DISS. ETH NO. 24593

# Characterization of novel methanol dehydrogenases and identification of the gene set required for methylotrophic growth

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
Andrea Maria Ochsner

MSc., ETH Zurich

born on 24.06.1989
citizen of
Escholzmatt-Marbach (LU)

accepted on the recommendation of
*Prof. Dr. Julia A. Vorholt*
*Prof. Dr. Beat Christen*
*Prof. Dr. Jörn Piel*

2017

# Table of Contents

# Abstract

Methylotrophy is the ability of microorganisms to use reduced one-carbon compounds, including methanol and methane, as the sole source of carbon and energy. The process has raised great interest from both basic and applied angles. Methylotrophs play an important role in the global carbon cycle by metabolizing methanol, pollutants such as dichloromethane, as well as the potent climate gas methane. Moreover, methylotrophs have been proposed as platform organisms for sugar-independent production of value-added chemicals. The plant-colonizing Alphaproteobacterium *Methylobacterium extorquens* is the most extensively studied methylotroph. Its biochemistry has been elucidated in the past 60 years, mostly by targeted approaches, but more recently also by complementary "omics" techniques such as proteomics and metabolomics.

This thesis presents new insights into the biochemistry of the first reaction of the methylotrophic metabolism, the oxidation of methanol to formaldehyde, and describes the essential genome of the model methylotroph *M. extorquens* PA1.

The initial metabolic step for growth on methanol is the oxidation of methanol catalyzed by methanol dehydrogenase (Mdh). Two fundamentally different enzyme systems exist in bacterial methylotrophs; a pyrroloquinoline quinone (PQQ)-dependent Mdh employed by Gram-negative methylotrophs and a nicotinamide adenine dinucleotide (NAD)-dependent Mdh employed by Gram-positive methylotrophs. The NAD-dependent Mdh of the model methylotroph *Bacillus methanolicus* shows a low catalytic efficiency *in vitro*, which is increased about 20-fold in the presence of the endogenous activator protein Act. The exact mechanism and the *in vivo* role of this activation remain poorly understood. Investigation of the specificity of the activation *in vitro* revealed that the catalytic efficiency of all investigated type III alcohol dehydrogenases is improved by Act, suggesting that the activation is not specific to methylotrophy, as previously assumed. Mdhs depending on the cofactor PQQ are several orders of magnitude more efficient in oxidizing methanol. In addition to the well-studied calcium-dependent PQQ-Mdh MxaFI, a novel type of PQQ-Mdh termed XoxF depending on rare-earth elements (REEs) instead of calcium has recently been identified. Intriguingly, the XoxF system was found to be more widespread in metagenomes as well as more abundantly expressed in natural habitats of methylotrophs, such as the phyllosphere. Investigation of the expression of the two systems in *M. extorquens* PA1 revealed that the REE lanthanum dominantly switches the expression from *mxaFI* to *xoxF*, suggesting that XoxF is the preferred system. Further investigation implied that the underlying regulatory cascade is more complex than assumed and involves one or more so-far unknown component(s).

In spite of the great interest in methylotrophs, the entire gene set necessary for growth on methanol has not yet been defined for any methylotroph. In this thesis, transposon sequencing (TnSeq) was employed to define all genetic elements of *M. extorquens* PA1 that are specifically required for growth

on methanol. This approach revealed almost 100 new genes that are essential or fitness-relevant for growth on methanol, illustrating knowledge gaps in methylotrophy. Surprisingly, one of these new methylotrophy genes was phosphoribulokinase that so far had been attributed a unique role in autotrophy. Follow-up experiments revealed its involvement in the regulation of one-carbon assimilation genes, suggesting a conserved role shared in autotrophy and methylotrophy.

Taken together, this thesis further characterized two divergent methanol dehydrogenase systems and defined the entire gene set required for growth of *M. extorquens* on methanol. The latter represents a valuable resource for future research towards a systems-level understanding of methylotrophy.

# Zusammenfassung

Methylotrophie bezeichnet die Fähigkeit von Mikroorganismen, reduzierte Ein-Kohlenstoff-verbindungen wie Methanol und Methan als alleinige Quelle von Kohlenstoff und Energie zu verwenden. Dieser Prozess ist für die Grundlagenforschung als auch angewandte Forschung von grossem Interesse. Methylotrophe spielen eine wichtige Rolle im globalen Kohlenstoffkreislauf, indem sie Methanol, schädliche Lösungsmittel wie Dichlormethan und das Treibhausgas Methan umwandeln. Darüber hinaus sind Methylotrophe potentielle Plattformorganismen für die Zucker-unabhängige Herstellung von wertschöpfenden Chemikalien. Das Pflanzen-besiedelnde Alphaproteobakterium *Methylobacterium extorquens* ist der am besten untersuchte Methylotrophe und wesentliche Beiträge zur Aufklärung seiner Biochemie wurden in den letzten 60 Jahren geleistet. Dazu wurden vor allem gezielte Ansätze, kürzlich aber auch vermehrt komplementäre systemweite Methoden wie die Proteomik und Metabolomik, verwendet.

Diese Arbeit präsentiert neue Einblicke in die Biochemie der ersten Reaktion des methylotrophen Stoffwechsels, der Oxidation von Methanol zu Formaldehyd, und beschreibt das essentielle Methylotrophie-Genom des Modell-Methylotrophen *M. extorquens* PA1.

Der einleitende Schritt für Wachstum auf Methanol, die Oxidation von Methanol, wird von einer Methanol Dehydrogenase (Mdh) katalysiert. Zwei grundlegend verschiedene Enzymsysteme existieren in bakteriellen Methylotrophen: Gram-negative Methylotrophe benutzen eine Pyrrolochinolinchinon (PQQ)-abhängige Mdh, während Gram-positive Methylotrophe eine Nicotinamidadenindinukleotid (NAD)-abhängige Mdh verwenden. Die NAD-abhängige Mdh von *Bacillus methanolicus* weist eine niedrige katalytische Effizienz auf, die jedoch in der Anwesenheit des endogenen Aktivator-Proteins Act ungefähr 20-fach gesteigert wird. Der exakte Mechanismus dieser Aktivierung und deren Rolle in der Zelle sind noch nicht verstanden. Untersuchungen zur Spezifität dieser Aktivierung zeigten, dass die katalytische Effizienz aller getesteten Typ III Alkohol Dehydrogenasen verbessert werden kann. Dies weist darauf hin, dass der Aktivierungsprozess nicht wie bis anhin angenommen, spezifisch für die Methylotrophie ist. PQQ-abhängige Mdhs sind um mehrere Grössenordnungen effizienter. Zusätzlich zur gut studierten Kalzium-abhängigen PQQ-Mdh MxaFI wurde kürzlich ein neuer Typ von PQQ-Mdhs entdeckt. Diese XoxF genannten Enzyme verwenden statt Kalzium Seltene Erden in ihrem aktiven Zentrum. Interessanterweise konnte in Metagenomstudien gezeigt werden, dass XoxF weiter verbreitet ist und im natürlichen Lebensraum der Organismen, z.B. in der Phyllosphäre, auch verstärkt hergestellt wird. Untersuchungen zur Expression von beiden PQQ-Mdh Systemen in *M. extorquens* PA1 ergaben, dass die Zugabe von Lanthan den Wechsel von MxaFI zu XoxF herbeiführt. Dies deutet darauf hin, dass das XoxF-Enzymsystem das bevorzugte der beiden ist. Weiterführende Untersuchungen deuteten an, dass die

zugrundeliegende Regulationskaskade komplex ist und eine oder mehrere noch unbekannte Komponenten beteiligt sind.

Ungeachtet des grossen Interessens an Methylotrophen, wurde der vollständige Satz an für die Methylotrophie benötigten Genen bisher nicht ermittelt. In dieser Arbeit wurde Transposonmutagenese verbunden mit Illumina-Sequenzierung dazu verwendet, alle genetischen Elemente von *M. extorquens* für das Wachstum in Gegenwart von Methanol, zu identifizieren. Dieser Ansatz hat insgesamt beinahe 100 bislang nicht charakterisierte Gene enthüllt, die wichtig für Wachstum auf Methanol sind und hat dabei veranschaulicht, dass es noch Wissenslücken gibt. Überraschenderweise kodiert eines dieser neuen Methylotrophie-Gene für das Enzym Phosphoribulokinase, dem man bisher lediglich eine einzigartige Rolle in der Autotrophie zugeschrieben hat. Weiterführende Experimente haben ergeben, dass das Enzym an der Regulierung von Genen, die zur Assimilation von Ein-Kohlenstoffverbindungen benötigt werden, beteiligt ist. Dies weist auf eine in Autotrophen und Methylotrophen konservierte Funktion hin.

Zusammenfassend hat diese Arbeit die beiden grundsätzlich verschiedenen Methanol Dehydrogenase Systeme weiter untersucht und den gesamten genetischen Satz, der von *M. extorquens* für Wachstum auf Methanol benötigt wird, bestimmt. Letzterer stellt eine wertvolle Quelle für weiterführende Forschung, mit dem Ziel des systemweiten Verständnisses von Methylotrophie, dar.

## Chapter 1:

# Introduction

## 1.1 Aerobic methylotrophy and its importance

Methylotrophy is the ability to grow at the expense of reduced one-carbon compounds, such as methanol or methane, as sole carbon source. Unlike autotrophs, which need to generate energy from a secondary source (such as light or reduced inorganic compounds), methylotrophs obtain the energy required for growth from their reduced substrate. Bacterial methylotrophs are not restricted to one phylogenetic group, but are spread across different phyla and employ different metabolic solutions (see section 1.2), suggesting that methylotrophy has evolved more than once. Methylotrophs can be divided into two classes, obligate methylotrophs that are specialized for growth on one-carbon substrates and facultative methylotrophs that are also able to grow on multicarbon substrates.

Methylotrophs have been isolated in a myriad of habitats including freshwater lakes (Chistoserdova *et al*, 2013), soil (Doronina *et al*, 1996), and the phyllosphere (Knief *et al*, 2010). The latter is abundantly colonized by methylotrophs, especially from the genus *Methylobacterium*. In this habitat they benefit from methanol produced during cell wall metabolism of the plant (Delmotte *et al*, 2009; Vorholt, 2012). The ability to utilize methanol has been shown to be an advantage during competitive plant colonization (Sy *et al*, 2005; Abanda-Nkpwatt *et al*, 2006). Methylotrophs play an important role in the global carbon cycle and are involved in the degradation of pollutants (Gälli & Leisinger, 1985) as well as of compounds involved in climate change. Methanotrophs, a subclass of methylotrophs able to utilize methane, were estimated to capture up to 95% of methane produced in biogenic and non-biogenic processes (Chistoserdova, 2015).

In addition to their importance in the environment, methylotrophs have raised interest as platform organisms in sugar-independent biotechnology. This is especially true for methanol as a substrate, which is already an important building block for chemical industry (Bertau *et al*, 2014) and a "methanol economy" has been suggested (Olah, 2013). Currently most methanol is produced from fossil materials, but the potential for production from sustainable, as well as renewable resources exists (Schrader *et al*, 2009).

## 1.2 The biochemical basis of methylotrophy

Methylotrophy refers to the growth on a large number of reduced carbon sources lacking carbon-carbon bonds with oxidation levels of –IV (methane), -II (e.g. methanol, methylamine, chloromethane) and 0 (e.g. dichloromethane). To allow growth on reduced one-carbon as sole carbon and energy source, the methylotroph has to fulfill two basic requirements, the assimilation of carbon and the production of energy. Different solutions are found in nature and can simplistically be divided into the following metabolic steps; (1) oxidation of primary one-carbon substrate to formaldehyde (exceptions

see below), (2) oxidation of formaldehyde to $CO_2$, and (3) assimilation of one-carbon into biomass (Chistoserdova, 2011) (Fig. 1).

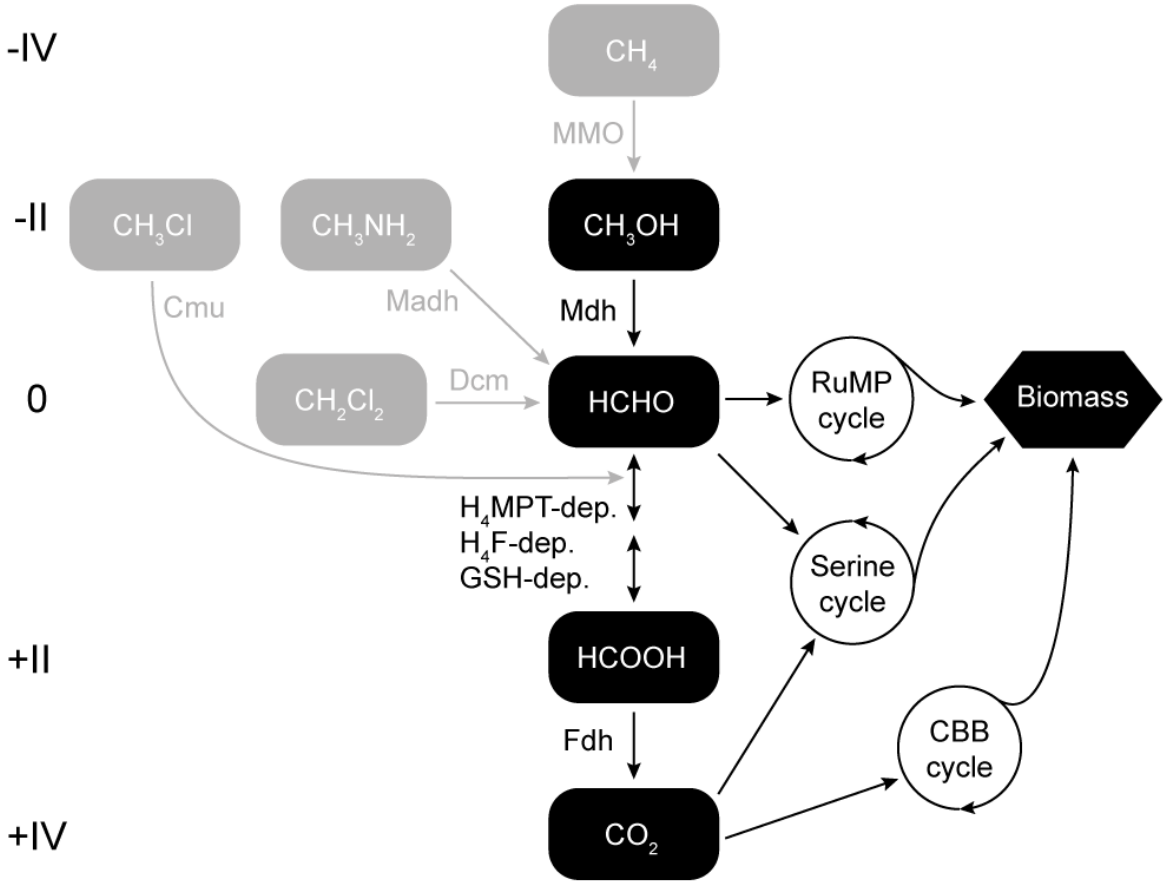


**Figure 1:** Overview of methylotrophic metabolism. $CH_4$, methane; $CH_3OH$, methanol; $CH_3NH_2$, methylamine; $CH_3Cl$, chloromethane; $CH_2Cl_2$, dichloromethane; HCHO, formaldehyde; HCOOH, formate; MMO, methane monooxygenase; Cmu, chloromethane methyltransferase; Madh, methylamine dehydrogenase; Mdh, methanol dehydrogenase; Dcm, dichloromethane dehalogenase; RuMP cycle, ribulose-monophosphate cycle; CBB cycle, Calvin-Benson-Bassham cycle; $H_4MPT$, tetrahydromethanopterin; $H_4F$, tetrahydrofolate; GSH, glutathione; Fdh, formate dehydrogenase.

**Primary oxidation of reduced one-carbon compounds**

Growth on the most reduced one-carbon compound methane ($CH_4$), also referred to as methanotrophy, starts with the oxidation of methane to methanol catalyzed by methane monooxygenase (MMO), for which two forms are known, a membrane-bound and a soluble one (Semrau *et al*, 2010).

Methanol ($CH_3OH$) is oxidized to formaldehyde by methanol dehydrogenase (Mdh). The two profoundly different systems dependent on different redox cofactors used by Gram-negative

(pyrroloquinoline quinone, PQQ) and Gram-positive methylotrophs (nicotinamide adenine dinucleotide, NAD) are described in section 1.3.

Methylamine ($CH_3NH_2$) is oxidized to formaldehyde either by methylamine dehydrogenase (Madh) (Shirai *et al*, 1978) or alternatively using the *N*-Methylglutamate pathway (Nayak & Marx, 2014a). Chloromethane ($CH_3Cl$) is metabolized by methyl-transfer to tetrahydrofolate ($H_4F$) resulting in methyl-$H_4F$ which is further oxidized to formate using an oxidative $H_4F$ pathway (Studer *et al*, 2002). Dichloromethane ($CH_2Cl_2$) is dehalogenated twice in a glutathione-dependent reaction (Muller *et al*, 2011). A more comprehensive list of carbon sources that are metabolized using methylotrophic pathways is given in (Chistoserdova, 2011).

**Dissimilation of formaldehyde to $CO_2$**

The formaldehyde produced in the first reaction is highly toxic for all living organisms and keeping its levels low is crucial for survival (Vorholt, 2002). A way to deal with the intermediate formaldehyde produced in the first step is the oxidation to $CO_2$. However, careful regulation of this process is crucial, because intermediates (or formaldehyde itself) are the starting points of most one-carbon assimilation cycles (see below). Several pathways catalyzing the oxidation of formaldehyde to formate and eventually $CO_2$ exist.

Pathways using a thiol as one-carbon carrier are widespread. The most common pathway depends on the antioxidant glutathione (GSH) and is not restricted to methylotrophs. It is found in methylotrophs such as *Paracoccus denitrificans* (Ras *et al*, 1995), as well as in many non-methylotrophs such as such as *Escherichia coli*, where it is involved in the detoxification of formaldehyde (Gonzalez *et al*, 2006). A similar pathway based on bacillithiol (BSH) (Gaballa *et al*, 2010), a functional GSH-analog of *Bacillus* species, was recently postulated in the methylotroph *B. methanolicus* (Müller *et al*, 2015b).

Another pathway is based on tetrahydrofolate ($H_4F$), which is present in most forms of life and involved in the interconversion of one-carbon units needed for the biosynthesis of essential building block such as purines (Maden, 2000). Two variants of the $H_4F$-dependent pathway exist in methylotrophs. The first variant relying on two enzymes including the multifunctional enzyme FolD is also present in non-methylotrophs. In methylotrophs, this variant seems to be strictly required for methylotrophic growth on substrates degraded by methyl-transfer such as chloromethane (Studer *et al*, 2002). A second variant relying on three separate enzymes is found in some methylotrophs such as *M. extorquens* (see below), where evidence points towards operation in the reductive direction (Crowther *et al*, 2008), therefore being part of assimilation rather than dissimilation (see section 1.4).

Another one-carbon interconversion pathway depends on tetrahydromethanopterin ($H_4MPT$), which is structurally similar to $H_4F$. It was first described in methanogenic Archaea, where it is involved in the reduction of $CO_2$ to methane (Escalante-Semerena *et al*, 1984). Surprisingly the genes were

subsequently discovered in the methylotroph *M. extorquens* AM1 where the pathway was found to operate in the reverse direction, oxidizing formaldehyde to formate (Chistoserdova *et al*, 1998). In spite of structural similarity to $H_4F$, the pathway contains distinct intermediates and relies on specific enzymes including the involvement of a second one-carbon carrier (see section 1.4).

The final product of all above-mentioned pathways (if they run in the oxidative direction) is formate, which is further oxidized to $CO_2$ by formate dehydrogenase. Various formate dehydrogenases differing in their cellular localization and cofactor-requirements are known (Vorholt & Thauer, 2002).

A special solution for dissimilation of formaldehyde is represented by the oxidative branch of the ribulose-monophosphate (RuMP) cycle, which is closely connected to assimilation (see below). In this pathway, one-carbon is released as $CO_2$ circumventing the need of formate dehydrogenases.

**Assimilation of one-carbon into biomass**

To divide, a cell has to rebuild itself using the building blocks of life. In the case of an organism growing on a one-carbon substrate, every carbon-carbon bond has to be newly formed. Similar to almost all autotrophic $CO_2$ fixation pathways, all known pathways for carbon assimilation in methylotrophs are based on fixing the one-carbon unit to a multicarbon acceptor molecule. To ensure continuous growth, the acceptor molecule has to be regenerated; therefore, all assimilation pathways are cyclic. The carbon output of the cycle is a small multicarbon unit that is further metabolized to generate all building blocks required for growth. Three different cycles for the assimilation of one-carbon into biomass are known in methylotrophs.

The ribulose-monophosphate (RuMP)-cycle directly incorporates formaldehyde by condensation with ribulose-5-phosphate generating hexulose-6-phosphate. Hexulose-6-phosphate is further metabolized, using reactions shared with glycolysis and the pentose phosphate pathway, to regenerate the acceptor ribulose-5-phosphate and to produce a C3-output (dihydroxyacetone phosphate/ glyceraldehyde-3-phosphate) every three turns. Alternatively, the assimilated one-carbon unit is oxidized to $CO_2$ using the oxidative branch of the RuMP cycle, to gain two reduction equivalents per one-carbon.

The serine cycle incorporates one-carbon in the form of methylene-$H_4F$ by condensation with glycine producing serine. Serine is then converted to glyoxylate and acetyl-CoA via a reaction sequence including one $CO_2$-fixation step. Glyoxylate is used to regenerate glycine to close the cycle and acetyl-CoA is the output unit of the cycle. The involved reactions are found in common pathways for glycine biosynthesis, glycolysis or the tricarboxylic acid (TCA) cycle. For the production of all building blocks, intermediates have to be withdrawn from the cycle resulting in its depletion. To ensure continuous operation, intermediates of the cycle have to be replenished from acetyl-CoA. Even though this could be achieved by the well-known glyoxylate shunt, most methylotrophs seem to rely on a

complex series of conversions of CoA-intermediates, the ethylmalonyl-CoA pathway (Erb *et al*, 2007; Peyraud *et al*, 2009) (see section 1.4).

The Calvin-Benson-Bassham (CBB) cycle is used by some methylotrophs to fix $CO_2$ produced by dissimilation reduced one-carbon (and exogenous $CO_2$) and utilize the produced reduction equivalents to fuel this energy demanding process. This metabolism can also be seen as a form of chemoorgano-autotrophy, where the organic compound is only used as electron donor similar to the inorganic compound in chemolitho-autotrophy. Many CBB-methylotrophs are also able to grow as true autotrophs, such as *Paracoccus denitrificans*, which is able to use methanol as "electron-donor", but also hydrogen or thiosulfate (Harms *et al*, 1996).

## 1.3 The enzymes catalyzing the initial step of methanol metabolism

### Pyrroloquinoline quinone (PQQ)-dependent methanol dehydrogenase

In all known Gram-negative methylotrophs, methanol oxidation is catalyzed by a periplasmatic pyrroloquinoline quinone (PQQ)-dependent methanol dehydrogenase (PQQ-Mdh). One of the best-studied PQQ-Mdh is the heterotetrameric MxaFI of *M. extorquens* AM1 (Williams *et al*, 2005). In addition to the prosthetic group PQQ, the enzyme contains a $Ca^{2+}$ that acts as a Lewis-acid polarizing the C5=O bond of the PQQ to facilitate the nucleophilic attack of the substrate hydride (Anthony & Williams, 2003). The reduced PQQ ($PQQH_2$) produced during the oxidation of methanol transfers its electrons to a specific cytochrome *c* (cyt $c_L$) and then via a second cytochrome *c* to the respiratory chain (Anthony, 1992). The genes encoding the two subunits as well as cyt $c_L$ are clustered with other genes required for the generation of a functional enzyme, such as genes involved in the insertion of the $Ca^{2+}$ ion (Richardson & Anthony, 1992; Morris *et al*, 1995).

A homolog of MxaF, termed “XoxF”, was first discovered in the genome of the methylotroph *Paracoccus denitrificans* (Harms *et al*, 1996) and subsequently in the vast majority of Gram-negative methylotrophs (Keltjens *et al*, 2014). The enzyme is expressed at low levels under laboratory conditions and was shown to possess low methanol and formaldehyde oxidation activity under these conditions (Schmidt *et al*, 2010). Surprisingly, XoxF is required for growth on methanol due to its role in the regulation of *mxaFI* expression (Skovran *et al*, 2011). Moreover, XoxF was identified as highly abundant in the phyllosphere (Delmotte *et al*, 2009) and a knockout was shown to have a disadvantage during plant colonization (Schmidt *et al*, 2010). A crucial step towards the unraveling of the elusive function of XoxF enzymes was the discovery that the addition of the rare-earth element (REE) lanthanum restores growth of an *M. extorquens* mutant lacking the classical PQQ-Mdh MxaFI (Nakagawa *et al*, 2012). Subsequent studies confirmed that REEs indeed replace $Ca^{2+}$ in the active site

of the enzyme (Pol *et al*, 2014), marking the discovery of the first enzyme specifically depending on REEs (Chistoserdova, 2016).

**Nicotinamide adenine dinucleotide (NAD)-dependent methanol dehydrogenases**

A different type of methanol dehydrogenase was discovered in Gram-positive methylotrophs, such as *Bacillus methanolicus* (Arfman *et al*, 1989). This enzyme belongs to the type III NAD(P)-dependent alcohol dehydrogenases (NAD-Mdh) (de Vries *et al*, 1992), is homodecameric and each subunit contains 1 $Zn^{2+}$ as well as 1 or 2 $Mg^{2+}$ ions (Vonck *et al*, 1991) and a non-covalently, but tightly bound NAD(H) (Arfman *et al*, 1997). Compared to PQQ-Mdhs, NAD-Mdhs show a several orders of magnitude lower substrate affinity and reaction rate (Krog *et al*, 2013). In spite of the inferior catalytic properties and the reaction being thermodynamically unfavorable (Whitaker *et al*, 2015), the enzyme supports high growth rates in *B. methanolicus* at 50°C (Müller *et al*, 2015b). Further investigation revealed the presence of an endogenous activator protein, termed Act, that increases the *in vitro* activity at low methanol concentrations roughly 40-fold (Arfman *et al*, 1991). Furthermore, this activation was only observed for the oxidation of methanol to formaldehyde, but not the reverse reaction (Arfman *et al*, 1997). Sequence analysis of Act revealed high homology to the so-called Nudix hydrolases (Kloosterman *et al*, 2002). Enzymes of this family hydrolyze various substrates containing a nucleoside diphosphate moiety, such as GDP-sugars and ADP-ribose (McLennan, 2006). Activity assays with purified Act, revealed the hydrolysis of several substrates including $NAD^+$, which is hydrolyzed to AMP and nicotinamide mononucleotide (NMN). Intriguingly, Act also catalyzes the hydrolysis of Mdh-bound $NAD^+$ leaving only the AMP-moiety bound, which was suggested to improve the catalytic efficiency of the enzyme (Kloosterman *et al*, 2002). The *in vivo* relevance of this activation and the requirement of Act for growth on methanol remain unknown.

## 1.4 *Methylobacterium extorquens* - a model methylotroph

Ever since its isolation more than 50 years ago (Peel & Quayle, 1961), *M. extorquens* AM1 (formerly *Pseudomonas* AM1) has been a model organism for methylotrophy (Anthony, 2011). In addition to the one-carbon substrates methanol and methylamine, it grows on several multicarbon sources including succinate (C4) and acetate (C2), making it a facultative methylotroph (Large *et al*, 1961; Schneider *et al*, 2012). Its facultative metabolism is a useful trait, because it enables comparative studies under methylotrophic and non-methylotrophic growth conditions.

Many pathways and enzymes involved in methylotrophy that are known to be widespread in methylotrophs, have initially been discovered in *M. extorquens* (Chistoserdova, 2011). The innovative work of Quayle and colleagues in the 1960ies led to the elucidation of the serine cycle as the first

methylotrophy-specific pathway for one-carbon assimilation (Large *et al*, 1961, 1962a, 1962b; Large & Quayle, 1963). It was recognized early on that the organism lacks the activity of isocitrate lyase, the key enzyme of the glycoxylate shunt, which would be the simplest solution for regenerating depleted intermediates of the serine cycle (Dunstan & Anthony, 1973). Nevertheless, the exact sequence of the pathway, which uses a series of CoA-linked intermediates has not been solved until more recently (Erb *et al*, 2007; Peyraud *et al*, 2009). Another surprising discovery was the operation of the $H_4$MPT-dependent pathway for formaldehyde dissimilation (see below), which was thought to be unique to methanogenic Archaea, where it operates in the reverse direction (Chistoserdova *et al*, 1998).

Apart from *M. extorquens* AM1, more recently strain PA1, a closely related strain with identical 16S rRNA gene sequence, gains interest as a major model strain. Strain PA1 which was recently isolated from the *Arabidopsis* phyllosphere (Knief *et al*, 2010) and thus lacks a long history of domestication in the laboratory (Carroll *et al*, 2014). It was shown to have an overall conserved one-carbon metabolism, but superior growth rates on many substrates including methanol (Nayak & Marx, 2014b). In addition, it has a simpler genome structure (Marx *et al*, 2012) and is more suitable for transposon mutagenesis (Metzger *et al*, 2013).

**The methylotrophic pathways of *M. extorquens***

*M. extorquens* employs a PQQ-dependent methanol dehydrogenase (refer to section 1.3 for details). The $Ca^{2+}$-dependent enzyme MxaFI of *M. extorquens* was studied in great detail (reviewed in (Anthony & Williams, 2003)). In addition, one or several XoxF-type PQQ-Mdhs, two in AM1 and one in PA1, are encoded on the genome.

The formaldehyde produced by methanol dehydrogenase is further oxidized to formate using the $H_4$MPT-dependent pathway (Chistoserdova *et al*, 1998) (Fig. 2). The condensation of formaldehyde with $H_4$MPT can occur spontaneously (Escalante-Semerena *et al*, 1984), but is catalyzed by an enzyme termed formaldehyde-activating enzyme (Fae) in *M. extoquens* (Acharya *et al*, 2005; Vorholt *et al*, 2000). Mutants lacking *fae* are unable to grow on methanol and more sensitive to formaldehyde when growing on an alternative carbon source, suggesting that the spontaneous reaction is not sufficient for detoxification (Vorholt *et al*, 2000). The resulting methylene-$H_4$MPT is oxidized to methenyl-$H_4$MPT by one of two NAD(P)-dependent dehydrogenase MtdA and MtdB (Vorholt *et al*, 1998). The cyclic methenyl-$H_4$MPT is then hydrolyzed to formyl-$H_4$MPT by the cyclohydrolase Mch (Pomper *et al*, 1999). The formyl group is subsequently transferred to methylofuran (MYFR) (Hemmann *et al*, 2016) by the formyltransferase/hydrolase complex (Fhc) (Pomper & Vorholt, 2001). Unlike initially assumed based on the reaction of the methanogenic homolog, the complex catalyzes the hydrolysis of formyl-MYFR to formate, but not the oxidation to $CO_2$ (Pomper *et al*, 2002).

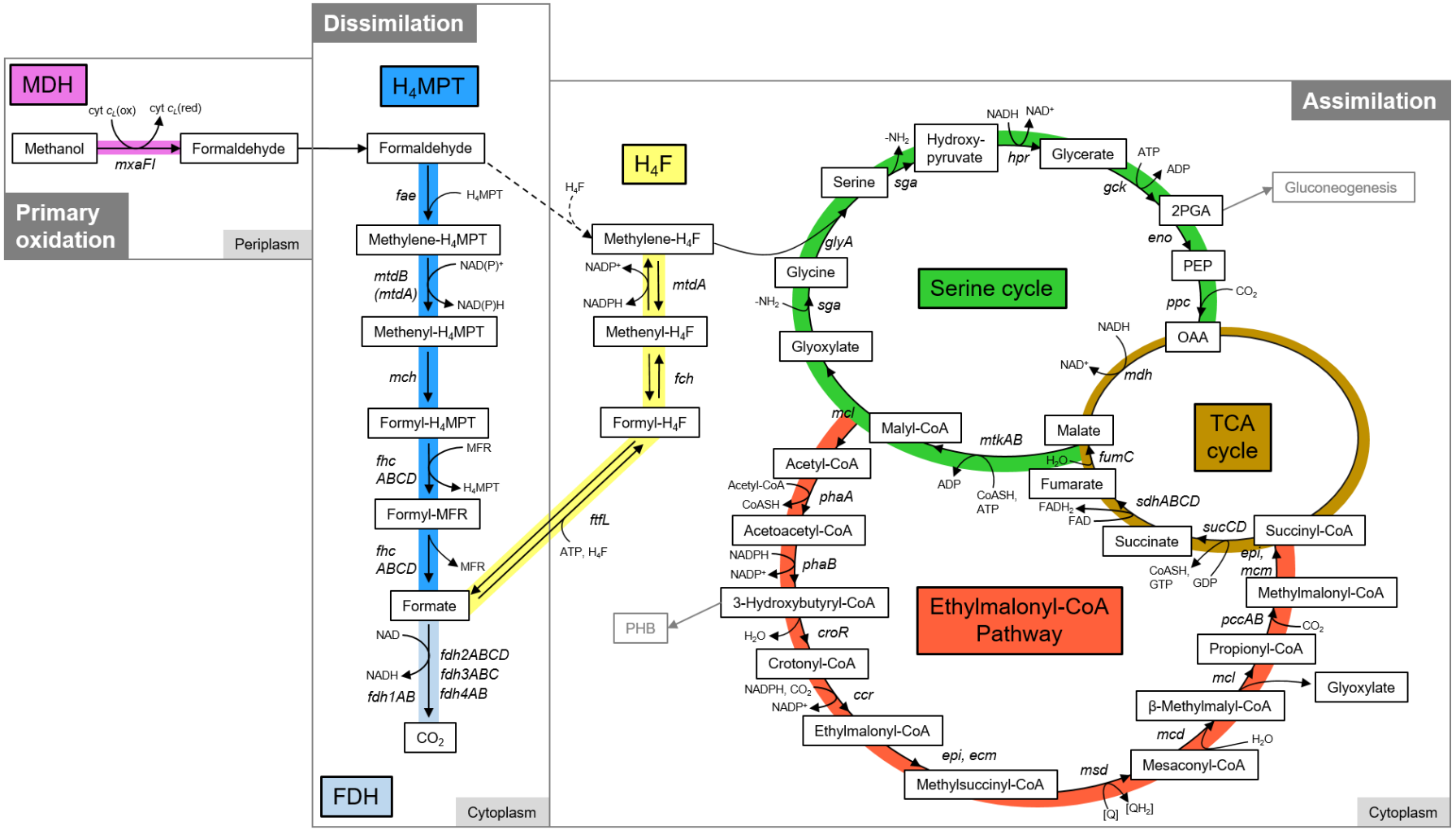


**Figure 2:** Central methylotrophic metabolism of *M. extorquens*. The coloring marks the different pathways. Genes encoding for the corresponding enzymes are shown in italics. This figure is taken from Ochsner *et al* (2014).

The released formate is either assimilated using the $H_4F$-dependent pathway operating in reductive direction or further oxidized to $CO_2$ by a formate dehydrogenase. Investigation of this metabolic step in *M. extorquens* led to the discovery of four formate dehydrogenases (Fdh1-4). Single knockouts have no (Δ*fdh1*/*2*/*3*) or only a weak (Δ*fdh4*) growth defect, while a triple knockout (Δ*fdh1-3*) accumulates formate and a quadruple mutant (Δ*fdh1-4*) is completely unable to grow on methanol (Chistoserdova *et al*, 2004, 2007). This high degree of redundancy points towards formate as an important metabolic intermediate.

One-carbon units are assimilated using the serine cycle, which starts with the condensation of methylene-$H_4F$ with glycine to form serine (Fig. 2). This reaction is catalyzed by serine hydroxymethyltransferase GlyA, which is not methylotrophy-specific but present in most organism where it runs in the opposite direction to generate methylene-$H_4F$ for biosynthetic purposes (Maden, 2000). The formed serine is then metabolized via C3 intermediates partly shared with gluconeogenesis (2-phosphoglycerate and phosphoenolpyruvate (PEP)). The carboxylation of PEP by PEP carboxylase (Ppc) results in C4-units shared with the TCA cycle (oxaloacetate and malate) which are then linked to CoA to form malyl-CoA. Malyl-CoA is split by malyl-CoA lyase (Mcl) resulting in glyoxylate which is used to regenerate the acceptor glycine, and the output molecule acetyl-CoA (Anthony, 1982).

The one-carbon unit methylene-$H_4F$ can be produced either directly by spontaneous condensation of formaldehyde and $H_4F$ or by an indirect route via formate. The indirect route starts by condensation of formate with $H_4F$ to form formyl-$H_4F$ followed by reductive operation of the $H_4F$-dependent pathways to form methylene-$H_4F$ (Marx *et al*, 2003). Even though the indirect route uses one additional ATP per one-carbon unit, recent evidence points to the operation of the indirect pathway in *M. extorquens*, further supporting the role of formate as a branching point (Crowther *et al*, 2008).

Due to the withdrawal of intermediates from the serine cycle for biosynthetic purposes, the cycle has to be replenished from the output molecule acetyl-CoA. The simplest solution is presented by the widespread glyoxylate shunt (Kornberg & Krebs, 1957) that could be used to produce a surplus of glyoxylate by shortcutting the decarboxylation steps of the TCA cycle. However, *M. extorquens* does not employ this pathway, but rather relies on a complex series of CoA-linked transformations (Korotkova *et al*, 2002, 2005). The complete pathway, termed the EMCP, has only been solved recently (Alber *et al*, 2006; Erb *et al*, 2007; Meister *et al*, 2005) and was subsequently shown to operate in *M. extorquens* (Peyraud *et al*, 2009). The pathway starts with the condensation of two acetyl-CoA units produced in the serine cycle followed by a series of CoA-linked transformations including one carboxylation step to produce methylmalyl-CoA. Methylmalyl-CoA is cleaved to propionyl-CoA and glyoxylate, which replenishes the serine cycle. Priopionyl-CoA is carboxylated to eventually re-enter the TCA cycle at the stage of succinyl-CoA closing the cycle. Unlike the serine cycle, the EMCP contains several specific reactions including crotonyl-CoA-reductase/carboxylase (Ccr) (Erb *et al*, 2007) and ethylmalonyl-CoA mutase (Ecm) (Erb *et al*, 2008).

## 1.5 Aims of this thesis

Methylotrophy has sparked great interest from a basic research viewpoint due to the involvement of methylotrophs in the global carbon cycle as well as from an applied perspective as promising process for sugar-independent biotechnology. In spite of this, the set of genetic requisites enabling methylotrophic growth has not yet been defined and many fundamental questions remain unanswered. The goal of this thesis was to improve the understanding of the initial reaction of methylotrophy both in Gram-negative and Gram-positive methylotrophs, which use fundamentally different systems and to define the methylotrophy genome of the model methylotroph *M. extorquens* PA1.

The NAD-dependent methanol dehydrogenase catalyzing methanol oxidation in Gram-positive methylotrophs such as the promising production host *B. methanolicus* is poorly understood. The catalytic efficiency of the isolated enzyme is low, but is improved by the activator protein Act. The specificity of this activation and its *in vivo* role remained unknown. In chapter 2, novel insights into the activation of NAD-Mdhs by Act are provided showing that activation is not specific, but widespread among type III alcohol dehydrogenases. Furthermore, attempts towards investigation of the *in vivo* role of Act by a synthetic biology approach are described.

In contrast, the PQQ-dependent Mdhs of Gram-negative methylotrophs have been studied in great detail over the last 50 years. A homolog of the initially discovered $Ca^{2+}$-dependent enzyme has recently been shown to depend on rare-earth elements. Like many known methylotrophs, *M. extorquens* encodes for both types of PQQ-Mdhs. In chapter 3, new insights into the regulation of PQQ-Mdhs in absence and presence of REEs and into the underlying regulatory cascade are presented.

Chapter 4 describes the determination of the entire gene set required for methylotrophy in *M. extorquens* by transposon sequencing. This approach uncovered almost 100 new genes specifically involved in growth on methanol including a novel core regulator for one-carbon assimilation.

General conclusions and future perspectives are presented in chapter 5.

# Chapter 2:

# *In vitro* activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases and attempts at *in vivo* investigation

Major parts of this chapter (corresponding to result parts 3.4.1 to 3.4.4) have been published in:

"*In vitro* activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases is more widespread than assumed"

Andrea M. Ochsner*, Jonas E. N. Müller*, Carlos A. Mora, Julia A. Vorholt

*Contributed equally to this paper

FEBS Letters 588 (2014) 2993–2999, doi: 10.1016/j.febslet.2014.06.008

**Author contributions**

A.M.O., J.E.N.M., and J.A.V. designed the research; A.M.O., J.E.N.M, and C.A.M. performed research; A.M.O. and J.E.N.M. analyzed data; A.M.O., J.E.N.M., and J.A.V. wrote the paper with input from all authors. Experiments corresponding to results part 3.4.5 were performed by Ralph Nüssli and A.M.O.

# Chapter 2: *In vitro* activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases and attempts at *in vivo* investigation

## 2.1 Abstract

In the Gram-positive methylotroph *Bacillus methanolicus*, methanol oxidation is catalyzed by an NAD-dependent methanol dehydrogenase (Mdh) that belongs to the type III alcohol dehydrogenase (Adh) family. It was previously shown that the *in vitro* activity of *B. methanolicus* Mdh is increased by the endogenous activator protein Act, a Nudix hydrolase. Here we show that this feature is not unique, but more widespread among type III Adhs in combination with Act or other Act-like Nudix hydrolases. In addition, we studied the effect of site directed mutations in the predicted active site of Mdh and two other type III Adhs with regard to activity and activation by Act. Furthermore, we present attempts to investigate the *in vivo* significance of the activation by Act by heterologous expression in the well-studied methylotroph *Methylobacterium extorquens*.

## 2.2 Introduction

Alcohol dehydrogenases (Adh) are a group of dehydrogenase enzymes that are present in numerous organisms and facilitate the interconversion between alcohols and aldehydes or ketones. They often use nicotinamide adenine dinucleotide ($NAD^+$ to NADH) as a cosubstrate. However, there are also quinoenzymes that require pyrroloquinoline quinone (PQQ) as an enzyme-bound electron acceptor. A typical example for this latter enzyme type is the methanol dehydrogenase (Mdh) of methylotrophic proteobacteria [1], while in Gram-positive methylotrophs, such as the thermophilic *Bacillus methanolicus*, NAD-dependent Mdhs are used [2]. Three NAD-dependent Mdhs are found in *B. methanolicus* MGA3: Mdh, Mdh2 and Mdh3 [3]. Mdh is encoded on the naturally occurring plasmid pBM19, which also harbors other genes encoding enzymes involved in methylotrophy, and is the principle enzyme for methanol oxidation [3]. The role of the two additional Mdh paralogs (termed Mdh2 and Mdh3) is not yet clear [3,4]. These enzymes are encoded on the chromosome and exhibit a sequence identity of 96% to each other and approximately 60% to Mdh [5]. Mdh was shown to have a decameric structure, with one $Zn^{2+}$ and one to two $Mg^{2+}$ ions bound to each subunit. Based on homology, this enzyme was classified as a type III alcohol dehydrogenase [6]. Mdh was proposed to follow a ping-pong reaction mechanism in which the electrons from methanol are first transferred to a

tightly bound NAD cofactor, which is subsequently re-oxidized by a free $NAD^+$ to complete the catalytic cycle [7]. The *in vitro* activity of Mdh was shown to be increased up to 40-fold by the endogenous activator protein Act [8], which is chromosomally encoded [5]. Act was found to contain a motif ($GX_5EX_7REUXEEXGU$) that is conserved in a family of enzymes that hydrolyze nucleoside diphosphates linked to a moiety X (Nudix) [9]. The substrates of these so-called Nudix hydrolases include nucleotide sugars, nucleoside-triphosphates and dinucleotide coenzymes [10]. Enzymes of this family are present in all classes of organisms and their function is believed to mainly involve house-keeping activities, such as the regulation of glycogen formation by the hydrolysis of ADP-glucose (e.g. NudF from *Escherichia coli*) or the degradation of the potentially mutagenic 8-oxo-dGTP (*E. coli* MutT) [9]. Act was shown to hydrolyze $NAD^+$ to AMP and $NMN^+$ and, with higher efficiency, ADP-ribose to AMP and ribose-5-phosphate [11]. The significance of this cleavage reaction is not entirely clear; notably, however, it was found that the incubation of Mdh with Act results in AMP that is bound to the enzyme [11]. It was suggested that bound $NAD^+$ is hydrolyzed by Act with a much higher efficiency than the free cofactor [11]. The detected increase in Mdh activity can be explained by a change of its ping-pong mechanism to a faster ternary complex mechanism in which free $NAD^+$ directly participates in the reaction by accepting electrons from methanol [12]. It was also proposed that the mutation of an active site serine (S97G), which results in the loss of the ability to tightly bind NAD(H), mimics the activating effect of Act on the $V_{max}$ [12]. In addition to its activating effect on the $V_{max}$, it was found that the presence of Act lowers the $K_m$ for methanol [5]. Later, it was shown that the other two NAD-dependent Mdhs from *B. methanolicus*, Mdh2 and Mdh3, are also activated by Act [5]. The *in vivo* relevance of Act for methylotrophic growth is currently unknown and difficult to investigate due to a lack of genetic tools for directed knockout mutagenesis in *B. methanolicus*.

In this study, we investigated the specificity of the activation of Mdh by Act by testing various NAD-dependent Adhs and Nudix hydrolases. In addition, we present novel insight into the active site of Mdh, Mdh2 and another type III Adh by structure prediction and site-directed mutagenesis. Our results contribute to the understanding of methanol oxidation in Gram-positive methylotrophs and are of relevance when considering *in vitro* applications of type III Adhs.

## 2.3 Results and discussion

### 2.3.1 Act increases the catalytic efficiencies of *B. methanolicus* Mdh and Mdh2

*B. methanolicus* MGA3 Mdh and Mdh2 activities were previously shown to be increased by Act under alkaline pH conditions [5]. However, because neutral intracellular pH can be assumed for *B. methanolicus* [13,14], we decided to additionally study the enzymes at pH 7.4 to represent more physiological conditions. The enzyme activities of Mdh and Mdh2 were determined using methanol as

a substrate at pH 9.0 and pH 7.4. As previously shown, the activities of both enzymes were higher at pH 9.0 compared to pH 7.4 [5]. The activating effect of Act was observed at both pH values (Tab. 1A). Also as previously shown, Act not only increased the $V_{max}$ of the enzymes but also decreased the $K_m$ for methanol [5]. The decrease in the $K_m$ was more pronounced for Mdh than for Mdh2 at both pHs. The increase in the $V_{max}$ was found to be similar at both pH values for Mdh but was more pronounced at pH 7.4 for Mdh2 (previously shown in [5]). To describe the overall effect of Act, the catalytic efficiency $k_{cat}/K_m$ was calculated. Act had a positive effect on the catalytic efficiency of Mdh and Mdh2 under all tested conditions. For Mdh, the increase was 23-fold at pH 7.4 and 50-fold at pH 9.0, for Mdh2, a 20-fold increase was observed at pH 7.4 and a 9-fold increase was observed at pH 9.0.

**Table 1.** Kinetic parameters for Mdh and Mdh2 with and without Act at pH 7.4 and 9.0 with (A) methanol or (B) ethanol as a substrate. Means ± 95 % confidence intervals are shown.

| **(A)** | **pH 7.4** | | | **pH 9.0** | | |
|---|---|---|---|---|---|---|
| | $V_{max}$ [mU/mg] | $K_m^{MeOH}$ [mM] | $k_{cat}/K_m^{MeOH}$ [$M^{-1}s^{-1}$] | $V_{max}$ [mU/mg] | $K_m^{MeOH}$ [mM] | $k_{cat}/K_m^{MeOH}$ [$M^{-1}s^{-1}$] |
| **Mdh** | 129±10 | 349±72 | 0.3 | 151±8 | 150±25 | 0.7 |
| **Mdh+Act** | 253±28 | 25±9 | 6.8 | 474±32 | 9±2 | 35.3 |
| **Mdh2** | 43±4 | 733±177 | 0.04 | 151±12 | 416±97 | 0.3 |
| **Mdh2+Act** | 317±23 | 255±45 | 0.8 | 394±16 | 96±12 | 2.8 |

| **(B)** | **pH 7.4** | | | **pH 9.0** | | |
|---|---|---|---|---|---|---|
| | $V_{max}$ [mU/mg] | $K_m^{EtOH}$ [mM] | $k_{cat}/K_m^{EtOH}$ [$M^{-1}s^{-1}$] | $V_{max}$ [mU/mg] | $K_m^{EtOH}$ [mM] | $k_{cat}/K_m^{EtOH}$ [$M^{-1}s^{-1}$] |
| **Mdh** | 208±17 | 225±47 | 0.6 | 457±36 | 224±51 | 1.4 |
| **Mdh+Act** | 566±49 | 51±17 | 7.4 | 3441±475 | 161±64 | 14.3 |

Mdh and Mdh2 have a diverse alcohol substrate spectrum, and activation by Act has been observed for ethanol as well [5]. To determine if the effect of Act is substrate independent, the Mdh activity with and without Act was measured for different substrates at both pH values (Fig. 1). Act showed an activating effect on Mdh activity with all of the tested substrates. The extent of the activation, however, was dependent on the substrate. To better understand this effect, Mdh was characterized kinetically with ethanol as a substrate at pH 7.4 and 9.0. The observed effect of Act on the $V_{max}$ was similar to the effect with methanol, but the effect on the $K_m$ was much smaller. Therefore, the resulting change in the $k_{cat}/K_m$ was smaller, 12-fold compared to 27-fold at pH 7.4 and 10-fold compared to 52-fold at pH 9.0 (Tab. 1B).

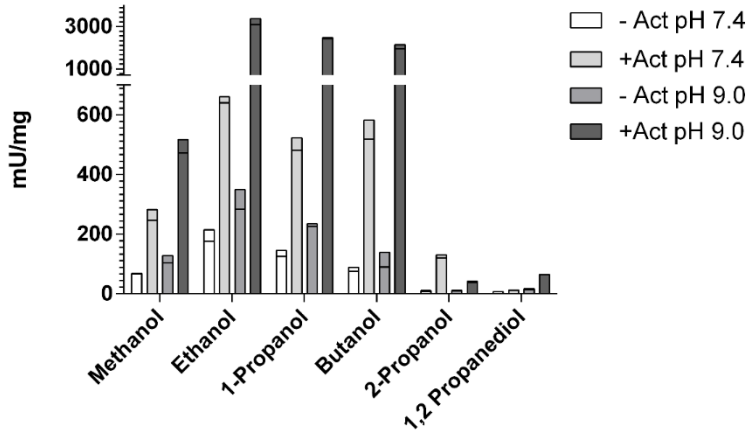


**Figure 1.** Activity of Mdh with and without Act at pH 7.4 and 9.0 with different substrates. One-point measurements were recorded under standard assay conditions (see Materials and methods). Duplicate measurements are shown.

From these findings, it can be concluded that Mdh is optimized for the conversion of methanol and that the effect of Act further improves this selectivity. While Act increases the $V_{max}$ for both methanol and ethanol to a similar extent, the decrease in the $K_m$ is much more pronounced for methanol.

### 2.3.2 Other type III alcohol dehydrogenases are also activated by Act

Because all three Mdh paralogs from *B. methanolicus* are activated by Act [5], we tested whether similar enzymes from other organisms are also affected. Mdh homologs were identified using a BLAST search and four of the top hits that also showed high homology to Mdh2 were chosen for more detailed analysis. The selected enzymes were Adhs from *Desulfotomaculum kuznetsovii* (AdhDK), *Bacillus coagulans* (AdhBC), *Lysinibacillus fusiformis* (AdhLF) and *Lysinibacillus sphaericus* (AdhLS). Of these donor organisms, only *D. kuznetsovii* is known to grow on methanol, but it is unknown if the Adh studied here is the one responsible for methanol oxidation in this bacterium [15]. The four enzymes harbor a sequence identity of 50-60% to Mdh and 60-70% to Mdh2, and all appear to belong to the family of type III alcohol dehydrogenases (Fig. 2A). All of the enzymes were purified and tested for their activities and for activation by Act with methanol and ethanol at pH 7.4 and 9.0 (Fig. 3).

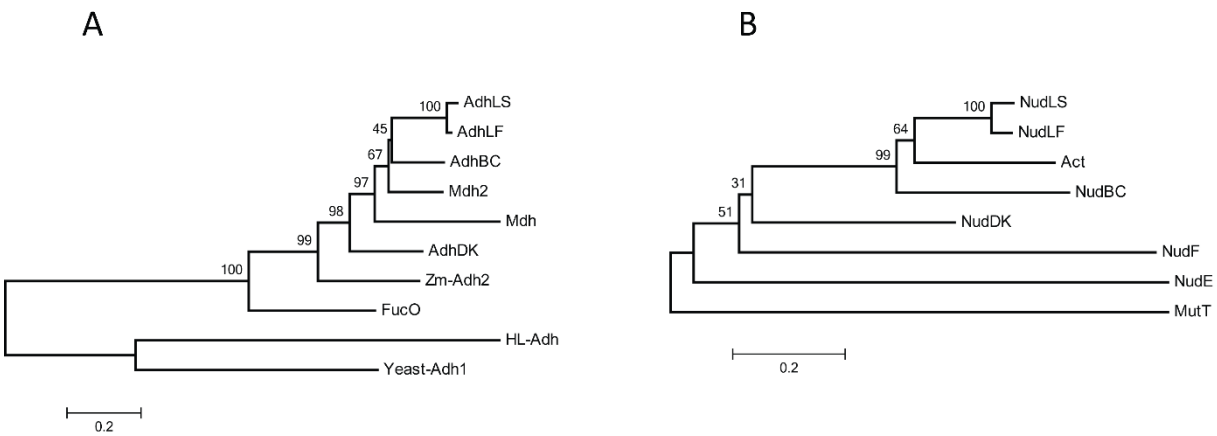


**Figure 2.** Phylogenetic tree of (A) Adhs and (B) Nudix hydrolases used in this study. The protein sequences were aligned (BLOSUM matrix), and trees were created (neighbor-joining; Poisson model) using Mega 5 [16]. Bootstrap values (500 replications) are shown. The alignments are listed in Fig. S1 (Adhs) and Fig. S2 (Nudix hydrolases). Abbreviations: HL-Adh (Horse-liver Adh), Zm-Adh2 (*Z. mobilis* Adh2).

All of the tested Mdh homologs showed activity with both ethanol and methanol and were activated by Act (Fig. 3). Similar to Mdh and Mdh2, all of the enzymes showed a higher activity at pH 9.0 than at pH 7.4 and were more active with ethanol than with methanol. Of all tested enzymes, AdhDK with Act showed the highest activity with methanol as a substrate (Fig. 3), supporting the assumed role of this enzyme in methanol metabolism [15]. AdhBC showed, by far, the highest activity with ethanol as a substrate, suggesting that this Adh is optimized for ethanol.

The enzyme showing the highest activity with methanol, AdhDK, was characterized kinetically with methanol as substrate at pH 7.4 (Tab. 2). Act increased the catalytic efficiency of AdhDK by a factor of 62. The resulting catalytic efficiency is almost identical to that of Mdh, with a higher $K_m$ but also a higher $V_{max}$.

This data shows that activation by Act is not restricted to the Mdhs of *B. methanolicus* but is widespread among type III Adhs. To test whether Act activates enzymes outside of this family, two type I Adhs, horse liver Adh and yeast Adh, were tested. Both showed, in addition to their expected activity with ethanol [17], activity with methanol, but Act was not able to increase the activity of either (data not shown).

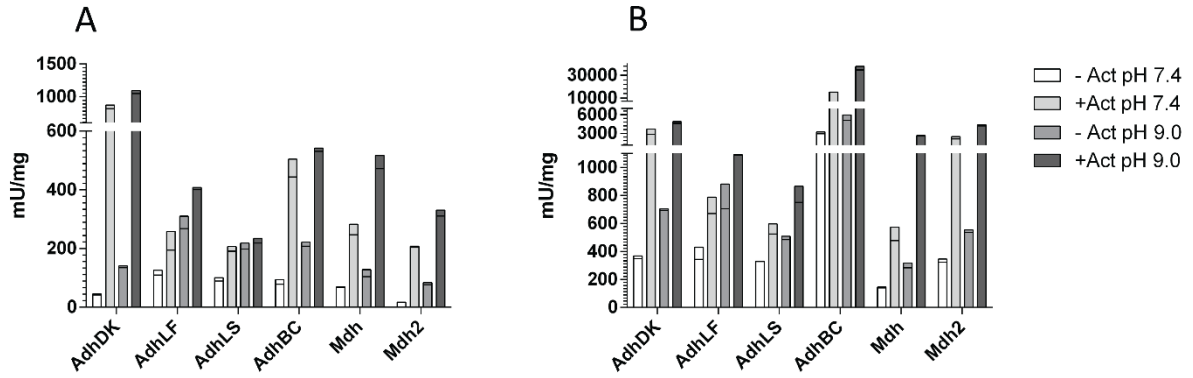


**Figure 3.** Activities of different Adhs (incl. Mdhs) with and without Act at pH 7.4 and 9.0 with (A) methanol or (B) ethanol as a substrate under standard assay conditions. Duplicate measurements are shown.

**Table 2.** Kinetic parameters for AdhDK with and without Act at pH 7.4 with methanol as a substrate. Means ± 95 % confidence intervals are shown.

| | $V_{max}$ [mU/mg] | $K_m^{MeOH}$ [mM] | $k_{cat}/K_m^{MeOH}$ [$M^{-1}s^{-1}$] |
|---|---|---|---|
| **AdhDK** | 76±7 | 446±113 | 0.1 |
| **AdhDK+Act** | 962±33 | 91±10 | 7.2 |

### 2.3.3 Other Nudix hydrolases can activate Adhs to the same extent as Act

Act belongs to the enzyme family of Nudix hydrolases and uses $NAD^+$ (in addition to ADP-ribose) as a substrate [11]. Because it was previously shown that NudF from *Bacillus subtilis* can mimic the activation by Act [5], it seems likely that this feature is more general than assumed. Nudix proteins are found in the vast majority of organisms [9], including the Adh donor organisms, where we identified Act homologs (40-60% sequence identity) using a BLAST search: NudDK (*D. kuznetsovii*), NudBC (*B. coagulans*), NudLF (*L. fusiformis*) and NudLS (*L. sphaericus*) (Fig. 2B). All of these enzymes harbor an amino acid motif that is indicative of the Nudix hydrolase family. They are either annotated as Nudix hydrolases (NudBC and NudDK) or ADP-ribose pyrophosphatases (NudLF and NudLS), but their specificities were not yet determined experimentally. In addition, we included NudF and NudE from *E. coli* (20-30% sequence identity) in our analysis, which are both members of the Nudix hydrolase family [9]. NudE was previously described to exhibit a rather broad substrate spectrum, including ADP-ribose and NADH, which is preferred over $NAD^+$ [18]. NudF has been shown to specifically hydrolyze ADP-sugars, including ADP-ribose [10], and to our knowledge, $NAD^+$ has not yet been examined as a substrate for this enzyme. All of the selected Nudix family proteins were tested

for their ability to activate the different Adhs, i.e., Mdh, Mdh2, AdhDK, AdhBC, AdhLF and AdhLS. Except for NudE, all of the Nudix enzymes were able to activate Mdh and all other Adhs (Fig. 4). Notably, the extent of the observed activation was independent of the activating Nudix hydrolase and was only dependent on the Adh. The ability of a Nudix enzyme to activate Adhs is likely linked to its ability to hydrolyze $NAD^+$, which seems to correlate with its ability to hydrolyze ADP-ribose. It has been previously proposed that the bound cofactor is hydrolyzed to achieve Mdh activation [7]. However, our finding that various Nudix enzymes exhibit the same degree of activation on Mdh/Adh enzymes questions this model, which is based on protein-protein interaction. The reason for the lack of activation by NudE remains unclear and may be linked to the preference of this enzyme for NADH rather than $NAD^+$ [18]. Taken together, our findings show that the *in vitro* activation of Adhs is not limited to Act but is also possible with numerous homolog Nudix hydrolases.

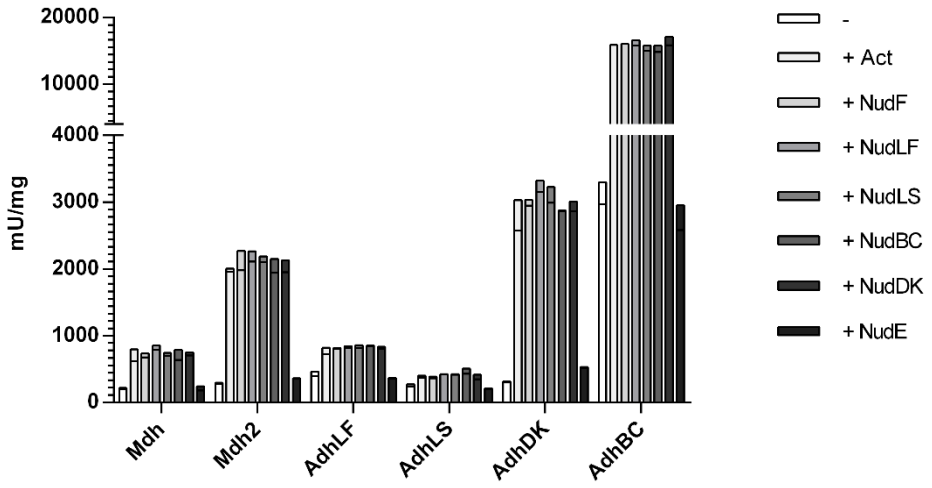


**Figure 4.** Activities of different Adhs in combination with different Nudix hydrolases at pH 7.4 using ethanol as a substrate under standard assay conditions. Duplicate measurements are shown.

### 2.3.4 Site-directed Mdh and Adh mutants indicate common active site residues that affect overall activity, Act dependency and pyridine-nucleotide specificity

It was previously shown that the active site mutation S97G of Mdh of *B. methanolicus* C1 mimics the effect of Act on Mdh activity and leads to complete insensitivity towards activation by Act [12]. To link the proposed activation mechanism [12] to the Adhs described above that were also activated by Act, mutants were generated for Mdh (S98G) and Mdh2 (S101G) of *B. methanolicus* MGA3 as well as AdhDK (S103G). The activities of these mutants with methanol or ethanol were measured at pH 7.4 and pH 9.0 with and without the addition of Act (Fig. 5). Using methanol as a substrate, the Mdh mutant showed increased activity compared to the wild-type enzyme as previously described [12] and was even higher than that of Act-activated Mdh at pH 9.0. Contrary to previous findings, the addition

of Act still had an activating effect on Mdh$^{S98G}$ at both pH values. Interestingly, for Mdh2 and AdhDK, the mutation caused a loss of activity as well as insensitivity towards Act (Fig. 5A).

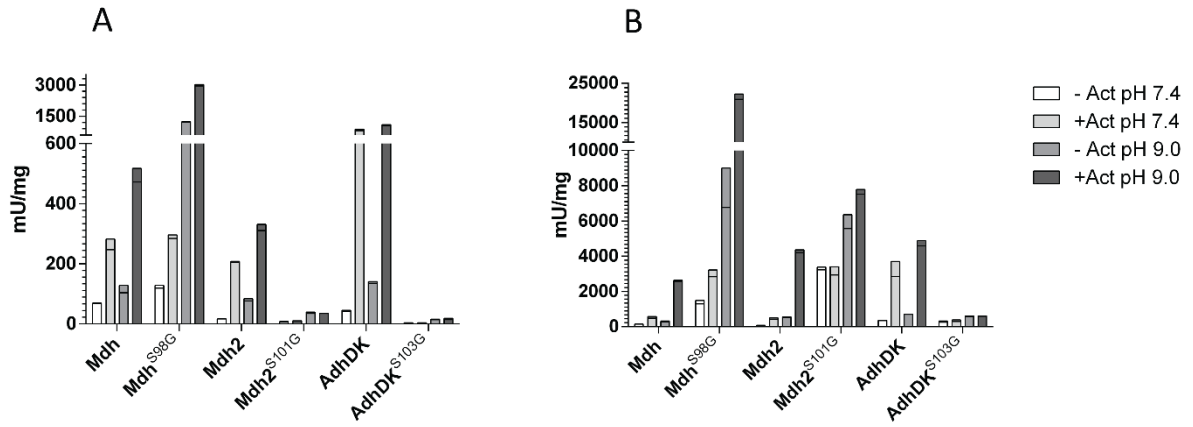


**Figure 5.** Activities of the activation (S97G) mutants of Mdh, Mdh2 and AdhDK with and without Act at pH 7.4 and 9.0 with (A) methanol or (B) ethanol as a substrate. The enzyme activities were determined under standard assay conditions. Duplicate measurements are shown.

For Mdh$^{S98G}$, the findings with ethanol as a substrate were similar to those with methanol; however, a generally increased activity was observed with ethanol compared to methanol. Notably, Mdh2$^{S101G}$ and AdhDK$^{S103G}$ were still insensitive towards activation by Act with ethanol as a substrate; however, the measured activities almost reached or even exceeded the level of the activated wild-type enzyme. To investigate the active site mutation in greater detail, the kinetics of Mdh$^{S98G}$ were measured with methanol at pH 7.4 with and without Act (Tab. 3A).

The observed $V_{max}$ was about twice that of the activated wild-type enzyme, but the $K_m$ was increased even compared to the wild-type enzyme without the addition of Act. The addition of Act to Mdh$^{S98G}$ led to a doubling of the $V_{max}$ but only a slight reduction in the $K_m$. The resulting catalytic efficiency was similar to that of wild-type Mdh in the absence of Act. These results suggest that the active site mutation only partly mimics Act activation. The lack of a tightly bound NAD may lead to a change of mechanism, from a ping-pong mechanism to a faster ternary complex mechanism, and therefore to an increase in the $V_{max}$. However, the positive effect of Act on the $K_m$ for methanol was lacking. This result suggests that the AMP moiety of the cofactor, as it is bound after treatment with Act, is required for efficient substrate binding (decrease in the $K_m$). For Mdh2$^{S101G}$ and AdhDK$^{S103G}$ the apparent loss of activity on methanol could be explained by an even stronger increase in the (already high) $K_m$ towards methanol. The increased activities of the mutants with ethanol might be due to a less severe effect on the $K_m$ and therefore an overall positive effect of the mutation under the tested substrate concentrations.

**Table 3.** Kinetic parameters of the (A) activation and (B) NADP-using mutants of Mdh at pH 7.4 with methanol as a substrate. The Mdh D38G mutant was measured with $NADP^+$. The parameters of the wild-type enzyme are shown in (C) for comparison. Means ± 95 % confidence intervals are shown.

| **(A)** | $V_{max}$ [mU/mg] | $K_m^{MeOH}$ [mM] | $k_{cat}/K_m^{MeOH}$ [$M^{-1}s^{-1}$] |
|---|---|---|---|
| **$Mdh^{S98G}$** | 440±53 | 1151±274 | 0.3 |
| **$Mdh^{S98G}$+Act** | 819±82 | 847±190 | 0.7 |
| **(B)** | | | |
| **$Mdh^{D38G}$ ($NADP^+$)** | 46±2 | 379±53 | 0.1 |
| **$Mdh^{D38G}$+Act ($NADP^+$)** | 128±8 | 137±16 | 0.6 |
| **(C)** | | | |
| **Mdh** | 129±10 | 349±72 | 0.3 |
| **Mdh+Act** | 253±28 | 25±9 | 6.8 |

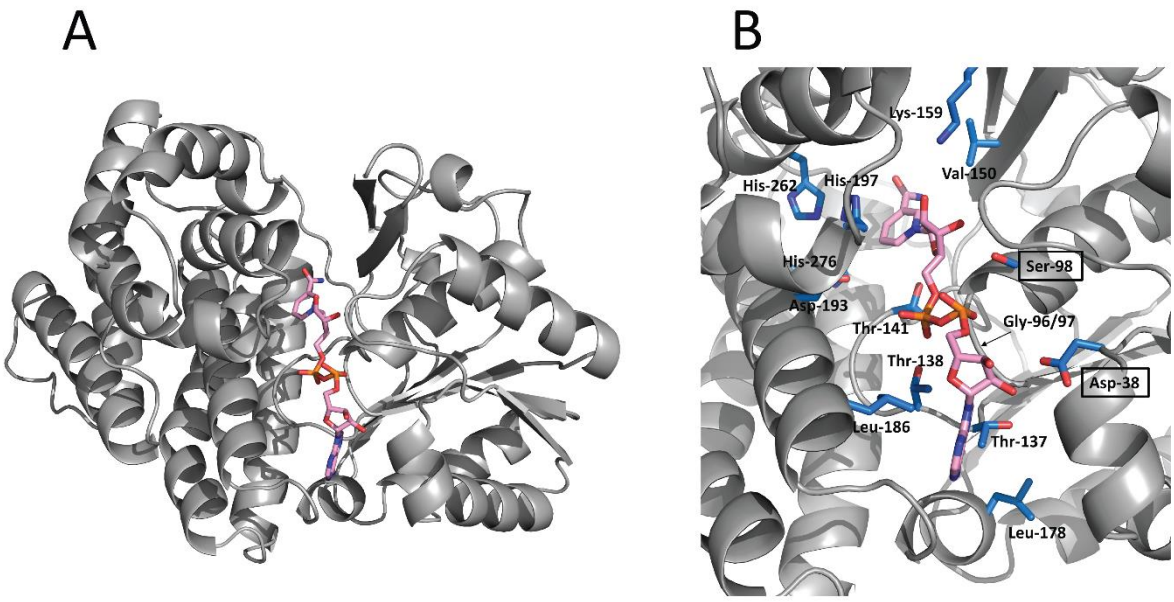


**Figure 6.** Predicted model structure of Mdh. (A) Monomer of Mdh with $NAD^+$. (B) Zoom-in of Mdh active site. The amino acids predicted to be involved in cofactor and metal (His-197, His-262, His-276) binding are shown. The framed residues were mutated in this study.

All of the tested Mdh and Adh enzymes are strictly NAD-dependent and cannot use $NADP^+$ as a cofactor. Thus, we aimed to determine whether cofactor specificity can be altered by site-directed mutagenesis and whether the effect of such mutations is conserved among the different enzymes. To this end, the 3D structures of Mdh, Mdh2 and AdhDK were predicted in silico using published type III alcohol dehydrogenase structures, such as *E. coli* lactaldehyde reductase (FucO) and *Z. mobilis* Adh2

(Fig. 6). By docking the NAD(H) cofactor to the Mdh model, an aspartate (Asp-38) was identified that potentially presents a steric hindrance for the additional phosphate group (analogous to FucO [19]) (Fig. 6B). Consequently, the respective aspartate was mutated to glycine in Mdh (D38G), Mdh2 (D41G) and AdhDK (D43G). The resulting enzyme derivatives were able to use $NADP^+$ and $NAD^+$ with similar activity (Fig. 7). Interestingly, all of the mutants were also activated by Act when using $NADP^+$ as a co-substrate; however, the extent of activation was slightly lower compared to that of the mutant and wild-type enzymes with $NAD^+$.

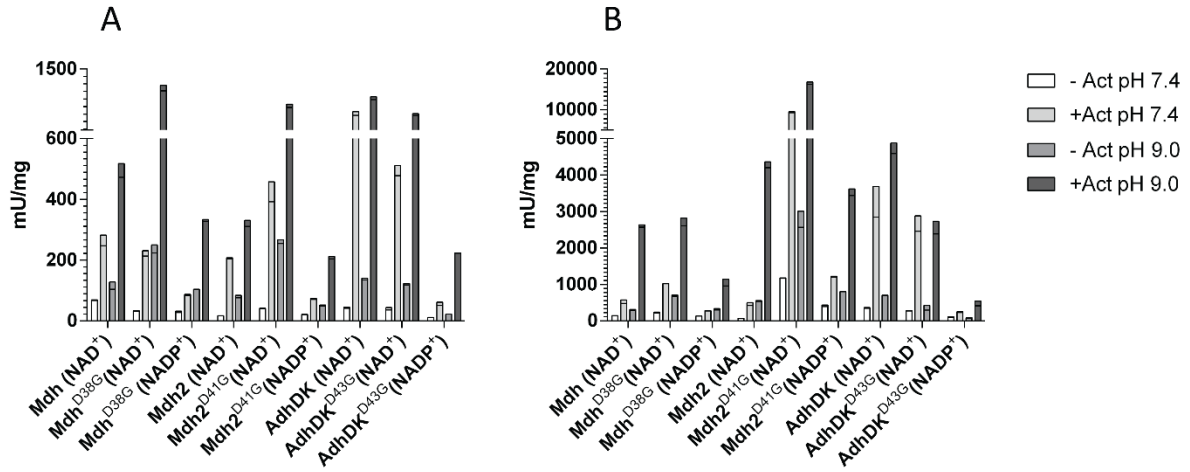


**Figure 7.** Activities of NADP-using mutants of Mdh, Mdh2 and AdhDK with and without Act at pH 7.4 and 9.0 with $NAD^+$ or $NADP^+$ as a co-substrate with (A) methanol or (B) ethanol. The enzyme activities were determined under standard assay conditions. Duplicate measurements are shown. The wild-type enzyme activities with $NADP^+$ were below 0.5% of NAD-dependent activities. (data not shown).

Because Mdh is the main enzyme for methanol conversion in *B. methanolicus*, the kinetics of the NADP-using mutant were investigated in further detail (Tab. 3B). The $V_{max}$ with $NADP^+$ was decreased compared to the wild-type enzyme with $NAD^+$. The effect of Act on the $V_{max}$ was the same as for the wild-type enzyme, but the effect on the $K_m$ was much lower. The successful mutation of all of the studied enzymes reflects the overall conservation of type III alcohol dehydrogenases and shows the relatedness of the Adh/Mdh enzymes analyzed in this study. It is interesting that this mutation does not result in an apparent loss of activity with $NAD^+$ compared to the wild-type enzyme, suggesting that evolutionary pressure conserved the ability of Adh/Mdh enzymes to use $NAD^+$ rather than $NADP^+$ by maintaining an enzyme with steric hindrance for the latter.

### 2.3.5 Attempts to investigate the *in vivo* role of Act by complementation of *Methylobacterium extorquens* mutants

The *in vivo* role of the activation of Mdh by Act is poorly understood and difficult to investigate in the natural host due to lacking genetic tools. As an alternative, we aimed at complementing *Methylobacterium extorquens* PA1 strains lacking their PQQ-dependent methanol dehydrogenases (PQQ-Mdh) MxaFI and XoxF. For the complementation, both *mdh* and *mdh2* (with or without *act*) from *B. methanolicus* strain MGA3 as well as the related strain PB1 [5] were expressed using a constitutive promoter (see Materials and Methods). Expression was successful, albeit at low levels and with poor protein solubility that was even further reduced when *act* was co-expressed (data not shown). NAD-dependent methanol dehydrogenase activities in the very low mU-range could be measured for some constructs, but no consistent beneficial effect of Act could be observed (Fig. S3). Growth above background level was not observed for any of the complemented strains (Fig. S4). Neither genes codon-optimized for *E. coli* nor for *M. extorquens* did result in a significant improve of enzyme activity or in growth on methanol (data not shown). Since no growth on methanol as sole carbon source was possible, the complemented strains were also tested for an improved growth yield under co-consumption conditions. To this end, the strains were grown on a mixture of succinate and methanol, which were shown to be co-consumed [20]. None of the strains showed a significantly higher yield compared to the negative control under the tested conditions (Fig. S5A). To generate a more sensitive readout, labelling experiments with $^{13}$C-methanol and unlabeled succinate were performed with two strains harboring the two combinations that showed the highest *in vitro* activities (*mdh2* + *act* from PB1 and *mdh* from MGA3) (Fig. S5B). Under co-consumption conditions, labelling incorporation is only expected in metabolites containing a one-carbon incorporation step during their biosynthesis. Both investigated strains had higher labelling than the empty control in most investigated metabolites, but the level of the wildtype was not reached. For example, ATP showed 7.9 and 7.0% total labelling compared to 4.5% in the negative control and 18.8% in the wildtype. These results confirm the activity of the NAD-dependent Mdhs *in vivo* and could offer a growth-independent approach for *in vivo* investigation of Mdh and Act.

## 2.4 Conclusion

In this study, we show that the observed activating effect of Act on *B. methanolicus* Mdhs is not unique but possible with numerous interchangeable partners belonging to type III Adhs and Act homologs belonging to the Nudix family. Because most of the tested enzymes do not originate from methylotrophs, the proposed specific role of this activation in methylotrophy is questionable. By testing different substrates, we were able to show that at least for Mdh, the main enzyme involved in

methanol oxidation in *B. methanolicus*, activation by Act specifically improves the catalytic efficiency towards methanol. In general, activation by Act seems to have two independent effects. On one hand, it increases the reaction speed ($V_{max}$) by factors that are dependent on the Adh, most likely by switching from a ping-pong mechanism to a ternary complex reaction mechanism. On the other hand, it improves the substrate affinity ($K_m$) to an extent dependent on the substrate and the nature of the Adh enzyme. All of the Nudix hydrolases that were able to increase Adh activity did so by the same factor as Act, suggesting a common mechanism. In order to obtain a better understanding of the observed phenomenon, mutants that were capable of using $NADP^+$ as a cofactor were successfully created based on *in silico* models. Interestingly, these mutants could be activated by Act independently of the cofactor used. The biological role of this activation is yet to be proven in *B. methanolicus* and other organisms. Nevertheless, the activating effect of Act and its Nudix family homologs on the activity of different Adhs might be useful for *in vitro* industrial applications of type III Adhs. The *in vivo* investigation by heterologous expression in *M. extorquens* PA1 was hampered by low solubility in the host. Nonetheless, low Mdh activity was detectable, providing a starting point for further *in vivo* experiments.

## 2.5 Materials and methods

### 2.5.1 Construction of expression plasmids and site directed mutagenesis

pET21a plasmids containing *B. methanolicus* MGA3 *mdh* (Uniprot no. I3DTM5), *mdh2* (I3E949) and *act* (I3EA59) are described [5] and were received from SINTEF, Norway. A number of additional genes for Adhs and Nudix hydrolases from a variety of organisms were amplified from genomic DNA or previously constructed vectors and cloned into pET21a expression vector (Thermo Fisher Scientific, Reinach, Switzerland). Primers were designed using CloneManager 9 (Scientific & Educational Software) and synthesized at Microsynth (Balgach, Switzerland). The reverse primer was used to mutate the stop-codon in order to attach the C-terminal $His_6$-tag encoded on the pET21a plasmid. *adhBC* (*B. coagulans* 36D1, G2TN45), *adhLF* (*L. fusiformis* ZC1, D7WPP7), and *adhLS* (*L. sphaericus* C3-41, B1HX72) were first amplified from previously constructed plasmids based on the pSEVA424 backbone [21] and cloned into pET21a. *AdhDK* (*D. kuznetsovii* 17, F6CJW2), *nudBC* (*B. coagulans*, G2TKZ3), *nudLF* (*L. fusiformis*, D7WXY3), *nudLS* (*L. sphaericus*, B1HRL7), and *nudDK* (*D. kuznetsovii*, F6CJI1) were amplified from genomic DNA. For expression of *E. coli* K-12 *nudF* (Q93K97) and *nudE* (P45799), expression strains containing the according plasmids were isolated from the ASKA collection [22] and directly used for overexpression. Site directed mutagenesis of $Mdh^{S98G}$, $Mdh2^{S101G}$, $Mdh^{D38G}$, $Mdh2^{D41G}$ and $AdhDK^{D43G}$ was done by mutation of the pET21a constructs using the Stratagene QuikChange mutagenesis protocol (Agilent Technologies). Primers

were designed using PrimerX (http://www.bioinformatics.org/primerx/). AdhDK$^{S103G}$ was produced using overlap PCR. All primers are listed in Tab. S1.

### 2.5.2 Protein expression and purification

*E. coli* BL21 (DE3) cells containing pET21a were grown at 37°C in 400 mL LB medium containing 100 µg/mL ampicillin, or for ASKA expression strains 25 µg/mL chloramphenicol. Cultures were induced at $OD_{600}$ = 0.4-0.8 using 0.1 mM of isopropyl-β-thiogalactoside (IPTG). Cells were harvested after 5-6 hours of expression and pellets were resuspended in 1 mL 100 mM potassium phosphate buffer pH 7.4 (phosphate buffer) supplemented with protease inhibitor (EDTA-free) (Roche, Rotkreuz, Switzerland). Cell lysis was done by passing samples through a French pressure cell press (SLM instruments, Thermo Fisher Scientific) three times at approximately 1000 Psi. Cell debris and membrane fractions were removed by ultracentrifugation (140'000 g for 60 min at 4°C). The supernatant was incubated with 0.5 mL pre-equilibrated Protino Ni-NTA Agarose (Macherey-Nagel, Oensingen, Switzerland) for 1 hour. Washing was done using 10 mL buffer A (phosphate buffer, 20 mM Imidazole) and elution was done with 4 mL buffer B (phosphate buffer, 300 mM imidazole, 2 mM β-mercaptoethanol). The elution fraction containing the protein of interest was concentrated using Amicon Ultra-4 (10 kDa) (Millipore, Zug, Switzerland) washed once with phosphate buffer and then filled up to 1.5 mL using the same buffer. Protein concentration was measured using the Pierce BCA protein assay kit (Thermo Fisher Scientific) with bovine serum albumin (BSA) (Sigma-Aldrich, Buchs, Switzerland) as a standard. Purity of the obtained proteins was confirmed to be greater than 90 % for all enzymes using SDS-PAGE (data not shown).

### 2.5.3 Enzyme assay and determination of kinetic parameters

The *in vitro* Adh assay was performed in 1 mL pre-heated 100 mM MOPS-HCl buffer pH 7.4 or pH 9.0 (adjusted at 50°C) and 5 mM $MgSO_4$. The standard assay contained 500 µM $NAD^+$ (Sigma-Aldrich) or $NADP^+$ (Sigma-Aldrich), 0.5-150 µg Adh (purified as describe above or commercial for horse-liver Adh (55689, Sigma-Aldrich) and yeast Adh (A3263, Sigma-Aldrich)), 15-50 µg Act or another Nudix hydrolase (if indicated) and was started using 500 mM of alcohol substrate (Methanol, Ethanol, 1-Propanol, 2-Propanol, 1-Butanol or Propane-1,3-diol). Activities were calculated from maximal slopes of NAD(P)H production at 340 nm measured by a Cary 50 Bio UV-Visible spectrophotometer (Varian, Steinhausen, Switzerland) heated to 50°C. One unit (U) of activity was defined as the amount of enzyme that is required to process 1 µmol of substrate per minute. The Michaelis-Menten parameters were determined by measuring activities at varying alcohol (methanol or ethanol) concentrations (0-2000 mM) keeping the cofactor at saturation (500 µM). In total 15-20

data points were measured for each kinetic determination and were fitted using Prism 5 (GraphPad Software). Catalytic efficiency $k_{cat}/K_m$ of enzymes was calculated under the assumption that each subunit contains one active site. Activation factors were defined as the activity measured in the presence of Act divided by the activity determined in the absence of Act.

### 2.5.4 Structure predictions

Structure predictions were done using protein model portal (http://www.proteinmodelportal.org/) [23]. For the figures the structure prediction of the M4T server (http://manaslu.aecom.yu.edu/M4T/) was used [24]. For Mdh the used template structures were lactaldehyde reductase (FucO) from *E. coli* (1RRM) and Adh2 from *Zymomonas mobilis* (3OX4). Figures were prepared using the PyMOL Molecular Graphics System (Version 1.4.1, Schrodinger LLC).

### 2.5.5 Sequence alignments and phylogenetic trees

Alignments and phylogenetic trees were generated using MEGA5. ClustalW alignments were done using the BLOSUM protein weight matrix. Phylogenetic trees were calculated by neighbor-joining using the Poisson model. The Bootstrap method (500 replications) was used as a phylogenetic test. Sequence alignments were visualized using BioEdit [25].

### 2.5.6 Heterologous expression in *M. extorquens* PA1

*M. extorquens* PA1 was grown at 28°C in minimal medium with the following composition: 20.7 mM phosphate buffer, 30.29 mM $NH_4Cl$, 0.81 mM $MgSO_4$, trace elements (40.3 µM $FeSO_4$, 15.65 µM $ZnSO_4$, 12.61 µM $CoCl_2$, 5.09 µM $MnCl_2$, 16.17 µM $H_3BO_3$, 1.65 µM $Na_2MoO_4$, 1.2 µM $CuSO_4$ and 20.41 µM $CaCl_2$), and methanol (123 mM, 'MMM') or succinate (30.8 mM, 'MMS'). For solid minimal medium 15 g/L agar was added.

Gene deletions of *mxaF* (Mext_4150) and *xoxF* (Mext_1809) were generated using pK18_*mobsacB* [26]. Plasmids for overexpression of NAD-dependent Mdhs were based on pTE105 which harbors the constitutive P*tuf* promoter [27]. During amplification, the ribosomal binding site (RBS) 'AAGGAG' was added. If *act* was part of the construct, a second RBS was included with it.

Growth experiments were performed in 96-well plate format using a Tecan Infinity M200 Pro spectrophotometer (Tecan). Pre-cultures were grown overnight in 20 mL MMS plus methanol (123 mM), washed two times minimal medium without carbon source and adjusted to an $OD_{600}$ of 0.5. 20 µL was used to inoculate 180 µL medium (1.1x carbon source to compensate dilution by inoculation) in 96-well plates (ThermoFischer Scientific, Nuclon 96 Flat Bottom Transparent Polystyrol) to a start

$OD_{600}$ of 0.05. Plate was incubated at 28°C while shaking (linear) with an amplitude of 1 mm for 522 s and $OD_{600}$ was measured every 10 min using a bandwidth of 9 nm and 25 flashes.

For the $^{13}C$ labelling experiment, cells were inoculated in minimal medium (phosphate concentration reduced to 10.4 mM) containing 30.8 mM naturally labelled succinate and 123 mM $^{13}C$-methanol to an $OD_{600}$ of 0.03 and incubated at 28°C until an $OD_{600}$ of 1 was reached. One OD-unit (1mL OD1) was sampled by filtration through a 0.2 µm regenerated cellulose filter (pre-washed with 50°C MilliQ water), followed by washing with MilliQ water at 20°C and transfer of filter to Schott bottles containing 8 mL boiling MilliQ water. After several brief sonication steps, the bottles were cooled down on ice and the extract was transferred into Falcon tubes for subsequent lyophilization. The freeze-dried samples were resuspended to 1 µg/µL (assuming an OD to CDW correlation: 0.27 mg/mL OD1), spun down for 10 min at 20'000 g and 4°C, and 20 µL were mixed with 80 µL ammonium acetate buffer pH 8.0 containing 1.6 mM tributylamine. Samples were measures using nanoscale ion-pair reversed phase HPLC-MS analysis was conducted as described in [28]. Data analysis and calculation of natural labelling was performed using eMZed [29].

## 2.6 References


1. Anthony, C., and Williams, P. (2003). The structure and mechanism of methanol dehydrogenase. Biochim Biophys Acta *1647*, 18–23.
2. Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. Env Microbiol *13*, 2603–2622.
3. Heggeset, T.M., Krog, A., Balzer, S., Wentzel, A., Ellingsen, T.E., and Brautaset, T. (2012). Genome sequence of thermotolerant *Bacillus methanolicus*: features and regulation related to methylotrophy and production of L-lysine and L-glutamate from methanol. Appl Env Microbiol *78*, 5170–5181.
4. Müller, J.E.N., Litsanov, B., Bortfeld-Miller, M., Trachsel, C., Grossmann, J., Brautaset, T., and Vorholt, J.A. (2014). Proteomic analysis of the thermophilic methylotroph *Bacillus methanolicus* MGA3. Proteomics *14*, 725–737.
5. Krog, A., Heggeset, T.M., Müller, J.E.N., Kupper, C.E., Schneider, O., Vorholt, J.A., Ellingsen, T.E., and Brautaset, T. (2013). Methylotrophic *Bacillus methanolicus* encodes two chromosomal and one plasmid born $NAD^+$ dependent methanol dehydrogenase paralogs with different catalytic and biochemical properties. PLoS One *8*, e59188.
6. Vonck, J., Arfman, N., De Vries, G.E., Van Beeumen, J., Van Bruggen, E.F., and Dijkhuizen, L. (1991). Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus* sp. C1. J Biol Chem *266*, 3949–3954.
7. Arfman, N., Hektor, H.J., Bystrykh, L. V, Govorukhina, N.I., Dijkhuizen, L., and Frank, J.

(1997). Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from *Bacillus methanolicus*. Eur J Biochem *244*, 426–433.

8. Arfman, N., Van Beeumen, J., De Vries, G.E., Harder, W., and Dijkhuizen, L. (1991). Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant *Bacillus* spp. J Biol Chem *266*, 3955–3960.
9. McLennan, A.G. (2006). The Nudix hydrolase superfamily. Cell Mol Life Sci *63*, 123–143.
10. Mildvan, A.S., Xia, Z., Azurmendi, H.F., Saraswat, V., Legler, P.M., Massiah, M.A., Gabelli, S.B., Bianchet, M.A., Kang, L.W., and Amzel, L.M. (2005). Structures and mechanisms of Nudix hydrolases. Arch Biochem Biophys *433*, 129–143.
11. Kloosterman, H., Vrijbloed, J.W., and Dijkhuizen, L. (2002). Molecular, biochemical, and functional characterization of a Nudix hydrolase protein that stimulates the activity of a nicotinoprotein alcohol dehydrogenase. J Biol Chem *277*, 34785–34792.
12. Hektor, H.J., Kloosterman, H., and Dijkhuizen, L. (2002). Identification of a magnesium-dependent NAD(P)(H)-binding domain in the nicotinoprotein methanol dehydrogenase from Bacillus methanolicus. J Biol Chem *277*, 46966–46973.
13. Markert, B., Stolzenberger, J., Brautaset, T., and Wendisch, V.F. (2014). Characterization of two transketolases encoded on the chromosome and the plasmid pBM19 of the facultative ribulose monophosphate cycle methylotroph *Bacillus methanolicus*. BMC Microbiol *14*, 7.
14. Stolzenberger, J., Lindner, S.N., Persicke, M., Brautaset, T., and Wendisch, V.F. (2013). Characterization of fructose 1,6-bisphosphatase and sedoheptulose 1,7-bisphosphatase from the facultative ribulose monophosphate cycle methylotroph *Bacillus methanolicus*. J Bacteriol *195*, 5112–5122.
15. Visser, M., Worm, P., Muyzer, G., Pereira, I.A., Schaap, P.J., Plugge, C.M., Kuever, J., Parshina, S.N., Nazina, T.N., Ivanova, A.E., *et al.* (2013). Genome analysis of *Desulfotomaculum kuznetsovii* strain 17T reveals a physiological similarity with *Pelotomaculum thermopropionicum* strain SIT. Stand Genomic Sci *8*, 69–87.
16. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol *28*, 2731–2739.
17. Reid, M.F., and Fewson, C.A. (1994). Molecular characterization of microbial alcohol dehydrogenases. Crit Rev Microbiol *20*, 13–56.
18. O'Handley, S.F., Frick, D.N., Dunn, C.A., and Bessman, M.J. (1998). Orf186 represents a new member of the Nudix hydrolases, active on adenosine(5')triphospho(5')adenosine, ADP-ribose, and NADH. J Biol Chem *273*, 3192–3197.
19. Montella, C., Bellsolell, L., Perez-Luque, R., Badia, J., Baldoma, L., Coll, M., and Aguilar, J. (2005). Crystal structure of an iron-dependent group III dehydrogenase that interconverts L-

lactaldehyde and L-1,2-propanediol in *Escherichia coli*. J Bacteriol *187*, 4957–4966.

20. Peyraud, R., Kiefer, P., Christen, P., Portais, J.C., and Vorholt, J.A. (2012). Co-consumption of methanol and succinate by *Methylobacterium extorquens* AM1. PLoS One *7*, e48271.
21. Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las Heras, A., Paez-Espino, A.D., Durante-Rodriguez, G., Kim, J., Nikel, P.I., *et al.* (2013). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res *41*, D666-75.
22. 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.
23. Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., and Schwede, T. (2013). The Protein Model Portal--a comprehensive resource for protein structure and model information. Database *2013*
24. Fernandez-Fuentes, N., Madrid-Aliste, C.J., Rai, B.K., Fajardo, J.E., and Fiser, A. (2007). M4T: a comparative protein structure modeling server. Nucleic Acids Res *35*, W363-8.
25. Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser *41*, 95–98.
26. Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and 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.
27. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J.A., and Erb, T.J. (2014). 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 *17*, 430-443
28. 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
29. Kiefer, P., Schmitt, U., and Vorholt, J.A. (2013). eMZed: An open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. Bioinformatics *29*, 963–964.

## 2.7 Supplementary material

**Table S1:** List of all the primers used in this study. Restriction site are underlinde (restriction enzymes are shown in brackets) and mutations of stop-codon are shown in small letters.

| | |
|---|---|
| AdhBC | Fwd AAATCTAGAAAGGAGAGGGGAAAGATG (XbaI)<br>Rev AAACTCGAGCTTTGACATCGCTGCTTTTATAATC (XhoI) |
| AdhLF | Fwd AAACATATGTCAGACGTTCTAAAGC (NdeI)<br>Rev TTTAAGCTTTgAAGAGAGGGCAACTG (HindIII) |
| AdhLS | Fwd AAACATATGTCAGACGTTCTAAAGC (NdeI)<br>Rev TTTTAAGCTTTgATGAAAGAGCAGCAACTG (HindIII) |
| AdhDK | Fwd AAACATATGGCATTGGGTGAACAGGTTTACG (NdeI)<br>Rev AAAAAGCTTaTACATTGCCTGCCGGTAGATGTC (HindIII) |
| NudBC | Fwd AGAACATATGGCCGGGTAACATGCATAAATATAAATACG (NdeI)<br>Rev AGAAGCGGCCGCTCcTTTTCGCGCCTCTTCAAGCATC (NotI) |
| NudLF | Fwd AGAACATATGCAAGCGAGGTTATCGTCAATGAAAAAG (NdeI)<br>Rev AGAAGCGGCCGCTcATGCTAAAAGATTATTTTCTAAATTTGCA GCTAGCCATAG (NotI) |
| NudLS | Fwd AGAACATATGGCGAGGTTATCGTCTATGAAAAAGTTTG (NdeI)<br>RevAGAAGCGGCCGCCTcAAGATTATTTTCTAAATAGGTAGCAAGCCA AAG (NotI) |
| NudDK | Fwd AGAACATATGATGTCCGCATTAATCGAAAAGAAGCTG (NdeI)<br>Rev AGAAGCGGCCGCTTTTgAAACCTTTTCCCGGGCCAAAAG (NotI) |
| Mdh D38G | Fwd CGCTTATCGTTACAGgcGCATTCCTTCACAG<br>Rev CTGTGAAGGAATgcGCCTGTAACGATAAGCG |
| Mdh S98G | Fwd CGGTGGAGGTgGCTCTCACG<br>Rev CGTGAGAGCcACCTCCACCG |
| Mdh2 D41G | Fwd CTTTATTAGTTACAGgcGCTGGTCTTCACGG<br>Rev CCGTGAAGACCAGCgcCTGTAACTAATAAAG |
| Mdh2 S101G | Fwd GGCGGCGGAgGTTCACATGA<br>Rev TCATGTGAACcTCCGCCGCC |
| AdhDK D43G | Fwd CTCATTGTTACGGgCGCCTACCTGGC<br>Rev GCCAGGTAGGCGcCCGTAACAATGAG |
| AdhDK S103G | Fwd1 AGATATACATATGGCATTGGGTGAACAGGTTTACG (NdeI)<br>Rev1 AATCGTGGGAGCcGCCGCCGCCCAGG<br>Fwd2 CTGGGCGGCgGCGGCTCCCACGATTG<br>Rev2 TTGTTAGCAGCCGGATCTCAGTGGTG (NotI on amplified backbone) |

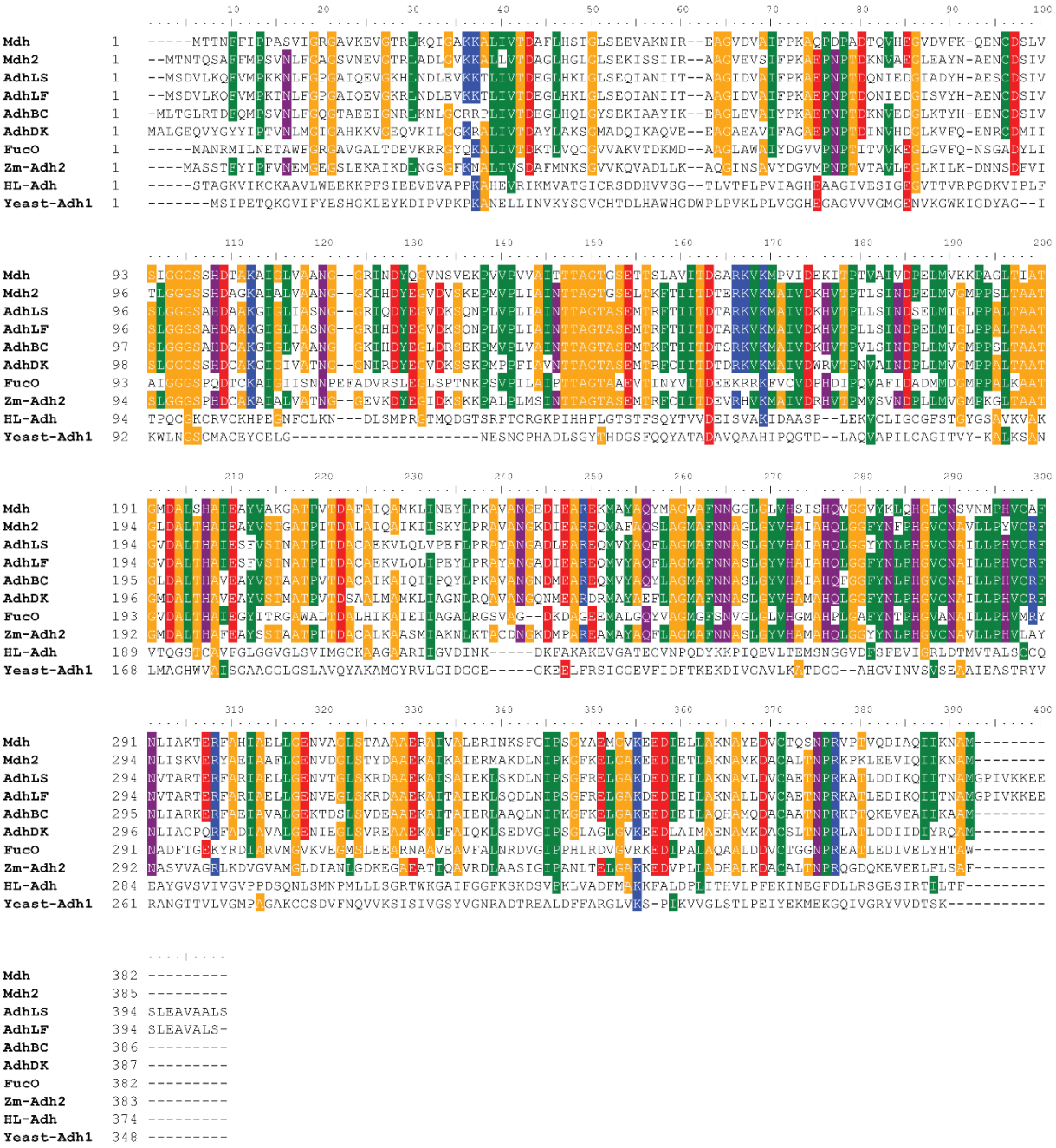


**Figure S1:** Protein sequence alignments of alcohol dehydrogenases performed with BioEdit. Amino acids with at least 60% identity are colored (red for negatively charged, blue for positively charged, orange for small nonpolar, green for hydrophobic and magenta for polar). Abbreviations: HL-Adh (Horse-liver Adh), Zm-Adh2 (*Z. mobilis* Adh2).

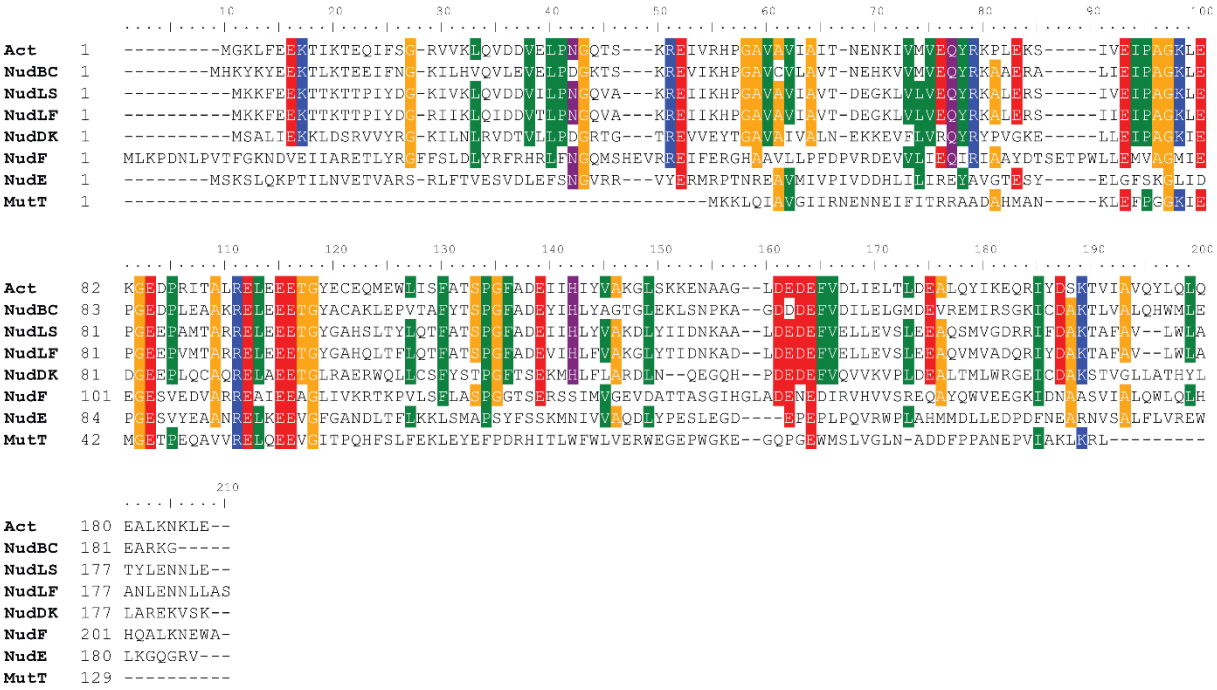


**Figure S2:** Protein sequence alignments of Nudix hydrolases performed with BioEdit. Amino acids with at least 60% identity are colored (red for negatively charged, blue for positively charged, orange for small nonpolar, green for hydrophobic and magenta for polar).

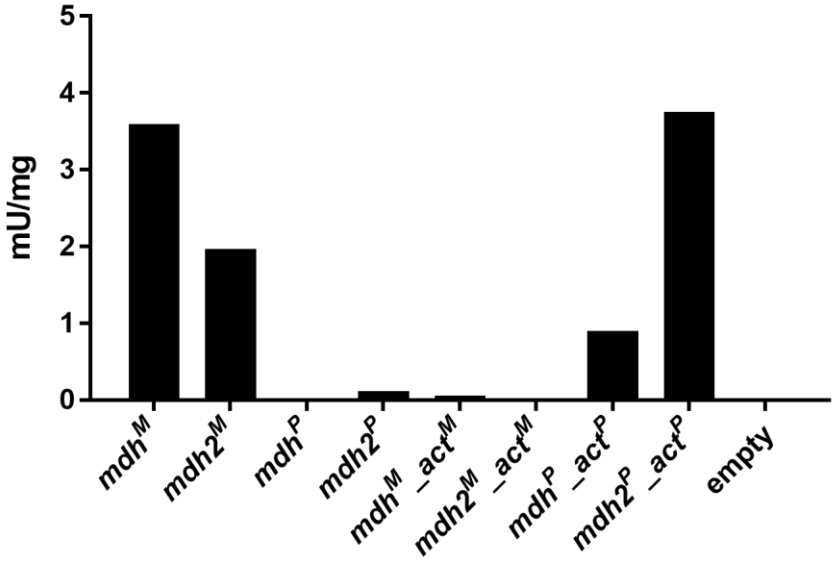


**Figure S3:** *In vitro* NAD-dependent methanol dehydrogenase activities in *M. extorquens* cell lysates at 28°C. All strains were wildtype strains harboring pTE105 with different insert(s) and were grown in minimal medium containing succinate (n = 1). M = strain MGA3, P = Strain PB1.

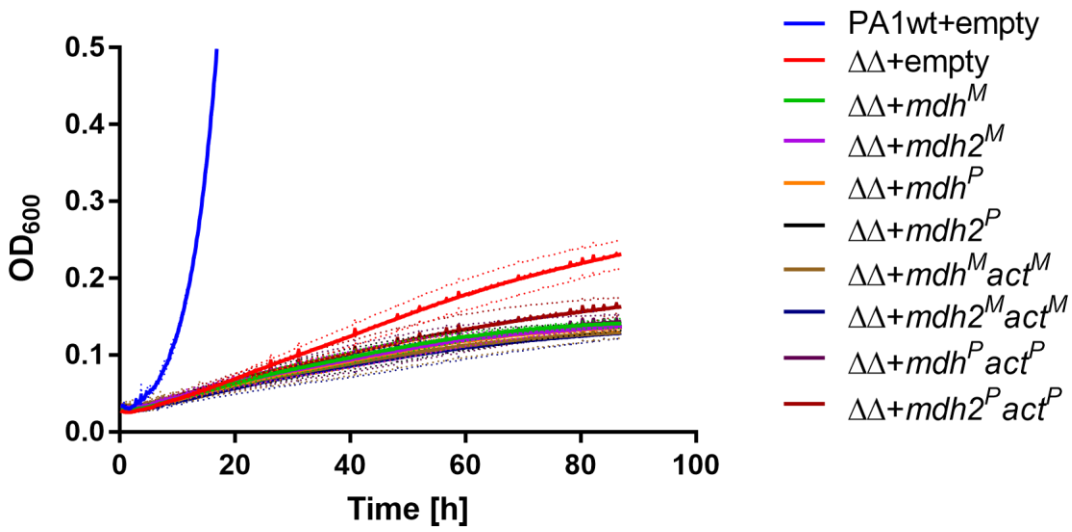


**Figure S4:** Growth of *M. extorquens* wildtype ('PA1wt') and Δ*mxaF*Δ*xoxF* ('ΔΔ') complemented with pTE105 plasmids without insert ('empty') or harboring different NAD-dependent Mdhs in combination with Act (M = strain MGA3, P = strain PB1) on methanol.

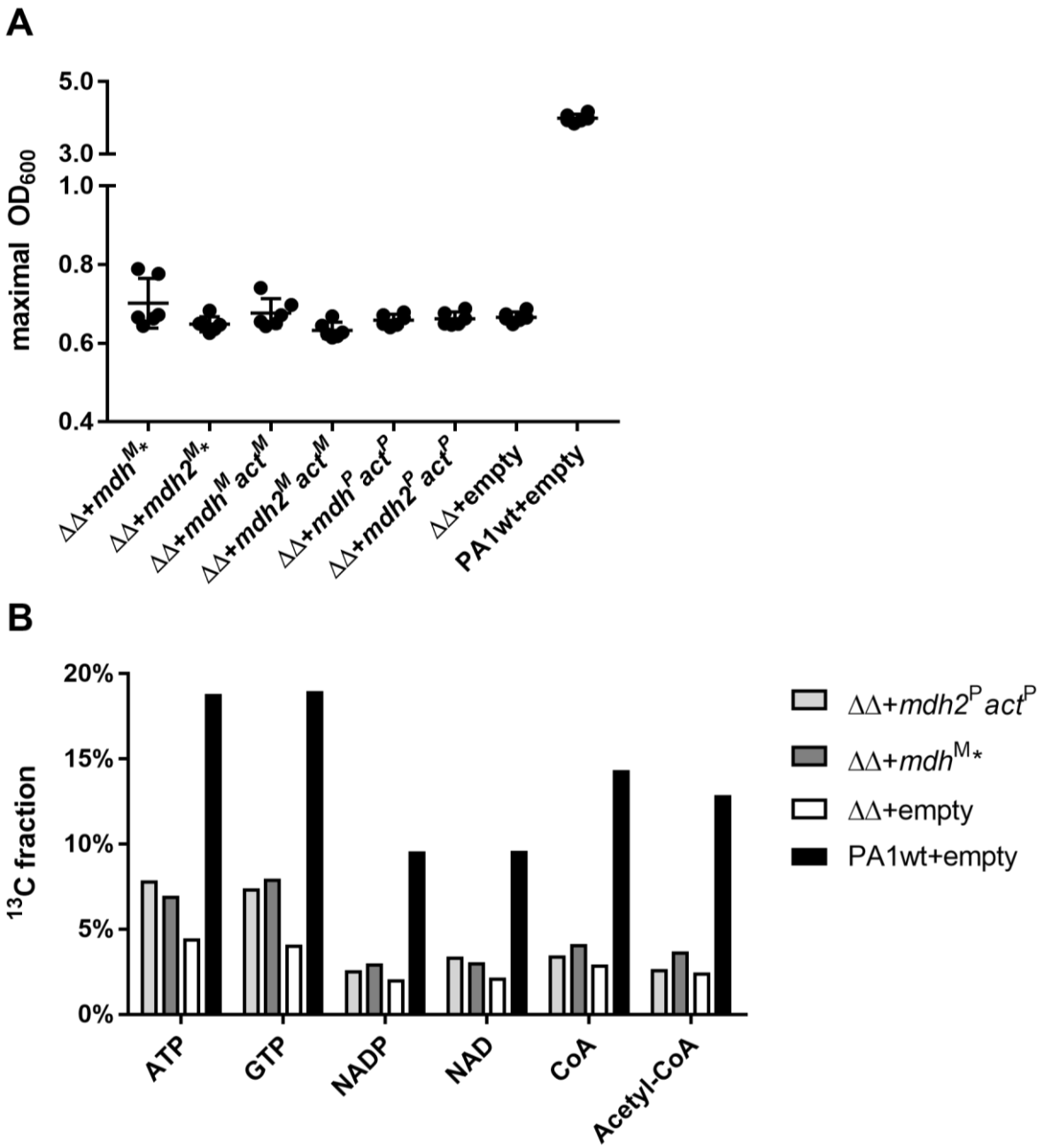


**Figure S5:** Investigation of engineered strains under co-consumption conditions. (A) Final optical densities of *M. extorquens* wildtype ('PA1wt') and Δ*mxaF*Δ*xoxF* ('ΔΔ') complemented with pTE105 plasmids without insert ('empty') or harboring different NAD-dependent Mdhs in combination with Act (M = strain MGA2, P = strain PB1, * = codon-optimized for *E. coli*.) grown in minimal medium containing limiting amounts of succinate (5 mM) and an excess of methanol (123 mM) (n = 5). (B) Labelling incorporation from $^{13}$C-methanol into a selection of metabolites that incorporate C1-units (C1-$H_4F$ or $CO_2$) during their biosynthesis. *M. extorquens* wildtype ('PA1wt') or Δ*mxaF*Δ*xoxF* ('ΔΔ') (harboring pTE105 plasmids with or without insert(s)) was grown to $OD_{600}$ 1 in minimal medium containing naturally labelled succinate and $^{13}$C-labelled methanol (n = 1).

## Chapter 3:

# Novel insights into the regulation of methanol dehydrogenases in *Methylobacterium extorquens* PA1

Andrea M. Ochsner, Ralph Nüssli, Julia A. Vorholt

**Author contributions**

A.M.O. and J.A.V. designed the research; A.M.O. and R.N. performed research; A.M.O. analyzed data; A.M.O. and J.A.V. wrote the manuscript.

# Chapter 3: Novel insights into the regulation of methanol dehydrogenases in *M. extorquens*

Authors: Andrea M. Ochsner, Ralph Nüssli, Julia A. Vorholt

Institute of Microbiology, ETH Zurich, Switzerland

## 3.1 Summary

The model methylotroph *Methylobacterium extorquens* employs a pyrroloquinoline quinone (PQQ)-dependent methanol dehydrogenase (Mdh) to catalyze the initial step of methanol conversion. It was recently found that the organism possesses a second rare-earth element (REE)-dependent PQQ-Mdh system in addition to the $Ca^{2+}$-dependent MxaFI discovered 50 years ago. Interestingly, the presence of REEs does not only induce the expression of the genes for REE-dependent XoxF, but also represses the expression of *mxaFI*, suggesting that XoxF is the preferably used Mdh system. The regulatory cascade underlying the REE-dependent switch is not yet completely understood. We investigated the regulation of genes encoding PQQ-Mdhs in *M. extorquens* PA1 in response to the REE lanthanum ($La^{3+}$). Using reporter assays and a suppressor screen, we were able to show that $La^{3+}$ is sensed in the absence of XoxF. This is in contrast to the current working model that suggested XoxF as the $La^{3+}$-sensor and adds an additional layer of complexity to the regulatory network.

## 3.2 Introduction

Methylotrophy is the ability of microorganisms to grow on reduced C1-compounds such as methanol as their sole source of carbon and energy. Methylotrophs are of interest due to their important role in the global carbon cycle [1] and as potential hosts for sugar-independent biotechnological production [2]. The Alphaproteobacterium *Metyhlobacterium extorquens* is one of the best studied methylotrophs and its suitability for production of value-added chemicals has been shown [3]. The first step of methylotrophy is the oxidation of methanol to formaldehyde. In *M. extorquens*, this reaction is catalyzed by a methanol dehydrogenase that depends on the cofactor pyrroloquinoline quinone (PQQ) [4]. The initially discovered enzyme is a heterotetramer encoded by *mxaFI* that carries a $Ca^{2+}$ ion at its active site [5–7]. Apart from these genes, multiple additional genes are required for a functional methanol dehydrogenase, such as genes involved in the insertion of the $Ca^{2+}$ ion [8–10]. Genome sequencing revealed a close homolog (~50% identity) of *mxaF*, termed *xoxF*, but its function remained elusive for a long time [11,12]. Metagenomics and metaproteomics suggested that XoxF

plays an equally or even more important role in the environment than MxaFI [13,14]. Further research showed that the XoxF enzyme is a rare-earth element (REE)-dependent PQQ-Mdh that has a REE instead of a $Ca^{2+}$ ion in its active site [15,16]. The simpler XoxF enzyme lacks a small subunit and is encoded in a cluster containing a total of three genes that are thought to be sufficient for functional production of XoxF [17]. The discovery of XoxF as $La^{3+}$-dependent enzyme marks the first discovery of an enzyme that specifically depends on REEs as a cofactor [18]. In addition to the presence of paralogues in model methylotrophs, some bacteria exclusively harbor *xoxF* and may have been missed in past enrichments for methylotrophs in the absence of REEs [16]. Shortly after the discovery of the REE-dependent Mdh XoxF, REE-dependent multicarbon alcohol dehydrogenases were discovered in *M*. *extorquens* AM1 [19] and subsequently also in the non-methylotroph *Pseudomonas putida* [20].

The presence of XoxF and MxaFI in the same organism provokes the question on the regulation of the two enzymes. Interestingly, XoxF does not just represent an auxiliary second methanol dehydrogenase system, but replaces MxaFI in the presence of suitable REEs [21]. The regulatory cascade underlying this switch of Mdh system in response to REE is poorly understood. In the methanotroph *Methylomicrobium buryatense* the two sensor kinases MxaY and MxaB were found to be involved in the $La^{3+}$-dependent switch of methanol dehydrogenase expression [22,23]. *M*. *extorquens* does not possess MxaY, but instead seems to rely on other regulators. The two-component systems MxbDM and MxcQE are implicated in the regulation the expression of the *mxa* cluster and mutational analysis suggests that MxcQE is upstream of MxbDM in the cascade [24]. Furthermore, XoxF was, in addition to its role as REE-dependent Mdh, shown to be involved in this regulatory cascade in the absence of REEs [25]. In a recent study in *M*. *extorquens* AM1, XoxF was also suggested to be involved in $La^{3+}$-sensing [21]. According to the proposed model, the inactive apo-form of XoxF (without $La^{3+}$) binds to the sensor kinase MxcQ, which then activates the expression of the two-component system MxbDM via its cognate response regulator MxcE. After sensing a second signal, MxbDM activates P*mxa* and at the same time represses P*xox*. When XoxF is converted to its active form by $La^{3+}$-binding or the *xoxF* gene is knocked out, the activating signal disappears and the expression is switched from *mxa* to *xox* expression. However, experimental data confirming the role of XoxF as REE-sensor is still lacking.

In this study we investigated the lanthanide-dependent regulation of methanol oxidation systems in *M*. *extorquens* PA1, a novel model strain isolated from the *Arabidopsis* phyllosphere [26,27] using reporter assays, a suppressor screen and enzyme activity assays. In line with results reported for strain AM1, the addition of Lanthanum ($La^{3+}$) switches the expression from *mxa* to *xox* in a dose-dependent manner. Furthermore, our data suggest that the cascade regulating the switch of Mdh expression is more complex than previously hypothesized and involves at least one additional $La^{3+}$-sensor.

## 3.3 Results

### 3.3.1 XoxF is required for $La^{3+}$-dependent growth on methanol and is essential for the expression of the *mxa* cluster in *M. extorquens* PA1

To confirm the presence of a REE-dependent Mdh in *M. extorquens* PA1, we investigated the growth of a strain lacking the gene encoding the large subunit of the classical methanol dehydrogenase MxaFI (Δ*mxaF*). The strain was not able to grow on methanol unless the rare-earth element Lanthanum ($La^{3+}$) was present (Figure 1, Table S1), as it was reported in strain AM1 [15]. Unlike strain AM1, strain PA1 only encodes one XoxF-type REE-dependent PQQ-Mdh, which thus was essential for growth on methanol in the presence of $La^{3+}$ in the Δ*mxaF* strain. Notably, the *xoxF* gene was also essential for growth on the standard methanol medium. This is in line with findings in strain AM1, where a strain lacking both *xoxF1* and *xoxF2* was shown to be unable to grow on methanol [25]. Intriguingly, the *mxaF* and *xoxF* knockouts as well as the double knockout (Δ*mxaF*Δ*xoxF*, data not shown) all showed very slow growth on methanol in both presence and absence of $La^{3+}$, suggesting the presence of an additional low-activity methanol dehydrogenase.

To further investigate the expression of the two Mdh-systems, we analyzed the activities of P*mxa* and P*xox* in the wildtype and both knockout strains using *mCherry* fusions. Since the knockout strains do not grow on methanol, these experiments were done on a combination of succinate and methanol (see Figure S1). In the wildtype strain grown in the absence of $La^{3+}$, the *mxa* promoter was active, while the *xox* promoter activity was low (~3% of P*mxa*). The expression pattern of P*mxa* and P*xox* in the Δ*mxaF* strain was identical to the wildtype, but the pattern was changed dramatically in the *xoxF* knockout. In the Δ*xoxF* strain, P*mxa* was turned off and P*xox* was induced to a level of about 60% of the *mxa* promoter in the wildtype (Figure 2). These results are in line with findings in strain AM1 [25] and suggest that XoxF activates the expression of the *mxa* cluster and inhibits the expression of its own cluster either directly or via other factors.

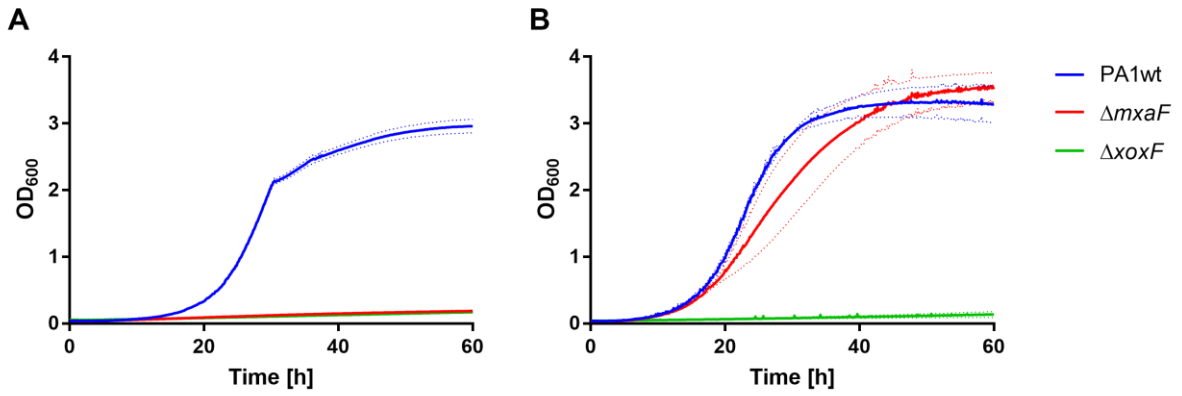


**Figure 1:** Growth of wildtype ('PA1wt'), Δ*mxaF*, and Δ*xoxF* on (A) the standard methanol medium and (B) methanol medium containing 30 µM $La^{3+}$. Mean and standard deviation (dotted lines) are shown (n = 3). Growth rates are listed in Table S1.

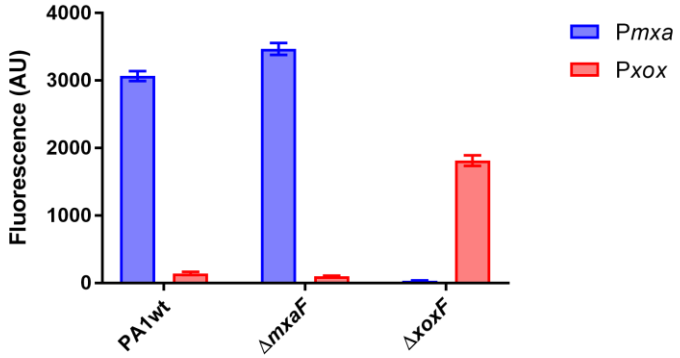


**Figure 2:** Promoter activities of P*mxa* and P*xox* in the wildtype (PA1wt) and knockouts in *mxaF* (Δ*mxaF*) or *xoxF* (Δ*xoxF*). All strains were grown on succinate with addition of methanol (n = 4).

### 3.3.2 $La^{3+}$ induces the expression of the *xox* cluster in a dose-dependent manner

In the absence of $La^{3+}$, *xoxF* was barely expressed (see above), but the fact that XoxF is required for growth on methanol in the presence of $La^{3+}$ (in the Δ*mxaF* strain) raises the question if $La^{3+}$ induces its expression. Indeed, $La^{3+}$ strongly induced P*xox* in the wildtype and at the same time inhibited P*mxa* (Figure 3A). Furthermore, closer inspection of the activity of the *xox* promoter with different $La^{3+}$ concentrations showed that full activation was almost achieved with 3 µM $La^{3+}$, with approximately linear titratability at concentrations between 0.5 and 3 µM. This mostly agrees with recent findings in strain AM1 [21], except that higher $La^{3+}$ concentrations were required for full activation in strain PA1. The finding that the expression is switched to *xoxF* in the presence of $La^{3+}$, notwithstanding the presence of $Ca^{2+}$, implies that the $La^{3+}$-dependent XoxF is favored over the $Ca^{2+}$-dependent MxaFI. The titratability of P*xox* suggests that the bacterium adjusts the expression levels of its two Mdh systems in response to the available $La^{3+}$ concentration.

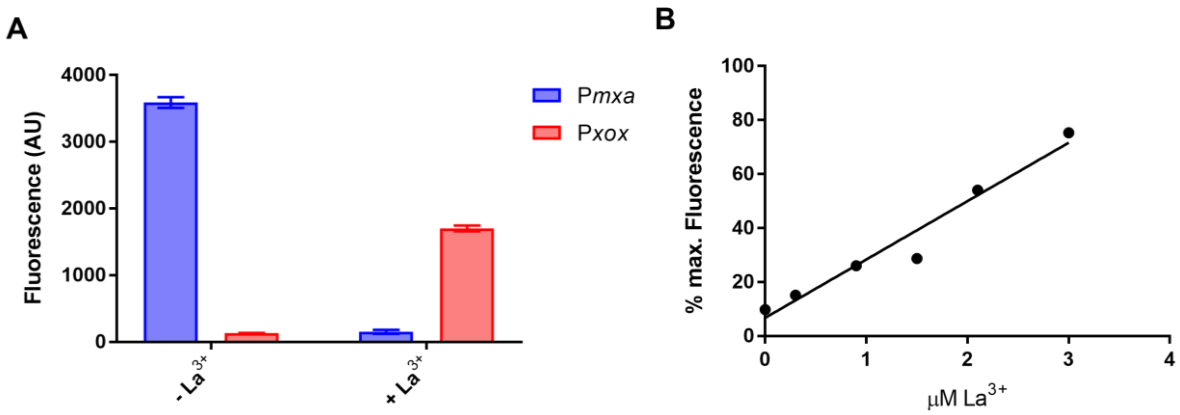


**Figure 3:** (A) Promoter activities of P*mxa* and P*xox* in the wildtype grown in methanol medium (containing $Ca^{2+}$) in the absence (-$La^{3+}$) or presence (+$La^{3+}$) of 30 µM $La^{3+}$ (n = 4). (B) $La^{3+}$ titration of P*xox* of the wildtype grown in succinate plus methanol, given in percent of maximal P*xox* activity (n = 1).

### 3.3.3 Suppressor mutants of Δ*xoxF* occur readily and express the *mxa* cluster again

As discussed above XoxF seems to regulate the expression of the *mxa* cluster as well as its own expression. In addition to XoxF, two additional two-component systems, MxcQE and MxbDM are involved in the signaling cascade regulating *mxa* expression, but the details are not completely understood [25]. In order to investigate the signaling cascade downstream of XoxF, we performed a suppressor screen to identify potential mutations that would overcome the effect of a *xoxF* knockout. We were indeed able to identify several suppressor mutants after prolonged incubation of Δ*xoxF* on methanol (Figure 4A, Table S1), as previously observed in strain AM1 [25].

We investigated two of these suppressor mutants in more detail. One of the suppressors ("Δ*xoxF*_sup1") showed wildtype-like growth, while the other one showed a prolonged lag-phase that was variable between experiments ("Δ*xoxF*_sup2"). Next, we confirmed that the suppressor phenotype was due to re-appearance of methanol dehydrogenase activity (Figure 4B). Investigation of the methanol dehydrogenase activity in these suppressor strains showed wildtype-like activities for the two investigated strains grown on methanol (without $La^{3+}$). The activity on succinate was roughly wildtype-level in Δ*xoxF*_sup1, while no activity was measurable in the *mxaF* or *xoxF* knockouts and surprisingly not in Δ*xoxF*_sup2 either. To confirm that the measured activity in the suppressor mutants was due to MxaFI and not another methanol dehydrogenase, activities of P*mxa* and P*xox* were measured. Indeed, the P*mxa* was active again during growth on methanol in both suppressors (Figure 4C), but in line with the Mdh assay, P*mxa* was very low on succinate (plus methanol) in Δ*xoxF*_sup2 (Figure 4D). Like in the wildtype, the *xox* promoter was not induced in the absence of $La^{3+}$ in Δ*xoxF*_sup1 on both methanol and succinate (plus methanol), but not so in Δ*xoxF*_sup2 where it has

some residual activity on methanol and approximately half-maximal activity on succinate (plus methanol).

Taken together, these results indicate restored wildtype Mdh regulation in Δ*xoxF*_sup1 and an intermediate regulation in Δ*xoxF*_sup2. Δ*xoxF*_sup2 behaves differently under selective (methanol) than under non-selective (succinate plus methanol) conditions. This could be either an intriguing regulatory phenotype or due to accumulation of further mutation(s) in the strain during growth on methanol, but not on succinate. The latter hypothesis is backed by the prolonged lag-phase of Δ*xoxF*_sup2 on methanol, which could be due to emergence of "double" suppressor mutants. Further experiments such as re-inoculation of the emerging population and sequencing will be required to determine whether the genotype of Δ*xoxF*_sup2 is changed.

### 3.3.4 Two independent Δ*xoxF* suppressors showed distinct mutations in *mxbD*

To pinpoint the suppressor mutations, we sequenced the genome Δ*xoxF*_sup1. This revealed a duplication of seven amino acids (21 bases) in *mxbD* (Mext_1822), encoding a sensor histidine kinase that was previously shown to be involved in the regulation of *mxa* expression. Sanger sequencing confirmed the mutation in Δ*xoxF*_sup1 and furthermore revealed a point mutation in *mxbD* of Δ*xoxF*_sup2 (Ala324Pro). Protein sequence analysis of MxbD showed that the duplication in Δ*xoxF*_sup1 at the end of the predicted HAMP (Histidine kinases, Adenylate cyclases, Methyl accepting proteins, Phosphatases) domain that was shown to be involved in the signal transduction from the periplasmatic sensor domain to the cytosolic catalytic domain in other systems [28]. The Ala324Pro mutation of Δ*xoxF*_sup2 is located in the histidine kinase (HisKA) domain in a predicted alpha-helix close to phosphor-acceptor histidine (His308). To confirm that the observed mutations are causal for the suppressor phenotype, we recreated the suppressor mutants by knocking out *mxbD* in the Δ*xoxF* background and re-introducing mutated copies of *mxbD* under control of the native promoter P*mxb* on a plasmid. The Δ*xoxF* Δ*mxbD* strains complemented with the two suppressor copies of *mxbD* showed growth on methanol plates, while strains complemented with empty plasmids or plasmids harboring wildtype copies of *mxbD* did not grow (see Figure S2).

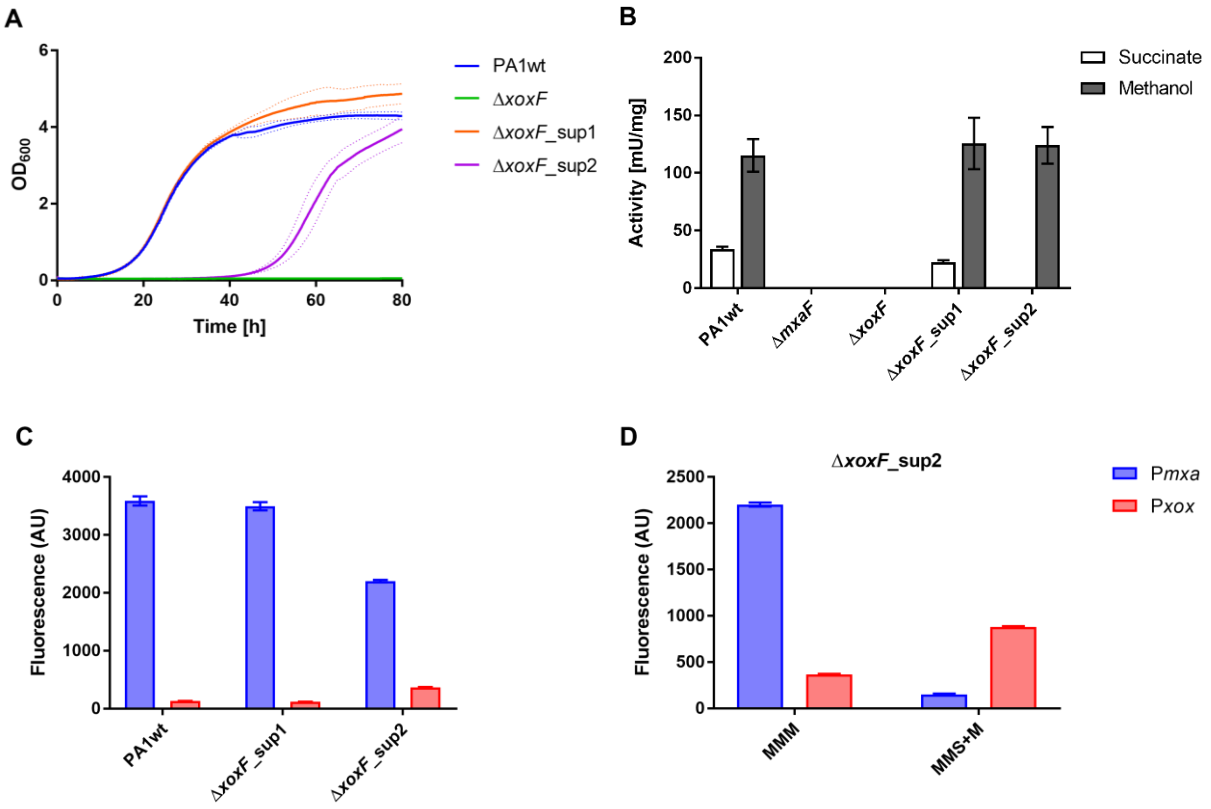


**Figure 4:** Phenotypes of Δ*xoxF* suppressor mutants in the absence of $La^{3+}$. (A) Growth of wildtype and two Δ*xoxF* suppressor mutants on methanol minimal medium (n = 3). Growth rates are listed in Table S1. (B) Methanol dehydrogenase activities measured in the wildtype, knockouts in *mxaF* or *xoxF*, and two Δ*xoxF* suppressor mutants grown on succinate or methanol. Knockout strains were only measured on succinate, because no growth occurs on methanol (n = 3, technical). (C) Promoter activities of P*mxa* and P*xox* in the wildtype and two Δ*xoxF* suppressor grown in methanol medium (n = 4). (D) Promoter activities of P*mxa* and P*xox* in Δ*xoxF*_sup2 grown in methanol ("MMM") or succinate medium containing methanol ("MMS+M") (n = 4).

### 3.3.5 Effect of $La^{3+}$ on the Δ*xoxF* suppressors shows that $La^{3+}$ is sensed in the absence of XoxF

Suppressor strains provide a unique opportunity to investigate $La^{3+}$ regulation in the absence of XoxF, which was suggested to be the $La^{3+}$-sensor. This is otherwise challenging due to the requirement of *xoxF* for the expression of *mxa* and therefore growth on methanol, also in the absence of $La^{3+}$. In the wildtype, promoter activity of P*mxa* was turned off in the presence of $La^{3+}$, while P*xox* was turned on (see above) and normal growth is supported by either MxaFI or XoxF. In the case of the Δ*xoxF* suppressor mutants, growth can only be achieved by MxaFI (or potentially another Mdh), since *xoxF* is knocked out. Therefore, wildtype-like $La^{3+}$ regulation leading to a switch to P*xox* expression would result in very slow or no growth. Wildtype-like growth in the presence of $La^{3+}$ would suggest a lack of

$La^{3+}$ regulation/sensing. The growth of both *xoxF* suppressor mutants on methanol in the presence of $La^{3+}$ was delayed compared to growth in the absence of $La^{3+}$ to variable extents depending on the experiment (Figure 5A, Table S1). Δ*xoxF*_sup2 was not analyzed further at this point due to the likely accumulation of mutations under selective conditions (see above). In the Δ*xoxF*_sup1 strain, P*xox* activity was low in the absence of $La^{3+}$ and only increased slightly in the presence of $La^{3+}$. This observation is in line with a phenotype reverted to wildtype that lost responsiveness to $La^{3+}$. On the other hand, P*mxa* activity reached wildtype-levels in the absence of $La^{3+}$, but was decreased by half in the presence of $La^{3+}$, showing that $La^{3+}$ is still sensed to some extent by Δ*xoxF*_sup1 (Figure 5B).

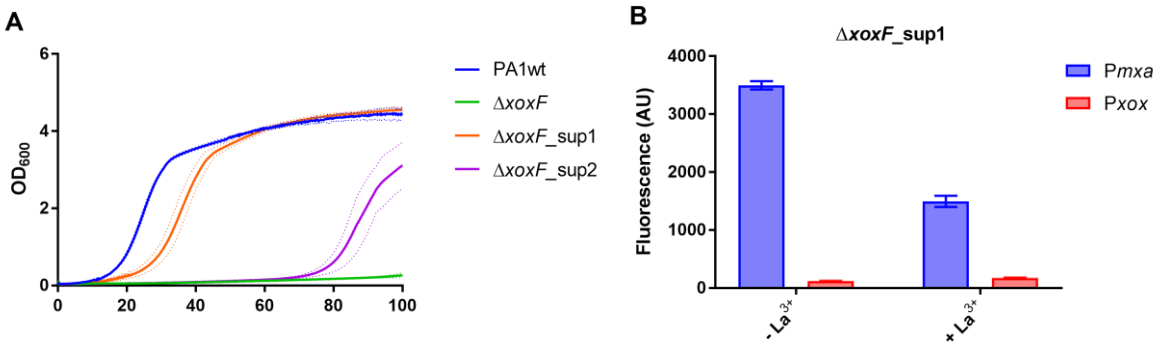


**Figure 5:** Effect of $La^{3+}$ on Δ*xoxF* suppressor mutants. (A) Growth of wildtype, the *xoxF* knockout and the two Δ*xoxF* suppressor mutants on methanol minimal medium (n = 3). Growth rates are listed in Table S1. (B) Promoter activities of P*xox* and P*mxa* in the presence and absence of 30 µM $La^{3+}$ in the two Δ*xoxF* suppressor mutants grown on methanol (n = 4).

## 3.4 Discussion

In this study, we showed that *M. extorquens* strain PA1 is able to grow on methanol in the presence of the rare-earth element $La^{3+}$ and that XoxF is responsible for the observed growth, confirming that the findings in strain AM1 [15,21] also hold true for strain PA1. Furthermore, as described for the two XoxF proteins in strain AM1 [25], the single XoxF protein of strain PA1 is involved in the regulation of both the regulation of *mxa* cluster expression and expression of its own gene cluster.

The fact that *M. extorquens* PA1 Δ*xoxF*Δ*mxaF* double mutant still shows slow residual growth on methanol in both the absence and presence of $La^{3+}$ suggests the presence of another, unknown methanol oxidation capacity. Recently growth of the AM1 Δ*mxaF*Δ*xoxF1*Δ*xoxF2* strain on methanol in the presence of $La^{3+}$ was attributed to the $La^{3+}$/PQQ-dependent ethanol dehydrogenase *exaF* [19]. In contrast to this study, we also observed background growth in the absence of $La^{3+}$ and growth was slower in the case of PA1, suggesting that it is mediated by another activity.

In the presence of $La^{3+}$, expression is switched from *mxa* to *xox* in a dose-dependent manner, notwithstanding the presence of $Ca^{2+}$. This implies that the $La^{3+}$-dependent XoxF is favored over the

$Ca^{2+}$-dependent MxaFI and that the bacterium adjusts the expression level of its two Mdh systems in response to the available $La^{3+}$ concentration. Compared to a recent study with strain AM1 [21], the PA1 *xox* promoter required roughly ten times higher $La^{3+}$ concentrations (>3 µM instead of 0.25 µM) for full activity. A possible reason for this is that we use a phosphate-based medium which was identified as problematic due to formation of insoluble lanthanum phosphate ($LaO_4P$) [20], while Vu *et al*. used a PIPES-based medium [29].

To date it is not completely understood how the $La^{3+}$-dependent regulation of Mdh expression is achieved. A recently published model suggests XoxF as $La^{3+}$-sensor that in its inactive form (no $La^{3+}$ bound) serves as an activating signal sensed by the sensor kinase MxcQ which eventually leads to the activation of *mxa* expression [21]. Direct evidence for the role of XoxF as $La^{3+}$-sensor is difficult to obtain due to *xoxF* being essential for grow on methanol (even in the absence of $La^{3+}$) and because of the effect of a *xoxF* knockout on P*mxa* (and P*xox*) being identical to the effect of $La^{3+}$ addition. Δ*xoxF* suppressor mutants offer a valuable tool to investigate the regulatory cascade. Such suppressors have already been described in strain AM1, but the responsible mutation has not been identified [25].

We sequenced two PA1 Δ*xoxF* suppressor mutants and found causative mutations in the sensor kinase gene *mxbD* in both cases. Δ*xoxF*_sup1 reproducibly showed wildtype-like growth and P*mxa* expression levels on methanol in the absence of $La^{3+}$, while Δ*xoxF*_sup2 seems to be in an intermediary state and prone to further mutate under selective conditions. Since we cannot exclude changes in the genotype of Δ*xoxF*_sup2 depending on the growth conditions, rendering comparisons between conditions impossible, we focused on Δ*xoxF*_sup1.

If XoxF is the only $La^{3+}$-sensor in the cell (suggested in the above-described model, [21]), the addition of $La^{3+}$ to Δ*xoxF*_sup1 (wildtype-state in absence of XoxF) should have no effect on growth or Mdh expression whatsoever. Surprisingly, when $La^{3+}$ was added to the medium Δ*xoxF*_sup1 showed a growth delay and decreased activity of P*mxa*. This implies that XoxF is not the only $La^{3+}$ sensor and that the mutated MxbD is partially but not fully irresponsive to $La^{3+}$. Our results strengthen the hypothesis that $La^{3+}$ is sensed at the level of MxbD or above.

Further experiments are required to understand the complex regulatory cascade of the two methanol dehydrogenase systems in response to $La^{3+}$. Our data suggest that the regulation is even more complex than previously assumed and likely involves additional components that are yet to be identified.

**Acknowledgements**

We would like to thank Catherine Aquino from the functional genomics center Zürich (FGCZ) for next-generation sequencing.

## 3.5 Methods

### 3.5.1 Strains and growth conditions

*M. extorquens* PA1 was grown at 28°C in minimal medium containing 20.7 mM phosphate buffer, 30.29 mM $NH_4Cl$, 0.81 mM $MgSO_4$, trace elements (40.3 µM $FeSO_4$, 15.65 µM $ZnSO_4$, 12.61 µM $CoCl_2$, 5.09 µM $MnCl_2$, 16.17 µM $H_3BO_3$, 1.65 µM $Na_2MoO_4$, 1.2 µM $CuSO_4$ and 20.41 µM $CaCl_2$), and methanol (123 mM) or succinate (30.83 mM). Depending on the experiment, 30 µM $LaCl_3$ heptahydrate (Sigma-Aldrich, Cat#262072) was added. For plates 15 g/L agar was added. *Escherichia coli* DH5α was used for molecular work and was cultured at 37°C in LB medium. If appropriate, tetracycline (10 µg/mL) or kanamycin (50 µg/mL) were added for selection.

### 3.5.2 Construction of knockouts, expression plasmids and promoter fusions.

Gene deletions of *mxaF* (Mext_4150) and *xoxF* (Mext_1809) were generated using pK18_*mobsacB* [30], while the gene deletion of *mxbD* (Mext_1822) was performed using pCM433 [31].
Plasmids for expression of *mxbD* were constructed based on pTE100 [32] by amplifying the *mxbD* gene including its promoter (265 bp upstream) from the corresponding strains (wildtype, Δ*xoxF*_sup or Δ*xoxF*_sup2) and inserting it using the EcoRI and XbaI sites.
The promoter fusion plasmid of P*xox* was constructed based on pTE100_mChe [32] by cloning 264 bp upstream of *xoxF* into the EcoRI and XbaI sites of the vector. The promoter fusion plasmid of P*mxa* was constructed based on pTE102 [32] by cloning the *mCherry* gene of pTE100_mChe into the SpeI and KpnI sites of the pTE102 vector.
Electrocompetent *M. extorquens* PA1 cells were prepared using the following protocol; exponentially growing (OD600 between 1 and 3) were incubate on ice for 30 min, spun down at 4000 g at 4°C for 15 min. The pellet was washed twice with 1 volume ice-cold sterile MilliQ water, followed by 0.5 volumes ice-cold sterile 10% glycerol and finally resuspended in 0.01 volumes 10% glycerol. Competent PA1 cells were stored in 100 µL aliquots at -80°C.
Electroporation of (suicide) plasmids into competent *M. extorquens* PA1 cells was performed at 1.8 (2.15) kV, followed by regeneration in nutrient broth (NB) without NaCl (Sigma-Aldrich) at 28°C for at least two (five) hours and plating on appropriate minimal medium plates.

### 3.5.3 Growth assays in microtiter plates

Overnight pre-cultures were grown in 20 mL MMS with additional 123 mM methanol in shake flasks. To remove the remaining substrate, cells were spun down at 3220 g at 28°C for 15 min, the pellet was resuspended in minimal medium without carbon source (MM), spun down again (see above) and

finally resuspended in MM. Next, the OD600 was adjusted to 0.5 to inoculate the main cultures in 96-well plates (ThemoFischer Scientific Nunclon 96 Flat Bottom Transparent Polystyrol). 180 µL of medium containing the respective carbon source plus 10% to compensate dilution by inoculation were pre-dispensed and inoculated with 20 µL of the washed culture (start OD600 = 0.05). OD600 was measured every 10 min using a Tecan Infinity M200 Pro spectrophotometer (Tecan) with a bandwidth of 9 nm and 25 flashes. The plates were shaken with 1 mm amplitude for 522 s between measurements while incubating at 28°C.

### 3.5.4 Promoter fluorescence assays

Fluorescence quantification was performed either online in 96-well format or as single measurements in cuvettes. The 96-well setup was as described above except that black plates with a clear bottom (CELLSTAR, µClear black) in combination with a lid from a tissue culture test plate (f-base) and that fluorescence was measured every 10 min in addition to OD600. The following settings were used for the fluorescence measurement: excitation at 554 nm, emission at 610 nm, a Z-position of 17706 µm, and a gain of 100. The shaking time was reduced to 442 s to accommodate the additional measurement. Fluorescence was compared at an OD600 of 0.28. The fluorescence value of the negative control (wt + pTE714_empty) was subtracted except if the initial fluorescence value of the well was higher than the negative control (due to promoter expression on medium used for pre-cultures, i.e. P*mxa* pre-grown on MMM, main culture in MMM+$La^{3+}$).
For measurements in the single cuvette setup, cells were grown until an OD600 of 1, put on ice and fluorescence was quantified in a quartz cuvette (HellmaAnalytics, 105.251-QS, 3 x 3 mm) using a Cary Eclipse fluorescence spectrometer (Varian) with excitation at 560 nm and excitation at 610 nm.

### 3.5.5 Methanol dehydrogenase activity assay

The activity of PQQ-dependent methanol dehydrogenase was measured as described previously [4]. The assay mix contained 100 mM Tris-HCl (pH 9.0), 1.1 mM phenanzine ethosulfate (PES), 43.3 µM 2,6-dichlorophenol-indophenol (DCPIP), 1 mM KCN, 5.3 mM methanol and 15 mM $NH_4Cl$. The reaction was started with the addition of $NH_4Cl$ and the decrease of absorbance at 600 nm (due to reduction of DCPIP) was measured over time using a Cary 50 Bio UV-VIS spectrometer (Varian). The slope (after subtracting the background before $NH_4Cl$ addition) was used to calculate enzyme activity in µmol $min^{-1}$ $mg^{-1}$ using the law of Beer-Lambert ($\varepsilon_{DCPIP}$ = 19100 $M^{-1}cm^{-1}$).

### 3.5.6 Whole genome sequencing

For whole genome sequencing of Δ*xoxF*_sup1, the genomic DNA was extracted using the MasterPure™ DNA and RNA Purification Kit (epicentre) and sequenced using paired-end Illumina HiSeq sequencing at the Functional Genomics Centre Zurich.

## 3.6 References


1. Chistoserdova, L. (2015). Methylotrophs in natural habitats: current insights through metagenomics. Appl Microbiol Biotechnol *99*, 5763–5779.
2. Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M. V, Marx, A., and Vorholt, J.A. (2009). Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. Trends Biotechnol *27*, 107–115.
3. Ochsner, A.M., Sonntag, F., Buchhaupt, M., Schrader, J., and Vorholt, J.A. (2014). *Methylobacterium extorquens*: methylotrophy and biotechnological applications. Appl Microbiol Biotechnol *99*, 517-534.
4. Anthony, C., and Zatman, L.J. (1964). The microbial oxidation of methanol. 2. The methanol-oxidizing enzyme of *Pseudomonas* sp. M 27. Biochem J *92*, 614–621.
5. Nunn, D.N., and Lidstrom, M.E. (1986). Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1. J Bacteriol *166*, 581–590.
6. Nunn, D.N., Day, D., and Anthony, C. (1989). The second subunit of methanol dehydrogenase of *Methylobacterium extorquens* AM1. Biochem J *260*, 857–862.
7. Richardson, I.W., and Anthony, C. (1992). Characterization of mutant forms of the quinoprotein methanol dehydrogenase lacking an essential calcium ion. Biochem J *287.3*, 709–715.
8. Amaratunga, K., Goodwin, P.M., O'Connor, C.D., and Anthony, C. (1997). The methanol oxidation genes *mxaFJGIR*(*S*)*ACKLD* in *Methylobacterium extorquens*. FEMS Microbiol Lett *146*, 31–38.
9. Morris, C.J., Kim, Y.M., Perkins, K.E., and Lidstrom, M.E. (1995). Identification and nucleotide sequences of *mxaA*, *mxaC*, *mxaK*, *mxaL*, and *mxaD* genes from *Methylobacterium extorquens* AM1. J Bacteriol *177*, 6825–6831.
10. 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.
11. Chistoserdova, L., and Lidstrom, M.E. (1997). Molecular and mutational analysis of a DNA

region separating two methylotrophy gene clusters in *Methylobacterium extorquens* AM1. Microbiology *143.5*, 1729–1736.

12. Schmidt, S., Christen, P., Kiefer, P., and Vorholt, J.A. (2010). Functional investigation of methanol dehydrogenase-like protein XoxF in *Methylobacterium extorquens* AM1. Microbiology *156*, 2575–2586.
13. Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., and Vorholt, J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc Natl Acad Sci USA *106*, 16428–16433.
14. Taubert, M., Grob, C., Howat, A.M., Burns, O.J., Dixon, J.L., Chen, Y., and Murrell, J.C. (2015). XoxF encoding an alternative methanol dehydrogenase is widespread in coastal marine environments. Environ Microbiol *17*, 3937–3948.
15. Nakagawa, T., Mitsui, R., Tani, A., Sasa, K., Tashiro, S., Iwama, T., Hayakawa, T., and Kawai, K. (2012). A catalytic role of XoxF1 as $La^{3+}$-dependent methanol dehydrogenase in *Methylobacterium extorquens* strain AM1. PLoS One *7*, e50480.
16. Pol, A., Barends, T.R., Dietl, A., Khadem, A.F., Eygensteyn, J., Jetten, M.S., and Op den Camp, H.J. (2014). Rare earth metals are essential for methanotrophic life in volcanic mudpots. Env Microbiol *16*, 255–264.
17. Keltjens, J.T., Pol, A., Reimann, J., and Op den Camp, H.J. (2014). PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. Appl Microbiol Biotechnol *98*, 6163–6183.
18. Chistoserdova, L. (2016). Lanthanides: New life metals? World J Microbiol Biotechnol *32*, 1–7.
19. Good, N.M., Vu, H.N., Suriano, C.J., Subuyuj, G.A., Skovran, E., and Martinez-Gomez, N.C. (2016). Pyrroloquinoline quinone ethanol dehydrogenase in *Methylobacterium extorquens* AM1 extends lanthanide-dependent metabolism to multicarbon substrates. J Bacteriol *198*, 3109–3118.
20. Wehrmann, M., Billard, P., Martin-Meriadec, A., Zegeye, A., and Klebensberger, J. (2017). Functional role of lanthanides in enzymatic activity and transcriptional regulation of pyrroloquinoline quinone-dependent alcohol dehydrogenases in *Pseudomonas putida* KT2440. mBio *8*, 1–14.
21. Vu, H.N., Subuyuj, G.A., Vijayakumar, S., Good, N.M., Martinez-Gomez, N.C., and Skovran, E. (2016). Lanthanide-dependent regulation of methanol oxidation systems in *Methylobacterium extorquens* AM1 and their contribution to methanol growth. J Bacteriol *198.8*, 1250-1259.
22. Chu, F., Beck, D.A.C., and Lidstrom, M.E. (2016). MxaY regulates the lanthanide-mediated methanol dehydrogenase switch in *Methylomicrobium buryatense*. PeerJ *4*, e2435.

23. Chu, F., and Lidstrom, M.E. (2016). XoxF acts as the predominant methanol dehydrogenase in the type I methanotroph *Methylomicrobium buryatense*. J Bacteriol *198*, 1317–1325.
24. Springer, A.L., Morris, C.J., and Lidstrom, M.E. (1997). Molecular analysis of *mxbD* and *mxbM*, a putative sensor-regulator pair required for oxidation of methanol in *Methylobacterium extorquens* AM1. Microbiology *143.5*, 1737–1744.
25. Skovran, E., Palmer, A.D., Rountree, A.M., Good, N.M., and Lidstrom, M.E. (2011). XoxF is required for expression of methanol dehydrogenase in *Methylobacterium extorquens* AM1. J Bacteriol *193*, 6032–6038.
26. 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.
27. 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.
28. Aravind, L., and Ponting, C.P. (1999). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. FEMS Microbiol Lett *176*, 111–116.
29. 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.
30. Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and 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.
31. Marx, C.J. (2008). Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. BMC Res Notes *1*, 1.
32. Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J.A., and Erb, T.J. (2014). 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 *17* 430-443.

## 3.7 Supplementary material

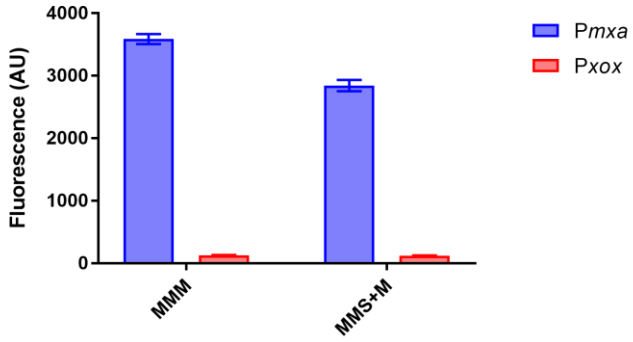


**Figure S1:** Comparison of P*mxa* and P*xox* activity in the wildtype grown on minimal medium containing methanol ('MMM') and succinate plus methanol without $La^{3+}$ ('MMS+M') (n = 4).

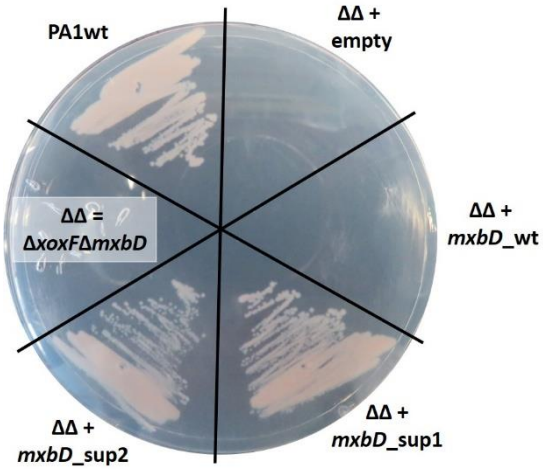


**Figure S2:** Growth of *M. extorquens* Δ*xoxF*Δ*mxbD* strains complemented with either an empty pTE100 plasmid ('ΔΔ+empty') or pTE100 plasmids containing wildtype ('ΔΔ+*mxbD*_wt'), sup1 ('ΔΔ+*mxbD*_sup1'), and sup2 ('ΔΔ+*mxbD*_wt') versions of *mxbD* on a methanol minimal medium plate without $La^{3+}$.

**Table S1:** Growth rates (μ) of wildtype, knockout and suppressor strains grown on methanol minimal medium in the absence of $La^{3+}$ ('MMM') or containing 30 μM $La^{3+}$ ('MMM+$La^{3+}$') (n = 3).

| **Strain** | **Growth rate (μ) $h^{-1}$** | |
|---|---|---|
| | MMM | MMM+$La^{3+}$ |
| PA1wt | 0.1921±0.0007 | 0.205±0.002 |
| Δ*mxaF* | 0.0388±0.0002 | 0.200±0.001 |
| Δ*xoxF* | 0.0190±0.0001 | 0.0207±0.0005 |
| Δ*xoxF*_sup1 | 0.194±0.003 | 0.137±0.004 |
| Δ*xoxF*_sup2 | 0.171±0.006 | 0.134±0.009 |

## Chapter 4:

# Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy

Andrea M. Ochsner, Matthias Christen, Lucas Hemmerle, Rémi Peyraud, Julia A. Vorholt

Current Biology 27.17 (2017) 2579-2588, doi: 10.1016/j.cub.2017.07.025

**Author contributions**

A.M.O., M.C., B.C. and J.A.V. designed the research; A.M.O., M.C., and L.H. performed research; R.P. contributed analytical tools; A.M.O., M.C., B.C., L.H. and R.P. analyzed data; A.M.O. and J.A.V. wrote the paper with input from all authors.

# Chapter 4: Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy

Andrea M. Ochsner[1], Matthias Christen[2], Lucas Hemmerle[1], Rémi Peyraud[3], Beat Christen[2], Julia A. Vorholt[1]

[1]Institute of Microbiology, Department of Biology, ETH Zurich, Vladimir-Prelog-Weg 1-5/10, 8093 Zurich, Switzerland

[2]Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Auguste-Piccard-Hof 1, 8093 Zurich, Switzerland

[3]LIPM, Université de Toulouse, INRA, CNRS, 24 chemin de Borde Rouge, 31326 Castanet-Tolosan, France

## 4.1 Abstract

Methylotrophy is the ability of organisms to grow at the expense of reduced one-carbon compounds such as methanol or methane. Here, we used transposon sequencing combining hyper-saturated transposon mutagenesis with high-throughput sequencing to define the essential methylotrophy genome of *Methylobacterium extorquens* PA1, a model methylotroph. To distinguish genomic regions required for growth only on methanol from general required genes, we contrasted growth on methanol with growth on succinate, a non-methylotrophic reference substrate. About 500'000 insertions were mapped for each condition, resulting in a median insertion distance of five base pairs. We identified 147 genes and 76 genes as specific for growth on methanol and succinate, respectively, and a set of 590 genes as required under both growth conditions. For the integration of metabolic functions, we reconstructed a genome-scale metabolic model and performed *in silico* essentiality analysis. In total, the approach uncovered 95 previously not described genes as crucial for methylotrophy, including genes involved in respiration, carbon metabolism, transport, and regulation. Strikingly, regardless of the absence of the Calvin cycle in the methylotroph, the screen led to the identification of the gene for phosphoribulokinase as essential during growth on methanol but not on succinate. Genetic experiments in addition to metabolomics and proteomics revealed that phosphoribulokinase serves a key regulatory function. Our data support a model according to which ribulose-1,5-bisphosphate is an essential metabolite that induces a transcriptional regulator driving one-carbon assimilation.

## 4.2 Introduction

Methylotrophic bacteria are able to use reduced one-carbon compounds such as methanol as sole sources of carbon and energy. This process raises interest from a fundamental point of view to identify the requisites for methylotrophy and thus achieve an understanding of the enzymes and pathways that allow energy generation from methanol as well as incorporation of reduced one-carbon to build up all the cell constituents [1]. Interest in methylotrophy is also given by the ecological importance of methylotrophs as a crucial part of the global carbon cycle [2]. Moreover, methylotrophy has sparked great demand from an applied perspective for the conversion of highly abundant one-carbon compounds to cell biomass for feed [3] and to value added products for chemical industry either by natural methylotrophs [4,5] or more recently also by platform organisms via means of synthetic biology [6,7] including the design of new pathways [8].

*Methylobacterium extorquens* AM1 has been the primary model strain to study aerobic methylotrophy since its isolation in 1961 [9]. Strain AM1, as other *Methylobacterium* species, is a facultative methylotroph and thus also able to use multi-carbon sources instead of methanol. This ability provides means to distinguish features that are specific for methylotrophy from those generally required for growth. Substantial insights into methylotrophy have been generated over the past decades based on rational, targeted approaches [5,10] and a limited throughput transposon screen [11]. From a systems perspective, the core metabolism of *M. extorquens* has an unusual topology built from an amalgam of enzymes that are unique to methylotrophy, enzymes shared with methanogenic archaea and ubiquitous enzymes. Core enzymes and pathways involved in methanol consumption comprise methanol dehydrogenase enzymes [12] and the tetrahydromethanopterin ($H_4$MPT) pathway [13,14] for one-carbon dissimilation, as well as the serine cycle [15–18] and ethylmalonyl-CoA pathway [19,20] for one-carbon assimilation. These pathways are tightly connected and it has been emphasized that the metabolic network operates as a unique but highly fragile process to achieve one-carbon dissimilation and assimilation [21].

Strain PA1 is phylogenetically closely related to strain AM1 (100% 16 S rRNA gene identity) and was isolated from the phyllosphere of *Arabidopsis thaliana* [22]. Due to several advantages over AM1, including faster growth on methanol, a simpler genome structure [23,24] and better suitability for transposon mutagenesis [25], it has been proposed as a new model strain for methylotrophy [26]. Furthermore, it is a competitive plant colonizer [22] making it a suitable model to study methylotrophy in its natural environment.

The complete gene set essential for growth on methanol has not yet been experimentally determined for a methylotrophic bacterium and it is currently unknown, how many genes in total contribute to methylotrophy as a complex metabolic trait. In this study, we employed transposon mutagenesis coupled to next generation sequencing, i.e. TnSeq [27,28], to identify all genes required for growth of *M.*

*extorquens* PA1 on methanol and, for comparison, succinate. To further integrate data on gene essentiality, we performed mass-spectrometry (MS)-based proteomics and generated and applied a genome-scale model. In-depth analysis of one of the newly found genes essential for methylotrophy revealed a metabolic and regulatory link to a core function of methylotrophy, one-carbon assimilation.

## 4.3 Results and discussion

### 4.3.1 Identification of conditionally required genes by transposon sequencing

To identify genes that are generally required for growth on minimal medium and genes that are specifically required for methylotrophy, we conducted an essentiality screen on methanol and on succinate using high-throughput transposon mutagenesis. A suicide plasmid encoding a Tn5 transposon [29] was introduced into *M. extorquens* PA1 by electroporation. Mutants that carry the Tn5 insertion in a genomic region required for growth on the selected condition are not able to divide, while mutants carrying the Tn5 insertion in a fitness-relevant region grow more slowly. Integration sites were amplified in parallel using a semi-arbitrary PCR [28], followed by sequencing, and mapped to the genome (Figure 1). In total, we determined 486'577 insertions on methanol and 531'410 insertions on succinate, corresponding to a mean transposon insertion density of 11.2 and 10.3 base pairs, respectively (Data S1 and S2) throughout the genome (Figure 2). Depending on the number of Tn5 insertions and their position, we assigned each ORF to one of three categories: essential, fitness (both also referred to as required) or non-essential (see Materials and methods).

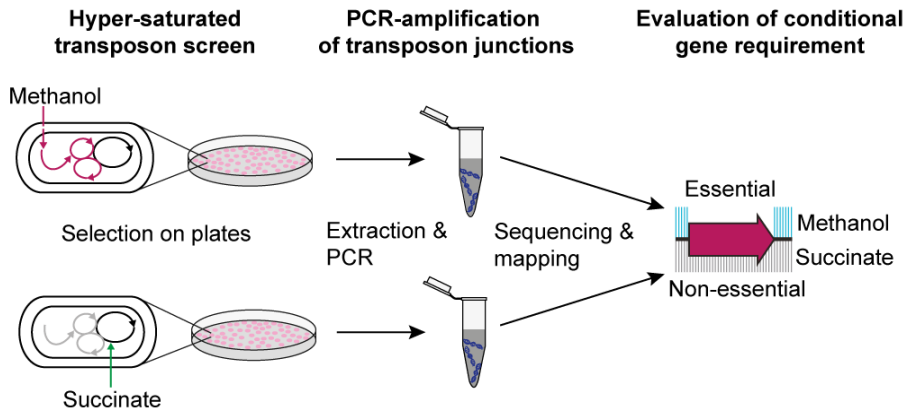


**Figure 1. Experimental setup of transposon sequencing.** Colonies containing random transposon insertions formed on methanol or succinate agar plates are washed off, DNA is extracted, and insertion sites are amplified using a semi-arbitrary nested PCR. The products are then sequenced using paired-end Illumina sequencing, the reads are mapped to the genome, and every gene is rated as "essential" (or "fitness-relevant") or "non-essential" on methanol as well as on succinate, here exemplarily shown for a hypothetic gene.

Overall, the accuracy of the screen is exemplified by the "re-identification" of 52 genes previously known to be essential for growth on methanol (see below). Additionally, we chose a subset of hitherto uncharacterized genes for site-directed mutagenesis and could confirm growth defects (Table S1).
By comparing gene essentiality data sets for the two profiled conditions, we distinguished four main classes of genes: (i) genes that are not required in either condition, (ii) genes that are required in both conditions, (iii) genes that are specific for growth on methanol and (iv) genes that are specific for growth on succinate. The largest part of the genome is not required in either growth condition (83%). From the 813 genes required for growth on methanol and/or succinate, most are shared (590 genes) (Figure 3, Data S1), reflecting the need of the cell to sustain cellular core functions on the minimal medium independent of the growth substrate. Additionally, we noticed that a higher number of genes was required exclusively during growth on methanol (147 genes) compared to succinate (76 genes). This finding is consistent with succinate entering directly into central metabolism, i.e. the TCA cycle, while an extended and specific gene set is needed to connect the entry point of methanol to central metabolism [10,21] and to detoxify formaldehyde, which is an intermediate in methanol oxidation.

### 4.3.2 The core genome

A large fraction of essential genes required for growth on both methanol and succinate comprise housekeeping functions in replication, transcription, translation and cell division, as expected. The largest number of these shared genes encode metabolic enzymes (Figure 3). To relate their function within the metabolic network of the cell and to establish a basis for the analysis of the metabolic trait of methylotrophy, we built a genome-scale metabolic model for *M. extorquens* PA1 (Data S3), manually adapted from the model of strain AM1 [21] (see Materials and methods). The *M. extorquens* PA1 model contains 867 genes (895 reactions) which corresponds to 18% of the total gene number. As expected for growth on minimal medium, the vast majority of genes associated with pathways for amino acid and cofactor biosynthesis are required under both conditions (Figure 4).
We noted, however, several genes for biosynthetic routes that were required during growth on succinate but not methanol (Data S1). One striking set of genes encodes enzymes involved in the biosynthesis of the coenzyme lipoic acid (*lipAB*, Mext_2799 & Mext_4211). Lipoic acid is a coenzyme used by the glycine cleavage system (*gcv*), pyruvate dehydrogenase (*pdh*) and 2-oxoglutarate dehydrogenase (*suc*), which are, notably, all not required during growth on methanol either (Figure 4). Another two genes essential on succinate only encode enzymes involved in serine biosynthesis, phosphoserine phosphatase (*serB*, Mext_2655) and 3-phosphoglycerate dehydrogenase (*serA*, Mext_0213). The requirement of these two genes highlights the need for a classical serine biosynthesis pathway on succinate, in contrast to serine being part of central metabolism and its unique biosynthesis during growth on methanol [21,30].

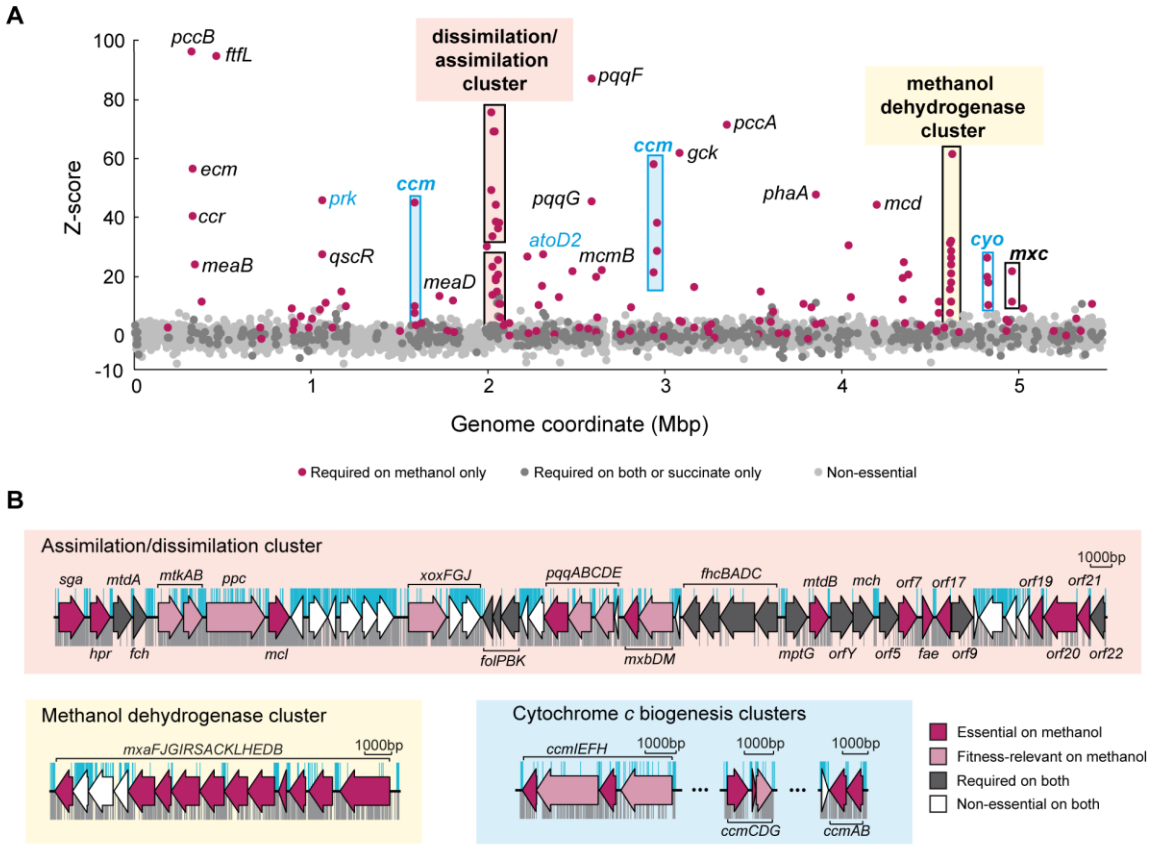


**Figure 2. Genome-wide identification of methylotrophy genes.** A) The Z-score, a measure to compare insertion densities between the two conditions, is plotted for every gene in the genome. Genes that have fewer insertions on methanol have a positive value, whereas genes that have fewer insertions on succinate have a negative value. Only the section between -10 and 100 is shown to highlight the methylotrophy genes, for the entire figure refer to Figure S2. The color of the dot represents the assigned class: non-essential under both conditions (light gray), required under both conditions or only on succinate (dark gray), and required on methanol only (pink). Known genes and clusters (marked with boxes) are shown in black, while some newly identified genes and clusters are shown in blue. All labelled genes are listed with full names in Data S1. B) Insertions detected in selected clusters highlighted in A). Insertions on methanol are shown in blue and insertions on succinate in gray. Arrow length corresponds to gene length and arrow color to essentiality data as determined in the screen. Genes that are only essential or fitness-relevant on methanol are shown in pink (dark and light), genes that are required under both conditions are shown in dark gray, and genes that are not required under either condition are shown in white. See also Figure S1 and S2.

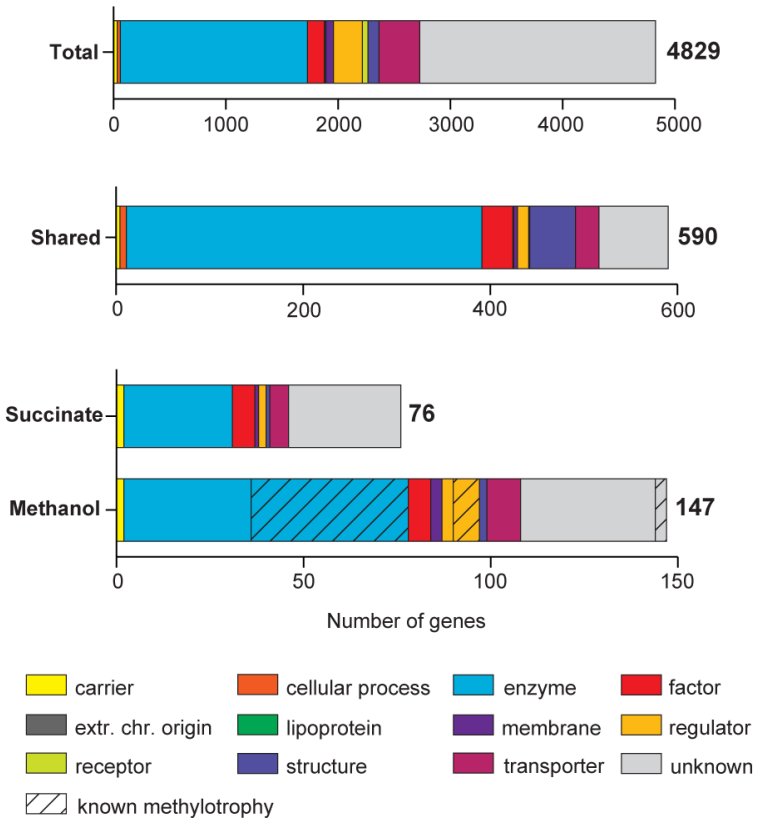


**Figure 3. Functional categorization of shared and specific required genes.** The number of genes belonging to each functional category assigned by the MicroScope platform (see Materials and methods) is shown. "Total" shows the distribution of all ORFs in the genome, "shared" shows the distribution of all genes that are either essential or fitness-relevant on both substrates, while "succinate" and "methanol" show the distributions of genes only essential or fitness-relevant on succinate or methanol, respectively. For the methanol-specific genes, the known genes are shown as hatched parts. See also Data S1.

It can be expected that essential genes are expressed under the chosen growth conditions. To check this and more generally to correlate gene essentiality data with protein abundance data, we applied MS-based proteomics and quantified roughly 1500 proteins (Data S4). The fraction of quantified proteins encoded by the genes that are required under both growth conditions, was significantly higher (437 from 590, 74%) compared to all quantified proteins (1447 from 4845, 30%) and even higher compared to those not required under either conditions (883 from 4017, 22%) (Figure S1). The lower percentage of quantified proteins encoded by non-essential genes might indicate a function that is relevant only under certain environmental conditions.

### 4.3.3 Genome-wide identification of methylotrophy genes

As previously mentioned, we identified 147 genes that are specifically required during growth on methanol (Figure 3, Data S1). First, we analyzed the overall distribution of these genes throughout the genome by calculating a Z-score corresponding to the enrichment of insertions under the two conditions for every gene (see Materials and methods) and plotted it along the genome coordinate (Figure 2A, S2). Two genomic regions essential for methylotrophy stand out, representing the methanol dehydrogenase cluster [31] with 13.9 kb and a large carbon dissimilation/assimilation cluster with almost 50 kb [13] (for insertions in these two large clusters see Figure 2B). The genes in these clusters encode for enzymes and pathways of the known methylotrophy network (Figure 4) including methanol dehydrogenase, the $H_4$MPT- and tetrahydrofolate ($H_4$F)-dependent pathways (partially, however, also required on succinate), and the serine cycle. In addition, we found 17 smaller clusters required for growth on methanol scattered throughout the genome, as well as 75 orphan methylotrophy genes (Figure 2A). The presence of several clusters and numerous orphan genes suggests methylotrophy was not acquired by recent horizontal gene transfer. Instead, this genomic arrangement points to an ancient origin of features, including the acquisition of individual genes and/or repurposing of endogenous genes. This observation is in-line with the deepest rooting of some of the key enzymes involved in methylotrophy [32,33].

Moreover, our genome-wide essentiality analysis revealed 95 previously not described 'methylotrophy genes' (Table S2). Integration with the proteomics data set (Data S4) revealed that only 17% of the detected encoded proteins (6 proteins of 35) were present in higher amounts under methylotrophic compared to non methylotrophic conditions in contrast to about half of known methylotrophy genes that were upregulated under methylotrophic conditions. This suggests that identification of genes via comparative expression analysis had almost reached its limit regarding the identification of novel candidate genes essential for growth on methanol.

In total, 34 of the newly identified 'methylotrophy genes' are categorized as enzymes. A prominent constitutive gene encodes phosphoribulokinase (*prk*, Mext_0980), the investigation of which is described in detail below. Another newly discovered enzyme required for growth on methanol is 5-formyl-$H_4$F-cyclo-ligase (*5fcl*, Mext_2363) that catalyzes the recycling of 5-formyl-$H_4$F. In contrast to 10-formyl-$H_4$F, an intermediate required for one-carbon assimilation, 5-formyl-$H_4$F is known to be an inhibitor of folate-dependent reactions [34]; therefore, its accumulation has to be prevented. We validated the importance of *5fcl* for growth on methanol by targeted knockout and confirmed a significantly reduced growth rate ($t_D$ = 7.89±0.02 h, wildtype: $t_D$ = 3.33±0.01 h, Table S1). The need for *5fcl* can likely be explained by the higher activity of serine hydroxymethyl transferase (GlyA) (Figure 4) on methanol compared to succinate because the enzyme forms 5-formyl-$H_4$F as a side product [35]. Elevated GlyA activity in cell extracts from methanol-grown cells was previously observed [36] and is consistent with higher protein abundance (Data S4). Another gene required specifically during growth on methanol encodes succinyl-CoA:acetoacetyl-CoA transferase (*atoD2*, Mext_2071). The enzyme is

potentially involved in the degradation of polyhydroxybutyrate (PHB). The result of the screen is also reflected in the severe growth defect on methanol of a respective site-directed knockout strain ($t_D$ = 7.56±0.05 h) and strong induction of the encoded protein during growth on methanol compared to succinate (Data S4).

Furthermore, 15 'methylotrophy' genes are involved in respiration (overview in Figure S3). The cytochrome *c* biogenesis system (*ccm*, Mext_1405-7, 2612+4, 2634-5) was strictly required for growth on methanol but not succinate (Figure 2B). We confirmed this result by a site-directed knockout of *ccmAB*. The mutant was unable to grow on methanol and grew normally on succinate (Table S1). In addition, cytochrome *o*-dependent terminal ubiquinol oxidase (*cyoABCD*, Mext_4334-7) was identified as fitness-relevant in both the screen and upon inspection of the respective site-directed mutant ($t_D$ = 4.14±0.03 h) (Table S1). The Ccm system, responsible for the production of functional cytochrome *c*, is expected to be required during methylotrophic growth to allow the function of methanol dehydrogenase [37]. The dispensability for growth on succinate might be explained by electron transfer to one of the terminal ubiquinol oxidases directly from the ubiquinol pool instead of via the *bc1* complex and cytochrome *c* to a terminal cytochrome *c* oxidase. Indeed, the *bc1* complex is non-essential during growth on either methanol or succinate in-line with flux balance analysis (FBA) (Data S1). The need for one of the terminal ubiquinol oxidases during growth on methanol is, however, surprising, since the electron flow from the ubiquinol pool to cytochrome c might circumvent the need of terminal ubiquinol oxidases. One possible explanation is the operation of a branched respiratory chain during growth on methanol. Such an electron flow might diminish a bottleneck via cytochrome *c* and may also be relevant to provide plasticity in case of co-substrate utilization conditions [38]; alternatively, the protein complex could be involved in regulation as described *in Pseudomonas putida* [39].

A total of 10 genes required during growth on methanol are categorized as regulators (Figure 3). Seven of these are involved in the transcriptional regulation of methanol dehydrogenase, serine cycle enzymes, and PHB biosynthesis. One of the three new genes encoding for potential regulators (*glnG*, Mext_2820) might be involved in the regulation of nitrogen metabolism and glutamine biosynthesis together with two other new methylotrophy genes (*glnD*, Mext_0335 and *glnF*, Mext 2501). Alternatively, the proteins could be involved in the regulation of other genes, for example PHB biosynthesis, as it was shown for GlnDKZ in *Azospirillum brasilense* [40]. Further investigation will be required to pinpoint the role of this regulatory cascade in methylotrophy.

Other new genes required for growth on methanol include 9 predicted transporters, and 36 proteins of unknown function. Site-directed knockouts of two genes of unknown functions showed a strong (Mext_1983, $t_D$ = 10.44±0.08 h) and weak (Mext_1769, $t_D$=4.23±0.01 h) fitness-phenotype on methanol (Table S1). Mext_1983 belongs to uncharacterized protein family UPF0262, while Mext_1769 cannot be assigned to any protein family, but is located closely (24 kb upstream) to the dissimilation/assimilation cluster. Further experiments are required to elucidate their role in

methylotrophy. Taken together, the screen revealed a number of genes, previously not known to be involved in methylotrophy.

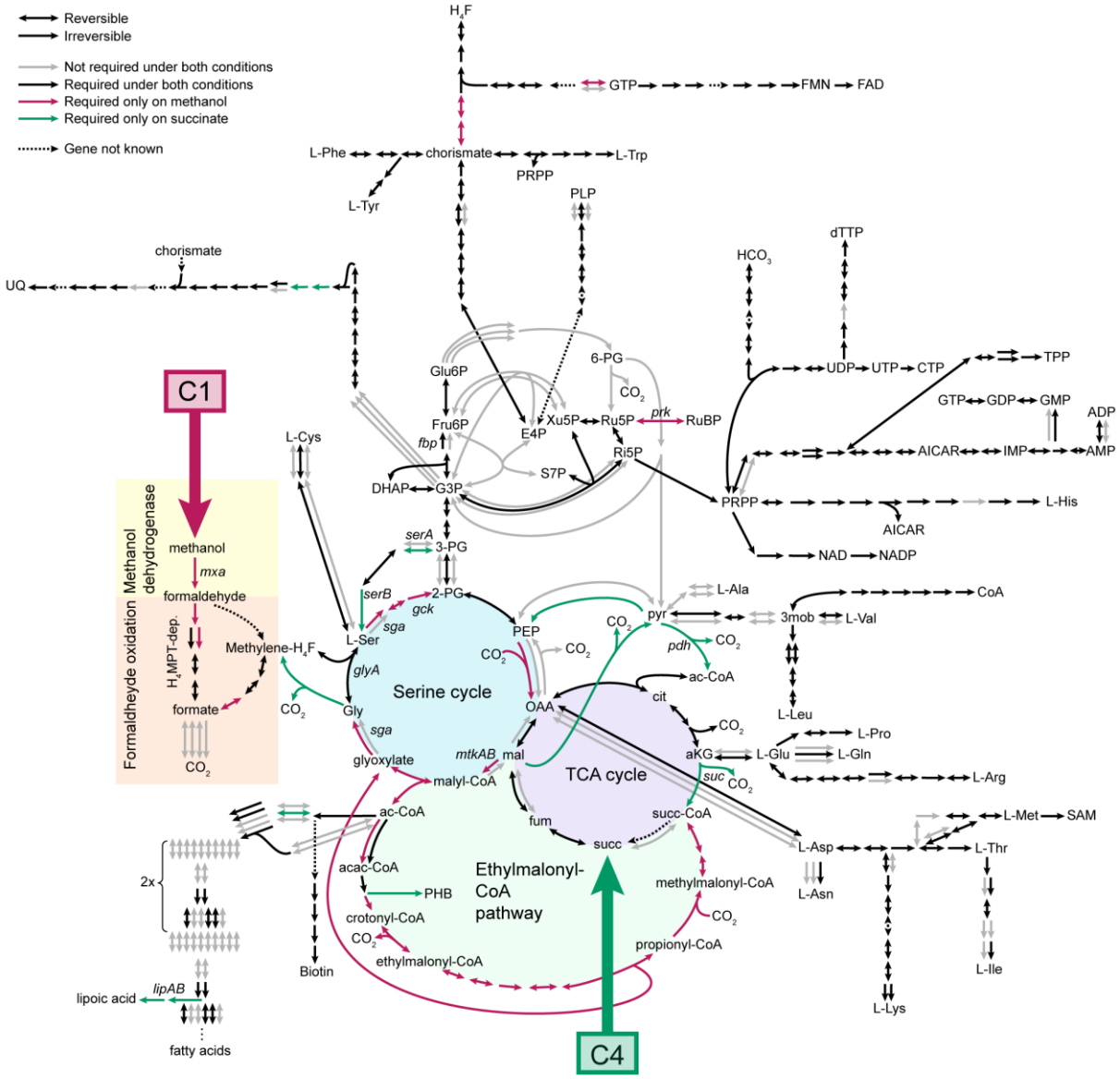


**Figure 4. Integration of conditional gene essentiality using the genome-scale model.** The central part of the genome-scale model is shown. Reversible reactions are shown as double arrows, reactions that show gene redundancy in the model are shown with parallel arrows, and reactions that have no gene assigned are shown as hatched arrows. Gene requirement is indicated by colors; gray for reactions not required under either conditions, black for reactions required (essential or fitness-relevant) under both conditions and pink or green for genes required only on methanol or succinate, respectively. Lipoic acid (*lipAB*) biosynthesis is not in the model and was added manually. See also Figure S3 and Data S3.

### 4.3.4 Phosphoribulokinase is required for methylotrophy

Interestingly, our screen revealed a gene coding for phosphoribulokinase (*prk*, Mext_0980) as a prominent new 'methylotrophy gene'. Not a single insertion in the gene was detected during growth on methanol (Figure 5A, Data S1). We confirmed the failure to grow on methanol by a directed knockout, while the growth rate on succinate remained unaltered (Table S1). A low Prk activity was detected in an earlier study in *M. extorquens* AM1 [36]; however, the requirement of the gene remained unknown. Prk catalyzes the ATP-dependent phosphorylation of ribulose-5-phosphate (Ru5P) to ribulose-1,5-bisphosphate (RuBP), which together with RuBP carboxylase/oxygenase (RuBisCO) are known as unique reactions of the Calvin cycle [41]. Notably, *M. extorquens* PA1, as well as other *Methylobacterium* species, is lacking RuBisCO and thus a functional Calvin cycle.

To validate the functionality of Prk *in vivo*, we conducted a complementation experiment, in which we used the *M. extorquens* PA1 *prk* knockout strain and introduced *prk* from the well-studied Calvin cycle methylotroph *Paracoccus denitrificans* (Prk identity 69%). Indeed, growth on methanol could be restored to wildtype level ($t_D$ = 3.34±0.01 h) (see also Figure 5B). In contrast, complementation with the endogenous *M. extorquens* PA1 Prk possessing an active site point mutation (Asp-42 to Ala) corresponding to the *Rhodobacter sphaeroides* enzyme [42], was unable to restore growth (Figure 5B, Table S1) suggesting that an active Prk enzyme is required during growth on methanol.

Next, we conducted LC-MS-based metabolomics to identify the product of the Prk reaction, RuBP, in *M. extorquens* PA1 cells. We detected RuBP (780±40 nmol/g CDW), when cells were grown on methanol. In contrast, the metabolite could barely be detected in cells grown on succinate (~2 nmol/g CDW) (Figure 5C). This finding is striking because Prk is present under methylotrophic and non-methylotrophic growth conditions and not significantly regulated at the protein level itself (Data S4, Figure S1). The presence of RuBP under methylotrophic conditions may point to an allosteric regulation of Prk and/or different metabolic fluxes and C5 sugar pools. Due to the absence of RuBisCO in *M. extorquens*, the fate of RuBP remains unclear. Pool turnover might be achieved by an unknown enzymatic reaction or by dephosphorylation, e.g. due to the side-activity of fructose-1,6-bisphosphate (FBP) phosphatase [43], the gene for which is located upstream of *prk* in *M. extorquens*. Another hypothesis is that RuBP is not further converted but is produced at such low rate that it dilutes during doubling. To investigate the fate of RuBP in *M. extorquens* PA1, we applied dynamic labelling experiments using $^{13}$C-labelled methanol. We found a relatively fast label incorporation into RuBP ($t_{1/2}$ = ~150 s), albeit at a slower rate compared to serine cycle intermediates (e.g. $t_{1/2}$ = ~10 s for phosphoglycerate). Closer inspection of potential products generated from RuBP was performed by defining a solution space based on possible chemical transformations [44]. However, no putative metabolite could be identified in the dynamic labeling data sets. These data do not disprove further conversion of RuBP, but are principally congruent with the hypothesis that RuBP is reverted back to

Ru5P, thus representing an essential metabolic dead-end product, which may for example fulfill regulatory functions.

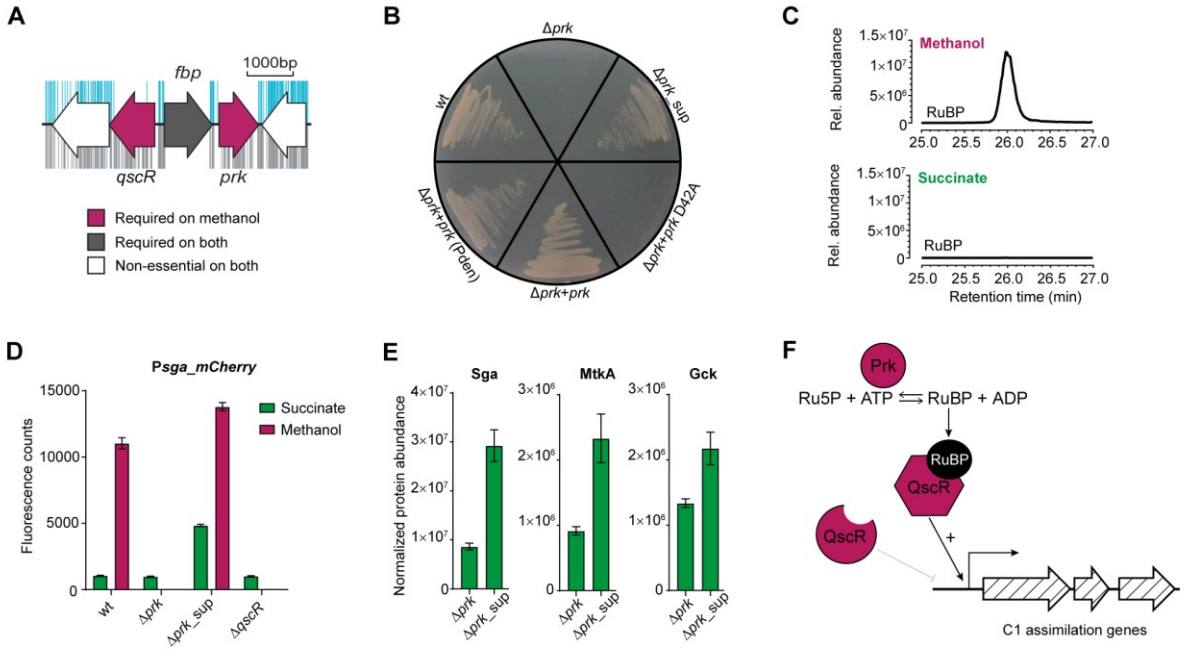


**Figure 5. Analysis of phosphoribulokinase.** A) Transposon insertions into the phosphoribulokinase gene (*prk*) and the neighboring genes encoding for the serine cycle regulator (*qscR*) and fructose-1,6-bisphosphate phosphatase (*fbp*). B) Growth of wildtype ("wt"), *prk* knockout ("Δ*prk*"), Δ*prk* suppressor ("Δ*prk*_sup"), and complementation of Δ*prk* with *prk* from *P. denitrificans* ("Δ*prk*+*prk* (Pden)"), endogenous *prk* ("Δ*prk*+*prk*"), or endogenous *prk* carrying an active-site mutation ("Δ*prk*+*prk* D42A") on a methanol plate. C) Extracted ion chromatogram of ribulose-1,5-bisphosphate (RuBP) in the wildtype strain grown on methanol or succinate. D) Investigation of serine glyoxylate aminotransferase (*sga*) expression level using fluorescence quantification of a reporter construct of P*sga* with *mCherry* on succinate (green) and methanol (pink) in different backgrounds; wildtype (wt), knockouts in *prk* (Δ*prk*) and *qscR* (Δ*qscR*), and the reconstructed Δ*prk* suppressor strain that carries a mutation in *qscR* (Δ*prk*_sup). Δ*prk* and Δ*qscR* were only measured on succinate since they show no growth on methanol. Mean ± SD is shown (n = 8). E) Protein abundance of three serine cycle proteins in the Δ*prk* strain compared to the Δ*prk* suppressor strain grown on succinate; serine glyoxylate aminotransferase (Sga), malate thiokinase subunit A (MtkA), and glycerate kinase (Gck). Mean ± SD is shown (n = 4). F) Model for the function of Prk in methylotrophy. See also Figure S4, Table S1 and Data S4.

### 4.3.5 A regulatory role of phosphoribulokinase for one-carbon assimilation

The *prk* gene is co-localized not only with *fbp*, but also with *qscR* (Figure 5A), the gene encoding a known regulator of serine cycle enzymes (e.g. serine-glyoxylate aminotransferase (Sga), Figure 4) [36], which are key to carbon assimilation in *M. extorquens* and are strongly upregulated during

methylotrophic growth (Data S4). The co-localization with *prk* (and other Calvin cycle genes) is also often observed for the related Calvin cycle regulator *cbbR* [45]. QscR and CbbR both belong to the family of LysR-type transcriptional regulators (LTTR) which are induced by metabolic intermediates, often from the pathway they are regulating [46]. For QscR, formyl-$H_4F$ has been identified as an activating metabolite [47]; for CbbRs it was shown that the regulators are activated by NADPH, ATP, FBP and RuBP in different organisms [45]. Although it has been shown that the expression level of *prk* is not regulated by QscR [36], it is possible that allosteric regulation of QscR by RuBP, the product of Prk, could provide a link to QscR and thus methylotrophy.

To gain insights into the precise function of Prk in methylotrophy, we selected for suppressor mutations that rescue the growth defect on methanol. We were indeed able to isolate Δ*prk* suppressor mutants that grew on methanol, albeit at a slower growth rate than the wildtype (Table S1). Based on our hypothesis that RuBP might activate QscR, we sequenced the *qscR* gene in two independently isolated suppressor mutants. Indeed, sequencing revealed a deletion of one glycine residue (in $Gly_{90\text{-}94}$) located at the interface of the predicted hinge-region and the regulatory domain (based on alignment with the crystallized LTTR-family member CbnR [48]). To show whether the mutation in QscR is responsible and sufficient for the observed suppressor phenotype, i.e. growth on methanol, we exchanged the genomic *qscR* with the mutated version in the Δ*prk* background strain. Indeed, methylotrophy could be restored to similar extent compared to the observed initial suppressor mutants (Δ*prk*_sup $t_D$ = 5.08±0.02 h). We then further substantiated the link between Prk and QscR by reporter gene assays. We fused the promoter of the largest serine cycle cluster (containing *sga*, *hpr*, *mtdA*, *fch* [36], Figure 2B) with *mCherry*, which allowed us to use fluorescence as a readout to detect *sga* promoter activity in different strains. We confirmed the higher activity of the promoter in the reconstructed Δ*prk* suppressor mutant (Figure 5D). In addition, we conducted MS-based proteomics and compared the suppressor strain with the Δ*prk* strain. The data revealed higher abundances of nine proteins known to be under the control of QscR on succinate (Figure 5E, Data S4); however, these did not reach the level of the wildtype growing on methanol (Figure S4), suggesting a partially constitutively active QscR, in-line with the *qscR* suppressor phenotype described above. Notably, such mutations have also been described for other regulators of the LTTR family [45]. Moreover, the overproduction of Prk (~30-fold) did not lead to significantly increased abundance of serine cycle proteins during growth on succinate (Data S4 and Figure S4), further supporting the hypothesis that Prk is activated at the post-translational level. Besides confirming the link between Prk and QscR, the proteome data also allowed to enlarge the known QscR-regulon by additional genes, i.e. formate-tetrahydrofolate ligase *ftfL* and oxalyl-CoA reductase *panE2* (Mext_2139) [49] including its neighboring genes (Mext_2138/40) (Data S4). Due to the central regulatory function of QscR, the regulator is also of interest for re-engineering of *M. extorquens* as was

suggested recently [50]. Thus, tuning RuBP levels via Prk activity may provide an alternative engineering path for the production of value-added products.

Taken together, our data present experimental evidence for a regulatory link between Prk and QscR. According to our model, Prk activity results in an increased RuBP pool, which in turn serves as a metabolic signal to induce one-carbon assimilation *in vivo* (Figure 5F). The allosteric regulation of Prk by NADH (activation) as well as AMP (inhibition) [51] may provide an intriguing network of metabolic signals balancing energy metabolism and carbon assimilation.

## 4.4 Conclusion

In this study, we used a transposon sequencing approach to identify the entire set of genes required for growth on methanol in *M. extorquens* PA1. Comparing deviations in gene essentiality between methylotrophic and non-methylotrophic growth conditions allowed us to pinpoint genes required exclusively for one-carbon metabolism of methanol. Among the genes that are only necessary during growth on methanol are known methylotrophy genes, validating the approach, but also 95 genes that were previously unknown. A number of these specific 'methylotrophy genes' are involved in respiration, emphasizing the large metabolic difference in the oxidation of methanol vs. multi-carbon substrates. Another gene previously not known to be involved in methylotrophy is phosphoribulokinase, which was surprising due to the unique role it has been attributed to in autotrophy. We found that Prk is functionally linked to the serine cycle regulator QscR, a CbbR-ortholog, likely due to the reaction product RuBP serving as co-inducer. Although the signals regulating Prk activity in *M. extorquens* remain to be identified, we speculate that the enzyme is allosterically regulated as described for other bacterial Prks. Our data point to a conserved Prk-dependent regulatory mechanism for one-carbon assimilation in autotrophs and serine cycle methylotrophs. The regulatory role of Prk in conjunction with CbbR/QscR might thus have been conserved during evolution beyond autotrophs and extend to methylotrophs that do not operate the Calvin cycle. Our results thus enlarge the essentiality and known role of Prk in autotrophy to methylotrophy.

Taken together, the generated essentiality data extend our understanding of methylotrophy. At the same time, the data provide a basis for metabolic engineering to convert methanol into value-added products in *M. extorquens* itself [50,52–54] and beyond for transplanting methylotrophy to established platform organisms [6,7].

**Acknowledgements**

We thank Claudia Fortes, Jonas Grossmann and Christian Trachsel from the Functional Genomics Center Zurich (FGCZ) for their support on the proteomics experiments, Catharine Aquino from the

FGCZ for performing next-generation sequencing, and Patrick Kiefer for support with LC-MS measurements and data analysis. This study was supported by a grant from the Swiss SystemsX.ch initiative within the framework of the ERA-Net ERASysAPP, MetApp to JAV, a grant 31003A_166476 from the Swiss National Science Foundation to BC, and by grants from the Eidgenössische Technische Hochschule (ETH) Zürich (ETH-08 16-1 to BC and ETH-41 14-2 to JAV).

## 4.5 Materials and methods

### 4.5.1 Key resource table

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Chemicals, Peptides, and Recombinant Proteins** | | |
| Nutrient broth (NB) without NaCl | Sigma-Aldrich | Cat#17181 |
| $^{13}$C-labelled methanol | Cambridge Isotope Laboratories | Cat#CLM359 |
| Ribulose-1,5-bisphosphate (RuBP) | Sigma-Aldrich | Cat#83895 |
| cOmplete™, Mini EDTA-free Protease Inhibitor Cocktail | Sigma-Aldrich | Cat#11836170001 |
| Tris(2-carboxylethyl)phosphine (TCEP) | Sigma-Aldrich | Cat#75259-10G |
| Iodoacetamide (IAA) | Sigma-Aldrich | Cat#I6125-100G |
| Sequencing grade modified trypsin | Promega | Cat#V5113 |
| iRT peptide kit WR | Biognosys | Prod#Ki-3002 |
| **Critical Commercial Assays** | | |
| Pierce™ BCA protein assay kit | Thermo Fischer Scientific | Cat#23225 |
| **Experimental Models: Organisms/Strains** | | |
| *M. extorquens* PA1 / wild-type | [22] | DSM #23939 |
| *M. extorquens* Δ*prk*::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* Δ*cyoA*::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* ΔMext_1769::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* ΔMext_1983::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* Δ*5fcl*::GmR (fwd) | This paper | N/A |
| *M. extorquens* Δ*atoD2*::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* Δ*ccmAB*::GmR (fwd/rev) | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) pTE105_*prk*_PA1 | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) pTE105_*prk*_PA1_D42A | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) pTE105_*prk*_Pden | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) pTE105 | This paper | N/A |

| | | |
|---|---|---|
| *M. extorquens* PA1 pTE105_*prk*_PA1 | This paper | N/A |
| *M. extorquens* PA1 pTE105 | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) suppressor (natural) | This paper | N/A |
| *M. extorquens* Δ*prk*::GmR (fwd) suppressor (artificial) | This paper | N/A |
| **Deposited Data** | | |
| Raw transposon insertion data | This paper | Data set S2 |
| Proteomics raw data | This paper | ProteomeXchange: PXD006834 |
| **Oligonucleotides** | | |
| Primers used in this study | This paper | Table S3 |
| **Recombinant DNA** | | |
| pTn5_gent_14N | [29] | N/A |
| pREDSIX | [55] | Addgene #74103 |
| pTE105 | [56] | Addgene #59400 |
| pTE100_mChe | [56] | N/A |
| pTE105_*prk*_PA1 | This paper | N/A |
| pTE105_*prk*_PA1_D42A | This paper | N/A |
| pTE105_*prk* (Pden) | This paper | N/A |
| pTE100_P*sga*_mChe | This paper | N/A |
| **Software and Algorithms** | | |
| eMZed2 | [57] | http://emzed.ethz.ch/ |
| DynaMet pipeline | [58] | NA |
| Progenesis LCMS v.3.0.6039.34628 | Nonlinear Dynamics | http://www.nonlinear.com/progenesis/qi-for-proteomics/ |
| Mascot Server 2.5.1.3 | Matrix Science | NA |
| Scaffold version 4.4.1.1 | Proteome Software | http://www.proteomesoftware.com/products/scaffold/ |
| MicroScope platform | [59] | http://www.genoscope.cns.fr/agc/microscope/ |
| InParanoid 4.1 | [60] | http://inparanoid.sbc.su.se/ |
| FlexFlux | [61] | http://lipm-bioinfo.toulouse.inra.fr/flexflux/ |
| OptFlux 3.2.6 | [62] | http://optflux.org/ |
| GraphPad Prism version 7.02 | GraphPad Software, Inc | NA |
| MATLAB version 8.3. | MathWorks | NA |

| **Other** | | |
|---|---|---|
| Dr. Maisch Reprosil-Gold 129, 1.9 µM, 50 x 2 mm | Morvay Analytik | Cat#r119.9g.s0502 |
| Dr. Maisch Reprosil-Gold 120, 3.0 µm, 100 mm x 0.1 mm | Morvay Analytik | Cat#r13.9g.s10 |
| Nunclon 96 Flat Bottom Transparent Polystyrol | Thermo Fischer Scientific | Cat#167008 |
| CELLSTAR 96 well µClear black | Greiner Bio-One | Cat#655087 |
| Tissue culture test plate (f-base) | TPP | Cat#92096 |
| Zirconia/Silica Beads 0.1 mm | BioSpec Products | Cat#11079101z |
| Sep-Pak Vac C18 | Waters Corporation | Cat#WAT023590 |
| SPIN Module and Recovery Tubes | MP Biomedicals (Life Sciences) | Cat#112080800 |

### 4.5.2 Experimental model and subject details

All strains used in this study are listed in the key resource table. *M. extorquens* PA1 was grown at 28°C on minimal medium containing 20.7 mM phosphate buffer, 30.29 mM $NH_4Cl$, 0.81 mM $MgSO_4$ and the following trace elements; 40.3 µM $FeSO_4$, 15.65 µM $ZnSO_4$, 12.61 µM $CoCl_2$, 5.09 µM $MnCl_2$, 16.17 µM $H_3BO_3$, 1.65 µM $Na_2MoO_4$, 1.2 µM $CuSO_4$ and 20.41 µM $CaCl_2$, methanol (123 mM, "MMM") or succinate (30.83 mM, "MMS") were added as carbon source and for plates 15 g/L agar was added. Plasmids were introduced into *M. extorquens* PA1 by electroporation as described previously [63]. Electrocompetent cells were prepared by incubation cells grown to OD1-3 on ice for 30 min, spinning at 4000 g at 4°C for 15 min, washing the pellet twice with 1 volume ice-cold sterile MilliQ water, once with 0.5 volumes ice-cold sterile 10% glycerol and resuspending in 0.01 volumes 10% glycerol. Competent cells were stored in 100 µL aliquots at -80°C.

(Suicide) plasmids were electroporated into competent *M. extorquens* PA1 cells at 1.8 (2.15) kV, the cells were regenerated for at least 2 (5) hours at 28°C in nutrient broth (NB) without NaCl and cells were plated on appropriate minimal medium plates.

*Escherichia coli* DH5α was used for construction and amplification of all plasmids and was cultured in LB medium at 37°C. When appropriate, medium was supplemented with gentamicin (20 µg/mL), tetracycline (10 µg/mL), kanamycin (50 µg/mL) or ampicillin (100 µg/mL).

### 4.5.3 Construction of knockouts and plasmids

All plasmids and primers are listed in the key resource table and Table S3, respectively. Gene deletion mutants were generated using pREDSIX. Upstream and downstream regions of roughly 750 bp were amplified by PCR and cloned into pREDSIX and the gentamicin (or kanamycin)-resistance cassette was

inserted between the regions using the introduced SpeI site (HR1_fwd/rev, HR2_fwd/rev). Both possible insertion directions were selected, the resulting plasmids were transformed into electrocompetent *M. extorquens* PA1 cells (see below) and cells were plated on appropriate minimal medium plates. Subsequently, colonies that did not show fluorescence (i.e. the ones that have lost the *mCherry* encoded on the plasmids backbone) were restreaked and the gene knockout was confirmed by colony PCR using primers outside the homologous regions (check_fwd/rev) combined with primers specific for the antibiotics cassette (GmR_1/2).

Plasmids for *prk* overexpression were constructed based on pTE105 by amplifying *prk* from PA1 (Mext_0980) or *P. denitrificans* PD1222 (Pden_1696) and cloning the respective gene downstream of the P*tuF* promoter using the introduced SpeI and KpnI sites (*prk*_PA1_fwd/rev, *prk*_Pden_fwd/rev). During amplification, a ribosomal binding site (AAGGAGA) and an additional 8 bp were added upstream of the genes using the forward primer. The D42A mutation in Prk was introduced using the Stratagene QuikChange mutagenesis protocol (Agilent Technologies). The corresponding primers were designed using PrimerX (http://www.bioinformatics.org/primerx/) (*prk*_PA1_D42A_fwd/rev).

Promoter fusion plasmids of P*sga* were constructed using pTE100_mChe by amplifying the 305 bp region ranging from 557 to 252 bp upstream of the *sga* (Mext_1795) start codon (based on the QscR-biding sites described in [47]) and cloning upstream of *mCherry* using the introduced EcoRI and XbaI sites (P*sga*_fwd/rev).

### 4.5.4 Transposon mutagenesis, sequencing and analysis

Library construction and DNA extraction. The optimized transposon plasmid pTn5_gent_14N including a barcode [29] was transformed into electrocompetent *M. extorquens* PA1 cells (see above), plated on minimal medium plates containing either methanol or succinate and gentamycin (20 µg/mL), and incubated until colonies formed (3 to 4 days). After incubation, mutant colonies (10'000-15'000/plate) were washed-off from each plate and resuspended in 5 mL 20 mM phosphate buffer pH 7.1. Cell suspensions of two plates were pooled and DNA extracted by spinning down cells, resuspending pellet in 60 µL 5 M guanidinium thiocyonate, incubating for 5 min at 80°C, adding 20 µL 10 M ammonium acetate and spinning down (5 min at 20'000 g). 60 µL of 2-propanol was added to the resulting supernatant, mixed by inverting, spun down again (5 min at 20'000 g), the supernatant was removed, the pellet washed twice with 140 µL 70% ethanol, air-dried and resuspended in 50 µL 1 mM Tris-HCl pH 8.5 by heating to 80°C and vortexing. The concentration of DNA was measured using dsDNA-specific QuantiFluor ONE dsDNA dye with a Quantus fluorometer (Promega).

Amplification of transposon junctions. Transposon junctions were amplified using Taq Polymerase mix (Promega GoTaq) in 10 µL reactions in 384 well plates in two rounds as described previously [28]. In the first round, a semi-arbitrary PCR using 1 µL (1.5 µL for succinate) purified gDNA and the

transposon-specific M13 primer in combination with four barcoded arbitrary primers that differed in their 3' pentanucleotide sequence (ACGCC, TCGCC, CGAGG, CTCGC) [29] was performed. The following PCR program was used for the amplification; (1) 94°C for 3 min, (2) 94°C for 30 s, (3) 42°C for 30 s, slope -1°C/cycle, (4) 72°C for 1 min, go to step 2, 6 times (5) 94°C for 30 s, (6) 58°C for 30 s, (7) 72°C for 1 min, go to step 5, 25 times, (8) 72°C for 3 min, (9) 12° hold.

An aliquot (1 µL) of the first round served as template for the second round of amplification using the Illumina paired-end primers PE1.0 and PE2.0. The following PCR program was applied for the amplification; (1) 94°C for 3 min, (2) 94°C for 30 s, (3) 64°C for 30 s, (4) 72°C for 1 min, go back to step 2, 30 times, (5) 72°C for 3 min, (7) 12°C hold.

All reaction of one condition amplified with the same arbitrary primer were pooled, resulting in two times four pools which were then separated by electrophoresis on a 2% agarose gel. DNA fragments with sizes ranging from roughly 200 to 700 bp were excised and purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). DNA concentrations were measured using a Nanodrop spectrometer and equal amounts were pooled for sequencing.

Illumina sequencing. Samples were sequenced using paired-end 125 bp reads on a HiSeq Illumina instrument using Illumina sequencing chemistry v4. Primers PE1.0 and PE2.0 were used. Standard base-calling from raw images using the genome analyzer software suite OLB (Illumina) was performed. To calculate crosstalk matrixes and to calibrate phasing parameters, a phiX reference spike-in was used. Sequencing was performed at the Functional Genomics Center Zürich.

Sequence analysis and mapping of reads. A custom sequence analysis pipeline based on Python, Biopython [64], bwa [65], and Matlab (as described previously [28,29]) was used to process raw read, align sequences and analyze genomic insertion sites. Reads with a perfect match to the 15 bp end sequence of the transposon (GTGTATAAGAGACAG) were selected for further analysis. Conditions were demultiplexed using the barcode introduced by the arbitrary primer and subsequently the adaptor sequences, the transposon-specific sequence as well as the arbitrary primer sequence were trimmed.

Reads with genomic inserts larger than 15 bp were mapped to the *M. extorquens* PA1 genome NC_010172.1 [24], only considering reads with no mismatches in the first 15 bp that were shorter than 500 bp, and that could be unambiguously mapped. The insertion position was defined as the first genomic base when reading out of the I-end of the transposon.

Essentiality analysis of protein coding sequences. The number and position of insertions disrupting each ORF was determined using a custom Matlab script and insertions in the stop codon as well as insertions in the +1 frame (assembly of functional protein from two polypeptides could still be possible) were excluded. The length of the non-disrupted 5' region, by considering the first not in-frame insertion after the start codon if a second insertion was less than 100 bp apart (otherwise the distance to the next insertion was considered). Furthermore, the largest internal non-disrupted segment was determined with similar criteria. ORFs were categorized as essential if the non-disrupted 5' region covered at least 60%

of the ORF or if the largest non-disrupted internal region covered at least 60% and the insertion density was below one insertion every 50 bp (in case of mis-annotated ORFs). ORFs were categorized as fitness-relevant if they showed less than one insertion every 25 bp and a non-disrupted 5' region covering less than 60% of the gene.

A Z-score metric was employed to compare the number of insertions in both conditions in every ORF. The expected number of insertions on succinate (Succ) was calculated according to

$$Ins_{Succ}^{Exp} = \frac{Ins_{MeOH}^{Meas}}{Ins_{MeOH}^{Total}} * Ins_{Succ}^{Total}$$

where $Ins_{MeOH}^{Meas}$ corresponds to the number of insertion in the respective ORF on methanol and $Ins_{MeOH}^{Total}$ and $Ins_{Succ}^{Total}$ to the total numbers of insertions (in coding regions) measured on methanol and succinate, respectively.

The Z-score was calculated according to

$$Z_{MeOH} = \frac{Ins_{Succ}^{Exp} - Ins_{Succ}^{Meas}}{\sqrt{Ins_{Succ}^{Exp} * (1 - \frac{Ins_{MeOH}^{Meas}}{Ins_{MeOH}^{Total}})}}$$

where $Ins_{Succ}^{Meas}$ is the number of insertions in the respective ORF on succinate.

If the ORF tolerates less insertions on methanol, $Z_{MeOH}$ is negative and if it tolerates less on succinate, it is positive. Since the Z-scores is asymmetric, a second Z-score was calculated by exchanging all values of succinate and methanol ($Z_{Succ}$) and a combined Z-score ($Z_{comb}$) was generated by replacing all positive $Z_{MeOH}$ with $-Z_{Succ}$ (a positive value, because the affected ORFs show a negative $Z_{Succ}$). For graphical representation $Z_{comb}$ was inverted to give positive values for genes that are more important on methanol and zero insertions were set to one to get a (underestimated) Z-score for genes that otherwise could not be plotted at all. A $Z_{MeOH}$ or $Z_{Succ}$ below -10 was used as an additional criterion to rate the corresponding ORFs as fitness-relevant on methanol or succinate, respectively (marked as "fitness-z" in Data S1).

### 4.5.5 Growth- and promoter fluorescence assays

Pre-cultures were performed in 20 mL MMS with additional 123 mM methanol in shake flasks overnight. To remove the remaining substrate, cells were spun down at 3220 g at 28°C for 15 min, washed once with and then resuspended in minimal medium without carbon source. Afterwards $OD_{600}$ was adjusted to 0.5 for inoculation of the main culture in 96-well plates (Nunclon). 180 µL of medium containing one third of the standard amount of the respective carbon source (+10% to compensate dilution by inoculation, 45.1 mM) were inoculated with 20 µL of the washed culture to a start $OD_{600}$ of 0.05. Plates were wrapped in parafilm to prevent evaporation; methanol and succinate experiments were performed separately to prevent cross-contamination with methanol via gas-phase. $OD_{600}$ was measured

using a Tecan Infinity M200 Pro spectrophotometer (Tecan) every 10 minutes with a bandwidth of 9 nm and 25 flashes. Between measurements, the plates were shaken with 1 mm amplitude for 522 s while incubating at 28°C.

For the fluorescence assays, the protocol was essentially as described above with the only changes being the type of plate and the settings of the spectrometer. The main culture was grown in black plates with a clear bottom (µClear black) in combination with a lid from a tissue culture test plate (f-base) and fluorescence (excitation: 554 nm, emission: 610 nm, Z-position: 17706 µm, gain: 100) was measured in addition to $OD_{600}$, which reduced the possible shaking time in between measurements to 442 s. The fluorescence was then compared at identical $OD_{600}$ values.

### 4.5.6 Metabolomics

Dynamic $^{13}C$ labeling experiment. Main cultures were grown in MMM with half the standard amount of methanol (61.5 mM) to an $OD_{600}$ between 1 and 2. For each time point, 0.5 mL of the culture were mixed with 4.5 mL MMM with 123 mM $^{13}C$ labelled methanol and was incubated for 0, 5, 10, 20, 30, 60, 120, 300, and 600 s in a 50 mL falcon tube at 28°C while shaking. Subsequently, medium was removed by filtering through a polyether sulfone (PESU) 0.2 µm filter (Sartorius Stedim), pre-washed with 50°C MilliQ water, applying vacuum. The filter was then washed with 5 mL 28°C warm MilliQ water, transferred into 8 mL -20°C cold quenching solution (60% acetonitrile, 20% methanol, 20% 0.5 M formic acid) incubated on ice for 10 min, and briefly sonified. Afterwards, the filters were removed, the samples frozen in liquid nitrogen and lyophilized overnight. Lyophilized samples were resuspended in MilliQ to a final biomass concentration of 0.5 µg/µL (assuming a CDW correlation of 0.27 mg/mL $OD_{600}$ 1), spun down for 10 min at 20'000 g and supernatant was diluted to 0.1 µg/µL and analyzed using LC-MS (UHPLC method, see below). Untargeted data analysis was performed using the DynaMet pipeline (see Key Resource Table).

Metabolite quantification. Pre-cultures were grown in 20 mL MMS with additional 123 mM methanol overnight and main cultures were inoculated to an $OD_{600}$ of 0.015-0.025 in triplicates and grown until and $OD_{600}$ of 1. Subsequently, 0.5 mL were sampled as described above, with the only difference being that after transferring the filter to the quenching solution, 67.5 µL of fully $^{13}C$-labeled internal standard of strain PA1 (ISTD, 2 mg CDW/mL) which corresponds to the same biomass as the sampled 0.5 $OD_{600}$-units. For LC-MS analysis (nLC method, see below), the resuspended extract (0.5 µg/µL) is diluted to 0.125 µg/µL in 100 µM tributylamine (TBA) pH 9.0 (see below).

For quantification, fully labeled internal standard (ISTD) was mixed in the different ratios with naturally labelled chemical ribulose-1,5-bisphosphate (RuBP) standard. Chemical standard (unlabeled, M0) and ISTD (fully labelled, M+5) RuBP peaks were integrated using eMZed. The resulting linear equation was used to determine the amount of RuBP in the samples using the ratio of unlabeled RuBP to fully

labeled RuBP (ISTD). For samples with ratios below the range of the standard curve, the RuBP concentration was roughly estimated using the less abundant M+4 peak (resulting from incompletely labelled ISTD) as follows; the ratio of the RuBP M+4 and M+5 peaks was determined (in a total of 12 samples) and the theoretical concentration of M+4 was determined. The concentration of RuBP in the samples was than estimated by comparing the areas of the M0 and the M+4 peaks.

LC-MS measurements. Samples were analyzed either using microscale ultra-high-performance liquid chromatography mass spectrometry (UHPLC-HRMS) with a Dionex UltiMATE 3000 (Thermo Scientific) system or a nanoscale HPLC-HRMS with an nLC-ultra (Eksigent) system hyphenated to a Q Exactive Plus (Thermo Scientific) instrument by a heated electrospray-ionization (HESI) probe as described [66]. With the UHPLC system, metabolites were separated on a C18 column (Dr. Maisch Reprosil-Gold 129) using 1.7 mM TBA (dissolved in 1.5 mM acetic acid, pH adjust to 9.0 with ammonium hydroxide) as solvent A. Elution was performed using methanol (solvent B) with a multistep gradient (% B); 0 min, 3%; 9 min, 48%; 13 min, 90%; 15.3 min, 3%; 17.3 min, 3%. 10 µL sample were injected and a flow of 500 µL/min was used.

With the nLC system, metabolites were separated on a C18 column (Dr. Maisch Reprosil-Gold 120) using 100 µM TBA (dissolved in 100 µM acetic acid, pH adjust to 9.0 with ammonium hydroxide) with 3 % methanol as solvent A. Elution was performed using 1:1 (V:V) mixture of 2-propanol and methanol (solvent B) with a multistep gradient (% B); 0 min, 0%; 3 min, 0%; 35 min, 12%; 36 min, 90%; 48 min, 90%; 49 min, 0%; 60 min, 0%. 1 µL sample were injected and a flow of 400 nL/min was used. Mass acquisition was performed in negative Fourier transform mass spectrometry (FTMS) using full MS scan mode with an m/z-range of $150 \leq m/z \leq 1000$.

### 4.5.7 Proteomics

Sampling and cell lysis. All strains were grown as independent pre- and main cultures to an $OD_{600}$ 1.0±0.2. Subsequently, four $OD_{600}$-units (corresponding to 4 mL $OD_{600}$ 1) were sampled by spinning down for 15 min at 4°C and 3220 g, washing with 4 mL 10 mM $MgCl_2$, resuspending in 1 mL 10 mM $MgCl_2$, spinning down again, and shock-freezing pellets in liquid nitrogen.

Preparation of Samples for MS. Bacterial cell pellets were dissolved in lysis buffer containing 100 mM ammonium bicarbonate, 8 M Urea and 1x protease inhibitor cocktail. Bacterial cells were lysed with a combination of indirect sonication (2 x 1 min and 1 x 30 s, 100% amplitude, 0.8 cycle time) in a VialTweeter (HIFU, Hielscher) and bead beating for 15 min at 30 Hz (TissueLyser II, QIAGEN) using 0.1 mm silica beads. Insoluble parts were removed by centrifugation using SPIN module recovery tubes at 13'000 g for 15 min at 4°C and protein concentration of supernatant was determined using BCA assay kit according to manufactures instructions. Protein disulfide bonds were reduced by addition of 5 mM tris(2-carboxylethyl)phosphine (TCEP) and incubation for 30 min at 37°C and cysteine residues were

alkylated by adding 10 mM iodoacetamide (IAA) and incubation for 30 min in the dark at room temperature. Samples were subsequently diluted 1:5 with freshly prepared 50 mM ammonium bicarbonate buffer. Sequencing grade modified trypsin was added at an enzyme to protein ratio of 1:50 and protein digestion was carried out overnight at 37°C with shaking at 300 rpm. Subsequently, trypsin was inactivated by addition of 1% formic acid and incubation at 95°C for 10 min followed by centrifugation at 20000 g for 10 min. The supernatant was subsequently desalted using Sep-Pak Vac C18 reversed phase columns, pre-washed with 100% methanol. For the C18 clean up the columns were activated using buffer A containing 80% acetonitrile (ACN) and 0.1% formic acid (FA) and equilibrated with buffer B containing 3% ACN and 0.1% FA. After careful loading of samples, columns were washed 4 times using buffer B before samples were eluted by gravity flow with buffer C containing 50% ACN and 0.1% FA. Eluted samples were dried under vacuum and re-solubilized in 3% ACN and 0.1% FA to a final concentration of 0.1-1.0 mg/mL. To create an artificial alignment reference that contains the most feature information, equal amounts of all samples were pooled. Finally, iRT peptides were added to each sample.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Mass spectrometry analysis of peptide samples was performed on an EASY-nLC 1000 system (Thermo Fischer Scientific) coupled to a Q Exactive mass spectrometer (Thermo Fisher Scientific). The chromatographic separation was performed using an ACN/water solvent system containing two channels with 0.1% (v/v) formic acid for channel A and 0.1% (v/v) formic acid, 99.9% (v/v) acetonitrile for channel B. For each sample 3 µL of peptides was loaded on an EASY-Spray C18 LC column (75 µm × 500 mm, Thermo Fischer Scientific) heated to 50°C and eluted at a flow rate of 300 nL/min by a gradient from 2% to 30% B in 115 min, 47% B in 4 min, and 98% B in 10 min. The mass spectrometer was configured to acquire mass spectra in data-dependent mode, with an automatic switch between MS and MS/MS scans using a top 12 method. Full-scan MS spectra were acquired in the Orbitrap analyzer with a mass range of 300-1700 m/z and a resolution of 70000 with an automated gain control (AGC) target value of $3 \times 10^6$. HCD peptide fragments (isolation window 2 m/z) were obtained using a normalized collision energy of 25 with an AGC target value of $5 \times 10^4$ at 35000 resolution. To avoid multiple scans of dominant ions, dynamic exclusion was set to 30 s. Sample measurements were acquired using internal lock mass calibration on m/z 371.10124 and 445.12003.

Label-Free Protein Quantification and Protein Identification. The acquired raw MS files were loaded into the commercial software package Progenesis LCMS using the High Mass Accuracy Instrument option. Automatic alignment was performed using the run containing the most features, automatically chosen by Progenesis. In the aligning step, 3−5 vectors along the retention time gradient were manually seeded. From each Progenesis peptide ion (default sensitivity in peak picking) a maximum of the top five tandem mass spectra were exported using the charge deconvolution and deisotoping option and a maximum number of 200 peaks per MS/MS. The Mascot generic file (.mgf) was searched using the

Mascot Server against a forward and reversed protein sequence database containing the 4829 annotated proteins of *M. extorquens* PA1 (https://www.ncbi.nlm.nih.gov/nuccore/NC_010172.1) concatenated with 6721 yeast proteins, 260 known mass spectrometry contaminants, and the 11 iRT peptides. Parameters for precursor tolerance and fragment ion tolerance were set to ±10 ppm and ±0.05 Da, respectively. Trypsin was used as the protein-cleaving enzyme, and one missed cleavage was allowed. Carbamidomethylation of Cys was set as fixed modification, and oxidation of Met was set as variable. The Mascot results were loaded into Scaffold using 5% peptide and 10% protein false discovery rate (FDR). The Scaffold Spectrum Report was exported and imported back into Progenesis LCMS. Normalization was performed using all proteins except for Prk (WP_012252675.1), which was knocked out or overexpressed in the different conditions. For quantification, all proteins identified with at least two unique peptide ions were assessed. A false discovery rate on quantified proteins was estimated at 0.16%. Proteins were grouped with Progenesis and the normalized abundance from the three most abundant peptide ions (relative quantification using Hi-3) from the same protein group were averaged together individually for each sample. For statistical testing, a One-way ANOVA was applied on the normalized protein abundance. The cutoffs for significant regulation were $p<0.05$ and a fold-change larger than 1.5.

### 4.5.8 *In silico* metabolic network analysis

Genome-scale model reconstruction. The existing model of *M. extroquens* AM1 [21] was standardized according to the BiGG Models standards (http://bigg.ucsd.edu). In detail, all reaction and metabolite identifiers were changed according to the standards. Next, the standardized genome-scale model of AM1 was transferred to PA1 strain using the protocol described by Peyraud et al. [67] Briefly, orthologues of all AM1 genes utilized in the model were identified in strain PA1 using the InParanoid software with a BLOSUM45 matrix for prokaryotes using a cut-off of 50 bits and 0.50 sequence coverage. Reactions associated to a least one common orthologue in AM1 were assigned to the PA1 model. The 20 reactions associated with genes that had no orthologue in PA1 were deleted. In addition, the genetic redundancy was reduced for 55 reactions due to a smaller number of genes in PA1. PA1 genes that have no homolog in AM1, but were assigned to be an enzyme according to the genome annotation [23] were manually investigated and were added to the model if a specific function could be assigned (11 new reactions). The model was converted to Systems Biology Markup Language (SBML) and is available in Data S3.
Flux balance analysis. Gene essentiality analysis was performed using the software FlexFlux. Flux balance analysis (FBA) of the wildtype PA1, as well as of single gene knockouts were performed with biomass production (R_Biomass_PA1_core) as the objective function and methanol or succinate as the sole source of carbon and energy as constraint (listed in Data S1). For visualization of the model and of fluxes, the OptFlux software was used.

### 4.5.9 Quantification and statistical analysis

GraphPad Prism, MATLAB, Progenesis and Microsoft Excel 2016 were used for data analysis and visualization. Figures were composed using Adobe Illustrator CC.

At least four biological replicates were performed for growth curves, proteomics and promoter quantification. Mean ± standard deviation (SD) or standard error (SE) are shown as indicated in the legend of the corresponding figure or table.

### 4.5.10 Data and software availability

The raw transposon insertion data is available as Supplemental information.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [68] partner repository with the dataset identifier PXD006834.

## 4.6 References


1. Chistoserdova, L., Kalyuzhnaya, M.G., and Lidstrom, M.E. (2009). The expanding world of methylotrophic metabolism. Annu Rev Microbiol *63*, 477–499.
2. Chistoserdova, L. (2015). Methylotrophs in natural habitats: current insights through metagenomics. Appl Microbiol Biotechnol *99*, 5763–5779.
3. Tlusty, M., Rhyne, A., Szczebak, J.T., Bourque, B., Bowen, J.L., Burr, G., Marx, C.J., and Feinberg, L. (2017). A transdisciplinary approach to the initial validation of a single cell protein as an alternative protein source for use in aquafeeds. PeerJ *5*, e3170.
4. Schrader, J., Schilling, M., Holtmann, D., Sell, D., Filho, M. V, Marx, A., and Vorholt, J.A. (2009). Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. Trends Biotechnol *27*, 107–115.
5. Ochsner, A.M., Sonntag, F., Buchhaupt, M., Schrader, J., and Vorholt, J.A. (2014). *Methylobacterium extorquens*: methylotrophy and biotechnological applications. Appl Microbiol Biotechnol *99*, 517-534.
6. 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., *et al.* (2015). Engineering *Escherichia coli* for methanol conversion. Metab Eng *28*, 190–201.
7. Whitaker, W.B., Sandoval, N.R., Bennett, R.K., Fast, A.G., and Papoutsakis, E.T. (2015). Synthetic methylotrophy: engineering the production of biofuels and chemicals based on the biology of aerobic methanol utilization. Curr Opin Biotechnol *33*, 165–175.
8. Bogorad, I.W., Chen, C.T., Theisen, M.K., Wu, T.Y., Schlenz, A.R., Lam, A.T., and Liao, J.C.

(2014). Building carbon-carbon bonds using a biocatalytic methanol condensation cycle. Proc Natl Acad Sci USA *111*, 15928–15933.

9. Peel, D., and Quayle, J.R. (1961). Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM1. Biochem J *81*, 465–469.
10. Chistoserdova, L., Chen, S.W., Lapidus, A., and Lidstrom, M.E. (2003). Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. J Bacteriol *185*, 2980–2987.
11. Marx, C.J., O'Brien, B.N., Breezee, J., and Lidstrom, M.E. (2003). Novel methylotrophy genes of *Methylobacterium extorquens* AM1 identified by using transposon mutagenesis including a putative dihydromethanopterin reductase. J Bacteriol *185*, 669–673.
12. Anthony, C. (1986). Bacterial oxidation of methane and methanol. Adv Microb Physiol *27*, 113–210.
13. Chistoserdova, L., Vorholt, J.A., Thauer, R.K., and Lidstrom, M.E. (1998). C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. Science *281*, 99–102.
14. Vorholt, J.A., Marx, C.J., Lidstrom, M.E., and Thauer, R.K. (2000). Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. J Bacteriol *182*, 6645–6650.
15. Large, P.J., Peel, D., and Quayle, J.R. (1961). Microbial growth on C1 compounds. 2. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM1, and methanol-grown *Hyphomicrobium vulgare*. Biochem J *81*, 470–480.
16. Large, P.J., Peel, D., and Quayle, J.R. (1962). Microbial growth on C1 compounds. 3. Distribution of radioactivity in metabolites of methanol-grown *Pseudomonas* AM1 after incubation with [$^{14}$C]methanol and [$^{14}$C]bicarbonate. Biochem J *82*, 483–488.
17. Large, P.J., Peel, D., and Quayle, J.R. (1962). Microbial growth on C1 compounds. 4. Carboxylation of phosphoenolpyruvate in methanol-grown *Pseudomonas* AM1. Biochem J *85*, 243–250.
18. Large, P.J., and Quayle, J.R. (1963). Microbial growth on C1 compounds. 5. Enzyme activities in extracts of *Pseudomonas* AM1. Biochem J *87*, 386–396.
19. Erb, T.J., Berg, I.A., Brecht, V., Muller, M., Fuchs, G., and 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*, 10631–10636.
20. Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J.C., and Vorholt, J.A. (2009). Demonstration of the ethylmalonyl-CoA pathway by using $^{13}$C metabolomics. Proc Natl Acad Sci USA *106*, 4846–4851.
21. 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 Syst Biol *5*, 189.

22. 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.
23. Vuilleumier, S., Chistoserdova, L., Lee, M.C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., *et al.* (2009). *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. PLoS One *4*, e5584.
24. 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., *et al.* (2012). Complete genome sequences of six strains of the genus *Methylobacterium*. J Bacteriol *194*, 4746–4748.
25. 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.
26. 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.
27. van Opijnen, T., Bodi, K.L., and Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods *6*, 767–772.
28. Christen, B., Abeliuk, E., Collier, J.M., Kalogeraki, V.S., Passarelli, B., Coller, J.A., Fero, M.J., McAdams, H.H., and Shapiro, L. (2011). The essential genome of a bacterium. Mol Syst Biol *7*, 528.
29. Christen, M., Beusch, C., Bösch, Y., Cerletti, D., Flores-Tinoco, C.E., Del Medico, L., Tschan, F., and Christen, B. (2016). Quantitative selection analysis of bacteriophage φCbK susceptibility in *Caulobacter crescentus*. J Mol Biol *428*, 419–430.
30. Harder, W., and Quayle, J.R. (1971). The biosynthesis of serine and glycine in *Pseudomonas* AM1 with special reference to growth on carbon sources other than C1 compounds. Biochem J *121*, 753–762.
31. Amaratunga, K., Goodwin, P.M., O'Connor, C.D., and Anthony, C. (1997). The methanol oxidation genes *mxaFJGIR*(*S*)*ACKLD* in *Methylobacterium extorquens*. FEMS Microbiol Lett *146*, 31–38.
32. Chistoserdova, L. (2016). Wide distribution of genes for tetrahydromethanopterin/ methanofuran-linked C1 transfer reactions argues for their presence in the common ancestor of bacteria and archaea. Front Microbiol *7*, 1–5.
33. Weiss, M.C., Sousa, F.L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., and Martin, W.F. (2016). The physiology and habitat of the last universal common ancestor. Nat

Microbiol *1*, 16116.

34. Meier, C., Carter, L.G., Winter, G., Owens, R.J., Stuart, D.I., and Esnouf, R.M. (2007). Structure of 5-formyltetrahydrofolate cyclo-ligase from *Bacillus anthracis* (BA4489). Acta Crystallogr. Sect F Struct Biol Cryst Commun *63*, 168–172.
35. Stover, P., and Schirchs, V. (1990). Serine hydroxymethyltransferase catalyzes the hydrolysis of 5,10-methenyltetrahydrofolate to 5-formyltetrahydrofolate. J Biol Chem *265*, 14227–14233.
36. Kalyuzhnaya, M.G., and 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*, 1229–1235.
37. Anthony, C. (1975). The microbial metabolism of C1 compounds. The cytochromes of *Pseudomonas* AM1. Biochem J *146*, 289–298.
38. Peyraud, R., Kiefer, P., Christen, P., Portais, J.C., and Vorholt, J.A. (2012). Co-consumption of methanol and succinate by *Methylobacterium extorquens* AM1. PLoS One *7*, e48271.
39. Morales, G., Ugidos, A., and Rojo, F. (2006). Inactivation of the *Pseudomonas putida* cytochrome *o* ubiquinol oxidase leads to a significant change in the transcriptome and to increased expression of the CIO and *cbb3*-1 terminal oxidases. Environ Microbiol *8*, 1764–1774.
40. Sun, J., Van Dommelen, A., Van Impe, J., and Vanderleyden, J. (2002). Involvement of *glnB*, *glnZ*, and *glnD* genes in the regulation of poly-3-hydroxybutyrate biosynthesis by ammonia in *Azospirillum brasilense* Sp7. Appl Environ Microbiol *68*, 985–988.
41. Tabita, F.R. (1980). Pyridine nucleotide control and subunit structure of phosphoribulokinase from photosynthetic bacteria. J Bacteriol *143*, 1275–1280.
42. Charlier, H.A., Runquist, J.A, and Miziorko, H.M. (1994). Evidence supporting catalytic roles for aspartate residues in phosphoribulokinase. Biochemistry *33*, 9343–50.
43. Donahue, J.L., Bownas, J.L., Niehaus, W.G., and Larson, T.J. (2000). Purification and characterization of *glpX*-encoded fructose-1,6-bisphosphatase, a new enzyme of the glycerol 3-phosphate regulon of *Escherichia coli*. J Bacteriol *182*, 5624–5627.
44. Breitling, R., Ritchie, S., Goodenowe, D., Stewart, M.L., and Barrett, M.P. (2006). *Ab initio* prediction of metabolic networks using Fourier transform mass spectrometry data. Metabolomics *2*, 155–164.
45. Dangel, A., and Tabita, F.R. (2015). CbbR, the master regulator for microbial carbon dioxide fixation. J Bacteriol *197*, 3488–3498.
46. Maddocks, S.E., and Oyston, P.C.F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology *154*, 3609–3623.
47. Kalyuzhnaya, M.G., and Lidstrom, M.E. (2005). QscR-mediated transcriptional activation of serine cycle genes in *Methylobacterium extorquens* AM1. J Bacteriol *187*, 7511–7517.
48. Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K., and Senda, T. (2003). Crystal

structure of a full-length LysR-type transcriptional regulator, CbnR: Unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. J Mol Biol *328*, 555–566.

49. Schneider, K., Skovran, E., and Vorholt, J.A. (2012). Oxalyl-coenzyme A reduction to glyoxylate is the preferred route of oxalate assimilation in *Methylobacterium extorquens* AM1. J Bacteriol *194*, 3144–3155.
50. Liang, W.-F., Cui, L.-Y., Cui, J.-Y., Yu, K.-W., Yang, S., Zhang, C., and Xing, X.-H. (2016). Biosensor-assisted transcriptional regulator engineering for *Methylobacterium extorquens* AM1 to improve mevalonate synthesis by increasing the acetyl-CoA supply. Metab Eng *39*, 1–10.
51. Rindt, K.-P., and Ohmann, E. (1969). NADH and AMP as allosteric effectors of ribulose-5-phosphate kinase in *Rhodopseudomonas spheroides*. Biochem Biophys Res Commun *36*, 357–364.
52. 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 α-humulene from methanol. Metab Eng *32*, 82–94.
53. Sonntag, F., Müller, J.E.N., Kiefer, P., Vorholt, J.A., Schrader, J., and Buchhaupt, M. (2015). 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*, 3407–3419.
54. Rohde, M.-T., Tischer, S., Harms, H., and Rohwerder, T. (2017). Production of 2-hydroxyisobutyric acid from methanol by *Methylobacterium extorquens* AM1 expressing (*R*)-3-hydroxybutyryl coenzyme A-isomerizing enzymes. Appl Environ Microbiol *83*, e02622-16.
55. Ledermann, R., Strebel, S., Kampik, C., and Fischer, H.-M. (2016). Versatile vectors for efficient mutagenesis of *Bradyrhizobium diazoefficiens* and other Alphaproteobacteria. Appl Environ Microbiol *82*, 2791-2799.
56. 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.
57. Kiefer, P., Schmitt, U., and Vorholt, J.A. (2013). eMZed: An open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. Bioinformatics *29*, 963–964.
58. Kiefer, P., Schmitt, U., Müller, J.E.N., Hartl, J., Meyer, F., Ryffel, F., and Vorholt, J.A. (2015). DynaMet: A fully automated pipeline for dynamic LC-MS data. Anal Chem *87*, 9679–9686.
59. Vallenet, D., Engelen, S., Mornico, D., Cruveiller, S., Fleury, L., Lajus, A., Rouy, Z., Roche, D., Salvignol, G., Scarpelli, C., *et al.* (2009). MicroScope: A platform for microbial genome

annotation and comparative genomics. Database *2009*, 1–12.

60. Sonnhammer, E.L.L., and Östlund, G. (2015). InParanoid 8: Orthology analysis between 273 proteomes, mostly eukaryotic. Nucleic Acids Res *43*, D234–D239.
61. Marmiesse, L., Peyraud, R., and Cottret, L. (2015). FlexFlux: combining metabolic flux and regulatory network analyses. BMC Syst Biol. *9*, 93.
62. Rocha, I., Maia, P., Evangelista, P., Vilaça, P., Soares, S., Pinto, J.P., Nielsen, J., Patil, K.R., Ferreira, E.C., and Rocha, M. (2010). OptFlux: an open-source software platform for *in silico* metabolic engineering. BMC Syst Biol *4*, 45.
63. 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.
64. Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., *et al.* (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.
65. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.
66. Hartl, J., Kiefer, P., Meyer, F., and Vorholt, J.A. (2017). Longevity of major coenzymes allows minimal *de novo* synthesis in microorganisms. Nat Microbiol *2*, 17073.
67. Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., and Genin, S. (2016). A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. PLoS Pathog *12*, 1–25.
68. Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., *et al.* (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res *44*, D447–D456.

## 4.7 Supplementary material

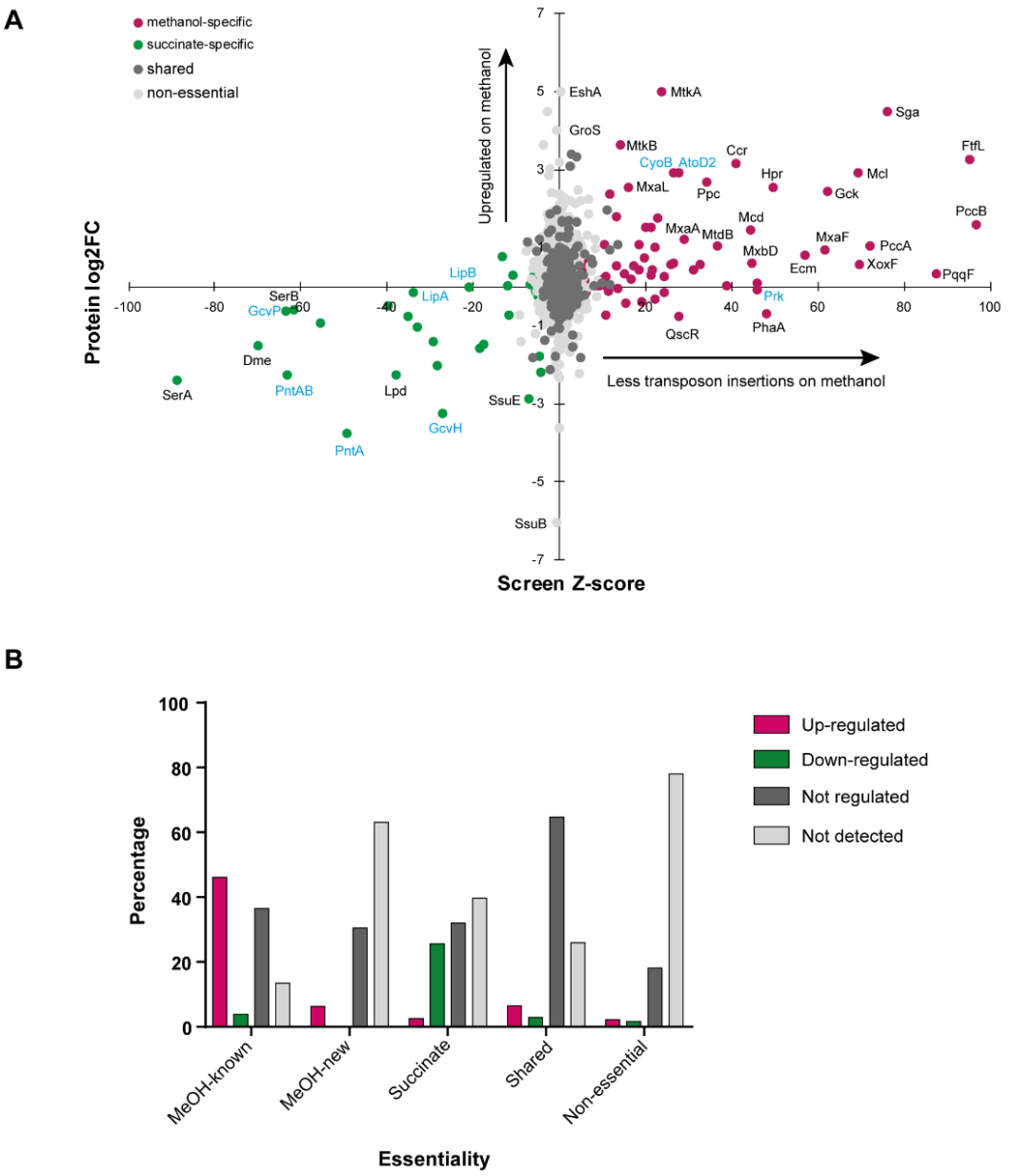


**Figure S1. Correlation of gene essentiality with protein abundance changes**

A) All genes encoding for proteins that were detected in the proteomes (with at least two unique peptides) are plotted with their Z-score along the x-axis and the log2 fold-change in protein level (no p-value cutoff was used) along the y-axis. The color of the dot represents the assigned class: non-essential under both conditions (light gray), required under both conditions (dark gray), required on succinate only (green), and required on methanol only (pink). Known "methylotrophy" proteins are shown in black, while some newly identified "methylotrophy" proteins (encoded by new methylotrophy genes) are shown in blue. B) Percentage of proteins that are up-regulated, down-regulated, not regulated, or not detected in each category of gene requirement. The genes that are specifically required on methanol are divided into known and new methylotrophy genes.

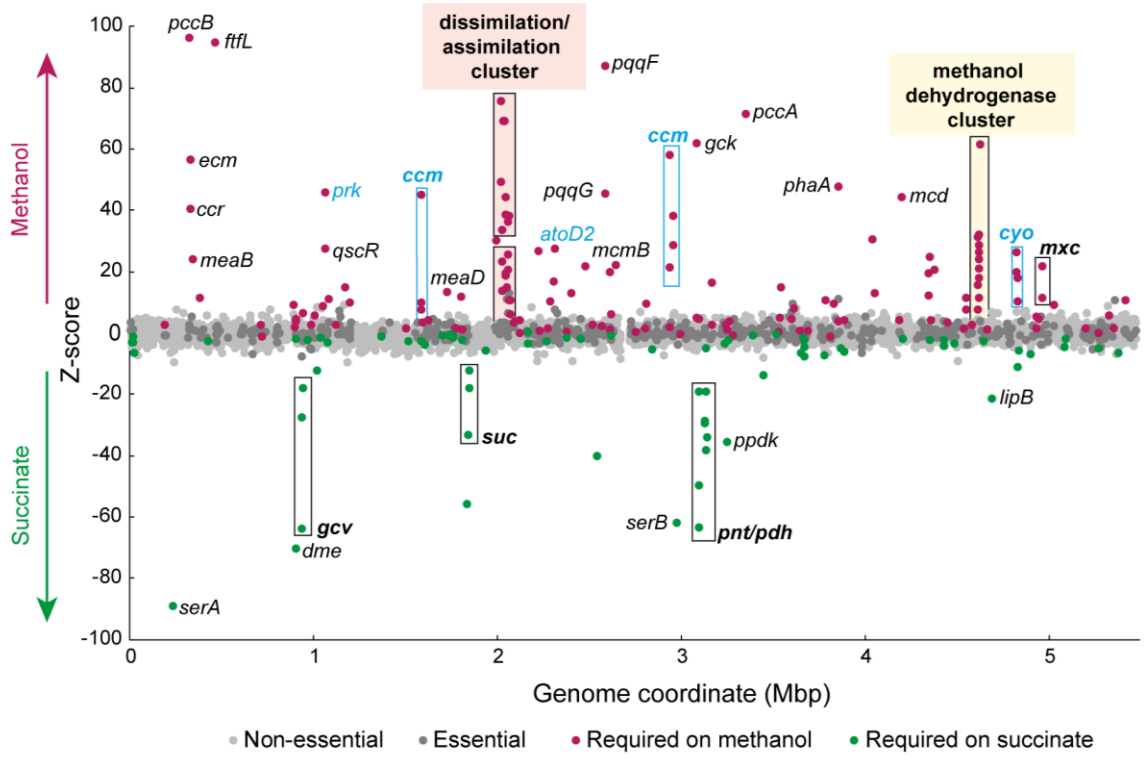


**Figure S2. Global essentiality data analysis**

The Z-score (see Materials and methods), a measure to compare insertion densities between the two conditions, is plotted for every gene in the genome. Genes that have fewer insertions on methanol than on succinate have a positive value, whereas genes that have more insertions on succinate than methanol have a negative value. The color of the dot represents the assigned class: non-essential under both conditions (light gray), required under both conditions (dark gray), only on succinate (green) or only on methanol (pink). Known genes and clusters (marked with boxes) are shown in black, while newly identified genes and clusters are shown in blue. All labelled genes are listed with full names in Data set S1.

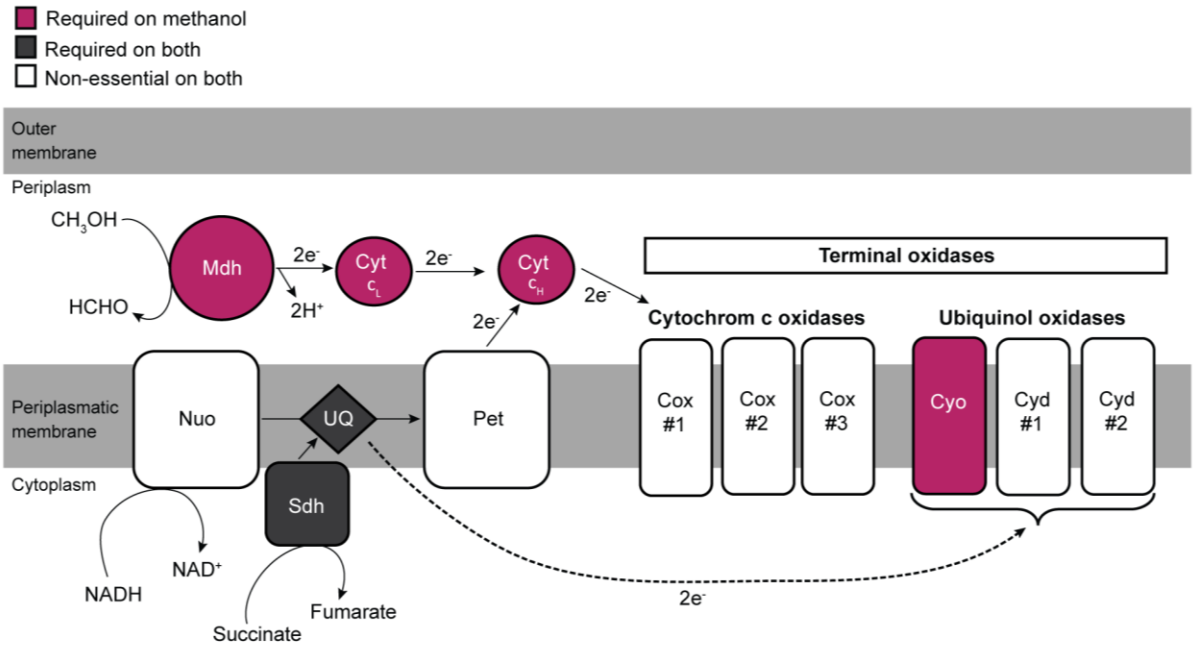


**Figure S3. Schematic representation of the respiratory chain in *M. extorquens* PA1**

Parts are colored according to their essentiality; required on methanol (pink), required on both (black), non-essential on both (white). The following abbreviations were used; Mdh = methanol dehydrogenase (Mext_4137-50), Cyt cL = methylotrophy-specific cytochrome *c* (Mext_4148), CytcH = general cytochrome *c* (Mext_0924), Nuo = NADH dehydrogenase (Mext_ 1072-085+101), UQ = ubiquinol pool, Sdh = succinate dehydrogenase (Mext_3600-1), Pet = cytochrome *bc1* complex (Mext_2542-4), Cox = cytochrome *c* dependent oxidases (#1: Mext_0094-6, #2: Mext_0222-5, #3: Mext_3258+62-3), Cyo = cytochrome *o* dependent ubiquinol oxidase (Mext_4334-7), Cyd = cytochrome *d* dependent ubiquinol oxidases (#1: Mext_1360-1, #2: Mext_3833-4).

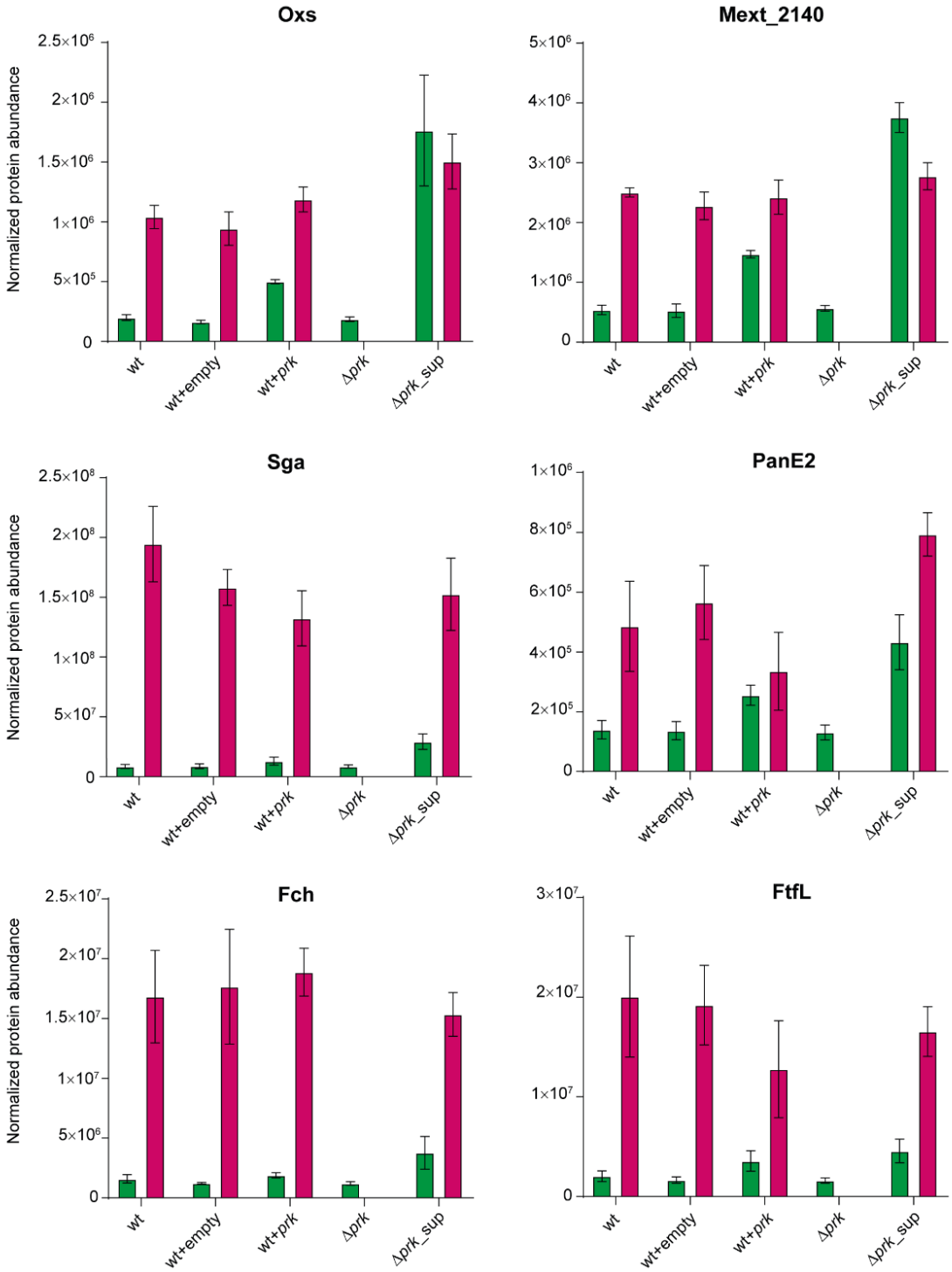


**Figure S4. Abundance of six QscR-regulated proteins**

Protein abundance of the top six QscR-regulated proteins (highest fold-change, $p<0.01$). Oxs, oxalyl-CoA synthetase, Mext_2140, neighboring gene of *oxs* and *panE2* annotated as "AMP-dependent synthetase", Sga, serine glyoxylate aminotransferase, PanE2, oxalyl-CoA reductase, Fch, methenyl-$H_4F$ cyclohydrolase, FtfL, formyl-$H_4F$ ligase grown on succinate (green) and methanol (pink) in different strain backgrounds; wildtype (wt), wt with empty pTE105 plasmid (wt+empty), wt with pTE105 containing *prk* (wt+*prk*), *prk* knockout (Δ*prk*), and the reconstructed Δ*prk* suppressor strain that carries

a mutation in *qscR* (Δ*prk*_sup). Δ*prk* was only measured on succinate since it shows no growth on methanol. Mean ± SD is shown (n = 4).

**Table S1. Growth rates of knockouts and complementation strains**

Growth rates and doubling times were determined for both orientations of the inserted gentamicin resistance cassette (GmR). Mean ±SE (n = 4) is given. Pden = *P. denitrificans*.

| **Strain** | **Growth rate ($h^{-1}$)** | **Doubling time (h)** |
|---|---|---|
| | **Methanol** | |
| PA1wt | 0.2081±0.0009 | 3.31±0.01 |
| Δ*prk* | nm*<br>nm** | nm*<br>nm** |
| Δ*prk*+empty | nm* | nm* |
| Δ*prk*+*prk*_PA1 | 0.1960±0.0008* | 3.52±0.01* |
| Δ*prk*+*prk*_Pden | 0.2070±0.0006* | 3.34±0.01* |
| Δ*prk*+*prk*_PA1_D42A | nm* | nm* |
| Δ*prk* _sup (artificial) | 0.1365±0.0006* | 5.08±0.02* |
| Δ*prk*_sup (natural) | 0.1336±0.0005* | 5.18±0.02* |
| Δ*cyoA* | 0.166±0.001*<br>0.144±0.002** | 4.14±0.03*<br>4.77±0.06** |
| ΔMext_1769 | 0.1662±0.0006*<br>0.1485±0.0007** | 4.23±0.01*<br>4.64±0.02** |
| ΔMext_1983 | 0.0659±0.0005*<br>0.0621±0.0002** | 10.44±0.08*<br>11.12±0.04** |
| Δ*5fcl* | 0.0876±0.0002* | 7.89±0.02* |
| Δ*atoD2* | 0.0910±0.0006*<br>0.0862±0.0006** | 7.56±0.05*<br>8.04±0.06* |
| Δ*ccmAB* | nm*<br>nm** | nm*<br>nm** |
| | **Succinate** | |
| PA1wt | 0.251±0.001 | 2.76±0.01 |
| Δ*prk* | 0.254±0.001*<br>0.2512±0.0007** | 2.73±0.01*<br>2.759±0.008** |

*GmR-fwd, **GmR-rev, nm = not measurable

**Table S2. New methylotrophy genes identified in the screen**

Annotation according to [59] with some adjustments. The Z-score for methanol is shown (see Materials and methods).

| Category | | | Length (bp) | Requirement methanol | Z score |
|---|---|---|---|---|---|
| **Carrier** | | | | | |
| Mext_0924 | *cycA* | cytochrome c class I | 362 | essential | 5.93 |
| Mext_3277 | *grxC* | glutaredoxin 3 | 257 | essential | 0.78 |
| **Enzyme** | | | | | |
| Mext_0335 | *glnD* | PII uridylyl-transferase | 2786 | fitness-z | 11.51 |
| Mext_0961 | | alcohol dehydrogenase | 980 | fitness-cost | 8.81 |
| Mext_0980 | *prk* | phosphoribulokinase | 860 | essential | n.d. |
| Mext_1023 | | peptidase M23B | 2093 | fitness-cost | 2.72 |
| Mext_1405 | *cycL* | cytochrome *c* biogenesis protein | 461 | essential | 7.87 |
| Mext_1406 | *cycK* | cytochrome *c* assembly protein | 2000 | fitness-cost | 45.34 |
| Mext_1585 | *parA* | cobyrinic acid ac-diamide synthase | 860 | fitness-cost | 1.54 |
| Mext_2042 | *hss* | homospermidine synthase | 1448 | fitness-z | 10.48 |
| Mext_2067 | *ppk* | polyphosphate kinase | 2366 | fitness-cost | 16.99 |
| Mext_2139 | *panE* | 2-dehydropantoate 2-reductase | 1007 | fitness-cost | 13.04 |
| Mext_2614 | *ccmG* | periplasmic protein thiol-disulfide oxidoreductase | 620 | fitness-cost | 21.41 |
| Mext_2822 | *ntrY* | PAS sensor protein | 2399 | fitness-cost | 2.93 |
| Mext_3181 | *ccrfM* | DNA methylase N-4/N-6 domain-containing protein | 1196 | fitness-cost | 5.33 |
| Mext_3651 | | leucyl aminopeptidase | 1493 | fitness-cost | 30.90 |
| Mext_3906 | *pabA* | glutamine amidotransferase of anthranilate synthase | 629 | essential | 19.49 |
| Mext_3907 | | paraaminobenzoate synthase subunit I | 1418 | essential | 25.13 |
| Mext_4335 | *cyoC* | cyt *o* ubiquinol oxidase subunit III | 647 | fitness-z | 19.88 |
| Mext_4336 | *cyoB* | cyt *o* ubiquinol oxidase subunit I | 2003 | fitness-cost | 26.43 |
| Mext_4337 | *cyoA* | ubiquinol oxidase subunit II | 1163 | fitness-z | 18.29 |
| Mext_4641 | *mraW* | S-adenosyl-methyltransferase | 1055 | fitness-cost | 1.59 |
| Mext_4768 | *ruvB* | Holliday junction DNA helicase | 1070 | essential | 5.84 |
| Mext_0903 | | SirA family protein | 284 | essential | 2.70 |
| Mext_2034 | | NAD-dependent epimerase/dehydratase | 923 | fitness-cost | 1.66 |
| Mext_2071 | *atoD2* | coenzyme A transferase | 1982 | fitness-cost | 27.53 |
| Mext_2363 | | 5-formyltetrahydrofolate cyclo-ligase | 590 | essential | 6.28 |
| Mext_2762 | *recJ* | single-stranded-DNA-specific exonuclease | 1835 | fitness-cost | 4.74 |
| Mext_3319 | | Holliday junction resolvase YqgF | 482 | fitness-cost | 0.77 |
| Mext_3662 | | 1A family penicillin-binding protein | 2501 | fitness-z | 13.21 |
| Mext_3905 | | class IV aminotransferase | 857 | essential | 12.22 |
| Mext_4004 | | glycosyl transferase family protein | 983 | fitness-cost | 3.57 |
| Mext_4430 | | nucleotidyl transferase | 746 | fitness-cost | 1.73 |

| | | | | | |
|---|---|---|---|---|---|
| Mext_4440 | *addA* | double-strand break repair helicase | 3443 | fitness-cost | 4.68 |
| Mext_4502 | | bifunctional transaldolase/ phosoglucose isomerase | 2825 | fitness-cost | 9.19 |
| Mext_4787 | | FAD linked oxidase domain-containing protein | 1190 | fitness-cost | 1.69 |
| **Factor** | | | | | |
| Mext_1001 | *rpoH* | RNA polymerase factor sigma-32 | 893 | fitness-z | 11.33 |
| Mext_1407 | *cycJ* | CcmE/CycJ protein | 533 | essential | 10.20 |
| Mext_1408 | *cycH* | cytochrome c-type biogenesis protein CcmI | 1658 | fitness-cost | 3.57 |
| Mext_2501 | *rpoN* | RNA polymerase sigma-54 factor | 1535 | fitness-z | 9.78 |
| Mext_1607 | | cytochrome oxidase assembly | 1148 | fitness-cost | 12.16 |
| Mext_3427 | | hypothetical protein | 263 | essential | -0.92 |
| **Unknown** | | | | | |
| Mext_1865 | | zinc-binding CMP/dCMP deaminase | 1466 | fitness-cost | 3.78 |
| Mext_0173 | | hypothetical protein | 1544 | fitness-cost | 2.72 |
| Mext_0650 | | hypothetical protein | 389 | essential | 2.80 |
| Mext_0818 | | hypothetical protein | 305 | essential | 2.12 |
| Mext_0821 | | hypothetical protein | 1052 | essential | 4.87 |
| Mext_0822 | | flagellar export protein FliJ | 404 | essential | 3.91 |
| Mext_0858 | | NLP/P60 protein | 860 | fitness-cost | 6.83 |
| Mext_1347 | | hypothetical protein | 326 | fitness-cost | 1.74 |
| Mext_1613 | | hypothetical protein | 239 | essential | 1.22 |
| Mext_1769 | | hypothetical protein | 521 | essential | 30.54 |
| Mext_1870 | | hypothetical protein | 581 | fitness-cost | 3.88 |
| Mext_1897 | | hypothetical protein | 197 | essential | 0.13 |
| Mext_1983 | | hypothetical protein | 524 | essential | 26.93 |
| Mext_1997 | | hypothetical protein | 236 | essential | 0.95 |
| Mext_2115 | | hypothetical protein | 266 | essential | 0.57 |
| Mext_2311 | | HpcH/HpaI aldolase | 662 | fitness-cost | 2.45 |
| Mext_2446 | | hypothetical protein | 491 | essential | 0.48 |
| Mext_2498 | | hypothetical protein | 758 | fitness-cost | 1.81 |
| Mext_2673 | | hypothetical protein | 326 | essential | -0.25 |
| Mext_2746 | | hypothetical protein | 542 | essential | 5.19 |
| Mext_2890 | | hypothetical protein | 257 | essential | 3.02 |
| Mext_2898 | | hypothetical protein | 188 | essential | 1.52 |
| Mext_2911 | | hypothetical protein | 218 | essential | 3.88 |
| Mext_2935 | | hypothetical protein | 188 | essential | -0.59 |
| Mext_3178 | | hypothetical protein | 476 | fitness-cost | 0.38 |
| Mext_3240 | | hypothetical protein | 164 | essential | 4.61 |
| Mext_3248 | | hypothetical protein | 944 | fitness-cost | 8.08 |
| Mext_3413 | | hypothetical protein | 692 | essential | 10.83 |
| Mext_3503 | | hypothetical protein | 221 | fitness-cost | 4.27 |
| Mext_3771 | | hypothetical protein | 281 | essential | 4.27 |
| Mext_3915 | | hypothetical protein | 560 | fitness-cost | 4.49 |

| | | | | | |
|---|---|---|---|---|---|
| Mext_4073 | | hypothetical protein | 566 | fitness-cost | 1.73 |
| Mext_4086 | | hypothetical protein | 284 | essential | 7.99 |
| Mext_4115 | | hypothetical protein | 332 | fitness-cost | 2.70 |
| Mext_4189 | | hypothetical protein | 200 | essential | 1.24 |
| Mext_4431 | | hypothetical protein | 1649 | fitness-cost | 5.60 |
| **Membrane component** | | | | | |
| Mext_4843 | *omp* | peptidoglycan-associated lipoprotein | 512 | fitness-z | 10.73 |
| Mext_0816 | | capsule polysaccharide export protein-like protein | 1244 | fitness-cost | 9.51 |
| Mext_2613 | *ccmD* | heme exporter protein | 152 | essential | 1.91 |
| **Regulator** | | | | | |
| Mext_1893 | *cspE* | DNA-binding cold-shock protein | 209 | essential | 4.47 |
| Mext_2820 | *glnG* | nitrogen regulation protein NR(I) | 1472 | fitness-z | 16.42 |
| Mext_1093 | | LysR substrate-binding protein | 890 | fitness-cost | 10.10 |
| **Transporter** | | | | | |
| Mext_2612 | *ccmC* | heme exporter protein | 806 | essential | 58.30 |
| Mext_2634 | *ccmB* | heme exporter protein | 665 | essential | 38.45 |
| Mext_2635 | *ccmA* | heme exporter protein | 647 | essential | 28.91 |
| Mext_4338 | | putative transmembrane transporter protein | 1418 | fitness-z | 10.50 |
| Mext_4714 | *pstC* | phosphate ABC transporter permease | 974 | essential | 0.14 |
| Mext_1443 | | cobalt transporter subunit CbtA | 767 | essential | 4.47 |
| Mext_2356 | | lipid ABC transporter ATPase/inner membrane protein | 1856 | essential | 19.92 |
| Mext_3186 | | porin | 1661 | fitness-z | 15.02 |
| Mext_4085 | | capsule polysaccharide export protein-like protein | 1346 | fitness-cost | 11.83 |
| **Structure** | | | | | |
| Mext_0653 | *rpsP* | 30S ribosomal protein S16 | 362 | essential | -0.87 |
| Mext_2355 | *rpmE* | 50S ribosomal protein L31 | 254 | fitness-cost | 1.51 |

**Table S3. Primers used in this study**

Restriction sites are underlined and ribosomal binding sites (RBS) are shown in bold.

| **Knockout construction** | |
|---|---|
| **Δ*prk*** | |
| Δ*prk*_HR1_fwd_NsiI | GTTCAGATGCATGCTGCCGGCTTCGTC |
| Δ*prk*_HR1_rev_SpeI | GCTAAGACTAGTGGGCCGAGTGCTTGG |
| Δ*prk*_HR2_fwd_SpeI | GGAATGACTAGTTCGTCCGGCTCCCTC |
| Δ*prk*_HR2_rev_MunI | TCAGTACAATTGCCGACGGCGTGTGC |
| Δ*prk*_check_fwd | GCATCCGCCCAGTGACG |
| Δ*prk*_check_rev | ACCGGGAGCCCTACTACC |
| **Δ*cyoA*** | |
| Δ*cyoA*_HR1_fwd_NsiI | GTTCAGATGCATGCCCGTGACGTAGTCTTTG |
| Δ*cyoA*_HR1_rev_SpeI | GCTAAGACTAGTTCTGAGCAGCATCATGG |
| Δ*cyoA*_HR2_fwd_SpeI | GGAAGCTGACTAGTGTTGAAACTCGTCCGTAAATCC |
| Δ*cyoA*_HR2_rev_MunI | TCAGTACAATTGCGCCCTCGATTGGAAATGG |
| Δ*cyoA*_check_fwd | CCAGCCAGAACGGGATCG |
| Δ*cyoA*_check_rev | ATCGAGCCCGGCAAGC |
| **ΔMext_1769** | |
| ΔMext_1769_HR1_fwd_NsiI | GTTCAGATGCATGTCGCGCAGGATCAGGTTC |
| ΔMext_1769_HR1_rev_SpeI | GCTAAGACTAGTTCCGCTCTCGCAAAGCACTC |
| ΔMext_1769_HR2_fwd_SpeI | GGACTGACTAGTGACCGGCAGCGTCCC |
| ΔMext_1769_HR2_rev_MunI | TCAGTACAATTGACCGCGCCGGAATTCG |
| ΔMext_1769_check_fwd | TTCTGCGCCGCCATGTC |
| ΔMext_1769_check_rev | CGTGGCTTGACGAAACG |
| **ΔMext_1983** | |
| ΔMext_1983_HR1_fwd_NsiI | GTTCAGATGCATACATCGCGGTAGCCGTC |
| ΔMext_1983_HR1_rev_SpeI | GCTAAGACTAGTGATCGCGGCTTTATCCTGCCTG |
| ΔMext_1983_HR2_fwd_SpeI | GGACTGACTAGTGCCGGCGCCTTCTC |
| ΔMext_1983_HR2_rev_MunI | TCAGTACAATTGGTTTGGGCAGGTTGGCATC |
| ΔMext_1983_check_fwd | GGCCCTCACACCTTCAGC |
| ΔMext_1983_check_rev | CCTACGGCACGGAAACCATC |
| **Δ*5fcl*** | |
| Δ*5fcl*_HR1_fwd_NsiI | GTTCAGATGCATGGCGCACGGGCTC |
| Δ*5fcl*_HR1_rev_SpeI | GCTAAGACTAGTACTCGCATCGGCTCTTTC |
| Δ*5fcl*_HR2_fwd_SpeI | GGAATGACTAGTGAGCCGTTGGAGTGTCG |
| Δ*5fcl*_HR2_rev_MunI | TCAGTACAATTGCACCGGCGAGGTCAAC |
| Δ*5fcl*_check_fwd | CGCCATGCCGCACAAC |
| Δ*5fcl*_check_rev | CGCTGCCGAACTCGATGC |
| **Δ*atoD2*** | |

| | |
|---|---|
| Δ*atoD2*_HR1_fwd_NsiI | GTTCAGATGCATAGCGCGTTGTAGATGACGTG |
| Δ*atoD2*_HR1_rev_SpeI | GCTAAGACTAGTCCGCCGTCCTCCACTTTC |
| Δ*atoD2*_HR2_fwd_SpeI | GGACTGACTAGTGGATCGCGAAAGGCTTCCG |
| Δ*atoD2*_HR2_rev_MunI | TCAGTACAATTGCCCAATCGGGCTTCGACTG |
| Δ*atoD2*_check_fwd | CGGGTAGGAGACCCAGTAGG |
| Δ*atoD2*_check_rev | CGGTGAAGGCGCGAAGC |
| **Δ*ccmAB*** | |
| Δ*ccmAB*_HR1_fwd_NsiI | GTTAGGATGCATATCATTGCCACCTCGTCTAC |
| Δ*ccmAB*_HR1_rev_SpeI | GGAATGACTAGTTTCGCCGGATCTACTTATCG |
| Δ*ccmAB*_HR2_fwd_SpeI | TTGGAAACTAGTGGCCTTCGATGAGGTTG |
| Δ*ccmAB*_HR2_rev_MunI | GCTGAACAATTGTGGCGAGCTTTGAAGCTG |
| Δ*ccmAB*_check_fwd | GGGCTCACGCAGTCGATG |
| Δ*ccmAB*_check_rev | GCCTCGGGAATGGAGTAGTAAGTG |
| **Gentamicin (GmR)-specific primers for confirmation of knockouts** | |
| GmR_1 | GCCTTGATGTTACCCGAGAG |
| GmR_2 | GCGAAGTAATCGCAACATCC |
| **Cloning of P*sga* into pTE100_mChe** | |
| P*sga*_fwd_EcoRI | TTAAGAATTCCGCTTTGGGCGGTCTCC |
| P*sga*_rev_XbaI | GGCCTCTAGATCGCCTCCAAAGGACCGAAC |
| **Cloning of *prk* into pTE105** | |
| *prk*_PA1_fwd_RBS_SpeI | TAGCACTAGT**AAGGAGA**TCGAGCATATGTCGGCGCGTCATCCC |
| *prk*_PA1_rev_KpnI | ATTCGGTACCAGCCGGACGAGGCTCAC |
| *prk*_Pden_fwd_RBS_SpeI | TCGCACTAGT**AAGGAGA**GGACAACCATGAGCAAGAAAC |
| *prk*_Pden_rev_KpnI | ATTCGGTACCTCAGGCGCGTTTCGATTC |
| **Mutation of *prk*** | |
| *prk*_PA1_D42A_fwd | CTACATCGAGGGCGCCGGCTTCCATGCG |
| *prk*_PA1_D42A_rev | CGCATGGAAGCCGGCGCCCTCGATGTAG |

Supplemental Data S1-S4 can be found online at http://dx.doi.org/10.1016/j.cub.2017.07.025

## Chapter 5:

# Discussion and outlook

Microorganisms able to grow at the expense of reduced one-carbon substrates have been first observed more than a century ago (Söhngen, 1906) and initial elucidation of their metabolism followed in the 1960s. The first solution to methylotrophic growth discovered was a form of chemoautotrophic metabolism based on the CBB cycle (Quayle & Keech, 1959, 1963), which was discovered a few years earlier by Calvin and colleagues (Bassham *et al*, 1953) and was therefore not a *per se* new assimilation metabolism. The isolation of *Pseudomonas* AM1 (later renamed to *M. extorquens* AM1) paved the way to the discovery of the first truly methylotrophic pathway for assimilation of one-carbon (Peel & Quayle, 1961). Initial elucidation of its metabolism mainly relied on elegant $^{14}C$-incorporation experiments to identify the metabolites involved in the metabolism of methanol (Large *et al*, 1962a). Subsequent biochemical analysis revealed the enzymes catalyzing the individual steps (Large *et al*, 1962b; Large & Quayle, 1963; Anthony, 1970). Classical genetic approaches led to the discovery of gene regions of interest, which were further dissected using complementation strategies (Nunn & Lidstrom, 1986). With the sequencing of the complete genome of *M. extorquens* AM1 (Chistoserdova *et al*, 2003; Vuilleumier *et al*, 2009) genome-wide approaches, such as transcriptomics and proteomics became possible (Bosch *et al*, 2008; Laukel *et al*, 2004; Okubo *et al*, 2007). These studies confirmed the differential expression of many known genes and revealed various so-far unknown components. Furthermore, targeted inactivation of genes encoding potentially interesting functions or co-localized with known methylotrophy genes revealed a few dozen genes specifically required for growth on methanol (Chistoserdova *et al*, 2003). This extensively studied methylotroph only represents one of many metabolic solutions to growth on reduced one-carbon units (Chistoserdova, 2011). One of the less studied methylotrophs, which has sparked interest as a biotechnological production strain, is the Gram-positive *B. methanolicus* (Brautaset *et al*, 2007). In terms of methylotrophic pathways, it represents a complementary model to *M. extorquens*. Like all known Gram-positive methylotrophs, it lacks the well-studied PQQ-dependent Mdh and instead relies on an enzyme using NAD as a cofactor (Arfman & Dijkhuizen, 1990).

The goal of this thesis was to improve the understanding of the important initial step of methylotrophy (chapters 2&3) as well as to generate new insights into methylotrophy as a whole by defining the methylotrophy genome of *M. extorquens* PA1 (chapter 4).

The initial reaction that has to be fulfilled by any methylotroph growing on methanol is the oxidation of methanol to formaldehyde. In the case of the Gram-positive *B. methanolicus*, this reaction is catalyzed by an NAD-dependent Mdh, whose *in vitro* activity is stimulated by an endogenous activator protein. The activation of NAD-Mdhs by the so-called Act is poorly understood. In chapter 2, it was shown that the *in vitro* activation is not limited to the partners NAD-Mdh and Act. All investigated alcohol dehydrogenases belonging to the type III family could be activated by Act. Furthermore, activation could be achieved by several NUDIX hydrolases, likely depending on their ability to hydrolyze $NAD^+$ *in vitro*. The *in vivo* role of the activation by Act is currently unknown. The

hydrolysis of free $NAD^+$ in the cell would most likely be detrimental for the cell due to depletion of NAD(H) pools. In contrast, the cleavage of Mdh-bound $NAD^+$ by Act as suggested by Kloosterman *et al* (2002) would be a more plausible scenario *in vivo*. One would expect that the hydrolysis of an $NAD^+$ molecule directly in the active site of a second enzyme would require very specific protein-protein interaction. Our results, however, show that the activation is not specific, but possible with interchangeable partners, making specific protein-protein interactions unlikely.

The *in vivo* essentiality of Act for methylotrophy in *B. methanolicus* can currently not be studied due to lack of genetic tools. A recent proteome study in *B. methanolicus* MGA3 showed that Act is not upregulated on methanol (Müller *et al*, 2014). Notably, the same study only observed slight upregulation of methanol dehydrogenase itself, which is in line with only weak regulation of Mdh gene expression in other methylotrophs including *M. extorquens* (Okubo *et al*, 2007). An alternative way to study the activation *in vivo* is presented by a synthetic biology approach. We used the Gram-negative model methylotroph *M. extorquens* to probe exchangeability of its PQQ-Mdh with different NAD-Mdhs from two *B. methanolicus* strains. This setup presents a unique possibility to investigate the *in vivo* function of NAD-Mdh (and its activator Act) in an organism where the downstream pathways for methanol metabolism are already in place.

As the growth of a strain lacking PQQ-Mdh could not be complemented in either the absence or presence of Act, no conclusions could be drawn from these experiments. The most likely reason was low soluble production of NAD-Mdhs, which was in some cases even further decreased when *act* was co-expressed. More fundamental causes for lack of growth could be the accumulation of intracellular formaldehyde, which is otherwise only produced in the periplasm, or a redox imbalance. Coping with the latter would require adaptations such as the upregulation of transhydrogenase (interconverting NADH and NADPH), which is strongly downregulated on methanol compared to succinate (Bosch *et al*, 2008) or adjustments in the expression of respiratory chain complexes. Possible future directions could be expression optimization using a titratable promoter or by testing other NAD-dependent methanol or alcohol dehydrogenases that might prove more suitable for expression in *M. extorquens*. Once a slow-growing *M. extorquens* employing a NAD-dependent Mdh is obtained, laboratory evolution could be used to optimize the strain as well as the enzyme. The study of the mutations accumulating in the genome of *M. extorquens* could give valuable insights into necessary adaptations. Furthermore, an NAD-Mdh optimized for a mesophilic bacterium as well as the knowledge of Act-requirement *in vivo* would be valuable for current endeavors to transfer methylotrophy into platform organisms such as *E. coli* (Müller *et al*, 2015a) or *Corynebacterium glutanicum* (Leßmeier *et al*, 2015). In spite of their inferior catalytic properties, all current synthetic methylotrophy approaches rely on NAD-dependent Mdhs and not on the efficient, but more complex PQQ-Mdhs.

Functional production of the model PQQ-Mdh MxaFI requires more than 10 proteins. The recently discovered REE-dependent PQQ-Mdh XoxF seems to be simpler, only depending on three genes (Pol

*et al*, 2014). The genomes of many model methylotrophs encode for both enzyme systems (Chistoserdova, 2011). In chapter 3, we investigated the lanthanum ($La^{3+}$)-dependent regulation of methanol oxidation in *M. extorquens* PA1. In line with recently published results for strain AM1 (Vu *et al*, 2016), the promoter activity of P*xox* is increased in the presence of $La^{3+}$, while the activity of P*mxa* is strongly decreased. Furthermore, XoxF is able to sustain wildtype-like growth rates of a PA1 strain lacking the classical MxaFI on methanol if $La^{3+}$ is present. The dominant switch from MxaFI to XoxF when $La^{3+}$ is added indicates that XoxF is not just an auxiliary enzyme, but is the preferred Mdh under growth conditions where REEs are available. Reasons could be the simpler structure of the system and the superior catalytic parameters including improved substrate affinity (Keltjens *et al*, 2014). The latter might give a crucial advantage under environmental conditions where methanol is likely limiting. This is supported by an increasing number of studies including meta-proteomics and -genomics highlighting the relevance of XoxF in the environment (Delmotte *et al*, 2009; Chistoserdova, 2016). In addition to their prevalence, XoxF proteins also show a much higher phylogenetic diversity compared to MxaF, which only corresponds to one branch in a phylogenetic tree of PQQ-Mdhs (Chistoserdova, 2016). Taken together this suggests that XoxF and not MxaFI is the ancestral enzyme. Importantly, all known MxaFI-methylotrophs also encode for at least one XoxF, while the number of known methylotrophs exclusively encoding for XoxF is constantly increasing (Chistoserdova, 2011). Methylotrophs that only depend on XoxF might be limited to environments where sufficient REEs are present, while possessing both MxaFI and XoxF could lead to more flexibility. Titratability of the *xoxF* promoter of *M. extorquens* PA1 at intermediate $La^{3+}$ concentrations suggests that the expression of the two systems is tuned in response to REE availability. Several components were implicated in the underlying regulatory cascade including XoxF itself as the REE-sensor (Vu *et al*, 2016). Our results show that $La^{3+}$ still provokes a response in cells lacking *xoxF* suggesting the presence of at least one additional $La^{3+}$-sensor. Further research is required to elucidate the regulatory cascade and to understand the role of XoxF in a larger context, but accumulating evidence points to a more important role of XoxF than previously anticipated.

In spite of their obvious differences, the two divergent solutions for methanol oxidation in Gram-negative and Gram-positive methylotrophs are able to sustain fast growth rates. It is currently unknown to what extent the organism has to adjust to accommodate the enzymes and deal with potential inherit difficulties.

Mdh catalyzes the first reaction in a long cascade needed for growth on methanol as sole carbon source. Many involved components have been identified by targeted approaches in the last few decades, indicating a complex metabolic and regulatory network. The entirety of required components, however, had hitherto not been determined for any methylotroph. In chapter 4, we used transposon sequencing to define the whole gene set required for growth of *M. extorquens* PA1 on methanol. This revealed almost 100 new genes that are either essential or fitness-relevant for growth on methanol, but

not the multi-carbon substrate succinate. The identification of such a high number of novel methylotrophy genes was surprising in the light of over 60 years of research. Among the newly identified methylotrophy genes was the gene encoding phosphoribulokinase (Prk). This finding was surprising because Prk, together with ribulose-bisphosphate carboxylase/oxygenase (RuBisCO), are commonly known as the specific reactions of the CBB cycle for carbon dioxide fixation. The presence of the *prk* gene in a bacterium lacking RuBisCO, could be attributed to evolution from a common ancestor of autotrophs and methylotrophs. Dedicated experiments showed a connection between Prk and the serine cycle regulator QscR. Our data suggests that DNA-binding of QscR, similarly to the related CBB-regulator CbbR (Dangel & Tabita, 2015), is activated by ribulose-1,5-bisphosphate (RuBP), the product of the reaction catalyzed by Prk. In *M. extorquens*, the poolsize of RuBP was almost 400-fold higher in cells growing on methanol compared to succinate, suggesting that Prk activity is highly regulated. Since this does not seem to happen on protein abundance level, allosteric regulation of its activity is likely. The activity of many bacterial Prks was shown to be activated by NADH and inhibited by AMP *in vitro* (Tabita, 1980). Sensing of NADH and AMP levels through Prk might represent a way to regulate costly expression of one-carbon assimilation genes in response to the energy state of the cell. Further investigation is required to determine the signals regulating Prk activity in *M. extorquens*. Tuning the activity of Prk in the cell would be both interesting for engineering purposes and for further investigation of the regulation *in vivo*. One approach could be the manipulation of intracellular NADH and/or AMP levels, for example by addition of different substrates or via overexpression of certain enzymes. Alternatively, Prk itself could either be modified by site-directed amino acid changes or be replaced with a differently regulated cyanobacterial Prk. Furthermore, the Δ*prk* suppressor mutants that partially activate expression of one-carbon assimilation genes even in the absence of methanol represent valuable models to study the impact of a deregulation. Growth under laboratory conditions on an excess of succinate was wildtype-like, but the lack of regulation could prove disadvantageous under other single or mixed substrate conditions or during a substrate switch. These experiments could provide important hints towards the elucidation of this unusual regulatory mechanism.

The essentiality assessment of every gene in the genome of *M. extorquens* PA1 represents a valuable resource for the identification of previously undescribed components involved in methylotrophy. In addition to the identification of Prk as a novel core component of the cascade regulating one-carbon metabolism, there are still dozens of new methylotrophy genes that have yet to be assigned a role in this complex process, including genes encoding for potential transporters and regulators, as well as many genes of unknown function. The discovery of Prk illustrates that regulation-guided identification of genes involved in a certain process might overlook central components. On the other hand, essentiality-guided identification, such as TnSeq, is unable to predict important factors if there is redundancy (e.g. the four formate dehydrogenases of *M. extorquens* which are highly regulated). To

get systems-level understanding of an organisms, both approaches combined with detailed biochemical investigation is crucial.

# 6. References


Abanda-Nkpwatt D, Müsch M, Tschiersch J, Boettner M & Schwab W (2006) Molecular interaction between *Methylobacterium extorquens* and seedlings: Growth promotion, methanol consumption, and localization of the methanol emission site. *J Exp Bot* **57:** 4025–4032

Acharya P, Goenrich M, Hagemeier CH, Demmer U, Vorholt JA, Thauer RK & Ermler U (2005) How an enzyme binds the C1 carrier tetrahydromethanopterin. Structure of the tetrahydromethanopterin-dependent formaldehyde-activating enzyme (Fae) from *Methylobacterium extorquens* AM1. *J Biol Chem* **280:** 13712–13719

Alber BE, Spanheimer R, Ebenau-Jehle C & Fuchs G (2006) Study of an alternate glyoxylate cycle for acetate assimilation by *Rhodobacter sphaeroides*. *Mol Microbiol* **61:** 297–309

Anthony C (1970) Cytochrome c and the oxidation of C1 compounds in *Pseudomonas* AM1. *Biochem J* **119:** 54P–55P

Anthony C (1982) *The biochemistry of methylotrophs*. Academic Press London

Anthony C (1992) The c-type cytochromes of methylotrophic bacteria. *Biochim Biophys Acta* **1099:** 1–15

Anthony C (2011) How half a century of research was required to understand bacterial growth on C1 and C2 compounds; the story of the serine cycle and the ethylmalonyl-CoA pathway. *Sci Prog* **94:** 109–137

Anthony C & Williams P (2003) The structure and mechanism of methanol dehydrogenase. *Biochim Biophys Acta* **1647:** 18–23

Arfman N, Van Beeumen J, De Vries GE, Harder W & Dijkhuizen L (1991) Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant *Bacillus* spp. *J Biol Chem* **266:** 3955–3960

Arfman N & Dijkhuizen L (1990) Methanol dehydrogenase from thermotolerant methylotroph *Bacillus* C1. *Methods Enzym* **188:** 223–226

Arfman N, Hektor HJ, Bystrykh L V, Govorukhina NI, Dijkhuizen L & Frank J (1997) Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from *Bacillus methanolicus*. *Eur J Biochem* **244:** 426–433

Arfman N, Watling EM, Clement W, van Oosterwijk RJ, de Vries GE, Harder W, Attwood MM & Dijkhuizen L (1989) Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. *Arch Microbiol* **152:** 280–288

Bassham JA, Benson AA, Kay LD, Harris AZ, Wilson AT & Calvin M (1953) The path of carbon in photosynthesis. XXI. The cyclic regeneration of carbon dioxide acceptor. *J Amer chem Soc*

Bertau M, Offermanns H, Plass L, Schmidt F & Wernicke H-J (2014) *Methanol: The basic chemical*

*and energy feedstock of the future*. Springer Berlin Heidelberg

Bosch G, Skovran E, Xia Q, Wang T, Taub F, Miller JA, Lidstrom ME & Hackett M (2008) Comprehensive proteomics of *Methylobacterium extorquens* AM1 metabolism under single carbon and nonmethylotrophic conditions. *Proteomics* **8:** 3494–3505

Brautaset T, Jakobsen OM, Josefsen KD, Flickinger MC & Ellingsen TE (2007) *Bacillus methanolicus*: a candidate for industrial production of amino acids from methanol at 50°C. *Appl Microbiol Biotechnol* **74:** 22–34

Carroll SM, Xue KS & Marx CJ (2014) Laboratory divergence of *Methylobacterium extorquens* AM1 through unintended domestication and past selection for antibiotic resistance. *BMC Microbiol* **14:** 2

Chistoserdova L (2011) Modularity of methylotrophy, revisited. *Env. Microbiol* **13:** 2603–2622

Chistoserdova L (2015) Methylotrophs in natural habitats: current insights through metagenomics. *Appl Microbiol Biotechnol* **99:** 5763–5779

Chistoserdova L (2016) Lanthanides: New life metals? *World J Microbiol Biotechnol* **32:** 1–7

Chistoserdova L, Chen SW, Lapidus A & Lidstrom ME (2003) Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. *J Bacteriol* **185:** 2980–2987

Chistoserdova L, Crowther GJ, Vorholt JA, Skovran E, Portais JC & Lidstrom ME (2007) Identification of a fourth formate dehydrogenase in *Methylobacterium extorquens* AM1 and confirmation of the essential role of formate oxidation in methylotrophy. *J Bacteriol* **189:** 9076–9081

Chistoserdova L, Kalyuzhnaya MG & Lidstrom ME (2013) Cycling single-carbon compounds: from omics to novel concepts. *Microbe* **8:** 395–400

Chistoserdova L, Laukel M, Portais JC, Vorholt JA & Lidstrom ME (2004) Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J Bacteriol* **186:** 22–28

Chistoserdova L, Vorholt JA, Thauer RK & Lidstrom ME (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **281:** 99–102

Crowther GJ, Kosaly G & Lidstrom ME (2008) Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *J Bacteriol* **190:** 5057–5062

Dangel A & Tabita FR (2015) CbbR, the master regulator for microbial carbon dioxide fixation. *J Bacteriol* **197:** 3488–3498

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, von Mering C & Vorholt JA (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci USA* **106:** 16428–16433

Doronina N V, Sokolov AP & Trotsenko YA (1996) Isolation and initial characterization of aerobic chloromethane-utilizing bacteria. *FEMS Microbiol Lett* **142:** 179–183

Dunstan PM & Anthony C (1973) Microbial metabolism of C1 and C2 compounds. The role of acetate during growth of *Pseudomonas* AM1 on C1 compounds, ethanol and β-hydroxybutyrate. *Biochem J* **132:** 797–801

Erb TJ, Berg IA, Brecht V, Muller M, Fuchs G & Alber BE (2007) Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc Natl Acad Sci USA* **104:** 10631–10636

Erb TJ, Retey J, Fuchs G & Alber BE (2008) Ethylmalonyl-CoA mutase from *Rhodobacter sphaeroides* defines a new subclade of coenzyme $B_{12}$-dependent acyl-CoA mutases. *J Biol Chem* **283:** 32283–32293

Escalante-Semerena JC, Rinehart Jr. KL & Wolfe RS (1984) Tetrahydromethanopterin, a carbon carrier in methanogenesis. *J Biol Chem* **259:** 9447–9455

Gaballa A, Newton GL, Antelmann H, Parsonage D, Upton H, Rawat M, Claiborne A, Fahey RC & Helmann JD (2010) Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in Bacilli. *Proc Natl Acad Sci USA* **107:** 6482–6486

Gälli R & Leisinger T (1985) Specialized bacterial strains for the removal of dichloromethane from industrial waste. *Conserv Recycl* **8:** 91–100

Gonzalez CF, Proudfoot M, Brown G, Korniyenko Y, Mori H, Savchenko A V & Yakunin AF (2006) Molecular basis of formaldehyde detoxification. Characterization of two S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J Biol Chem* **281:** 14514–22

Harms N, Ras J, Koning S, Reijnders WNM, Stouthamer AH & van Spanning RJM (1996) Genetics of C1 metabolism regulation in *Paracoccus denitrificans*. *Microbial growth on C1 compounds.* Springer Netherlands 126–132

Hemmann JL, Saurel O, Ochsner AM, Stodden BK, Kiefer P, Milon A & Vorholt JA (2016) The one-carbon carrier methylofuran from *Methylobacterium extorquens* AM1 contains a large number of α- and γ-Linked glutamic acid residues. *J Biol Chem* **291:** 9042–9051

Keltjens JT, Pol A, Reimann J & Op den Camp HJ (2014) PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl Microbiol Biotechnol* **98:** 6163–6183

Kloosterman H, Vrijbloed JW & Dijkhuizen L (2002) Molecular, biochemical, and functional characterization of a Nudix hydrolase protein that stimulates the activity of a nicotinoprotein alcohol dehydrogenase. *J Biol Chem* **277:** 34785–34792

Knief C, Frances L & Vorholt JA (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

Kornberg HL & Krebs HA (1957) Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature* **179:** 988–991

Korotkova N, Chistoserdova L, Kuksa V & Lidstrom ME (2002) Glyoxylate regeneration pathway in

the methylotroph *Methylobacterium extorquens* AM1. *J Bacteriol* **184:** 1750–1758

Korotkova N, Lidstrom ME & Chistoserdova L (2005) Identification of genes involved in the glyoxylate regeneration cycle in *Methylobacterium extorquens* AM1, including two new genes, *meaC* and *meaD*. *J Bacteriol* **187:** 1523–1526

Krog A, Heggeset TM, Müller JEN, Kupper CE, Schneider O, Vorholt JA, Ellingsen TE & Brautaset T (2013) Methylotrophic *Bacillus methanolicus* encodes two chromosomal and one plasmid born $NAD^+$ dependent methanol dehydrogenase paralogs with different catalytic and biochemical properties. *PLoS One* **8:** e59188

Large PJ, Peel D & Quayle JR (1961) Microbial growth on C1 compounds. 2. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM1, and methanol-grown Hyphomicrobium vulgare. *Biochem J* **81:** 470–480

Large PJ, Peel D & Quayle JR (1962a) Microbial growth on C1 compounds. 3. Distribution of radioactivity in metabolites of methanol-grown *Pseudomonas* AM1 after incubation with [$^{14}$C]methanol and [$^{14}$C]bicarbonate. *Biochem J* **82:** 483–488

Large PJ, Peel D & Quayle JR (1962b) Microbial growth on C1 compounds. 4. Carboxylation of phosphoenolpyruvate in methanol-grown *Pseudomonas* AM1. *Biochem J* **85:** 243–250

Large PJ & Quayle JR (1963) Microbial growth on C1 compounds. 5. Enzyme activities in extracts of *Pseudomonas* AM1. *Biochem J* **87:** 386–396

Laukel M, Rossignol M, Borderies G, Volker U & Vorholt JA (2004) Comparison of the proteome of *Methylobacterium extorquens* AM1 grown under methylotrophic and nonmethylotrophic conditions. *Proteomics* **4:** 1247–1264

Leßmeier L, Pfeifenschneider J, Carnicer M, Heux S, Portais J-C & Wendisch VF (2015) Production of carbon-13-labeled cadaverine by engineered *Corynebacterium glutamicum* using carbon-13-labeled methanol as co-substrate. *Appl Microbiol Biotechnol* **99:** 10163–10176

Maden BE (2000) Tetrahydrofolate and tetrahydromethanopterin compared: functionally distinct carriers in C1 metabolism. *Biochem J* **350 Pt 3:** 609–629

Marx CJ, Bringel F, Chistoserdova L, Moulin L, Farhan Ul Haque M, Fleischman DE, Gruffaz C, Jourand P, Knief C, Lee MC, Muller EE, Nadalig T, Peyraud R, Roselli S, Russ L, Goodwin LA, Ivanova N, Kyrpides N, Lajus A, Land ML, et al (2012) Complete genome sequences of six strains of the genus *Methylobacterium*. *J Bacteriol* **194:** 4746–4748

Marx CJ, Laukel M, Vorholt JA & Lidstrom ME (2003) Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *J Bacteriol* **185:** 7169–7175

McLennan AG (2006) The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63:** 123–143

Meister M, Saum S, Alber BE & Fuchs G (2005) L-malyl-coenzyme A/β-methylmalyl-coenzyme A lyase is involved in acetate assimilation of the isocitrate lyase-negative bacterium *Rhodobacter*

*capsulatus*. *J Bacteriol* **187:** 1415–1425

Metzger LC, Francez-Charlot A & Vorholt JA (2013) Single-domain response regulator involved in the general stress response of *Methylobacterium extorquens*. *Microbiology* **159:** 1067–1076

Morris CJ, Kim YM, Perkins KE & Lidstrom ME (1995) Identification and nucleotide sequences of *mxaA*, *mxaC*, *mxaK*, *mxaL*, and *mxaD* genes from *Methylobacterium extorquens* AM1. *J Bacteriol* **177:** 6825–6831

Muller EEL, Hourcade E, Louhichi-Jelail Y, Hammann P, Vuilleumier S & Bringel F (2011) Functional genomics of dichloromethane utilization in *Methylobacterium extorquens* DM4. *Env Microbiol* **13:** 2518–2535

Müller JEN, Litsanov B, Bortfeld-Miller M, Trachsel C, Grossmann J, Brautaset T & Vorholt JA (2014) Proteomic analysis of the thermophilic methylotroph *Bacillus methanolicus* MGA3. *Proteomics* **14:** 725–737

Müller JEN, Meyer F, Litsanov B, Kiefer P, Potthoff E, Heux S, Quax WJ, Wendisch VF, Brautaset T, Portais JC & Vorholt JA (2015a) Engineering *Escherichia coli* for methanol conversion. *Metab Eng* **28:** 190–201

Müller JEN, Meyer F, Litsanov B, Kiefer P & Vorholt JA (2015b) Core pathways operating during methylotrophy of *Bacillus methanolicus* MGA3 and induction of a bacillithiol-dependent detoxification pathway upon formaldehyde stress. *Mol Microbiol* **98:** 1089–1100

Nakagawa T, Mitsui R, Tani A, Sasa K, Tashiro S, Iwama T, Hayakawa T & Kawai K (2012) A catalytic role of XoxF1 as $La^{3+}$-dependent methanol dehydrogenase in *Methylobacterium extorquens* strain AM1. *PLoS One* **7:** e50480

Nayak DD & Marx CJ (2014a) Methylamine utilization via the N-methylglutamate pathway in *Methylobacterium extorquens* PA1 involves a novel flow of carbon through C1 assimilation and dissimilation pathways. *J Bacteriol* **196:** 4130-4139

Nayak DD & Marx CJ (2014b) Genetic and phenotypic comparison of facultative methylotrophy between *Methylobacterium extorquens* strains PA1 and AM1. *PLoS One* **9:** e107887

Nunn DN & Lidstrom ME (1986) Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1. *J Bacteriol* **166:** 581–590

Ochsner AM, Sonntag F, Buchhaupt M, Schrader J & Vorholt JA (2014) *Methylobacterium extorquens*: methylotrophy and biotechnological applications. *Appl Microbiol Biotechnol* **99:** 517-534

Okubo Y, Skovran E, Guo X, Sivam D & Lidstrom ME (2007) Implementation of microarrays for *Methylobacterium extorquens* AM1. *OMICS* **11:** 325–340

Olah GA (2013) Towards oil independence through renewable methanol chemistry. *Angew Chem Int Edit* **52:** 104–107

Peel D & Quayle JR (1961) Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM 1. *Biochem J* **81:** 465–469

Peyraud R, Kiefer P, Christen P, Massou S, Portais JC & Vorholt JA (2009) Demonstration of the ethylmalonyl-CoA pathway by using $^{13}C$ metabolomics. *Proc Natl Acad Sci USA* **106:** 4846–4851

Pol A, Barends TR, Dietl A, Khadem AF, Eygensteyn J, Jetten MS & Op den Camp HJ (2014) Rare earth metals are essential for methanotrophic life in volcanic mudpots. *Env Microbiol* **16:** 255–264

Pomper BK, Saurel O, Milon A & Vorholt JA (2002) Generation of formate by the formyltransferase/hydrolase complex (Fhc) from *Methylobacterium extorquens* AM1. *FEBS Lett* **523:** 133–137

Pomper BK & Vorholt JA (2001) Characterization of the formyltransferase from *Methylobacterium extorquens* AM1. *Eur J Biochem* **268:** 4769–4775

Pomper BK, Vorholt JA, Chistoserdova L, Lidstrom ME & Thauer RK (1999) A methenyl tetrahydromethanopterin cyclohydrolase and a methenyl tetrahydrofolate cyclohydrolase in Methylobacterium extorquens AM1. *Eur J Biochem* **261:** 475–480

Quayle JR & Keech DB (1959) Carbon assimilation by *Pseudomonas oxalaticus* (OX1). 1. Formate and carbon dioxide utilization during growth on formate. *Biochem J* **72:** 631–637

Quayle JR & Keech DB (1963) Carbon assimilation by *Pseudomonas oxalaticus* (OX1). 2. Formate and carbon dioxide utilization by cell-free extracts of the organism grown on formate. *Biochem J* **87:** 368–373

Ras J, Van Ophem PW, Reijnders WN, Van Spanning RJ, Duine JA, Stouthamer AH & Harms N (1995) Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. *J Bacteriol* **177:** 247–51

Richardson IW & Anthony C (1992) Characterization of mutant forms of the quinoprotein methanol dehydrogenase lacking an essential calcium ion. *Biochem J* **287.3:** 709–715

Schmidt S, Christen P, Kiefer P & Vorholt JA (2010) Functional investigation of methanol dehydrogenase-like protein XoxF in *Methylobacterium extorquens* AM1. *Microbiology* **156:** 2575–2586

Schneider K, Peyraud R, Kiefer P, Christen P, Delmotte N, Massou S, Portais JC & Vorholt JA (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:** 757–766

Schrader J, Schilling M, Holtmann D, Sell D, Filho M V, Marx A & Vorholt JA (2009) Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends Biotechnol* **27:** 107–115

Semrau JD, DiSpirito AA & Yoon S (2010) Methanotrophs and copper. *FEMS Microbiol Rev* **34:** 496–531

Shirai S, Matsumoto T & Tobari J (1978) Methylamine dehydrogenase of *Pseudomonas* AM1. A subunit enzyme. *J Biochem* **83:** 1599–1607

Skovran E, Palmer AD, Rountree AM, Good NM & Lidstrom ME (2011) XoxF is required for expression of methanol dehydrogenase in *Methylobacterium extorquens* AM1. *J Bacteriol* **193:** 6032–6038

Söhngen NL (1906) Über Bakterien, welche Methan als Kohlenstoffnahrung und Energiequelle gerbrauchen. *Zentrabl Bakteriol Parasitenk Infekt* **15:** 513–517

Studer A, McAnulla C, Buchele R, Leisinger T & Vuilleumier S (2002) Chloromethane-induced genes define a third C1 utilization pathway in *Methylobacterium chloromethanicum* CM4. *J Bacteriol* **184:** 3476–3484

Sy A, Timmers AC, Knief C & Vorholt JA (2005) Methylotrophic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. *Appl Env Microbiol* **71:** 7245–7252

Tabita FR (1980) Pyridine nucleotide control and subunit structure of phosphoribulokinase from photosynthetic bacteria. *J Bacteriol* **143:** 1275–1280

Vonck J, Arfman N, De Vries GE, Van Beeumen J, Van Bruggen EF & Dijkhuizen L (1991) Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus* sp. C1. *J Biol Chem* **266:** 3949–3954

Vorholt JA (2002) Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch Microbiol* **178:** 239–249

Vorholt JA (2012) Microbial life in the phyllosphere. *Nat Rev Microbiol* **10:** 828–840

Vorholt JA, Chistoserdova L, Lidstrom ME & Thauer RK (1998) The NADP-dependent methylene tetrahydromethanopterin dehydrogenase in *Methylobacterium extorquens* AM1. *J Bacteriol* **180:** 5351–5356

Vorholt JA, Marx CJ, Lidstrom ME & Thauer RK (2000) Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. *J Bacteriol* **182:** 6645–6650

Vorholt JA & Thauer RK (2002) Molybdenum and tungsten enzymes in C1 metabolism. *Met Ions Biol Syst* **39:** 571–619

de Vries GE, Arfman N, Terpstra P & Dijkhuizen L (1992) Cloning, expression, and sequence analysis of the *Bacillus methanolicus* C1 methanol dehydrogenase gene. *J Bacteriol* **174:** 5346–5353

Vu HN, Subuyuj GA, Vijayakumar S, Good NM, Martinez-Gomez NC & Skovran E (2016) Lanthanide-dependent regulation of methanol oxidation systems in *Methylobacterium extorquens*

AM1 and their contribution to methanol growth. *J Bacteriol* **198:** JB.00937-15

Vuilleumier S, Chistoserdova L, Lee MC, 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, et al (2009) Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* **4:** e5584

Whitaker WB, Sandoval NR, Bennett RK, Fast AG & Papoutsakis ET (2015) Synthetic methylotrophy: engineering the production of biofuels and chemicals based on the biology of aerobic methanol utilization. *Curr Opin Biotechnol* **33:** 165–175

Williams PA, Coates L, Mohammed F, Gill R, Erskine PT, Coker A, Wood SP, Anthony C & Cooper JB (2005) The atomic resolution structure of methanol dehydrogenase from *Methylobacterium extorquens*. *Acta Crystallogr D* **61:** 75–79

# 7. Acknowledgements

First, I would like to thank Prof Dr. Julia Vorholt for giving me the opportunity to work on this exciting project and for her support in these years. Thank you for trusting me and giving me a lot of freedom, but also for guiding and encouraging me when needed.

I would also like to thank my PhD thesis committee, Prof. Dr. Beat Christen and Prof. Dr. Jörn Piel for their interest in my work and for offering valuable advice to advance the project.

Additionally, I would like to thank my collaborators Prof. Dr. Beat Christen, Dr. Matthias Christen and Dr. Rémi Peyraud for teaching me a lot.

Furthermore I would like to thank the former and current members of the metabolism subgroup (a.k.a. Club Met), especially Dr. Jonas Ernst Norbert Müller for introducing me to the exciting field of methylotrophy when I joined the lab as a Semester student and Fabian Meyer for a lot of things including his contagious enthusiasm. In addition, I would like to thank all members of the PROMYSE and METAPP projects for many fruitful and enjoyable meetings.

In addition, I would like to thank all previous and current members of the Vorholt and Fischer groups, as well as the Erb group for such an enjoyable working atmosphere in the lab and especially for the countless hours of fun that we shared inside and outside of the lab, which was a continuous source of motivation. In this regard, I would like to thank especially Dr. Christine Vogel, Dr. Dominik Peter and Lisa Gottschlich.

I am also grateful to all Bachelor and Master Students that I supervised during my PhD for their scientific contributions, for brining excitement to the lab, and also for teaching me a lot personally. Special thanks go to Ralph Nüssli who put a lot of work into the NAD-Mdh exchange project both during his Master thesis and afterwards as an intern.

Finally, I would like to thank my family and Philipp for supporting me for all these years and for always being there when I needed them.