specific origins facilitates cloning and further manipulation as shuttle vectors. Lastly, the different incompatibility groups of *repABC* operons ensures that multiple mini-chromosomes can be maintained in a single cell without cross-talk. CHAPTER I of this thesis expands upon the work Döhlemann *et al*. (2017) by characterizing new *repABC* cassettes suited for *M. extorquens* specifically.

1.7 Aims of this thesis

The aim of this thesis was to expand and re-engineer the metabolic network of *M. extorquens* AM1 to alter its heterotrophic status. This was achieved through the implementation of heterologous and synthetic pathways into its central carbon metabolism.

In CHAPTER I, we describe novel inducible promoters for *M. extorquens* AM1 based on the LacI-*lacO* system. These are tight before induction and exhibit good dynamic range with different final expression levels upon induction. Importantly, the expression level achieved with these promoters exceeds that of the P_{mxaF} promoter making them the strongest promoters described for this organism. We identified *repABC* cassettes and developed them as new extrachromosomal elements, termed 'mini-chromosomes', for *M. extorquens* AM1. These maintain a stable unit copy number and faithful inheritance even in the absence of selection. Multiple *repABC* replicons can be combined into one cell where they will act as fully independent mini-chromosomes. An unstable replicon, Mex-CM4, showed the behavior needed to establish CRISPR-Cas and similar techniques for which transient expression and fast extinction of the replicon are prerequisites.

In CHAPTER II, we closed the native pentose phosphate pathway with the CBB cycle in *M. extorquens* AM1. This engineered organism should gain cellular carbon via CO_2 fixation while conserving energy through methanol oxidation – since C_1 assimilation was blocked – to allow (chemo)organoautotrophic growth. Introduction of the CBB cycle resulted in a positive growth phenotype linked to an active RubisCO and was further characterized by ^{13}C-metabolomics and whole-cell proteomics.

In CHAPTER III, we rationally engineered *M. extorquens* AM1 to grow on the C_5 sugar ribose. Growth on this novel carbon source was further improved by adaptive laboratory evolution (ALE). The mutations accumulated across different parental lines were identified by whole genome sequencing and further characterized. Ribose metabolism of evolved clones was combined with the CBB cycle for simultaneous CO_2 fixation and acetate was provided as an additional energy source. Adaptive laboratory evolution in chemostat bioreactors was performed under limiting ribose concentrations with the

intent of generating an organoautotrophic strain. Further strain engineering is required for *M. extorquens* to fully synthesize its biomass from CO_2 fixation by the CBB cycle.

In CHAPTER IV, we explored the *in vivo* implementation of a synthetic CO_2 fixation cycle: the **c**rotonyl-CoA/**et**hylmalonyl-**C**oA/**h**ydroxybutyryl-CoA (CETCH) cycle. It was designed as a novel CO_2 fixation cycle and realized *in vitro* using purified enzymes in our laboratory in a previous study[41]. In the current study, we identified and characterized candidates for every step of the CETCH cycle in *M. extorquens* AM1 and various selection strains were created to restore growth on C_1 and C_2 carbon sources with a complete CETCH cycle.

References

1. Schönheit, P., Buckel, W., and Martin, W. F. (2016) On the Origin of Heterotrophy, *Trends Microbiol. 24*, 12-25.
2. Peel, D., and Quayle, J. R. (1961) Microbial growth on C_1 compounds. 1. Isolation and characterization of *Pseudomonas* AM1, *Biochem. J. 81*, 465-469.
3. Bousfield, I. L., and Green, P. N. (1985) Reclassification of Bacteria of the Genus Protomonas Urakami and Komagata 1984 in the Genus Methylobacterium (Patt, Cole, and Hanson) Emend. Green and Bousfield 1983, *Int. J. Syst. Bacteriol. 35*, 209.
4. Green, P. N., and Ardley, J. K. (2018) Review of the genus *Methylobacterium* and closely related organisms: a proposal that some *Methylobacterium* species be reclassified into a new genus, *Methylorubrum* gen. nov., *Int. J. Syst. Evol. Microbiol. 68*, 2727-2748.
5. Green, P. N. (2006) Methylobacterium, In *Prokaryotes*.
6. Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C., Peyraud, R., Robinson, D. G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C. J., Vorholt, J. A., Olson, M. V., Kaul, R., Weissenbach, J., Medigue, C., and Lidstrom, M. E. (2009) Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources, *PloS one 4*, e5584.
7. Erb, T. J., Berg, I. A., Brecht, V., Muller, M., Fuchs, G., and Alber, B. E. (2007) Synthesis of C_5-dicarboxylic acids from C_2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway, *Proceedings of the National Academy of Sciences of the United States of America 104*, 10631-10636.
8. Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., and Vorholt, J. A. (2009) Demonstration of the ethylmalonyl-CoA pathway by using 13C metabolomics, *Proc. Natl. Acad. Sci. USA 106*, 4846-4851.
9. Smejkalova, H., Erb, T. J., and Fuchs, G. (2010) Methanol assimilation in *Methylobacterium extorquens* AM1: demonstration of all enzymes and their regulation, *PloS one 5*.
10. Chistoserdov, A. Y., Chistoserdova, L. V., McIntire, W. S., and Lidstrom, M. E. (1994) Genetic organization of the *mau* gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of *mau* mutants, *J. Bacteriol. 176*, 4052-4065.
11. Toyama, H., Anthony, C., and Lidstrom, M. E. (1998) Construction of insertion and deletion *mxa* mutants of *Methylobacterium extorquens* AM1 by electroporation, *FEMS Microbiol. Lett. 166*, 1-7.
12. Marx, C. J., and Lidstrom, M. E. (2001) Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria, *Microbiology 147*, 2065-2075.
13. Marx, C. J., and Lidstrom, M. E. (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria, *BioTechniques 33*, 1062-1067.
14. Marx, C. J., O'Brien, B. N., Breezee, J., and Lidstrom, M. E. (2003) Novel methylotrophy genes of *Methylobacterium extorquen*s AM1 identified by using transposon

mutagenesis including a putative dihydromethanopterin reductase, *J. Bacteriol. 185*, 669-673.
15. Marx, C. J. (2008) Development of a broad-host-range sacB-based vector for unmarked allelic exchange, *BMC Res Notes 1*, 1.
16. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., and Erb, T. J. (2015) A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria, *ACS Synth. Biol. 4*, 430-443.
17. Carrillo, M., Wagner, M., Petit, F., Dransfeld, A., Becker, A., and Erb, T. J. (2019) Design and Control of Extrachromosomal Elements in *Methylorubrum extorquens* AM1, *ACS Synth. Biol. 8*, 2451-2456.
18. Sonntag, F., Kroner, C., Lubuta, P., Peyraud, R., Horst, A., Buchhaupt, M., and Schrader, J. (2015) Engineering *Methylobacterium extorquens* for de novo synthesis of the sesquiterpenoid alpha-humulene from methanol, *Metab. Eng. 32*, 82-94.
19. Yang, Y. M., Chen, W. J., Yang, J., Zhou, Y. M., Hu, B., Zhang, M., Zhu, L. P., Wang, G. Y., and Yang, S. (2017) Production of 3-hydroxypropionic acid in engineered *Methylobacterium extorquens* AM1 and its reassimilation through a reductive route, *Microbial cell factories 16*, 179.
20. Hu, B., and Lidstrom, M. E. (2014) Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production, *Biotechnology for biofuels 7*, 156.
21. Ochsner, A. M., Sonntag, F., Buchhaupt, M., Schrader, J., and Vorholt, J. A. (2015) *Methylobacterium extorquens*: methylotrophy and biotechnological applications, *Appl. Microbiol. Biotechnol. 99*, 517-534.
22. Berg, I. A. (2011) Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways, *Appl. Environ. Microbiol. 77*, 1925-1936.
23. Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., and Milo, R. (2017) Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points, *eLife 6*.
24. Ljungdahl, L. G., and Wood, H. G. (1969) Total synthesis of acetate from CO_2 by heterotrophic bacteria, *Annu. Rev. Microbiol. 23*, 515-538.
25. Ljungdahl, L. G. (1986) The autotrophic pathway of acetate synthesis in acetogenic bacteria, *Annu. Rev. Microbiol. 40*, 415-450.
26. Wood, H. G. (1991) Life with CO or CO_2 and H_2 as a source of carbon and energy, *FASEB J. 5*, 156-163.
27. Calvin, M., and Benson, A. A. (1948) The Path of Carbon in Photosynthesis, *Science 107*, 476-480.
28. Bassham, J. A., Benson, A. A., and Calvin, M. (1950) The Path of Carbon in Photosynthesis. VIII. The Role of Malic Acid, *J. Biol. Chem. 185*, 781-787.
29. Bassham, J. A., Benson, A. A., Kay, L. D., Harris, A. Z., Wilson, A. T., and Calvin, M. (1954) The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor, *J. Am. Chem. Soc. 76*, 1760-1770.
30. Cleland, W. W., Andrews, T. J., Gutteridge, S., Hartman, F. C., and Lorimer, G. H. (1998) Mechanism of Rubisco: The Carbamate as General Base, *Chem. Rev. 98*, 549-562.
31. Bowes, G., Ogren, W. L., and Hageman, R. H. (1971) Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase, *Biochem. Biophys. Res. Commun. 45*, 716-722.

32. Eisenhut, M., Kahlon, S., Hasse, D., Ewald, R., Lieman-Hurwitz, J., Ogawa, T., Ruth, W., Bauwe, H., Kaplan, A., and Hagemann, M. (2006) The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria, *Plant Physiol. 142*, 333-342.
33. Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M. (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants, *Proceedings of the National Academy of Sciences of the United States of America 105*, 17199-17204.
34. Antonovsky, N., Gleizer, S. s., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D. G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., and Milo, R. (2016) Sugar Synthesis from CO_2 in *Escherichia coli*, *Cell 166*, 115-125.
35. Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., and Milo, R. (2019) Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO_2, *Cell 179*, 1255-1263 e1212.
36. Schada von Borzyskowski, L., Carrillo, M., Leupold, S., Glatter, T., Kiefer, P., Weishaupt, R., Heinemann, M., and Erb, T. J. (2018) An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1, *Metab. Eng. 47*, 423-433.
37. Gassler, T., Sauer, M., Gasser, B., Egermeier, M., Troyer, C., Causon, T., Hann, S., Mattanovich, D., and Steiger, M. G. (2019) The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO_2, *Nat. Biotechnol.*
38. Herz, E., Antonovsky, N., Bar-On, Y., Davidi, D., Gleizer, S., Prywes, N., Noda-Garcia, L., Lyn Frisch, K., Zohar, Y., Wernick, D. G., Savidor, A., Barenholz, U., and Milo, R. (2017) The genetic basis for the adaptation of *E. coli* to sugar synthesis from CO_2, *Nature communications 8*, 1705.
39. Russmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., Drexler, H., Steiger, M., Troyer, C., Al Chalabi, A., Krebiehl, G., Sonntag, D., Zellnig, G., Daum, G., Graf, A. B., Altmann, F., Koellensperger, G., Hann, S., Sauer, M., Mattanovich, D., and Gasser, B. (2015) Systems-level organization of yeast methylotrophic lifestyle, *BMC Biol. 13*, 80.
40. Bar-Even, A., Noor, E., Lewis, N. E., and Milo, R. (2010) Design and analysis of synthetic carbon fixation pathways, *Proceedings of the National Academy of Sciences of the United States of America 107*, 8889-8894.
41. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S., and Erb, T. J. (2016) A synthetic pathway for the fixation of carbon dioxide in vitro, *Science 354*, 900-904.
42. Eaton, R. W. (1997) *p*-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate, *J. Bacteriol. 179*, 3171-3180.
43. Kaczmarczyk, A., Vorholt, J. A., and Francez-Charlot, A. (2013) Cumate-inducible gene expression system for sphingomonads and other *Alphaproteobacteria*, *Appl. Environ. Microbiol. 79*, 6795-6802.
44. Choi, Y. J., Morel, L., Bourque, D., Mullick, A., Massie, B., and Miguez, C. B. (2006) Bestowing inducibility on the cloned methanol dehydrogenase promoter (P_{mxaF}) of

Methylobacterium extorquens by applying regulatory elements of *Pseudomonas putida* F1, *Appl. Environ. Microbiol. 72*, 7723-7729.
45. Chou, H. H., and Marx, C. J. (2012) Optimization of gene expression through divergent mutational paths, *Cell reports 1*, 133-140.
46. Chubiz, L. M., Purswani, J., Carroll, S. M., and Marx, C. J. (2013) A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*, *BMC Res Notes 6*, 183.
47. Müller-Hill, B. (1996) *The lac Operon : a short history of a genetic paradigm*, Walter de Gruyter, Berlin ; New York.
48. Mueller-Hill, B., Rickenberg, H. V., and Wallenfels, K. (1964) Specificity of the Induction of the Enzymes of the Lac Operon in *Escherichia Coli*, *J. Mol. Biol. 10*, 303-318.
49. Monod, J. (1956) Remarks on the mechanism of enzyme induction, In *Enzymes: units of biological structure and function* (Gaebler, O. H., Ed.), pp 7-28, Academic Press, Inc., New York.
50. Datta, N., and Hedges, R. W. (1972) Host ranges of R factors, *J. Gen. Microbiol. 70*, 453-460.
51. Top, E. M., Van Daele, P., De Saeyer, N., and Forney, L. J. (1998) Enhancement of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation in soil by dissemination of catabolic plasmids, *Antonie Van Leeuwenhoek 73*, 87-94.
52. Jain, A., and Srivastava, P. (2013) Broad host range plasmids, *FEMS Microbiol. Lett. 348*, 87-96.
53. Morris, C. J., and Lidstrom, M. E. (1992) Cloning of a methanol-inducible moxF promoter and its analysis in moxB mutants of *Methylobacterium extorquens* AM1rif, *J. Bacteriol. 174*, 4444-4449.
54. Metzger, L. C., Francez-Charlot, A., and Vorholt, J. A. (2013) Single-domain response regulator involved in the general stress response of *Methylobacterium extorquens*, *Microbiology 159*, 1067-1076.
55. Choi, Y. J., Bourque, D., Morel, L., Groleau, D., and Miguez, C. B. (2006) Multicopy integration and expression of heterologous genes in *Methylobacterium extorquens* ATCC 55366, *Appl. Environ. Microbiol. 72*, 753-759.
56. Novick, R. P. (1987) Plasmid Incompatibility, *Microbiological reviews 51*, 381-395.
57. Harrison, P. W., Lower, R. P., Kim, N. K., and Young, J. P. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid, *Trends Microbiol. 18*, 141-148.
58. Touchon, M., and Rocha, E. P. (2016) Coevolution of the Organization and Structure of Prokaryotic Genomes, *Cold Spring Harbor perspectives in biology 8*, a018168.
59. diCenzo, G. C., and Finan, T. M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution, *Microbiol. Mol. Biol. Rev. 81*.
60. Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010) Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*, *Nat. Rev. Microbiol. 8*, 163-170.
61. Ebersbach, G., and Gerdes, K. (2005) Plasmid segregation mechanisms, *Annu. Rev. Genet. 39*, 453-479.
62. Fournes, F., Val, M. E., Skovgaard, O., and Mazel, D. (2018) Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes, *Frontiers in microbiology 9*, 1833.
63. MacLean, A. M., Finan, T. M., and Sadowsky, M. J. (2007) Genomes of the symbiotic nitrogen-fixing bacteria of legumes, *Plant Physiol. 144*, 615-622.

64. Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005) Vibrios commonly possess two chromosomes, *J. Bacteriol. 187*, 752-757.
65. Cevallos, M. A., Cervantes-Rivera, R., and Gutierrez-Rios, R. M. (2008) The repABC plasmid family, *Plasmid 60*, 19-37.
66. Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012) The ABCs of plasmid replication and segregation, *Nat. Rev. Microbiol. 10*, 755-765.
67. Pinto, U. M., Flores-Mireles, A. L., Costa, E. D., and Winans, S. C. (2011) RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *repC* and functions only in *cis*, *Mol. Microbiol. 81*, 1593-1606.
68. Cervantes-Rivera, R., Pedraza-Lopez, F., Perez-Segura, G., and Cevallos, M. A. (2011) The replication origin of a repABC plasmid, *BMC Microbiol. 11*, 158.
69. Williams, D. R., and Thomas, C. M. (1992) Active partitioning of bacterial plasmids, *J. Gen. Microbiol. 138*, 1-16.
70. Ramirez-Romero, M. A., Soberon, N., Perez-Oseguera, A., Tellez-Sosa, J., and Cevallos, M. A. (2000) Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid, *J. Bacteriol. 182*, 3117-3124.
71. MacLellan, S. R., Zaheer, R., Sartor, A. L., MacLean, A. M., and Finan, T. M. (2006) Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons, *Mol. Microbiol. 59*, 1559-1575.
72. Ramirez-Romero, M. A., Tellez-Sosa, J., Barrios, H., Perez-Oseguera, A., Rosas, V., and Cevallos, M. A. (2001) RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon, *Mol. Microbiol. 42*, 195-204.
73. Pappas, K. M., and Winans, S. C. (2003) The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon, *Mol. Microbiol. 49*, 441-455.
74. Gonzalez, V., Santamaria, R. I., Bustos, P., Hernandez-Gonzalez, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S. C., Ramirez, M. A., Jimenez-Jacinto, V., Collado-Vides, J., and Davila, G. (2006) The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons, *Proceedings of the National Academy of Sciences of the United States of America 103*, 3834-3839.
75. Young, J. P., Crossman, L. C., Johnston, A. W., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., Wexler, M., Curson, A. R., Todd, J. D., Poole, P. S., Mauchline, T. H., East, A. K., Quail, M. A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., Whitehead, S., and Parkhill, J. (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components, *Genome biology 7*, R34.
76. Bartosik, D., Baj, J., and Wlodarczyk, M. (1998) Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1, *Microbiology 144 (Pt 11)*, 3149-3157.
77. Soberon, N., Venkova-Canova, T., Ramirez-Romero, M. A., Tellez-Sosa, J., and Cevallos, M. A. (2004) Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*, *Plasmid 51*, 203-216.
78. Döhlemann, J., Wagner, M., Happel, C., Carrillo, M., Sobetzko, P., Erb, T. J., Thanbichler, M., and Becker, A. (2017) A Family of Single Copy repABC-Type

Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium Sinorhizobium meliloti, *ACS Synth. Biol. 6*, 968-984.

PART TWO: RESULTS

CHAPTER I

Design and control of extrachromosomal elements in *Methylorubrum extorquens* AM1[‡]

[‡] Published in: **ACS Synthetic Biology** 2019, 8: 2451-2456.

Design and control of extrachromosomal elements in *Methylorubrum extorquens* AM1

Martina Carrillo[†], Marcel Wagner[‡], Florian Petit[†], Amelie Dransfeld[†, ‡], Anke Becker[‡], Tobias J. Erb*[,†, ‡]

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

[‡]LOEWE Center for Synthetic Microbiology, 35043 Marburg, Germany

ACS Synthetic Biology 2019, 8: 2451-2456.

Author Contributions

T.J.E., A.B., M.W., and M.C. designed the research. M.C., M.W., F.P., and A.D. performed the experimental work. M.C. and M.W. analyzed or provided data. M.C. and T.J.E. wrote the manuscript.

Abstract

Genetic tools are a prerequisite to engineer cellular factories for synthetic biology and biotechnology. *Methylorubrum extorquens* AM1 is an important platform organism of a future C_1-bioeconomy. However, its application is currently limited by the availability of genetic tools. Here we systematically tested *repABC* regions to maintain extrachromosal DNA in *M. extorquens*. We used three elements to construct mini-chromosomes that are stably inherited at single copy number and can be shuttled between *Escherichia coli* and *M. extorquens*. These mini-chromosomes are compatible among each other and with high-copy number plasmids of *M. extorquens*. We also developed a set of inducible promoters of wide expression range reaching levels exceeding those currently available, notably the P_{mxaF}-promoter. In summary, we provide a set of tools to control the dynamic expression and copy number of genetic elements in *M. extorquens*, which opens new ways to unleash the metabolic and biotechnological potential of this organism for future applications.

INTRODUCTION

Methylorubrum extorquens AM1 (formerly *Methylobacterium extorquens*) has been used as a model organism to study methylotrophy, *i.e.* growth on C_1 carbon sources, such as methanol and formate, since its isolation in 1961[1]. The genome of *M. extorquens* AM1 has been fully sequenced and a genome-scale metabolic network was reconstructed for rational engineering[2, 3]. The unique ability of *M. extorquens* to grow on C_1-units has made it an increasingly relevant organism for biotechnology in a methanol- and formate-based bioeconomy[4]. Production of bulk and value-added chemicals such as mevalonate, α-humulene, 3-hydroxypropionate, and 1-butanol has already been realized in *M. extorquens*[5-8]. For the further development of *M. extorquens* as a platform organism in C_1-biotechnology, a broad set of genetic tools are required. Several basic tools are available: one replicative plasmid, *cre-loxP* and *sacB*-based suicide vectors for allelic replacement, and genetic parts including constitutive and, to a lesser extent, inducible promoters[9-14]. Yet, this basic set of genetic tools is far from being complete and many applications are challenged by the fact that more elaborate genetic tools are missing for this organism.

One particular challenge is the lack of dynamic control of gene expression in *M. extorquens*. The P_{lac} promoter and its derivatives, which are widely used in many bacteria, confer only weak or leaky expression in *M. extorquens*[10]. As an alternative to P_{lac}, cumate-inducible promoters have been created by combining the CuO-CymR system from *Pseudomonas putida* F1 to native P_{mxaF} and $P_{META1p2148}$ promoters[12, 14]. While these hybrid promoters are strong upon induction, overexpression of the CymR repressor itself is toxic to the cells causing severe growth defects.

Another, equally important challenge is the availability and control over extrachromosomal genetic elements. Only one broad host range plasmid has been reported which replicates in *M. extorquens* without causing severe growth defects[9, 10]. This plasmid (referred to as oriV-traJ' henceforth) has several disadvantages: a multiple and varying copy number, limited compatibility with other extrachromosomal genetic elements, and the need for constant selective pressure; which restricts its application potential. Genomic integration would circumvent using a plasmid. However, the transformation efficiency is several orders of magnitude lower for genomic integration events and neutral sites for integration have not been characterized in this organism. It also requires the construction of a suicide vector with the desired DNA cargo flanked by homologous regions (≥0.5 kb) and subsequent verification of integration at the desired locus. All of these steps are work intensive and time-consuming.

Many Alphaproteobacteria have multi-partite genomes composed of a main chromosome and one or more megaplasmids. The replication of megaplasmids is integrated into the bacterial cell cycle and occurs once to ensure faithful transmission of genetic information[15, 16]. A small region within these megaplasmids, termed the *repABC* cassette,

drives vertical transmission of these secondary replicons. *repA* and *repB* encode for partitioning proteins analogous to the well-known ParA and ParB, while RepC is a replication initiator protein. The origin of replication resides within the *repC* coding sequence whose expression is downregulated by a short, non-translated counter-transcribed RNA (ctRNA) typically located in the *repB*/*repC* intergenic region[15-17]. Based on this distinctive region, a new family of *repABC*-type shuttle vectors was successfully designed for *Sinorhizobium meliloti* mimicking the characteristic features of secondary replicons *i.e.*, single copy number and stable propagation[18]. Four of these *repABC* regions were shown to be able to replicate in *M. extorquens* but they affected the growth of the organism[18].

Here, to expand the limited genetic toolbox of *M. extorquens* AM1 we developed a set of inducible, orthogonal promoters of different strengths, which can be dynamically controlled by IPTG. Furthermore, we systematically tested new *repABC* regions to establish a set of extrachromosomal elements ('mini-chromosomes') that are faithfully inherited by daughter cells and compatible with each other. This work provides the tools for the extensive genetic engineering of *M. extorquens* AM1 in the future.

RESULTS AND DISCUSSION

Realization of tight, IPTG-inducible promoters with a dynamic range.

To overcome the problem of weak and leaky expression from IPTG-inducible promoters in *M. extorquens*, we sought to develop different *lacO*-controlled promoters that are tight and show a wide range of expression levels after addition of the inducer.

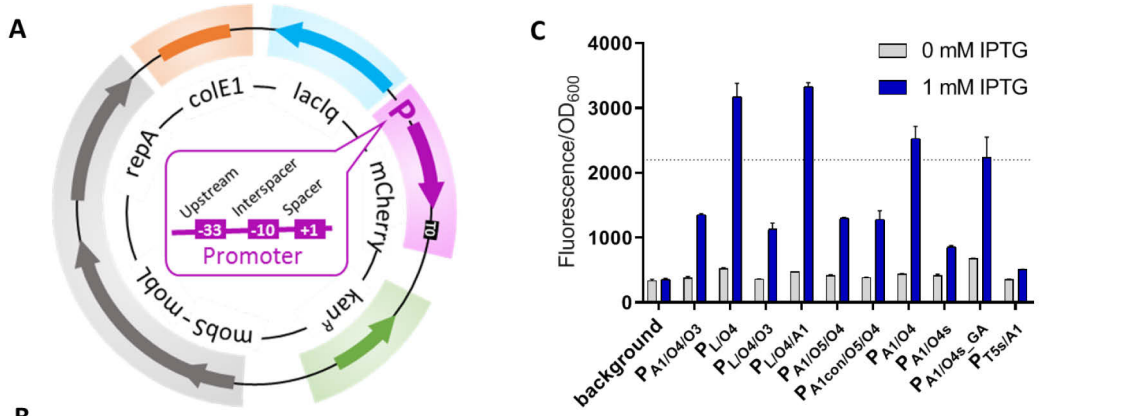
The $P_{A1/O4/O3}$ promoter[19, 20], also known as $P_{A1lacO-1}$, was functional in *M. extorquens* with 27-fold induction of the fluorescence reporter gene mCherry (Fig. 1C). The $P_{A1/O4}$ and $P_{A1/O4s}$ promoters[21] were also functional and showed 21- and 6-fold induction, respectively. A single point mutation within the O4s region of $P_{A1/O4s}$, resulting in $P_{A1/O4s\_GA}$, showed increased mCherry expression at the expense of leakiness without inducer.

The T5 promoter has been used for many years for protein overexpression in *E. coli* [22-24]. The $P_{T5\text{-}lac}$ promoter from pCA24N was weakly functional in *M. extorquens* but there was no difference in the fluorescent signal between uninduced and induced states (data not shown). We modified the upstream region and replaced *lacO4* for a *lacO4s* sequence resulting in promoter $P_{T5s/A1}$ that is tightly repressed but only weakly induced (11-fold). While $P_{T5s/A1}$ is not suitable for gene overexpression, it might be interesting for applications where a protein is toxic if expressed at high levels or constitutively within the cell.

We also created three $P_{L\text{-}lacO}$ hybrid promoters, $P_{L/O4}$, $P_{L/O4/O3}$ and $P_{L/O4/A1}$, which are all strongly repressed in the absence of inducer. Upon induction, $P_{L/O4}$ and $P_{L/O4/A1}$ exhibit 15-

and 23-fold induction of mCherry, respectively. Notably, these two promoters provide mCherry fluorescence levels that are approximately 1.5-fold higher than those of the P_{mxaF} promoter, which drives 9% of soluble protein expression in *M. extorquens*[25] (Fig. 1C, dashed line shows P_{mxaF} levels). The dynamic range of the $P_{L/O4/A1}$ promoter is given in Fig. S1. See Table S4 and S5 for more details on all inducible promoters tested in this study.

In summary, we created and identified a set of IPTG-inducible promoters for *M. extorquens* that range between 6 and 36-fold induction (Table S4). Additionally, our P_{A1}- and P_L-derived promoters range in maximum strength between 9% and 166% of the strong P_{mxaF} promoter, which opens new possibilities for the controlled overexpression of proteins. Thus far, the highest expression level reported from an inducible promoter in *M. extorquens* was 33% of the P_{mxaF} promoter[13]. To facilitate the use of these new promoters by the scientific community, we created empty vectors featuring the respective promoter and a multiple cloning site with an optional N-terminal Strep-II tag compatible with the Methylobrick system[9] and pET vectors (Novagen).



B

Promoter	Upstream region	-33	Interspacer	-10	Spacer	+1
$P_{A1/O4/O3}$	TTTATCAAAAAGAGTG	TTGACT	TGTGAGCGGATAACAAT	GATACT	TAGATTC	AATTGTGAGCGGATAACAATTTCACACA
$P_{L/O4}$	GATAAATTATCTCTGGCGGTG	TTGACA	TGTGAGCGGATAACAAT	GATACT	GAGCAC	ATCAGCAGGACGCACTGAC
$P_{L/O4/O3}$	GATAAATTATCTCTGGCGGTG	TTGACA	TGTGAGCGGATAACAAT	GATACT	TAGATTC	AATTGTGAGCGGATAACAATTTCACACAG
$P_{L/O4/A1}$	GATAAATTATCTCTGGCGGTG	TTGACA	TGTGAGCGGATAACAAT	GATACT	GAGCAC	ATCGAGAGGGACACGGCGA
$P_{A1/O5/O4}$	AATTGTGAGCGGTAACAATTTAG	TTGACT	TGTGAGCGGATAACAAT	GATACT	TAGATTC	ATCGAGAGGGACACGGCGA
$P_{A1con/O5/O4}$	AATTGTGAGCGGTAACAATTTAG	TTGACT	TGTGAGCGGATAACAAT	TATAAT	TAGATTC	ATCGAGAGGGACACGGCGA
$P_{A1/O4}$	TTTATCAAAAAGAGTG	TTGACT	TGTGAGCGGATAACAAT	GATACT	TAGATTC	ATCGAGAGGGACACGGCGA
$P_{A1/O4s}$	TTTATCAAAAAGAGTG	TTGACT	TGTGAGCGCTCACAATT	GATACT	TAGATTC	ATCGAGAGGGACACGGCGA
$P_{A1/O4s\_GA}$	TTTATCAAAAAGAGTG	TTGACT	TGTGAGCACTCACAATT	GATACT	TAGATTC	ATCGAGAGGGACACGGCGA
$P_{T5s/A1}$	TTTATCAAAAAGAGTG	TTGCTT	TGTGAGCGCTCACAATT	TATAAT	TAGATTC	AATTGTGAGCGGATAACAATTTCACACAG

Fig. 1 IPTG-inducible promoters for *M. extorquens*. Map of pIND4-derived plasmid for testing new promoters expressing mCherry (**A**). Sequence of hybrid promoters constructed and tested (**B**), *lacO* sequences are shown in red. The point mutation within the *lacO4s* region is underlined. Fluorescence/OD_{600} before and after addition of 1 mM IPTG (**C**). The dashed line marks fluorescence/OD_{600} signal obtained using the strong P_{mxaF} promoter.

Identification of suitable *repABC* regions for use in *M. extorquens* AM1.

To establish artificial mini-chromosomal elements in *M. extorquens*, we tested eleven *repABC* regions from six different organisms. We introduced these regions into pK18mob2, which is a "suicide" vector unable to replicate in *M. extorquens* (Fig 2A and S2). All *repABC* regions were maintained as independent replicons in *M. extorquens.* The *repABC*-based replicons were recovered for restriction analysis and sequencing of the *repABC* region to confirm their integrity. The only exception was Mnod-2 from *Methylobacterium nodulans*, for which no colonies were obtained after electroporation, likely due to incompatibility of the Mnod-2 *repABC* region with the native genomic system.

Next, we determined the growth behavior of *M. extorquens* carrying each replicon under selective conditions, as well as replicon stability after 96 hours under non-selective conditions (Fig 2B). We observed a general trend that replicon stability and growth behavior of *M. extorquens* were correlated. Mnod-1, also from *M. nodulans*, was very unstable, showed the lowest doubling time (Fig. 2B) and a long lag phase (data not shown), suggesting that *repABC* cassettes from this organism are generally not well compatible with *M. extorquens*. Similarly, Nham-2a and Nham-2b from *Nitrobacter hamburgensis* are very unstable and affected the growth. Replicons originating from *Oligotropha carboxidovorans* were also unstable in the absence of selective pressure, in particular Ocar-1 and Ocar-2a, discouraging the use of these elements for genetic manipulations of *M. extorquens*.

The *repABC* cassette Mex-CM4 originating from *M. extorquens* CM4 showed an interesting behavior. Mex-CM4 allowed fast doubling of *M. extorquens* AM1 under selective conditions, while cells quickly lost the replicon in the absence of antibiotic pressure (Fig. 1B and S4). The *repABC* cassettes of Mex-CM4 and Mex-DM4 (discussed below) are 98% identical to each other, except for the fact that Mex-DM4 has a longer region downstream of *repC* (Fig. 1B and Fig. S2). It might be the case that Mex-CM4 lacks some *parS* sites required for a faithful transmission of the respective replicon to the daughter cells. This behavior makes Mex-CM4 an interesting system to establish CRISPR-Cas in *M. extorquens* where transient expression followed by fast extinction of the genetic element is a desired feature.

The replicons Mex-DM4 (originating from *M. extorquens* DM4), Mrad-JCM (from *M. radiotolerans* JCM2831), and Nham-3 (from *N. hamburgensis* X-14 plasmid 3) showed the highest stability with 97%, 86%, and 77% of cells still harboring the plasmid after 96 hours without antibiotic selection, respectively. The replicons were isolated from kanamycin resistant cells after 96 hours without selective pressure and their integrity was verified by restriction analysis (data not shown). These three replicons allowed fast doubling times of *M. extorquens*, reaching almost wildtype-like growth behavior, and were present at a copy number of 1 (Table 1), indicating that they are good candidates for the construction of mini-chromosomes.

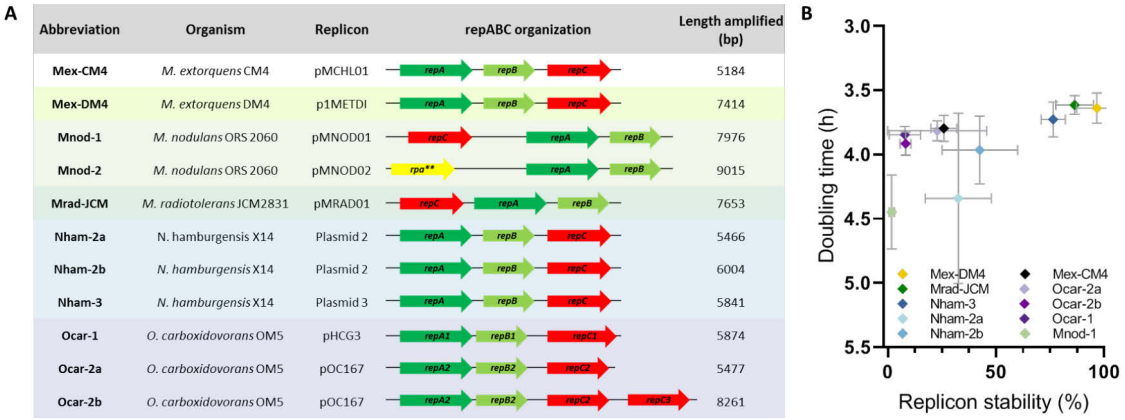


Fig. 2 *repABC* regions tested in *M. extorquens*. (**A**) Size and operon organization. (**B**) Doubling time versus replicon stability (% of kanamycin resistant colonies) of *repABC* regions in the suicide vector pK18mob2. Mean and SD from three biological replicates.

Table 1 Copy number of stable *repABC* cassettes in *M. extorquens*. Mean ± SD from three biological replicates.

Replicon	Copy number (Mean ± SD)
oriV-traJ' (pTE101)	8 ± 1.21
Mex-DM4 (pAD1)	1 ± 0.17
Mrad-JCM (pAD3)	1 ± 0.13
Nham-3 (pAD6)	1 ± 0.09

Assembly and testing of mini-chromosomes in *M. extorquens* AM1.

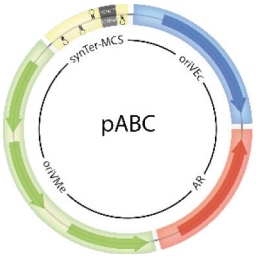
Next, we used the stable *repABC* regions identified in the pK18mob2 system plus Mex-CM4 to construct mini-chromosomes that can be shuttled between *M. extorquens* and *E. coli*. The pABC vector concept and assembly strategy is based on four basic modules each with several standardized parts (see below). Each part is flanked by standardized linker sequences for ligase chain reaction LCR-based assembly. This enables a fast and convenient build-up of custom-made replicons for individual purposes. Additional modules, if necessary, can be implemented with ease. The position and orientation of each module (Scheme Table 2) was designed to minimize crosstalk from adjacent regions. The MCS and the promoter of the *repABC* operon are insulated by flanking transcriptional terminators. The antibiotic resistance cassette is located downstream and in the same orientation as the *repABC* transcriptional unit to avoid reverse transcriptional read-through.

We decided to assemble several mini-chromosomes from the four basic modules, i.e.: (1) a *repABC*-based origin for *M. extorquens* AM1 (Mex-DM4, Nham-3, Mrad-JCM, Mex-CM4), (2) an *E. coli* origin of replication (pMB1, p15A, or pSC101*), (3) a minimal antibiotic

resistance cassette (Kanamycin, Gentamicin, Tetracycline), and (4) a multiple cloning site (MCS1, MCS2, MCS3) flanked by synthetic terminators (Table 2). When we assembled and tested the final mini-chromosomes, our constructs showed similar or higher growth rates of transformed *M. extorquens* compared to replicative plasmids. Replicon stability of the mini-chromosomes was confirmed via flow cytometry, showing similar results as before (Fig. S3 and S4). We also assembled pABC mini-chromosomes with a *mob* site for conjugation and the Cre recombinase under control of constitutive (P_{coxB}) and inducible ($P_{L/O4/A1}$) promoters as additional fifth modules (see Table S2).

The modular construction of the mini-chromosomes allowed us to test inter-compatibility of replicons while avoiding extended regions of high sequence identity which could lead to homologous recombination events. Nham-3 was compatible with Mex-DM4 and Mrad-JCM, while the two latter were not compatible with each other (Tables 2 and 3). In addition, we also tested the compatibility of Mex-DM4, Mrad-JCM, and Nham-3 with the pIND4 and pTE101 vectors, respectively. All three mini-chromosomes were fully compatible with those plasmids, further expanding the box of available and compatible extrachromosomal genetic elements in *M. extorquens* AM1. Replicon integrity was verified by restriction analysis from single and double compatibility test of pABCs and/or replicative plasmids. There were no instances of recombination, proving all compatible replicons functioned autonomously (data not shown). In all cases tested, double transformations were possible, adding to the convenience of these tools for their use in genetic engineering of *M. extorquens* AM1.

Table 2 Modular construction of *repABC*-based mini-chromosomes.

Name	oriVMe	oriVEc	Antibiotic Resistance	SynTer-MCS	Build up Scheme
pABC-DM4K	Mex-DM4	pMB1	Kanamycin	1	
pABC-JCMK	Mrad-JCM	pMB1	Kanamycin	1	
pABC-X14K	Nham-3	pMB1	Kanamycin	1	
pABC-CM4K	Mex-CM4	pMB1	Kanamycin	1	
pABC-DM4G	Mex-DM4	p15A	Gentamicin	2	
pABC-X14G	Nham-3	p15A	Gentamicin	1	
pABC-DM4T	Mex-DM4	p15A	Tetracycline	2	
pABC-X14T	Nham-3	pSC101*	Tetracycline	3	

Abbreviations: oriVMe – *repABC*-based origin of replication for *M. extorquens* AM1, oriVEc – origin of replication for *E. coli*, SynTer-MCS – multiple cloning site flanked by synthetic terminators.

Table 3 Doubling times (h) of *M. extorquens* with *repABC*-based mini-chromosomes in single and double compatibility tests. Mean ± SD from three biological replicates[a].

Origin of replication	Mex-CM4	Mex-DM4	Mrad-JCM	Nham-3	oriV-traJ'	pMG160
CM2720 (strain)	4.11 ± 0.02	3.90 ± 0.01	3.85 ± 0.03	3.94 ± 0.02	4.09 ± 0.05	4.31 ± 0.17
Mex-DM4			X	4.41 ± 0.04	4.15 ± 0.03	4.46 ± 0.06
Mrad-JCM				5.44 ± 0.31	4.68 ± 0.21	5.48 ± 0.27
Nham-3					4.43 ± 0.04	4.49 ± 0.03
oriV-traJ'						5.54 ± 0.40

[a]X = not compatible. pTE100/pTE101 = oriV-traJ', pIND4/pTE1841 = pMG160. Areas were shaded to avoid redundancy.

CONCLUSIONS

Here, we provide a set of novel promoters for *M. extorquens* AM1 that are tight and show a dynamic range of different expression levels upon induction; even exceeding the promoter elements available thus far. We also identified a set of *repABC* regions that allowed us to construct different mini-chromosomal elements that can be stably maintained at single copy number and in different combinations with each other. In addition, we also provide *repABC* regions that allow transient expression in the absence of antibiotic selection, which is a prerequisite to establish CRISPR-Cas-based methods in this organism in the future. Altogether, these tools expand genetic tools available for the engineering of *M. extorquens*, an important platform organism for a sustainable C_1-biotechnology. The genetic tools developed in this study are freely available to the community and hopefully leverage the metabolic and biotechnological potential of *M. extorquens* AM1.

MATERIALS AND METHODS

Strains and cultivation conditions

M. extorquens AM1 strains (Table S1) were grown at 30°C in minimal medium with 123 mM methanol[26]. Antibiotics were used accordingly: kanamycin 35 µg/mL for *M. extorquens* or 50 µg/mL for *E. coli*, gentamicin 7.5 µg/mL for *M. extorquens* or 8-15 µg/mL for *E. coli*, tetracycline 4-10 µg/mL for both, chloramphenicol 34 µg/mL for *E. coli*, ampicillin 100 µg/mL for *E. coli*. *E. coli* TOP10 Δ*dapA* strains were supplemented with 0.3 µM diaminopimelic acid (DAP). *E. coli* was grown at 37°C in LB medium. *E. coli* TOP10 and DH5α (Thermo Scientific™) were used for construction and amplification of all plasmids in this study. Solid medium had 1.5% (w/v) select agar.

Strain construction

Plasmids (Table S2) were transferred into *M*. *extorquens* by electroporation or tri-parental mating[27]. The *dapA* gene of *E. coli* TOP10 was disrupted by an FRT-flanked chloramphenicol resistance cassette, amplified from pKD3 with MC261+MC262 primers, using the Quick and Easy *E. coli* Gene Deletion kit (Gene Bridges K006-GVO-GB) according to manufacturer's instructions. In brief, 20µL of recipient *M. extorquens* from a well-grown culture were spotted on non-selective medium and incubated at 30°C for 24 h. Donor and helper *E. coli* TOP10 Δ*dapA* strains were spotted on top of the recipient cells and incubated at 30°C for 24 h. pRK2013 was used as helper plasmid. Spotted cells were recovered and plated at appropriate dilutions on selective medium. Donor and helper *E. coli* TOP10 Δ*dapA* cannot grow without DAP. Single colonies were screened for chromosomal integration of pTE1899 by colony PCR.

DNA manipulation and plasmid construction

Refer to Table S2 for detailed information on the construction of each plasmid (Supporting Information). Standard molecular techniques were used for amplification, purification, cloning and transformation of DNA[28]. Point mutations were generated by QuikChange® Site-Directed mutagenesis (Stratagene, La Jolla, USA). Ligase chain reaction (LCR) and primer hybridization reactions were performed as published before[18]. T4 Polynucleotide Kinase, FastAP, and FastDigest restriction enzymes were obtained from Thermo Scientific, Q5 DNA polymerase was obtained from NEB, and used according to manufacturer's instructions. DNA oligos were obtained from Eurofins Genomics (Table S3). Plasmid isolation and PCR product purification was performed with NucleoSpin® Plasmid and NucleoSpin® Gel and PCR Clean-up kits (Macherey Nagel), E.Z.N.A. Plasmid Mini Kit (Omega Bio-Tek), or illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences) according to manufacturer's instructions.

Growth in 96-well plates

Cultures were inoculated from a single colony or glycerol stocks into Erlenmeyer flasks with 20 mL of the appropriate medium. Cells from late exponential phase were diluted in fresh medium at an OD_{600}≈0.05 and 180 µl of cell suspension was aliquoted into Nunclon™ Delta Surface (Thermo Scientific #167008) 96-well plates. The temperature was kept constant at 30°C and OD_{600} was recorded every 30 min using a Tecan Infinite M200Pro (Tecan, Männedorf, Switzerland). Data was analyzed using the GraphPad Prism 7 software.

qPCR-based copy number determination

M. extorquens strains were grown in selective medium until OD_{600} 1.0, adjusted to 1.38 × 10^6 cells (≙ 10 ng DNA/µL) and boiled at 95°C for 15 min to serve as templates. qPCR was carried out in a qTOWER Thermal Cycler (Analytik Jena, Germany) using the Takyon No ROX SYBR 2X MasterMix blue dTTP. Reactions were performed according to the manufacturer's instructions in a 5µL volume. *M. extorquens katA*::pTE1179 was used as a reference strain. Replicon copy number was calculated according to Lee *et al*., 2006[29] using primer sets JD253+254 and MC147+148.

Replicon stability

M. extorquens carrying pK18mob2-*repABC* derivatives were grown until OD_{600}=1.0-1.5 in selective medium. Cell suspensions were diluted every 24 hours to an OD_{600}≈0.02 in non-selective medium. Dilution series were plated on agar with and without kanamycin. Antibiotic resistance was correlated with the presence of the assessed plasmid. The inheritance stability of selected *repABC* replicons expressing mCherry as mini-chromosomes was assessed via flow cytometry on a BD LSRFortessaTM SORP flow-cytometer (BD Biosciences, NJ, USA). Fluorescence was detected using a 561 nm laser at 100 mW and a 610/20 bandpass filter[30, 31]. Forward and side scatter values were monitored using a 488 nm laser at 100 mW. The acquired data was analyzed using FACSDiva™ software v8.0 (BD Biosciences). The percentage of fluorescent and non-fluorescent cells was determined from 30,000 gated events each.

ABBREVIATIONS

SD – standard deviation, h – hours, IPTG – Isopropyl β-D-1-thiogalactopyranoside.

ACKNOWLEDGMENTS

We thank S. González Sierra for supporting the FACS analysis and M. Seletskiy for technical assistance. We thank J. Armitage for providing us with pIND4 and S. Vuilleumier for providing us with *M. extorquens* CM4 and *M. nodulans* ORS 2060 strains. This work was created in the framework of the CHIRAMET program (German Ministry of Education and

Research) and received support from the SYNMIKRO (LOEWE program, State of Hesse, Germany), TRR 174 (German Research Foundation), the European Research Council (Grant No. 637675 “SYBORG”), the European Union Coordination and Support Action ‘BioRoboost’ (H2020 CSA Grant No. 820699), as well as the Max Planck Society.

REFERENCES

1. Peel, D., and Quayle, J. R. (1961) Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM1, *Biochem. J. 81*, 465-469.
2. Peyraud, R., Schneider, K., Kiefer, P., Massou, S., Vorholt, J. A., and Portais, J. C. (2011) Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1, *BMC systems biology 5*, 189.
3. Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C., Peyraud, R., Robinson, D. G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C. J., Vorholt, J. A., Olson, M. V., Kaul, R., Weissenbach, J., Medigue, C., and Lidstrom, M. E. (2009) *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources, *PLoS One 4*, e5584.
4. Ochsner, A. M., Sonntag, F., Buchhaupt, M., Schrader, J., and Vorholt, J. A. (2015) *Methylobacterium extorquens*: methylotrophy and biotechnological applications, *Appl Microbiol Biotechnol 99*, 517-534.
5. Liang, W. F., Cui, L. Y., Cui, J. Y., Yu, K. W., Yang, S., Wang, T. M., Guan, C. G., Zhang, C., and Xing, X. H. (2017) Biosensor-assisted transcriptional regulator engineering for *Methylobacterium extorquens* AM1 to improve mevalonate synthesis by increasing the acetyl-CoA supply, *Metab. Eng. 39*, 159-168.
6. Sonntag, F., Kroner, C., Lubuta, P., Peyraud, R., Horst, A., Buchhaupt, M., and Schrader, J. (2015) Engineering *Methylobacterium extorquens* for de novo synthesis of the sesquiterpenoid alpha-humulene from methanol, *Metab. Eng. 32*, 82-94.
7. Hu, B., and Lidstrom, M. E. (2014) Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production, *Biotechnol Biofuels 7*, 156.
8. Yang, Y. M., Chen, W. J., Yang, J., Zhou, Y. M., Hu, B., Zhang, M., Zhu, L. P., Wang, G. Y., and Yang, S. (2017) Production of 3-hydroxypropionic acid in engineered *Methylobacterium extorquens* AM1 and its reassimilation through a reductive route, *Microbial cell factories 16*, 179.
9. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., and Erb, T. J. (2015) A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria, *ACS Synth. Biol. 4*, 430-443.
10. Marx, C. J., and Lidstrom, M. E. (2001) Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria, *Microbiology-Sgm 147*, 2065-2075.

11. Marx, C. J., and Lidstrom, M. E. (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria, *Biotechniques 33*, 1062-1067.
12. Kaczmarczyk, A., Vorholt, J. A., and Francez-Charlot, A. (2013) Cumate-inducible gene expression system for sphingomonads and other *Alphaproteobacteria*, *Appl Environ Microbiol 79*, 6795-6802.
13. Chubiz, L. M., Purswani, J., Carroll, S. M., and Marx, C. J. (2013) A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*, *BMC research notes 6*, 183.
14. Choi, Y. J., Morel, L., Bourque, D., Mullick, A., Massie, B., and Miguez, C. B. (2006) Bestowing inducibility on the cloned methanol dehydrogenase promoter (PmxaF) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1, *Appl. Environ. Microbiol. 72*, 7723-7729.
15. Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012) The ABCs of plasmid replication and segregation, *Nat. Rev. Microbiol. 10*, 755-765.
16. Fournes, F., Val, M. E., Skovgaard, O., and Mazel, D. (2018) Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes, *Front Microbiol 9*, 1833.
17. Cevallos, M. A., Cervantes-Rivera, R., and Gutierrez-Rios, R. M. (2008) The repABC plasmid family, *Plasmid 60*, 19-37.
18. Döhlemann, J., Wagner, M., Happel, C., Carrillo, M., Sobetzko, P., Erb, T. J., Thanbichler, M., and Becker, A. (2017) A Family of Single Copy repABC-Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium *Sinorhizobium meliloti*, *ACS Synth. Biol. 6*, 968-984.
19. Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements, *Nucleic Acids Res. 25*, 1203-1210.
20. Ind, A. C., Porter, S. L., Brown, M. T., Byles, E. D., de Beyer, J. A., Godfrey, S. A., and Armitage, J. P. (2009) Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*, *Appl. Environ. Microbiol. 75*, 6613-6615.
21. Lanzer, M., and Bujard, H. (1988) Promoters largely determine the efficiency of repressor action, *Proc. Natl. Acad. Sci. USA 85*, 8973-8977.
22. Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M. T., and Dobberstein, B. (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo, *Methods Enzymol 155*, 416-433.
23. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research, *DNA Res. 12*, 291-299.
24. Gentz, R., and Bujard, H. (1985) Promoters recognized by *Escherichia coli* RNA polymerase selected by function: highly efficient promoters from bacteriophage T5, *J Bacteriol 164*, 70-77.
25. Liu, Q., Kirchhoff, J. R., Faehnle, C. R., Viola, R. E., and Hudson, R. A. (2006) A rapid method for the purification of methanol dehydrogenase from *Methylobacterium extorquens*, *Protein Expr Purif 46*, 316-320.

26. Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., and Vorholt, J. A. (2009) Demonstration of the ethylmalonyl-CoA pathway by using 13C metabolomics, *Proc. Natl. Acad. Sci. USA 106*, 4846-4851.
27. Toyama, H., Anthony, C., and Lidstrom, M. E. (1998) Construction of insertion and deletion mxa mutants of *Methylobacterium extorquens* AM1 by electroporation, *FEMS Microbiol Lett 166*, 1-7.
28. Sambook, J., and Russel, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Lee, C., Kim, J., Shin, S. G., and Hwang, S. (2006) Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*, *J. Biotechnol. 123*, 273-280.
30. Piatkevich, K. D., and Verkhusha, V. V. (2011) Guide to red fluorescent proteins and biosensors for flow cytometry, *Methods Cell Biol 102*, 431-461.
31. Telford, W. G., Hawley, T., Subach, F., Verkhusha, V., and Hawley, R. G. (2012) Flow cytometry of fluorescent proteins, *Methods 57*, 318-330.

SUPPLEMENTARY INFORMATION

Contents of the material supplied as Supporting Information:

Table S1 Strains used in this study

Strains	Description	Reference
CM2720	*Methylorubrum extorquens* AM1 strain deficient in cellulose production	Delaney *et al.*, 2013[1]
SmCreΔhsdR	*Sinorhizobium meliloti* Rm1021 *cre* expression strain with *hsdR* deletion, *tauX*:: *cre-tetRA* (Tc^R)	Döhlemann *et al.*, 2016[2]
Escherichia coli DH5α	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-	Thermo Fischer Scientific
E. coli TOP10	F- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG*	Thermo Fischer Scientific
E. coli TOP10 Δ*dapA*	F- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG dapA::FRT-cat-FRT*; DAP deficient	This study

DAP – diaminopimelic acid

Table S2 Plasmids used in this study

Plasmids	Description	Cloning strategy	Reference
pRK2013	Mobilization helper plasmid, Km^R	n. a.	Figurski and Helinski, 1979[4]
pK18mob2	Suicide vector, Km^R	n. a.	Tauch *et al*., 1998[5]
pKD3	FRT-flanked chloramphenicol resistance cassette, Amp^R Cm^R	n. a.	Datsenko *et al*., 2000[6]
pACYC184	Carrying p15A replication origin, Cm^R, Tc^R	n. a.	Chang and Cohen, 1978[7]
pXG-10	Carrying pSC101* replication origin, Cm^R	n. a.	Urban and Vogel, 2007[8]
pIND4	pMG160 origin, $P_{A1/O4/O3}$ promoter, empty vector, Km^R	n. a.	Ind *et al*., 2009[9]
pTE1830	pMG160 origin, $P_{A1/O4/O3}$ promoter, empty vector, Gm^R	PCR products of prMC276+MC278 from pIND4, prMC228+MC277 from pIND4, and prMC274+MC275 from pLAR-Gm (*aacC1*) were assembled via Gibson assembly.	This study
pTE1841	pMG160 origin, $P_{A1/O4/O3}$ promoter, empty vector, Tc^R	NheI, NotI 5.283 bp fragment from pIND4, PCR products of pr276+295 from pIND4, prMC311+MC277 from pIND4, and prMC296+MC297 from pTE100 (*tetA*) were assembled via Gibson assembly.	This study
pTE100	Promoter-less, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al*, 2015[10]
pTE101	Promoter-less, oriV-traJ' origin, Km^R	n. a.	Schada v. B. *et al*, 2015[10]
pTE102	P_{mxaF} promoter, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al*, 2015[10]
pTE102-mChe	P_{mxaF} promoter, mCherry, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al*, 2015[10]
pTE104-mChe	P_{coxB} promoter, mCherry, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al*, 2015[10]
pLoriVSm	Unloaded library vector for *repABC* regions, pK18mob2-derivative, Km^R	n. a.	Döhlemann *et al*., 2017[11]
pLsynTer-1	Library plasmid providing *synTer1*-MCS, Km^R	n. a.	Döhlemann *et al*., 2017[11]

pLsynTer-2	Library plasmid providing *synTer2*-MCS, Km^R	n. a.	Döhlemann *et al*., 2017[11]
pLsynTer-3	Library plasmid providing *synTer3*-MCS, Km^R	n. a.	Döhlemann *et al*., 2017[11]
pLAR-Km	Library plasmid providing a kanamycin resistance cassette, Km^R	n. a.	Döhlemann *et al*., 2017[11]
pLAR-Gm	Library plasmid providing a gentamicin resistance cassette Km^R, Gm^R	n. a.	Döhlemann *et al*., 2017[11]
pLAR-Tc	Library plasmid providing a tetracyclin resistance cassette Km^R, Tc^R	n. a.	Döhlemann *et al*., 2017[11]
pLoriT	Unloaded library vector for oriT parts, pK18mob2-derivative, Km^R	n. a.	Döhlemann *et al*., 2017[11]
pLoriT-1	Library plasmid providing a *mob* site	n. a.	Döhlemann *et al*., 2017[11]
pAD1	pLoriVSm with *repABC*_Mex-DM4, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD7+8 from *Methylorubrum extorquens* DM4 gDNA.	This study
pAD2	pLoriVSm with *repABC*_Ocar-2a, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD13+14 from *Oligotropha carboxidovorans* OM5 gDNA.	This study
pAD3	pLoriVSm with *repABC*_Mrad-JCM, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD9+10 from *Methylobacterium radiotolerans* JCM2831 gDNA.	This study
pAD4	pLoriVSm with *repABC*_Ocar-1, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD11+12 from *Oligotropha carboxidovorans* OM5 gDNA.	This study
pAD5	pLoriVSm with *repABC*_Nham-2a, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD3+4 from *Nitrobacter hamburgensis X14* gDNA.	This study
pAD6	pLoriVSm with *repABC*_Nham-3, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD5+6 from *Nitrobacter hamburgensis X14* gDNA.	This study

pAD7	pLoriVSm with *repABC*_Ocar-2b, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD13+15 from *Oligotropha carboxidovorans* OM5 gDNA.	This study
pAD8	pLoriVSm with *repABC*_Mex-CM4, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD24+25 from *Methylorubrum extorquens* CM4 gDNA.	This study
pAD9	pLoriVSm with *repABC*_Mnod-1, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD26+27 from *Methylobacterium nodulans* ORS 2060 gDNA.	This study
pAD10	pLoriVSm with *repABC*_Mnod-2, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD28+29 from *Methylobacterium nodulans* ORS 2060 gDNA.	This study
pAD11	pLoriVSm with *repABC*_Nham-2b, Km^R	Backbone: ScaI linearized and dephosphorylated pLoriVSm; Insert: T4 PNK phosphorylated PCR product of prAD1+2 from *Nitrobacter hamburgensis X14* gDNA.	This study
pMW216	Km^R pMB1 origin; *repABC*_Mex-DM4; MCS 1	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD1 (*repABC*_Mex-DM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW217	Km^R pMB1 origin; *repABC*_Mrad-JCM; MCS 1	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD3 (*repABC*_Mrad-JCM), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW218	Km^R pMB1 origin; *repABC*_Nham-3; MCS 1	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD6 (*repABC*_Nham-3), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW219	Km^R pMB1 origin; *repABC*_Mex-CM4; MCS 1	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD8 (*repABC*_Mex-CM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study

pMW220	Gm^R p15A origin; *repABC*_Mex-DM4 ; MCS 2	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Gm (P_{min2}-*aacC1*), p15A_for + p15A_rev from pACYC184 (p15A origin), oriVSm_for + oriVSm_rev from pAD1 (*repABC*_Mex-DM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-2 (MCS2) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW221	Gm^R p15A origin; *repABC*_Nham-3; MCS 1	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Gm (P_{min2}-*aacC1*), p15A_for + p15A_rev from pACYC184 (p15A origin), oriVSm_for + oriVSm_rev from pAD6 (*repABC*_Nham-3), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW223	Tc^R p15A origin; *repABC*_Mex-DM4; MCS 2	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Tc (P_{min2}-*tetR*), p15A_for + p15A_rev from pACYC184 (p15A origin), oriVSm_for + oriVSm_rev from pAD1 (*repABC*_Mex-DM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-2 (MCS2) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW224	Tc^R pSC101* origin; *repABC*_Nham-3; MCS 3	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Tc (P_{min2}-*tetR*), pSC101*_for + pSC101*_rev from pXG-10 (pSC101* origin), oriVSm_for + oriVSm_rev from pAD6 (*repABC*_Nham-3), synTer-MCS_for + synTer-MCS_rev from pLsynTer-3 (MCS3) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.	This study
pMW231	Library plasmid providing the gene for the Cre recombinase under the control of $P_{L/O4/A1}$	Backbone: ScaI linearized and dephosphorylated pLoriT; Insert: T4 PNK phosphorylated PCR products of prMW1041+1042 from pTE1887 and prMW1028+1029 from SmCreΔhsdR gDNA. Three- component ligation with T4 DNA ligase.	This study
pMW232	Library plasmid providing the gene for the Cre recombinase under the control of P_{coxB}	Backbone: ScaI linearized and dephosphorylated pLoriT; Insert: T4 PNK phosphorylated PCR products of prMW1043+1044 from pTE104-mChe and prMW1028+1029 from SmCreΔhsdR gDNA. Three- component ligation with T4 DNA ligase.	This study
pMW233	Gm^R; pMB1 origin; *repABC*_Mex-CM4; MCS 1; oriT	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Gm (P_{min2}-*aacC1*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD8 (*repABC*_Mex-CM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1), mob_for + mob_rev from pLoriT-1 (*mob* site) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study

pMW234	Gm^R; pMB1 origin; *repABC*_Nham-3; MCS 2; oriT	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Gm (P_{min2}-*aacC1*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD6 (*repABC*_Nham-3), synTer-MCS_for + synTer-MCS_rev from pLsynTer-2 (MCS2), mob_for + mob_rev from pLoriT-1 (*mob* site) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study
pMW235	Tc^R; pMB1 origin; *repABC*_Mex-DM4; MCS 3; oriT	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Tc (P_{min2}-*tetR*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD1 (*repABC*_Mex-DM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-3 (MCS3), mob_for + mob_rev from pLoriT-1 (*mob* site) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study
pMW236	Gm^R; pMB1 origin; *repABC*_Mrad-JCM; MCS 1; oriT	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Gm (P_{min2}-*aacC1*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD3 (*repABC*_Mrad-JCM), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1), mob_for + mob_rev from pLoriT-1 (*mob* site) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study
pMW237	Km^R; pMB1 origin; *repABC*_Mex-CM4; MCS 1; P_{coxB}-*cre*	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD8 (*repABC*_Mex-CM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1), mob_for + mob_rev from pMW232 (P_{coxB}-*cre*) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study
pMW238	Km^R; pMB1 origin; *repABC*_Mex-CM4; MCS 1; $P_{L/O4/A1}$-*cre*	Phosphorylated PCR products of primers AR_for + AR_rev from pLAR-Km (P_{min2}-*nptII*), pMB1_for + pMB1_rev from pK18mob2 (pMB1 origin), oriVSm_for + oriVSm_rev from pAD8 (*repABC*_Mex-CM4), synTer-MCS_for + synTer-MCS_rev from pLsynTer-1 (MCS1), mob_for + mob_rev from pMW231 ($P_{L/O4/A1}$-*cre*) were *in vitro* assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.	This study
pTE1179	pK18mob2 with homologous region of *katA* (META1p3421), Km^R	Backbone: XbaI, KpnI linearized pK18mob2; Insert: XbaI, KpnI digest of PCR product from *M. extorquens* AM1 gDNA prMC157+158. Ligation with T4 DNA ligase.	This study
pTE1870	$P_{A1/O4/O3}$-mCherry_AAV protein degradation tag, Km^R	Backbone: BsrGI, HindIII digested pTE1853; Insert: primer hybridization product prMC398+399. Ligation with T4 DNA ligase.	This study

pTE1875	P_{mxaF}-mCherry_AAV protein degradation tag, Tc^R	Backbone: SpeI, KpnI lin pTE102; Insert: XbaI, KpnI 786bp fragment from pTE1870. Ligation with T4 DNA ligase.	This study
pTE1899	pTE1179 with P_{mxaF}-mCherry_AAV tag expression cassette, Km^R	Backbone: XbaI linearized pTE1179; Insert: XbaI, SpeI 1.3 kb fragment of pTE1875. Ligation with T4 DNA ligase. Product 1.	This study
pTE1825	pIND4-derivative with $P_{A1/O4/O3}$-mCherry	Backbone: HindIII, XbaI linearized pIND4; Insert: HindIII, XbaI 740 bp fragment of pTE102-mChe. Ligation with T4 DNA ligase.	This study
pTE1852	KpnI site at bp 1292 removed by SNP mutation, Km^R	QuikChange mutagenesis of pTE1825 with prMC348+349.	This study
pTE1853	$P_{A1/O4/O3}$-mCherry; KpnI sites removed, Km^R	QuikChange mutagenesis of pTE1852 with prMC353+354.	This study
pTE1855	$P_{L/O4}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-8. Ligation with T4 DNA ligase.	This study
pTE1856	$P_{L/O4/O3}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-367+369+370. Ligation with T4 DNA ligase.	This study
pTE1877	$P_{L/O4/A1}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-366+422+421. Ligation with T4 DNA ligase.	This study
pTE1878	$P_{A1/O5/O4}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+381+383+378+419+420. Ligation with T4 DNA ligase.	This study
pTE1879	$P_{A1con/O5/O4}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+377+381+383+419+423. Ligation with T4 DNA ligase	This study
pTE1880	$P_{A1/O4}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+376+380+382+419+420. Ligation with T4 DNA ligase.	This study
pTE1881	$P_{A1/O4s}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+379+380+382+419+424. Ligation with T4 DNA ligase.	This study
pTE1882	$P_{A1/O4s\_GA}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+380+419+424+425+426. Ligation with T4 DNA ligase.	This study
pTE1863	$P_{T5s/A1URS/O3}$-mCherry, Km^R	Backbone: XbaI, AatII linearized pTE1853; Insert: primer hybridization product prMC361-363+369+373-375+380. Ligation with T4 DNA ligase.	This study
pTE2714	$P_{A1/O4/O3}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1853; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1885	$P_{L/O4}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1855; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1886	$P_{L/O4/O3}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1856; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study

pTE1887	$P_{L/O4/A1}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1877; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1888	$P_{A1/O5/O4}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1878; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1889	$P_{A1con/O5/O4}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1879; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1890	$P_{A1/O4}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1880; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1891	$P_{A1/O4s}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1881; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1892	$P_{A1/O4s\_GA}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1882; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE1893	$P_{T5s/A1}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	Backbone: XbaI, HindIII cut pTE1863; Insert: primer hybridization product prMC404-408. Ligation with T4 DNA ligase.	This study
pTE2704	P_{mxaF}-mCherry_AAV; *repABC*_Mex-DM4, Km^R	Backbone: ScaI, PacI linearized pMW216; Insert: ScaI, PacI digest of PCR product prMC437+438 of pTE1875. Ligation with T4 DNA ligase.	This study
pTE2705	P_{mxaF}-mCherry_AAV; *repABC*_Mrad-JCM, Km^R	Backbone: ScaI, PacI linearized pMW217; Insert: ScaI, PacI digest of PCR product prMC437+438 of pTE1875. Ligation with T4 DNA ligase.	This study
pTE2706	P_{mxaF}-mCherry_AAV; *repABC*_Nham-3, Km^R	Backbone: ScaI, PacI linearized pMW218; Insert: ScaI, PacI digest of PCR product prMC437+438 of pTE1875. Ligation with T4 DNA ligase.	This study
pTE2707	P_{mxaF}-mCherry_AAV; *repABC*_Mex-CM4, Km^R	Backbone: ScaI, PacI linearized pMW219; Insert: ScaI, PacI digest of PCR product prMC437+438 of pTE1875. Ligation with T4 DNA ligase.	This study

n. a. – not applicable, gDNA – genomic DNA, Amp^R – ampicillin resistance, Cm^R – chloramphenicol resistance, Gm^R – gentamicin resistance, Km^R – kanamycin resistance, Tc^R – tetracycline resistance, pr – primers

Table S3 Primers used in this study.

Primer #	Sequence
MC147	ATTCAACTTCCCGCACGAGT
MC148	GGCACCTTGTTGATCATCGC
MC157	TGATCTAGAGATTCTGACGACCCGTCAGG
MC158	TATGGTACCTCGCGCTGGTTCGAATAGAC
MC228	GAAACGCCTGGTATCTTTATAGTC
MC261	TGCCATACCAAACGTACCATTGAGACACTTGTTTGCACAGAGGATGGCCCTGGGAATTAGCCATGGTC
MC262	AGACCCAGTTCCTTACATGCCCATTTCACCGGGATTGGATTGGGTTCGACTCAGTCTTGAGCGATTGTGTAGG
MC274	ACAGTAATACAAGGGGTGTTATGTTACGCAGCAGCAACG
MC275	CAACCAATTAACCAATTCTGTTAGGTGGCGGTACTTGGG
MC276	AACACCCCTTGTATTACTG
MC277	CAGAATTGGTTAATTGGTTG
MC278	ATAAAGATACCAGGCGTTTC
MC295	TCGACGCGGGCCGAGCTTTG
MC296	CAACCAATTAACCAATTCTGATCAGCGATCGGCTCGTTGC
MC297	ACAGTAATACAAGGGGTGTTTCATGCTTGACACTTTATCAC
MC311	TTATTGGTGAGAATCCAAGC
MC348	CGCGCGAATTGCAGTTACCATTTATCAGGG
MC349	CCCTGATAAATGGTAACTGCAATTCGCGCG
MC352	CGAGCTCGGTACTGACGTAGCCCAGC
MC353	GCTGGGCTACGTCAGTACCGAGCTCG
MC361	CTAAGAAACCATTATTATCATGAC
MC362	ATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC
MC363	CCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGT
MC364	TTCACCTCGAGAAAAGATAAATTATCTCTGGCGGTGTTGACATGTG
MC365	ATCTTTTCTCGAGGTGAAGACGAAAGGGCCTCGTGATACG
MC366	AGTATCATTGTTATCCGCTCACATGTCAACACCGCCAGAGATAATTT
MC367	CTAGAGTCAGTGCGTCCTGCTGATGTGCTC
MC368	AGCGGATAACAATGATACTGAGCACATCAGCAGGACGCACTGACT
MC369	CTAGACTGTGTGAAATTGTTATCCGCTCACAATTGAATCTA
MC370	AGCGGATAACAATGATACTTAGATTCAATTGTGAGCGGATAACAATTTCACACAGT
MC373	AGCGCTCACAATTTATAATTAGATTCAATTGTGAGCGGATAACAATTTCACACAGT
MC375	ATTATAAATTGTGAGCGCTCACAAAGCAACACTCTTTTTG
MC376	AGTATCATTGTTATCCGCTCACAAGTCAACACTCTTTTTG
MC377	ATTATAATTGTTATCCGCTCACAAGTCAACTAAATTGTTACCGCTCA
MC378	AGTATCATTGTTATCCGCTCACAAGTCAACTAAATTGTTACCGCTCA
MC379	AGTATCAATTGTGAGCGCTCACAAGTCAACACTCTTTTTG
MC380	ATAAATTTTCTCGAGGTGAAGACGAAAGGGCCTCGTGATACG
MC381	CAATTTTTTCTCGAGGTGAAGACGAAAGGGCCTCGTGATACG
MC382	TTCACCTCGAGAAAATTTATCAAAAAGAGTGTTGACTTGTGAGCG
MC383	TTCACCTCGAGAAAAAATTGTGAGCGGTAACAATTTAGTTGACTTGTGAGCG
MC398	GTACAAGGCCGCGAACGACGAAAACTACGCCGCGGCCGTCTGAACTAGTCCTGCAGGTACCTTAAGA
MC399	AGCTTCTTAAGGTACCTGCAGGACTAGTTCAGACGGCCGCGGCGTAGTTTTCGTCGTTCGCGGCCTT
MC404	CTAGAATTAAAGAGGAGAAATTAACCATGGCGAGCTGGAGCCATCCGCAGTTCGAA
MC405	AGCTCGCCATGGTTAATTTCTCCTCTTTAATT
MC406	AAGATCGAAGGCCGCCATATGTCGGGCGGATCCTGAACTAGTCCTGCAGGTACCGTA
MC407	TCCGCCCGACATATGGCGGCCTTCGATCTTTTCGAACTGCGGATGGCTCC
MC408	AGCTTACGGTACCTGCAGGACTAGTTCAGGA
MC419	CTAGAGTCGCCGTGTCCCTCTCGATGAATCTA
MC420	GATAACAATGATACTTAGATTCATCGAGAGGGACACGGCGACT
MC421	CTAGAGTCGCCGTGTCCCTCTCGATGTGCTC
MC422	AGCGGATAACAATGATACTGAGCACATCGAGAGGGACACGGCGACT

MC423	GATAACAATTATAATTAGATTCATCGAGAGGGACACGGCGACT
MC424	GATAACAATTATAATTAGATTCATCGAGAGGGACACGGCGACT
MC425	AGTATCAATTGTGAGTGCTCACAAGTCAACACTCTTTTTG
MC426	TTCACCTCGAGAAAATTTATCAAAAAGAGTGTTGACTTGTGAGCA
MC437	CTGACAGTACTCAAGCTAGCTTCCCGCTTGGTC
MC438	CGCTTAATTAACGACGGCCAGTGAATTAGG
AD1	ACCCGATCGCGAAAACTTCC
AD2	CTCGAGGGTTGGGCGAAAG
AD3	ATGGATCCGAAATCGCCCAG
AD4	AACTTCCTTCGTAGCGGTC
AD5	AACGAAGTTCGGCCATGACC
AD6	GGCGAACGCGAGGAGAAATT
AD7	ATGCTGCGGATCGCTAAAAAGA
AD8	TTCGGCTTGGTCTCGATCA
AD9	AGGTCTCGGAAAGACATCCG
AD10	CGGATCCAGGTGTTCAGGAA
AD11	ATCATTTGCGGCTCGCCTAG
AD12	GTCCAGACCATCGGGCTCTA
AD13	GGAAAGCCTGGTGCTGAAAT
AD14	TTTGTCGTCCTTGAAGGGATAGTT
AD15	CGGATCATCGATCGCATTCTG
AD24	ACCATCAAGAACTGCCCGTT
AD25	GTGACGGAATGACGCCGTAA
AD26	ATCTCCATTCACACCGCTGCTT
AD27	TTTGTCCGCGAGGTCTTCGTT
AD28	GTGCTTCGTCATGGCTCCAT
AD29	AATGACCTCTGACCGGCGTTA
MW125	GCGGTCGACGTTATGGAGCAGCAACGATGTT
MW126	GCGCTGCAGAGTACTTTAGGTGGCGGTACTTGGGTCG
MW1028	GCGCATATGTCCAATTTACTGACCGTACACCAAAAT
MW1029	GCTAGCGAATTCTAGTCGCCATC
MW1041	CGCCATATGTAATTTCTCCTCTTTAATTCTAGAGTC
MW1042	TCGCGCTAACTTACATTAATTG
MW1043	CGTATCCCCAGAGGCAGCCAAATTGC
MW1044	GCGCATATGCCCGCTTGGCTCCCCTGGTG
MW1046	ATAACTTCGTATATGGTATTATATACGAACGGTAG
MW1047	TCGACTACCGTTCGTATATAATACCATATACGAAGTTAT
MW1048	GTACCGTTCGTATAGCATACATTATACGAAGTTAT
MW1049	ATAACTTCGTATAATGTATGCTATACGAACGGTACTGCA
JD253	CTCGCAGAGCAGGATTCCCGTT
JD254	GGCAGGATAGGTGAAGTAGGCCCA
AR_for	GACCTTTTCTCCGACGAATAGA
AR_rev	GTCTTATCTGAAAGTTGTGCCTG
oriVSm_for	GAAACTGTCACTGGTCCGT
oriVSm_rev	TTCAGTTACGATAGAGTTCCACG
synTer-MCS_for	CTATTGAAGGAACACTGTATCTCG
synTer-MCS_rev	GTCAACCCGCTTACACTC
pMB1_for	TACTACTGTTTCAGACTGGCGTAATCACTCAGTAGATCAAAGGATCTTCTTGAGATCCTTTTTTTC
pMB1_rev	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTCGCGTTGCTGGCGTTTTTC
p15A_for	TACTACTGTTTCAGACTGGCGTAATCACTCAGTCTACATTTGAAGAGATAAATTGCACTG
p15A_rev	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTTAGCGGAGTGTATACTGGC
pSC101*_for	TACTACTGTTTCAGACTGGCGTAATCACTCAGTCCTAGGGTACGGGTTTTGC
pSC101*_rev	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTGACAGTAAGACGGGTAAGCCT

mob_for	TGTGACGATAAGTTCCCTACTG
mob_rev	CCTCTATTGATAACGGGTGACA
BO1	AACTTCTCGTGGAACTCTATCGTAACTGAAGTCTATTGAAGGAACACTGTATCTCGGTCA
BO2	GTTATCAGAAGAGTGTAAGCGGGTTGACTATACTACTGTTTCAGACTGGCGTAATCACTC
BO3	GGCAATCACTGAGTAAGACTATTCTGTCCAGACCTTTTCTCCGACGAATAGAGTAACAGA
BO3a	GGCAATCACTGAGTAAGACTATTCTGTCCATGTGACGATAAGTTCCCTACTGACAGAATC
BO3b	GGTTACAGTGTCACCCGTTATCAATAGAGGGACCTTTTCTCCGACGAATAGAGTAACAGA
BO4	GATTGGTCAGGCACAACTTTCAGATAAGACGAAACTGTCACTGGTCCGTTGATACAATCC

Table S4 Fold induction of IPTG-inducible promoters in *M. extorquens* AM1. Mean ± SD.

Promoter	Fluorescence/OD_{600}		Fold induction	Maximum strength relative to the P_{mxaF} promoter (%)[†]
	0 mM IPTG	1 mM IPTG		
background	339 ± 14	355 ± 13	n. a.	n. a.
P_{mxaF}	2152 ± 16	n. a.	n. a.	100
$P_{A1/O4/O3}$	376 ± 18	1355 ± 11	27	55.6
$P_{L/O4}$	526 ± 8	3172 ± 147	15	156.8
$P_{L/O4/O3}$	361 ± 4	1132 ± 66	36	43.2
$P_{L/O4/A1}$	471 ± 1	3332 ± 43	23	165.7
$P_{A1/O5/O4}$	416 ±8	1297 ± 14	12	52.4
$P_{A1con/O5/O4}$	384 ± 6	1286 ± 91	21	51.8
$P_{A1/O4}$	441 ± 4	2531 ± 131	21	121.1
$P_{A1/O4s}$	420 ± 14	855 ± 17	6	27.8
$P_{A1/O4s\_GA}$	678 ± 5	2235 ± 224	6	104.6
$P_{T5s/A1}$	353 ± 6	510 ± 4	11	8.6

[†]Values after induction with 1 mM IPTG. n. a. – not applicable.

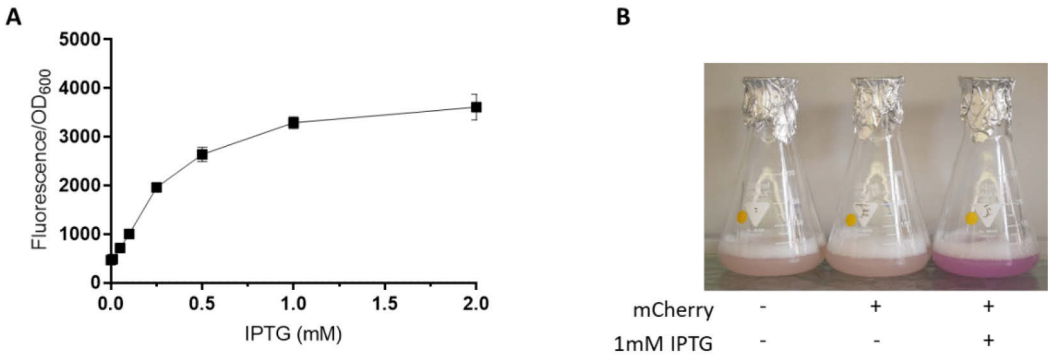


Fig. S1 $P_{L/O4/A1}$ promoter in *M. extorquens* AM1. Dynamic range of expression with 0-2 mM IPTG (**A**). Cultures expressing ± mCherry before and after induction with 1 mM IPTG; culture of an empty vector is shown as a reference (**B**).

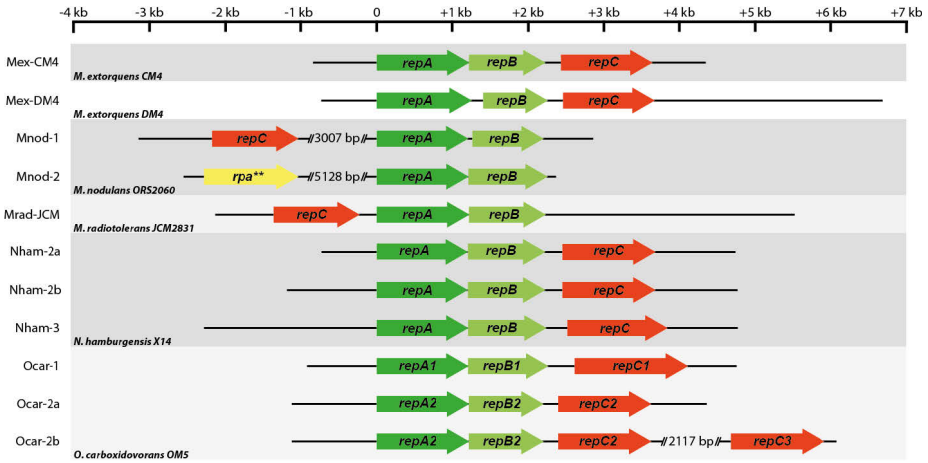


Fig. S2 Overview of the individual *repABC* regions tested in this study. All *repABC* regions are shown with the start codon of *repA* aligned at the zero mark. The black lines depict the regions amplified outside of the *repABC* coding sequences to include necessary *parS* sites. A scale bar is given as a reference.

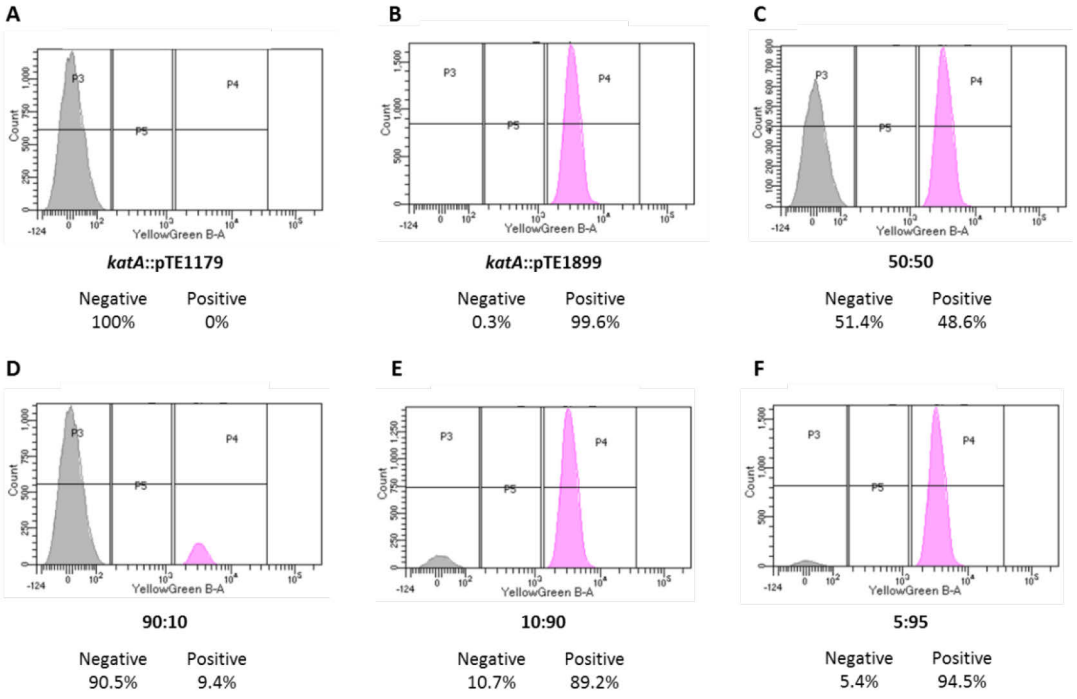


Fig. S3 Verification of flow cytometry sensitivity. Control strains were created by integrating the suicide vector pK18mob2 with (*katA*::pTE1899) and without (*katA*::pTE1179) a mCherry expression cassette into the chromosome of *M. extorquens*. Gates for were adjusted to distinguish between fluorescent (shown in pink) and non-fluorescent (shown in grey) cells using *katA*::pTE1179 (**A**), *katA*::pTE1899 (**B**), and mixtures thereof (**C-F**). 30,000 events were recorded per sample.

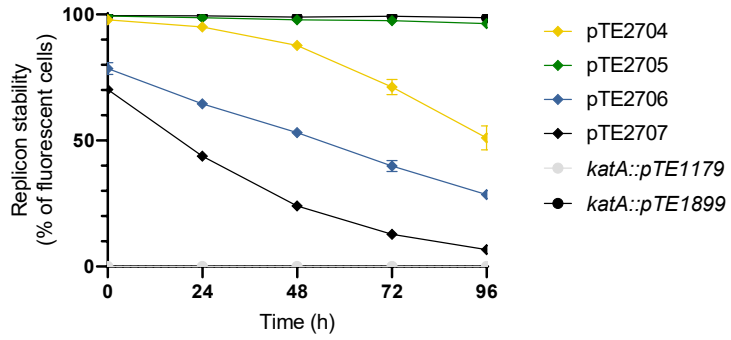


Fig. S4 Replicon stability measured by flow cytometry of mini-chromosomes with Mex-DM4, Mrad-JCM, Nham-3, and Mex-CM4 *repABC* regions (pTE2704-7, respectively) and an mCherry expression cassette. Control strains were created by integrating the suicide vector pK18mob2 with (*katA*::pTE1899) and without (*katA*::pTE1179) an mCherry expression cassette into the chromosome of *M. extorquens* AM1. 30,000 events were recorded per sample.

Table S5 DNA sequence of inducible promoters characterized in this study[a].

Promoter	Sequence
$P_{A1/O4/O3}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**TTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAACAATGATACT TAGATTCAATTGTGAGCGGATAACAATTTCACACA**TCTAGA
$P_{L/O4}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**GATAAATTATCTCTGGCGGTGTTGACATGTGAGCGGATAACAATG ATACTGAGCACATCAGCAGGACGCACTGAC**TCTAGA
$P_{L/O4/O3}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**GATAAATTATCTCTGGCGGTGTTGACATGTGAGCGGATAACAATG ATACTTAGATTCAATTGTGAGCGGATAACAATTTCACACAG**TCTAGA
$P_{L/O4/A1}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**GATAAATTATCTCTGGCGGTGTTGACATGTGAGCGGATAACAATG ATACTGAGCACATCGAGAGGGACACGGCGA**CTCTAGA
$P_{A1/O5/O4}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**AATTGTGAGCGGTAACAATTTAGTTGACTTGTGAGCGGATAACAA TGATACTTAGATTCATCGAGAGGGACACGGCGA**CTCTAGA
$P_{A1con/O5/O4}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**AATTGTGAGCGGTAACAATTTAGTTGACTTGTGAGCGGATAACAA TTATAATTAGATTCATCGAGAGGGACACGGCGA**CTCTAGA
$P_{A1/O4}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**TTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAACAATGATACT TAGATTCATCGAGAGGGACACGGCGA**CTCTAGA
$P_{A1/O4s}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**TTTATCAAAAAGAGTGTTGACTTGTGAGCGCTCACAATTGATACTT AGATTCATCGAGAGGGACACGGCGA**CTCTAGA
$P_{A1/O4s\_GA}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**TTTATCAAAAAGAGTGTTGACTTGTGAGCACTCACAATTGATACTT AGATTCATCGAGAGGGACACGGCGA**CTCTAGA
$P_{T5s/A1}$	GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTT CGTCTTCACCTCGAGAAAA**TTTATCAAAAAGAGTGTTGCTTTGTGAGCGCTCACAATTTATAATT AGATTCAATTGTGAGCGGATAACAATTTCACACAG**TCTAGA

[a]The promoter region is shown in **bold**, AatII and XbaI restriction sites are underlined. The surrounding sequences were included only to facilitate primer hybridization.

REFERENCES FOR SUPPLEMENTARY INFORMATION

1. Delaney, N. F., Kaczmarek, M. E., Ward, L. M., Swanson, P. K., Lee, M. C., and Marx, C. J. (2013) Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*, *PLoS One 8*, e62957.
2. Dohlemann, J., Brennecke, M., and Becker, A. (2016) Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/loxP site-specific recombination, *J. Biotechnol. 233*, 160-170.
3. Figurski, D. H., and Helinski, D. R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans, *Proc. Natl. Acad. Sci. USA 76*, 1648-1652.
4. Tauch, A., Zheng, Z., Puhler, A., and Kalinowski, J. (1998) Corynebacterium striatum chloramphenicol resistance transposon Tn5564: genetic organization and transposition in *Corynebacterium glutamicum*, *Plasmid 40*, 126-139.
5. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc. Natl. Acad. Sci. USA 97*, 6640-6645.
6. Chang, A. C., and Cohen, S. N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid, *J. Bacteriol. 134*, 1141-1156.
7. Urban, J. H., and Vogel, J. (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo, *Nucleic Acids Res. 35*, 1018-1037.
8. Ind, A. C., Porter, S. L., Brown, M. T., Byles, E. D., de Beyer, J. A., Godfrey, S. A., and Armitage, J. P. (2009) Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*, *Appl. Environ. Microbiol. 75*, 6613-6615.
9. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., and Erb, T. J. (2015) A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria, *ACS Synth. Biol. 4*, 430-443.
10. Döhlemann, J., Wagner, M., Happel, C., Carrillo, M., Sobetzko, P., Erb, T. J., Thanbichler, M., and Becker, A. (2017) A Family of Single Copy repABC-Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium *Sinorhizobium meliloti*, *ACS Synth. Biol. 6*, 968-984.

CHAPTER II

Initial engineering of the Calvin-Benson-Bassham (CBB) cycle in *Methylorubrum extorquens* AM1[‡]

[‡]Published in: **Metabolic Engineering** 47 (2018) 423–433.

An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1

Lennart Schada von Borzyskowski[a,b], Martina Carrillo[a], Simeon Leupold[d], Timo Glatter[a], Patrick Kiefer[b], Ramon Weishaupt[b], Matthias Heinemann[d], Tobias J. Erb[a,c]

[a]Max Planck Institute for Terrestrial Microbiology, Department of Biochemistry and Synthetic Metabolism, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany
[b]Institute of Microbiology, ETH Zurich, Vladimir-Prelog-Weg 4, 8093 Zurich, Switzerland
[c]SYNMIKRO, LOEWE Center for Synthetic Microbiology, Universität Marburg, 35043 Marburg, Germany
[d]Molecular Systems Biology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Metabolic Engineering 2018, 47: 423–433.

Author contributions

T.J.E. conceived the project; L.S.v.B. and T.J.E. designed research, analyzed data and wrote the manuscript with input from all authors; L.S.v.B., M.C. and R.W. performed the experiments; T.G. performed LC-MS measurements for proteomics and analyzed the data; P.K. performed LC-MS measurements for metabolic tracer analysis and provided advice and resources for data analysis; and S.L. and M.H. performed in silico analyses

Abstract

Organisms are either heterotrophic or autotrophic, meaning that they cover their carbon requirements by assimilating organic compounds or by fixing inorganic carbon dioxide (CO_2). The conversion of a heterotrophic organism into an autotrophic one by metabolic engineering is a long-standing goal in synthetic biology and biotechnology, because it ultimately allows for the production of value-added compounds from CO_2. The heterotrophic Alphaproteobacterium *Methylobacterium extorquens* AM1 is a platform organism for a future C_1-based bioeconomy. Here we show that *M. extorquens* AM1 provides unique advantages for establishing synthetic autotrophy, because energy metabolism and biomass formation can be effectively separated from each other in the organism. We designed and realized an engineered strain of *M. extorquens* AM1 that can use the C_1 compound methanol for energy acquisition and forms biomass from CO_2 by implementation of a heterologous Calvin-Benson-Bassham (CBB) cycle. We demonstrate that the heterologous CBB cycle is active, confers a distinct phenotype, and strongly increases viability of the engineered strain. Metabolic ^{13}C-tracer analysis demonstrates the functional operation of the heterologous CBB cycle in *M. extorquens* AM1 and comparative proteomics of the engineered strain show that the host cell reacts to the

implementation of the CBB cycle in a plastic way. While the heterologous CBB cycle is not able to support full autotrophic growth of *M. extorquens* AM1, our study represents a further advancement in the design and realization of synthetic autotrophic organisms.

1. Introduction

All cellular life forms need to assimilate carbon for the synthesis of organic molecules that comprise their biomass. This fundamental feature of life is achieved by one of two distinct metabolic modes: Heterotrophy, i.e., the assimilation of reduced organic carbon compounds, or autotrophy, i.e., the fixation of inorganic carbon dioxide (CO_2) into biomass. The latter process is typically energized by light (photo-autotrophy) or chemical energy (chemo-autotrophy).

In the past, biotechnology has mainly capitalized on heterotrophic metabolism for the production of value-added chemicals from simpler carbon compounds, such as sugars or amino acids (Dellomonaco et al., 2011; Enquist-Newman et al., 2014; Galanie et al., 2015; Paddon et al., 2013). The direct utilization of CO_2 as carbon feedstock in biotechnology, however, has gained considerable interest recently. Metabolic engineering of autotrophs could create novel routes that allow the direct conversion of CO_2 into valuable chemicals (Bar-Even et al., 2010; Liu et al., 2016; Nichols et al., 2015; Schwander et al., 2016). When coupled to photosynthesis, hydrogen gas, reduced chemical compounds or electricity, these routes would not only provide a sustainable carbon capture and conversion strategy for biotechnology, but also circumvent the need for organic carbon from plant biomass, with its low photosynthetic efficiency (typically < 1%) and ethical concerns as it directly competes with food production (Fargione et al., 2008; Sims et al., 2010).

For production of value-added compounds from CO_2, two different approaches can be envisioned. One approach is the implementation of biosynthetic production pathways into naturally existing carbon-fixing organisms, such as cyanobacteria and algae. The other option is to convert a conventional heterotrophic production strain into a ‘synthetic autotroph’. Compared to the first approach, the latter strategy is more radical. Implementing autotrophy into a native heterotroph requires a complete remodeling of its central carbon and energy metabolism, which is comparable to a ‘metabolic heart transplantation’ (Bar-Even et al., 2010). Therefore, the successful realization of a ‘synthetic autotroph’ will be a milestone in synthetic biology that probes the plasticity and modularity of metabolism to realize a ‘plug-and-play principle’ in living organisms.

Previous attempts of engineering a synthetic autotrophic organism largely focused on the model bacterium *Escherichia coli*. In a pioneering study, all enzymes of the 3-hydroxypropionate bicycle for CO_2 fixation were functionally expressed in *E. coli*. However, neither autotrophic growth nor a CO_2-fixing phenotype of the engineered strain was achieved (Mattozzi et al., 2013). Other groups have focused on implementing the

Calvin-Benson-Bassham cycle (CBB cycle) for CO_2 fixation in this organism. Overexpression of the CBB cycle's key enzymes, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and phosphoribulokinase (Prk) in *E. coli* resulted in a significant reduction of CO_2 emissions (Zhuang and Li, 2013), and even allowed the biosynthesis of sugars from CO_2 (Antonovsky et al., 2016). Although these studies represent important steps towards realizing 'synthetic autotrophy', all CO_2-fixing *E. coli* strains created so far still depend on addition of a multi-carbon compound that is produced from biomass feedstocks.

Here, we sought to move one step further by engineering an organism that is able to satisfy its needs for energy and biomass solely from single-carbon sources that are not derived from plant feedstocks. To this end, we chose the Alphaproteobacterium *Methylobacterium extorquens* AM1 (Peel and Quayle, 1961) as chassis organism. This pink-pigmented, aerobic organism can grow heterotrophically on multicarbon compounds as well as on the reduced single carbon compound methanol. Methanol is projected to play a key role in future bioeconomies (Bertau et al., 2014; Olah, 2013; Schrader et al., 2009), and *M. extorquens* AM1 is considered a biotechnological platform organism of the future. Several tools for genetic and metabolic engineering of *M. extorquens* AM1 are available (Choi et al., 2006; Chubiz et al., 2013; Marx, 2008; Marx and Lidstrom, 2001, 2002, 2004; Schada von Borzyskowski et al., 2015) and different *Methylobacterium* strains have already been developed to produce value-added chemicals from methanol (Hofer et al., 2010; Hu and Lidstrom, 2014; Ochsner et al., 2015; Orita et al., 2014; Sonntag et al., 2014, 2015a, 2015b). Expanding the metabolic capacity of *M. extorquens* AM1 to form biomass from the ubiquitous raw material CO_2, while using the highly reduced energy carrier methanol solely as an electron source, is a plausible and highly desirable goal.

In this study, we established a heterologous CBB cycle for CO_2 fixation in *M. extorquens* AM1. We engineered the central carbon metabolism of *M. extorquens* AM1 to separate energy acquisition from biomass formation. In our engineered strain, electrons are provided from methanol oxidation, while biomass is formed by CO_2 fixation via the heterologous CBB cycle. We demonstrate operation of the heterologous CBB cycle by growth phenotyping, ^{13}C-tracer analysis, and whole-cell proteomics. Our results reveal a positive phenotype upon expression of the CBB cycle as well as significant metabolic and proteomic changes in the engineered strain. Our work represents a stepping stone towards creating a fully synthetic autotrophic organism.

2. Material and methods

2.1. Bacterial strains, culture conditions, and plasmid delivery

If not mentioned otherwise, wild-type (WT) *M. extorquens* AM1 (Peel and Quayle, 1961) was used. *E. coli* DH5α was used for construction and amplification of all plasmids used in this work. *M. extorquens* AM1 was cultured at 30 °C in minimal medium (Peyraud et al.,

2009) supplemented with either 124 mM methanol or 30.8 mM succinate, or in Nutrient Broth (NB) without additional NaCl (SigmaAldrich, St. Louis, MO, USA). *E. coli* DH5α was cultured in LB medium at 37 °C. When appropriate, media were supplemented with tetracycline or kanamycin at a concentration of 10 or 50 μg mL^{-1}, respectively. Competent cells of *M. extorquens* AM1 WT and all mutant strains were made according to a published procedure, and electroporation was used for plasmid delivery into *M. extorquens* AM1 as described previously (Toyama et al., 1998).

2.2. Plasmid construction

All oligonucleotides used in this study are listed in **SI Table 1**. All plasmids used and created in this study are listed in **SI Table 2**. Cloning was performed according to standard protocols (Sambrook and Russell, 2001). Restriction enzymes and T4 DNA ligase were obtained from NEB (Frankfurt am Main, Germany). FastAP thermosensitive alkaline phosphatase was obtained from Thermo Scientific (St. Leon-Rot, Germany). PCRs were conducted with Phusion polymerase (Thermo Scientific, St. Leon-Rot, Germany) following the recommendations of the manufacturer using the protocol for GC-rich templates. Correct plasmid sequences were verified by Sanger sequencing (Eurofins, Ebersberg, Germany). pCM80 (Marx and Lidstrom, 2001) was cut with BamHI and EcoRI, the PCR product of RuBisCO (amplified from genomic DNA of *Rhodospirillum rubrum* S 1 with the primers R_Rbc_fw and R_Rbc_rv) was digested with the same enzymes and ligated into the dephosphorylated vector backbone to generate pTE92. pTE92 was cut with HindIII and BamHI, the PCR product of Prk from *Methylobacterium extorquens* AM1 (generated with primers M_Prk_fw and M_Prk_rv) was digested with the same enzymes and ligated into the dephosphorylated vector backbone to generate pTE94. pTE95 was generated by site-directed mutagenesis using pTE94 as template and the primers K191M_QC_fw and K191M_QC_rv. Prk was PCR-amplified from genomic DNA of *Synechococcus elongatus* PCC 7942 with the primers S_Prk_fw and S_Prk_rv. The resulting PCR product was digested with HindIII and BglII and ligated into pTE92 that had been digested with HindIII and BamHI to generate pTE96. The large and small subunits of RuBisCO (RuBisCO LSU + SSU) as well as Prk were PCR-amplified from genomic DNA of *Paracoccus denitrificans* DSM 413 with the primers P_RbcL_fw, P_RbcL_rv, P_RbcS_fw, P_RbcS_rv, P_Prk_fw and P_Prk_rv respectively. The PCR products were digested with NdeI and EcoRI and ligated into pET28b that had been digested with the same enzymes to generate pTE245, pTE246, and pTE248, respectively. The three His-tagged ORFs were excised from these vectors with XbaI and EcoRI and ligated into pTE100 that had been digested with XbaI and MunI to generate pTE500, pTE502, and pTE504. His-tagged Rbc LSU was excised from pTE500 with XbaI and KpnI and ligated into pTE102 that had been digested with SpeI and KpnI to generate pTE519. His-tagged Prk was excised from pTE504 with XbaI and KpnI and ligated into pTE519 that had been digested with SpeI and KpnI to generate pTE535. His-tagged Rbc SSU was excised from pTE502 with XbaI and KpnI and ligated into pTE535 that had

been digested with SpeI and KpnI to generate pTE550. The flanking regions of *glyA* from *M. extorquens* AM1 were PCR amplified with the primers glyA_up_fw, glyA_up_rv, glyA_do_fw and glyA_do_rv. The PCR products of the up- and downstream regions were used as templates to generate a fusion PCR product by overlap PCR with the primers glyA_up_fw and glyA_do_rv. The overlap PCR product was digested with EcoRI and XbaI and ligated into pK18-mob-sacB (Schäfer et al., 1994), which had been digested with the same enzymes, to generate pTE210.

2.3. Strain construction

All strains used and created in this study are listed in SI Table 2. *M. extorquens* AM1 Δ*glyA* was made from *M. extorquens* AM1 WT by transforming pTE210 into electrocompetent cells and screening for kanamycin resistance and sucrose sensitivity. The deletion was confirmed by diagnostic PCR and sequencing of the PCR product. *M. extorquens* AM1 Δ*cel* Δ*ftfL*::Kan was made from *M. extorquens* AM1 Δ*cel* (strain CM2720) (Delaney et al., 2013) by introduction of the *ftfL* deletion using pCM216 (Marx et al., 2003). The transformants were screened for kanamycin resistance and tetracycline sensitivity. The deletion was confirmed by diagnostic PCR and sequencing of the PCR product.

2.4. Cell lysis and determination of protein concentration

M. extorquens AM1 cultures were harvested in mid-exponential phase and aliquoted in 400 mg samples in 800 µL 100 mM Tris pH 7.8. Cells were lysed by sonication on ice with a Sonopuls HD 200 sonicator (Bandelin, Berlin, Germany), using 30% of the maximal amplitude for 15 s, which was repeated 5 times. The samples were centrifuged at 40,000 g and 4 °C for 60 min and the supernatant was transferred into new tubes. The total protein concentration of the supernatant was quantified by Bradford assay (Bradford, 1976) with bovine serum albumin (BSA) as standard and analyzed on 12.5% SDS PAGE gels (Laemmli, 1970), applying 20 µg protein per lane.

2.5. Spectrophotometric assays to measure RuBisCO and Prk activities

The activity of RuBisCO in cell-free extracts was measured by a continuous spectrophotometric assay determining NADH absorption at 340 nm on a Cary 50 UV/Vis spectrophotometer at 30 °C according to published procedures (Racker, 1957, 1962). The cell-free extract of a pCM80 transformant was used to correct for the background reaction in all assays. The activity of Prk in cell-free extracts was measured by modifying the same assay, using 0.6 mM ribulose 5-phosphate as substrate for Prk and adding purified RuBisCO from *R. rubrum* S 1 (a kind gift from Ryan Farmer and Bob Tabita, Ohio State

University) in excess, so that activity of Prk could be measured by coupling it to the downstream reaction of RuBisCO.

2.6. Growth assays

2.6.1. Assays in 96-well plates

All strains of *M. extorquens* AM1 were pre-grown at 30 °C in minimal medium containing succinate as sole carbon source until an OD_{600} of approximately 1.5. Then, cells were harvested, washed with minimal medium containing methanol and used to inoculate growth cultures of 180 µL minimal medium containing methanol in 96-well plates (Nunclon Delta Surface, Thermo Fisher Scientific, Darmstadt, Germany) sealed with parafilm. Growth at 30 °C was monitored in at least three technical replicates and at least two biological replicates at 600 nm in a Tecan Infinite M200Pro reader connected to a Tecan Gas Control Module (Tecan, Männedorf, Switzerland). Data was analyzed using the GraphPad Prism 6 software.

2.6.2. Assays in bioreactor

M. extorquens AM1 Δ*cel* Δ*ftfL*::Kan pTE94 was pre-grown at 30 °C in 200 mL minimal medium containing succinate as sole carbon source in an atmosphere with 5% CO_2 until an OD_{600} of approximately 1.5. Then, cells were harvested, washed with minimal medium containing methanol and used to inoculate a DASGIP SR0700ODLS bioreactor (Eppendorf, Hamburg, Germany) to an initial OD_{600} of approximately 0.2. During assembly of the bioreactor, all parts and tubes were rinsed with methanol and water in order to avoid contamination with ethanol and other substances. The bioreactor contained an initial volume of 700 mL, was stirred at 500 rpm and continuously gassed with 5% CO_2 in air at a flow rate of 15 L h^{-1}. pH was kept constant at 7.0 by addition of 3 M NH_4OH and 1 M H_2SO_4, and temperature was kept at 30 °C. Samples were taken via the sampling port of the bioreactor and OD_{600} was measured manually on a Ultrospec 3000 photospectrometer (Pharmacia Biotech (now part of GE Healthcare), Uppsala, Sweden).

2.7. Quantification of CFU/mL

Four replicate cultures each of *M. extorquens* AM1 Δ*cel* Δ*ftfL*::Kan pTE94 and pTE95 were pre-grown at 30 °C in minimal medium containing succinate as sole carbon source in an atmosphere with 5% CO_2 until an OD_{600} of approximately 1.5. Then, cells were harvested, washed with minimal medium containing methanol, and used to inoculate cultures of 25 mL minimal medium containing methanol to an OD_{600} of 0.2, which were incubated in an atmosphere with 5% CO_2. Every 50 h and also 8 h after inoculation, samples were taken. Dilution series of these samples were spread on plates (1.5% agar) containing minimal

medium with 30.8 mM succinate and tetracycline and incubated for 96 h at 30 °C, after which the resulting colonies were counted manually.

2.8. ^{13}C-tracer analysis-based metabolomics

Cultures of *M. extorquens* AM1 Δ*cel* Δ*ftfL*::Kan pTE94 and pTE95 were pre-grown at 30 °C in minimal medium containing succinate as sole carbon source in an atmosphere with 5% CO_2 until the late exponential phase. Then, cells were harvested, washed with minimal medium containing methanol and used to inoculate cultures of 30 mL minimal medium containing methanol. The cultures were incubated for 4 h in a defined atmosphere with 5% CO_2 at natural ^{13}C abundance. These cultures were then quickly collected by centrifugation (1 min, 10,000 g), the supernatant was discarded, and the pellet was washed with 30 mL medium (pre-warmed to 30 °C) containing no carbon source. The cultures were again pelleted (1 min, 10,000 g), the supernatant was discarded, and cells were re-suspended in 30 mL fresh medium containing either 124 mM ^{13}C-methanol or 8 mM ^{13}C-bicarbonate and 124 mM methanol at natural ^{13}C abundance. Cultures were then transferred into 100 mL shake flasks and incubated in an atmosphere containing 5% CO_2 at natural ^{13}C abundance when containing ^{13}C-methanol or incubated in a VLSB18 shaking water bath (VWR, Darmstadt, Germany) kept at 30 °C when containing ^{13}C-bicarbonate. This marked the starting point of the labeling experiment. To analyze the dynamic labeling process of metabolites, samples were taken after 0, 1, 5, 10, 15, 30, 45, 60 and 90 min. In short, each sample was transferred onto a polyethersulfone (PESU) 0.2 µm filter (Sartorius Stedim, Göttingen, Germany) pre-washed with an excess amount of 60 °C hot water. By applying vacuum, the medium was removed and the cells were washed with 5 mL 30 °C warm MilliQ water. The filter with cells was transferred into 100 mL Schott bottles filled with 8 mL of - 20 °C cold quenching solution (60% acetonitrile/20% methanol/20% 0.5 M formic acid) and kept on ice for 15 min. The bottles were sonicated every 5 min for 20 s using a Bransonic M2800-E sonication bath (VWR, Darmstadt, Germany). Afterwards, the filters were removed, samples were transferred into prepared 50 mL falcon tubes, frozen in liquid nitrogen, and lyophilized overnight. Sampled biomass amount was approximately 1 mg of cell dry weight. Lyophilized samples were resuspended in MilliQ water to give a final biomass concentration of 1 µg/µL. The resuspended samples were centrifuged at 20,000 g for 15 min at 4 °C to remove insoluble particles. Prior to HPLC-MS analysis, 10 µL of each sample was mixed with 90 µL of solvent A (see below). Nanoscale ion-pair reversed phase HPLC–MS analysis of the resuspended samples was performed as described previously (Müller et al., 2015), with the following modifications. Solvent A was 233 µM tributylamine in 234 µM acetic acid adjusted to pH 9.0 with NH_4OH; subsequently, 3% methanol were added to the solution. Solvent B was a 1:1 (V:V) mixture of 2-propanol and methanol. The applied gradient was as follows: 0 min (A 100%, B 0%); 3 min (A 100%, B 0%); 35 min (A 88%, B 12%), 36 min (A 10%, B 90%), 48 min (A 10%, B 90%), 49 min (A 100%, B 0%), 60 min (A 100%, B 0%). The flow rate was

kept constant at 400 µL/min throughout the gradient. Obtained MS-data were analyzed using eMZed 2.24.6 (Kiefer et al., 2013; emzed.ethz.ch).

2.9. Whole-cell shotgun proteomics using liquid chromatography-mass spectrometry (LC-MS/MS)

Four replicate cultures each of *M. extorquens* AM1 Δ*cel*, Δ*cel* pTE94, Δ*cel* Δ*ftfL*::Kan pTE94 and Δ*cel* Δ*ftfL*::Kan pTE95 were pre-grown at 30 °C in minimal medium containing succinate as sole carbon source in an atmosphere with 5% CO_2 until the late exponential phase. At this point, the first samples for proteomic analysis were taken. Then, cells were harvested, washed with minimal medium containing methanol and used to inoculate cultures of 200 mL minimal medium containing methanol, to an OD_{600} of approx. 0.2 and incubated in an atmosphere with 5% CO_2. After a sufficient incubation time (20 h for both of the Δ*cel* strains, 65 h for both of the Δ*cel* Δ*ftfL* strains) with these carbon sources, samples for proteomic analysis were taken. Samples were washed once with ice-cold PBS and cell pellets were flash-frozen in liquid nitrogen and stored at - 80 °C until further processing. Sample preparation was carried out as described previously (Glatter et al., 2015), with the following modifications: 2% sodium lauroyl sarcosinate (SLS) was used instead of 2% sodium deoxycholate (SDC), and quenching with N-acetylcysteine as well as digestion with LysC were omitted. LC-MS/MS analysis of digested lysates was performed on a Thermo QExactive Plus mass spectrometer (Thermo Scientific), which was connected to an electrospray ionsource (Thermo Scientific). Peptide separation was carried out using an Ultimate 3000 RSLCnano (Thermo Scientific) equipped with a RP-HPLC column (75 µm x 35 cm) packed in-house with C18 resin (1.9 µm; Dr. Maisch). The following separating gradient was used: 98% solvent A (0.15% formic acid) and 2% solvent B (80% acetonitrile, 0.15% formic acid) to 32% solvent B over 175 min and to 50% B for additional 20 min at a flow rate of 300 nl/min. The data acquisition mode was set to obtain one high resolution MS scan at a resolution of 70,000 full width at half maximum (at m/z 200) followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS attempts, the charged state screening modus was enabled to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 s. The ion accumulation time was set to 50 ms (MS) and 50 ms at 17,500 resolution (MS/MS). The automatic gain control (AGC) was set to 3 × 106 for MS survey scan and 1 × 105 for MS/MS scans. Label-free quantification of raw data and data evaluation was performed as described previously (Ahrné et al., 2013; Glatter et al., 2015). In short, the raw data was imported into Progenesis (Nonlinear Dynamics, version 2.0) and processed data was further evaluated using SafeQuant. For database search, the protein database for *M. extorquens* AM1 was downloaded from Uniprot (www.uniprot.org, download date: April 2016; RuBisCO from *R. rubrum* S 1 (Uniprot-ID P04718) was added manually) and search was performed using the decoy strategy. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues); two missed cleavages were allowed;

carbamidomethylation (C) was set as fixed modification; oxidation (M) as variable modification. The mass tolerance was set to 10 ppm for precursor ions and 0.02 Da for fragment ions. The heatmap shown in **Fig. 5** was made with GProX (Rigbolt et al., 2011).

2.10. In silico modelling

The in silico predictions of the biomass yields were done using flux balance analysis (FBA). To this end, the previously published genome scale metabolic reconstruction of *M. extorquens* AM1, iRP911, consisting of 1139 reactions and 977 metabolites, was used (Peyraud et al., 2011). To account for the introduced RuBisCO gene, one reaction, representing the conversion of D-Ribulose 1,5-bisphosphate and CO_2 to two molecules of 3-Phospho-D-glycerate, was added to the model. The FBA problem of the form

$$\boldsymbol{maximize}\ v_{BiomassSynthesis}$$

$$S_{ij} * v_j = 0$$

$$\boldsymbol{subject\ to}\ v_j^{lo} \leq v_j \leq v_j^{up}$$

where S_{ij} is the stoichiometric matrix and v_j the flux through the reaction *j*, was then implemented using GAMS 2.8.2 and solved using Cplex 12.7.0.0. The lower and upper bounds of v_j were used as specified in iRP911 and only the exchange of O_2, CO_2, H_2O, trace elements (NH_3, Pi, SO_4^{2-}, Na^+, Fe^{2+}, K^+, $Mg2^+$, Cu^{2+}, Mn^{2+}, Zn^{2+}, Cl^-) and the uptake of methanol was allowed. To realize the knockout mutants, the lower and upper bound of the respective reactions was set to 0. For Δ*glyA*, the reaction R0015 was set to 0, for Δ*ftfL* the reaction R0012 and for Δ*ccr* the reaction R0032. In case of Δ*ftfL* the glycine cleavage complex (R0243/R0244/R0245) and in case of Δ*ccr* the butyryl-CoA carboxylase (R0034) had to be additionally blocked to correctly predict the lethal phenotype, which had been determined in earlier experimental studies (Chistoserdova and Lidstrom, 1996; Marx et al., 2003).

3. Results

3.1. Engineering strategy to realize a synthetic CO_2-fixing M. extorquens AM1 strain

To convert the heterotroph *M. extorquens* AM1 into a synthetic autotroph, we designed an experimental strategy that is based on the decoupling of energy acquisition and carbon assimilation in this organism (Fig. 1). When *M. extorquens* AM1 grows on methanol, 84% of this single carbon compound is completely oxidized to CO_2 to acquire energy, while the remaining 16% of methanol are assimilated into biomass (Fig. 1A) through the serine cycle and a number of interconnected pathways (Peyraud et al., 2011). Deletion of genes essential for methanol assimilation yields mutant strains of *M. extorquens* AM1 that cannot grow on methanol anymore, because they are not able to convert it into biomass. Yet, these mutant strains are still able to draw energy and reducing power from methanol

oxidation. When these mutants are equipped with a functional CO_2 fixation pathway, carbon assimilation in *M. extorquens* AM1 should be restored, but from CO_2 and not from reduced organic carbon as in the wild-type (WT) strain, resulting in an autotrophic phenotype (Fig. 1B).

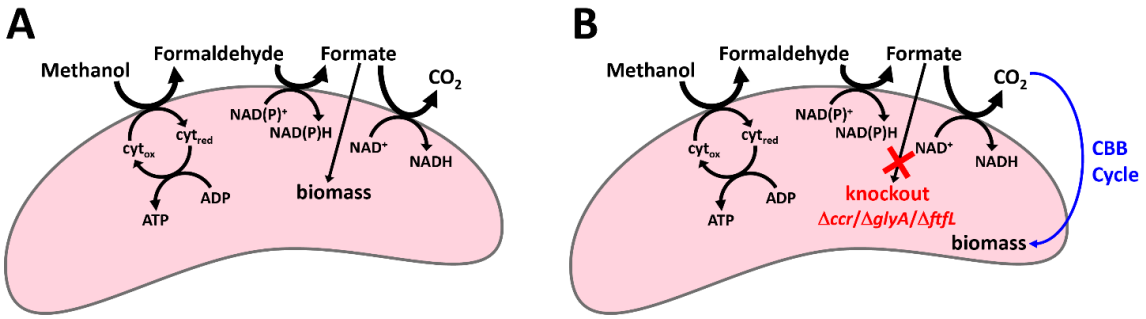


Fig. 1. Engineering strategy for the conversion of *M. extorquens* AM1 into a synthetic autotroph. (A) During WT growth on methanol, this C1 compound is fully oxidized to CO_2 for energy acquisition (i.e., the generation of ATP and reducing equivalents), while carbon for biomass formation is assimilated at the level of formate. (B) In this project, the assimilation of formate is prevented by gene knockouts, so that the cell is forced to utilize CO_2 for biomass formation via the heterologous CBB cycle, while energy acquisition from methanol oxidation still takes place.

First, we identified several potential target genes of the central carbon metabolism in *M. extorquens* AM1 (see Supplementary Fig. S1) that would allow us to largely decouple energy acquisition from carbon assimilation. These included (i) *ccr*, encoding crotonyl-CoA carboxylase/reductase, the key enzyme of the ethylmalonyl-CoA pathway (EMCP) for glyoxylate regeneration (Erb et al., 2007); (ii) *glyA*, encoding serine hydroxymethyltransferase, a key enzyme of the serine cycle (Chistoserdova and Lidstrom, 1994) and (iii) *ftfL*, encoding formate tetrahydrofolate ligase, the enzyme linking methanol oxidation and carbon assimilation (Marx et al., 2003). While the three genes are essential for growth on methanol (Chistoserdova and Lidstrom, 1994, 1996; Marx et al., 2003), the corresponding knock out strains are still able to grow on succinate as carbon source, allowing for genetic manipulations and convenient handling of the corresponding mutants.

To equip *M. extorquens* AM1 with an autotrophic CO_2-fixation module, we decided to implement the CBB cycle in this organism. Of all naturally evolved autotrophic CO_2 fixation pathways (Berg et al., 2010; Fuchs, 2011), the CBB cycle is the one that is most closely connected to proteobacterial central carbon metabolism and requires a minimal number of additional genes to be introduced. Based on genomic analysis, *M. extorquens* AM1 encodes all but one gene of the CBB cycle, namely RuBisCO, so that in principle the heterologous expression of only one protein should be sufficient to establish synthetic autotrophy in this organism.

We used a previously published stoichiometric metabolic network model (Peyraud et al., 2011) and Flux Balance Analysis (FBA) to validate and predict the outcome of this design strategy. Indeed, all three deletion strains (Δ*ccr*, Δ*glyA*, Δ*ftfL*) were predicted to be unable to grow on methanol (see Supplementary Table S3), in line with previously published results for these mutants. Upon implementation of the CBB cycle, the metabolic model predicted recovery of growth on methanol for the *glyA* and *ftfL* deletion strains, but not for the *ccr* deletion strain (see detailed results in Supplementary Table S3). In order to evaluate the output of these in silico analyses, we decided to test the three knockout strains as hosts for implementation of the synthetic CBB cycle in *M. extorquens* AM1.

3.2. Realization of a heterologous CBB cycle for CO_2 fixation in M. extorquens AM1

To establish a functional CBB cycle, we first evaluated different RuBisCO candidates for their functionality in *M. extorquens* AM1. We tested a type I RuBisCO from *Paracoccus denitrificans* DSM 413 as well as a type II RuBisCO from *Rhodospirillum rubrum* S 1. Both enzymes had been previously purified and biochemically characterized (Bowien, 1977; Schloss et al., 1979). Moreover, both enzymes are from Alphaproteobacteria closely related to *M. extorquens* AM1, with a similar codon usage and GC content, which makes their heterologous expression in *M. extorquens* AM1 favorable. Plasmid-based expression and solubility of the two enzymes in *M. extorquens* AM1 was confirmed by SDS-PAGE gel analysis (see Supplementary Fig. S2). When assayed in cell extracts of *M. extorquens* AM1, however, only the type II RuBisCO of *R. rubrum* S 1 was active with a specific CO_2 fixation rate of approximately 100 mU/mg total protein. Therefore, we decided to use this enzyme for our heterologous CBB cycle.

Although *M. extorquens* AM1 has a phosphoribulokinase (Prk) homolog encoded in its genome (META1p0758), only very low activity of this enzyme could be measured in extracts of methanol-grown cells (Kalyuzhnaya and Lidstrom, 2003). To realize a sufficiently active CBB cycle we decided to overexpress Prk in addition to RuBisCO. We tested the well-characterized Prk from the cyanobacterium *Synechococcus elongatus* PCC 7942 (Kobayashi et al., 2003), as well as the native homolog from *M. extorquens* AM1. While the cyanobacterial Prk could not be functionally expressed in *M. extorquens* AM1, the native homolog was highly soluble, as judged by SDS-PAGE gel analysis (see Supplementary Fig. S2). Activity of *M. extorquens* AM1 Prk was demonstrated by addition of purified *R. rubrum* S 1 RuBisCO to cell extracts, which allowed us to successfully reconstitute the reaction sequence from ribulose 5-phosphate to glycerate 3-phosphate.

Based on these results, we generated a plasmid-borne carbon fixation module that encoded Prk and RuBisCO in one operon under the control of the strong *mxaF* promoter (Anderson et al., 1990). As negative control for the carbon fixation module, we created a version of the same plasmid encoding Prk as well as a catalytically inactive RuBisCO variant. In this RuBisCO mutant, an essential active site lysine, which is carbamylated in

the WT enzyme, was substituted by a methionine, rendering the enzyme inactive (Cleland et al., 1998; Mueller-Cajar et al., 2007). While this RuBisCO active site mutant does not allow for functional CO_2 fixation, it still applies the same burden of protein expression on the host cell, minimizing any epistatic effects. We used this construct as negative control plasmid for all subsequent experiments.

3.3. Screening of M. extorquens AM1 deletion strains with the CBB cycle identifies a distinct phenotype

We next tested the effect of the heterologous CBB cycle in the background of the three mutant strains, Δ*ccr*, Δ*glyA*, and Δ*ftfL*. To this end, the mutant strains were transformed with the plasmid encoding the CBB cycle enzymes RuBisCO and Prk. Precultures were grown on succinate minimal medium, then cells were washed, transferred into methanol minimal medium and the optical density (OD_{600}) of the cultures was monitored over time in 96-well plates.

Under atmospheric CO_2 concentrations, none of the strains tested exhibited an observable phenotype. It is well known that under ambient atmosphere RuBisCO shows a strong side reaction with oxygen that causes accumulation of the toxic compound 2-phosphoglycolate (Bowes et al., 1971). However, when tested in an atmosphere containing 5% CO_2, that effectively suppresses the oxygenase side reaction of RuBisCO, the ΔftfL strain showed a distinct phenotype (Fig. 2). The OD_{600} of Δ*ftfL* carrying the carbon fixation module slowly increased by approximately 50% over the course of about 150 h after transfer into methanol medium, before it stopped. From the increase in OD_{600} during the initial 100 h, we estimated a doubling time of 230 ± 10 h. Notably, neither the Δ*glyA* (in contrast to the in silico prediction) nor the Δ*ccr* strains (in agreement with the in silico prediction) expressing the CBB cycle, or any of the strains transformed with the empty plasmid or the negative control plasmid with the inactive RuBisCO exhibited this phenotype under the same conditions.

The observed phenotype of the Δ*ftfL* strain overexpressing RuBisCO and Prk was reproducible over many biological replicates (n > 10), measured in separate experiments and with different layouts of the 96-well plates in order to exclude any artifacts. However, in several cases, we observed clumping of the cells, which is a well-known problem in *M. extorquens* AM1 liquid cultures (Delaney et al., 2013). Therefore, the *ftfL* deletion was introduced into the previously described Δ*cel* strain that shows a reduced clumping (Delaney et al., 2013). The resulting Δ*cel* Δ*ftfL* double mutant strain was used for all further experiments. In 96-well plate assays, the same phenotype observed for Δ*ftfL* could be reproduced multiple times (n > 10) when the Δ*cel* Δ*ftfL* strain was equipped with the carbon fixation module. Again, the observed phenotype was highly reproducible (see Supplementary Fig. S3) and did also not depend on the starting conditions, for example the initial OD_{600} (see Supplementary Fig. S4). A similar phenotype was observed for the

same strain when incubated in a bioreactor on methanol medium gassed with 5% CO_2 (see Supplementary Fig. S5). Under these conditions, the increase in OD_{600} ceased after approximately 100 h.

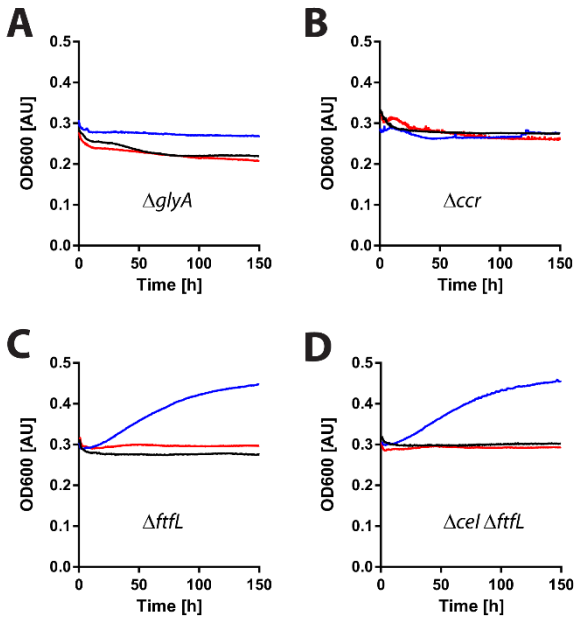


Fig. 2. Growth curves of engineered *M. extorquens* AM1 strains on methanol in a 5% CO_2 atmosphere. The deletion strains of *glyA* (A), *ccr* (B) and *ftfL* (C) are transformed with an empty plasmid backbone (black), a plasmid for the overexpression of RuBisCO and Prk (blue) or a plasmid for the overexpression of catalytically inactive RuBisCO and Prk (red). Growth assays with the double deletion strain Δ*cel* Δ*ftfL* yielded similar results as with Δ*ftfL* (D). Shown are representative curves of at least three technical and at least two biological replicates; for the strains Δ*ftfL* and Δ*cel* Δ*ftfL*, more than 10 biological replicates each were carried out.

In contrast to the phenotype observed when overexpressing RuBisCO and Prk in the Δ*cel* Δ*ftfL* background, the Δ*cel* strain carrying the same plasmid showed strongly impaired growth on methanol and CO_2 (see Supplementary Fig. S6 and Supplementary Note 1). These experiments demonstrated that a heterologous CBB cycle had a positive effect only in the *M. extorquens* AM1 Δ*cel* Δ*ftfL* background.

3.4. Presence of the CBB cycle increases viability of M. extorquens AM1 Δcel ΔftfL

Although the OD_{600} of *M. extorquens* AM1 Δ*cel* Δ*ftfL* expressing the CBB cycle increased on methanol minimal medium with 5% CO_2, the cultures stagnated reproducibly after about 150 h. Transfer in fresh methanol medium did not stimulate a further increase in OD_{600}, independent of whether the transfer took place during the first 150 h or in the stagnation phase (data not shown). Clearly, the Δ*cel* Δ*ftfL* cells engineered with the CBB cycle were not able to sustain continuous growth under these conditions, but seemed to

survive longer. Thus, we investigated whether the presence or absence of the heterologous CBB cycle actually correlated to a difference in the number of viable cells.

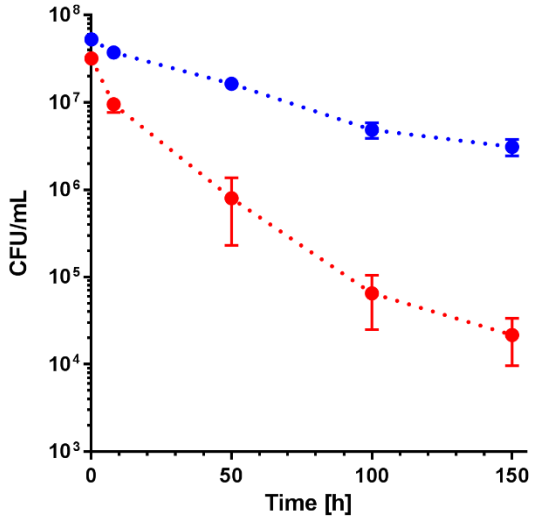


Fig. 3. Time course of CFU/mL during incubation with methanol medium and 5% CO_2. After 150 h, about 100-fold more cells are re-culturable of the strain Δ*cel* Δ*ftfL* with the CBB cycle (blue circles) compared to the negative control strain Δ*cel* Δ*ftfL* with the inactive CBB cycle (red circles). The average of four biological replicates is shown; error bars denote standard deviation.

Because direct cell counting does not yield information about the viability of the cells, we quantified the number of colony forming units (CFU)/mL as a proxy for cell viability. We compared *M. extorquens* AM1 Δ*cel* Δ*ftfL* overexpressing the functional CBB cycle to the negative control that expressed the catalytically inactive RuBisCO. Samples were taken every 50 h, as well as 8 h after inoculation, and plated on agar plates containing succinate minimal medium to quantify viable cells as CFU/mL. While both cultures started from the same initial CFU/mL upon transfer on methanol, the number of viable cells dropped rapidly in the negative control with the non-functional RuBisCO. After 150 h, less than 0.07% of the initial number of CFU/mL were recovered on succinate medium. In contrast, approximately 7% of the initial CFU/mL could be recovered after 150 h for the cells that overexpressed the functional CBB cycle (Fig. 3). The result that not all cells were recovered is in line with the observation that when switching from methanol to succinate medium, only a fraction of cells actively continues to grow (Strovas and Lidstrom, 2009; Strovas et al., 2007). In summary, cells containing a functional carbon fixation module remained more viable by two orders of magnitude compared to the negative control, suggesting that the 100-fold increased viability was directly linked to an actively operating CBB cycle.

3.5. ^{13}C-tracer analysis demonstrates functional CO_2 fixation in M. extorquens AM1 via the CBB cycle

Next, we sought to demonstrate functionality of the heterologous CBB cycle by ^{13}C-tracer analysis. For these experiments, *M. extorquens* AM1 Δ*cel* Δ*ftfL* cultures overexpressing either the functional CBB cycle or the negative control were pre-grown in minimal medium containing succinate, washed, and then transferred into minimal medium containing methanol in an atmosphere with 5% CO_2. 4 h after the transfer into methanol medium, aliquots of the cultures were taken and incubated with either ^{13}C-methanol in an atmosphere with 5% non-labeled CO_2, or non-labeled methanol plus ^{13}C-bicarbonate. Then, samples were taken and the time course of ^{13}C-tracer incorporation into metabolites was monitored over 90 min by HPLC-MS. No incorporation of ^{13}C-carbon into metabolites was measured from ^{13}C-methanol in any of the engineered strains (data not shown), confirming that methanol cannot enter central carbon metabolism in the *ftfL* mutant background. When incubated with ^{13}C-bicarbonate, however, we observed incorporation of label into central carbon metabolites in the Δ*cel* Δ*ftfL* strain carrying a functional RuBisCO.

During methylotrophic growth of *M. extorquens* AM1 WT, 50% of carbon is derived from CO_2 that is assimilated via phosphoenolpyruvate (PEP) carboxylase in the serine cycle (Crowther et al., 2008; Large et al., 1961; Peyraud et al., 2011). Notably, CO_2 fixation via PEP carboxylase cannot proceed in the same way in the *ftfL* deletion mutant. Interruption of formate-tetrahydrofolate synthesis prevents continuous operation of the serine cycle, which should result in decreased label incorporation from ^{13}C-bicarbonate via PEP carboxylase. In fact, between 10% and 20% label incorporation into PEP was observed in the *ftfL* mutant background (Fig. 4).

On the other hand, expression of a functional CBB cycle in the background of the *ftfL* mutant should result in the incorporation of bicarbonate-derived ^{13}C-label into CBB cycle metabolites. In agreement with these expectations, we observed ^{13}C-labeled metabolites in the strain engineered with the heterologous CBB cycle. These included phosphoglycerates, pentose mono- and bisphosphates, as well as hexose- and heptose monophosphates, which are all intermediates of the CBB cycle. Label incorporation proceeded rapidly over the course of the initial 30 min to then become slower and stagnate at 20–30% ^{13}C-labeled carbon for most CBB cycle metabolites (Fig. 4). Notably, fully labeled fractions of all important metabolites of the CBB cycle (e.g., five-fold labeled pentose mono- and bisphosphates, seven-fold labeled sedoheptulose 7-phosphate, etc.) appeared over the course of the labeling experiment, clearly indicating that the CBB cycle turned multiple times and was active in the strain that carried a functional RuBisCO. The observed labeling patterns and the maximum incorporation of label (20–30%) is in line with results of other studies that also aimed at implementing novel central carbon metabolic features into host organisms, e.g. the engineering of a heterologous ribulose monophosphate cycle in E. coli (Müller et al., 2015).

The ^{13}C-tracer incorporation observed in the strain engineered with the heterologous CBB cycle was strongly reduced in the negative control expressing the inactive RuBisCO (Fig. 4). We only observed incorporation into PEP, which is well in line with the residual effect of PEP carboxylase in the interrupted serine cycle. Moreover, we also detected some increasing incorporation of ^{13}C-label into pentose mono- and –bisphosphates. Note that ribulose 1,5-bisphosphate (RuBP) can still be formed in the negative control strain that overexpresses Prk. However, due to inactive RuBisCO, this compound is a metabolic dead end in this strain and accumulates over time. Since RuBP cannot be further metabolized by RuBisCO and is not utilized for biosynthesis, ^{13}C-label in this metabolite can accumulate stronger than in other metabolites, which are drained for biosynthesis or further converted in their respective pathways. With increasing amounts of RuBP being synthesized by Prk, the reaction equilibrium with ribulose 5-phosphate (Ru5P) shifts towards the latter compound, so that ^{13}C-label also accumulates in Ru5P, and, consequently, the pentose phosphate pool. Taken together, the ^{13}C-tracer analysis data strongly suggested that the CBB cycle was operating in the *M. extorquens* AM1 Δ*ftfL* strain equipped with Prk and RuBisCO.

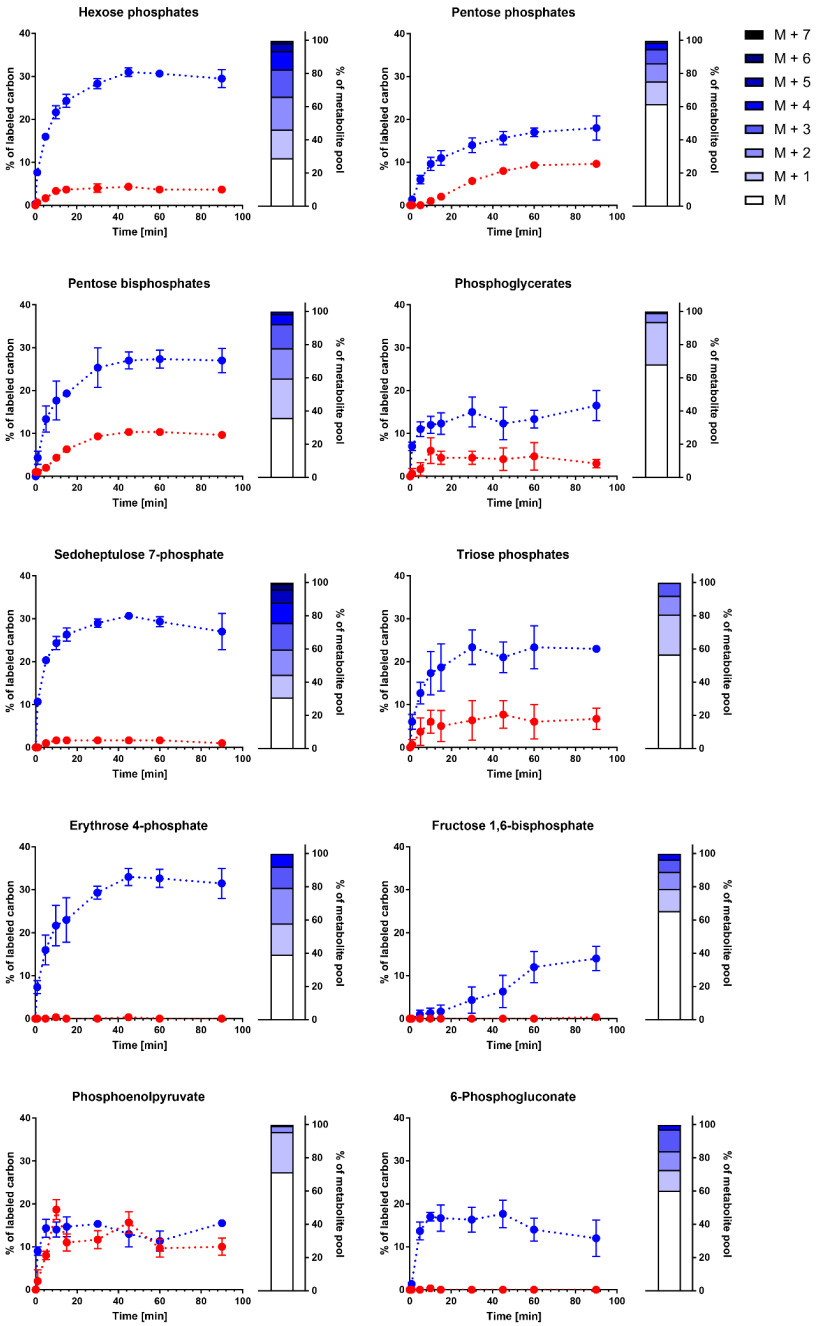


Fig. 4. ^{13}C-tracer analysis in engineered *M. extorquens* AM1 strains with the heterologous CBB cycle from ^{13}C-bicarbonate. ^{13}C-label incorporation in selected metabolites over the course of 90 min is shown 4 h after transfer to minimal medium containing methanol and 5% CO_2. Blue circles: Δ*cel* Δ*ftfL* with the CBB cycle; red circles: Δ*cel* Δ*ftfL* with the inactive CBB cycle. The average of three biological replicates is shown; error bars denote standard deviation. The adjacent bar graphs show the mass isotopologue distribution of the respective metabolite in strain Δ*cel* Δ*ftfL* with the CBB cycle after 60 min incubation with ^{13}C-bicarbonate. The average of three biological replicates is shown; error bars are not shown to increase clarity. Hexose phosphates include glucose 6-phosphate, fructose 6-phosphate and glucose 1-phosphate pools. Pentose phosphates include xylose 5-phosphate, ribose 5-phosphate, xylulose 5-phosphate, and ribulose 5-phosphate pools. Pentose bisphosphates include ribulose 1,5-bisphosphate and ribose 1,5-bisphosphate pools. Phosphoglycerates include 2-phosphoglycerate and 3-phosphoglycerate pools. Triose phosphates include dihydroxyacetonephosphate and glyceraldehyde 3-phosphate pools.

3.6. Comparative proteome analysis reveals specific changes in response to introduction of the CBB cycle

Finally, to identify further adaptions that allowed to establish the observed phenotype we aimed at investigating the effect of the synthetic CBB cycle onto the proteome of *M. extorquens* AM1. To this end, we analysed the proteome of four different strains upon a substrate switch from succinate to methanol. Our analysis included the *M. extorquens* AM1 Δ*cel* ('wild-type') strain with three engineered variants: Δ*cel* expressing Prk and RuBisCO, Δ*cel* Δ*ftfL* expressing Prk and RuBisCO and Δ*cel* Δ*ftfL* expressing Prk and inactive RuBisCO.

We first grew all strains on succinate and withdrew samples for proteome analysis during exponential phase (succinate condition). Subsequently, we transferred the cells to minimal medium containing methanol as sole carbon source and incubated them in an atmosphere containing 5% CO_2. After a sufficient incubation time (20 h for both of the Δ*cel* strains, 65 h for both of the Δ*cel* Δ*ftfL* strains), we withdrew an additional sample for analysis (methanol condition). We performed whole-cell shotgun proteomics and quantified the change in the proteome of each individual strain in response to the switch from the succinate to the methanol condition (Fig. 5, individual columns). Finally, we compared this data between the four strains. Our analysis identified striking differences in the expression levels of central carbon metabolic enzymes between the four different strains.

It is known that upon switching *M. extorquens* AM1 from succinate to methanol as carbon source, enzymes of the TCA cycle are downregulated, while enzymes of the EMCP become upregulated (Schneider et al., 2012). This pattern was confirmed for the two Δ*cel* strains. In contrast, TCA cycle enzyme expression levels, as well as expression levels of EMCP enzymes did not change in the same way in both Δ*cel* Δ*ftfL* deletion strains during the substrate switch. These changes indicated that the deletion of *ftfL* itself had some general effect on the central carbon metabolism of *M. extorquens* AM1 and that this effect was independent of the presence of an active or inactive CBB cycle.

In addition, PQQ-dependent methanol dehydrogenase and its cytochrome electron acceptor (MxaFGI) are known to be upregulated upon the switch from succinate to methanol (Bosch et al., 2008; Okubo et al., 2007). This pattern was confirmed for both Δ*cel* strains, but notably also the Δ*cel* Δ*ftfL* strain equipped with the active CBB cycle. In contrast, MxaFGI was not upregulated in the strain Δ*cel* Δ*ftfL* carrying the inactive CBB cycle during the substrate switch. This suggests that the *ftfL* deletion strain is able to generate sufficient levels of ATP through methanol oxidation only in the presence of an active CBB cycle, which is in line with the observed phenotype and the increased viability of this strain compared to the inactive CBB cycle control, as described above. The expression levels of RuBisCO and Prk barely decreased between the two sampling time points in both Δ*cel* Δ*ftfL* strains (see Supplementary File 1 for details), indicating that the

CBB cycle was active over the course of the whole experiment.

Another striking difference between the four strains was the specific upregulation of a two-subunit transketolase. This enzyme was strongly upregulated only in Δ*cel* Δ*ftfL* carrying the functional CBB cycle, while its expression levels were not significantly influenced by the substrate switch in all other strains. Transketolase is a key enzyme in the CBB cycle for interconversion of sugar phosphates. The specific upregulation of this enzyme in the Δ*cel* Δ*ftfL* strain with the active CBB cycle is expected to increase (or stabilize) flux through the synthetic carbon fixation pathway.

Along the same lines, several enzymes at the beginning of biosynthetic pathways that branch off from the CBB cycle were downregulated only in Δ*cel* Δ*ftfL* expressing the active CBB cycle. This holds true for GlmU, which produces amino sugars from fructose 6-phosphate, PurC and PurK, which operate in the purine biosynthetic pathway that starts from ribose 5-phosphate, RibB, which catalyzes the first step in riboflavin biosynthesis starting from ribulose 5-phosphate, as well as AroG, the enzyme for the first step in the biosynthesis of aromatic amino acids, which utilizes PEP and erythrose 4-phosphate, another metabolite of the CBB cycle, as substrates. The expression levels of all these enzymes in the three other strains were either not significantly changed or barely (0 > Log2 fold change > -0.5) changed upon the switch from succinate to methanol (see Supplementary File 1 for details). The specific downregulation of those five enzymes that drain carbon from the sugar phosphate pools into biosynthesis probably contribute to a more stable operation of the synthetic CBB cycle in the Δ*cel* Δ*ftfL* strain with the functional CBB cycle module. In summary, these experiments identified a specific reaction of the proteome in the presence of an active CBB cycle under selective pressure that explains the observed phenotype and the prolonged viability described above.

4. Discussion

This work aimed at establishing synthetic autotrophy in *M. extorquens* AM1. To that end, we decoupled energy metabolism and carbon assimilation in *M. extorquens* AM1 and equipped the organism with a heterologous CBB cycle. This strategy allowed the organism to acquire energy by methanol oxidation, while it became dependent on CO_2 fixation for biomass synthesis.

To establish a transiently functional CBB cycle for CO_2 fixation, it was sufficient to overexpress only two enzymes in *M. extorquens* AM1 – RuBisCO and Prk. Although *M. extorquens* AM1 encodes an endogenous Prk, this enzyme only performs a regulatory role (Ochsner et al., 2017) and needs to be overexpressed to operate a heterologous CBB cycle in *M. extorquens* AM1. Notably, overexpression of Prk in both the presence and the absence of active RuBisCO was not toxic to *M. extorquens* AM1. This is in stark contrast to

E. coli, which is known to be sensitive to ribulose 1,5-bisphosphate formation when overexpressing spinach Prk in the absence of RuBisCO (Hudson et al., 1992).

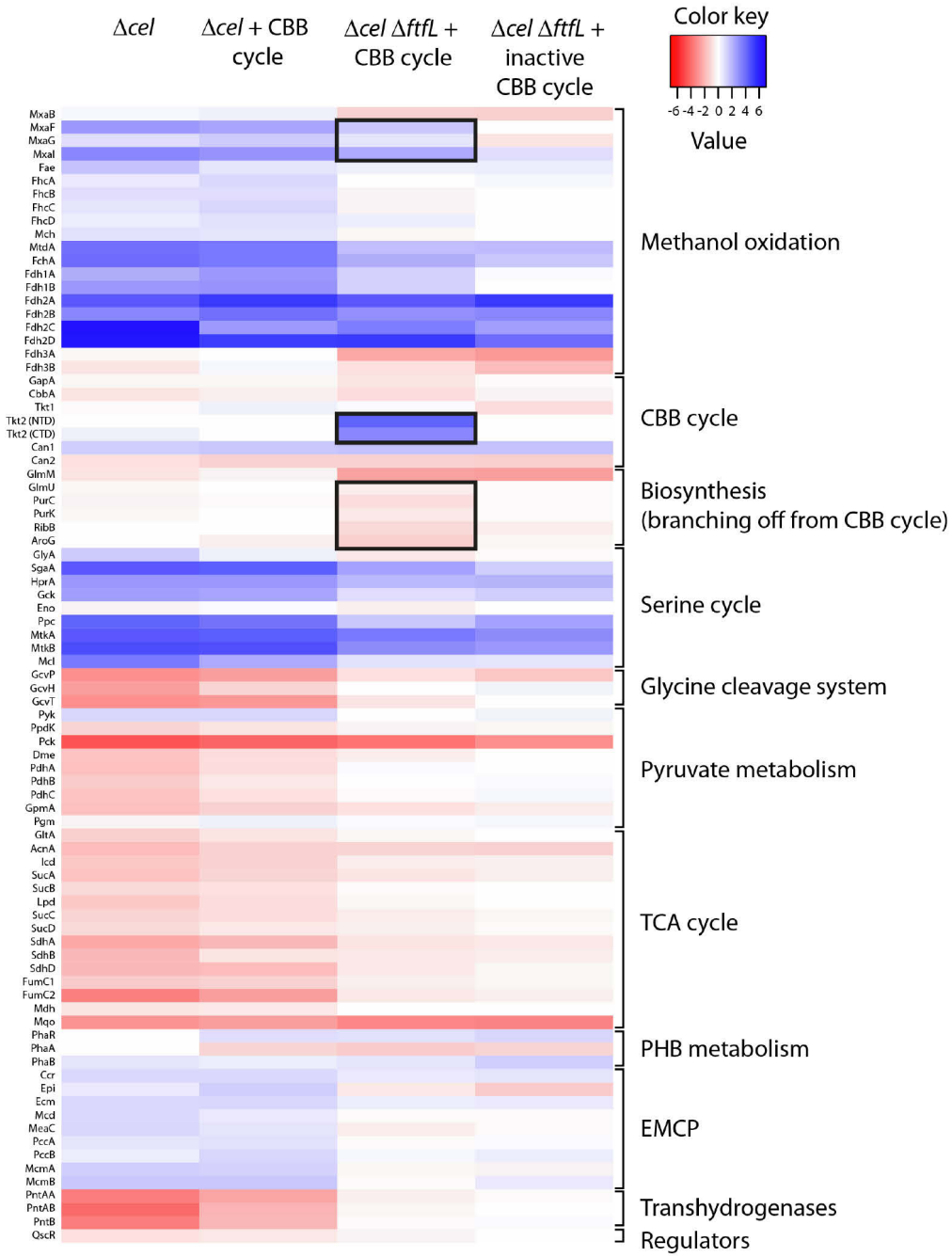


Fig. 5. Whole-cell shotgun proteomics of engineered *M. extorquens* AM1 strains. The heatmap shows the Log2 fold changes in the expression levels of selected proteins upon a substrate switch from succinate to methanol. Proteins that are discussed in the main text are denoted by a black frame. For detailed values and significance of the data see Supplementary File 1.

Expression of the CBB cycle in *M. extorquens* AM1 Δ*cel* Δ*ftfL* resulted in a distinct phenotype that strongly increased viability of the engineered cells compared to the

negative control. ^{13}C-tracer analysis revealed multiple incorporation of labeled CO_2 into intermediates of the CBB cycle, indicating that the synthetic CO_2 fixation module is indeed functional and can turn multiple times. Proteomics analysis showed distinct changes in the proteome of *M. extorquens* AM1 expressing the functional CBB cycle, which might stabilize flux through the heterologous CO_2 fixation pathway.

While a disruption between methanol oxidation and serine cycle (Δ*ftfL*) can be partially rescued by the heterologous CBB cycle, this is not the case when the metabolic network that is involved in carbon assimilation of tetrahydrofolate-linked C1 units (H_4F-C1) via the serine cycle (Δ*glyA*) or the EMCP (Δ*ccr*) is disrupted. Most likely in the Δ*ftfL* strain the H_4F-C1 units required for biosynthetic pathways can be generated by reverse activity of serine hydroxymethyltransferase (encoded by *glyA*) or by operation of the glycine cleavage complex. In contrast, the *glyA* deletion strain would require a horse shoe-like operation of the serine cycle together with the glycine cleavage complex to generate serine, glycine and H_4F-C1 units. Although predicted to be viable with a carbon fixation pathway by *in silico* analysis, this metabolic mode is apparently not supported by a *glyA* deletion strain expressing a heterologous CBB cycle. Similarly, interruption of glyoxylate regeneration caused by deletion of *ccr* cannot be compensated by the heterologous CBB cycle, probably because the heterologous CBB cycle cannot provide sufficient precursors for synthesis of glyoxylate and subsequently of glycine and serine.

Despite the promising results obtained in this study, no continuous growth of strains engineered with the CBB cycle could be observed on methanol and CO_2 thus far. A major challenge in establishing synthetic metabolic cycles is to delicately balance the continuous operation of a cycle with the requirement to constantly drain metabolites from the cycle into biosynthetic pathways (Barenholz et al., 2017). Too little flux into biosynthesis will slow down growth, while too much flux into biosynthetic routes will inevitably bring the cycle to a hold. These problems were also identified in recent efforts that aimed at establishing artificial methylotrophy in *E. coli* (Müller et al., 2015), as well as implementing a heterologous CBB cycle for sugar synthesis from CO_2 in *E. coli* (Antonovsky et al., 2016). For the latter, it was demonstrated that a reduced catalytic capacity of enzymes that drain intermediates from the CBB cycle into biosynthesis was a key factor in establishing a stably operating heterologous CBB cycle in *E. coli* (Herz et al., 2017). The same challenge of balancing metabolic fluxes exists in the case of engineering a heterologous CBB cycle in *M. extorquens* AM1. A sufficient pool of ribulose 1,5-bisphosphate needs to be maintained to ensure continuous carbon fixation, while at the same time precursors need to be drained for biomass formation (see Supplementary Figure S7). Along these lines, it is interesting to note that we observed a specific upregulation of enzymes that increase flux and a specific downregulation of enzymes that drain intermediates from the CBB cycle upon expression of Prk and RuBisCO in *M. extorquens* AM1. However, these changes in enzyme expression levels do not seem to be sufficient to sustain a long-lasting operation of the heterologous carbon fixation pathway. An obvious next step towards sustained

growth of our engineered organism would be to improve flux through the synthetic CBB cycle by rational approaches as well as by directed evolution, in a similar way as described (Antonovsky et al., 2016), with a special focus on mutations that decrease the catalytic capacity of enzymes that affect the efflux of CBB cycle metabolites from the autocatalytic carbon fixation pathway towards biosynthesis (Herz et al., 2017).

In summary, our study represents another step forward in the generation of synthetic autotrophic organisms by rational metabolic engineering. Using *M. extorquens* AM1, a platform organism for a future C1 bioindustry, we could demonstrate that it is in principle possible to completely separate energy metabolism and carbon assimilation in this bacterium to create a synthetic organoautotroph that can use the C1 carbon source methanol for energy acquisition, but is dependent on CO_2 for the formation of biomass. Our results open the possibility to further improve synthetic autotrophy in *M. extorquens* AM1. At the same time, they lay the foundation for further explorations of the modularity of our approach. For instance, instead of the CBB cycle a non-natural carbon fixation pathway (Bar-Even et al., 2010; Schwander et al., 2016) could be introduced into *M. extorquens* AM1. Furthermore, replacing methanol as energy source with hydrogen, light or electricity could allow to create versatile synthetic autotrophs for biotechnology in the future.

Acknowledgements

We are grateful to Philipp Christen for skillful assistance with LC-MS sample measurement, to Andreas Kautz for help with running bioreactors and to Ryan Farmer and Bob Tabita for kindly providing purified RuBisCO from *R. rubrum* S 1. We are greatly indebted to Julia A. Vorholt for hosting the project in her laboratory at ETH Zurich during its initial stages as well as for helpful discussions.

Funding

This work was supported by ETH Zurich Grant ETH-41/12-2, an IMPRS fellowship for M.C., and the Max Planck Society.

References

Ahrné, E., Molzahn, L., Glatter, T., Schmidt, A., 2013. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. Proteomics. 13**,** 2567-78.

Anderson, D. J., Morris, C. J., Nunn, D. N., Anthony, C., Lidstrom, M. E., 1990. Nucleotide sequence of the *Methylobacterium extorquens* AM1 *moxF* and *moxJ* genes involved in methanol oxidation. Gene. 90**,** 173-6.

Antonovsky, N., Gleizer, S., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D. G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., Milo, R., 2016. Sugar synthesis from CO_2 in *Escherichia coli*. Cell. 166**,** 115-25.

Bar-Even, A., Noor, E., Lewis, N. E., Milo, R., 2010. Design and analysis of synthetic carbon fixation pathways. Proc. Natl. Acad. Sci. USA. 107**,** 8889-94.

Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., Milo, R., 2017. Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. eLife. 6.

Berg, I. A., Kockelkorn, D., Ramos-Vera, W. H., Say, R. F., Zarzycki, J., Hugler, M., Alber, B. E., Fuchs, G., 2010. Autotrophic carbon fixation in archaea. Nat. Rev. Microbiol. 8**,** 447-60.

Bertau, M., Offermanns, H., Plass, L., Schmidt, F., Wernicke, H.-J., 2014. Methanol: The basic chemical and energy feedstock of the future. Springer-Verlag Berlin, Heidelberg.

Bosch, G., Skovran, E., Xia, Q., Wang, T., Taub, F., Miller, J. A., Lidstrom, M. E., Hackett, M., 2008. Comprehensive proteomics of *Methylobacterium extorquens* AM1 metabolism under single carbon and nonmethylotrophic conditions. Proteomics. 8**,** 3494-505.

Bowes, G., Ogren, W. L., Hageman, R. H., 1971. Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase. Biochem. Biophys. Res. Commun. 45**,** 716-22.

Bowien, B., 1977. D-Ribulose 1,5-bisphosphate carboxylase from *Paracoccus denitrificans*. FEMS Microbiol. Lett. 2**,** 263-266.

Bradford, M. M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72**,** 248-54.

Chistoserdova, L. V., Lidstrom, M. E., 1994. Genetics of the serine cycle in *Methylobacterium extorquens* AM1: cloning, sequence, mutation, and physiological effect of *glyA*, the gene for serine hydroxymethyltransferase. J. Bacteriol. 176**,** 6759-62.

Chistoserdova, L. V., Lidstrom, M. E., 1996. Molecular characterization of a chromosomal region involved in the oxidation of acetyl-CoA to glyoxylate in the isocitrate-lyase-negative methylotroph *Methylobacterium extorquens* AM1. Microbiology. 142**,** 1459-68.

Choi, Y. J., Morel, L., Bourque, D., Mullick, A., Massie, B., Miguez, C. B., 2006. Bestowing inducibility on the cloned methanol dehydrogenase promoter (*P*mxaF) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1. Appl. Environ. Microbiol. 72**,** 7723-9.

Chubiz, L. M., Purswani, J., Carroll, S. M., Marx, C. J., 2013. A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*. BMC Res. Notes. 6**,** 183.

Cleland, W. W., Andrews, T. J., Gutteridge, S., Hartman, F. C., Lorimer, G. H., 1998. Mechanism of RuBisCO: The carbamate as general base. Chem. Rev. 98**,** 549-562.

Crowther, G. J., Kosaly, G., Lidstrom, M. E., 2008. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. J. Bacteriol. 190**,** 5057-62.

Delaney, N. F., Kaczmarek, M. E., Ward, L. M., Swanson, P. K., Lee, M. C., Marx, C. J., 2013. Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. PloS one. 8**,** e62957.

Dellomonaco, C., Clomburg, J. M., Miller, E. N., Gonzalez, R., 2011. Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. Nature. 476**,** 355-9.

Enquist-Newman, M., Faust, A. M., Bravo, D. D., Santos, C. N., Raisner, R. M., Hanel, A., Sarvabhowman, P., Le, C., Regitsky, D. D., Cooper, S. R., Peereboom, L., Clark, A., Martinez, Y., Goldsmith, J., Cho, M. Y., Donohoue, P. D., Luo, L., Lamberson, B., Tamrakar, P., Kim, E. J., Villari, J. L., Gill, A., Tripathi, S. A., Karamchedu, P., Paredes, C. J., Rajgarhia, V., Kotlar, H. K., Bailey, R. B., Miller, D. J., Ohler, N. L., Swimmer, C., Yoshikuni, Y., 2014. Efficient ethanol production from brown macroalgae sugars by a synthetic yeast platform. Nature. 505**,** 239-43.

Erb, T. J., Berg, I. A., Brecht, V., Müller, M., Fuchs, G., Alber, B. E., 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. Proc. Natl. Acad. Sci. USA. 104**,** 10631-6.

Fargione, J., Hill, J., Tilman, D., Polasky, S., Hawthorne, P., 2008. Land clearing and the biofuel carbon debt. Science. 319**,** 1235-8.

Fuchs, G., 2011. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu. Rev. Microbiol. 65**,** 631-58.

Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., Smolke, C. D., 2015. Complete biosynthesis of opioids in yeast. Science. 349**,** 1095-100.

Glatter, T., Ahrné, E., Schmidt, A., 2015. Comparison of different sample preparation protocols reveals lysis buffer-specific extraction biases in gram-negative bacteria and human cells. J. Proteome Res. 14**,** 4472-85.

Herz, E., Antonovsky, N., Bar-On, Y., Davidi, D., Gleizer, S., Prywes, N., Noda-Garcia, L., Lyn Frisch, K., Zohar, Y., Wernick, D. G., Savidor, A., Barenholz, U., Milo, R., 2017. The genetic basis for the adaptation of *E. coli* to sugar synthesis from CO_2. Nat. Commun. 8**,** 1705.

Hofer, P., Choi, Y. J., Osborne, M. J., Miguez, C. B., Vermette, P., Groleau, D., 2010. Production of functionalized polyhydroxyalkanoates by genetically modified *Methylobacterium extorquens* strains. Microb. Cell Fact. 9**,** 70.

Hu, B., Lidstrom, M. E., 2014. Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production. Biotechnol. Biofuels. 7**,** 156.

Hudson, G. S., Morell, M. K., Arvidsson, Y. B. C., Andrews, T. J., 1992. Synthesis of spinach phosphoribulokinase and ribulose 1,5-bisphosphate in *Escherichia coli*. Aust. J. Plant Physiol. 19**,** 213-221.

Kalyuzhnaya, M. G., Lidstrom, M. E., 2003. QscR, a LysR-type transcriptional regulator and CbbR homolog, is involved in regulation of the serine cycle genes in *Methylobacterium extorquens* AM1. J. Bacteriol. 185**,** 1229-35.

Kiefer, P., Schmitt, U., Vorholt, J. A., 2013. eMZed: an open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. Bioinformatics. 29**,** 963-4.

Kobayashi, D., Tamoi, M., Iwaki, T., Shigeoka, S., Wadano, A., 2003. Molecular characterization and redox regulation of phosphoribulokinase from the cyanobacterium *Synechococcus* sp. PCC 7942. Plant Cell Physiol. 44**,** 269-76.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227**,** 680-5.

Large, P. J., Peel, D., Quayle, J. R., 1961. Microbial growth on C1 compounds. 2. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM 1, and methanol-grown *Hyphomicrobium vulgare*. The Biochemical journal. 81**,** 470-80.

Liu, C., Colon, B. C., Ziesack, M., Silver, P. A., Nocera, D. G., 2016. Water splitting-biosynthetic system with CO_2 reduction efficiencies exceeding photosynthesis. Science. 352**,** 1210-3.

Marx, C. J., 2008. Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. BMC Res. Notes. 1**,** 1.

Marx, C. J., Laukel, M., Vorholt, J. A., Lidstrom, M. E., 2003. Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. J. Bacteriol. 185**,** 7169-75.

Marx, C. J., Lidstrom, M. E., 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other gram-negative bacteria. Microbiology. 147**,** 2065-75.

Marx, C. J., Lidstrom, M. E., 2002. Broad-host-range *cre-lox* system for antibiotic marker recycling in gram-negative bacteria. BioTechniques. 33**,** 1062-7.

Marx, C. J., Lidstrom, M. E., 2004. Development of an insertional expression vector system for *Methylobacterium extorquens* AM1 and generation of null mutants lacking *mtdA* and/or *fch*. Microbiology. 150**,** 9-19.

Mattozzi, M. D., Ziesack, M., Voges, M. J., Silver, P. A., Way, J. C., 2013. Expression of the sub-pathways of the *Chloroflexus aurantiacus* 3-hydroxypropionate carbon fixation bicycle in *E. coli*: Toward horizontal transfer of autotrophic growth. Metab. Eng. 16**,** 130-9.

Mueller-Cajar, O., Morell, M., Whitney, S. M., 2007. Directed evolution of RuBisCO in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. Biochemistry. 46**,** 14067-74.

Müller, J. E. N., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., Quax, W. J., Wendisch, V. F., Brautaset, T., Portais, J. C., Vorholt, J. A., 2015. Engineering *Escherichia coli* for methanol conversion. Metab. Eng. 28**,** 190-201.

Nichols, E. M., Gallagher, J. J., Liu, C., Su, Y., Resasco, J., Yu, Y., Sun, Y., Yang, P., Chang, M. C., Chang, C. J., 2015. Hybrid bioinorganic approach to solar-to-chemical conversion. Proc. Natl. Acad. Sci. USA. 112**,** 11461-6.

Ochsner, A. M., Christen, M., Hemmerle, L., Peyraud, R., Christen, B., Vorholt, J. A., 2017. Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy. Current Biology. 27**,** 2579-2588.

Ochsner, A. M., Sonntag, F., Buchhaupt, M., Schrader, J., Vorholt, J. A., 2015. *Methylobacterium extorquens*: Methylotrophy and biotechnological applications. Appl. Microbiol. Biotechnol. 99**,** 517-34.

Okubo, Y., Skovran, E., Guo, X., Sivam, D., Lidstrom, M. E., 2007. Implementation of microarrays for *Methylobacterium extorquens* AM1. Omics : a journal of integrative biology. 11**,** 325-40.

Olah, G. A., 2013. Towards oil independence through renewable methanol chemistry. Angew. Chem. Int. Ed. 52**,** 104-7.

Orita, I., Nishikawa, K., Nakamura, S., Fukui, T., 2014. Biosynthesis of polyhydroxyalkanoate copolymers from methanol by *Methylobacterium extorquens* AM1 and the engineered strains under cobalt-deficient conditions. Appl. Microbiol. Biotechnol. 98**,** 3715-25.

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., Newman, J. D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 496**,** 528-32.

Peel, D., Quayle, J. R., 1961. Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM 1. Biochem. J. 81**,** 465-9.

Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., Vorholt, J. A., 2009. Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. Proc. Natl. Acad. Sci. USA. 106**,** 4846-51.

Peyraud, R., Schneider, K., Kiefer, P., Massou, S., Vorholt, J. A., Portais, J. C., 2011. Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1. BMC Syst. Biol. 5**,** 189.

Racker, E., 1957. The reductive pentose phosphate cycle. I. Phosphoribulokinase and ribulose diphosphate carboxylase. Arch. Biochem. Biophys. 69**,** 300-10.

Racker, E., 1962. Ribulose diphosphate carboxylase from spinach leaves. Method. Enzymol. 5**,** 266-270.

Rigbolt, K. T., Vanselow, J. T., Blagoev, B., 2011. GProX, a user-friendly platform for bioinformatics analysis and visualization of quantitative proteomics data. Mol. Cell. Proteomics. 10**,** O110.007450.

Sambrook, J., Russell, D. W., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., Erb, T. J., 2015. A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other Alphaproteobacteria. ACS Synth. Biol. 4**,** 430-43.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., Pühler, A., 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: Selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene. 145**,** 69-73.

Schloss, J. V., Phares, E. F., Long, M. V., Norton, I. L., Stringer, C. D., Hartman, F. C., 1979. Isolation, characterization, and crystallization of ribulose bisphosphate carboxylase from autotrophically grown *Rhodospirillum rubrum*. J. Bacteriol. 137**,** 490-501.

Schneider, K., Peyraud, R., Kiefer, P., Christen, P., Delmotte, N., Massou, S., Portais, J. C., Vorholt, J. A., 2012. The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by *Methylobacterium extorquens* AM1 during growth on acetate. J. Biol. Chem. 287**,** 757-66.

Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M. V., Marx, A., Vorholt, J. A., 2009. Methanol-based industrial biotechnology: Current status and future perspectives of methylotrophic bacteria. Trends Biotechnol. 27**,** 107-15.

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

Sims, R. E., Mabee, W., Saddler, J. N., Taylor, M., 2010. An overview of second generation biofuel technologies. Bioresour. Technol. 101**,** 1570-80.

Sonntag, F., Buchhaupt, M., Schrader, J., 2014. Thioesterases for ethylmalonyl-CoA pathway-derived dicarboxylic acid production in *Methylobacterium extorquens* AM1. Appl. Microbiol. Biotechnol. 98**,** 4533-44.

Sonntag, F., Kroner, C., Lubuta, P., Peyraud, R., Horst, A., Buchhaupt, M., Schrader, J., 2015a. Engineering *Methylobacterium extorquens* for *de novo* synthesis of the sesquiterpenoid alpha-humulene from methanol. Metab. Eng. 32**,** 82-94.

Sonntag, F., Müller, J. E. N., Kiefer, P., Vorholt, J. A., Schrader, J., Buchhaupt, M., 2015b. High-level production of ethylmalonyl-CoA pathway-derived dicarboxylic acids by *Methylobacterium extorquens* under cobalt-deficient conditions and by polyhydroxybutyrate negative strains. Appl. Microbiol. Biotechnol. 99**,** 3407-19.

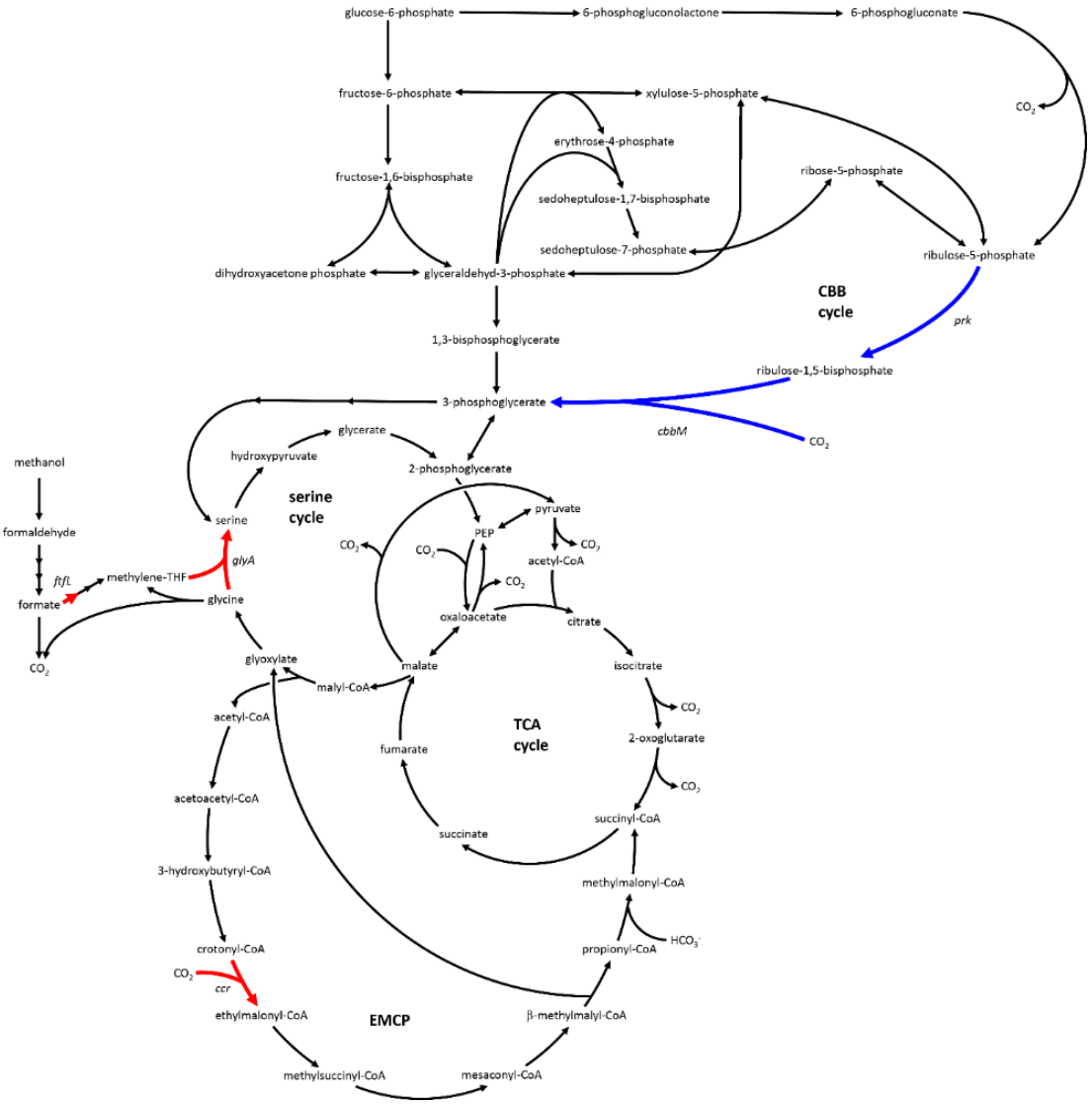
Strovas, T. J., Lidstrom, M. E., 2009. Population heterogeneity in *Methylobacterium extorquens* AM1. Microbiology. 155**,** 2040-8.

Strovas, T. J., Sauter, L. M., Guo, X., Lidstrom, M. E., 2007. Cell-to-cell heterogeneity in growth rate and gene expression in *Methylobacterium extorquens* AM1. J. Bacteriol. 189**,** 7127-33.

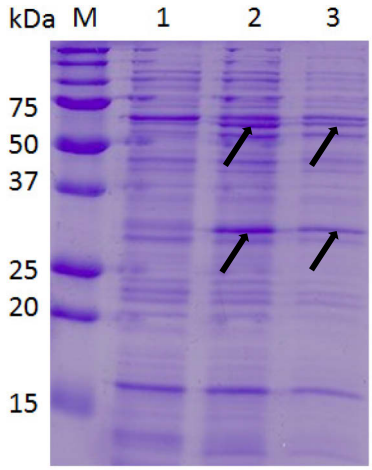
Toyama, H., Anthony, C., Lidstrom, M. E., 1998. Construction of insertion and deletion *mxa* mutants of *Methylobacterium extorquens* AM1 by electroporation. FEMS Microbiol. Lett. 166**,** 1-7.

Zhuang, Z. Y., Li, S. Y., 2013. RuBisCO-based engineered *Escherichia coli* for *in situ* carbon dioxide recycling. Bioresour. Technol. 150**,** 79-88.

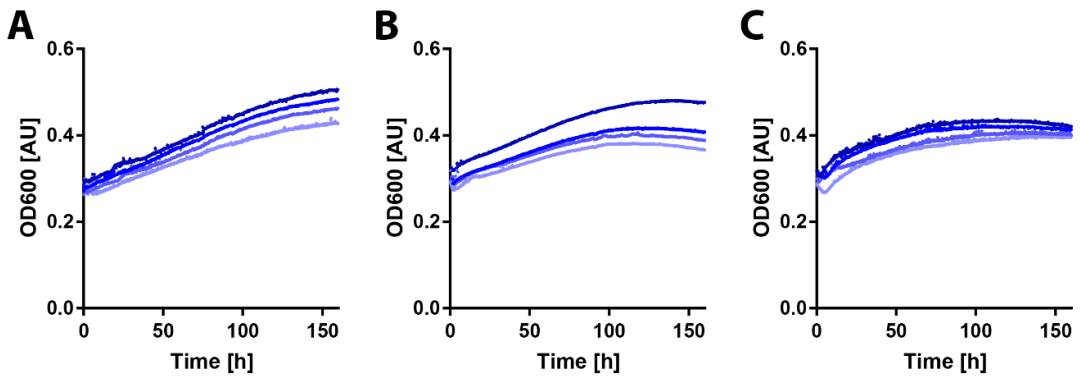
Supporting Information



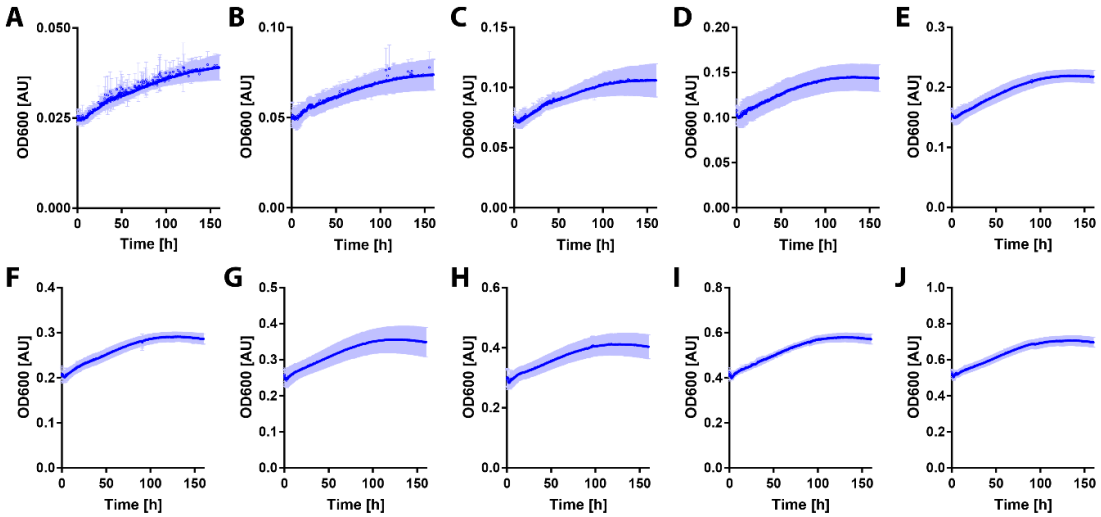
Supplementary Figure S1: Central carbon metabolism of *M. extorquens* AM1 with a heterologous Calvin-Benson-Bassham (CBB) cycle. The three reactions that were abolished by deletion of the respective genes *glyA*, *ccr* or *ftfL* are indicated by red arrows, while the reactions that are catalyzed by the enzymes phosphoribulokinase (encoded by *prk*) and RuBisCO (encoded by *cbbM*) are indicated by blue arrows.



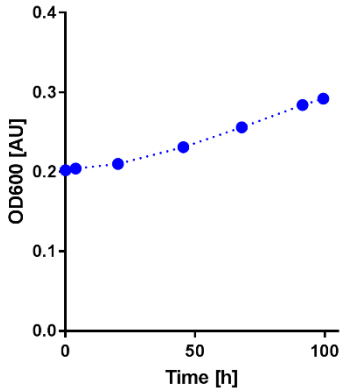
Supplementary Figure S2: SDS-PAGE gel demonstrating the overexpression and solubility of RuBisCO from *R. rubrum* S 1 and Prk from *M. extorquens* AM1 in *M. extorquens* AM1. Whole-cell lysate of *M. extorquens* AM1 transformed with pCM80, containing 20 ug total protein, was loaded in lane 1 as a negative control. Whole-cell lysate of *M. extorquens* AM1 transformed with pTE94, containing 20 ug total protein, was loaded in lane 2. The soluble fraction of cell lysate of *M. extorquens* AM1 transformed with pTE94, containing 20 ug total protein, was loaded in lane 3. RuBisCO (52 kDa) is indicated by the upper arrow in lane 2 and 3 and Prk (34 kDa) is indicated by the lower arrow in lane 2 and 3.



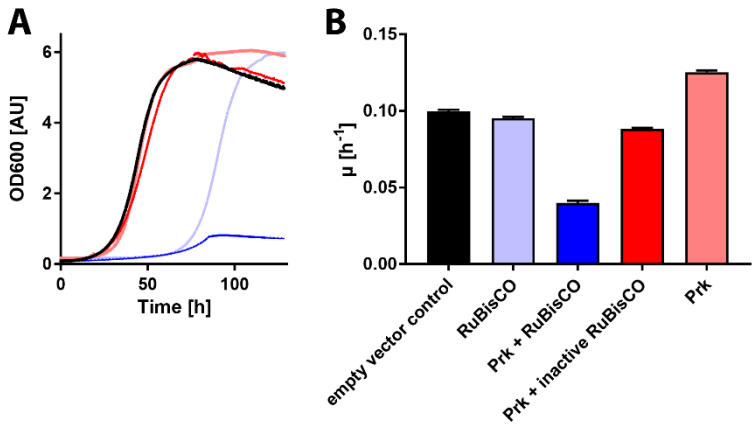
Supplementary Figure S3: Growth curves of *M. extorquens* AM1 Δ*cel* Δ*ftfL* overexpressing RuBisCO and Prk on methanol in a 5% CO_2 atmosphere. Shown are three individual experiments inoculated from different precultures (= biological replicates; A, B, C). Cultures were inoculated to an initial OD_{600} of 0.3, and growth was monitored in 96-well plates over the course of 150 h. Four individual growth curves from different wells (= technical replicates) are shown per experiment, colored in different shades of blue.



Supplementary Figure S4: Growth curves of *M. extorquens* AM1 Δ*cel* Δ*ftfL* overexpressing RuBisCO and Prk on methanol in a 5% CO_2 atmosphere. Cultures were inoculated to different initial OD_{600} values, and growth was monitored in 96-well plates over the course of 150 h. The initial OD_{600} values were 0.025 (A), 0.05 (B), 0.075 (C), 0.1 (D), 0.15 (E), 0.2 (F), 0.25 (G), 0.3 (H), 0.4 (I) and 0.5 (J). The average curve of at least six technical replicates is shown in blue, with the respective 95% confidence interval in light blue.



Supplementary Figure S5: Growth curve of *M. extorquens* AM1 Δ*cel* Δ*ftfL* overexpressing RuBisCO and Prk in a bioreactor on minimal medium containing 123 mM methanol with constant gassing of 5% CO2 in air. Samples were taken at the indicated timepoints and OD600 was measured manually on a photospectrometer.



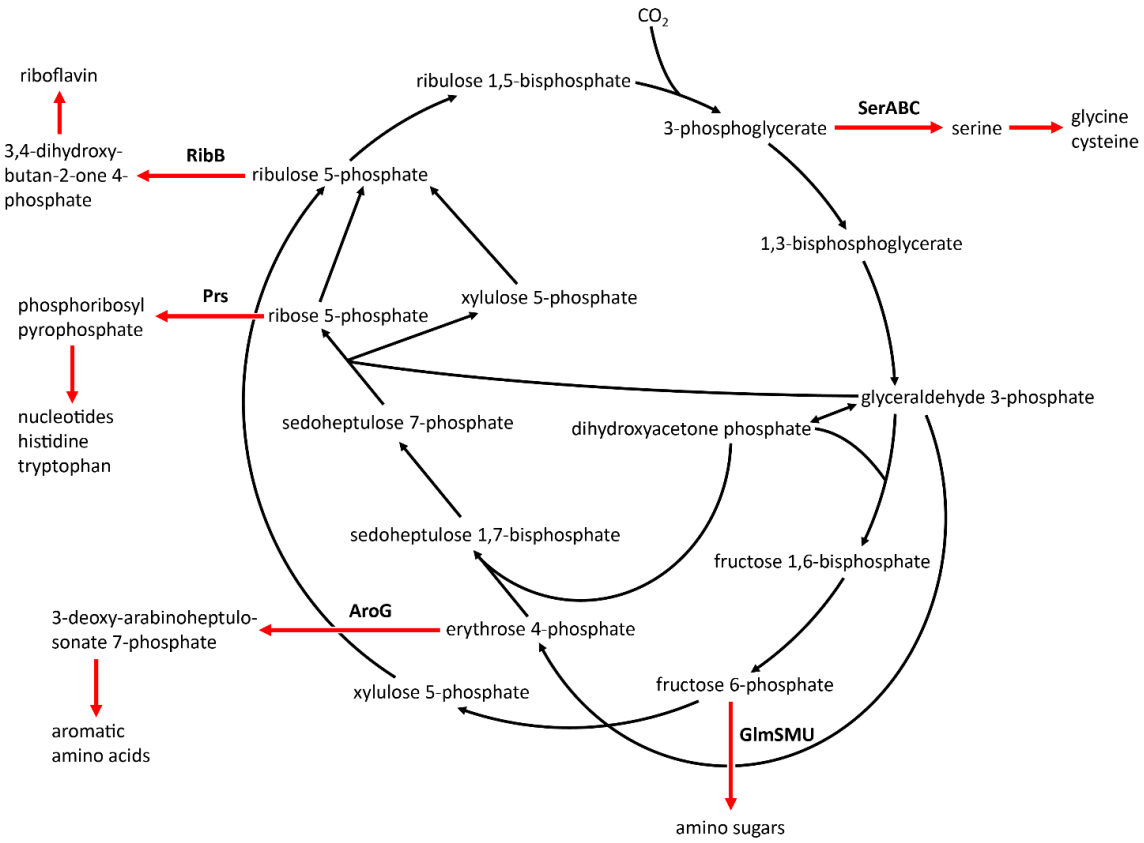
Supplementary Figure S6: Growth behaviour of *M. extorquens* AM1 Δ*cel* with different plasmids for the overexpression of Prk + RuBisCO (blue), Prk + inactive RuBisCO (red), RuBisCO alone (light blue), Prk alone (light red) as well as an empty vector control (black) on methanol in a 5% CO2 atmosphere. Shown are representative growth curves (A) and the average values for specific growth rates (B), calculated from at least 20 technical replicates. Error bars represent standard deviation.

6.1 Supplementary Note 1

When growing the Δ*cel* strain transformed with different plasmids on methanol in an atmosphere containing 5% CO_2, it becomes evident that without the overexpression of any proteins, the growth rate (μ = 0.10 h^{-1}) is significantly lower than the growth rate of *M. extorquens* AM1 WT on methanol (μ = 0.16 h^{-1}) (Kiefer et al., 2009). This effect can be rationalized by the excess amount of carbon dioxide both in the gas and the liquid phase, which is apparently detrimental to growth of *M. extorquens* AM1.

When overexpressing RuBisCO alone, the growth rate (μ = 0.095 h^{-1}) is very close to the empty plasmid control strain, most likely due to lack of a high amount of ribulose 1,5-bisphosphate in the cell, meaning that the substrate of RuBisCO is only available in low concentrations and no effect of the heterologous enzyme activity can be seen. However, upon overexpression of both RuBisCO and Prk, both growth rate (μ = 0.04 h^{-1}) and final OD_{600} of the strain are strongly decreased (see Supplementary Figure S6), suggesting that the possibility to operate the CBB cycle considerably perturbs methylotrophic metabolism of *M. extorquens* AM1.

It is further confirmed that both functional RuBisCO and Prk are required for this effect, since the growth rate of the Δ*cel* strain overexpressing inactive RuBisCO and Prk (μ = 0.088 h^{-1}) is only slightly lower than the growth rate of the empty plasmid control strain, and final OD_{600} is essentially unchanged. Finally, overexpression of Prk alone has a positive effect on growth rate (μ = 0.125 h^{-1}). In this case, accumulation of ribulose 1,5-bisphosphate, which is a key regulatory metabolite in methylotrophy (Ochsner et al., 2017), seems to be beneficial for optimal growth of *M. extorquens* AM1 on methanol and 5% CO_2.



Supplementary Figure S7: Branching points for biosynthesis from the CBB cycle. Reactions of the CBB cycle are denoted by black arrows, while reactions of biosynthetic pathways are denoted by red arrows. The enzymes that convert metabolites of the CBB cycle into biomass precursors are denoted next to the arrows.

Supplementary Table S1: Primers used in this study

target	name	sequence[a]	restriction site
Mext PRK	M_PRK_fw	5‘-CTC**AAGCTT**GAAGGGAGAGAAGCCCATGGTC-3‘	HindIII
Mext PRK	M_PRK_rv	5‘-GA**GGATCC**GAGGCTCACGCCAG-3‘	BamHI
Pden PRK	P_PRK_fw	5‘-GCGGCCA**CATATG**AGCAAGAAACACCCGATCATTTCC-3‘	NdeI
Pden PRK	P_PRK_rv	5‘-GCGGACT**GAATTC**TCAGGCGCGTTTCGATTCGC-3‘	EcoRI
Syn PRK	S_PRK_fw	5‘-GAGCGCT**AAGCTT**GAAGGGAGAGAAGCCCATGAGCAAGCCAG-3‘	HindIII
Syn PRK	S_PRK_rv	5‘-GGCGGCA**AGATCT**CTATACGCTCGCGGCGAC-3‘	BglII
Rrub RubisCO	R_RBC_fw	5‘-GAC**GGATTC**ATCTAGGGAGAGTCCACCGATGGACCAG-3‘	BamHI
Rrub RubisCO	R_RBC_rv	5‘-CT**GAATTC**AGGCGGTTACGCCGGAAGGG-3‘	EcoRI
Pden RBC LSU	P_RBCL_fw	5‘-GCGGCCA**CATATG**AACGAGATGAGCAAATCCGAAATCAC-3‘	NdeI
Pden RBC LSU	P_RBCL_rv	5‘-GGCGACT**GAATTC**TTAGCTGACGGACGCGGTCGGAAC-3‘	EcoRI
Pden RBC SSU	P_RBCS_fw	5‘-GCGGCCA**CATATG**CGTATCACTCAGGGCTGCTTTTCC-3‘	NdeI
Pden RBC SSU	P_RBCS_rv	5‘-GGCGACT**GAATTC**TCAGCCGACCAGCTCATGCGTATAG-3‘	EcoRI
Rrub RubisCO	K191M_QC_fw	5‘-GCGGCGACTTCATCATGAACGACGAGCCC-3‘	-
Rrub RubisCO	K191M_QC_rv	5‘-GGGCTCGTCGTTCATGATGAAGTCGCCGCCCAG-3‘	-
glyA_up	glyA_up_fw	5‘-GTTAT**GAATTC**GCGCCAAGCGCGTGC-3‘	EcoRI
glyA_up	glyA_up_rv	5‘-TCAGGCGAGCGTAGCCTCTTTTCGTTG-3‘	-
glyA_down	glyA_do_fw	5‘-CGTAAACAACGAAAAGAGGCTACGCTCGCCTGACTAAAAACACCTCCG-3‘	-
glyA_down	glyA_do_rv	5‘-GATGGCC**TCTAGA**CCCAGACGCTCTCAAAAG-3‘	XbaI

[a] Nucleotides in bold and underlined are recognition sites for endonuclease restriction enzymes.

Supplementary Table S2: Strains and plasmids used in this study

strain or plasmid	genotype or relevant features[a]	source or ref
E. coli DH5	supE44, ΔlacU169 (Φ80lacZDM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1	Life Technologies
M. extorquens AM1	wild-type	Peel and Quayle (1961)
M. extorquens AM1 61ADH	Δ*ccr*; Km^R	Chistoserdova and Lidstrom (1996)
M. extorquens AM1 CM216K.1	Δ*ftfL*; Km^R	Marx et al. (2003)
M. extorquens AM1 CM2720	Δ*cel*	Delaney et al. (2013)
M. extorquens AM1 LS1	Δ*glyA*	this work
M. extorquens AM1 LS2	Δ*cel* Δ*ftfL*; Km^R	this work
pET28b	*E. coli* expression vector, T7 promoter, Km^R	Novagen (Merck)
pCM80	*M. extorquens* AM1 expression vector, pCM62 derivative, *mxaF* promoter; Tc^R	Marx and Lidstrom (2001)
pTE100	*M. extorquens* AM1 expression vector, pCM80 derivative, promoterless; Tc^R	Schada von Borzyskowski et al. (2015)
pTE102	*M. extorquens* AM1 expression vector, pCM100 derivative, *mxaF* promoter; Tc^R	Schada von Borzyskowski et al. (2015)
pTE92	pCM80-based, Rrub RubisCO, Tc^R	this work
pTE94	pCM80-based, Mext PRK and Rrub RubisCO, Tc^R	this work
pTE95	pCM80-based, Mext PRK and Rrub RubisCO K191M, Tc^R	this work
pTE96	pCM80-based, Syn PRK and Rrub RubisCO, Tc^R	this work
pTE245	pET28b-based, Pden RubisCO LSU, Km^R	this work
pTE246	pET28b-based, Pden RubisCO SSU, Km^R	this work
pTE248	pET28b-based, Pden PRK, Km^R	this work
pTE500	pTE100-based, Pden RubisCO LSU, Tc^R	this work
pTE502	pTE100-based, Pden RubisCO SSU, Tc^R	this work
pTE504	pTE100-based, Pden PRK, Tc^R	this work
pTE519	pTE102-based, Pden RubisCO LSU, Tc^R	this work
pTE535	pTE102-based, Pden RubisCO LSU and Pden PRK, Tc^R	this work
pTE550	pTE102-based, Pden RubisCO LSU+SSU and Pden PRK, Tc^R	this work
pK18-mob-sacB	allelic exchange vector, *sacB* gene, Km^R	Schäfer et al. (1994)
pCM216	pCM184-based, *ftfL* flanking regions, Km^R, Tc^R, Amp^R	Marx et al. (2003)
pTE210	pK18-mob-sacB-based, *glyA* flanking regions, Km^R	this work

[a] Km^R, kanamycin resistance; Tc^R, tetracycline resistance; Amp^R, ampicillin resistance

Supplementary Table S3: Results of *in silico* analyses of the biomass yield of engineered strains of *M. extorquens* AM1 and their potential recovery due to implementation of the CBB cycle using Flux Balance Analysis (FBA). iRP911, a stoichiometric metabolic network model of *M. extorquens* AM1 was used and the RuBisCO reaction was manually added to close the CBB cycle. To realize the KO mutants the lower and upper bound of the flux through the corresponding reactions was set to 0.

Parameters: **Objective function: Biomass synthesis** **Uptake of organic substances = 0, except for methanol = 15.1 mmol g_{DW}^{-1} h^{-1}** **NGAM = 9.5 mmol g_{DW}^{-1} h^{-1}; R0002 = 0**			
Strain	**Modified bounds for respective reactions**	**$Y_{BM/MeOH}$ [g_{BM} g_{MeOH}^{-1}]**	**% of WT $Y_{BM/MeOH}$**
WT	RuBisCO = 0	0.444	100%
WT + RuBisCO	---	0.444	100%
Δ*glyA*	R0015 = 0; RuBisCO = 0	0	0%
Δ*glyA* + RuBisCO	R0015 = 0	0.390	88%
Δ*ftfL*	R0012 = 0; RuBisCO = 0	0.320	72%
Δ*ftfL*, no glycine cleavage system (Δ*gcvH/P/T)*	R0012 = 0; R0243/R0244/R0245 = 0; RuBisCO = 0	0	0%
Δ*ftfL*, no glycine cleavage system (Δ*gcvH/P/T)* + RuBisCO	R0012 = 0; R0243/R0244/R0245 = 0	0.388	87%
no glycine cleavage system (Δ*gcvH/P/T)*	R0243/R0244/R0245 = 0; RuBisCO = 0	0.444	100%
no glycine cleavage system (Δ*gcvH/P/T)* + RuBisCO	R0243/R0244/R0245 = 0	0.444	100%
Δ*ccr*	R0032 = 0; RuBisCO = 0	0.437	98%
Δ*ccr,* no butyryl-CoA carboxylase	R0032 = 0; R034 = 0; RuBisCO = 0	0	0%
Δ*ccr,* no butyryl-CoA carboxylase + RuBisCO	R0032 = 0; R034 = 0	0	0%
no butyryl-CoA carboxylase	R034 = 0; RuBisCO = 0	0.444	100%
no butyryl-CoA carboxylase + RuBisCO	R034 = 0	0.444	100%

References for Supporting Information

Chistoserdova, L. V., Lidstrom, M. E., 1996. Molecular characterization of a chromosomal region involved in the oxidation of acetyl-CoA to glyoxylate in the isocitrate-lyase-negative methylotroph *Methylobacterium extorquens* AM1. Microbiology. 142**,** 1459-68.

Delaney, N. F., Kaczmarek, M. E., Ward, L. M., Swanson, P. K., Lee, M. C., Marx, C. J., 2013. Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. PloS one. 8**,** e62957.

Kiefer, P., Buchhaupt, M., Christen, P., Kaup, B., Schrader, J., Vorholt, J. A., 2009. Metabolite profiling uncovers plasmid-induced cobalt limitation under methylotrophic growth conditions. PloS one. 4**,** e7831.

Marx, C. J., Laukel, M., Vorholt, J. A., Lidstrom, M. E., 2003. Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. J. Bacteriol. 185**,** 7169-75.

Marx, C. J., Lidstrom, M. E., 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other gram-negative bacteria. Microbiology. 147**,** 2065-75.

Ochsner, A. M., Christen, M., Hemmerle, L., Peyraud, R., Christen, B., Vorholt, J. A., 2017. Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy. Curr. Biol. 27**,** 2579-2588.

Peel, D., Quayle, J. R., 1961. Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM 1. Biochem. J. 81**,** 465-9.

Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., Erb, T. J., 2015. A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other Alphaproteobacteria. ACS Synth. Biol. 4**,** 430-43.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., Pühler, A., 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: Selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene. 145**,** 69-73.

CHAPTER III

Engineering and evolving a CO_2 fixing ribose metabolism in *Methylorubrum extorquens* AM1‡

‡Manuscript draft for submission.

Engineering and evolving a CO_2 fixing ribose metabolism in *Methylorubrum extorquens* AM1

Martina Carrillo[a], Luise Peeck[a], Melanie Klose[a], Vincent Hervé[b], Andreas Brune[b], Tobias J. Erb[a,c]

[a]Department of Biochemistry and Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany
[b]Research Group Insect Gut Microbiology and Symbiosis, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany
[c]SYNMIKRO, LOEWE Center for Synthetic Microbiology, Universität Marburg, 35043 Marburg, Germany

Manuscript draft for submission.

Author contributions

T.J.E. conceived the project; M. C. and T. J. E. designed the research and wrote the manuscript with input from all authors; M. C., L. P., and M. K. performed experiments, provided and analyzed data. V. H. performed hybrid genome assemblies.

Abstract

Methylorubrum extorquens AM1 is a heterotrophic organism capable of utilizing methanol and formate (C_1) and various organic acids (C_2-C_4), amongst others, as carbon sources. Yet, it does not grow on C_5 sugars. In this study, we sought to expand the metabolic network of *M. extorquens* AM1 to include C_5 sugars, *i.e.* ribose, as carbon source. This goal was achieved first by implementation of a ribose assimilation module in a rational engineering approach followed by adaptive laboratory evolution to obtain robust growth on C_5 sugars. Finally, our strains were further engineering towards a carbon-positive ribose metabolism by implementation of the Calvin Benson Bassham cycle for co-fixation of CO_2.

1. Introduction

Methylorubrum extorquens AM1 (formerly *Methylobacterium*) is a plant-associated bacterium with a versatile metabolism and most well-known for its ability to grow on C_1 compounds such as formate, methanol, and methylamine as sole carbon and energy sources (Peel and Quayle, 1961). Several bulk and value-added chemicals have been produced in *M. extorquens* including the sesquisterpenoid α-humulene (Sonntag et al., 2015), 3-hydroxypropionate (Yang et al., 2017), and 1-butanol (Hu and Lidstrom, 2014). The central carbon metabolism of *M. extorquens* features unique pathways and metabolites, such as the ethylmalonyl-CoA pathway (EMCP), which make this organism an interesting target for the production of carotenoids, mono- and di-carboxylic acids, as well as the implementation of synthetic CO_2 fixation cycles (Schwander et al., 2016), compared to more canonical model organisms, such as *E. coli* and *Pichia pastoris*.

To expand the biotechnological potential of *M. extorquens* AM1, there have been several efforts to radically engineer the organism's central carbon metabolism. A functional glyoxylate cycle has been recently introduced, which allowed to drain intermediates from the EMCP during growth on acetate as carbon source (Schada von Borzyskowski et al., 2018b). Introduction of a Calvin Benson Bassham (CBB) module resulted in a positive growth phenotype on methanol that was linked to an active and fully functional RubisCO (Schada von Borzyskowski et al., 2018a). Altogether, these efforts are first steps to create customized microbial platforms for different applications. Here we sought to further expand on these efforts by converting *M. extorquens* AM1 into an organism that is able to use C_5 sugars as carbon source.

We conceived a strategy to achieve this goal in several steps by (i) introducing a C_5 sugar assimilation module into *M. extorquens* AM1, (ii) adaptive laboratory evolution (ALE) to create a strain with robust growth on C_5 sugars, and (iii) further engineering the strain towards co-fixation of CO_2 to create a carbon-positive ribose metabolism.

2. Materials and Methods

2.1. Strains and cultivation conditions

M. extorquens AM1 strains (see Table **S1**) were routinely grown at 30°C in minimal medium (Peyraud et al., 2009) supplemented with 123 mM methanol, 30.8 mM sodium succinate, 10 mM sodium acetate, or 0-10 mM D-ribose (Sigma Aldrich R9629)**,** as stated. Plasmids (see Table **S2**) were transferred into *M*. *extorquens* by electroporation (Toyama et al., 1998). Antibiotics for selection purposes were used accordingly: kanamycin 35 for *M. extorquens* or 50 µg mL^{-1} for *E. coli*, gentamicin 15 µg mL^{-1} for *E. coli*, tetracycline 10 µg mL^{-1} for both, chloramphenicol 34 µg mL^{-1} for *E. coli*, ampicillin 100 µg mL^{-1} for *E. coli*. Tetracycline was prepared fresh in water to avoid contamination of ethanol as a carbon source. When necessary, *M. extorquens* was grown in a 5% CO_2 atmosphere using INFORS HT Minitron incubators (Bottmingen, Switzerland). *E. coli* was routinely grown at 37°C in LB medium (Miller formula), except for those with plasmids containing the *sacB* gene

which were grown in Lennox-LB. *E. coli* TOP10 and DH5α (Thermo Scientific™) cells were used for construction and amplification of all plasmids in this study. For solid medium 1.5% (w/v) select agar was added.

2.2. DNA manipulation and plasmid construction

Refer to Table **S2** for detailed information on the construction of each plasmid in the Supporting Information. Standard molecular cloning techniques were used for amplification, purification, cloning and transformation of DNA (Sambrook and Russel, 2001). T4 Polynucleotide Kinase, FastAP, and FastDigest restriction enzymes were obtained from Thermo Scientific and used according to manufacturer's instructions. Q5 DNA polymerase was obtained from New England Biolabs (NEB) and used according to manufacturer's instructions. DNA oligos were obtained from Eurofins Genomics and Sigma-Aldrich (see Table **S3**). Plasmid isolation and PCR product purification was performed with NucleoSpin® Plasmid and NucleoSpin® Gel and PCR Clean-up kits (Macherey Nagel), according to manufacturer's instructions.

2.3. Strain construction

2.3.1. *phaA*::0 knockout

The *phaA* (META1p3700) gene was knocked out by allelic exchange with a shortened peptide where the 17 first amino acids are fused to the last 18 amino acids thus keeping expression of downstream genes intact using pTE1839. The first recombination event was selected for on solid medium with succinate and tetracycline thus creating an intermediate strain confirmed by colony PCR (as described in Section 2.4) using primers MC315+MC310+OS14. The second recombination event occurred with counter selection of the intermediate strain on solid medium with succinate and 5% sucrose without tetracycline. Single colonies were screened for tetracycline sensitivity, inability to grow on C_1/C_2 carbon sources, and the shortened peptide *phaA*::0 was confirmed by colony PCR with primers MC315+MC165 and sequencing.

2.3.2. Truncated TobG_R200* allele

The *tobG* (META2p0012) gene was replaced by a truncated allele with a stop codon at the 200th codon triplet and a 363 bp deletion using pTE2712 in CM2720. The first recombination event was selected on solid methanol medium with tetracycline and confirmed by colony PCR using primers MC485+453. The second recombination event was counter selected on solid methanol medium with 5% sucrose and without tetracycline. Single colonies were screened for tetracycline sensitivity and truncated allele was confirmed by colony PCR with primers MC485+MC486 and sequencing yielding CM2720 TobG_R200*.

2.4. Colony PCR from *M. extorquens*

Colony PCR was carried out by suspending cells from a single colony in 35µL of a 10% (v/v) solution of Chelex®100 (Sigma C7901; 50-100 mesh) and boiling cell suspensions at 95°C for 15 min. Boiled cell lysates were cooled on ice, spun down, and 1.5-2µL was used as

template DNA. Phusion Green Hot Start II High-Fidelity PCR Master Mix (Thermo Scientific F566S) was used for amplification. Each reaction contained 1X Master mix, template DNA as described above, 1X Q5 GC enhancer (NEB), and primers (0.05 µM each) in a final volume of 40 µL per reaction. The thermocycling conditions were as follows: initial denaturation at 98°C for 5 min, then 26 cycles of: 98°C 30 sec, 55°C for 20 sec, and 72°C for 30 s/kb; and a final extension step at 72°C for 10 min.

2.5. Growth in 96-well plates

Cultures were inoculated from a single colony or from glycerol stocks into Erlenmeyer flasks with the appropriate medium as a preculture. Cells from late exponential phase were collected, washed once, and resuspended in fresh medium at an OD_{600} of 0.2. Nunclon Delta Surface 96-well plates (Thermo Fisher Scientific, Darmstadt, Germany) were filled with 180 µl of cell suspension in each well. OD_{600} was recorded every 30 min using a Tecan Infinite M200Pro (Tecan, Männedorf, Switzerland). The temperature was kept constant at 30°C and when necessary a 5% CO_2 atmosphere was maintained using the Tecan Gas Control Module (Tecan, Männedorf, Switzerland). Data was analyzed using GraphPad Prism 7.

2.6. Adaptive laboratory evolution by serial transfers

The parental strain CM2720 + pTE1154 (P_{coxB}-*rbsK-glf*), termed GL0, was grown in 100 mL of methanol medium as a preculture and was transferred to medium with 7.5 mM ribose at an OD_{600} ≈ 0.2. This main culture was split into six parallel cultures termed GL1-6 each with 100 mL of medium in 500 mL Erlenmeyer flasks. As previously mentioned, tetracycline used for plasmid selection was dissolved in water to avoid ethanol as a second carbon source. At regular intervals, culture was removed for OD_{600} measurement as well as supernatant samples for HPLC. Once the cultures reached OD_{600} ≈ 0.8-1.0, they were transferred to fresh ribose medium to an OD ≈ 0.2. After each transfer, the cultures were renamed to GL1a-6a, GL1a-6b and so on, respectively, and cells from the previous transfer were stored as glycerol stocks. Eventually, the ribose concentration in the medium was reduced to 6.25 mM and the cultures were serially transferred until the growth rate on ribose improved significantly. Cells from glycerol stocks were plated on the appropriate solid medium and incubated at 30°C until single colonies appeared. Single colonies were then picked and cultured on methanol or succinate liquid mineral medium, respectively, with tetracycline as non-selective conditions. Individual clones from different stages of ALE were stored as glycerol stocks for further study and resequencing.

2.7. Growth and evolution using DASGIP bioreactors

The DASGIP Bioblock system (Eppendorf, Hamburg, Germany) was used with DASGIP SR0700ODLS vessels. The headplate was assembled as follows (clockwise with temperature valve as a reference): temperature valve, condenser unit, OD sensor, septum, DO sensor, sampling port + acid & base inlets, pH sensor, and sparger. Additional inlets for medium influx and culture removal were connected via bent stainless-steel flanges (DASIP part no. 78107363) connected to the side arms of the vessel. All

parameters were controlled using the DAScontrol 5 software. Appropriate *M. extorquens* strains were first grown in succinate minimal medium and then transferred to 1 L of minimal medium with 10 mM acetate and 4.5-7.5 mM ribose with shaking in Erlenmeyer flasks in an INFORS HT Minitron incubator and a 5% CO_2 atmosphere. Cells were harvested to inoculate a DASGIP vessels with 700 mL of fresh medium to an initial OD_{600} of 0.5-1.0. The temperature was kept constant at 30°C and cold water flowed through the condenser unit to prevent evaporation. The pH of the medium was kept at 7.0 by automatic addition of 0.5 M H_2SO_4 and 1 M NH_4OH. The culture was stirred at 350-400 rpm using two 6-blade Rushton blades. The GA4 gas mixing module was used to deliver 5% CO_2 in air to the culture at a constant flow rate of 8-12 L h^{-1} via an L-shaped sparger. A quasi-chemostat operation was achieved adjusting the flow rate of pumps controlling fresh medium influx and cell culture removal. In the early stages of evolution, the cells were placed under stringent and less stringent conditions by addition of small amounts of succinate to the vessel. After approx. 50 days, succinate was no longer added, the flow rate was increased, and the ribose concentration was lowered gradually. Acetate and 5% CO_2 sparging were kept constant as energy and carbon sources, respectively. Cell culture was sampled at regular intervals and stored as glycerol stocks.

2.8. Whole genome (re-)sequencing of evolved clones

2.8.1. Short-read Illumina sequencing and analysis

Genomic DNA of selected strains was isolated using the Nucleo®Spin Microbial DNA kit (Macherey Nagel). Eppendorf DNA LoBind Tubes were used at all times (VWR 525-0130). DNA concentrations were measured using the Qubit dsDNA BR Assay Kit (Thermo Fischer Scientific Q32853). Libraries were prepared using 250 ng of isolated genomic DNA following the NEBNext® Ultra™ II FS DNA Library Prep with Sample Purification Beads (New England Biolabs (NEB) E6177S). NEBNext® Multiplex Oligos for Illumina® (NEB E7500) were used. An approximate insert size distribution of 200-350 bps was selected using the provided beads according to manufacturer's instructions. Adaptor-ligated DNA was enriched through 4 cycles of PCR. Quality and size distribution of the libraries was assessed on a Bioanalyzer High Sensitivity DNA Analysis chip. Prepared libraries were sequenced with the MiniSeq High Output Reagent Kit (300-cycles) (Illumina FC-420-1003) on a MiniSeq System. Sequence analysis was carried out using CLC Genomic Workbench 11.0 (QIAGEN). Reads were mapped against the *M. extorquens* AM1 genome (Vuilleumier et al., 2009) using GenBank Accession files CP001510 (main chromosome), CP001511 (megaplasmid), CP001512 (p1META1), CP001513 (p2META1), CP001514 (p3META1), and pTE1154 (Ribose assimilation plasmid) as reference replicons. Variants were detected with the Basic Variant Detection tool (min. coverage = 10; min. count = 8; min. frequency = 50%) and filtered against parental control reads.

2.8.2. Long-read PacBio sequencing and Hybrid genome assemblies

PacBio sequencing of selected clones was performed by BaseClear (Leiden, The Netherlands). Raw Illumina reads were quality-filtered using the illumina-utils library v1.4.1 (Eren et al., 2013). PacBio reads were converted from bam format to fastq format using BEDTools v2.25 (Quinlan and Hall, 2010). First, the PacBio reads were corrected

using LoRDEC v0.9 (Salmela and Rivals, 2014) and the quality-filtered Illumina reads. Then the corrected PacBio reads were assembled with Canu v1.8 (Koren et al., 2017). The Illumina reads were assembled using SPAdes version 3.13.1 (Antipov et al., 2016b) and the PacBio assembly as trusted contigs. Similarly, plasmids were assembled with the plasmidSPAdes algorithm (Antipov et al., 2016a). Finally, assembled genomes were polished using Pilon v1.23 (Walker et al., 2014).

2.9. Biochemical characterization of TIM WT and G77D mutant

2.9.1. Overexpression and protein purification

Triose phosphate isomerase (TIM) wild-type and its G77D mutant were overexpressed in CM2720 using pTE1897 or pTE1898, respectively. Cells from pre cultures were used to inoculate 2L of methanol medium with kanamycin to a starting OD_{600} of 0.05. Expression was induced with 1 mM of IPTG at $OD_{600} \approx 0.1$. Cells were harvested by centrifugation (4500 rcf, 10 min, 4°C) once they reached $OD_{600} \geq 4$ and the cell pellet was frozen at -20°C. Cell pellets were resuspended in 20 mL of 50mM HEPES pH 7.5, 150mM NaCl, 10 mM $MgCl_2$, 10 µg/mL DNAse I (AppliChem A3778). Cells were lysed by sonication, the lysate was cleared by ultracentrifugation (100,000 rcf, 60 min, 4°C), and filtered through a 0.45 µm filter. The cleared lysate was loaded on a 1 mL StrepTrap HP (GE Healthcare). Unbound protein was removed with 20 mL of 50mM HEPES pH 7.5, 150mM NaCl. The protein was eluted in 50mM HEPES pH 7.5, 150mM NaCl, 2.5 mM desthiobiotin. Elution fractions were pooled and concentrated in Amicon® Ultra 4 mL centrifugal filters with a 10K cut-off (Merck). The buffer was exchanged during concentration to 20mM Triethanolamine-HCl pH 7.6. Concentration was determined on a NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA) using the extinction coefficient at 280 nm, as calculated by ProtParam (http://expasy.org). Enzyme purity was confirmed by SDS-PAGE. The purified proteins were stored in 20% (v/v) glycerol at −80°C.

2.9.2. Enzyme kinetic measurements

TIM activity was measured spectrophotometrically following NADH oxidation (ε = 6.22 mM^{-1} cm^{-1}) through the isomerization of glyceraldehyde 3-phosphate (GA3P) to dihydroxyacetone 3-phosphate (DHAP) coupled to commercial α-glycerophosphate dehydrogenase (α-GDH; Sigma G6751). Assays were performed on a Cary-60 UV/Vis spectrophotometer (Agilent) using 10mm quartz curvettes (Hellma, Müllheim, Germany) at 30°C. The assay (120 µL) contained 100 mM Triethanolamine-HCl pH 7.6, 200 µM NADH, 40 µg of α-GDH, 0.53 µM BSA, and purified *M. extorquens* TIM (1 ng for WT, 59.5 ng for G77D mutant). Different concentrations of GA3P (0.17-6 mM) were used to start the reaction. Kinetic parameters were determined using purified proteins and six different substrate concentrations measured in triplicates. The data points were fitted using GraphPad Prism 7 to obtain k_{cat} and apparent K_M values.

2.10. Quantification of ribose by HPLC

Ribose from culture supernatants was quantified by HPLC using the Shimadzu Prominence LC-20A system. Samples (10µL injection volume) were separated through a Rezex ROA-Organic Acid (H^+ 8%) column (Part No. 00G-0138-E0 Phenomenex, Aschaffenburg,

Germany) at 55°C applying a flow rate of 0.8 mL min^{-1} with MiliQ water or 5 mM H_2SO_4 as the mobile phase. Ribose was detected using RID-20A Refractive Index Detector at 40°C. Ribose spiked into mineral medium exhibited a linear behavior between 0.5 and 22.5 mM which were used as standards for quantification. Peak height was calculated using the LabSolutions software.

3. Results and Discussion

3.1. Rational engineering of ribose consumption

M. extorquens AM1 is able to grow on various carbon sources, including C_1 compounds and a number of organic acids. However, growth on sugars and related substrates has never been reported for this organism (Knief et al., 2010; Peyraud et al., 2011). The first bottleneck for growth on ribose is its transport across the outer and inner membranes. While no ribose transporter could be identified in the genome of *M. extorquens* AM1, 15 putative porins and 13 putative carbohydrate facilitator homologs belonging to the major facilitator superfamily, are encoded by *M. extorquens* AM1.

Once inside the cell, the next step is activation of ribose into a corresponding C_5 sugar phosphate, *e.g.* ribose-5P (R5P), which can be assimilated via the pentose phosphate pathway (PPP). While *M. extorquens* AM1 lacks an obvious ribose kinase homolog for the activation of ribose into R5P, it encodes genes for the carbon rearrangement reactions of the PPP (Fig. **1A**), through which R5P could be further converted into glyceraldehyde 3-phosphate (GA3P) and fructose 6-phosphate (F6P). GA3P can enter central carbon metabolism, while the fate of F6P is unclear, as *M. extorquens* AM1 lacks a phosphofructokinase homolog. F6P can be converted to 6-phosphogluconate (6PGl) via the oxidative PPP and further assimilated to GA3P and pyruvate via the Entner-Doudoroff pathway (ED pathway) (Entner and Doudoroff, 1952) or decarboxylated to ribulose 5-phosphate (Ru5P) for a new round of re-shuffling (see Fig. **1A** and Table **S4**.).

To enable the intracellular activation of ribose, we first introduced pTE1145 encoding ribose kinase (*rbsK*) from *E. coli* into *M. extorquens*. This strain was able to grow on ribose as the sole carbon source (Fig. **1C**), while an empty vector control did not show any growth over the same time (Fig. **1B**). To test the effect of improved ribose uptake on growth, we introduced the sugar facilitator *glf* from *Zymomonas mobilis* in addition to *rbsK*. Glf is known to have a broad substrate spectrum including glucose, fructose, xylose and mannose, therefore we speculated that it would facilitate ribose transport as well (Parker et al., 1997; Weisser et al., 1995; Weisser et al., 1996). CM2720 + pTE1154 (expressing *rbsK* and *glf*) indeed improved growth on ribose at lower ribose concentrations (see Fig. **1D**), suggesting that sugar import was facilitated by Glf. Altogether these experiments showed that ribose metabolism can be engineered into *M. extorquens* AM1 by introduction of a ribose kinase and a sugar facilitator.

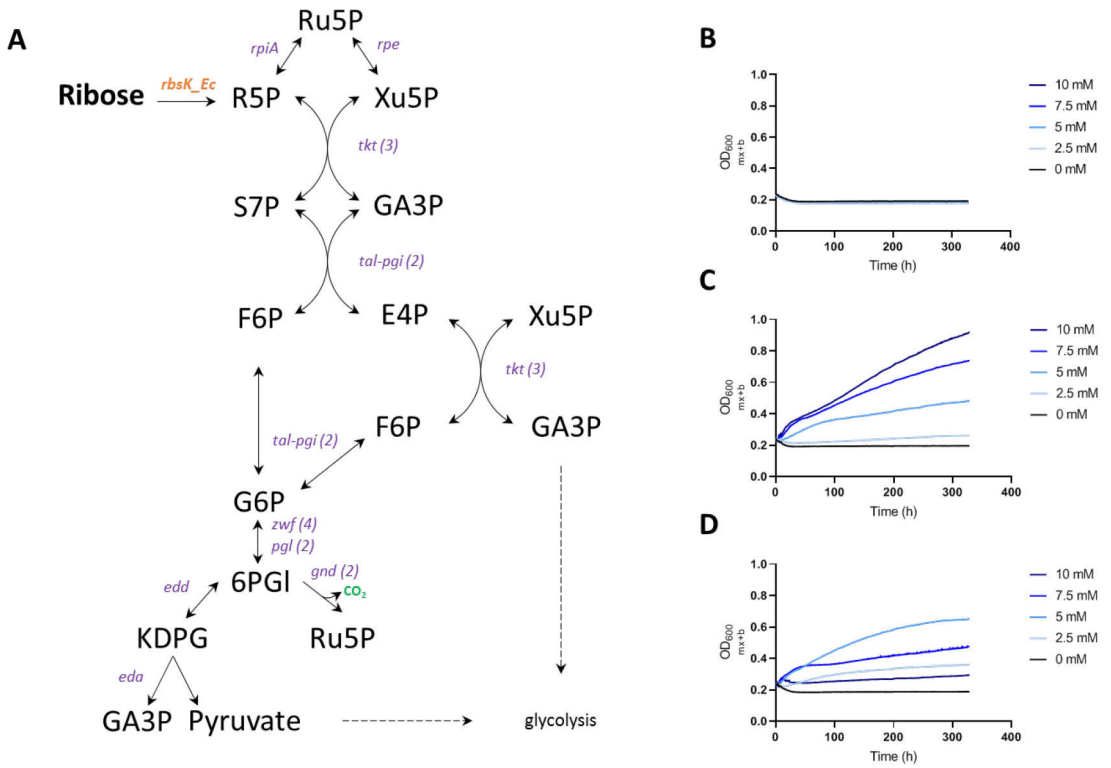


Fig. 1. Engineering ribose utilization in *M. extorquens* AM1. **A**) Metabolic map of sugar phosphate metabolism. Native genes are shown in purple, heterologous genes in orange. In case of multiple copies of a gene the number of copies is given in parenthesis. Refer to Table **S4** for more information on each gene. Growth curves of CM2720 with pTE100 (empty vector) (**B**), expressing *rbsK* from *E. coli* via pTE1145 (P_{coxB}-*rbsK*) (**C**), or *rbsK* from *E. coli* and *glf* from *Z. mobilis* via pTE1154 (P_{coxB}-*rbsK-glf*), (**D**) in mineral medium with 0-10 mM ribose. Abbreviations: R5P – ribose 5-phosphate, Ru5P – ribulose 5-phosphate, Xu5P – xylulose 5-phosphate, S7P – sedoheptulose 7-phosphate, GA3P – glyceraldehyde 3-phosphate, F6P – fructose 6-phosphate, E4P – erythrose 4-phosphate, G6P – glucose 6-phosphate, 6PGl – 6-phosphogluconate, KDPG – 2-keto-3-deoxy-6-phosphogluconate.

3.2. Adaptive laboratory evolution to improve growth on ribose

Having engineered a basic *M. extorquens* AM1 strain capable of growing on ribose as the sole carbon source, we further used ALE to improve growth rate. Six parallel cultures of *M. extorquens* CM2720 + pTE1145 (only *rbsK*) and CM2720 + pTE1154 (P_{coxB}-*rbsK-glf*), referred to as the parent strain GL0 henceforth, were transferred to and sub-cultured in mineral medium with ribose as the sole carbon source to apply selective pressure (Fig. **2**). At regular intervals, samples from the supernatant of each culture were analyzed by HPLC to verify that increasing cell density (OD_{600}) correlated with ribose consumption. While *M. extorquens* CM2720 + pTE1145 (only *rbsK*) strains did not show improved growth over the duration of the experiment, the growth rate of the *M. extorquens* CM2720 + pTE1154 strains increased significantly by a factor of 3 to 4, after approx. 30 generations.

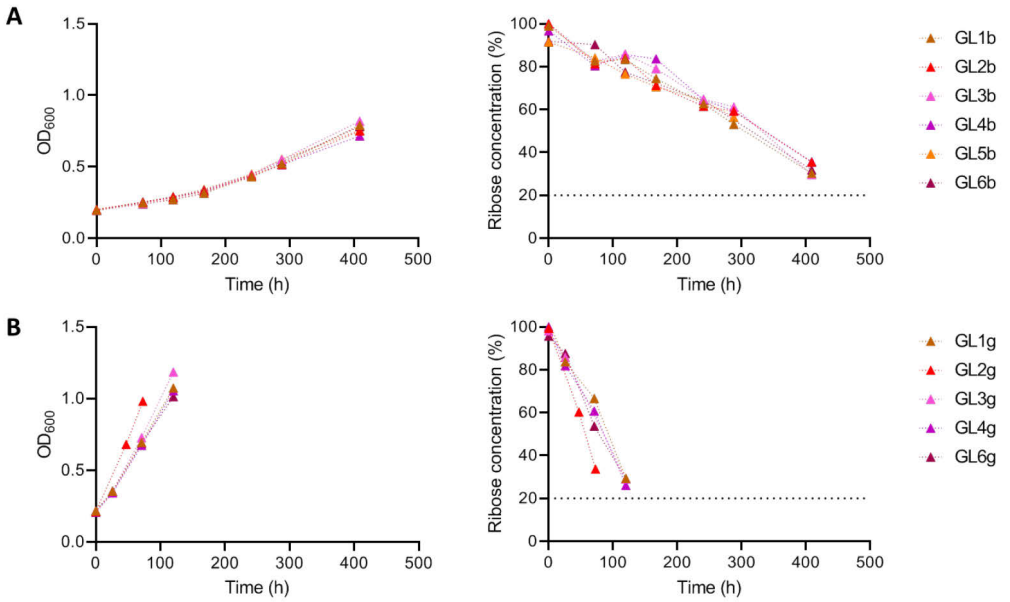


Fig. 2. Adaptive laboratory evolution (ALE) of *M. extorquens* AM1 towards improved ribose utilization via serial transfer. OD_{600} (left panel) and ribose concentration in supernatants (right panel) of GL1-6 from transfer "b" (**A**) and "g" (**B**).

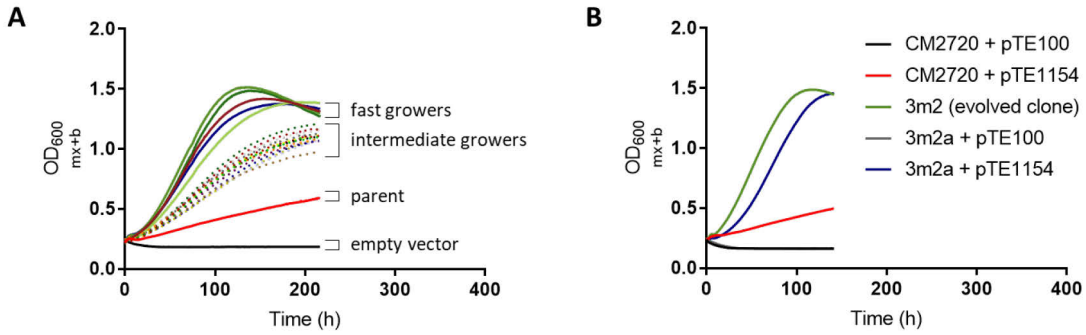


Fig. 3. Clones with improved growth on ribose. **A**) Growth profile of selected GL evolved clones in mineral medium with 6.25 mM ribose. The empty vector control (CM2720 + pTE100) is shown in black, the parent GL0 (CM2720 + pTE1154) is shown in red, evolved single clones of GL0 with intermediate growth are shown in dashed lines, best growing evolved clones are shown in solid lines. **B**) Growth profile of a cured evolved clone, 3m2a, and retransformed with an empty vector control (pTE100) or the ribose assimilation plasmid (pTE1154).

We isolated individual clones from these cultures and characterized their growth behavior on ribose. All clones tested showed improved growth on ribose as the sole carbon source compared to the parent GL0 and could be divided into two sub-populations: intermediate and fast growers (see Fig. **3A**). To test the genetic basis for improved growth on ribose of these strains, we cured 3m2, one of the fastest growing strains, of its plasmid pTE1154 (encoding *rbsK* and *glf*), resulting in strain 3m2a. We then re-transformed 3m2a with empty vector (pTE100) and the original pTE1154. Strain 3m2a + pTE1154 was able to grow on ribose nearly as fast as the evolved clone 3m2 (see Fig. **3B**), while strain 3m2a + pTE100

did not show growth on ribose. This strongly indicated that the growth phenotype was linked to changes in the genetic background of 3m2, which in turn improved growth rate of the introduced pathway.

Subsequently, we performed whole genome sequencing (WGS) on clones from three different lineages (1m3, 3m1, 3m2, 3m3, and 6m1) to elucidate the genetic changes responsible for the improved phenotype in our evolved clones. WGS revealed several single nucleotide polymorphisms (SNPs) in the evolved clones, which were absent in the parent (GL0) (Table **1**). Interestingly, all clones lost at least part of a recently characterized polyketide synthase (PKS) cluster located on the megaplasmid encoded by the genes *tobA-L*. This cluster was shown to produce toblerol compounds, which are apparently relevant for interactions between *Methylobacterium* species (Ueoka et al., 2018). While most of the clones lost ≈15 kb within this locus encompassing *tob(E)FGHIJKL* genes, one had an SNP resulting in an early stop codon within *tobG* (see **Table 1** and Fig. **S1**). This mutation was reverse engineered in the ancestral strain resulting in CM2720 TobG_R200* transformed with pTE100 (empty vector) or pTE1154 (P_{coxB}-*rbsK-glf*), however there was no significant growth improvement on ribose between TobG_R200* and the wildtype CM2720 (see **Fig. 4**), suggesting that this mutation is not (solely) responsible for the observed growth phenotype.

Table 1. Single nucleotide polymorphisms (SNPs) of GL evolved clones compared to the parent (GL0).

Ordered Locus (Gene)	Type of SNP[a]	Assigned function or protein domain	Evolved clone				
			1m3	3m1	3m2	3m3	6m1
META1p2066	R403S	Outer membrane protein assembly factor – BamA.					
META1p2915	Deletion M137*	DNA repair. Hypothetical protein.					
META1p4266	V68I	Putative CoxB-related protein.					
META1p4406 (*rpoC*)	R348H	RNA polymerase beta'-subunit.					
META1p5115 (*tpiA*)	G77D	Triosephosphate isomerase.					
META2p0010 (*tobE*)	Insertion N694*	Toblerol PKS cluster.					
META2p0012 (*tobG*)	L306H, D305G	Toblerol PKS cluster.					

[a]SNPs resulting in early stop codons are shown following the numbering of the original amino acid and an asterisk. Shaded areas denote if SNP mutation found in the corresponding evolved clone.

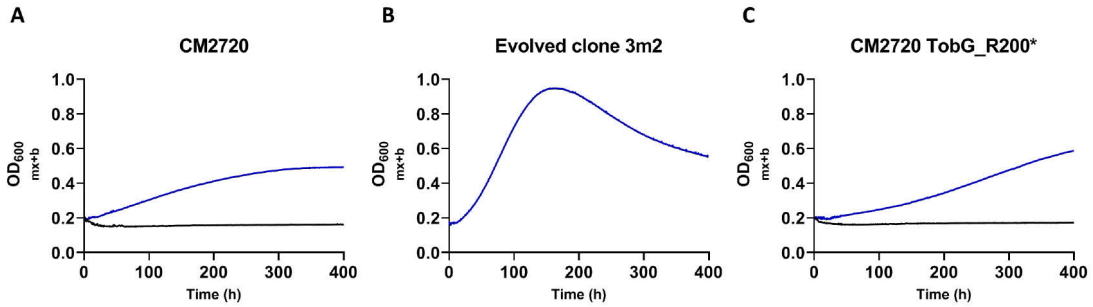
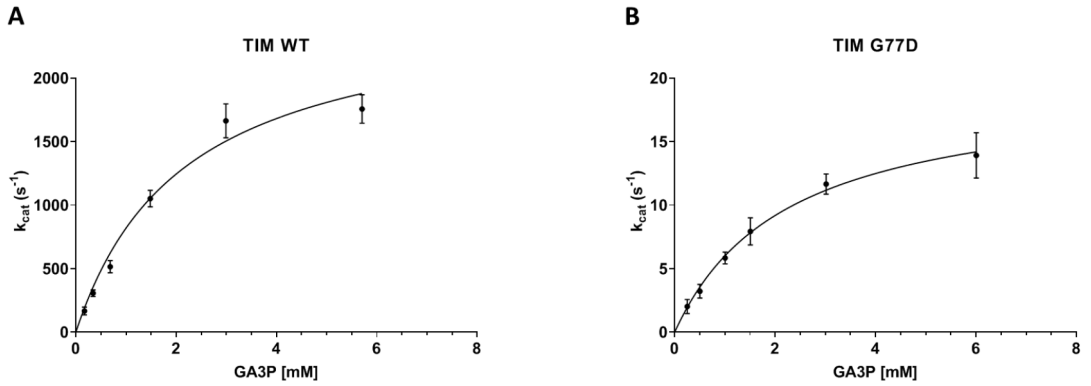


Fig. 4. Reverse engineering of TobG_R200* mutation in CM2720. Growth profiles on 4.5 mM ribose of parental strain CM2720 (**A**), evolved clone 3m2 (**B**), and CM2720 TobG_R200* (**C**). Strains with the empty vector (pTE100) are shown in black while strains with ribose assimilation plasmid (pTE1154) are shown in blue.

The remaining SNPs differed across the different lineages. Two of the evolved clones had missense mutations in *rpoC* (encoding RNA polymerase β' subunit), and a gene encoding a putative CoxB-related protein (META1p4266) (see Table **1**). These mutations might suggest a reorganization of transcription at a global and a specific level given that ribose assimilation genes (*rbsk-glf*) are under the control of the *coxB* promoter. Interestingly, one of the other evolved clones, 1m3, had a G77D mutation within the triose phosphate isomerase (TIM) gene. This amino acid is conserved in all TIM sequences and a G72A mutation in the corresponding amino acid causes severe defects in humans (Cabrera et al., 2018; Lolis et al., 1990; Watanabe et al., 1996). This glycine residue is part of a flexible loop at the dimerization interface which causes dimer instability if mutated and, therefore, lowers the specific activity of TIM (Cabrera et al., 2018). We overproduced and purified *M. extorquens* TIM wildtype and G77D mutant for biochemical characterization. As expected, the V_{max} of the G77D mutant was two orders of magnitude lower than the wildtype enzyme, while the K_M remained unaffected (Fig**. 5**). The decreased activity of the TIM G77D mutant points towards an altered GA3P metabolism in the 1m3 clone. It is tempting to speculate that by lowering TIM activity, the GA3P produced during the ribose assimilation will be shuffled faster by GA3P dehydrogenase into lower glycolysis and therefore biomass formation. Thus, resulting in enhanced assimilation of ribose. Since TIM is fully reversible in its isomerization of GA3P and DHAP, any changes in flux would depend on metabolite concentrations and the activity of enzymes surrounding this important metabolic branching point.



Enzyme	V_{max} (U mg^{-1} protein)	k_{cat} (s^{-1})	K_M (mM)	K_{cat}/K_M (M^{-1} s^{-1})	Relative activity (%)
TIM WT	5456 ± 412	2591 ± 196	2.16 ± 0.38	1.2×10^6	100
TIM G77D	41.12 ± 2.65	19.5 ± 1.3	2.25 ± 0.34	8.7×10^3	0.75

Fig. 5. Biochemical characterization of *M. extorquens* triosephosphate isomerase (TIM). Michaelis-Menten kinetics of TIM WT (**A**) and G77D mutant (**B**). The data points were analyzed using nonlinear regression in GraphPad Prism 7. Mean values ± standard deviation are shown.

3.3. Establishing ribose and CO_2 co-fixation via the CBB cycle

Having evolved *M. extorquens* AM1 for growth on ribose, we aimed at improving the carbon efficiency of our strain on ribose by implementing a CBB cycle module that would allow co-fixation of CO_2, and thus additional biomass formation, when extra energy is provided (see Fig. **6A**). To this end, we aimed at separating carbon assimilation and energy generation in our evolved strain, similar to an earlier study (Schada von Borzyskowski et al., 2018a). While in the previous study we had disabled carbon assimilation from methanol (at the level of *ftfL*), we now decided to block carbon assimilation from acetate.

In the wildtype, a part of the acetate is fully oxidized to CO_2 providing energy in the form of reducing equivalents through the TCA cycle, while carbon is assimilated into biomass via the ethylmalonyl-CoA pathway (EMCP) (Erb et al., 2007; Schneider et al., 2012; Smejkalova et al., 2010). Therefore, we deleted *phaA*, the first step of the EMCP, to disable carbon assimilation, while still enabling acetate oxidation and thus energy conservation. The *phaA* knockout rendered strain 3m2a unable to grow on acetate plus CO_2, both with and without ribose, as expected (Fig. **6B-C**). When 3m2a *phaA*::0 was transformed with pTE1145 (encoding *rbsK*), the strain grew in the ribose-amended medium, in line with the observation that RbsK enables assimilation of ribose in *M. extorquens*. Next, we introduced a CBB cycle module to allow co-fixation of CO_2 on ribose (see Fig. **6A**), similar as previously reported (Schada von Borzyskowski et al., 2018a). We combined native phosphoribulokinase (encoded by *prk*) from *M. extorquens* AM1 and *Rhodospirillum rubrum* RubisCO form II (encoded by *cbbM*) into one operon under the control of the strong P_{mxaF} promoter and added the ribose assimilation module (P_{coxB}-*rbsK*) downstream separated by the transcriptional terminator T_{rrnB}, creating plasmid pTE1157. Growth rate

and yield were strongly improved, when strain 3m2a *phaA*::0 was transformed with pTE1157 (encoding *rbsK* and a functional CBB module), indicating that the CBB module was active and allowed for the additional fixation of inorganic carbon into biomass. Notably, when we tested a CBB module with RubisCO variant K191M that renders the enzyme catalytically inactive (Cleland et al., 1998; Mueller-Cajar et al., 2007) (pTE1171), the corresponding strain showed almost no growth. In this strain ribose is converted by Prk into ribulose-1,5-bisphosphate (RuBP), a toxic dead-end product that cannot be further converted due to an inactive RubisCO, which independently confirmed that a functional CBB module is required for efficient growth of the 3m2a *phaA*::0 mutant on acetate, ribose and CO_2. We further aimed at enhancing activity of the CBB module (and growth rate) by introducing bifunctional FBP/SBPase from *Cupriavidus necator* (*cbbF$_c$*) on top of the CBB module and *rbsK* (pTE1821). This resulted in a slight growth improvement compared to pTE1157. A similar growth rate enhancing effect was observed, when we included a sugar facilitator gene, *glf*, downstream of *rbsK* (data not shown). Altogether, these experiments demonstrated that introduction of a functional CBB module in our ribose evolved strain enabled co-fixation of CO_2 into biomass, when an additional energy sourcewas provided.

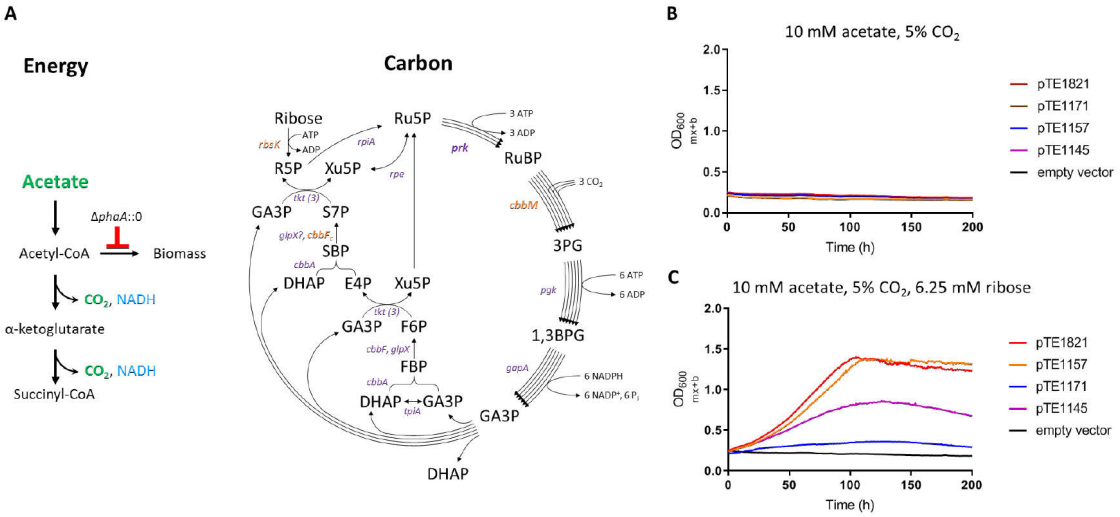


Fig. 6. Engineered CBB cycle for synthetic autotrophy in *M. extorquens*. **A**) Simplified scheme of the strategy used for achieving synthetic autotrophy with energy being generated via acetate oxidation in a *phaA*::0 strain background and carbon assimilation via CO_2 fixation of the heterologous CBB cycle. Native genes are shown in purple, heterologous genes in orange. In case of multiple copies of a gene, the number of copies is given in parenthesis. The reader should refer to Table **S4** for more information on each gene. Growth curves of 3m2a *phaA*::0 with different versions of the CBB cycle: empty vector; pTE1145 – P$_{coxB}$-*rbsK*; pTE1171 – P$_{mxaF}$-*prk*-*cbbM*_K191M-T$_{rrnB}$-P$_{coxB}$-*rbsK*; pTE1157 – P$_{mxaF}$-*prk*-*cbbM*-T$_{rrnB}$-P$_{coxB}$-*rbsK*; pTE1821 – P$_{mxaF}$-*prk*-*cbbM*-T$_{rrnB}$-P$_{coxB}$-*rbsK*-*cbbFc*. All grown on 10 mM acetate (energy), 5% CO_2, and either 0 mM ribose (**B**) or 6.25 mM ribose (**C**). Abbreviations: Ru5P – ribulose 5-phosphate, RuBP – ribulose 1,5-bisphosphate, 3PG – 3-phosphoglycerate, 1,3BPG – 1,3 bisphosphoglycerate, GA3P – glyceraldehyde 3-phosphate, DHAP – dihydroxyacetone phosphate, FBP – fructose 1,6-bisphosphate, F6P – fructose 6-phosphate, Xu5P – xylulose 5-phosphate, E4P – erythrose 4-phosphate, SBP – sedoheptulose 1,7-bisphosphate, S7P – sedoheptulose 7-phosphate, R5P – ribose 5-phosphate.

3.4. Adaptive laboratory evolution (ALE) in bioreactors

We sought to further evolve selected 3m2a *phaA*::0 +pTE1821 strains with an active CBB module. The idea was to limit ribose concentration in the medium to evolve the CBB module from a linear, CO_2-cofixing ribose assimilation module into a fully functional CBB cycle that allows biomass formation from CO_2 only. To that end, fresh medium with acetate and ribose was regularly fed to the bioreactors and air with 5% CO_2 was sparged through the cultures. The dilution rate was adjusted to keep the cell density constant. After >50 days, the culture in one of the bioreactors started growing faster (3-4x), which allowed to lower ribose concentration to 4.5 mM without affecting the growth rate. Single clones of the population were tested for improved growth on acetate, ribose and CO_2 compared to parental and empty vector controls (see Fig. **7A**). These clones grew faster than the parent, 3m2a *phaA*::0 +pTE1821. Sequencing revealed that all clones lost the CBB module (*prk-cbbM*) through an unwanted recombination event between the ribosomal binding sites (RBS) of *prk* and *rbsK*. As a result, *prk* and *cbbM* (needed for CO_2 fixation) were excised and *rbsK* (needed for ribose assimilation and previously transcribed from the P_{coxB} promoter) was placed under control of the stronger P_{mxaF} promoter (see Fig. **7B**). This resulted in improved growth on ribose, as a carbon and energy source, and acetate, as an energy source, while the dependency on CO_2 was completely lost.

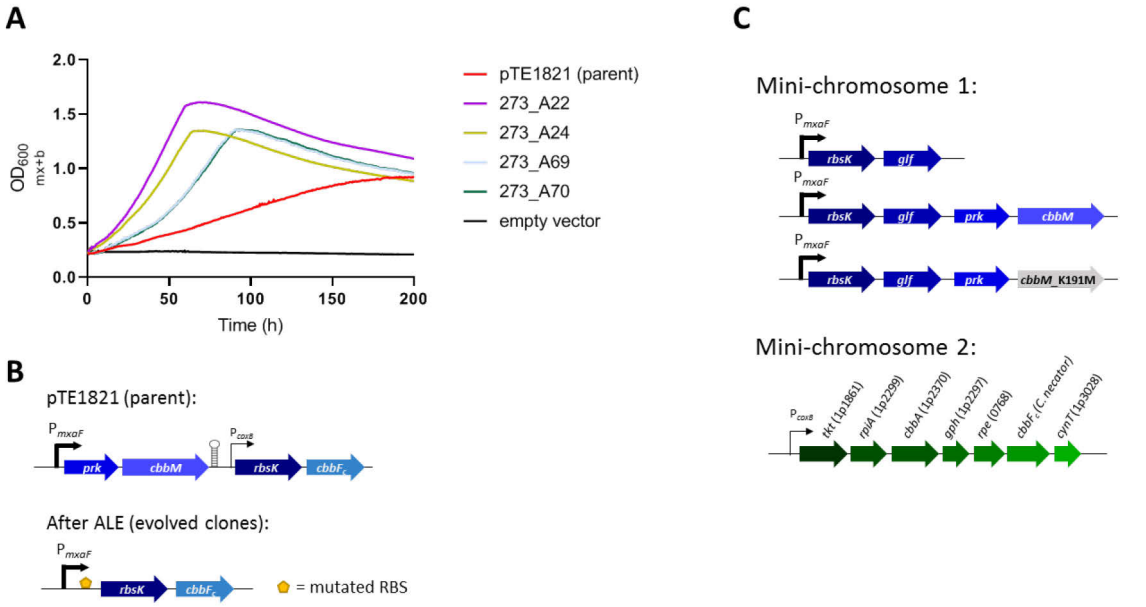


Fig. 7. Evolved clones after 61 days of ALE in DASGIP bioreactors. **A**) Growth profile of selected clones (A22, A24, A69, A70) on 10 mM acetate, 4.5 mM ribose and 5% CO_2. The empty vector and pTE1821 (parent strain) are given as a reference. **B**) Scheme of the pTE1821 plasmid belonging to the (parental) strain at the start of ALE in DASGIP bioreactors and the recombination event that resulted in improved growth on acetate and ribose. All evolved clones tested had a mutated RBS which served as the recombination site to excise *prk-cbbM* and enhance expression of *rbsK* via the upstream P_{mxaF} promoter. **C**) Proposed improved transcriptional units for ribose assimilation with engineered CBB cycle on a mini-chromosome and additional PPP/CBB cycle genes for overexpression on a separate mini-chromosome. Refer to Table **S4** for more information on each gene.

4. Conclusions and Outlook

Here we successfully engineered *M. extorquens* AM1 for the assimilation of ribose, a novel carbon source, by combining rational engineering and ALE. The evolved strains showed robust growth on ribose and were subjected to further metabolic engineering to create a strain that allows co-fixation of CO_2 with ribose as carbon source, when additional energy is provided (by acetate oxidation). The successful extension of the metabolic capabilities of *M. extorquens* AM1 in respect to C_5 sugars opens up interesting applications for the conversion of plant-derived material as feedstock in the future.

At the same time, our results represent a step towards establishing an autotrophic *M. extorquens* AM1 strain that grows with CO_2 as sole carbon source. Recently, a functional CBB module was introduced into *M. extorquens* AM1 (Schada von Borzyskowski et al., 2018a). While this strain was able to incorporate CO_2 into biomass (using methanol as an energy source), it was unable to grow continuously on CO_2 most likely due to an unbalanced flux through and out of the CBB cycle, which requires fine-tuning and is an important prerequisite for achieving robust growth based on autocatalytic cycles (Antonovsky et al., 2016; Barenholz et al., 2017; Gleizer et al., 2019; Herz et al., 2017). We note that our engineered strain now provides the possibility to constantly feed ribose to *M. extorquens* AM1 to keep the CBB module running during ALE, which was key for recent efforts to evolve a functional CBB cycle in *P. pastoris* and *E. coli*, respectively (Gassler et al., 2019; Gleizer et al., 2019).

The ALE experiment in which we aimed to further convert our strain into a fully autotrophic strain showcases that designing the right evolutionary pressure is a very delicate undertaking and prone to create escape mutants. Having learned from this experience, we now propose a modified genetic approach for a new round of ALE. First, we will utilize the newly developed mini-chromosomes for *M. extorquens* with the advantage of having a stable copy number as well as a faithful segregation system to daughter cells (Carrillo et al., 2019). All experiments thus far were performed using a multi-copy plasmid, whose copy number might be up- or downregulated depending on the metabolic state of the cell leading to artefacts. Secondly, we will create new transcriptional units where ribose assimilation genes (*rbsK* ± *glf)* are already under the control of P_{mxaF}, followed by CBB cycle genes (*prk-cbbM*) to avoid escape mutants, as observed in the ALE experiment. Thirdly, we could upregulate related enzymes, which might be currently limiting flux through the CBB cycle, as inferred from the improvement in growth rate that we observed, when $cbbF_c$ (FBP/SBPase) was overexpressed (see Fig. **6B**). Potential target enzymes include: transketolase, R5P isomerase, FBP aldolase, 2-phosphoglycolate phosphatase, carbonic anhydrase, Ru5P epimerase, and FBP/SBPase. Given that these genes are spread out across the genome, we would probably introduce a second copy under the control of a constitutive promoter using a separate yet compatible mini-chromosome (see Fig. **7C**). Lastly, additional gene deletions in our chassis *M. extorquens* strain might be necessary to ensure ribose growth is not just improved by an active CBB cycle, but completely dependent on it. Potential knockouts include glucose 6-phosphate dehydrogenase (Zwf, 4 putative genes) and/or phosphoglucose isomerase

(Pgi, 2 putative genes as bifunctional transaldolase-pgi) as observed in previous studies in *E. coli* (Antonovsky et al., 2016; Gleizer et al., 2019; Herz et al., 2017). Altogether, these efforts might help to finally develop an autotrophic *M. extorquens* AM1 and serve as a platform for the development of artificial CO_2 fixation pathways in the future.

5. Acknowledgments and Funding

We kindly thank Prof. Sprenger for providing us with pZY507*glf-glk*. We are grateful to Dr. José Vicente Gomes-Filho for support in operating the MiniSeq System. We also thank Dr. Javier Serranía for support during read mapping. We thank Sonja Fleissner for technical assistance. This work was supported by an IMPRS fellowship for M.C. and the Max Planck Society.

6. Supporting Information

Table S1 Strains used in this study

Strains	Description	Reference
***E. coli* DH5α**	F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$	Thermo Fischer Scientific
***E. coli* TOP10**	F- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG*	Thermo Fischer Scientific
***M. extorquens* CM2720**	AM1 strain deficient in cellulose production	(Delaney et al., 2013)
GL0	CM2720 + pTE1154 (parent, unevolved), Tc^R	This study
GL 1m3	Evolved clone from GL0, (pTE1154); Tc^R	This study
3m1	Evolved clone from GL0, (pTE1154); Tc^R	This study
3m2	Evolved clone from GL0, (pTE1154); Tc^R	This study
3m3	Evolved clone from GL0, (pTE1154); Tc^R	This study
6m1	Evolved clone from GL0, (pTE1154); Tc^R	This study
3m2a	Evolved clone from GL0, cured of pTE1154 (Tc^S)	This study
3m2a *phaA*::0	AM1, ribose evolved, *phaA*::0, deficient in C_1 and C_2 growth	This study
273_parent (3m2a *phaA*::0 + pTE1821)	3m2a *phaA*::0 + pTE1821 (unevolved), Tc^R	This study
273_A22	Evolved clone from “273_parent”, Tc^R	This study
273_A24	Evolved clone from “273_parent”, Tc^R	This study
273_A69	Evolved clone from “273_parent”, Tc^R	This study
273_A70	Evolved clone from “273_parent”, Tc^R	This study
CM2720 TobG_R200*	Truncated *tobG* allele in CM2720	This study

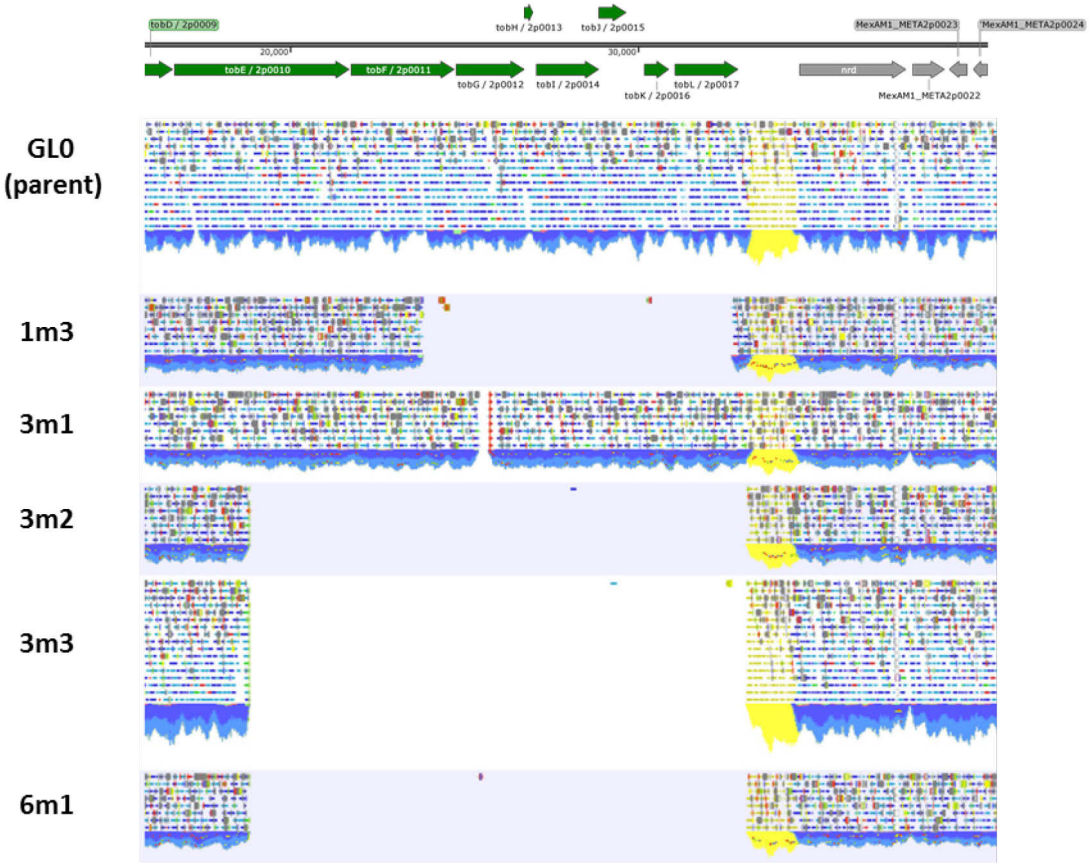


Fig. S1. Reads mapped to toblerol (*tob*) locus from the *M. extorquens* AM1 megaplasmid (GenBank: CP001511.1) of GL0 (CM2720 + pTE1154; unevolved parent) and ribose evolved clones 1m3, 3m1, 3m2, 3m2, and 6m1. All evolved clones lack part of the *tob* locus and clone 3m1 only lacks *tobG*; which was assumed to be the relevant mutation shared amongst all ribose evolved clones. Genes related to the *tob* locus are shown in green, neighboring genes in grey.

Table S2 Plasmids used in this study

Plasmids	Description	Cloning strategy	Reference
pET28b	$P_{T7\text{-}lacO}$ expression vector, N-terminal His tag, Km^R	n. a.	Novagen
pZY507*glf-glk*	Plasmid containing *glf* from Zymomonas mobilis	n. a. Sequencing revealed Glf_{R145A} mutation.	(Weisser et al., 1995)
pCM80	P_{mxaF} promoter, oriV-traJ' origin, Tc^R	n. a.	(Marx and Lidstrom, 2001)
pCM433	Allele exchange vector with *sacB* counter selection, Amp^R Tc^R Cm^R	n. a.	(Marx, 2008)
pTE680	pET16b-derivative with N-terminal Strep-II and C-terminal 6xHis tags, Amp^R	Backbone: NcoI, HindIII digested pET16b; Insert: NcoI, HindIII digested synthetic multiple cloning site bv1cut. Ligation with T4 DNA ligase.	Bastian Vögeli (unpublished results) / Scheffen *et al*.
pTE1885	$P_{L/O4/A1}$ promoter; empty MCS with N-terminal Strep-II tag, pMG160 origin, Km^R	n. a.	(Carrillo et al., 2019)
pTE100	Promoter-less, oriV-traJ' origin, Tc^R	n. a.	(Schada von Borzyskowski et al., 2015)
pTE104	P_{coxB} promoter, oriV-traJ' origin, Tc^R	n. a.	(Schada von Borzyskowski et al., 2015)
pTE94	P_{mxaF}-*prk-cbbM*, oriV-traJ' origin, Tc^R	n. a. Prk from *M. extorquens* AM1, CbbM from *R. rubrum* S1.	(Schada von Borzyskowski et al., 2018a)
pTE95	P_{mxaF}-*prk-cbbM*_K191M, oriV-traJ' origin, Tc^R	n. a. Prk from *M. extorquens* AM1, CbbM from *R. rubrum* S1. *cbbM*_K191M is a catalytically inactive mutant of *cbbM*.	(Schada von Borzyskowski et al., 2018a)
pTE1144	$P_{T7\text{-}lacO}$-RbsK expression vector, N-terminal His tag, Km^R	Backbone: NdeI, EcoRI linearized pET28b; Insert: NdeI, EcoRI digest of PCR product from *E. coli* K-12 gDNA pr. rbsK_NdeI_fw+ rbsK_EcoRI_fw. Ligation with T4 DNA ligase.	This study
pTE1145	P_{coxB}-*rbsK*, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI linearized pTE104; Insert: PCR product from pTE1144 prMC86+MC87. By Gibson assembly.	This study
pTE1122	pET16b-derivative expressing *AtGLYR1* containing an N-terminal Strep-II tag, Amp^R	Backbone: XbaI, EcoRI digested pTE680; Insert: XbaI, EcoRI digested AtGR1_opt (synthetic codon optimized gene for *M. extorquens* from Eurofins). Ligation with T4 DNA ligase.	This study / Carrillo *et al*. / Scheffen *et al*. (unpublished results)
pTE1152	pET16b-derivative expressing *glf* containing an N-terminal Strep-II tag, Amp^R	Backbone: 5.3 kb fragment of NdeI, SpeI digested pTE1122; Insert: NdeI, SpeI digest of PCR product from pZY507-glf-glk prMC98+MC99. Ligation with T4 DNA ligase.	This study

Plasmids	Description	Cloning strategy	Reference
pTE1154	P_{coxB}-*rbsK-glf*, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI linearized pTE1145. Insert: XbaI, KpnI digest of PCR of pTE1152 prT7 fw+MC99. Ligation with T4 DNA ligase.	This study
pTE1157	P_{mxaF}-*prk-cbbM*-T_{rrnB}-P_{coxB}-*rbsK,* oriV-traJ' origin, Tc^R	Backbone: EcoRI linearized and dephosphorylated pTE94; Inserts: PCR product of *E. coli* K-12 gDNA prMC115+MC116, PCR product of pTE1145 prMC117+MC114. By Gibson assembly.	This study
pTE1159	P_{mxaF}-*prk-cbbM*-T_{rrnB}-P_{coxB}-*rbsK-glf,* oriV-traJ' origin, Tc^R	Backbone: EcoRI linearized and dephosphorylated pTE94; Inserts: PCR product from *E. coli* K-12 gDNA prMC115+MC116, PCR product of pTE1154 prMC+MC. By Gibson assembly.	
pTE1171	P_{mxaF}-*prk-cbbM*_K191M-T_{rrnB}-P_{coxB}-*rbsK,* oriV-traJ' origin, Tc^R	Backbone: EcoRI linearized and dephosphorylated pTE95; Inserts: PCR product of *E. coli* K-12 gDNA prMC115+MC116, PCR product of pTE1145 prMC117+MC119, PCR product of pTE1145 prMC118 + MC114. By Gibson assembly.	This study
pTE1821	P_{mxaF}-*prk-cbbM*-T_{rrnB}-P_{coxB}-*rbsK-cbbFc_opt*, oriV-traJ' origin, Tc^R	Backbone: EcoRI, SpeI linearized pTE1157; Insert: MunI, SpeI 1.2 kb fragment of Strep *C. necator cbbFc* (*M. extorquens* codon optimized synthetic gene from Eurofins, see below). Ligation with T4 DNA ligase.	This study
pTE1839	Vector for allelic exchange with *phaA*::0 (META1p3700), shortened peptide has 17 Nter aa fused to 18 Cter aa, *sacB*, Amp^R Tc^R Cm^R	Backbone: BglII, SacI cut pCM433; Inserts: PCR products of 558 bps prMC305+MC306 and 703 bps prMC307+MC308 from *M. extorquens* gDNA. By Gibson assembly.	This study
pTE1897	$P_{L/O4/A1}$-TIM_{WT}, N-terminal Strep-II tag, pMG160 origin, Km^R	Backbone: NdeI, SpeI digested pTE1885; Insert: NdeI, SpeI digested PCR product pr430+431 of clone 3m2 gDNA. Ligation with T4 DNA ligase.	This study
pTE1898	$P_{L/O4/A1}$-TIM_{G77D}, N-terminal Strep-II tag, pMG160 origin, Km^R	Backbone: NdeI, SpeI digested pTE1885; Insert: NdeI, SpeI digested PCR product pr430+431 of clone 1m3 gDNA. Ligation with T4 DNA ligase.	This study
pTE2709	Vector with *tobG* locus of clone 3m1, Amp^R Tc^R Cm^R	Backbone: AatII, XhoI digested pCM433; Insert: PCR product of 3m1 gDNA with prMC446+MC447. By Gibson assembly. Correct sequence via primer walking showed IS3 family insertion in *tobG*.	This study
pTE2712	Vector for truncated TobG_R200* allelic exchange, *sacB*, Amp^R Tc^R Cm^R	Backbone: AatII, XhoI digested pCM433; Inserts: PCR products from pTE2709 prMC446+MC458 and MC459+MC447. Gibson assembly	This study
pTE2732	Vector with TIM_{G77D} for allelic exchange, *sacB*, Amp^R Tc^R Cm^R	Backbone: AatII, XhoI digested pCM433; Insert: PCR product prMC528+MC529 from clone 1m3 gDNA. By Gibson assembly.	This study

n. a. – not applicable, gDNA – genomic DNA, Amp^R – ampicillin resistance, Cm^R – chloramphenicol resistance, Km^R – kanamycin resistance, Tc^R – tetracycline resistance, pr – primers, aa – amino acids.

Strep *C. necator* CbbFc opt– Nter Strep-II tag codon optimized for *M. extorquens* AM1. Start and stop codons are underlined and shown in **bold**.

CTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACC**ATG**GCGAGCTGGAGCCATCCGCAGTTTGAAAAAATTGAAGGCCGCCATATGCCCGAAGTGCAACGGATGACACTGACCCAGTTT
CTTATCGAGGAGCGACGCCGGTATCCGGATGCGTCGGGAGGTTTTAATGGTCTGATTCTCAACGTCGCGATGGCCTGCAAGGAGATCGCGCGCGCGGTGGCCTTCGGCGCGTTGGGCGGGCTT
CATGGCAAGGCGTCGAATCAGGCCGGTGAGGCCGGCGCCGTGAACGTGCAGGGCGAAATCCAACAGAAGCTCGATGTCCTCAGCAACACCACGTTCCTCCGTGTGAACGAATGGGGCGGTTA
CCTCGCCGGTATGGCCTCGGAAGAAATGGAGGCGCCGTACCAGATTCCGGACCACTATCCCCGGGGCAAGTACCTCCTCGTTTTCGATCCCCTCGATGGGTCCTCCAACATCGACGTCAACGTCT
CGGTGGGGTCGATCTTCAGCGTGCTGCGCGCGCCCGAGGGCGCCAGTGCCGTCACGGAGCAGGACTTCCTCCAGCCGGGCAGCGCGCAGGTCGCCGCCGGCTATGCGCTGTATGGCCCAACG
ACCATGCTGGTCCTGACCGTCGGCAACGGCGTCAACGGCTTTACCCTGGACCCAAATCTGGGCGAGTTCTTCCTGACGCACCCGAACCTGCAGGTGCCGGCCGATACGCAGGAGTTCGCCATCA
ACGCCTCGAACTCGCGATTCTGGGAGGCCCCCGTGCAGCGTTACATCGCCGAATGCATGGCGGGCAAGTCGGGCCCGCGCGGCAAGGACTTCAACATGCGGTGGATCGCTAGCATGGTGGCC
GAGGCACACCGCATCCTGATGCGCGGCGGCGTGTTCATGTACCCCCGCGACTCTAAGGACCCAGCCAAGCCCGGCCGCCTGCGCCTCCTCTACGAGGCCAACCCTATCGCGTTTCTCATGGAGC
AGGCCGGCGGCCGCGCCTCGACCGGCCGCCAGACCCTGATGTCCGTCGCGCCCGGCGCCCTCCATCAGCGCATCGGCGTCATCTTCGGCAGCCGCAACGAGGTCGAGCGAATCGAGGGCTATC
ACACCGATCAGACCGACCCGGACCTCCCGTCCCCGCTCTTCAACGAGCGCAGCCTCTTCCGCGCGTCCGCC**TGA**CTAGTCCTGCAGGTACCTAATTCACTGGCCGTCGTT

Table S3 Primers used in this study.

Primer #	Sequence
MC86	AAGCTTAGATCTTGAAAGAAGGAGATATACCATGGGCAG
MC87	GACGGCCAGTGAATTAGGTACCTCTGACTAGTCGACGGAGCTCGAATTC
MC114	TTGTAAAACGACGGCCAGTG
MC115	AGCGCCCTTCCGGCGTAACCGCCTGGCCAGGCATCAAATAAAACG
MC116	GGCTGCCTCTGGGGATACGGTCTAGAGGAAGAGTTTGTAGAAACG
MC117	TTTTGCGTTTCTACAAACTCTTCCTCTAGACCGTATCCCCAGAG
MC118	ATCGGTACACCGGTTGATATCC
MC119	GATATCAACCGGTGTACCGATTGC
MC165	CTAGGTACCGCGAGAGCCTTAGATCAGAC
MC305	AGTGCCACCTGACGTCTAGAAGCGTGCGTCGTAGACGATG
MC306	ATGCAGAGCGTGACGGGCGTACGCGCCGCACC
MC307	TACGCCCGTCACGCTCTGCATCGGCGGCGGCATG
MC308	CGGCTGGATCCTCTAGTGAGCTGTGCAGGTGAACATCGAGTC
MC310	ATGGTCTTGCGGCGGACTTC
MC315	CCTGCTCGAAGATGATCTG
MC430	TATCATATGGCAAGCGAGGGACG
MC431	TGACTAGTAGGGCTCACCCCTCGTAAG
MC446	CGAAAAGTGCCACCTGACGTATACAGCGGAAGGGAATGGG
MC447	TACGTAGGGCCCGCGGCCGCCGATCAGCGTCTGTTAGTTCC
MC452	TTCAACACGCCCACCCAGAC
MC453	GCGTTTCCACCACTCCTTC
MC454	TTGCCCTGCGTTTGAGGATG
MC458	TTAGGGCCGTCCGATGCCGATGC
MC459	CGGCATCGGACGGCCCTAAGTTCGGGTTCCTGGATAC
MC485	GATGTCCTGACTCGCAACTTC
MC486	GAATGCCGAGTTCCTTTGAG
MC528	GAAAAGTGCCACCTGACGTCTGGCACTGCGAATGGACTG
MC529	TACGTAGGGCCCGCGGCCGCAGGGCTCACCCCTCGTAAG
MC532	GTCGACATCGGTTGATACG
MC540	AGGTCGCTGGTTCAAATCC
OS14	AGCCTGGTCGGCTGGATCCTC
rbsK_NdeI_fw	GACGCCTCATATGCAAAACGCAGGCAGCCTCG
rbsK_EcoRI_fw	GCCACTCGAATTCTCACCTCTGCCTGTCTAAAAATG
T7 fw	TAATACGACTCACTATAGGG

Table S4 Genes and corresponding gene products involved in the engineered CBB cycle of *M. extorquens* AM1

Gene name	Annotated gene product	Ordered locus (loci)	Organism of origin
eda	2-keto-3-deoxy-6-phosphogluconate aldolase	META1p1985	*M. extorquens* AM1 (native gene)
edd	6-phosphogluconate dehydratase	META1p1986	*M. extorquens* AM1 (native gene)
zwf	Glucose 6-phosphate 1-dehydrogenase	META1p0364, 1p1408, 1p1864, 1p2490	*M. extorquens* AM1 (native gene)
pgl	6-phosphogluconolactonase	META1p1866, 1p5284	*M. extorquens* AM1 (native gene)
tal-pgi	Bifunctional transaldolase/phosphoglucose isomerase	META1p1862, 1p4950	*M. extorquens* AM1 (native gene)
gnd	6-phosphogluconate dehydrogenase	META1p1863, 1p5283	*M. extorquens* AM1 (native gene)
cbbF	Fructose-1,6-bisphosphatase (sedoheptulose-1,7- bisphosphatase?)	META1p0757	*M. extorquens* AM1 (native gene)
glpX	Fructose-1,6-bisphosphatase (sedoheptulose-1,7- bisphosphatase?)	META1p1940	*M. extorquens* AM1 (native gene)
tkt	Transketolase	META1p1861, 1p2078-9, 1p4949	*M. extorquens* AM1 (native gene)
cbbA	Fructose-1,6-bisphosphate aldolase	META1p2370	*M. extorquens* AM1 (native gene)
tpiA	Triosephosphate isomerase	META1p5115	*M. extorquens* AM1 (native gene)
gapA	Glyceraldehyde-3-phosphate dehydrogenase	META1p2367	*M. extorquens* AM1 (native gene)
pgk	Phosphoglycerate kinase	META1p2369	*M. extorquens* AM1 (native gene)
gpmA	Phosphoglycerate mutase	META1p1116, 1p1218, 1p2472	*M. extorquens* AM1 (native gene)
prk	Phosphoribulokinase	META1p0758	*M. extorquens* AM1 (native gene)
rpiA	Ribose-5-phosphate isomerase	META1p2299	*M. extorquens* AM1 (native gene)
rpe	Ribulose-5-phosphate 3-epimerase	META1p0768	*M. extorquens* AM1 (native gene)
cynT	Carbonic anhydrase	META1p3028, (putative 1p1418 1p4490)	*M. extorquens* AM1 (native gene)
rbsK	Ribose kinase	JW3731	*E. coli* K-12
glf	Sugar facilitator	ZMO0366	*Zymomonas mobilis* CP4
$cbbF_c$	Bifunctional fructose-1,6-/Sedoheptulose-1,7-bisphosphatase	H16_B1390 (chromosomal)	*Cupriavidus necator* H16
cbbM	Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)	Rru_A2400	*Rhodospirrilum rubrum* S1

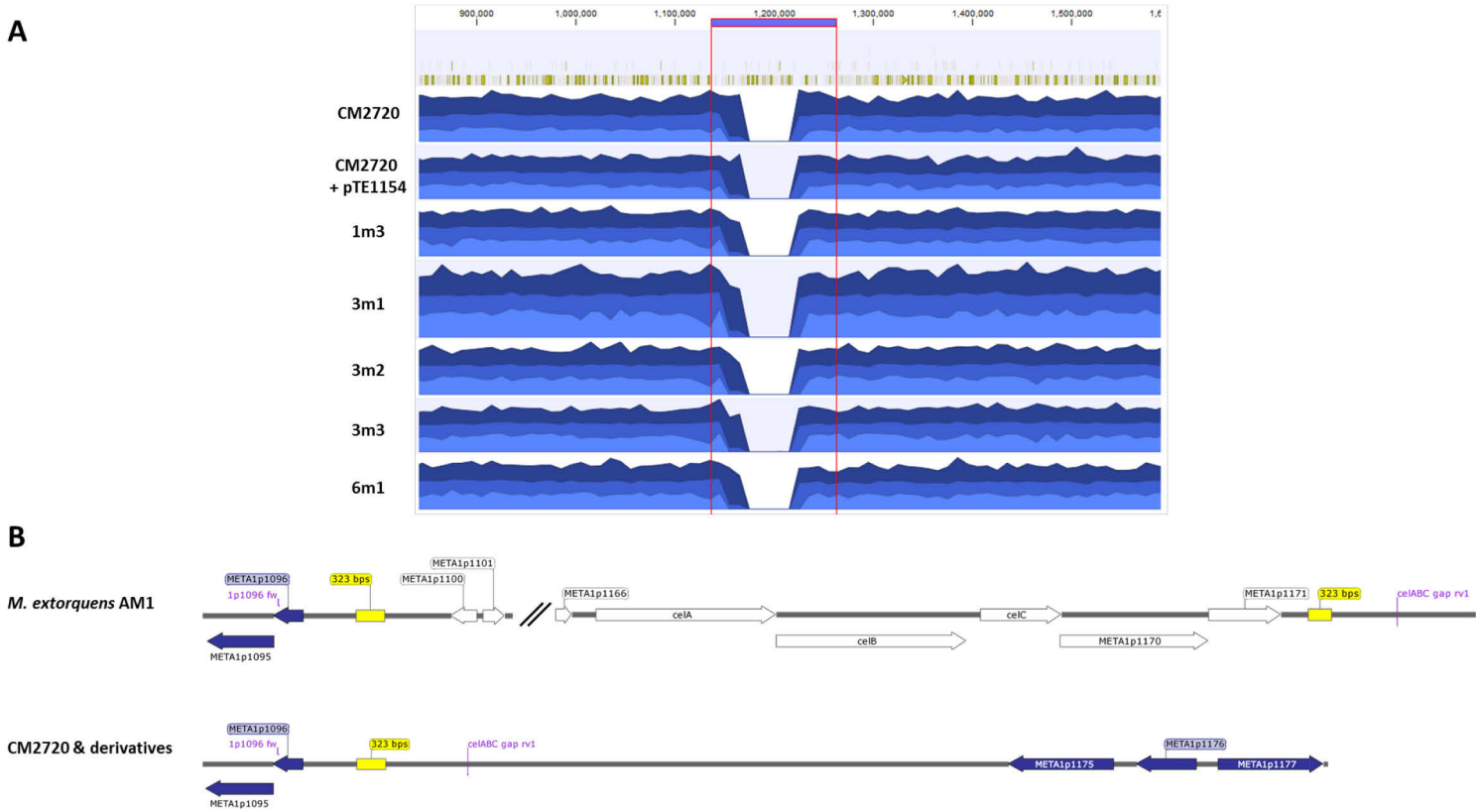


Fig. S2. Locus surrounding the *celABC* (META1p1167-69) operon in the main chromosome of *M. extorquens* AM1 (GenBank: CP001510). **A**) Illumina read coverage in CM2720, CM2720 + pTE1154 (GL0 parent) and ribose evolved clones 1m3, 3m1, 3m2, 3m3, 6m1. Lack of coverage suggests an (unexpected) large deletion in this region (≈ 77 kb) instead of the *celABC* operon deletion (≈ 8.3 kb) reported previously (Delaney et al., 2013). Given that this region is missing in CM2720 'wildtype' (Delaney et al., 2013) as well as all derivative strains suggests it did not occur during ALE. **B**) Genomic organization in wildtype *M. extorquens* AM1 and all CM2720-derivative strains. A 323 bp repetitive region (shown as a yellow box) was likely the template for recombination with *sacB* counterselection to yield a ≈ 77 kb deletion instead of the intended ≈ 8.3 kb deletion yielding CM2720 as performed by (Delaney et al., 2013). The correct sequence at this locus (shown below) was confirmed by colony PCR and Sanger sequencing using primers MC532 (1p1096 fw) and MC540 (celABC gap rv1) shown in purple for CM2720.

DNA sequence after ≈ 77 kb deletion near the *celABC* locus found in CM2720 and derivative strains. Primers MC532 and MC540 used for amplification and sequencing are shown in **purple**; the 323 bps repetitive region is shown in **yellow**; the start codon of META1p1096 is underlined.

GTCGACATCGGTTGATACGCCGCATGCAGCAGCGGCCGCCATGAGACGTTGACGGACGTCCAAGCTCGCCGGAGCAACCACGTATCCCGCCTCACGATCGCCGATCATGATCGAGGCGACGCG
TGGGGCCAAGTGCCAGTGCAGGCCGATCAGCGCACACAGCACGGCATCGAGTCTGTCCTGATCAGCCTTCCGCGGGGCTGCGATGCTGGCGAGGTCGTCAGCCCATGCTGAGAGCCCCGCGAT
CTGGTGCTCCTCGCTCAATCGCCGCATGGTAGCGACCACGTTGTGCCAATGAGCCAGCGCGAAGGTCTTCCGTCAGACTGGGTTTTAGCCCGGCGCGGTGAGCCGCCCGCAGAAGGCCAGATC
GAGTGGCGGAATTGTTAGCGCCAGGTGGTCGAATGTCCTGACCGCCCAAGTGGATGGTAGCGCGGCCCTCCTCAGGCAAAGGTTACACGTTCGAATCGTGTCGGGTGCACCATTCCTGAACTG
ATGTGTGAACGTTTTTCGGGTTGGGTGGATCTACGCGAGCGTGTCGCGGCTGCGGTGGCAGCAGGTGCCTTCTGCCATTGCGCCGCAGCCCACTTCGGGGTCAGCGTTTCGACTGCCAGCCGCT
GGTCGCAACGCTTCCGCCTCGAGGGTCAACTCTTCCCCAAGCCCAGGGGCGGCGATCACACCTCGAAGTGGATCAAGGCGTATGCCGCGCTTGATCCTGGCGATCTCCCAGCAGAAATCGCGTT
GACCTTGCCCCCTGTTTGCCCCCGTTCATAACCAGAGTCCGTTCTGGTTCTGGTCCTGTCGGGACCGGGTTGCAAACGCGTTCGGGGTGAGCCCGCCGAGGCTCGTGTGGGGGCGGCAGGTGTT
ATAGTCGACCCGCCAC**GCCTCGATGATGGTCCGGGCCGCCGGCAGGCTCCGGAACAGATGCTCGTTCAGGCACTCGTCGCGCAAGCGCCCGTTCAAGCTCTCGACAAAACCGTTCTGCTGCG**
GCTTGCCGGGCGCGATGTAATGCCACTCGACCGCACGCTCCTCCTGCCAGCGCAGGATGGCGTGCGAGGTCAGCTCGGTGCCGTTGTCCGAGACGATCATCAGCGGCTTGCCCCGGCCCTTG
ATGATCCGGTCGAGTTCACGCGTGACCCGCCGCCCGGACAGCGACGTGTCGACCACCAGCGCCAGGCACTCCCGCGTGCAGTCATCGACCACGACGAGGATGCGGAAGCGCCGCCCGTCGTC
GAGCGTGTCGGAGACGAAGTCGAGGCTCCAGCGCTGGTTCGGCTCCTGCGGCACCGCGGCGGGCGCTCGCGTGCCAAGTGCTCGCTTGCGACCTCCGCGCTTGCGCACCGACAGCCGCTCCTC
TCGGTAGAGCCGGAAGAGCTTCTTGTGGTTGAGAGCGAGGCCCTCCCGCCGCAGTAGGATGAGCAGGCGCCGGTAGCCGAACCGGCGGCGCTCGCCGGCCAGGCCGCGCAGGCGCACCCGC
ACAGCCGCGTCGTCACCACGGGTCGAGGCGTAGCGGTACGTCTTCGGCTCAAGGCCGATCAGCCCGCAGGCACGACGCTGCGAATAACCTTTCTCCTCGATGGCCCAGCTCACGGCCGTTCTCC
GTGCGCCGGGCGTCAGAAGTTTTTTCCCAGCGCCTCACGCAGCGTGGCCACGTCGAGCATCGCCTCGGCCAGGAGCTTCTTGAGCTTGCGGTTCTCCTCATCGAGCGCCTTCAGACGCCGCGCG
TCCGAGACCTCCATGCCGCCGAAGCGCGAGCGCCACGTGTAGAACGTCGCGTCGCTGATGCCGTGGCGGCGACAGATCTCAGCCACCGGCAGCCCAGCCTGCTGCTCCTTCAGGATGCCGATG
ATCTGCTCTTCGCTAAACCGGCTCTTCTTCATCGTCCGTCTCCTGCTCGACGGACCCTAGCTTCAGAATGAGGGCAGGCCAGGGGGCAAGGTCAGGACGACCGCGCTGACCGCGCGGCTCCTCG
CCCAGCCTACTCCGTGTGCAGCGCGAATGACAGAAAGCCCGCGGGCATGGTGGCCAGCGGGCTTGCGGTCGTCTGTGATGATGTCGGTTGGTTGGTGCGTGTGTTCGTGTGCGTTGGTTGCGG
GGGCT**GGATTTGAACCAGCGACCT**

List of common abbreviations:

KDPG – 2-keto-3-deoxy-6-phosphogluconate
G6P – glucose 6-phosphate
6PGl – 6-phosphogluconate
F6P – fructose 6-phosphate
E4P – erythrose 4-phosphate
FBP – fructose bisphosphate
GA3P – glyceraldehyde 3-phosphate
SBP – sedoheptulose bisphosphate
S7P – sedoheptulose 7-phosphate
DHAP – dihydroxyacetone phosphate
1,3 BPG – 1,3 bisphosphoglycerate
3PG – 3-phosphoglycerate
2PG – 2-phosphoglycerate
RuBP – ribulose bisphosphate
Ru5P – ribulose 5-phosphate
R5P – ribose 5-phosphate
Xu5P – xylulose 5-phosphate

7. References

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., Pevzner, P. A., 2016a. plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics. 32**,** 3380-3387.

Antipov, D., Korobeynikov, A., McLean, J. S., Pevzner, P. A., 2016b. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics. 32**,** 1009-15.

Antonovsky, N., Gleizer, S. s., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D. G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., Milo, R., 2016. Sugar Synthesis from CO_2 in *Escherichia coli*. Cell. 166**,** 115-25.

Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., Milo, R., 2017. Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. eLife. 6.

Cabrera, N., Torres-Larios, A., Garcia-Torres, I., Enriquez-Flores, S., Perez-Montfort, R., 2018. Differential effects on enzyme stability and kinetic parameters of mutants related to human triosephosphate isomerase deficiency. Biochimica et biophysica acta. General subjects. 1862**,** 1401-1409.

Carrillo, M., Wagner, M., Petit, F., Dransfeld, A., Becker, A., Erb, T. J., 2019. Design and Control of Extrachromosomal Elements in *Methylorubrum extorquens* AM1. ACS Synth. Biol. 8**,** 2451-2456.

Cleland, W. W., Andrews, T. J., Gutteridge, S., Hartman, F. C., Lorimer, G. H., 1998. Mechanism of Rubisco: The Carbamate as General Base. Chem. Rev. 98**,** 549-562.

Delaney, N. F., Kaczmarek, M. E., Ward, L. M., Swanson, P. K., Lee, M. C., Marx, C. J., 2013. Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. PloS one. 8**,** e62957.

Entner, N., Doudoroff, M., 1952. Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. J. Biol. Chem. 196**,** 853-62.

Erb, T. J., Berg, I. A., Brecht, V., Muller, M., Fuchs, G., Alber, B. E., 2007. Synthesis of C_5-dicarboxylic acids from C_2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proceedings of the National Academy of Sciences of the United States of America. 104**,** 10631-6.

Eren, A. M., Vineis, J. H., Morrison, H. G., Sogin, M. L., 2013. A filtering method to generate high quality short reads using illumina paired-end technology. PloS one. 8**,** e66643.

Gassler, T., Sauer, M., Gasser, B., Egermeier, M., Troyer, C., Causon, T., Hann, S., Mattanovich, D., Steiger, M. G., 2019. The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO_2. Nat. Biotechnol.

Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., Milo, R., 2019. Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO_2. Cell. 179**,** 1255-1263 e12.

Herz, E., Antonovsky, N., Bar-On, Y., Davidi, D., Gleizer, S., Prywes, N., Noda-Garcia, L., Lyn Frisch, K., Zohar, Y., Wernick, D. G., Savidor, A., Barenholz, U., Milo, R., 2017. The genetic basis for the adaptation of *E. coli* to sugar synthesis from CO_2. Nature communications. 8**,** 1705.

Hu, B., Lidstrom, M. E., 2014. Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production. Biotechnology for biofuels. 7**,** 156.

Knief, C., Frances, L., Vorholt, J. A., 2010. Competitiveness of diverse *Methylobacterium* strains in the phyllosphere of *Arabidopsis thaliana* and identification of representative models, including *M. extorquens* PA1. Microb. Ecol. 60**,** 440-52.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., Phillippy, A. M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27**,** 722-736.

Lolis, E., Alber, T., Davenport, R. C., Rose, D., Hartman, F. C., Petsko, G. A., 1990. Structure of yeast triosephosphate isomerase at 1.9-A resolution. Biochemistry. 29**,** 6609-18.

Marx, C. J., 2008. Development of a broad-host-range sacB-based vector for unmarked allelic exchange. BMC Res Notes. 1**,** 1.

Marx, C. J., Lidstrom, M. E., 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. Microbiology. 147**,** 2065-75.

Mueller-Cajar, O., Morell, M., Whitney, S. M., 2007. Directed evolution of rubisco in Escherichia coli reveals a specificity-determining hydrogen bond in the form II enzyme. Biochemistry. 46**,** 14067-74.

Parker, C., Peekhaus, N., Zhang, X., Conway, T., 1997. Kinetics of Sugar Transport and Phosphorylation Influence Glucose and Fructose Cometabolism by *Zymomonas mobilis*. Appl. Environ. Microbiol. 63**,** 3519-25.

Peel, D., Quayle, J. R., 1961. Microbial growth on C_1 compounds. 1. Isolation and characterization of *Pseudomonas* AM1. Biochem. J. 81**,** 465-9.

Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., Vorholt, J. A., 2009. Demonstration of the ethylmalonyl-CoA pathway by using 13C metabolomics. Proc. Natl. Acad. Sci. USA. 106**,** 4846-51.

Peyraud, R., Schneider, K., Kiefer, P., Massou, S., Vorholt, J. A., Portais, J. C., 2011. Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1. BMC systems biology. 5**,** 189.

Quinlan, A. R., Hall, I. M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26**,** 841-2.

Salmela, L., Rivals, E., 2014. LoRDEC: accurate and efficient long read error correction. Bioinformatics. 30**,** 3506-14.

Sambrook, J., Russel, D. W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schada von Borzyskowski, L., Carrillo, M., Leupold, S., Glatter, T., Kiefer, P., Weishaupt, R., Heinemann, M., Erb, T. J., 2018a. An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. Metab. Eng. 47**,** 423-433.

Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., Erb, T. J., 2015. A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria. ACS Synth. Biol. 4**,** 430-43.

Schada von Borzyskowski, L., Sonntag, F., Poschel, L., Vorholt, J. A., Schrader, J., Erb, T. J., Buchhaupt, M., 2018b. Replacing the Ethylmalonyl-CoA Pathway with the

Glyoxylate Shunt Provides Metabolic Flexibility in the Central Carbon Metabolism of *Methylobacterium extorquens* AM1. ACS Synth. Biol. 7**,** 86-97.

Schneider, K., Peyraud, R., Kiefer, P., Christen, P., Delmotte, N., Massou, S., Portais, J. C., Vorholt, J. A., 2012. The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by *Methylobacterium extorquens* AM1 during growth on acetate. J. Biol. Chem. 287**,** 757-66.

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

Smejkalova, H., Erb, T. J., Fuchs, G., 2010. Methanol assimilation in *Methylobacterium extorquens* AM1: demonstration of all enzymes and their regulation. PloS one. 5.

Sonntag, F., Kroner, C., Lubuta, P., Peyraud, R., Horst, A., Buchhaupt, M., Schrader, J., 2015. Engineering *Methylobacterium extorquens* for de novo synthesis of the sesquiterpenoid alpha-humulene from methanol. Metab. Eng. 32**,** 82-94.

Toyama, H., Anthony, C., Lidstrom, M. E., 1998. Construction of insertion and deletion *mxa* mutants of *Methylobacterium extorquens* AM1 by electroporation. FEMS Microbiol. Lett. 166**,** 1-7.

Ueoka, R., Bortfeld-Miller, M., Morinaka, B. I., Vorholt, J. A., Piel, J., 2018. Toblerols: Cyclopropanol-Containing Polyketide Modulators of Antibiosis in Methylobacteria. Angew. Chem. Int. Ed. Engl. 57**,** 977-981.

Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C., Peyraud, R., Robinson, D. G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C. J., Vorholt, J. A., Olson, M. V., Kaul, R., Weissenbach, J., Medigue, C., Lidstrom, M. E., 2009. Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. PloS one. 4**,** e5584.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., Earl, A. M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one. 9**,** e112963.

Watanabe, M., Zingg, B. C., Mohrenweiser, H. W., 1996. Molecular analysis of a series of alleles in humans with reduced activity at the triosephosphate isomerase locus. Am. J. Hum. Genet. 58**,** 308-16.

Weisser, P., Kramer, R., Sahm, H., Sprenger, G. A., 1995. Functional expression of the glucose transporter of *Zymomonas mobilis* leads to restoration of glucose and fructose uptake in *Escherichia coli* mutants and provides evidence for its facilitator action. J. Bacteriol. 177**,** 3351-4.

Weisser, P., Kramer, R., Sprenger, G. A., 1996. Expression of the *Escherichia coli pmi* gene, encoding phosphomannose-isomerase in *Zymomonas mobilis*, leads to utilization of mannose as a novel growth substrate, which can be used as a selective marker. Appl. Environ. Microbiol. 62**,** 4155-61.

Yang, Y. M., Chen, W. J., Yang, J., Zhou, Y. M., Hu, B., Zhang, M., Zhu, L. P., Wang, G. Y., Yang, S., 2017. Production of 3-hydroxypropionic acid in engineered *Methylobacterium extorquens* AM1 and its reassimilation through a reductive route. Microbial cell factories. 16**,** 179.

CHAPTER IV

Exploring in vivo implementation of synthetic carbon fixation cycles‡

‡First draft for publication.

Exploring *in vivo* implementation of synthetic carbon fixation cycles

Martina Carrillo[a], Thomas Schwander[a,b], Marieke Scheffen[a], Tobias J. Erb[a,c]

[a]Department of Biochemistry and Synthetic Metabolism, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany
[b]Competence Center for Biocatalysis, Institute of Chemistry and Biotechnology, Zürich University of Applied Sciences, Einsiedlerstrasse 31, 8820 Wädenswil, Switzerland.
[c]SYNMIKRO, LOEWE Center for Synthetic Microbiology, Universität Marburg, 35043 Marburg, Germany

First draft for publication.

Author contributions

T.J.E. conceived the project. M. C. and T. J. E. designed the research and wrote the manuscript with input from all authors. M. C., T. S., and M. S. performed experiments, analyzed or provided data.

Abstract

Our aim was to explore the *in vivo* implementation of synthetic CO_2 fixation cycles. We focus on the CETCH cycle that was recently described *in vitro* and we used the model organism *Methylorubrum extorquens* AM1 because it already utilizes a large number of the required enzymatic reactions within its native metabolism. In this study, we identify, characterize, and implement the missing steps of the CETCH cycles *in vivo*. Carbon acquisition exclusively via the CETCH cycle remains to be established and improved by adaptive laboratory evolution.

Introduction

While CO_2 is one of the drivers of global warming, it is also an overlooked source of carbon that could be exploited for a wealth of applications. Chemical CO_2 capture requires harsh conditions, yet nature has found mild and efficient solutions for biological CO_2 fixation using enzymes called carboxylases. Pathways that can incorporate inorganic carbon into central carbon metabolites as the sole carbon source enable organisms to grow autotrophically, *i. e.*, without assimilating organic carbon from their environment. Seven natural CO_2 fixation pathways, most of them cycles, have been described to date in phylogenetically and ecologically diverse organisms[1]. In this study, we focus on the implementation of a synthetic carbon fixation cycle.

The synthetic crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle relies on the carboxylation activity of crotonyl-CoA carboxylase/reductase (Ccr); an enzyme that uses reducing power to overcome the thermodynamic barrier for CO_2 fixation of enoyl-CoA esters[2]. Ccr was first characterized as part of the ethylmalonyl-CoA pathway (EMCP), an anaplerotic acetyl-CoA assimilation pathway common in α-proteobacteria lacking isocitrate lyase activity[3]. The two carboxylation steps of the EMCP allow the formation of glyoxylate, which is assimilated into biomass. Although the EMCP does not drive autotrophic growth in nature, its synthetic derivative CETCH could fulfill this role and was already successfully realized *in vitro*[2].

M. extorquens AM1 (formerly *Methylobacterium extorquens*) is a model organism whose genome has been sequenced[4], is genetically tractable[5, 6] and for which several genetic tools are available[7-12]. We utilize this bacterium for *in vivo* implementation as it relies on the EMCP for growth on C_1 and C_2 carbon sources. When growing on methanol (C_1) the EMCP assimilates acetyl-CoA formed by the serine cycle; the subsequently released glyoxylate serves as acceptor carbon backbone for C_1 unit assimilation (see Figure **1**). In the case of acetate (C_2), the EMCP is a linear pathway where the fixed glyoxylate is incorporated into malate for biomass formation (see Figure **2**). Knockouts upstream of *ccr* in the EMCP will abolish growth on C_1 and C_2 carbon sources. In turn, implementation of a functional CETCH cycle should restore growth of these strains. To achieve this goal we (i) characterized enzyme homologues to catalyze the reactions of the regeneration module for crotonyl-CoA (CO_2 acceptor), (ii) combined all steps in *M. extorquens* knockout strains for growth restoration on C_1/C_2 carbon sources, and (iii) explored co-feeding strategies selecting for a complete functional CETCH cycle for adaptive laboratory evolution (ALE) approaches in the future.

Results and Discussion

The implementation of the CETCH cycle into the *M. extorquens* metabolic network is described in Figures **1** and **2**. The first step was to find a succinyl-CoA reductase (Scr) that

would express well in *M. extorquens*. While the well-characterized homologue CkSucD of *Clostridium kluyveri* (ordered locus: CLK_3015) was used in the *in vitro* realization of this cycle[2, 13], we also investigated two other homologues, CdSucD from *Clostridioides difficile* 630 (ordered locus: CD630_23420) and PgSucD from *Porphymonas gingivalis* (ordered locus: PGN_0723), which, as far as we know, have not been biochemically characterized before. CkSucD had a lower k_{cat} of 2.5 ± 0.2 s^{-1} than previously reported, although it must be noted that it was purified from *Escherichia coli* instead of natively[13]. The *C. difficile* homologue had a k_{cat} of 28.1 ± 3.9 s^{-1} and the *P. gingivalis* homologue a k_{cat} of 2.7 ± 0.2 s^{-1} (see Figure **3**). Interestingly, both Clostridia homologues use NADPH as a cofactor, whereas PgSucD uses NADH to form succinic semialdehyde (SSA). All three homolog genes were expressed in *M. extorquens* using the inducible $P_{A1/O4/O3}$ promoter, with PgSucD showing the highest activity in cell lysates (see Figure **6a**).

The second step of the crotonyl-CoA regeneration module is the formation of 4-hydroxybutyrate (4HB) by a succinic semialdehyde reductase (Ssr). Initially, we used a codon-optimized version of the human enzyme, AKR7a2, as described previously[2, 14]. Although AKR7a2 can be produced in *E. coli*, as no eukaryotic post-translational modifications are needed, we could never detect Ssr activity in *M. extorquens* cell lysates expressing AKR7a2. We tested SSA reductases AtGR1 and Gox1801 from *Arabidopsis thaliana* and *Gluconobacter oxydans*, respectively, which showed good catalytic efficiencies in previous studies[15, 16]. Interestingly, AtGR1 uses NADPH as a cofactor while Gox1801 uses NADH and NADPH at equal rates. We expressed each under the control of a weak promoter, P_{fumC}, and detected 47 mU mg^{-1} total protein for AtGR1 and 226 mU mg^{-1} total protein for Gox1801 in *M. extorquens* cell lysates (see Figure **6b**).

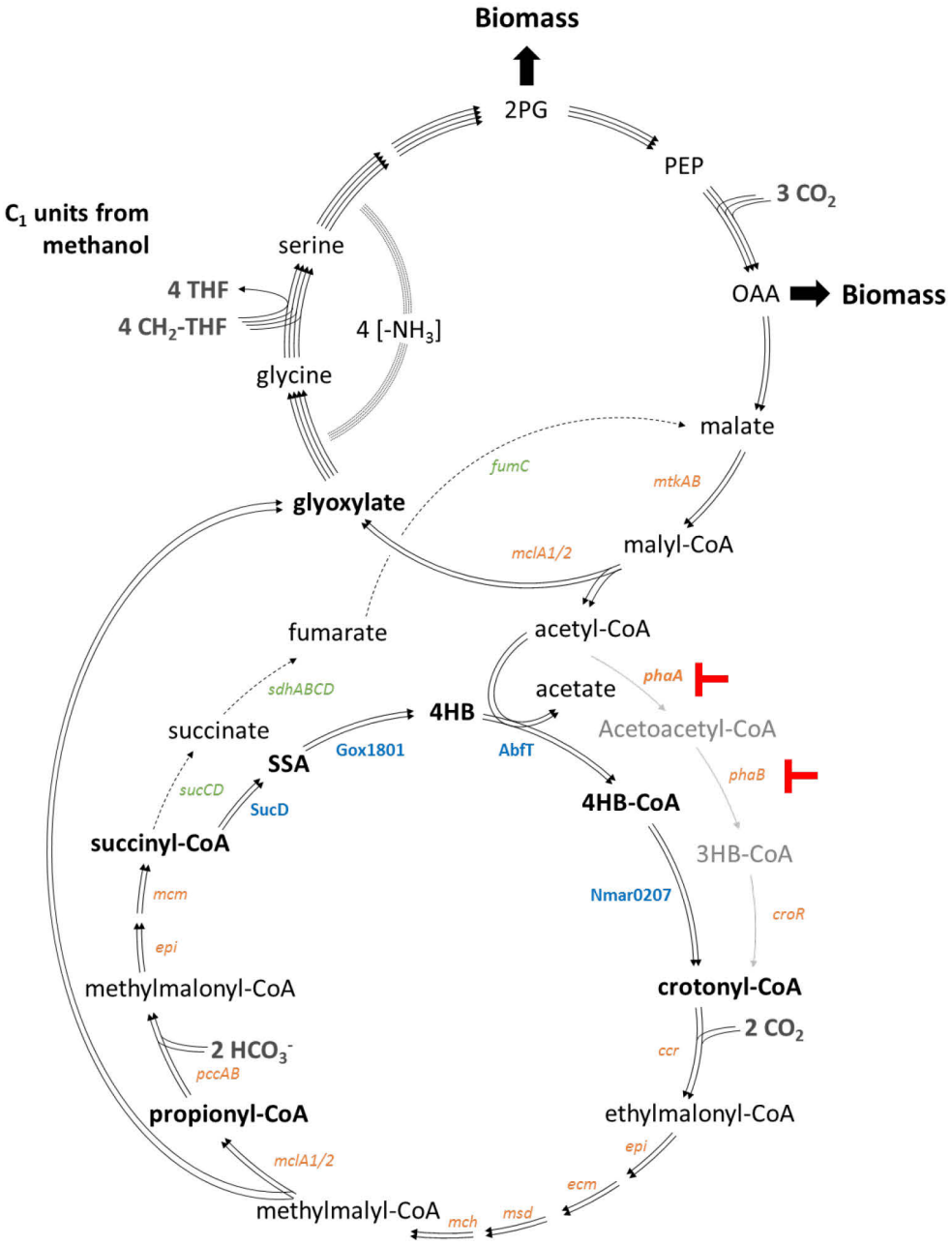


Figure 1. Implementation strategy of the CETCH cycle into the *M. extorquens* AM1 metabolic network during growth on methanol (C_1). A *phaA* or *phaB* knockout is proposed to prevent carbon assimilation into biomass. The regeneration module (shown in blue: SucD homologue, Gox1801, AbfT, Nmar0207) closes the CETCH cycle and should restore growth on methanol (C_1). Dashed lines represent metabolic reactions made redundant upon implementation of the CETCH cycle (genes in green). Abbreviations: CH2-THF: methylene tetrahydrofolate; THF: tetrahydrofolate; SSA: succinic semialdehyde, 4HB: 4-hydroxybutyrate, 4HB-CoA: 4-hydroxybutyryl-CoA. Genes: *phaA*: β-ketothiolase; *phaB*: acetoacetyl-CoA reductase; *croR*: crotonase; *ccr*: crotonyl-CoA carboxylase/reductase; *epi*: ethylmalonyl-CoA/methylmalonyl-CoA epimerase; *ecm*: ethylmalonyl-CoA mutase; *msd (ibd2)*: methylsuccinyl-CoA dehydrogenase; *mch*: mesaconyl-CoA hydratase, *mclA1/A2*: L-malyl-CoA/β-methylmalyl-CoA lyase; *pccAB*: propionyl-CoA carboxylase, *mcm*: methylmalonyl-CoA mutase; *mtkAB*: malate thiokinase; *sucCD*: succinyl-CoA synthetase; *sdhABCD*: succinate dehydrogenase complex; *fumC*: fumarase. For simplicity, cofactors are not shown.

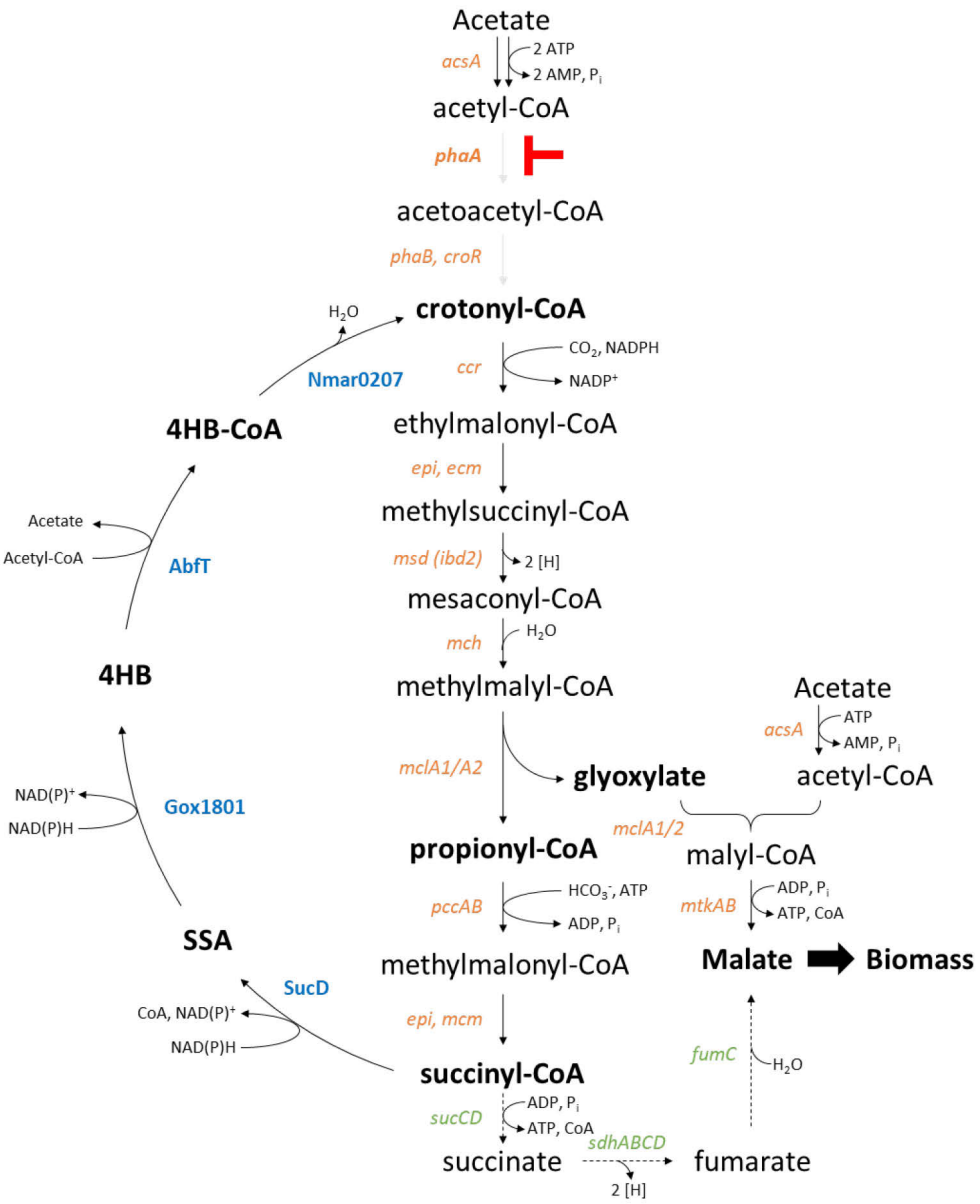
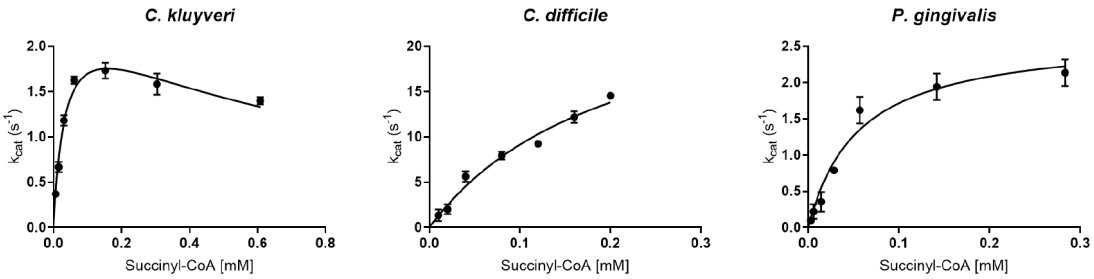


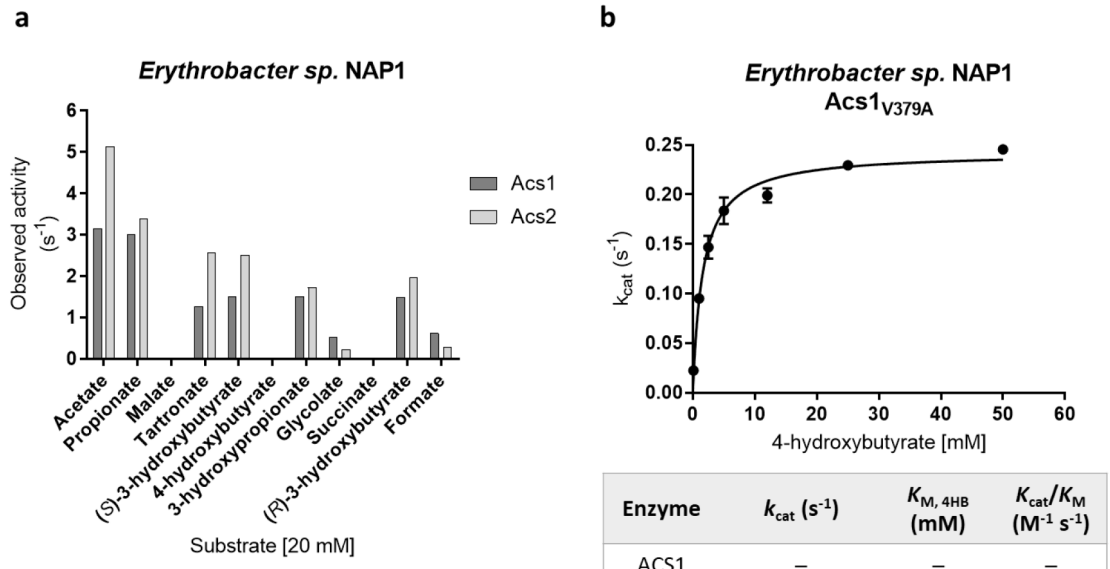
Figure 2. Implementation strategy of the CETCH cycle into the *M. extorquens* AM1 metabolic network during growth on acetate (C_2). A *phaA* knockout is proposed to prevent carbon assimilation into biomass. The regeneration module (shown in blue: SucD homologue, Gox1801, AbfT, Nmar0207) closes the CETCH cycle and should restore growth on acetate (C_2). Dashed lines represent metabolic reactions made redundant upon implementation of the CETCH cycle (genes in green). Abbreviations: SSA – succinic semialdehyde, 4HB - 4-hydroxybutyrate, 4HB-CoA – 4-hydroxybutyryl-CoA. Genes: *acsA*: acetyl-CoA synthetase; *phaA*: β-ketothiolase; *phaB*: acetoacetyl-CoA reductase; *croR*: crotonase; *ccr*: crotonyl-CoA carboxylase/reductase; *epi*: ethylmalonyl-CoA/methylmalonyl-CoA epimerase; *ecm*: ethylmalonyl-CoA mutase; *msd (ibd2)*: methylsuccinyl-CoA dehydrogenase; *mch*: mesaconyl-CoA hydratase, *mclA1/A2*: L-malyl-CoA/β-methylmalyl-CoA lyase; *pccAB*: propionyl-CoA carboxylase, *mcm*: methylmalonyl-CoA mutase; *mtkAB*: malate thiokinase; *sucCD*: succinyl-CoA synthetase; *sdhABCD*: succinate dehydrogenase complex; *fumC*: fumarase.



Organism	Enzyme	k_{cat} (s^{-1})	App. K_M (µM)	K_{cat}/K_M (M^{-1} s^{-1})	Cofactor
C. kluyveri	CkSucD (CKL_3015)	2.5 ± 0.2[a]	34 ± 6	7.29×10^4	NADPH
C. difficile	CdSucD (CD630_2342)	28.1 ± 3.9	208 ± 47	1.35×10^5	NADPH
P. gingivalis	PgSucD (PGN_0723)	2.7 ± 0.2	55 ± 12	4.84×10^4	NADH

Figure 3. Michaelis-Menten kinetics of Ssr homologues (SucD) from *C. kluyveri*, *C. difficile*, and *P. gingivalis*. The data points were analyzed using nonlinear regression in GraphPad Prism 7. [a]The *C. kluyveri* data set was analyzed for substrate inhibition and showed a K_i of 730 µM for succinyl-CoA. App. K_M – apparent K_M.

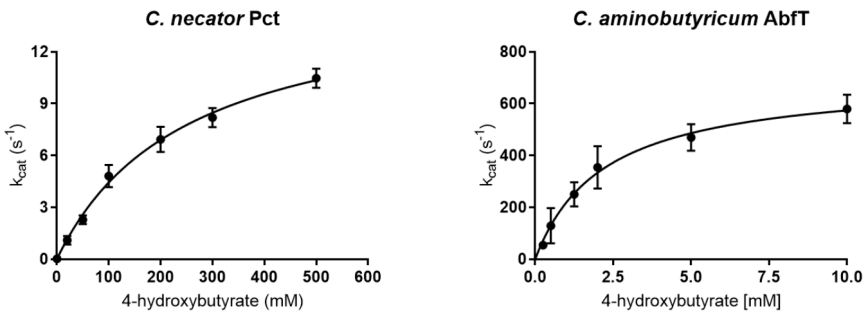
The activation of 4HB to the corresponding CoA ester can be carried out by a CoA synthetase or transferase. 4-hydroxybutyryl-CoA synthetases (Hbs) are exclusively found in the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle of autotrophic archaea. We first tested Nmar0206, an ADP-forming synthetase from *Nitrosopumilus maritimus*[17], but we could not detect activity in *M. extorquens* cell lysates (data not shown). Co-expression of *M. extorquens* and *E. coli* GroES-GroEL chaperons also did not yield any activity in cell lysates. We did not attempt to introduce more archaeal Hbs since those characterized originate from (hyper)thermophilic organisms[18]. Instead, we investigated various mesophilic CoA synthetases for promiscuity towards 4HB. First, we tested wildtype acyl-CoA synthetases: ScerAcs1p from *Saccharomyces cerevisiae*, EryAcs1 and EryAcs2 from *Erythrobacter sp*. NAP1, and EcPrpE from *E. coli*. However, we did not detect any activity towards 4HB. We also tested RlMatB from *Rhizobium leguminosarum*[19], *whose native substrate malonate is relatively bulky, hoping the active site might accommodate 4HB better. Neither wildtype MatB nor R298Q and R298K mutants showed any activity towards 4HB. The native malate thiokinase (MtkAB) was also tested for activity towards 4HB, but to no avail.* Interestingly, both Acs isozymes from *Erythrobacter sp.* NAP1 had a wide range of substrates, including 3-hydroxybutyrate but not 4HB, succinate, or malate (see Figure **4a**). A V379A mutation, shown to open the active site of *Salmonella enterica* Acs[20], was introduced into *Erythrobacter sp.* NAP1 Acs1, which allowed activation of 4HB with a k_{cat} of 0.243 ± 0.005 s^{-1} (see Figure **4b**). EryAcs1$_{V379A}$ has a catalytic efficiency of 150 M^{-1} s^{-1} with 4HB as substrate, which is not ideal for supporting growth, nevertheless, it was tested in *M. extorquens*. Under the weak promoter P_{fumC}, Hbs activity from EryAcs1$_{V379A}$ was not detectable over the background, while a stronger promoter caused cellular toxicity.



Enzyme	k_{cat} (s^{-1})	$K_{M, 4HB}$ (mM)	K_{cat}/K_M (M^{-1} s^{-1})
ACS1	–	–	–
$Acs1_{V379A}$	0.243 ± 0.005	1.6 ± 0.1	150

Figure 4. Putative acetyl-CoA synthetases of *Erythrobacter sp.* NAP1. Substrate specificity of Acs1 and Acs2 (**a**). Michaelis-Menten kinetics of $Acs1_{V379A}$ mutant towards 4-hydroxybutyrate (**b**). The data points were analyzed using nonlinear regression in GraphPad Prism 7. App. K_M – apparent K_M.

Next, we explored 4HB activation via the CoA transferases Pct from *Cupriavidus necator* and AbfT from *Clostridium aminobutyricum* (see Figure **5**). Pct showed a k_{cat} of 15.6 ± 0.9 s^{-1} albeit with an apparent K_M for 4HB of 250 ± 32 mM while AbfT had very good kinetic parameters - k_{cat} of 700 ± 50 s^{-1} and apparent K_M of 2.2 ± 0.4 mM - making it a promising candidate for *in vivo* implementation. Both transferases were characterized with excess acetyl-CoA as co-substrate although propionyl-CoA, amongst other CoA donors, is also accepted[21]. The *abfT* gene expressed very well in *M. extorquens*, with 33.3 U mg^{-1} total protein (with propionyl-CoA as donor) detected in cell lysates using the P_{coxB} promoter (see Figure **6c**).



Organism	Enzyme	k_{cat} (s^{-1})	App. K_M (mM)	K_{cat}/K_M (M^{-1} s^{-1})	CoA donor
C. necator	Pct	15.6 ± 0.9	250 ± 32	62	Acetyl-CoA
C. aminobutyricum	AbfT	700 ± 50	2.2 ± 0.4	3.18×10^5	Acetyl-CoA

Figure 5. Michaelis-Menten kinetics of CoA transferases towards 4HB from *C. necator* and *C. aminobutyricum* with acetyl-CoA provided as CoA donor. The data points were analyzed using nonlinear regression in GraphPad Prism 7. App. K_M – apparent K_M.

The final step of the regeneration module is the dehydration of 4-hydroxybutyryl-CoA (4HB-CoA) catalyzed by a 4-hydroxybutyryl-CoA dehydratase (Hbd), which yields crotonyl-CoA. Most Hbd homologues characterized thus far were very oxygen sensitive; however, the Hbd Nmar0207 from *N. maritimus* (ordered locus: Nmar_0207) was shown to be oxygen tolerant and originates from a mesophilic organism[17]. In *M. extorquens*, Hbd activity at 1.3 U mg^{-1} total protein was detected when *Nmar0207* was expressed under the control of the P_{mxaF} promoter (see Figure **6c**). This enzyme was therefore selected for further *in vivo* implementation attempts.

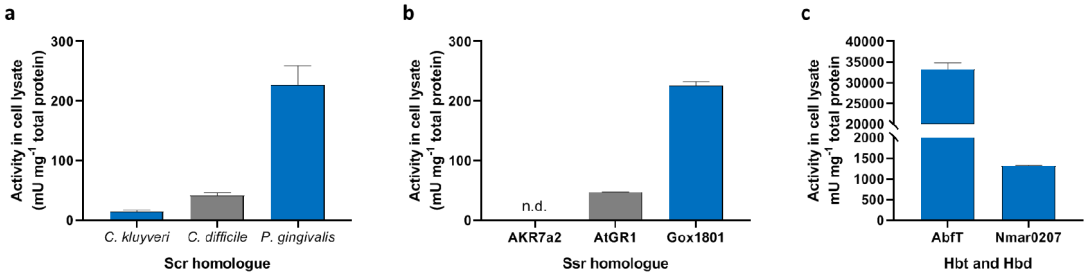


Figure 6. Specific activity of CETCH regeneration module enzymes in *M. extorquens* cell lysates. Genes encoding Scr homologues were expressed from the $P_{A1/O4/O3}$ promoter induced with 1 mM IPTG (**a**). Genes encoding Ssr homologues AKR7a2, AtGR1, and Gox1801 were expressed from the P_{fumC} promoter (**b**). Genes encoding the Hbt AbfT and Hbd Nmar0207 were expressed from the P_{coxB} and P_{mxaF} promoters, respectively (**c**). n.d. – not detected. Homologues shown in blue were used further.

Having identified genes and biochemically characterized the corresponding enzymes for each step of the regeneration module with suitable expression in *M. extorquens*, we proceeded to combine them in mutant strains deficient in growth on C_1/C_2 carbon sources. We first knocked out the *celABC* operon of previously generated *phaA*::kan and

phaB::kan mutants[22] to reduce clumping of cells for accurate long-term opticaldensity measurements, yielding Δ*cel* Δ*phaA*::kan and Δ*cel* Δ*phaB*::kan. These two strains were then equipped with a regeneration module consisting of CkSucD, Gox1801, AbfT, and Nmar0207 to restore growth on C_1/C_2 carbon sources. We fine-tuned gene expression for crotonyl-CoA formation via operon organization and by increasing the translation initiation rate (TIR) of CkSucD- which we believed to be limiting - from 12,000 to 19,000 (henceforth named +CETCH_12k and +CETCH_19k) according to the ribosomal binding site (RBS) calculator[23]. However, growth on methanol was not restored without the addition of external glyoxylate to the medium (see Figure **7**). Δ*cel* Δ*phaA*::kan +CETCH_12k and +CETCH_19k grew to higher cell densities compared to the empty vector control on methanol and 7.5 or 10 mM glyoxylate (see Figure **7a-c**). The adaptation period was also shorter; however, the doubling time (D_t) was still faster for the empty vector strain (see Table **1**). A similar behavior was observed for Δ*cel* Δ*phaB*::kan, where the +CETCH module significantly shortened the lag phase but did not improve the D_t. Interestingly, in this strain the lag phase was drastically shortened for +CETCH_12k and +CETCH_19k strains, but not for the empty vector, after re-inoculation in the same medium (see Figure **S1**), hinting at a need for the cells to adjust to the new environment. While these are already promising conditions for adaptive laboratory evolution (ALE), we wished to obtain a strain with tighter selection.

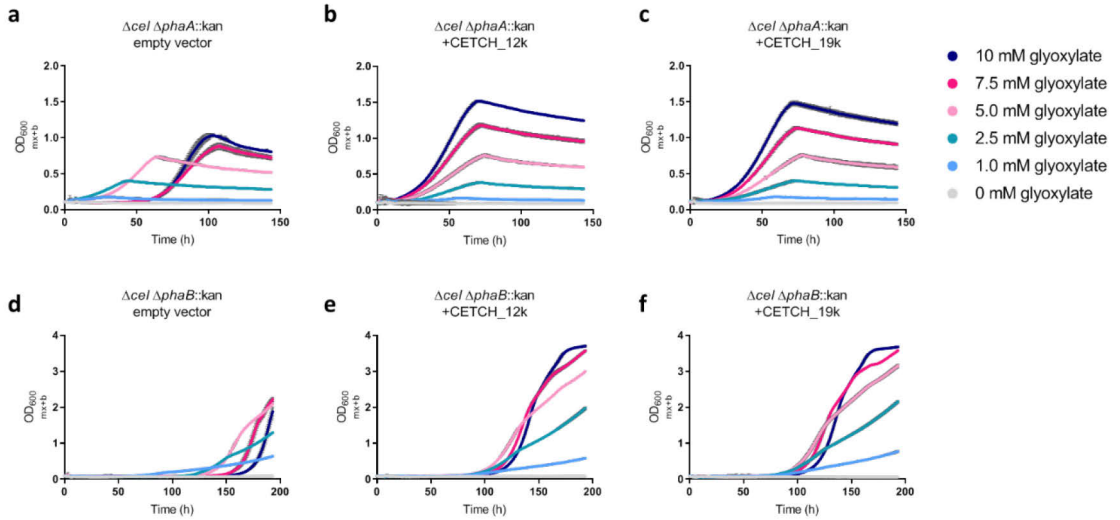


Figure 7. Growth profiles of Δ*cel* Δ*phaA*::kan (**a-c**) and Δ*cel* Δ*phaB*::kan (**d-f**) on 123 mM methanol and varying concentrations of glyoxylate. Empty vector (**a, d;** pTE1841), +CETCH_12k (**b, e;** pTE1843) and +CETCH_19k (**c, f;** pTE1845) cultures were all induced with 100 µM of IPTG. The RBS ahead of the gene encoding CkSucD was optimized for higher expression (12,000 or 19,000 TIR, respectively) according to the Salis RBS calculator v1.0.

Table 1. Comparison of the D_t (h) of *M. extorquens* strains grown on 123 mM methanol and glyoxylate.

Strain	External glyoxylate added			
	[2.5 mM]	[5 mM]	[7.5 mM]	[10 mM]
Δ*cel* Δ*phaA*::kan empty vector	21.15	16.35	12.87	10.79
Δ*cel* Δ*phaA*::kan +CETCH_12k	24.92	17.97	14.71	13.38
Δ*cel* Δ*phaA*::kan +CETCH_19k	25.38	17.71	14.00	13.02
Δ*cel* Δ*phaB*::kan empty vector	13.75	9.75	7.73	7.29
Δ*cel* Δ*phaB*::kan +CETCH_12k	16.50	12.21	9.58	9.44
Δ*cel* Δ*phaB*::kan +CETCH_19k	17.4	11.93	9.52	9.14

To that end, we tested compounds connected to the cycle to serve as helper carbon sources. 4HB and SSA would have been ideal candidates, since they are direct intermediates; however their supplementation did not support growth restoration via the CETCH cycle (data not shown). Propionate can be activated to propionyl-CoA and enter the lower part of the CETCH cycle, where it is interconverted to succinyl-CoA. *M. extorquens* CM2720 (D_t ≈ 19 h), and to a lesser extent *phaA* mutants (D_t > 40 h), can grow on 10 mM propionate (see Figure **S2** and Table **S4**). Yet, growth is abolished in *phaA* mutants in the presence of acetate as can be seen for empty vector strain (see Figure **8**, black). Δ*cel* Δ*phaA*::kan +CETCH_12k was able to grow on acetate supplemented with 3 mM propionate after a long adaptation period (see Figure **8a**, green). A control strain with Gox1801, AbfT and Nmar0207, but lacking SucD, the committing step of the crotonyl-CoA regeneration module, (termed w/o SucD_CETCH), grew faster and with a much shorter adaptation period (see Figure **8a**, pink). This suggests that propionate can be assimilated independently of a full CETCH module most likely due to AbfT-mediated activation of propionate. In this scenario, CkSucD represents a burden for the cell leading to the longer lag phase observed in +CETCH_12k strains. We attempted to keep carbon assimilation of propionate within the CETCH cycle by deleting the succinyl-CoA synthetase encoded by *sucCD* (META1p1540-41) in CM2720 *phaA*::0. The growth behavior of this strain when equipped with an empty vector, w/o SucD_CETCH, or +CETCH_12k, showed the same trend as before with an extended lag phase (see Figure **8b**). Given that AbfT does not show activity towards succinate[21], it is likely that the native malate thiokinase converts succinyl-CoA to succinate compensating for the loss of *sucCD* in this strain[24, 25]. In the current set-up, malate thiokinase is essential for assimilation of glyoxylate formed by the cycle (see Figure1 and Figure 2). Its activity could be circumvented by implementation of malate synthase from the glyoxylate cycle, which was established in various EMCP mutant strains[26], but not implemented in *phaA* or *phaB* deficient strains. As an alternative to *sucCD*, we knocked out the succinate dehydrogenase complex (encoded by *sdhABCD*) in

CM2720 *phaA*::0; however, this strain grew very poorly in the preculture on malate, and no growth was observed on acetate with or without propionate (see Figure **S3**).

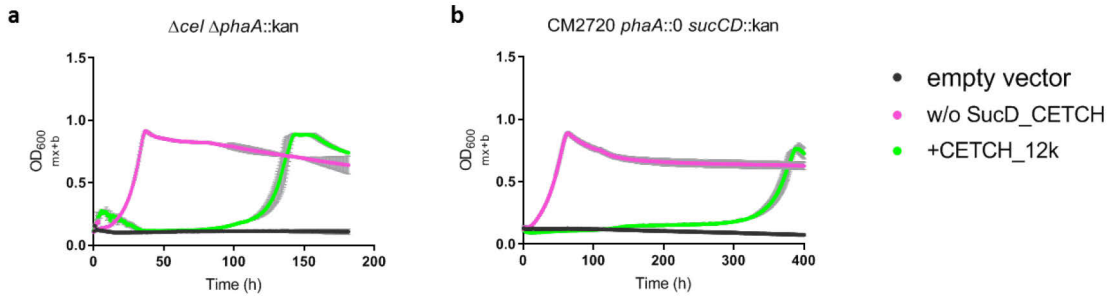


Figure 8. Growth profile of Δ*cel* Δ*phaA*::kan (**a**) and CM2720 *phaA*::0 *sucCD*::kan (**b**) grown on 10 mM acetate and 3 mM propionate. Empty vector (black, pTE1841), w/o SucD_CETCH (pink, pTE1858) and +CETCH_12k (green, pTE1843) cultures were induced with 1 mM of IPTG. The strain w/o SucD_CETCH contains Nmar0207, AbfT, and Gox1801 without any SucD. The RBS ahead of the gene encoding CkSucD in +CETCH_12k was optimized for 12,000 TIR according to the Salis RBS calculator v1.0.

Outlook

A tight selection strain based on propionate supplementation favoring a full CETCH module could not be identified so far. Implementation of the glyoxylate cycle or the β-hydroxyaspartate cycle[27] for glyoxylate assimilation in our selection strains could provide metabolic flexibility for successful implementation of the CETCH cycle. Transfer of the constructs created in this study to the recently established mini-chromosomes might also help by stabilizing gene expression without replicative burden[12]. In this study, methanol with glyoxylate supplementation in Δ*cel* Δ*phaA*::kan or Δ*cel* Δ*phaB*::kan was the most promising selection scheme since glyoxylate assimilation is independent of AbfT and the empty vector was not favored after re-inoculation. Overexpression of the native acetyl-CoA synthetase, which should be expressed at house-keeping levels, might also be beneficial to reactivate acetyl-CoA used by AbfT. Alternative feeding schemes with 3-hydroxybutyrate were not explored, but could also favor a full CETCH module. Any helper carbon sources, *e.g.* glyoxylate, propionate, etc., would be decreased over time during ALE to increase selective pressure until the CETCH module can provide cellular carbon from CO_2 fixation in *M. extorquens*.

Acknowledgements and Funding

We kindly thank Uwe Deppenmeier for providing us with pBBR1p264-gox1801-ST. We thank Karina Elgoff for cloning of pTE395. This work was supported by the European Research Council (Grant No. 637675 *SYBORG*), an IMPRS fellowship for M.C., and the Max Planck Society.

Materials and Methods

Materials

Succinic semialdehyde was purchased from Santa Cruz Biotechnology Inc. (Dallas, USA). Q5 DNA polymerase was obtained from NEB (Frankfurt am Main, Germany), FastAP and FastDigest restriction enzymes were obtained from Thermo Fischer Scientific (St. Leon-Rot, Germany), and used according to manufacturer's instructions. DNA oligos were obtained from Eurofins Genomics (Ebersberg, Germany) and Sigma-Aldrich (Darmstadt, Germany) (Table **S3**). Synthesized genes were obtained from Eurofins Genomics and GeneArt™ (Life Technologies GmbH, Darmstadt, Germany). Sanger sequencing was performed by Eurofins Genomics and Microsynth AG (Balgach, Switzerland).

Synthesis of CoA esters

Acetyl-CoA, succinyl-CoA and propionyl-CoA were synthesized from their respective anhydrides as described previously[28]. 4-hydroxybutyryl-CoA was synthesized enzymatically with purified Nmar0206 from 4-hydroxybutyrate, CoA, and ATP as described previously[17]. All CoA-thioesters were purified using a HPLC (1260 Infinity, Agilent Technologies GmbH) with a Gemini® 10µm NX-C18 110 Å Column (Phenomenex, Aschaffenburg, Germany) as described before[29]. The concentration of CoA-esters was quantified by determining the absorption at 260 nm (ε=16.4 mM^{-1} cm^{-1}). Sodium 4-hydroxybutyrate was synthesized from γ-Butyrolactone[17].

Bacterial strains and cultivation conditions

Bacterial strains used in this study are shown in Table S1. *M. extorquens* AM1 was routinely grown at 30°C in minimal medium with 123 mM methanol, 10 mM acetate, 30.8 mM succinate, 1-10 mM glyoxylate, 3-10 mM propionate, or 10 mM L-malate as specified. The corresponding sodium salts of acids were used. Minimal medium contained: 30.29 mM NH_4Cl, 0.81 mM $MgSO_4$, 20 mM K/Na-phosphate buffer pH 6.7 or 7.1, 1X Iron solution from a 1000-fold stock (final concentrations: 40.30 µM $Na_2EDTA{\cdot}2H_2O$, 10.79 µM $FeSO_4{\cdot}7H_2O$) and 1X trace element solution from a 1000-fold stock (final concentrations: 15.65 µM $ZnSO_4{\cdot}7H_2O$, 12.61 µM $CoCl_2{\cdot}6H_2O$, 5.09 µM $MnCl_2$, 16.17 µM H_3BO_3, 1.65 µM $Na_2MoO_4 \cdot 2H_2O$, 1.20 µM $CuSO_4{\cdot}5H_2O$, and 20.41 µM $CaCl_2{\cdot}2H_2O$) as described previously[30]. When appropriate, gene expression in *M. extorquens* was induced with 100-1000 µM IPTG. Antibiotics for selection were used accordingly: kanamycin 35 µg/mL for *M. extorquens* or 50 µg/mL for *E. coli*, gentamicin 7.5 µg/mL for *M. extorquens* or 30 µg/mL for *E. coli*, tetracycline 10 µg/mL for both, chloramphenicol 34 µg/mL for *E. coli*, ampicillin 100 µg/mL for *E. coli*, and streptomycin 50 µg/mL for *E. coli*. *E. coli* TOP10 and DH5α (Thermo Scientific) were routinely grown at 37°C in LB-Miller for cloning and plasmid amplification, except *sacB*-containing plasmids which were grown on LB-Lennox. Solid medium contained 1.5% (w/v) select agar.

DNA cloning and mutagenesis

Standard molecular techniques were used for amplification, cloning and transformation of DNA[31]. Detailed information on the construction of each plasmid is given in Table **S2**. Where stated, the TIR of an RBS was designed and/or calculated using the Salis RBS calculator v1.0 against *M. extorquens* AM1[23]. Point mutations were generated by QuikChange Site-Directed mutagenesis (Stratagene, La Jolla, USA). Oligos for side-directed mutagenesis were designed using PrimerX (https://www.bioinformatics.org/primerx/index.htm). Gibson isothermal reactions were performed as previously reported[32]. DNA isolation and purification were performed using NucleoSpin Plasmid and NucleoSpin Gel and PCR Clean-up kits (Macherey Nagel) according to manufacturer's instructions.

Strain construction

The *celABC* operon (META1p1167-9) was knocked out by allelic exchange with a synthetic construct containing the T_1T_2 terminator using pTE1191 in Δ*phaA*::kan and Δ*phaB*::kan[22]. The first recombination event was selected on solid medium with succinate and tetracycline thus creating an intermediate strain confirmed by colony PCR using primers MC195+MC196+MC35 and MC166+MC167. The second recombination event occurred with counter selection of the intermediate strain on solid medium with succinate and 5% sucrose without tetracycline. Single colonies were screened for tetracycline sensitivity, lack of clumping in liquid culture, and confirmed by colony PCR with primers MC195+MC196+MC209 and sequencing. Similarly, *phaA* (META1p3700) and *phaB* (META1p3701) genes were knocked out in CM2720[33] by allelic exchange with shortened peptides using pTE1839 and pTE1840, respectively. The first recombination event was selected on solid medium with succinate and tetracycline and confirmed by colony PCR using primers MC315+MC310+OS14 for pTE1839 and MC309+MC310+OS14 for pTE1840. The second recombination event occurred with counter selection of the intermediate strain on solid medium with succinate and 5% sucrose without tetracycline. Single colonies were screened for tetracycline sensitivity, inability to grow on C_1/C_2 carbon sources, and the shortened peptide was confirmed by colony PCR and sequencing with primers MC164+MC165. The succinate-CoA ligase encoded by *sucCD* (META1p1538-9) was knocked out by replacement with a loxP-flanked kanamycin resistance cassette using pTE1894. Disruption of the α-ketoglutarate dehydrogenase encoded by *sucAB* (META1p1540-41) or *sucCDAB* (META1p1538-41), using pTE1895 and pTE1896 respectively, was not possible in a *phaA*::0 strain background even with diaminopimelic acid (0.3 mM) supplementation. Colonies resistant to kanamycin and sensitive to tetracycline were screened for correct allelic replacement by colony PCR and sequencing using primers MC440+MC441+MC442. Succinate dehydrogenase encoded by *sdhABCD* (META1p3859-63) was knocked out by replacement with a loxP-flanked kanamycin resistance cassette using pTE2713. L-malate was used as a carbon source since succinate dehydrogenase activity is necessary for growth on succinate. Colonies resistant to

kanamycin and sensitive to tetracycline were screened for correct allelic replacement by colony PCR and sequencing using primers MC481+MC482+MC484 and MC481+MC483+MC484.

Colony PCR from M. extorquens

Colony PCR was carried out by suspending cells from a single colony in 20µL of water or 35µL of a 10% (v/v) solution of Chelex®100 (Sigma C7901; 50-100 mesh) and boiling cell suspensions at 95°C for 15 min. Boiled cell lysates were cooled on ice, spun down, and 1-2µL used as template DNA. Phusion Green Hot Start II High-Fidelity PCR Master Mix (Thermo Scientific F566S) was used for amplification. Each reaction contained 1X Master mix, template DNA as described above, 1X Q5 GC enhancer (NEB), and primers (0.05 µM each) in a final volume of 40 µL per reaction. The thermocycling conditions were as follows: initial denaturation at 98°C for 5 min, 26 cycles of: 98°C 30 sec, 55°C for 20 sec, and 72°C for 30 s/kb; and a final extension step at 72°C for 10 min. Annealing temperature was adjusted if necessary.

Expression and purification of recombinant proteins

Unless otherwise stated, proteins were produced in *E. coli* Rosetta™ 2 (DE3) pLysS (Novagen #71403) cells. Colonies from a fresh transformation were used to inoculate GLB medium (10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract, 0.4% (v/v) glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4) with the appropriate antibiotics. Expression cultures were grown at 37°C with shaking until an OD_{600} of 0.6-1.0, cooled to 25°C or 30°C and induced with 0.25 mM IPTG. Cells were harvested by centrifugation (4500 rcf, 10 min, 4°C) 4 h or 16 h after induction depending on the temperature after induction, respectively, and resuspended in the corresponding buffer (see below). If not used immediately, cell pellets were stored at −20°C. Nmar0206 (pET16b-Nmar0206) and Nmar0207 (pTE393) were purified as previously reported[2]. EryACS1 (pTE1007), EryACS2 (pTE1008), and $EryAcs1_{V379A}$ (pTE1417) were produced in *E. coli* BL21-AI with addition of 0.02% arabinose and 0.5 mM IPTG. PgSucD (pTE1851) was produced in *E. coli* TOP10 + pRARE induced with 1 mM IPTG.

His-tagged proteins: Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% (v/v) glycerol). Cells were lysed by sonication, the lysate was cleared by ultracentrifugation (50,000 rcf, 45 min, 4°C), and filtered through a 0.45 µm filter. The cleared lysate was loaded on a 1 mL HisTrapFF column (GE Healthcare, Freiburg, Germany). Unbound protein was removed with 20 ml of 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 75 mM imidazole. The protein was eluted in 20 ml of 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole. The elution fractions were pooled and desalted in 20 mM Tris-HCl pH 7.5, 150 mM NaCl with a 5mL HiTrap Desalting column (GE Healthcare).

StrepII-tagged proteins: Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol). Cells were lysed by sonication, the lysate was cleared by ultracentrifugation (50,000 rcf, 45 min, 4°C), and filtered through a 0.45 µm filter. The

cleared lysate was loaded on a 1 mL StrepTrap HP (GE Healthcare). Unbound protein was removed with 20 mL of 50 mM Tris-HCl pH 7.5, 150 mM NaCl. The protein was eluted in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2.5 mM desthiobiotin.

Elution fractions were pooled and concentrated in Amicon® Ultra 4 mL centrifugal filters with the appropriate cut-off (Merck). Concentration was determined on a NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA) using the extinction coefficient at 280 nm, as calculated by protparam (https://web.expasy.org/protparam/). Enzyme purity was confirmed by SDS-PAGE. The purified proteins were stored in 50% (v/v) glycerol at −20°C or −80°C.

Preparation of cell lysates

Appropriate cultures were harvested by centrifugation (3,000 rcf, 15 min, 4°C) and resuspended in 100 mM MOPS-KOH pH7.5, 10% (v/v) glycerol. If not used immediately, cell pellets were stored at −20°C. Subsequently, pellets were lysed by sonication. The crude extract was clarified by two centrifugation steps (17,000 rcf, 4°C, 15 and 30-60 min respectively). Total protein content clarified cell lysates was determined by the Bradford method using BSA as a standard[34].

Spectrophotometric assays

Spectrophotometic assays were performed on a Cary-60 UV/Vis spectrophotometer (Agilent) using 10mm quartz curvettes (Hellma, Müllheim, Germany). Kinetic parameters were determined using purified proteins where typically six different substrate concentrations were measured in triplicates and the data points were fitted using GraphPad Prism 7 to obtain apparent k_{cat} and K_M values. At least three triplicate values were measured from *M. extorquens* cell lysates. A unit is defined as the conversion of 1 µmol of substrate per min.

Activity assay of succinyl-CoA reductases (Scr)

For Michaelis-Menten kinetics:
CkSucD (CKL_3015): The activity was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 200 µM NADPH, and 1.44 µg purified CkSucD. Different concentrations of succinyl-CoA (6-600 µM) were used to start the reaction. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). No activity was observed with NADH.
CdSucD (CD630_23420): The activity was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 200 µM NADPH, 0.225 µg BSA and 0.225 µg purified CdSucD. Different concentrations of succinyl-CoA (10-200 µM) were used to start the reaction. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). No activity was observed with NADH.
PgSucD (PGN_0732): The activity was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 200 µM NADH, and 0.25 µg purified PgSucD.

Different concentrations of succinyl-CoA (2-283 µM) were used to start the reaction. NADH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). No activity was observed with NADPH.

In *M. extorquens* cell lysates: Scr activity in lysates was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 240 µM NADPH or NADH as needed, and varying amounts of the appropriate cell lysate. Succinyl-CoA (150 µM) was used to start the reaction. NAD(P)H consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). Background activity from the cell lysate was subtracted and values were normalized to total protein content.

Activity assay of SSA reductases (Ssr)

In *M. extorquens* cell lysates: Ssr activity in lysates was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 240 µM NADPH or NADH as needed, and varying amounts of the appropriate cell lysate. Succinic semialdehyde (0.1-30 mM) was used to start the reaction. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). Background activity from the cell lysate was subtracted and values were normalized to total protein content.

Activity assay of acyl-CoA synthetases

EryAcs1 and EryAcs2: The substrate specificity of EryAcs1 and EryAcs2 was determined using a coupled spectrophotometric assay for AMP formation. The assay was performed at 30°C in 100 mM MOPS-KOH pH 7.5 buffer containing 5 mM $MgCl_2$, 2.5 mM ATP, 1 mM CoA, 2.5 mM PEP, 0.4 mM NADH, 3 U pyruvate kinase, 4.6 U lactic dehydrogenase, 40 µg myokinase, and purified Acs1 or Acs2. The reactions were started with 20 mM of each carboxylic acid. NADH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}).

4HB:CoA synthetases (Hbs): Activation of 4HB was tested via 4HB-CoA product formation in a coupled assay using purified Nmar0207 and *Candida tropicalis* Etr1p[35] as coupling enzymes. Nmar0207 will dehydrate 4HB-CoA to crotonyl-CoA which is reduced using NADPH by Etr1p to butyryl-CoA. The assay was typically performed at 30°C in 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 5 mM ATP, 1 mM CoA, 200 µM NADPH, and excess purified Nmar0207 and Etr1p. The reactions were started with 4HB. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}).

EryAcs1$_{V379A}$: The Michaelis-Menten kinetic was measured as described for Hbs using purified Nmar0207 and Etr1p. The assay contained additionally 40.3 µg of purified EryAcs1$_{V379A}$ and 4HB (0.1-50 mM) was used to start the reaction.

In *M. extorquens* cell lysates: Hbs activity in lysates was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 5 mM ATP, 1 mM CoA, 240 µM NADPH, excess purified Nmar0207 and Etr1p, and varying amounts of the appropriate cell lysate. 4HB (20 mM) was used to start the reaction. NADPH consumption was followed

at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). Background activity from the cell lysate was subtracted and values were normalized to total protein content

Activity assay of 4HB:CoA transferases (Hbt)

4HB:CoA transferases (Hbt): Transfer of the CoA moiety to 4HB was tested via 4HB-CoA product formation in a coupled assay using purified Nmar0207 and *Candida tropicalis* Etr1p as coupling enzymes (see above). The assay was typically performed at 30°C in 100 mM MOPS-KOH pH 7.5 buffer containing 200 µM NADPH, 2 mM Acetyl-CoA, and excess purified Nmar0207 and Etr1p. For Pct 1.5 µg and for AbfT 3.9 ng of purified protein was added. The reactions were started with 4HB. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}).

In *M. extorquens* cell lysates: Hbt activity in lysates was determined at 30°C in 120 µl of 100 mM MOPS-KOH pH 7.5 buffer containing 10 mM $MgCl_2$, 3 mM propionyl-CoA, 240 µM NADPH, excess purified Nmar0207 and Etr1p, and varying amounts of the appropriate cell lysate. 4HB (20 mM) was used to start the reaction. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). Background activity from the cell lysate was subtracted and values were normalized to total protein content

Activity assay of 4HB-CoA dehydratase (Hbd)

In *M. extorquens* cell lysates: Hbd activity in lysates was determined at 30°C in 200 µL of 50 mM potassium phosphate buffer pH 7.5 containing 200 µM NADPH, excess purified Etr1p, and varying amounts of the appropriate cell lysate. 4HB-CoA (200 µM) was used to start the reaction. NADPH consumption was followed at 340 nm (ε = 6.22 mM^{-1} cm^{-1}). Background activity from the cell lysate was subtracted and values were normalized to total protein content.

Table S1: Strains used in this study.

Strain Name	Description	Reference
***Escherichia coli* DH5α**	F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$	Thermo Fischer Scientific
***E. coli* TOP10**	F- *mcrA* Δ*(mrr-hsdRMS-mcrBC) Φ80lacZΔM15* Δ*lacX74 recA1 araD139* Δ*(ara leu) 7697 galU galK rpsL* (Sm^R) *endA1 nupG*	Thermo Fischer Scientific
***E. coli* BL21-AI**	F$^-$*ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm araB*::T7RNAP-*tet*A	Thermo Fischer Scientific
***E. coli* Rosetta™ 2 (DE3) pLysS**	F$^-$ *ompT hsdS*$_{B(rB^- mB^-)}$ *gal dcm* (DE3) pLysSRARE2; Cm^R	Merck Millipore
Δ*phaA*::kan	*M. extorquens* AM1 *phaA*::kan (META1p3700); Km^R. Deficient in growth on C_1/C_2.	Korotkova & Lidstrom 2001[22]
Δ*phaB*::kan	*M. extorquens* AM1 *phaB*::kan (META1p3701); Km^R. Deficient in growth on C_1/C_2.	Korotkova & Lidstrom 2001[22]
Δ*cel* Δ*phaA*::kan	Δ*phaA*::kan with *celABC* (META1p1167-9) operon exchanged for T_1T_2; Km^R	This study
Δ*cel* Δ*phaB*::kan	Δ*phaB*::kan with *celABC* (META1p1167-9) operon exchanged for T_1T_2; Km^R	This study
CM2720	*Methylorubrum extorquens* AM1 strain deficient in cellulose production.	Delaney *et al.*, 2013[33]
CM2720 *phaA*::0	*phaA* (META1p3700) replaced by shortened peptide in CM2720.	This study
CM2720 *phaB*::0	*phaB* (META1p3701) replaced by shortened peptide in CM2720.	This study
CM2720 *phaA*::0 *sucCD*::kan	*sucCD* (META1p1538-9) replaced by loxP-kan-loxP in CM2720 *phaA*::0; Km^R	This study
CM2720 *phaA*::0 *sdhABCD*::kan	*sdhABCD* (META1p3859-63) replaced by loxP-kan-loxP CM2720 *phaA*::0; Km^R	This study

Sm^R – Streptomycin resistant, Spectinomycin sensitive; Cm^R – chloramphenicol resistant; Km^R – kanamycin resistant.

Table S2: Plasmids used in this study.

Plasmid Name	Description	Cloning strategy or comments	Reference
pBB541	*E. coli* GroEL GroES chaperones, $Spec^R/Sm^R$	n. a.	De Marco *et al*., 2007[36]
pEMatB-nHis	MatB in pET28b, Nter His tag, Km^R	n. a. MatB from *R. leguminosarum bv. trifolii*.	Townsend Lab (JHU)
pBBR1p264-gox1801-ST	Gox1801 from *G. oxydans*, Km^R	n. a.	Meyer *et al*., 2015[16]
pET-19b::pct	$P_{T7\text{-}lacO}$-Pct expression vector, Nter His tag, Amp^R	n. a.	Lindenkamp *et al*., 2013[37]
pCM184	Allelic exchange vector, Amp^R Tc^R Km^R	n. a. loxP-flanked kanamycin resistance cassette.	Marx and Lidstrom (2002)[8]
pCM433	Allelic exchange vector, Amp^R Tc^R Cm^R	n. a. *sacB* counterselection.	Marx (2008)[10]
pCA24N(-)-prpE	$P_{T5\text{-}lacO}$-PrpE expression vector, Nter His tag, Cm^R	n. a. ASKA collection minus GFP.	Kitagawa *et al*., 2005[38]
p2BP1	$P_{T7\text{-}lacO}$-AKR7a2 expression vector, Nter His tag, Amp^R	n. a.	Gift from Udo Oppermann
pTE100	Promoter-less, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al.*, 2015[11]
pTE102	P_{mxaF} promoter, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al.*, 2015[11]
pTE103	P_{fumC} promoter, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al.*, 2015[11]
pTE104	P_{coxB} promoter, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al.*, 2015[11]
pTE105	P_{tuf} promoter, oriV-traJ' origin, Tc^R	n. a.	Schada v. B. *et al.*, 2015[11]
pIND4	pMG160 origin, $P_{A1/O4/O3}$ promoter, empty vector, Km^R	n. a.	Ind *et al*., 2009[39]
pTE1841	$P_{A1/O4/O3}$ promoter, empty vector, pMG160 origin, Tc^R	n. a. KpnI sites surrounding *lacI* are still present.	Carrillo *et al*., 2019[12]
pTE1890	$P_{A1/O4}$ promoter; empty multiple cloning site with Strep-II tag, Km^R	n. a.	Carrillo *et al*., 2019[12]
pTE2714	$P_{A1/O4/O3}$ promoter, empty multiple cloning site with Strep-II tag, Km^R	n. a.	Carrillo *et al*., 2019[12]
pET16b-Nmar0206	$P_{T7\text{-}lacO}$-Nmar0206 expression vector, Nter His tag, Amp^R	n. a. Nmar0206 was codon-optimized for *E. coli*.	Könneke *et al*., 2014[17]
pET16b-Nm1309	$P_{T7\text{-}lacO}$-Nmar1309 expression vector, Nter His tag, Amp^R	n. a. Nmar1309 was codon-optimized for *E. coli*.	Könneke *et al*., 2014[17]

Plasmid Name	Description	Cloning strategy or comments	Reference
pTE260	$P_{T7\text{-}lacO}$-Etr1p expression vector, Nter His tag, Amp^R	n.a. *Candida tropicalis* Etr1p (GI: 29726216) codon-optimized for *M. extorquens*.	Rosenthal, Vögeli *et al*., 2015[40]
pTE380	Duet expression vector, Nter His tag CkSucD, *E. coli* YihU (untagged), $Spec^R/Sm^R$	n. a. CkSucD (CKL_3015) was codon-optimized for *E. coli*.	Schwander *et al*., 2016[2]
pTE393	$P_{T7\text{-}lacO}$ -Nmar0207_opt expression vector, Nter His tag, Amp^R	n.a. Nmar0207 was codon-optimized for *E. coli*.	Schwander *et al*., 2016[2]
pTE702	Codon-optimized Nmar0207 in pTE100, promoter-less, Tc^R	Backbone: AseI, MfeI digested pTE100; Insert: NdeI, EcoRI 1.628 insert of pTE393.	This study
pTE395	P_{mxaF}-Nmar0207_opt, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE102; Insert: XbaI, KpnI 1.688 bps insert of pTE702. Ligation with T4 DNA ligase.	This study
pTE680	pLEGO (A1); pET16b derivative with Nter Strep-II and Cter His tags, Amp^R	Backbone: NcoI, HindIII digested pET16b; Insert: NcoI, HindIII digested synthetic multiple cloning site bv1cut. Ligation with T4 DNA ligase.	Bastian Vögeli / Carrillo et al. / Scheffen et al. (unpublished results)
pTE1007	$P_{T7\text{-}lacO}$-EryAcs1 expression vector, Nter His tag, Amp^R	n. a. EryAcs1 (GI: 494288136) was codon-optimized for *E. coli* and cloned into pSEVA141 by the Joint Genome Institute (JGI).	JGI / This study / Scheffen et al (unpublished results)
pTE1008	$P_{T7\text{-}lacO}$-EryAcs2 expression vector, Nter His tag, Amp^R	n. a. EryAcs2 (GI: 494290475) was codon-optimized for *E. coli* and cloned into pSEVA141 by the JGI.	JGI / This study / Scheffen et al (unpublished results)
pTE1417	$P_{T7\text{-}lacO}$-EryAcs1$_{V379A}$ expression vector, Nter His tag, Amp^R	QuikChange mutagenesis of pTE1007 with prACS_Ery_V379A.	This study / Scheffen et al (unpublished results)
pTE1101	P_{coxB}-AKR7a2-CkSucD-T_{rrnB}; Tc^R	Backbone: HindIII, KpnI digested pTE104; Insert: HindIII, KpnI digested AKR7a2 and CkSucD operon. Ligation with T4 DNA ligase. Both genes codon optimized for *M. extorquens*.	This study
pTE1102	P_{coxB}-Nmar0206-T_{rrnB}; Tc^R	Backbone: HindIII, KpnI digested pTE100; Insert: HindIII, KpnI digested P_{coxB}-Nmar0206. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1108	P_{coxB}-AKR7a2-T_{rrnB}; Tc^R	XhoI digestion of pTE1101. 8.7 kb fragment self-ligated with T4 DNA ligation.	This study
pTE1109	P_{fumC}-Nmar0206-T_{rrnB}; Tc^R	Backbone: HindIII, KpnI digested pTE103; Insert: PCR product prMC40+MC35 of pTE1102. By Gibson assembly.	This study
pTE1110	P_{fumC}-AKR7a2-CkSucD-T_{rrnB}; Tc^R	Backbone: HindIII, KpnI digested pTE103; Insert: PCR product prMC40+MC35 of pTE1101. By Gibson assembly.	This study

Plasmid Name	Description	Cloning strategy or comments	Reference
pTE1119	P_{fumC}-Nmar0206-groS-groL (META1p3553-META1p3552); Tc^R	Backbone: SpeI, MunI digested pTE1109. Insert: AvrII, EcorI digested PCR product prMC53+MC54 from *M. extorquens* gDNA. Ligation with T4 DNA ligase.	This study
pTE1120	P_{fumC}-Nmar0206-groS-groL (META1p5240-META1p5239); Tc^R	Backbone: SpeI, MunI digested pTE1109. Insert: AvrII, EcorI digested PCR product prMC55+MC56 from *M. extorquens* gDNA. Ligation with T4 DNA ligase.	This study
pTE1121	P_{fumC}-Nmar0206-groES-groEL (*E. coli*); Tc^R	Backbone: SpeI, MunI digested pTE1109. Insert: SpeI, EcorI digested PCR product prMC57+MC58 from pBB541. Ligation with T4 DNA ligase.	This study
pTE1122	$P_{T7\text{-}lacO}$-AtGR1 expression vector, Nter Strep-II tag, Amp^R	Backbone: XbaI, EcoRI digested pTE680; Insert: XbaI, EcoRI AtGR1. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study / Carrillo et al / Scheffen et al. (unpublished results)
pTE1123	P_{fumC}-AtGR1, oriV-traJ' origin, Tc^R	Backbone: SpeI, PstI digested pTE103; Insert: XbaI, PstI AtGR1. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1125	$P_{T7\text{-}lacO}$-Gox1801 expression vector, Nter Strep-II tag, Amp^R	Backbone: NdeI, KpnI digested pTE680; Insert: NdeI, KpnI digested PCR product prMC60+MC61 of pBBR1p264-gox1801-ST. Ligation with T4 DNA ligase.	This study
pTE1126	P_{fumC}-Gox1801, oriV-traJ' origin, Tc^R	Backbone: 7.6 kb SpeI, KpnI digested pTE1123; Insert: XbaI, KpnI 0.9 kb insert of pTE1125. Ligation with T4 DNA ligase.	This study
pTE1127	P_{fumC}-CkSucD, oriV-traJ' origin, Tc^R	Backbone: HindIII, KpnI digested pTE103; Insert: PCR product prMC59+MC35 of pTE1101. By Gibson assembly.	
pTE1128	P_{fumC}-Pct, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE103; Insert: XbaI, KpnI digested PCR product prMC133+MC62 of pET-19b::pct. Ligation with T4 DNA ligase.	This study
pTE1130	P_{tuf}-AtGR1, oriV-traJ' origin, Tc^R	Backbone: SpeI, PstI digested pTE105; Insert: XbaI, PstI AtGR1. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1137	$P_{T7\text{-}lacO}$-ScAcs1p expression vector, Nter Strep-II tag, Amp^R	Backbone: NdeI, SpeI digested pTE680; Insert: NdeI, SpeI digested PCR product prMC65+MC66 of CEN.PK2 gDNA. Ligation with T4 DNA ligase.	This study
pTE1138	$P_{T7\text{-}lacO}$-AbfT expression vector, Nter His tag, Amp^R	Backbone: XbaI, KpnI digested pTE1122; Insert: XbaI, KpnI 1.4 kb insert of AbfT. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1140	P_{tuf}-Pct, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE105; Insert: XbaI, KpnI digested PCR product prMC133+MC62 of pET-19b::pct. Ligation with T4 DNA ligase.	This study
pTE1147	$P_{T7\text{-}lacO}$-$MatB_{R298Q}$ expression vector, Nter His tag, Km^R	QuikChange mutagenesis of pEMatB-nHis with prMC74+MC75.	This study

Plasmid Name	Description	Cloning strategy or comments	Reference
pTE1148	$P_{T7\text{-}lacO}$-$MatB_{R298K}$ expression vector, Nter His tag, Km^R	QuikChange mutagenesis of pEMatB-nHis with prMC76+MC77.	This study
pTE1155	P_{fumC}-$EryAcs1_{V379A}$, oriV-traJ' origin, Tc^R	Backbone: NcoI, SpeI digested pTE1123; Insert: NcoI, SpeI digested PCR product prMC133+105 of pTE1417. Ligation with T4 DNA ligase.	This study
pTE1156	P_{fumC}-Nmar1309, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE103; Inserts: PCR products prMC91+MC107 and prMC106+MC108 of pET16b-Nm1309. By Gibson assembly.	This study
pTE1163	P_{coxB}-AbfT, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE104; Insert: XbaI, KpnI 1.4 kb insert of AbfT. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1164	P_{tuf}-$EryAcs1_{V379A}$, oriV-traJ' origin, Tc^R	Backbone: NcoI, SpeI digested pTE1130; Insert: NcoI, SpeI digested PCR product prMC133+105 of pTE1417. Ligation with T4 DNA ligase.	This study
pTE1191	Vector for allelic exchange of *celABC* operon by T_1T_2 terminator; Amp^R, Tc^R, Cm^R	Backbone: AatII, SacI digested pCM433; Insert: PCR product prOS13+OS14 of $\Delta celT_1T_2$ synthetic construct. By Gibson assembly.	This study
pTE1814	P_{coxB}-CdSucD, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE104; Insert: XbaI, KpnI insert of CdSucD. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1815	P_{fumC}-CdSucD, oriV-traJ' origin, Tc^R	Backbone: SpeI, KpnI digested pTE103; Insert: XbaI, KpnI insert of CdSucD. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1816	Duet expression vector, Nter His tag CdSucD, *E. coli* YihU (untagged), $Spec^R$/Sm^R	Backbone: NcoI, SalI 4.5kb pTE380 fragment; Insert: NcoI, XhoI insert of CdSucD. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1832	Duet expression vector, Nter His tag CkSucD, *E. coli* YihU (untagged), $Spec^R$/Sm^R	Backbone: NcoI, SalI 4.5kb pTE380 fragment; Insert: NcoI, XhoI insert of CkSucD. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens* shown as $CkSucD_{Mex}$ in Table S5.	This study
pTE1839	Vector for allelic exchange of *phaA*::0 (META1p3700), shortened peptide has 17 Nter aa fused to 18 Cter aa; Amp^R, Tc^R, Cm^R	Backbone: BglII, SacI cut pCM433; Inserts: PCR products of 558 bps prMC305+MC306 and 703 bps prMC307+MC308 from *M. extorquens* gDNA. By Gibson assembly.	This study / Carrillo et al. (unpublished results)
pTE1840	Vector for allelic exchange of *phaB*::0 (META1p3701), shortened peptide has 6 Nter aa fused to 32 Cter aa; Amp^R, Tc^R, Cm^R	Backbone: BglII, SacI cut pCM433; Inserts: PCR products of 479 bps prMC299+MC300 and 645 bps prMC301+MC302 from *M. extorquens* gDNA. By Gibson assembly.	This study
pTE1843	$P_{A1/O4/O3}$-$CkSucD_{12k\_RBS}$-Nmar0207-AbfT-Gox1801, pMG160 origin, Tc^R	Backbone: XbaI, HindIII digested pTE1841; Insert: XbaI, HindIII 5.518 bp fragment of pTE1838. Ligation with T4 DNA ligase.	This study

Plasmid Name	Description	Cloning strategy or comments	Reference
pTE1845	$P_{A1/O4/O3}$-CkSucD$_{19k_RBS}$-Nmar0207-AbfT-Gox1801, pMG160 origin, Tc^R	Backbone: XbaI, BamHI 12.7 kb fragment of pTE1843; Insert: XbaI, BamHI digested PCR product pr324+325 of pTE1843. Ligation with T4 DNA ligase.	This study
pTE1851	$P_{A1/O4/O3}$-PgSucD, pMG160 origin, Tc^R	Backbone: XbaI, HindIII digested pTE1841; Insert: XbaI, HindIII 1.4 kb PgSucD insert. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE1858	$P_{A1/O4/O3}$-Nmar0207-AbfT-Gox1801, pMG160 origin, Tc^R	Backbone: XbaI, HindIII digested pTE1843; Insert: AvrIII, HindIII digested PCR product pr41+291 of pTE1843. Ligation with T4 DNA ligase.	This study
pTE1894	Disruption of *sucCD* (META1p1538-1539) genes via loxP-flanked kanamycin resistance cassette; Km^R Tc^R.	Backbone: EcoRI, SacI backbone of pCM184; Inserts: PCR product pr409+410 of *M. extorquens* gDNA, PCR product pr411+412 of *M. extorquens* gDNA, PCR product pr413+414 of pCM184. By Gibson assembly.	This study
pTE1895	Disruption of *sucAB* (META1p1540-1541) genes via loxP-flanked kanamycin resistance cassette; Km^R Tc^R.	Backbone: EcoRI, SacI backbone of pCM184; Inserts: pcr pr415+416 of *M. extorquens* gDNA, pcr pr417+418 of *M. extorquens* gDNA, pcr pr413+414 of pCM184. By Gibson assembly.	This study
pTE1896	Disruption of *sucCDAB* (META1p1538-) genes via loxP-flanked kanamycin resistance cassette; Km^R Tc^R.	Backbone: EcoRI, SacI backbone of pCM184; Inserts: PCR product pr409+410 of Mex gDNA, PCR product pr417+418 of Mex gDNA, PCR product pr413+414 of pCM184. By Gibson assembly.	This study
pTE2710	$P_{A1/O4/O3}$-CdSucD, pMG160 origin, Tc^R	Backbone: NcoI, HindIII digested pTE1841; Insert: NcoI, HindIII digested PCR product pr267+403 of CdSucD. Ligation with T4 DNA ligase. Gene codon optimized for *M. extorquens*.	This study
pTE2713	Deletion/disruption of *sdhABCD* (META1p3859-63 genes via loxP-flanked kanamycin resistance cassette; Km^R Tc^R	Backbone: EcoRI, SacI backbone of pCM184; Inserts: PCR product pr469+470 of *M. extorquens* gDNA, PCR product pr471+472 of *M. extorquens* gDNA, PCR product pr413+414 of pCM184. By Gibson assembly.	This study
pTE2715	$P_{A1/O4/O3}$ promoter, empty vector, pMG160 origin, Tc^R	. Backbone: 7.2kb EcoRI, HindIII fragment of pTE1841; Insert: 1.6 kb EcoRI, HindIII fragment of pTE2714. Ligation with T4 DNA ligase. KpnI is a unique cutter in the MCS	This study
pTE2716	$P_{A1/O4}$- CkSucD$_{100k_RBS}$, pMG160 origin, Km^R	Backbone: XbaI, KpnI digested pTE1890; Insert: XbaI, KpnI digested PCR product prMC473+MC476 of pTE1127. Ligation with T4 DNA ligase.	This study
pTE2717	$P_{A1/O4}$-CdSucD$_{100k_RBS}$, pMG160 origin, Km^R	Backbone: XbaI, KpnI digested pTE1890; Insert: XbaI, KpnI dig of pcr prMC474+MC289 of pTE2710. Ligation with T4 DNA ligase.	This study
pTE2718	$P_{A1/O4}$-PgSucD$_{100k_RBS}$, pMG160 origin, Km^R	Backbone: XbaI, KpnI digested pTE1890; Insert: XbaI, KpnI digested PCR product prMC475+MC289 of pTE1851. Ligation with T4 DNA ligase.	This study
pTE2720	$P_{A1/O4/O3}$-Nmar0207-AbfT-Gox1801$_{10k_RBS}$, pMG160 origin, Tc^R	Backbone: XbaI, PstI digested pTE2715; Inserts: PCR product pr57+477 of pTE1858, PCR product prMC478+MC479 of pTE1858. By Gibson assembly.	This study

Plasmid Name	Description	Cloning strategy or comments	Reference
pTE2721	$P_{A1/O4/O3}$-Nmar0207-AbfT-Gox1801$_{10k_RBS}$-$P_{A1/O4}$-CkSucD$_{100k_RBS}$, pMG160 origin, Tc^R	Backbone: AvrII, KpnI digested pTE2720; Inserts: AvrII, KpnI digested PCR product prMC480+MC289 of pTE2716. Ligation with T4 DNA ligase.	This study
pTE2722	$P_{A1/O4/O3}$-Nmar0207-AbfT-Gox1801$_{10k_RBS}$-$P_{A1/O4}$-CdSucD$_{100k_RBS}$, pMG160 origin, Tc^R	Backbone: AvrII, KpnI digested pTE2720; Inserts: AvrII, KpnI digested PCR product prMC480+MC289 of pTE2717. Ligation with T4 DNA ligase.	This study
pTE2723	$P_{A1/O4/O3}$-Nmar0207-AbfT-Gox1801$_{10k_RBS}$-$P_{A1/O4}$-PgSucD$_{100k_RBS}$, pMG160 origin, Tc^R	Backbone: AvrII, KpnI digested pTE2720; Inserts: AvrII, KpnI digested PCR product prMC480+MC289 of pTE2718. Ligation with T4 DNA ligase.	This study

n. a. – not applicable, gDNA – genomic DNA, Amp^R – ampicillin resistance, Cm^R – chloramphenicol resistance, Km^R – kanamycin resistance, $Spec^R$/Sm^R – spectinomycin and streptomycin resistance, Tc^R – tetracycline resistance, pr – primers, aa – amino acids, Nter – N terminus, Cter – C terminus.

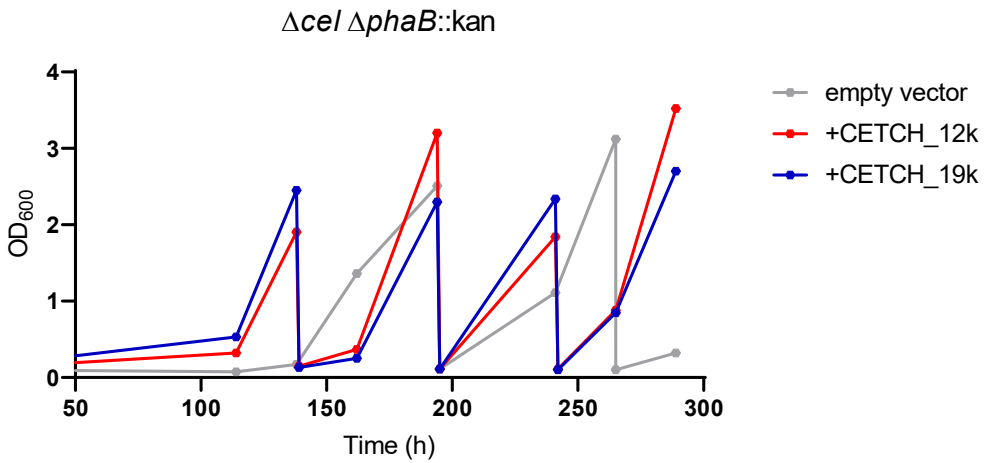
Table S3: DNA oligos used in this study‡.

Oligo Name	Sequence
MC26	CCATGGCAATTGAAGCTTCC
MC27	ATCTTCTTGGCCGCCTTCAC
MC31	ATGAGGAAACGGTCGAGAAG
MC32	CCAACTCGCCAAAGAGATAG
MC35	AACGACGGCCAGTGAATTAG
MC40	AGCCAACGAAGAGGCCACGAGCGTTTAACTTTAAGAAGGAG
MC41	GACAATCTGGTCTGTTTGTACCTAGGAAGAAGGAGATATAATTATGCGG
MC53	TGACGAATTCAGCCCAAACGACGGGAGTTC
MC54	AGCTCCTAGGCTTCAGGCGATCAGAAGTCC
MC55	AGCTGAATTCTGAAGACCAACCGCCATTCC
MC56	ATATCCTAGGTTAGAAGTCCATGCCGCCCATGC
MC57	GAGCGGATAACAATTTCACAC
MC58	TGCTACTAGTTTACATCATGCCGCCCATGC
MC59	AGCCAACGAAGAGGCCACGAGGTGGATACCTCGAGAATAAG
MC60	TGCACCATATGTCGAGTCCAAAGATCGG
MC61	CTGACGGTACCTGCGACTAGTTCATTTATGGGGAAGATTGGCAACG
MC62	CTGACGGTACCTGACGACTAGTGACAGGATTACAGGTGCAG
MC65	ACTGCATATGTCGCCCTCTGCCGTACAATC
MC66	CTGCCACTAGTCATTACAACTTGACCGAATC
MC74	CACGCCATTCTCGAGCAGTACGGCATGACGGAAAC
MC75	GTTTCCGTCATGCCGTACTGCTCGAGAATGGCGTG
MC76	GTCACGCCATTCTCGAGAAGTACGGCATGACGGAAAC
MC77	GTTTCCGTCATGCCGTACTTCTCGAGAATGGCGTGAC
MC83	GCTAGTTATTGCTCAGCGG
MC84	GCCTCGAGCAAGACGTTTCC
MC85	AACACTGGCAGAGCATTACG
MC90	AAATGGATGCGCGGCTGTTC
MC91	AAGCTTAGATCTTGAAGAAGGAGATATACCATGGG
MC105	TCAGACTAGTAGCCGGATCCTCGAGTTAAG
MC106	CGTCCATAGTGTGGTGATTG
MC107	CAATCACCACACTATGGACG
MC108	GACGGCCAGTGAATTAGGTACCTCTGACTATGTTTGACAGCTTATCATCG
MC133	TAATACGACTCACTATAGGG
MC164	ACGTCTAGAAATATCGCGCCGCCGAATGAG
MC165	CTAGGTACCGCGAGAGCCTTAGATCAGAC
MC166	AATGTGTAAAGGGCAAAGTG
MC167	CTGAACGGTCTGGTTATAGG
MC168	CGACACGGAAATGTTGAATAC
MC174	TCAGGTACCGCCATTGACGTTCACCTTCTC
MC175	GCATCTAGAGCACGCCTAGCGGGTATTTC
MC195	ACCAGCATCACCTCCAATTC
MC196	CGGCGTTGATGATCAGGCTC
MC201	GAGATGCATAGCAATGAAGTCTCTATCAAAGAAC
MC202	ACTCCTTAGGTTCCTAGGTCAGCCCCAGATCTCCTC
MC209	TGGCGAGCTTGAACCATTCC
MC218	AGTTCGTCCCGGTGTTCTTC
MC228	GAAACGCCTGGTATCTTTATAGTC
MC243	ACACCCGGAACCTCATTTCG
MC267	GTATCACGAGGCCCTTTCGTCT
MC289	TCCAAGCTCAGCTAATTAAG
MC291	GTCCAAGCTCAGCTAATTAAGCTTTCATTTATGGGGAAGATTG
MC299	AGTGCCACCTGACGTCTAGAGGACGACGAGAAGGTGAAC
MC300	ATCTCGTCGACGCGTTCCTGAGCCATAG
MC301	TATGGCTCAGGAACGCGTCGACGAGATCGCCCACGCGG

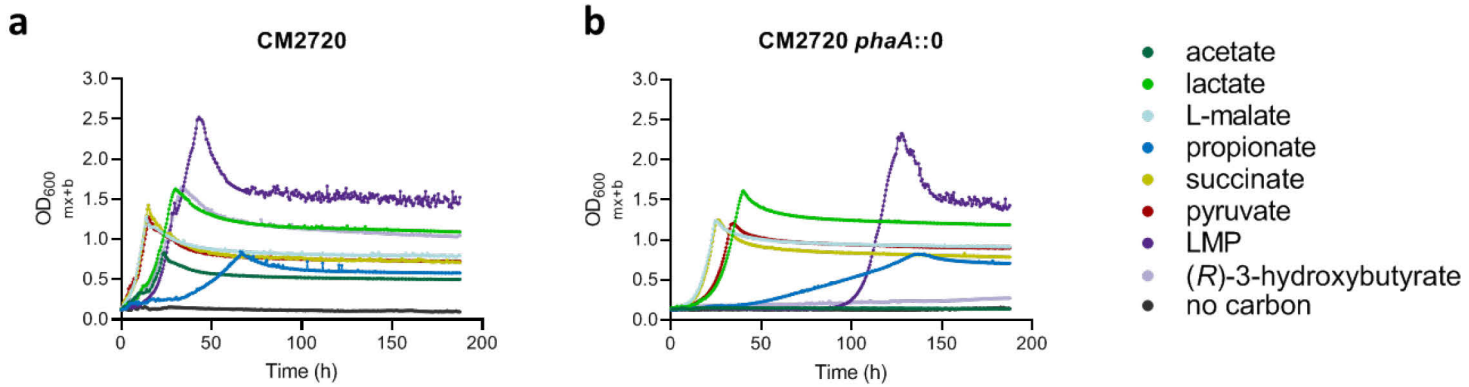
Oligo Name	Sequence
MC302	CGGCTGGATCCTCTAGTGAGCTGGATCGCGAAGAATTGCTC
MC305	AGTGCCACCTGACGTCTAGAAGCGTGCGTCGTAGACGATG
MC306	ATGCAGAGCGTGACGGGCGTACGCGCCGCACC
MC307	TACGCCCGTCACGCTCTGCATCGGCGGCGGCATG
MC308	CGGCTGGATCCTCTAGTGAGCTGTGCAGGTGAACATCGAGTC
MC309	GCCGTGAACAAGGACATGG
MC310	ATGGTCTTGCGGCGGACTTC
MC315	CCTGCTCGAAGATGATCTG
MC324	ACATCTAGATACGTTCGTTTAATTTAAGGTTTTTTATGGGCCATCACCACCATC
MC325	GACATTGGCCATGGGATCCTTATCG
MC403	AGTAAGCTTCTCGGTACCTGCAGGACTAG
MC409	CACCTGACGTCTAGATCTGAATTCTGGGCGCTGCAGAAGTTCTC
MC410	GAAGTTATGCGGCCGCCATATTTACGCGAGCTTGTCGTTG
MC411	CCCTACGTACGCGTGTTAACGCCGCAAGAAGCCGATGGTC
MC412	TCGGCTGGATCCTCTAGTGAGCTTCGAAGAAGCGCTGCCACTC
MC413	TATGGCGGCCGCATAACTTC
MC414	GTTAACACGCGTACGTAGGG
MC415	CACCTGACGTCTAGATCTGAATTCGCATCATGCCGGCCAACATC
MC416	GAAGTTATGCGGCCGCCATACTCTTCATCACCTGGACTTC
MC417	CCCTACGTACGCGTGTTAACCGCGGCCGATGATGTATCTG
MC418	TCGGCTGGATCCTCTAGTGAGCTTGCCCACGTCGATGCCCAAC
MC440	GGCCTGTTGAACAAGTCTGG
MC441	TCTCCATGCTGGAGCTGTTC
MC442	TCTCCTGGAAGCCGTAATGC
MC469	CACCTGACGTCTAGATCTGAATTCTCCTGTCGTCGGCGAGTTTC
MC470	GAAGTTATGCGGCCGCCATAATCTGTCTCATGCCAGTCTC
MC471	CCCTACGTACGCGTGTTAACGCTCTGCTTCAGGCCTATCG
MC472	TCGGCTGGATCCTCTAGTGAGCTAGGCGGCTTTCGACAGGAAC
MC473	GACTCTAGAACTGTCGCTAAAAGGAGGACATAATGGGCCATCACCACC
MC474	GACTCTAGAACTGTCGCTAAAAGGAGGACATAATGGGCAGCAGCCATC
MC475	GCTTCTAGAACTACGAGTCGTCAACAAACAATAACAAAATTAAAGGAGGTATTTTTATGGCCA GCTGGTCGCAC
MC476	AGAGGTACCTGCAGGACTAGTCCTTAGGTCAGCCCCAGATCTCCTC
MC478	AATAAACAGTATCAAGGAGGTAACTAATGGCGAGCTGGAGCCATC
MC479	AGCTCAGCTAATTAAGCTTACGGTACCTGCAGCCTAGGTCATTTATGGGGAAGATTG
MC480	GTACCTAGGGAGGCCCTTTCGTCTTCAC
MC481	TCTCAGCTATCTGCCGTATC
MC482	CGCGATGATTACGACGAAAG
MC483	CGACATGGCCCAGTTCAATC
MC484	AGCGAAACCTTGCAGCTTCC
ACS_Ery_V379A	CTGCGTCTGCTGGGCTCTGTAGGTGAGCCGATCAAC
OS12	ACCATCCATGAGCCGTATCG
OS13	CACATTTCCCCGAAAAGTGC
OS14	AGCCTGGTCGGCTGGATCCTC

‡Additional DNA oligos for sequencing purposes only are not shown.

SUPPLEMENTARY INFORMATION



Supplementary Figure S1. Growth profile of Δ*cel* Δ*phaB*::kan on 123 mM methanol and 7.5 mM glyoxylate. Empty vector (pTE1841), +CETCH_12k (pTE1843) and +CETCH_19k (pTE1845) cultures were all induced with 100 µM of IPTG. The cultures were re-inoculated into fresh medium at appropriate time points. The RBS of CkSucD expression was optimized for 12,000 or 19,000 TIR, respectively, according to the Salis RBS calculator v1.0.

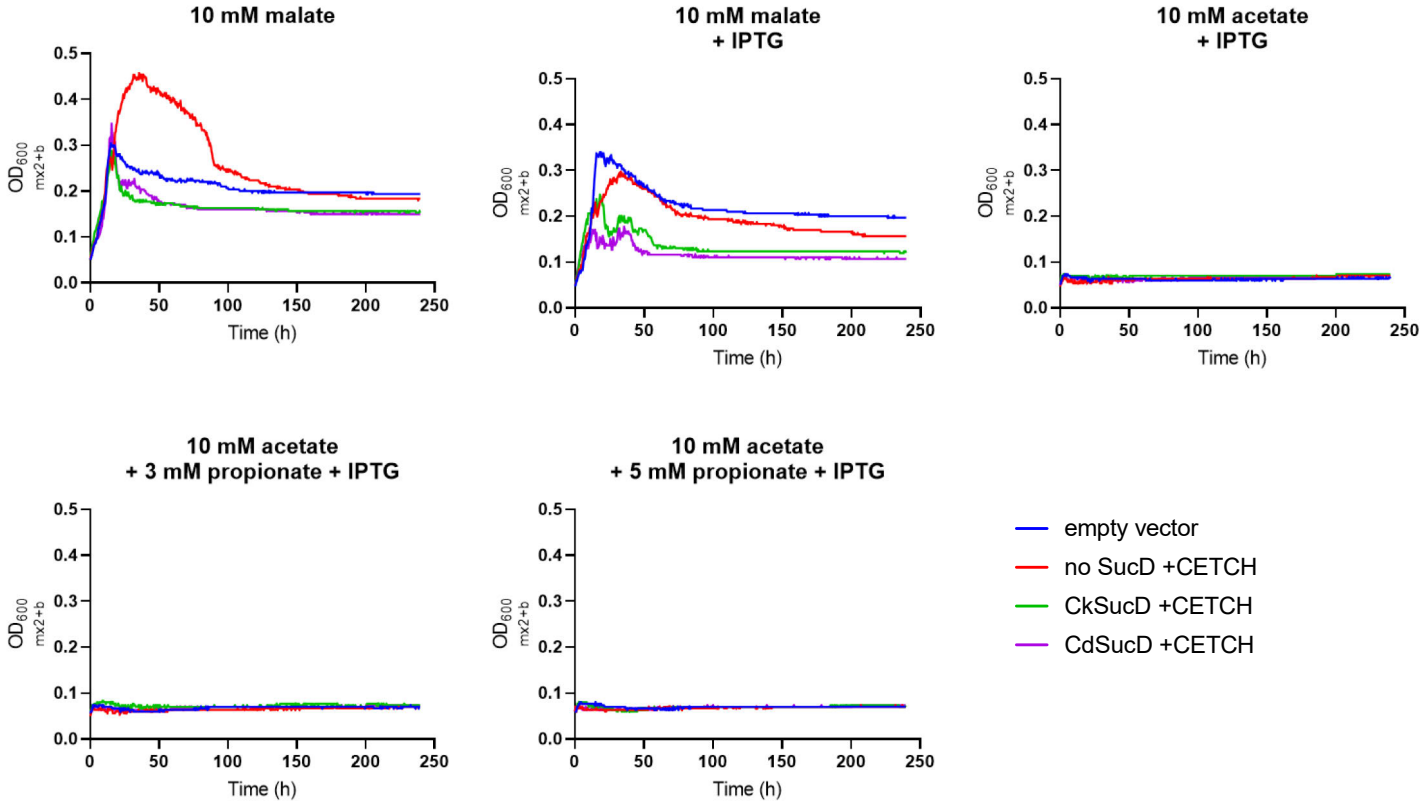


Supplementary Figure S2. Growth characterization CM2720 (**a**) and CM2720 *phaA* (**b**) on various organic acids. 10 mM of each carbon source was given using sodium salts when possible. LMP – lactate, L-malate and propionate (10 mM each). Sodium citrate was also tested but no growth was observed for either strain. As expected, *phaA*::0 is unable to grow on acetate. Growth of *phaA*::0 is also abolished on (*R*)-3-hydroxybutyrate (3HB) and impaired on propionate or LMP which was not expected and might suggest unknown regulatory effects.

Table S4: Doubling time (h) of CM2720 and its *phaA*::0 mutant on various carbon sources as shown in Supplementary Figure **S2**.

Carbon source	CM2720	CM2720 *phaA*::0
acetate	6.7	–
lactate	7.3	7.0
L-malate	4.1	5.3
propionate	19.8	40.9
succinate	4.1	5.2
pyruvate	4.4	6.7
LMP	5.0	6.4
(*R*)-3-hydroxyburyrate	8.2	n. d.

LMP: lactate, L-malate, and propionate; 10 mM each. No growth for CM2720 *phaA*::0 on acetate shown as –. n.d.: not determined.



Supplementary Figure S3. Growth characterization CM2720 *phaA*::0 *sdhABCD*::kan. Empty vector (pTE2715, blue), no SucD +CETCH (pTE2720, red), CkSucD +CETCH (pTE2721. green), and CdSucD +CETCH (pTE2722, purple). The growth of PgSucD +CETCH (pTE2723) is not shown given that it was severely impaired in the preculture on 10 mM malate due to higher expression of this homologue.

Table S5: Synthetic DNA or codon-optimized genes used in this study. Start and stop codons are underlined if applicable.

ΔcelT$_1$T$_2$

CACATTTCCCCGAAAAGTGCCACCTGACGTGAGCTCATGGTGTTTAAACCTACTCCTATCCCGGACAAGGCGTCGGCGCCCGCGTCACCGGCTACCAGCAGGAGGGCTCGCTGCTCACCGGCTA
CGAGTGGATCTGGCGGGAGGCGGCACTCGCCGGCTATGTCGGCTTCAACGTCCGCAGCAACCAGCTCTCGATCCCCGACCCCGGCAACCCGGTCGTCGGAACGGGAGTGGGCCTGAAGGTCG
CCGGCAACTTCTACGCAACGCCGACCGACCGGACCATGGTCTCGGCCTACGGCTCCTACTCGACCAAGTTCAACGCCTACTATTCCCGTTTCCGCGTCGGCTACATGGTCGCCGACGGCGTCTAT
ATCGGCCCGGAGGCGCTGTTCCTCGGCGACGACTTCTTCCGCCAGTACCGCGTCGGCGCCCATCTCTCGGGTTTGAGCTTCGGCCCCGTCCAGATGTCGCTCGCGGCGGGCTACGTGCGGGACC
GCGTCCAGGGCACGGGCTACTATTCCAGCATCGAGGCGCGCGCGAACTTTTAGAGCGCATTCCGACGGCGAGCAGGCCGCCGCTCGCGGCGGCGGCCGAAGTCAAGCGCTTGGATCAATGCG
GACAAGTTTAAACCCTATGGAAAAACGCCAGCAACGCGGCCTAGGTCCAATTGGTTCAATGCATATGCTCAGCTCGAGCGGCCGCGCCCTACGTACGCGTGTTAACCGGTTGACTAGTCCTGCA
GGTACCTAATTCACTGGCCGTCGTTTTACAAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAA
ATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTT
GCGTTTCTACAAACTCTTCCTCCCGGGTTAATTAAGCCCGGCATGGTCGATGAGAGCGCGCTACGCTACTACGCCTCGCAGAAGCAGACCGCTCGGATGAAGGCCGAGATCGCGCGGCTGAAG
CGGCTCTATCCCGGCTGGACCGAGCCGGGTGATCTCGACACCCTGCAACCGAGCCCGCCCGAAGAAGCGCCGCTGTGGGATCTGTTCACCGCCGGACGCTTCCAGGATCTGCGTGCAGCCATC
GCCGCGCGCCGCTCGCTCGAACCGAATTGGCAGCCCTCGGACGAACTCGCGCGCAAGCTCCGCCGGGCGGAGTTCCGCGCCAGCATCAAGGCCACCGCGGAGAAGAGCGGACCGGGCCGGA
ATGCCGACGAGGTGGTCGCCCTCTACCGCGCCGATTCCAGCGCCCTCGACACGACGGACCTCGAAAGTCTCTGGATCATCGCCGACGCGCTCGCCACGACGGGGGCGGGCGAGGACGCCTTCG
ATCTCTACAAGTCGGTCCTCGATGGTCCTGCCGATGCGGGCGCGCGGCTCGCCACGATCCAGAAGGCGATGAGCCATCTCAAGATGGATCAGGCCGACTTAATTAAGTAGCTCACTAGAGGATC
CAGCCGACCAGGCT

AbfT – Nter His tag codon optimized for *M. extorquens* AM1

TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCATGCATGATTGGAAGAAGATCTACGA
AGATCGCACTTGCACGGCTGATGAGGCCGTCAAGTCGATCAAGAGCGGGGATCGGGTTCTCTTCGCCCATTGCGTGGCCGAGCCGCCAGTCCTGGTAGAGGCTATGGTCGCGAATGCCGCGGC
ATATAAAAACGTCACCGTCAGTCACATGGTTACCCTCGGTAAGGGGGAGTATTCGAAGCCGGAATATAAGGAAAACTTCACGTTTGAAGGTTGGTTCACCAGCCCCTCTACCCGCGGTTCCATTG
CTGAAGGACACGGGCAGTTCGTCCCGGTGTTCTTCCACGAGGTGCCGAGCCTGATCCGGAAGGATATCTTCCACGTCGATGTCTTCATGGTGATGGTGTCCCCGCCCGATCACAATGGCTTCTGC
TGCGTCGGGGTCTCCTCCGACTACACCATGCAAGCCATCAAGAGCGCGAAGATCGTCCTCGCCGAAGTCAACGACCAGGTCCCTGTGGTCTACGGCGATACGTTCGTGCATGTCAGCGAGATCG
ATAAGTTCGTTGAGACCAGCCATCCGCTGCCCGAGATCGGCCTGCCCAAGATCGGCGAGGTCGAAGCCGCCATCGGCAAGCACTGTGCCAGCCTGATCGAGGACGGCTCGACCCTGCAATTGG
GCATCGGCGCCATCCCGGACGCCGTGCTTTCGCAGCTCAAAGACAAGAAGCATCTGGGTATCCACTCGGAGATGATCTCCGATGGCGTGGTGGATCTCTATGAAGCCGGCGTGATCGACTGCTC
CCAGAAGTCCATCGACAAGGGCAAGATGGCCATCACGTTCCTGATGGGCACCAAGCGCCTCTACGACTTCGCCGCCAACAACCCGAAAGTGGAGCTGAAGCCGGTGGACTACATCAACCACCC
GTCGGTGGTGGCCCAGTGCTCCAAGATGGTGTGCATCAATGCCTGCCTCCAGGTCGACTTCATGGGCCAGATCGTGAGTGACTCGATCGGCACGAAGCAGTTCTCCGGCGTGGGCGGCCAGGT
GGACTTCGTCCGCGGCGCCAGCATGTCGATTGACGGCAAGGGCAAGGCCATCATCGCGATGCCGTCGGTGGCCAAGAAGAAGGACGGCTCGATGATCAGCAAGATCGTCCCCTTTATCGACCA
CGGCGCGGCGGTCACCACCTCGCGCAACGACGCGGACTATGTCGTCACCGAGTACGGCATTGCGGAAATGAAAGGCAAGTCGCTCCAGGATCGTGCGCGCGCGCTCATTAACATCGCGCACCC
GGACTTCAAGGACGAGTTGAAGGCGGAGTTCGAGAAGCGGTTCAACGCGGCGTTCTGAGAATTCGAGCTCCGTCGACTAGTCCTGCAGGTACCTAATTCACTGGCCGTCGTTTTACAA

AtGR1 – Nter Strep-II tag codon optimized for *M. extorquens* AM1

TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGCGAGCTGGAGCCATCCGCAGTTTGAAAAAATTGAAGGCCGCCATATGGAGGTGGGTTTCCTCGGCCTCGGTATTATGGG
GAAGGCAATGTCGATGAATCTGCTGAAGAACGGATTTAAAGTCACCGTTTGGAACCGCACGCTCAGTAAGTGCGATGAGCTGGTCGAACACGGGGCGAGCGTCTGCGAAAGCCCCGCTGAGG
TCATCAAAAAGTGCAAGTACACGATCGCCATGCTCAGCGACCCGTGCGCGGCGCTCTCGGTGGTCTTCGATAAGGGCGGGGTGCTCGAACAGATCTGCGAGGGGAAGGGCTACATCGACATGT
CTACCGTCGATGCCGAAACCAGCCTCAAGATCAACGAGGCCATCACGGGCAAGGGCGGCCGCTTCGTCGAGGGCCCCGTCAGCGGCTCCAAGAAGCCCGCGGAGGACGGCCAGCTCATCATCC
TGGCGGCCGGCGATAAGGCCCTGTTCGAGGAGTCGATCCCGGCGTTCGACGTGCTTGGCAAGCGGTCCTTCTATCTGGGCCAGGTCGGCAACGGCGCCAAGATGAAGCTGATCGTGAACATGA
TCATGGGCTCGATGATGAACGCCTTCAGCGAGGGCCTGGTCTTGGCCGACAAGTCGGGCCTCTCGTCGGACACCCTGCTCGACATCCTCGACCTGGGCGCCATGACCAACCCCATGTTCAAGGG
CAAGGGCCCGTCCATGACCAAGTCCTCGTACCCGCCGGCGTTCCCGCTGAAGCATCAGCAGAAGGACATGCGCCTCGCGCTCGCGCTGGGCGACGAGAACGCCGTGTCCATGCCGGTGGCCGC
GGCCGCGAACGAGGCCTTCAAGAAGGCCCGCTCCCTCGGCCTGGGCGACCTCGACTTCTCGGCCGTGATCGAGGCCGTGAAGTTCTCGCGGGAGTGAACTAGTCCTGCAGGTACCGAATTC

Akr7a2-CkSucD operon – Nter His tags codon optimized for *M. extorquens* AM1

TGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCAATTGGTCGACCTCGAGTTAATTAACGTAAAGCTTGTCGACAATAATTTTGTTTAACTTTAAGAAGGAGA
TATACATATGCACCACCACCACCACCATTCGTCGGGCGTGGACCTGGGCACGGAGAACCTCTACTTCCAAAGCATGGAGATGGGTCGTCGAATGGATGCCCCCGCGTCCGCCGCGGCGGTTCGG
GCGTTCCTTGAGCGGGGGCACACCGAACTTGACACCGCTTTTATGTACAGCGACGGACAGTCGGAAACCATCTTGGGCGGTCTGGGCCTCGGGCTGGGCGGGGGGGGATTGCCGGGTGAAGAT
CGCCACAAAGGCGAATCCCTGGGACGGCAAAAGCCTCAAGCCGGATTCCGTGCGGTCGCAGCTGGAAACCTCCCTGAAGCGCCTCCAGTGCCCGCAGGTCGATCTCTTCTATCTGCATGCCCCG
GACCACGGCACACCCGTCGAAGAGACCCTGCACGCGTGCCAGCGCCTCCATCAGGAAGGCAAGTTCGTGGAGCTCGGCCTTTCCAACTATGCCTCGTGGGAGGTCGCCGAGATCTGCACGCTG
TGCAAGAGCAACGGCTGGATCTTGCCCACGGTGTATCAAGGCATGTACAACGCCACCACTCGCCAGGTCGAGACGGAGCTGTTCCCCTGCCTCCGCCACTTCGGCCTGCGTTTCTACGCCTACAA
CCCACTGGCGGGCGGGCTGCTCACAGGCAAATACAAGTACGAGGACAAGGATGGCAAGCAGCCGGTGGGCCGCTTCTTCGGTAACAGCTGGGCCGAGACCTACCGCAATCGCTTCTGGAAGG
AGCACCACTTCGAGGCCATCGCCTTGGTCGAAAAGGCCCTCCAGGCCGCGTACGGTGCCAGTGCACCGTCCGTGACCTCGGCGGCCCTCCGCTGGATGTACCATCATTCGCAGCTGCAAGGCGC
GCACGGCGACGCCGTGATCCTCGGCATGTCGTCGCTGGAGCAGCTGGAGCAGAACCTGGCCGCGACCGAGGAGGGCCCGCTGGAGCCGGCCGTCGTCGATGCGTTCAACCAGGCGTGGCACC
TCGTCGCTCACGAGTGCCCCAACTACTTCCGCTGACAGTTAAAGGTGGATACCTCGAGAATAAGGAGATATACCATGGGCCATCACCACCATCATCATCACCACCATCATTCGAGTGGACACATC
GAGGGTCGTCACATGAGCAATGAAGTCTCTATCAAAGAACTCATCGAGAAGGCGAAGGTCGCTCAGAAGAAGCTGGAAGCATATAGCCAAGAGCAGGTAGACGTTCTCGTGAAGGCCCTGGG
CAAGGTGGTGTACGACAACGCTGAGATGTTCGCCAAGGAAGCGGTCGAGGAGACTGAGATGGGCGTCTATGAGGACAAGGTCGCGAAGTGCCATCTGAAATCCGGGGCCATCTGGAACCACA
TTAAGGACAAAAAAACCGTCGGGATCATCAAGGAGGAGCCCGAACGGGCTCTGGTCTACGTGGCCAAGCCGAAGGGCGTGGTAGCCGCGACGACCCCGATCACGAACCCCGTCGTCACCCCG
ATGTGCAACGCTATGGCGGCGATCAAGGGCCGCAATACAATCATCGTGGCGCCCCACCCGAAGGCGAAGAAGGTGTCCGCCCATACCGTCGAGCTCATGAACGCCGAGCTCAAAAAGCTCGGC
GCGCCGGAGAACATCATCCAGATCGTCGAGGCCCCGTCACGGGAGGCCGCCAAGGAACTGATGGAATCGGCCGACGTCGTCATCGCCACCGGCGGCGCGGGCCGCGTGAAGGCGGCGTACA
GCAGCGGGCGCCCGGCGTACGGCGTCGGCCCGGGCAACTCGCAGGTGATTGTTGACAAGGGCTACGACTATAACAAGGCGGCCCAGGATATCATCACGGGCCGGAAGTATGACAACGGCATT
ATCTGCTCGTCGGAACAGAGTGTCATCGCGCCGGCCGAAGACTACGACAAGGTGATCGCCGCCTTCGTCGAGAACGGTGCCTTCTACGTGGAGGATGAGGAAACGGTCGAGAAGTTCCGCTCC
ACGCTGTTTAAAGACGGGAAGATCAACTCCAAGATCATCGGCAAGTCCGTCCAGATCATTGCCGATCTCGCCGGCGTCAAGGTGCCCGAGGGCACCAAGGTGATCGTGCTCAAGGGAAAGGGC
GCCGGCGAGAAGGACGTGTTGTGCAAGGAGAAAATGTGCCCCGTGCTCGTCGCCCTCAAGTACGATACCTTCGAGGAGGCCGTCGAGATCGCAATGGCAAACTACATGTACGAAGGTGCCGG
CCACACCGCCGGTATCCACTCGGACAACGACGAGAACATCAGGTACGCCGGCACGGTGCTGCCTATCTCGCGCCTTGTCGTGAACCAGCCGGCCACCACCGCCGGGGGCTCGTTCAACAACGG

CTTCAACCCGACCACCACTCTGGGCTGCGGGTCCTGGGGCCGCAACTCAATCAGCGAGAACCTTACCTACGAACACCTCATCAACGTGTCGCGCATCGGCTACTTCAACAAGGAGGCCAAGGTT
CCCTCGTACGAGGAGATCTGGGGCTGACTCGAGCAATTGATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATC
CGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTGCTAG
CGAATTCTGTACAACTAGTACATGCGGTACCGTGGGCGATCGCTCTAGAGCTAGCGAATTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC

EryACS1 – Nter His tag codon optimized for *E. coli*

CATATGACCGGTTTCGTAGAGCGTCCGGAACAGGCTCACACCCCGAACTGCACAGGTGTTCAATACGCAGCAATGTACGAGCGTTCCCTGGCTGATCCGGACGGCTTTTGGCTGGAACAGGCTA
AACGTCTGGACTGGACCCAGCAGCCGCGTAAGGGCGGCGAATGGTCTTACGACCCGGTTGACATAAAATGGTTCGCGGACGGTTCCCTGAACCTGTGTCATAACGCAGTAGACCGTCACCTGG
ATAGCCGCGGTGACACCCCGGCAATCATCTTCGAGCCGGATGACCCGGCTACTCCGTCTCGCACCCTTACCTATCGTCAGCTGCACTCCGAAGTTATTCACATGGCTAATGCTCTGAAAGCAATC
GGCGTAACCAAAGGCGAGCGTGTTACAATCTACATGCCGAACATCGTAGAAGGTGTTACCGCTATGCTGGCTTGCGCACGCCTGGGTGCGATCCACAGTGTAGTGTTCGGTGGTTTCTCTCCTG
AGGCTCTGGCCGGTCGTATCATCGACTGTGAATCTCGCTTCGTTGTGACTGCGGACGAAGGCAAACGCGGTGCAAAATCCGTTCCGCTGAAGGCTAACGTAGACGCTGCTCTGGAAGTTGAAG
GCGTAGACGTTACTGGTGTTCTGGTAGTACAGCATACCGGCCTGGCGGTTCCGATGACCGAGGGTCGCGACCACTGGTTCCATGAGGTTAAATCTGACGCTGACGTGCCGTGCGAAACTATGGC
AGCTGAAGACCCGCTGTTTATTCTGTACACAAGCGGCTCCACCGGTAAACCAAAAGGTGTTCTGCACACTACTGGTGGTTATGGTGTTTGGACTGCTACCACCTTCAGCTACATCTTCGACTACCA
GCCGGGTGAAGTGTTCTGGTGTACAGCAGACATCGGTTGGGTTACCGGTCATTCTTACATCGTATATGGTCCGCTGCAGAACGGTGCAACTCAGGTTCTGTTTGAAGGTGTACCAAACTACCCG
GACTTCGGTCGTTTCTGGGACGTTGTAGCTAAACACAAAGTTTCCATCCTCTACACTGCACCGACCGCTATCCGCGCGCTGATGCGTGAAGGTGATGACTATGTTACTTCACGTGACCGCAGTTCC
CTGCGTCTGCTGGGCTCTGTAGGTGAGCCGATCAACCCGGAGGCCTGGCGTTGGTACTTCGATGTTGTTGGTGAAGGTCGTTGTCCGATCATCGACACTTGGTGGCAGACCGAGACTGGCGGTT
GCATGATTACTACTCTGCCGGGTGCTCACGATATGAAGCCGGGCTCTGCAGGCCTGCCGATGTTCGGTATCCGTCCGCAGCTCGTTGACAACGACGGTGCAGTGCTGGACGGTGCAACAGAAG
GTAACCTGTGTATCACCCACTCTTGGCCGGGCCAGGCTCGCTCTGTTTACGGTGATCACGACCGCTTCGTACAGACTTACTTCTCCACCTACTCCGGCAAGTATTTCACTGGTGACGGTTGCAAAC
GTGACGAGGATGGCTACTATTGGATTACCGGTCGTGTGGACGACGTTATCAACGTTTCTGGCCATCGTATGGGTACTGCTGAGGTGGAAAGCGCACTGGTACTGCACCCGCAGGTAGCAGAAG
CAGCTGTGGTTGGTTACCCGCACGACGTTAAGGGCCAGGGTATCTACTGTTATGTAACCACCAACGCAGGCGTGGAAGGTTCTGACGAACTGTACCAGGAGCTGCGTGCCCACGTGCGCAAAG
AAATCGGCCCGATCGCAACTCCGGATCAGATCCAGTTCACTGACGGTCTGCCGAAGACGCGTTCCGGTAAAATTATGCGCCGTATCCTGCGTAAAGTGGCAGAAAACGACTATGGTTCTCTGGG
TGATACTTCCACCCTGGCGGACCCGTCTCTGGTGGATCGTCTGATTGAGGGCCGTCAGAAAACTTAACTCGAGGATCC

EryACS2 – Nter His tag codon optimized for *E. coli*

CATATGATCCATGCTGTTCAGCGTCCGCTGAGCTCTAACGCCACCCACTGCGTGACCGAGCAGTACGGCGAGCTGTACCGCCGTTCCGTAGAGACTCCTGATGCATTCTGGCTGGAACAGGCTG
AACGTCTGGACTGGTTCACCAAACCGGCAAAGGGCGGTGAGTGGTCTTATGACCCGGTGGATATCCAGTGGTTTGCTGACGGCACCCTGAACCTGTGTTATAATGCAGTAGACCGCCACCTGGA
AACCCGTGGTGACGACACTGCGATCATCTTCGAACCAGACGACCCGGCTTCCCCGAGTCGTACCCTGACCTACCGCGACCTGCATGGTGAGGTGGTACGCATGGCAAACGCACTGATGGCTATG
GGCGTAAAAAAAGGTGACCGTGTTACACTGTACATGCCGATGGTAGTTGAGGGTGTAACTGCAATGCTGGCGTGCGCACGCATCGGTGCAATCCACTCTGTGGTTTTCGGTGGTTTCAGCCCGG
AAGCTCTGGCAGGCCGCATCGAAGACTGCGGTAGCCGCTTCGTAGTGACCGCGGACCAGGGTCTGCGTGGCTCTAAGCGTGTTCCGCTCAAGGCTAACGTTGACGCAGCTCTGCAGATCGAAG
GCGCAGAGGTGGACGGTGTTCTGGTAGTGGCTCACACTGGCGCAGATGTTGCTATGACCGAAGGTCGCGACCACTGGTTCTCCGACCTGGCAAGCGACACTGACGTGCCTTGTGAAGTTATGC
AGGCAGAAGATCCGCTGTTCATCCTGTACACTTCTGGTTCCACTGGCAAACCAAAAGGTGTACTGCATACCACTGGTGGTTACGGTGTTTGGACTGCGACCACTTTCAACTACGTTTTCGACTATC

AGCCGGGTGAGGTATACTGGTGCACCGCTGACATCGGTTGGGTTACCGGCCACTCTTATATCGTTTATGGCCCGCTGATGAACGGCGCAACCTCCGTTGTTTTCGAAGGTGTGCCGAACTATCCG
GATCACGGTCGTTTCTGGGATGTAGTTGACAAGCACCAGGTTAACATTCTGTACACCGCTCCGACTGCTATCCGCGCTCTGATGCGCGAAGGTGACGAATTTGTGACCTCTCGCGACCGTAGCAG
CCTGCGTCTGCTCGGTTCTGTTGGTGAACCGATCAACCCGGAAGCATGGCGTTGGTACTTCGACGTAGTTGGTGAAGGCCGTTGCCCGATCGTAGACACCTGGTGGCAGACTGAGACCGGTGG
CGTGATGATTACTACCCTGCCGGCTGCTCACGACATGAAACCAGGCTCTGCAGGTCTGCCGTTTTTCGGTATCCAGCCGCAGCTGGTTGACAACGATGGTGGCGTTCTGGACGGTGCAGTCGAG
GGTAACCTGTGCATCACTGCCTCCTGGCCGGGCCAGGCACGTTCTGTATATGGTGACCATGAACGCTTTGAACAGACCTACTTCAGCACTTACCGTGGCAAGTACTTTACCGGCGACGGTTGCCG
CCGTGATGAGGACGGTTACTACTGGATCACCGGCCGCGTAGATGACGTTATCAACGTTTCTGGTCACCGTATGGGTACCGCGGAGGTTGAATCTGCTCTGGTGCTGCACGACAAAGTGGCAGA
AGCTGCAGTTGTGGGTTATCCACACGACATCAAGGGCCAGGGTATCTACTGCTACGTAACGCTGAACGCTGGCGAAGAAGGTTCTGAGGAGCTGTTCCGCGACCTGCGCATGCACGTTCGTAA
AGAAATCGGTCCGATCGCAACTCCGGACCACATCCAGTTCACTGATGGTCTGCCGAAGACCCGCTCCGGCAAAATTATGCGTCGTATCCTGCGCAAAGTTGCTGAAAACGACTACGGTTCTCTGG
GTGACACCTCCACCCTGGCAGACCCGTCTCTGGTGGACCGTCTGATCGAAGGCCGTCAGAACCGTTAACTCGAGGATCC

Nmar0206 – Nter His tag codon optimized for *M. extorquens* AM1

AAGCTTCCGTATCCCCAGAGGCAGCCAAATTGCCCGCTGCGCAACCCTGATACCGCGTCCAAGCGACGCTTGGGGGCGCTGTTGTGGTTGCGGTTTTTTGATCAGTATGGTTCTACGTCGCAACC
GCCCATCCGGCAGCGACACGGTCACGCCGGAACCGGCAAGGCCCGCGATACGATCGGGCAGTTCTGGGATTTGCGACGATCGTCGGCAAATTCAAGGGCAAACAAACCATAGGCACCAGGGG
AGCCAAGCGGGGACATGTCGACTAGCGTTTAACTTTAAGAAGGAGATATACCATGGGGCACCACCACCACCACCACCACCACCACCACTCGTCGGGTCACATCGAGGGCCGCCACATGACCGAT
TCTCCTATTCTCTCTCCGAAATCCATCGCGGTCATCGGCGCATCGGACAAACGTGGCTCAGTCGGCGCCACGATCACGTCGAACATCATGAACGGCTTCAAGGGCACGGTCTACCCTATCTCCCCC
ACCCGCGACACCGTTTTCTACAAGAAGGCTTACAAGTCCGTCCTGGACGTCCCGAAGTCCATTGACCTCGCCGTTATCGTGATCAAAAATACGCTTGTGACCCCGGTTCTTGAGGAATGCGGCAA
GAAGAAGATCAAGGGAGTGATTATCATCACAGCCGGCTTCAAGGAGGTCGATGAGGAAGGCGCTAAGCGCGAGCAGCAGGTGATCGACATCGCCAAGAAGTACAATATGCAAGTGGTCGGG
CCCAACTGCCTTGGCGTGATGAACCTCGACTCCAAAACCATGATGAACAGCACATTTCTCAAGGTCACGCCGAAGAGCGGAAAGATCGCCCTGGTCAGTCAAAGCGGCGCAATCTGCGCCGCGT
TGGTAGAAGACGCCTCCGCCCAGGGGATCGGCTTCTCCGCGGTGGTGTCGCTCGGCAACAAAGCGGTGATGTCTGAGGTCGATGTTCTGAAGATCCTCGCAAACCATAAGCAGACCGAGGTAA
TCGTCATGTATTTGGAGGACATGGGTGACGGCCAGGAGTTCCTGAAGGTTTGCAAGAACATCACGAAGAAGCTCAAAAAGCCTGTCCTCGTGTTGAAGTCAGGCCGCAGTCCCGAAGGCGCGA
AGGCCGCCATGTCGCACACGGGCGCCCTCATGGGCAGCGACGAGATCTACGACGCGCTCCTGAAACAATCCGGCGCAATCCGCGTTGACACAATGGAGGAGCTTTTCGACTATGCCACCGCGT
TCTCGAAGCAGCCCCTGCCTAGCAACGGCGATCTGGTCATCGTGTCTAACGCTGGAGGCCCGGCCATCATTTCCACCGACGCATGCTCCAAAGCGAAGATCAAGATGGCGGATATCACGAGCAT
TCGTAAAAAGATCGACGAAGTCATCCCGCCCTGGGGGAGCTCACGAAACCCGGTAGACATCGTCGGTGATGCTGACTTCAACCGCTTTCACAATGTGCTGGACCGTGTGCTCAAACACCCAAAG
GTCGGATCCGTGATCTCGATGTGCACGCCGAGCGGCACGCTGAACTATGACAAGCTCGCGGAAGTCATTGTGGAAATGTCGAAGAAGTACAAAAAGACCATGCTCGCCTCCCTGATGGGACTC
GATGAGGGGGTTACGAATAGGGAGATCCTCGCGGACGGCAACGTACCGTACTATACCTACGCCGAGGGCGCCATCCGCACCCTTGCCGCGATGATCAGGTTCTCCGACTGGGTCAAGAGCAGT
CCCGGCAAAATCACGAAGTTCAAGGTGAACAAGGCGAAAGCCAAAAAGATCTTTGACCAGGTGAAGAAGGAAAAGCGCCCGAACCTCCTTGAAGAGGAAGGGCAGGAGGTCCTCAAGGCGT
ACGGCCTCCCGCTCCCGAAATCCACCCTTGCCAAGAACGAGGCCGAGGCAGTGAAGGCGGCCAAGAAGATCGGTTACCCCGTCGTCATGAAGATCGCGAGCCCTCAGATCATCCATAAGAGCG
ATGCCGGCGGAGTCAAGGTAAACCTCACCAACGATGCCGAGGTCAAAGATGCGTTTAAGACCATCGTGAAGAACGCCAAAAAATATAACAAGAAGGCCGAGATCAAGGGCGTGCTCATTGTC
GAGATGGTGAAGGGGGGCAAGGAGCTCATCATCGGCTCGAAACTGGAACCCGGCTTCGGACCGGTCATCATGCTGGGCATGGGCGGCATCTACGTGGAGGTTCTCAAGGATGTGACGTTCAA
ACTCGCCCCGGTCACAGATAAGGAAGCGGACGACATGATCGCATCCATTAAGACCCAGAAACTTCTGCAAGGCGTGCGCGGCGAGAAACCGTCCGATATCGTAAAGCTCTCCGAATGCATCCA
GCGGCTGTCCCAGTTGGTCAGCGACTTCAAGGAGATCAAAGAGCTGGATATGAATCCTGTGTTGGTCATGGAAAAGGGCAAGGGCTGCCGAATCCTCGACGTGCGGATCGGTCTGTAGCAATT

GATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCC
GGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTGCTAGCGAATTCTGTACAACTAGTACATGCGGTACC

Nmar0207 – Nter His tag codon optimized for *E. coli* from Schwander *et al*., 2016[2]

GGATCCATGGCCAATGTCCTTAAAACCATCCGTAGTGGGGACGACTACATTGAGTCACTGCGTGGTCGTGACCTGAAAGTCTATCTCTTTGGCGAGTTGGTTAAGGAACCGGTTGATCATCCCAT
GATTCGCCCGTCGATTAACGCGGTAGCGGAAACCTACGACTTAGCTCTGCGTGAAGAAGCGCTTGCGTCAGCCGATTCGTCAATCACTGGTCTGAAAGTAAATCGCTTCCTGCATATCGCTGAAT
CCGCCGAGGATCTGGTTCTGCAAAACAAAATGCAGCGTAAACTGGGCCAGAATACCGGCACCTGTTTTCAGCGGTGTGTGGGCATGGATGCAATGAACTCGCTGCATAGTACTACCTTCGAAAT
CGATGAAAAGCACGGTACGGACTATCACAAACGGTTTCTGGAGTTCGTGAAGATGGTGCAGCAAGAGAATCTGGTGATTGGCGGAGCAATGACCGATCCGAAAGGCGATCGCAGCAAAGGTC
CCTCCGAACAAGATGATCCGGATCTGTTTACCCGCATTGTGGATACAGACGAGAAAGGAGTGTACGTGAGTGGCGCGAAAGCACACCAGACGGGCTGTATTAACAGCCATTGGATTATCTTAAT
GCCGACGATTCGTTTGACAGAAAGCGATAAAGATTGGGCGATTGTCGGTGCCATTCCAGCCGATGCGAAAGGCGTGACGTATATCTACGGTCGTCAAAGCTGCGATACTCGCAGCATGGAGGA
AGGGGATATCGATGATGGCAACGCGAAATTTGGTGGGCAGGAAGCGTTGATCATTCTGGATCGCGTTTTCATCCCGTGGGACAAAGTGTTTATGCATGGCGAATACGAATTTGCCAGCATGCTG
GTTGAACGCTTTACGTGCTATCATCGTCGCAGCTATGTGTGCAAAACGGGTCTGGGTGATGTACTGATCGGTGCAGCGGCTACCATTGCCGACTATAATGGCGTCCCTAAAGTTTCGCACATCAA
AGACAAAATTATCGAGATGACCCATCTCAACGAAACCATTTTCGCAGCGGGTATTGCATCCTCTCATCAAGGGCAGAAAATGAAATCAGGAGTATATCTGAACGATGATATGCTCGCCCAGGTCT
GCAAACACAACGTAACACGCTTCCCGTACGAAATCTCTCGCTTAGCACAGGACATTGCTGGCGGCCTGGTCGTGACGTTGCCTTCTGAGAAGGACTTTCGTCACCCAGAAGCCGGTCCTTTGCTT
AAGAAATATCTGGCGGGTCGCAAAGGCGTGGACGTTGAAAACCGCATGCGCATTCTGCGGTTAATCGAAAATATGACACTTGGCCGTAATGCTGTTGGATATCTGACCGAGAGTATGCATGGT
GCTGGGTCGCCACAGGCACAACGCATTCAGATTCAGCGTCAGATGCAAGTTGGGTATAAGAAGAATCTCGCGAAGAATTTAGCCGGCATTACTAACGATGTGGAAGAACCGAAAGAATCCAGC
GAATACTTCAAACGCGTCTTTAAAACCAAAGACTCTGTCCTGTAAGAATTC

CkSucD – Nter His tag codon optimized for *M. extorquens* AM1

ATTCTAGAAATTAGCTAATCTCGACCGAGGTATTCCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCTCGATGCATTCGAACGAAGTCAGCATCAAAGAGC
TCATCGAGAAGGCGAAAGTGGCCCAGAAGAAGCTCGAAGCGTACAGCCAAGAACAGGTCGACGTGCTGGTCAAGGCGCTCGGGAAGGTGGTCTACGACAATGCGGAAATGTTCGCCAAGGA
AGCCGTGGAGGAAACCGAAATGGGCGTCTACGAGGACAAGGTGGCCAAGTGCCACCTCAAGAGCGGCGCCATCTGGAACCACATCAAGGATAAGAAGACCGTGGGCATCATCAAAGAGGAG
CCGGAACGTGCCCTGGTCTACGTGGCGAAGCCCAAAGGCGTCGTCGCGGCGACCACGCCGATCACGAACCCGGTCGTGACGCCGATGTGCAACGCGATGGCGGCCATCAAGGGCCGTAACAC
CATCATCGTCGCCCCGCACCCGAAGGCCAAGAAAGTCTCGGCCCATACGGTGGAGCTCATGAACGCCGAGCTCAAGAAGCTCGGCGCGCCCGAGAACATCATCCAGATCGTGGAGGCGCCCAG
CCGGGAGGCCGCCAAGGAGCTGATGGAGAGCGCGGATGTCGTCATCGCCACCGGCGGCGCCGGGCGGGTCAAGGCGGCGTACAGCAGCGGCCGGCCGGCGTACGGGGTCGGCCCGGGTAA
CTCGCAGGTCATCGTGGACAAGGGCTACGACTACAACAAAGCCGCGCAGGACATCATCACCGGGCGGAAGTACGATAACGGGATCATCTGCTCGTCGGAGCAGTCGGTGATCGCGCCGGCCG
AGGACTACGATAAGGTGATCGCGGCGTTTGTGGAGAACGGCGCCTTCTATGTGGAGGACGAGGAGACCGTGGAGAAGTTCCGGTCCACGCTGTTCAAGGACGGCAAGATCAACTCCAAGATC
ATCGGTAAGTCGGTCCAGATCATCGCGGATCTGGCCGGCGTCAAGGTGCCGGAGGGCACCAAGGTCATCGTCCTGAAGGGCAAGGGGGCCGGCGAGAAGGACGTGCTCTGCAAAGAGAAGA
TGTGCCCCGTGCTCGTCGCGCTCAAGTACGACACCTTCGAGGAGGCCGTCGAAATCGCGATGGCGAATTATATGTATGAGGGCGCCGGCCACACGGCGGGCATCCATAGCGATAACGACGAGA

ACATCCGCTACGCCGGCACGGTCCTGCCCATCTCGCGCCTGGTCGTCAACCAGCCCGCGACCACCGCCGGCGGCTCGTTTAACAACGGCTTCAACCCGACCACGACCCTCGGCTGCGGCTCCTGG
GGCCGCAACTCCATCTCCGAGAATCTGACCTATGAACACCTGATCAACGTGTCCCGCATCGGTTATTTCAATAAAGAGGCGAAGGTCCCGTCCTACGAGGAAATCTGGGGCTGACTCGAGCTAC
CTAAGGTAGTAACTAGTCCTGCAGGTACCTAA

CdSucD$_{Mex}$ – Nter His tag codon optimized for *M. extorquens* AM1

TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCTCGATGCATGAAAAGGCTGTCGAGAATTT
CGAGGACTTGTCTAAGGAGTACATCAATGGGTACATCGAGCGCGCGCGTAAGGCACAACGCGAGTTCGAGTGCTACACGCAAGAGCAGGTGGACAAGATCGTCAAGATCGTCGGAAAGGTAG
TGTATTACAACGCCGAGTACCTTGCCAAGCTCGCCGTCGAAGAGACGGGCATGGGGGTCTATGAGGACAAGGTCGCGAAAAACAAGTCTAAGGCGAAGGTGATCTACAACAACTTGAAGGAC
AAGAAGAGTGTCGGAATCATCGATATCGATCGCGAAACCGGTATCACCAAAGTGGCCAAGCCGGTCGGCGTGGTCGCGGCGATTACCCCGTGCACCAACCCTATCGTGACACCGATGTCCAAC
GCGATGTTTGCGCTTAAGGGGCGGAACGCAATCATCATTACCCCCCACCACAAGGCGATCGGGTGCAGCACCAAGACCGTGGAGATGATCAACGAAGAGCTGGAAAAAATCGGCGCGCCGGA
GAACCTGATCCAGATCCTGGACCAGCAGAGCCGCGAGAACACCCGGAACCTCATTTCGTCCGCGGACGTCGTCATCGCCACCGGCGGTATGGGCATGGTGAAGGCGGCGTACTCCTCCGGCAA
GCCGGCGTTGGGTGTCGGCGCGGGCAACGTCCAGTGCATCATCGATCGGGACGTCGACATCAAGGAGGCGGTGCCGAAAATTATCGCCGGCCGCATCTTCGATAACGGCATCATTTGCTCCGG
CGAGCAGAGCGTCATTGTCGCCGAAGAGATGTTCGACAAGATCATGGACGAGTTCAAGAATAACAAGGGCTTCATCGTGCGGGACAAGGTGCAGAAGGAGGCGTTTCGCAATGCCATGTTCGT
CAATAAGTCGATGAACAAGGACGCTGTTGGGCAGTCGGTCCATACCATCGCCAAGATCGCCGGCGTCGAGATCCCCGAAGACACGAAGATCATCGTGATCGAGGCCGACGGCCCCGGCGAGG
AGGACATCATCGCCAAGGAAAAAATGTGCCCCGTGATTAGCGCCTATAAGTACAAATCCTTTGAGGAGGGCGTGGCCATCGCCAAAGCCAACCTCAACGTGGAGGGCAAGGGCCATTCGGTGT
CAATCCACTCGAACACGGTGAAGAACATCGAGTATGCCGGCGAGAACATCGAGGTGTCGCGCTTCGTGATTAACCAGTGCTGCGCCACCTCGGCGGGCGGCTCGTTCTTCAATGGCCTCGCCCC
GACCAATACCCTGGGCTGCGGCAGCTGGGGTAACAACTCTATTAGCGAGAATCTGGACTACAAACACCTGATCAATATCTCCCGCATCGCCTACTATATGCCGGAGAACGAGGTGCCGACGGAT
GAGGAGCTGTGGGGCTGACTCGAGCTACCTAAGGTAGTAACTAGTCCTGCAGGTACCT

PgSucD – Nter Strep-II tag codon optimized for *M. extorquens* AM1

CAATTGTGAGCGGATAACAATTTCACACATCTAGAGATCCGCAGTTAACCGAGTAAAGGAGGTAATATATGGCCAGCTGGTCGCACCCCCAGTTCGAGAAGATCGAAGGGCGGATGCATGAAA
TCAAGGAGATGGTGAGCCTGGCCCGCAAGGCCCAAAAGGAGTATCAGGCGACCCACAACCAGGAGGCCGTCGACAATATCTGCCGGGCCGCGGCGAAGGTCATCTATGAAAATGCGGCCATC
CTCGCGCGGGAAGCCGTGGACGAAACGGGTATGGGTGTGTACGAGCATAAGGTCGCCAAGAATCAGGGTAAGTCGAAGGGCGTGTGGTACAACCTGCATAACAAGAAGTCCGTCGGCATCCT
GAATATCGACGAGCGGACCGGCATGATCGAAATCGCGAAGCCCATCGGGGTGGTGGGGGCCGTCACGCCCACGACGAACCCGATCGTCACGCCCATGTCCAACATCATCTTCGCCCTCAAGAC
CTGCAACGCGATCATCATCGCGCCCCATCCGCGGTCGAAGAAGTGCAGCGCCCACGCGGTCCGCCTCATCAAGGAAGCGATCGCCCCGTTCAACGTGCCGGAGGGCATGGTCCAGATCATCGA
GGAACCCTCGATCGAGAAGACCCAGGAGCTGATGGGCGCGGTGGATGTGGTCGTCGCCACCGGCGGGATGGGCATGGTGAAGTCCGCCTACTCCAGCGGCAAGCCGAGCTTCGGCGTCGGG
GCGGGCAACGTCCAGGTGATCGTGGATAGCAACATCGACTTCGAGGCGGCCGCCGAGAAGATCATCACCGGCCGCGCGTTCGACAACGGCATCATCTGCAGCGGCGAGCAGAGCATCATCTAC
AACGAGGCGGACAAGGAGGCGGTGTTCACCGCGTTCCGCAACCACGGCGCGTACTTCTGCGACGAGGCGGAGGGCGATCGCGCCCGCGCGGCCATCTTCGAGAACGGCGCCATCGCCAAGG
ATGTGGTGGGCCAGTCCGTGGCGTTCATCGCCAAGAAGGCGAACATCAACATCCCGGAGGGCACCCGCATCCTCGTCGTCGAGGCGCGCGGCGTCGGCGCCGAGGACGTCATCTGCAAGGAG
AAGATGTGCCCGGTCATGTGCGCCCTGTCCTACAAGCATTTCGAGGAGGGCGTGGAGATCGCCCGCACCAACCTGGCCAACGAGGGCAACGGCCACACCTGCGCCATCCACTCGAACAACCAG

GCCCACATCATCCTCGCCGGCTCCGAGCTGACCGTGTCCCGCATCGTCGTCAACGCCCCGTCGGCCACCACCGCCGGCGGCCACATCCAGAACGGCCTCGCCGTCACGAACACCCTCGGCTGCG
GCTCGTGGGGCAACAACTCGATCTCGGAGAACTTCACGTACAAGCACCTGCTCAACATCTCGCGCATCGCCCCGCTGAACTCGTCGATCCACATCCCGGACGACAAGGAGATCTGGGAGCTCTG
ACCTAAGGACTAGTCCTGCAGGTACCTCTAAGCTTAATTAGCTGAGCTTGGA

References

1. Berg, I. A. (2011) Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways, *Appl. Environ. Microbiol. 77*, 1925-1936.
2. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S., and Erb, T. J. (2016) A synthetic pathway for the fixation of carbon dioxide in vitro, *Science 354*, 900-904.
3. Erb, T. J., Berg, I. A., Brecht, V., Muller, M., Fuchs, G., and Alber, B. E. (2007) Synthesis of C_5-dicarboxylic acids from C_2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway, *Proceedings of the National Academy of Sciences of the United States of America 104*, 10631-10636.
4. Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C., Peyraud, R., Robinson, D. G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C. J., Vorholt, J. A., Olson, M. V., Kaul, R., Weissenbach, J., Medigue, C., and Lidstrom, M. E. (2009) Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources, *PloS one 4*, e5584.
5. Chistoserdov, A. Y., Chistoserdova, L. V., McIntire, W. S., and Lidstrom, M. E. (1994) Genetic organization of the *mau* gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of *mau* mutants, *J. Bacteriol. 176*, 4052-4065.
6. Toyama, H., Anthony, C., and Lidstrom, M. E. (1998) Construction of insertion and deletion *mxa* mutants of *Methylobacterium extorquens* AM1 by electroporation, *FEMS Microbiol. Lett. 166*, 1-7.
7. Marx, C. J., and Lidstrom, M. E. (2001) Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria, *Microbiology 147*, 2065-2075.
8. Marx, C. J., and Lidstrom, M. E. (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria, *BioTechniques 33*, 1062-1067.
9. Marx, C. J., O'Brien, B. N., Breezee, J., and Lidstrom, M. E. (2003) Novel methylotrophy genes of *Methylobacterium extorquen*s AM1 identified by using transposon mutagenesis including a putative dihydromethanopterin reductase, *J. Bacteriol. 185*, 669-673.
10. Marx, C. J. (2008) Development of a broad-host-range sacB-based vector for unmarked allelic exchange, *BMC Res Notes 1*, 1.
11. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A., and Erb, T. J. (2015) A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria, *ACS Synth. Biol. 4*, 430-443.
12. Carrillo, M., Wagner, M., Petit, F., Dransfeld, A., Becker, A., and Erb, T. J. (2019) Design and Control of Extrachromosomal Elements in *Methylorubrum extorquens* AM1, *ACS Synth. Biol. 8*, 2451-2456.
13. Sohling, B., and Gottschalk, G. (1993) Purification and characterization of a coenzyme-A-dependent succinate-semialdehyde dehydrogenase from *Clostridium kluyveri*, *Eur. J. Biochem. 212*, 121-127.

14. Schaller, M., Schaffhauser, M., Sans, N., and Wermuth, B. (1999) Cloning and expression of succinic semialdehyde reductase from human brain. Identity with aflatoxin B1 aldehyde reductase, *Eur. J. Biochem. 265*, 1056-1060.
15. Hoover, G. J., Van Cauwenberghe, O. R., Breitkreuz, K. E., Clark, S. M., Merrill, A. R., and Shelp, B. J. (2007) Characteristics of an *Arabidopsis* glyoxylate reductase: general biochemical properties and substrate specificity for the recombinant protein, and developmental expression and implications for glyoxylate and succinic semialdehyde metabolism *in planta*, *Can. J. Bot./Rev. Can. Bot. 85*, 883–895.
16. Meyer, M., Schweiger, P., and Deppenmeier, U. (2015) Succinic semialdehyde reductase Gox1801 from *Gluconobacter oxydans* in comparison to other succinic semialdehyde-reducing enzymes, *Appl. Microbiol. Biotechnol. 99*, 3929-3939.
17. Könneke, M., Schubert, D. M., Brown, P. C., Hugler, M., Standfest, S., Schwander, T., Schada von Borzyskowski, L., Erb, T. J., Stahl, D. A., and Berg, I. A. (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation, *Proceedings of the National Academy of Sciences of the United States of America 111*, 8239-8244.
18. Berg, I. A., Ramos-Vera, W. H., Petri, A., Huber, H., and Fuchs, G. (2010) Study of the distribution of autotrophic CO_2 fixation cycles in Crenarchaeota, *Microbiology 156*, 256-269.
19. An, J. H., and Kim, Y. S. (1998) A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in *Rhizobium trifolii*--cloning, sequencing, and expression of the enzymes in *Escherichia coli*, *Eur. J. Biochem. 257*, 395-402.
20. Reger, A. S., Carney, J. M., and Gulick, A. M. (2007) Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase, *Biochemistry 46*, 6536-6546.
21. Scherf, U., and Buckel, W. (1991) Purification and properties of 4-hydroxybutyrate coenzyme A transferase from Clostridium aminobutyricum, *Appl. Environ. Microbiol. 57*, 2699-2702.
22. Korotkova, N., and Lidstrom, M. E. (2001) Connection between poly-beta-hydroxybutyrate biosynthesis and growth on C(1) and C(2) compounds in the methylotroph*Methylobacterium extorquens* AM1, *J. Bacteriol. 183*, 1038-1046.
23. Salis, H. M. (2011) The ribosome binding site calculator, *Methods Enzymol. 498*, 19-42.
24. Hersh, L. B. (1973) Malate adenosine triphosphate lyase. Separation of the reaction into a malate thiokinase and malyl coenzyme A lyase, *J. Biol. Chem. 248*, 7295-7303.
25. Hersh, L. B. (1974) Malate thiokinase. The reaction mechanism as determined by initial rate studies, *J. Biol. Chem. 249*, 6264-6271.
26. Schada von Borzyskowski, L., Sonntag, F., Poschel, L., Vorholt, J. A., Schrader, J., Erb, T. J., and Buchhaupt, M. (2018) Replacing the Ethylmalonyl-CoA Pathway with the Glyoxylate Shunt Provides Metabolic Flexibility in the Central Carbon Metabolism of *Methylobacterium extorquens* AM1, *ACS Synth. Biol. 7*, 86-97.
27. Schada von Borzyskowski, L., Severi, F., Kruger, K., Hermann, L., Gilardet, A., Sippel, F., Pommerenke, B., Claus, P., Cortina, N. S., Glatter, T., Zauner, S., Zarzycki, J., Fuchs, B. M., Bremer, E., Maier, U. G., Amann, R. I., and Erb, T. J. (2019) Marine

Proteobacteria metabolize glycolate via the beta-hydroxyaspartate cycle, *Nature 575*, 500-504.
28. Peter, D. M., Vogeli, B., Cortina, N. S., and Erb, T. J. (2016) A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters, *Molecules 21*, 517.
29. Rosenthal, R. G., Ebert, M. O., Kiefer, P., Peter, D. M., Vorholt, J. A., and Erb, T. J. (2014) Direct evidence for a covalent ene adduct intermediate in NAD(P)H-dependent enzymes, *Nat. Chem. Biol. 10*, 50-55.
30. Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., and Vorholt, J. A. (2009) Demonstration of the ethylmalonyl-CoA pathway by using 13C metabolomics, *Proc. Natl. Acad. Sci. USA 106*, 4846-4851.
31. Sambrook, J., and Russel, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
32. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat. Methods 6*, 343-345.
33. Delaney, N. F., Kaczmarek, M. E., Ward, L. M., Swanson, P. K., Lee, M. C., and Marx, C. J. (2013) Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*, *PloS one 8*, e62957.
34. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248-254.
35. Torkko, J. M., Koivuranta, K. T., Miinalainen, I. J., Yagi, A. I., Schmitz, W., Kastaniotis, A. J., Airenne, T. T., Gurvitz, A., and Hiltunen, K. J. (2001) *Candida tropicalis* Etr1p and *Saccharomyces cerevisiae* Ybr026p (Mrf1'p), 2-enoyl thioester reductases essential for mitochondrial respiratory competence, *Mol. Cell. Biol. 21*, 6243-6253.
36. de Marco, A. (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*, *Nature protocols 2*, 2632-2639.
37. Lindenkamp, N., Schurmann, M., and Steinbuchel, A. (2013) A propionate CoA-transferase of *Ralstonia eutropha* H16 with broad substrate specificity catalyzing the CoA thioester formation of various carboxylic acids, *Appl. Microbiol. Biotechnol. 97*, 7699-7709.
38. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research, *DNA Res. 12*, 291-299.
39. Ind, A. C., Porter, S. L., Brown, M. T., Byles, E. D., de Beyer, J. A., Godfrey, S. A., and Armitage, J. P. (2009) Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*, *Appl. Environ. Microbiol. 75*, 6613-6615.
40. Rosenthal, R. G., Vogeli, B., Quade, N., Capitani, G., Kiefer, P., Vorholt, J. A., Ebert, M. O., and Erb, T. J. (2015) The use of ene adducts to study and engineer enoyl-thioester reductases, *Nat. Chem. Biol. 11*, 398-400.

PART THREE: FINAL REMARKS

DISCUSSION

6.1 Beyond natural metabolism

The overarching goal of this thesis was to expand the metabolic potential of *M. extorquens* AM1. This organism can utilize many carbon sources heterotrophically, but not C_5 or C_6 sugars. This is unexpected, since its genome encodes for a complete pentose phosphate pathway (PPP), the Entner-Doudoroff (ED) pathway, and glycolytic/gluconeogenic pathways. Moreover, certain genes in the PPP show redundancy by being present in multiple copies[1]. We decided to engineer assimilation of ribose (C_5) in this organism, as it opens up interesting applications for the utilization of plant-derived material as feedstock and its conversion into biotechnologically relevant products in the future. To that end, overexpression of a heterologous ribose kinase gene enabled growth on ribose as the sole source of energy and carbon. Further rational engineering with a sugar facilitator and, lastly, adaptive laboratory evolution (ALE) led to robust growth on ribose by *M. extorquens* AM1 (CHAPTER III). While the genetic determinants remain to be elucidated, our evolved strains can still be further used. As *M. extorquens* wild type requires upper (sugar phosphate) metabolism solely to provide cellular building blocks, and not for the assimilation of carbon *per se*, no genetic studies or deletion strains have been generated, to the best of our knowledge. One study found that certain homolog genes of the PPP are differentially expressed on C_1 (methanol) *versus* C_4 (succinate) but did not pursue this further[2]. It would be very interesting to follow up on the evolved strains by transcriptomics and/or proteomics, comparing them to the parent to identify which homologues are active under these conditions. Subsequent biochemical characterization as done for triosephosphate isomerase (TIM; CHAPTER III) would also give new insights into the metabolic network of *M. extorquens* AM1. Currently, the fate of fructose 6-phosphate (F6P) in both engineered and evolved strains is unknown, since gene(s) encoding for a phosphofructokinase cannot be identified by homology searches. We assume that F6P formed during ribose assimilation enters the oxidative pentose pathway or the ED pathway. Furthermore, no mutations were found in genes belonging to related pathways after ALE, except for TIM (CHAPTER III). These genes are presumably not essential for biosynthesis and have been deleted in organisms such as *E. coli* and *Saccharomyces cerevisiae*[3, 4]. A more high-throughput method for gene disruption for *M. extorquens*, as discussed in the following sections, is necessary due to the number of targets as well as gene redundancy (two putative *tal-pgi* genes, two putative *zwf* genes, two putative *gnd* genes, and one copy of *edd* and *eda* each). In essence, the strains generated in CHAPTER III will enable us to probe the upper metabolic network on natural carbon sources (*i.e.* biosynthetic requirements) *versus* C_5 sugars (*i.e.* ribose assimilation specific) with growth as an easy read-out.

6.2 Synthetic autotrophy (again)

A very important, yet often overlooked, process during heterotrophic growth is energy conservation. That is because heterotrophic organisms generally utilize organic molecules both as energy and carbon sources. In (chemo)organoautotrophic growth, organic compounds are oxidized to produce reducing equivalents the cell needs and inorganic carbon is fixed to form biomass. In this thesis, we modified endogenous metabolism so energy could be conserved by the oxidation of organic compounds, and implemented CO_2 fixation by means of the natural, but heterologous, CBB (CHAPTERS II and III) and the artificial CETCH cycle (CHAPTER IV) in *M. extorquens* AM1.

To separate energy conservation from carbon assimilation, we took advantage of the native metabolic network and blocked assimilation at key steps via gene knockouts. During methylotrophy, only a fraction of C_1 units are assimilated via the serine cycle at the level of formate and the rest are fully oxidized for energy[5, 6]. Thus, the earliest step to block carbon assimilation from methanol without disrupting energy conservation is at the level of the formate tetrahydrofolate ligase (encoded by *ftfL*). Based on a metabolic network model, single deletions downstream of *ftfL*, namely the serine hydroxymethyltransferase (encoded by *glyA*) and crotonyl-CoA carboxylase/reductase (encoded by *ccr*) were also investigated. As expected, all mutant strains were unable to grow on methanol, yet only the Δ*ftfL* mutant showed a growth phenotype when the CBB cycle was introduced for carbon fixation hinting at (chemo)organoautotrophy (CHAPTER II). It was unexpected that the CBB cycle would fail to complement growth in the Δ*glyA* mutant, since it is at the start of the serine cycle and its gene product incorporates C_1 units onto glycine. One possible explanation is that an additional molecule of ATP and NADPH are invested, without any benefit since *glyA* is knocked out, in comparison to a disruption at the level of *ftfL*. Unknown regulatory processes or inhibition of enzymes by trapping tetrahydrofolate might also take place. The gene product encoded by *ccr* is located further downstream of *glyA* and therefore, drains more of the cell's resources without being able to assimilate carbon; making it a poor target in terms of energy conservation. In conclusion, we have shown that energy metabolism (*i.e.* methanol oxidation) can be separated from methanol assimilation (CHAPTER II).

Similarly, cellular energy can be generated from acetate oxidation as shown in CHAPTERS III and IV. If the ethylmalonyl-CoA pathway (EMCP) is disrupted, then C_2 units from acetate cannot be assimilated into biomass (as malate) and acetate will simply provide energy for the cell via the two oxidative decarboxylation steps of the TCA cycle. By combining ribose assimilation and CO_2 fixation via the CBB cycle, these engineered strains are dependent on ribose for energy. Hence, we used an EMCP mutant (disrupting the β-ketothiolase encoded by *phaA*) and provided acetate for extra energy (CHAPTER III). Only in this way, can (chemo)organoautotrophy theoretically be selected for in *M. extorquens* AM1 equipped with the CBB cycle. As mentioned in CHAPTER III, these strains managed to escape CO_2 fixation after ALE, exclusively assimilating ribose for carbon, but still consumed acetate

for energy. Additional gene deletions to tighten selection and constrain ribose assimilation only in the presence of an active CBB cycle are of interest before the next round of ALE.

With the aim of implementing the non-natural CETCH cycle into *M. extorquens* AM1, we disrupted the EMCP – abolishing growth on C_1/C_2 carbon source as they need glyoxylate as a carbon backbone – and aimed to complement growth by forming a closed autocatalytic cycle (CHAPTER IV). We investigated alternative feeding schemes since simply introducing and expressing the genes encoding for the crotonyl-CoA regeneration module did not restore growth. It is likely that the accumulation of certain metabolites, low overall flux through the regeneration module, and too many metabolites drained away by the endogenous metabolism result in low amounts of glyoxylate being formed from CO_2 fixation. This selection scheme was thought to be less stringent (compared to halting carbon assimilation at *ftfL*) since glyoxylate enters the endogenous central carbon metabolism without the need for a "glyoxylate assimilation module" such as the glycerate pathway as previously proposed[7]. This pathway is inherently disadvantageous as it releases one molecule of CO_2 for every two glyoxylate molecules assimilated[8, 9]. Combining the CETCH cycle with carbon-conserving glyoxylate (or glycolate) assimilation pathways can be explored in the future to promote carbon assimilation via the CETCH cycle in *M. extorquens* AM1[10, 11]. Alternately, the burden placed on CO_2 fixation by current selection schemes can be alleviated by selecting for amino acid biosynthesis instead of full biomass. In this scenario, a carbon and energy source (such as succinate) would be fed to a glyoxylate-glycine auxotrophic strain of *M. extorquens* AM1 and glyoxylate fixed through the CETCH cycle would provide glycine to the cell via a transaminase. Current optimization and identification of crucial bottlenecks of this cycle *in vitro* being performed in our laboratory will certainly help guide *in vivo* implementation further (C. Diehl, unpublished results)[7]. The generation of additional selection schemes by rational means as well as ALE will consolidate influx and efflux in balance, which has been shown to be crucial for keeping autocatalytic cycles up-and-running[12-15].

6.3 Importance of *repABC*-based mini-chromosomes.

Secondary replicons are widespread in Alphaproteobacteria and up to 10% of all bacterial genomes have a multipartite genome architecture[16, 17]. The *repABC* operons present in these replicons encode for the replication and segregation machinery. The gene product of the operon as well as with various regulatory elements give them chromosomal-like properties: unit copy number relative to the main chromosome, stable replication and partitioning even in the absence of selection[18-20]. Only recently, have natural *repABC* cassettes used in heterologous hosts such as *S. meliloti*[21] and *M. extorquens* AM1 (CHAPTER I) for synthetic biology to create artificial 'mini-chromosomes'.

Previous to our work, engineering of *M. extorquens* AM1 was limited by the fact that only one BHR origin of replication was suitable. By exploiting *repABC* regions from closely related organisms three stable mini-chromosomes (Mrad-JCM, Mex-DM4, and Nham-3) and one transiently stable mini-chromosome (Mex-CM4) were developed and characterized for *M. extorquens* AM1 (CHAPTER I). Mex-CM4 lacks additional *parS* site(s) required for faithful transmission in the absence of antibiotic selection and is thus perfectly suited to explore transient expression of proteins. In addition, we showed it was possible to combine multiple *repABC* mini-chromosomes in a single cell without growth impairments and all mini-chromosomes are compatible with BHR plasmids. It is important to highlight that these mini-chromosomes do not interfere with the native replicons of *M. extorquens* AM1. Moreover, each replicon behaves independently allowing more complex genetic engineering in the future. For example, dividing pathways are into modules, separately optimizing each module, and then seamlessly combining them. Importantly, the *repABC* region is relatively small (5-10 kb), thus facilitating their manipulation, yet they are is normally found in large replicons ranging from 0.3 to 2.6 Mb. Therefore, *repABC*--based mini-chromosomes will be able to withstand implementation of large DNA regions and metabolic pathways, a feat not possible with 'regular' plasmids.

Dynamic control over gene expression was also underdeveloped, mostly relying on the toxic CymR repressor and resulting in leakiness or weak expression after induction[22-24]. Therefore, we switched to the well-established LacI-*lacO* system[25]. By dividing promoters into their core elements, designing multiple parts for each element, and combining them together, we obtained ten novel promoters with distinct properties in terms of leakiness before induction and overall strength at full induction (CHAPTER I). The latter was compared to the P_{mxaF} promoter, which drives expression corresponding to 9% of soluble protein[26], and is therefore the strongest native promoter in this organism. Unlike previous studies[24], whose inducible promoters were at best 33% of the P_{mxaF} promoter, we identified promoters weaker, matching, and exceeding P_{mxaF}, making them the strongest promoters reported to date for *M. extorquens* AM1.

With these tools at hand, protein production and purification from *M. extorquens* AM1 (CHAPTER III) and *E. coli* (CHAPTER IV) for biochemical characterization was possible. Furthermore, they allowed us to express all genes of the CETCH cycle regeneration module dynamically so that they are only present when needed without representing an unnecessary burden to the cells (CHAPTER IV). Furthermore, these promoters are also suitable for related organisms: *Rhodobacter sphaeroides* 2.4.1 (M. Carrillo, unpublished results), and *M. extorquens* PA1 (J. Hemmann, personal communication). Beyond the LacI-*lacO* system, one possibility is to combine the elements we have identified with the tetracycline regulator-operator system[27, 28] (TetR-*tetO*) to overcome previous limitations; although, an inherent drawback is that tetracycline can no longer be used as a selectable marker[24]. Alternatively, newly optimized sensors can be used to have independent control over gene expression using 12 different small-molecule inducers[29]. Future work will focus

on using the tools developed here as a platform to test novel promoters and expression systems.

So far, our work was focused on the model organism *M. extorquens* AM1: however, the applicability of *repABC* mini-chromosomes extends to the hundreds of Alphaproteobacterial isolates with secondary replicons[19]. As we have shown, *repABC* cassettes from related organisms can confer chromosomal-like properties to heterologous hosts (CHAPTER I). Hence, suitable *repABC* cassettes can be identified for members of this class with multipartite genomes such as *Rhodobacter sphaeroides*, *Paracoccus denitrificans*, *Ruegeria pomeroyi*, and the many agriculturally relevant species of symbiotic nitrogen-fixing bacteria, whose genetic toolboxes would all benefit from mini-chromosomes. Even more remarkable, *M. extorquens* PA1[30], near-identical to AM1 but with a unified genome[31, 32], was able to replicate the mini-chromosomes from CHAPTER I (A. Ochsner, personal communication). This suggests that *repABC* mini-chromosomes are functional in related organisms lacking secondary replicons and could become universal DNA vehicles for members of the Alphaproteobacterial class. One example of particular interest in our laboratory is *Erythrobacter sp.* NAP1, which is so far genetically not tractable (I. Bernhardsgrütter, personal communication). In this context, it would be interesting to explore this possibility further.

6.4 Efficient genome editing tools for *M. extorquens* AM1

The *repABC* mini-chromosomes are an elegant solution for (over)expression of native and heterologous genes; however, gene deletions and other such modifications are not addressed. The current genome editing methods in *M. extorquens* AM1 entail using suicide vectors with homologous regions (> 0.5 kb each) with an antibiotic marker or sucrose counterselection[33, 34]. Both approaches are laborious and time-consuming as they rely on the native machinery for recombination and sucrose counterselection can still lead to wildtype revertants. More efficient genome editing tools have been developed for *E. coli* and other organisms with three prominent examples: (i) phage-derived recombinases[35-37], (ii) CRISPR-Cas9 assisted selection[38-40], and (iii) deaminase-mediated Target-AID editing[41, 42].

The RecET and λ-Red system from *E. coli* rac prophage and bacteriophage λ, respectively, employ an exonuclease, a single stranded DNA (ssDNA) binding protein, and a host-nuclease inhibitor protein to mediate site-specific genome editing using linear DNA with 36-60 nt homology arms[35, 36]. Both systems have been used extensively to make fast insertions, deletions and mutations in *E. coli* and a variety of other bacteria[43-47]. Preliminary results suggest that a codon-optimized version of the RecET system, transiently expressed from Mex-CM4 and under the control of an IPTG inducible promoter (CHAPTER I), can mediate recombination of linear PCR products with short homology arms in *M. extorquens* AM1 (M. Carrillo, unpublished results). Of course, further optimization

of the protocol (*e.g.* length of homology arms, preparation of competent cells, amount of DNA, induction of phage recombination proteins, etc.) is necessary. Although very effective, both the RecET and λ-Red systems still rely on an antibiotic resistance cassette for positive selection. In a two-step process, the RecET or λ-Red systems can be combined with sucrose counterselection, already established for *M. extorquens* AM1[34], for markerless modifications.

Clustered regularly interspaced palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins, part of bacterial adaptive immune systems[48, 49], have been repurposed as a powerful genome engineering tool using the Cas9 endonuclease from *Streptococcus pyogenes*[38, 39, 50]. A small RNA directs the Cas9 protein to the desired genomic locus ensuring targeted double-stranded breaks[49]. A template DNA for recombination is required as most prokaryotes lack the non-homologous end joining pathway of DNA repair otherwise resulting in cell death[51, 52]. Therefore, CRISPR-Cas9 in bacteria is used mostly for negative selection against unmodified cells in combination with the RecET or λ Red systems for recombination[40]. Recently, CRISPR interference (CRISPRi) for downregulation of targeted genes with a nuclease-deficient Cas9 (deactivated or dCas9) from *S. pyogenes* was established for *M. extorquens* AM1[53]. The effective downregulation of multiple targets shown in said study demonstrates that dCas9 is functional in this organism. Therefore, high-throughput, scarless genome editing with an active Cas9 variant and the RecET system should be straightforward to establish.

Lastly, Target-AID introduces cytosine-to-thymine substitutions instead of double stranded breaks making it much less toxic[41, 42]. Targeted modification with this method is achieved fusing the cytidine deaminase to dCas9 as well as an uracil DNA glycosylase inhibitor protein without the need for template DNA or any additional host-specific factors required. In conclusion, these genome editing tools should be readily applicable using the tools developed in CHAPTER I for *M. extorquens* AM1 and other Alphaproteobacteria.

6.5 Conclusions and Outlook

The aim of this thesis was to expand and re-engineer the metabolic network of *M. extorquens* AM1 to alter its growth capabilities. We were able to engineer growth on ribose as the sole carbon and energy source and growth was further improved by serial transfers. The metabolic fate of F6P produced during ribose assimilation, however, is still unclear. To uncover this, the ED pathway and both 6-phosphogluconate dehydrogenases should be deleted in ribose-evolved clones in the future.

We also aimed to engineer autotrophic growth in *M. extorquens* AM1, which is otherwise an obligate heterotrophic organism. The conversion of a heterotrophic organism into an autotrophic one is of interest because it ultimately allows for the production of value-

added compounds from CO_2. To that end, we halted carbon assimilation from methanol into biomass, while still allowing for energy conservation in the form of reducing equivalents, and equipped this strain with the CBB cycle for CO_2 fixation mimicking (chemo)organoautotrophic growth. We demonstrated functional operation of the heterologous CBB cycle through metabolic ^{13}C-tracer analysis as well as a positive growth phenotype but not continuous growth. Therefore, we introduced CO_2 fixation through the CBB cycle in ribose-evolved clones. We achieved robust continuous growth via CO_2 fixation, ribose assimilation, and energy conservation from acetate oxidation. However, the assimilation of organic and inorganic carbon into biomass was lost in escape mutations after ALE. Therefore, it is imperative to make ribose assimilation more tightly dependent on the CBB cycle. Deletion of phosphoglucose isomerase (2 putative bifunctional genes) and/or glucose 6-phosphate dehydrogenases (4 putative genes) are important targets in the future.

By making full use of the genetic tools developed in this thesis, a functional RecET system for efficient recombination of linear DNA, and later on CRISPR-Cas, can be established. This in turn would greatly speed up the creation of gene knockouts and make the systematic deletion of every open reading frame in *M. extorquens* AM1 feasible. Subsequent transfer of these tools to closely-related Alphaproteobacteria would also be possible. In summary, future development of genome editing techniques using the tools developed in this thesis would help uncover unknown gene functions and facilitate metabolic engineering to harness the biotechnological potential of *M. extorquens* AM1.

Nature has evolved various CO_2 fixation pathways but synthetic biology offers the possibility to circumvent these and create artificial ones such as the CETCH cycle. *M. extorquens* AM1 was an ideal organism to probe *in vivo* transplantation since many steps of the CETCH cycle are part of its native metabolism. We identified and implemented the missing steps corresponding to the crotonyl-CoA regeneration module. Selection schemes based on restoring growth on C_1/C_2 carbon sources of selected mutants still require further optimization. Future work could also explore restoring glycine auxotrophy via a transaminase from glyoxylate, the output of the CETCH cycle, as a less stringent selective pressure on CO_2 fixation. This thesis illustrates the complexity and importance of bridging the gap between *in silico* design and *in vivo* realization.

References

1. Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C., Peyraud, R., Robinson, D. G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C. J., Vorholt, J. A., Olson, M. V., Kaul, R., Weissenbach, J., Medigue, C., and Lidstrom, M. E. (2009) Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources, *PloS one 4*, e5584.
2. Skovran, E., Crowther, G. J., Guo, X., Yang, S., and Lidstrom, M. E. (2010) A systems biology approach uncovers cellular strategies used by *Methylobacterium extorquens* AM1 during the switch from multi- to single-carbon growth, *PloS one 5*, e14091.
3. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, *Mol. Syst. Biol. 2*, 2006 0008.
4. Giaever, G., and Nislow, C. (2014) The yeast deletion collection: a decade of functional genomics, *Genetics 197*, 451-465.
5. Crowther, G. J., Kosaly, G., and Lidstrom, M. E. (2008) Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1, *J. Bacteriol. 190*, 5057-5062.
6. Peyraud, R., Schneider, K., Kiefer, P., Massou, S., Vorholt, J. A., and Portais, J. C. (2011) Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1, *BMC systems biology 5*, 189.
7. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S., and Erb, T. J. (2016) A synthetic pathway for the fixation of carbon dioxide in vitro, *Science 354*, 900-904.
8. Barkulis, S. S., and Krakow, G. (1956) Conversion of glyoxylate to hydroxypyruvate by extracts of *Escherichia coli*, *Biochim. Biophys. Acta 21*, 593-594.
9. Hansen, R. W., and Hayashi, J. A. (1962) Glycolate metabolism in *Escherichia coli*, *J. Bacteriol. 83*, 679-687.
10. Trudeau, D. L., Edlich-Muth, C., Zarzycki, J., Scheffen, M., Goldsmith, M., Khersonsky, O., Avizemer, Z., Fleishman, S. J., Cotton, C. A. R., Erb, T. J., Tawfik, D. S., and Bar-Even, A. (2018) Design and in vitro realization of carbon-conserving photorespiration, *Proceedings of the National Academy of Sciences of the United States of America 115*, E11455-E11464.
11. Schada von Borzyskowski, L., Severi, F., Kruger, K., Hermann, L., Gilardet, A., Sippel, F., Pommerenke, B., Claus, P., Cortina, N. S., Glatter, T., Zauner, S., Zarzycki, J., Fuchs, B. M., Bremer, E., Maier, U. G., Amann, R. I., and Erb, T. J. (2019) Marine Proteobacteria metabolize glycolate via the beta-hydroxyaspartate cycle, *Nature 575*, 500-504.
12. Antonovsky, N., Gleizer, S. s., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D. G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., and Milo, R. (2016) Sugar Synthesis from CO_2 in *Escherichia coli*, *Cell 166*, 115-125.

13. Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., and Milo, R. (2017) Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points, *eLife 6*.
14. Herz, E., Antonovsky, N., Bar-On, Y., Davidi, D., Gleizer, S., Prywes, N., Noda-Garcia, L., Lyn Frisch, K., Zohar, Y., Wernick, D. G., Savidor, A., Barenholz, U., and Milo, R. (2017) The genetic basis for the adaptation of *E. coli* to sugar synthesis from CO_2, *Nature communications 8*, 1705.
15. Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., and Milo, R. (2019) Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO_2, *Cell 179*, 1255-1263 e1212.
16. Harrison, P. W., Lower, R. P., Kim, N. K., and Young, J. P. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid, *Trends Microbiol. 18*, 141-148.
17. diCenzo, G. C., and Finan, T. M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution, *Microbiol. Mol. Biol. Rev. 81*.
18. Cevallos, M. A., Cervantes-Rivera, R., and Gutierrez-Rios, R. M. (2008) The repABC plasmid family, *Plasmid 60*, 19-37.
19. Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012) The ABCs of plasmid replication and segregation, *Nat. Rev. Microbiol. 10*, 755-765.
20. Fournes, F., Val, M. E., Skovgaard, O., and Mazel, D. (2018) Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes, *Frontiers in microbiology 9*, 1833.
21. Dohlemann, J., Brennecke, M., and Becker, A. (2016) Cloning-free genome engineering in Sinorhizobium meliloti advances applications of Cre/loxP site-specific recombination, *J. Biotechnol. 233*, 160-170.
22. Choi, Y. J., Morel, L., Bourque, D., Mullick, A., Massie, B., and Miguez, C. B. (2006) Bestowing inducibility on the cloned methanol dehydrogenase promoter (P_{mxaF}) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1, *Appl. Environ. Microbiol. 72*, 7723-7729.
23. Kaczmarczyk, A., Vorholt, J. A., and Francez-Charlot, A. (2013) Cumate-inducible gene expression system for sphingomonads and other *Alphaproteobacteria*, *Appl. Environ. Microbiol. 79*, 6795-6802.
24. Chubiz, L. M., Purswani, J., Carroll, S. M., and Marx, C. J. (2013) A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*, *BMC Res Notes 6*, 183.
25. Müller-Hill, B. (1996) *The lac Operon : a short history of a genetic paradigm*, Walter de Gruyter, Berlin ; New York.
26. Liu, Q., Kirchhoff, J. R., Faehnle, C. R., Viola, R. E., and Hudson, R. A. (2006) A rapid method for the purification of methanol dehydrogenase from Methylobacterium extorquens, *Protein Expr Purif 46*, 316-320.
27. Meier, I., Wray, L. V., and Hillen, W. (1988) Differential regulation of the Tn*10*-encoded tetracycline resistance genes *tet*A and *tet*R by the tandem *tet* operators O_1 and O_2, *EMBO J. 7*, 567-572.
28. Hillen, W., and Berens, C. (1994) Mechanisms underlying expression of Tn*10* encoded tetracycline resistance, *Annu. Rev. Microbiol. 48*, 345-369.

29. Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J., and Voigt, C. A. (2019) *Escherichia coli* "Marionette" strains with 12 highly optimized small-molecule sensors, *Nat. Chem. Biol. 15*, 196-204.
30. Knief, C., Frances, L., and Vorholt, J. A. (2010) Competitiveness of diverse *Methylobacterium* strains in the phyllosphere of *Arabidopsis thaliana* and identification of representative models, including *M. extorquens* PA1, *Microb. Ecol. 60*, 440-452.
31. Marx, C. J., Bringel, F., Chistoserdova, L., Moulin, L., Farhan Ul Haque, M., Fleischman, D. E., Gruffaz, C., Jourand, P., Knief, C., Lee, M. C., Muller, E. E., Nadalig, T., Peyraud, R., Roselli, S., Russ, L., Goodwin, L. A., Ivanova, N., Kyrpides, N., Lajus, A., Land, M. L., Medigue, C., Mikhailova, N., Nolan, M., Woyke, T., Stolyar, S., Vorholt, J. A., and Vuilleumier, S. (2012) Complete genome sequences of six strains of the genus Methylobacterium, *J. Bacteriol. 194*, 4746-4748.
32. Nayak, D. D., and Marx, C. J. (2014) Genetic and phenotypic comparison of facultative methylotrophy between *Methylobacterium extorquens* strains PA1 and AM1, *PloS one 9*, e107887.
33. Marx, C. J., and Lidstrom, M. E. (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria, *BioTechniques 33*, 1062-1067.
34. Marx, C. J. (2008) Development of a broad-host-range sacB-based vector for unmarked allelic exchange, *BMC Res Notes 1*, 1.
35. Zhang, Y., Buchholz, F., Muyrers, J. P., and Stewart, A. F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*, *Nat. Genet. 20*, 123-128.
36. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products, *Proc. Natl. Acad. Sci. USA 97*, 6640-6645.
37. Murphy, K. C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*, *J. Bacteriol. 180*, 2063-2071.
38. Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems, *Nat. Biotechnol. 31*, 233-239.
39. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems, *Science 339*, 819-823.
40. Reisch, C. R., and Prather, K. L. (2015) The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*, *Scientific reports 5*, 15096.
41. Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K. Y., Shimatani, Z., and Kondo, A. (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems, *Science 353*.
42. Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H., and Kondo, A. (2018) Deaminase-mediated multiplex genome editing in Escherichia coli, *Nature microbiology 3*, 423-429.
43. Jensen, S. I., Lennen, R. M., Herrgard, M. J., and Nielsen, A. T. (2015) Seven gene deletions in seven days: Fast generation of Escherichia coli strains tolerant to acetate and osmotic stress, *Scientific reports 5*, 17874.

44. Katashkina, J. I., Hara, Y., Golubeva, L. I., Andreeva, I. G., Kuvaeva, T. M., and Mashko, S. V. (2009) Use of the lambda Red-recombineering method for genetic engineering of Pantoea ananatis, *BMC Mol. Biol. 10*, 34.
45. Hansen-Wester, I., and Hensel, M. (2002) Genome-based identification of chromosomal regions specific for Salmonella spp, *Infect. Immun. 70*, 2351-2360.
46. Ranallo, R. T., Barnoy, S., Thakkar, S., Urick, T., and Venkatesan, M. M. (2006) Developing live Shigella vaccines using lambda Red recombineering, *FEMS Immunol. Med. Microbiol. 47*, 462-469.
47. Wu, Y., Li, T., Cao, Q., Li, X., Zhang, Y., and Tan, X. (2017) RecET recombination system driving chromosomal target gene replacement in *Zymomonas mobilis*, *Electron. J. Biotechnol. 30*, 118-124.
48. Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes, *Science 321*, 960-964.
49. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, *Science 337*, 816-821.
50. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013) RNA-guided human genome engineering via Cas9, *Science 339*, 823-826.
51. Bowater, R., and Doherty, A. J. (2006) Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining, *PLoS Genet. 2*, e8.
52. Cui, L., and Bikard, D. (2016) Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*, *Nucleic Acids Res. 44*, 4243-4251.
53. Mo, X. H., Zhang, H., Wang, T. M., Zhang, C., Zhang, C., Xing, X. H., and Yang, S. (2020) Establishment of CRISPR interference in *Methylorubrum extorquens* and application of rapidly mining a new phytoene desaturase involved in carotenoid biosynthesis, *Appl. Microbiol. Biotechnol.*

ACKNOWLEDGEMENTS

I would like to thank my PhD supervisor **Tobias Erb** for trusting me with this project. I am very grateful for his guidance, encouragement, and constant support throughout. The same goes to my thesis advisory committee members **Prof. Anke Becker** and **Dr. Hannes Link**.

Thanks to all **Erblings** (past and present) who supported me during my PhD – your help, suggestions, and jokes all at the core of our lab. Thanks to **Jan** for being a general source of wisdom, always being willing to take extra time to answer my questions, and for putting out literal and figurative fires in the lab. Thanks to **Basti** for sharing his owl knowledge, his bench, and for our endless discussions about books. Thanks to **Marieke** for always making me look on the bright side of things. I am very grateful to **Gabo** for his motivational speeches and always cheering me on. Thanks to **Kyra** for her support in and outside of the lab. Thanks to **Simon** for brightening our lab and filling it with music. Thanks to **Thomi**, the first SYBORG, for teaching me how to purify enzymes and measure enzyme kinetics. Thanks to **Melanie** for all her work and ever willing to go the extra mile. I am very grateful to **Sonja** for her vital support during my PhD. Thanks to **Francesca** for her helpful feedback on this thesis. And of course, to my students **Luise**, **Amelie**, **Florian**, and **Mikhail** for their dedication and contributions to my projects.

Many, many thanks to my **friends** and **family**. To dearest **Robin** for all the fun times escaping the little bubble of Marburg. Un enorme abrazo a mis dos **Abuelitas** que siempre me han querido y apoyado con todo su amor. Muchas gracias **twina Anahi** por tu amistad incondicional todos estos años. Muchas gracias a mi **madre Margarita**, mis primos **Ernesto** y **Dani** y a mis dos **ñañas Cami** y **Manu**. A mi querido **papá Patricio** por su apoyo todos estos años, por motivarme y ayudarme a cumplir mis sueños.

CURRICULUM VITAE

Martina Carrillo

E-mail: martina.carrillo@mpi-marburg.mpg.de
Date of Birth: 17.08.1992
Nationality: Ecuadorean

Education and Research Experience

09/2015 – present	**PhD student**, Max-Planck-Institute for terrestrial Microbiology. Marburg, Germany. Principal Investigator: Prof. Dr. Tobias Erb. Project: "Implementation of CO_2 fixation pathways into *Methylorubrum extorquens* AM1".
09/2013 – 06/2015	**Master of Science in Molecular Life Sciences**, Jacobs University Bremen. Bremen, Germany. Graduated top 5% of class. Specialization areas: Molecular Biotechnology and Molecular Genetics.
07 – 09/2012	**Amgen Scholars Programme**, Cambridge, UK. Full-time intern the group of Prof. Dr. Alan Tunnacliffe at the University of Cambridge.
09/2010 – 06/2013	**Bachelor of Science in Biotechnology**, Jacobs University Bremen. Bremen, Germany. Graduated top 5% of class.
09/2003 – 05/2013	**International Baccalaureate Diploma**, Quito, Ecuador. American School of Quito. Degree focus on Natural Sciences.

Peer-reviewed Publications

Carrillo, M., Wagner, M., Petit, F., Dransfeld, A., Becker, A., Erb, T.J. Design and Control of Extrachromosomal Elements in *Methylorubrum extorquens* AM1. *ACS Synth. Biol*. 8 (**2019**): 2451-2456.

Schada v. B., L., **Carrillo, M.**, Leupold, S., Glatter, T., Kiefer, P., Weishaupt, R., Heinemann, M., Erb, T.J. An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. *Metab. Eng.* 47 (**2018**): 423-433.

Döhlemann, J., Wagner, M., Happel, C., **Carrillo, M.**, Sobetzko, P., Erb, T.J., Thanbichler, M., and Becker, A. A Family of Single Copy *repABC*-Type Shuttle Vectors Stably Maintained in the Alpha-Proteobacterium *Sinorhizobium meliloti*. *ACS Synth. Biol.* 6 (**2017**): 968-984

Klein, M., **Carrillo, M.**, Xiberras, J., Islam, Z.U., Swinnen, S., and Nevoigt, E. Towards the exploitation of glycerol's high reducing power in *Saccharomyces cerevisiae*-based bioprocesses. *Metab. Eng.* 38 (**2016**): 464-472.

Klein, M., Islam, Z.U., Knudsen, P.B., **Carrillo, M.**, Swinnen, S., Workman, M., and Nevoigt, E. The expression of glycerol facilitators from various yeast species improves

growth on glycerol of *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* 3 (**2016**): 252–257.

Barbosa, E.G., Crisp, A., Broadbent, S. E., **Carrillo, M.**, Boschetti, C., and Tunnacliffe, A. A functional difference between native and horizontally acquired genes in bdelloid rotifers. *Gene* 590 (**2016**): 186-191.

Swinnen, S., Klein, M., **Carrillo, M.**, McInnes, J., Nguyen, H.T.T., and Nevoigt, E. Re-evaluation of glycerol utilization in *Saccharomyces cerevisiae*: characterization of an isolate which grows on glycerol without supporting supplements. *Biotechnol. Biofuels* 6 (**2013**): 157.

Others

Languages English (fluent, C2), German (advanced, C1), Spanish (native), and French (beginner, A2).

Hobbies Football, indoor rock climbing.

Marburg, 23.04.2020