Metabolic flux analysis and metabolomics of Methylotrophs

Yanfen Fu

A dissertation

Submitted in partial fulfillment of the

Requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Mary E. Lidstrom, Chair

Caroline Harwood

Shaoyi Jiang

Program Authorized to Offer Degree:

Chemical Engineering

University of Washington

Abstract

Metabolic flux analysis and metabolomics of Methylotrophs

Yanfen Fu

Chair of the Supervisory Committee:

Frank Jungers Endowed Chair of Engineering Mary Lidstrom

Chemical Engineering and Microbiology Department

Methylotrophs are a group of microbes that use C1 compounds as the sole carbon and energy source, which are of interest to be used as platforms for bioconversion of C1 substrates to valuable chemicals and fuels. *Methylobacterium extorquens* AM1 is the best studied model organism for methylotrophs, and it uses the serine cycle for assimilation of C1 units. *Methylomicrobium buryatense* 5GB1 is a Type I methanotroph that uses the ribulose monophosphate (RuMP) pathway for assimilation, and it is also an industrially promising candidate for bio-conversion of methane into valuable chemicals. Metabolic flux analysis and metabolomics are two important approaches to investigate cell behavior at the level of metabolites and fluxes, which are integrative consequences of gene expression. These approaches will provide critical information for a rational strain design process in metabolic engineering iterations. In the first part of the thesis, the metabolic flux analysis approach was successfully applied to *M. exorquens* AM1 to investigate the metabolic network response to the trade-off between growth rate and biomass yield. Then the method was tailored to apply to *M. buryatense* 5GB1 for flux elucidation in core metabolism under methane growth to provide baseline information for better prediction of performance. At the end, the combined approaches were used to investigate the core metabolism of *M. buryatense* 5GB1 response to two substrates, methane and methanol.

This work has resulted in major new insights into methylotrophic metabolism. For the serine cycle methylotroph, it was demonstrated that the C3/C4 interconversion reactions, previously thought to be side reactions of little relevance to core metabolism, are critical to the ultimate values of growth rate and yield. This new insight now provides targets for strain engineering towards either maximum growth rate, or maximum yield. For the RuMP cycle methanotroph, it was determined first, that contrary to decades of

assumptions in the literature, the Type I methanotroph studied here operates a complete, oxidative TCA cycle. That finding changes the entire metabolic network balance, especially with regards to NADH generation, and this knowledge is essential for generating accurate predictive metabolic models. Second, it was determined that the partial serine cycle present in Type I methanotrophs does not play a significant role during growth on methane. However, it was discovered that it does play a role during growth on methanol. Growth on methanol was discovered to involve a substantially rearranged metabolic network compared to methane, in complete contrast to literature expectations. In this case, the TCA cycle is broken, NADH is generated mainly from formate oxidation, and flux shifts from the Embden-Meyer-Parnas pathway to the Entner-Douderoff pathway. In addition, substantial carbon is redirected to synthesis of glycogen and excretion of formate.

We have concluded that both metabolic flux analysis and metabolomics are extremely valuable tools in understanding the metabolic network of methylotrophs, and their use has corrected metabolic misunderstanding that has been in the literature for decades. These methods must be tailored for the C1 assimilation system, and the results presented in this thesis demonstrate an approach for the two main types of methylotrophs. These new techniques are now ready to use for investigating other methylotrophs in the future, and to serve as important tools for future metabolic engineering of methylotrophs for biotechnology applications.

ACKNOWLEDGMENTS

It has been a great experience for me. I have so many people on this list that I feel deeply grateful to. Without them, it will be a different story. First of all, I would like to thank my previous mentor, Jackie Shanks, in Iowa State University, who gave me the wonderful opportunity of studying metabolic flux analysis on *E.coli*. I would also like to thank another mentor and good friend, Song Yang, who spent a lot of time training me to get to know the C1 metabolism in methylotrophs. Secondly, I am so gratefully to everyone in Lidstrom lab (Aaron, Alexey, Amanda, Bo, David, Frances, Ginger, Janet, Jing, Marina, Mila, Mitch, Melissa, Sasha, Xin, Yu and Yue) and my student Yi, for the supportive environment where I feel encouraged to do my daily research work in lab. I have also received a lot of help in mass spectrum instrumentation from experts in mass spectrometry facility, Martin Sadilek and Laura T. Carlson. Many thanks to Brian Fitz as well as Christopher E. Freye for help with GC-FID instrumentation for methanol measurement. I would also like to thank all committee members for your support during all the exams. Last but not least, I would like to thank my advisor, mentor as well as role model, Mary E. Lidstrom, for offering the wonderful opportunity, for her support and encouragement, for all the inspiration she had brought to my work and life. I have learned and grown so much from the great experience that would be beneficial for life.

DEDICATION

To my husband, Ning, for his support, love, and fun outside school and work. To my little girl, Veronica, for the wonderful journey of motherhood. To my parents and sister, for their unconditional love and support.

Table of Contents

Chapter 1: Introduction

1.1 Methylotrophs

Methylotrophs are a group of microbes could use reduced carbon substrate (C1 substrate without carbon-carbon bond) such as methane, methanol, methylamine and formate as sole carbon source and energy source[1][2][3]. Furthermore, methanotrophs are a subgroup of methylotrphs that could grow on methane, which play an import role in global methane oxidation[1][4].

In recently years, a rising interests on using microbes for high value chemicals and biofuel production have been observed with the hope of building biorenewable system to substitute fossil fuel and petroleum derived substrate[5–9]. Under this context, methylotrophs are attractive biosytems for methane or methanol based bioconversion due to the fact that the conversion could happen near room temperature and pressure. They could be particularly useful for the production of single cell protein, vitamins, and biofuels from substrate which is cheap and abundant[10–13][14]. Meanwhile, due to the fact that methane is also a prime mitigation target for global warming alleviation, there's also possibility using methanotrophs for environmental remediation[15].

In order to facilitate and possibly speed up the biotechnology application with methylotrophs, basic research questions about the methylotrophic metabolism need to be answered especially on model organisms, such as *Methylobacterium extorquens* AM1 and *Methylomicrobium buryatense* 5G. With the advance of current genome sequence techniques and other 'omics' technique, it is much easier nowadays to conduct investigation on how metabolism works around specific bacteria. In general, two major carbon assimilation pathways occur in methylotrophs are the serine cycle pathway and ribulose monophosphate (RuMP) pathway. Luckily, the two methylotrophs included in this study deploy serine cycle and RuMP cycle respectively. In the following two sections, I would explain two different routes more in depth.

1.2 *Methylobacterium extorquens* AM1

Methylobacterium extorquens AM1 is the model organism for C1 metabolism investigation as well as methanol-based biotechnology application[16–18][14]. With the availability of genome sequence

data[19], genetic toolbox [20], genome scale model[21] as well as metabolomics and fluxomics technique[21][22], it offers great opportunities of rational strain development for biotechnology applications. A successful attempt to metabolic engineering AM1 for 1-butanol production using ethylmalonyl-CoA pathway conducted by Bo Hu was a great test case of using AM1 as a potential platform to produce high value chemical[23].

Yield is an important parameter to evaluate whether the bioprocess will be economically feasible. It is one of the parameters need to be optimized during strain development. When cell is under energy-limited growth, biomass yield and growth rate was determined by yield and rate of ATP production[24]. Constrained by thermodynamic principals, a trade-off between rate and yield should exist for heterotrophic organism theoretically[25]. However, the knowledge of how the central carbon metabolism responds to the trade-off is lacking in AM1. It will be very informative for strain development if the pathways responding to the trade-off could be identified. In Chapter 2, detailed metabolic flux maps were generated by ^{13}C metabolic flux analysis, and pathways correlated to high growth rate and high biomass yield were identified respectively. Mutant analysis supports the finding generated from metabolic flux analysis. Before getting into how the flux distribution through the central carbon metabolism in AM1, the central carbon metabolism network will be described in the following section.

1.2.1. Central carbon metabolism

The central carbon metabolism pathway was clearly laid out based on existing biochemistry knowledge[21]. The central carbon metabolism shown in Figure 1.1 in AM1 includes three major modules: tetrahydromethanopterin-dependent oxidation pathway, serine cycle, and ethylmalonyl-CoA pathway. Extracellular methanol needs to be transported into periplasm, and then converted into formaldehyde with methanol dehydrogenase. Formaldehyde will then be converted into formate through tetrahydromethanopterin-dependent oxidation pathway. At the first branch point, formate could either be oxidized to CO_2 to provide reducing power NADH catalyzed by formate dehydrogenase or assimilated through serine cycle to produce acetyl-CoA and C3, C4 intermediates compounds. Acetyl-CoA then entered ethylamlonyl-CoA cycle for glyoxylate regeneration to keep the assimilation serine cycle running. TCA cycle and reversed glycolysis pathway also need to be operating to produce enough precursors for

biomass synthesis. The metabolites shown in bold in Figure 1.1 are branch point metabolites, which were involved in more than one reaction. Metabolites with * are precursors for biomass synthesis.

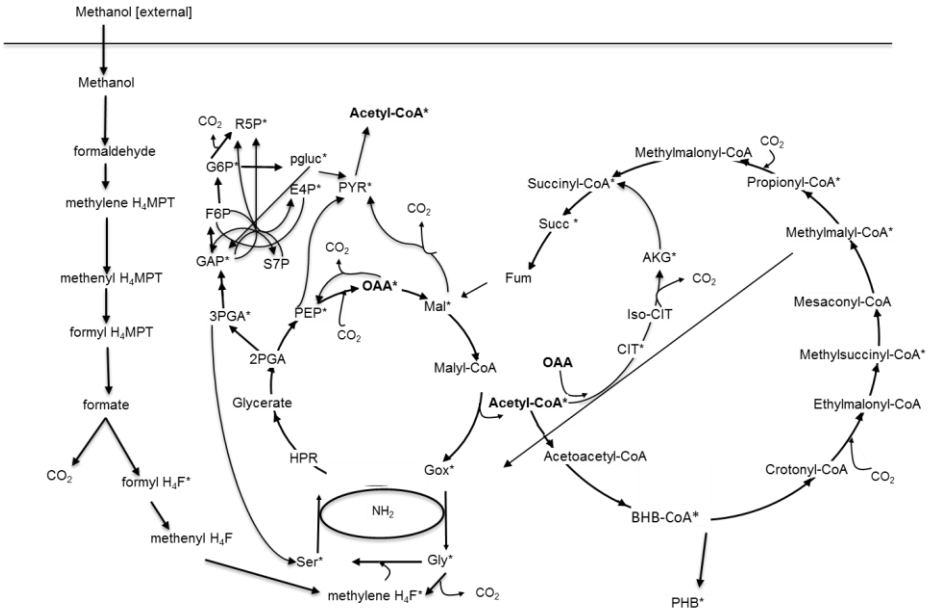


Figure 1.1 Central carbon metabolism in AM1 includes 3 modules: tetrahydromethanopterin-dependent oxidation pathway, serine cycle, and ethylmalonyl-CoA pathway

1.2.2. Central carbon metabolism response to environmental change and gene manipulation

Under methanotrophic growth, the ethylmalonyl-CoA pathway is operating to regenerate glyoxylate. Since AM1 is a facultative methylotroph, it could grow on multicarbon substrate, such as acetate or succinate. The central carbon metabolism shifted when the carbon source switched from C1 to multi-carbon compounds[26]. The flexibility of central carbon metabolism of AM1 provides capabilities of switching different carbon sources.

In summary, AM1 is a model organism for facultative type II methylotroph, with serine cycle as the major assimilation pathway for C1 unit to biomass. While on the other hand, Methylomicrobium buryatense 5G is an obligate type I methanotroph, which will be described in detail in the following section.

1.3 *Methylomicrobium buryatense* 5G

1.3.1 Central carbon metabolism

Aerobic methanotrophs oxidize methane to convert nature gas chemicals and liquid fuels via metabolic engineering. *Methylomicrobium buryatense* 5G, a haloalkaliphilic gammaproteobacterial (type I) methenotroph, is the strain of interest for industrial biotechnology because it is one of the best characterized methanotroph strain equipped with genetic manipulation tools[27]. In type I methanotrophs, ribulose mono phosphate cycle (RuMP) is the major assimilation pathway to incorporate C1 compound (formaldehyde) into central carbon metabolism through pyruvic acid node and dihydroxyacetone phosphate (DHAP), shown in Figure 1.2. Methane was oxidized into methanol with particulate methane monooxygenase (pMMO), and then converted into formaldehyde with methanol dehydrogenase. Formaldehyde will either be further oxidized into CO_2 to produce NADH or incorporated into RuMP cycle by reacting with ribulose-5-phosphate (R5P) to produce hexulose-6-phophate (H6P). Ribulose-5-phosphate needs to be regenerated from fructose-6-phosphate and glyceraldehyde phosphate to keep the cycle running. Two major variants of assimilation pathway are available to convert 6-carbon sugar phosphate into pyruvic acid and dihydroxyacetone phosphate, namely, EMP variant and ED variant, show in in Figure 1.2 as well. The net outcome of this RuMP cycle is incorporate 3 molecules of carbon in methane into 1 molecule of 3 carbon compound.

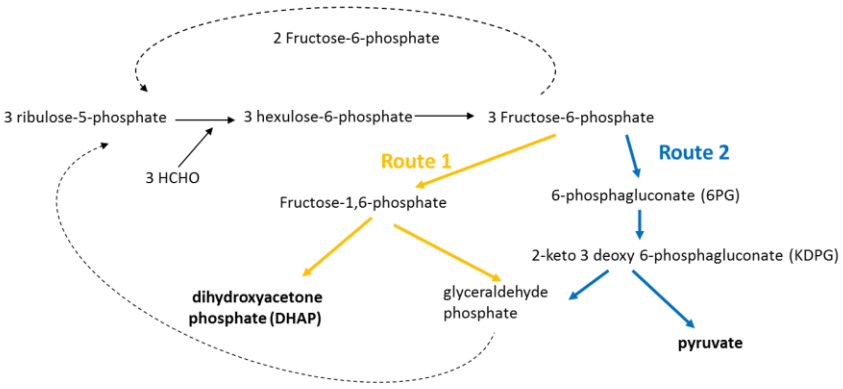


Figure 1.2 Methane assimilation using ribulose mono phosphate cycle (RuMP). Two variants are available to convert 6 carbon sugar phosphates into C3 compounds.

After pyruvate node, there are multiple carboxylation reactions to connect C3 compounds to C4 compounds, such as oxaloacetate node as well as malic acid node shown in Figure 1.3. Oxaloacetate (OAA) could be synthesized from PEP and pyruvate through carboxylation reaction. It could also interconvert with malate through malate dehydrogenase (MDH). Malate could be synthesized through pyruvate carboxylation, malate dehydrogenase, and fumarase from fumarate. Acetyl-CoA (AcCoA) is produced either by pyruvate dehydrogenase or malyl-CoA lyase (MCL), and then enters TCA cycle with OAA. It's also an important precursor for amino acid biosynthesis or fatty acid biosynthesis.

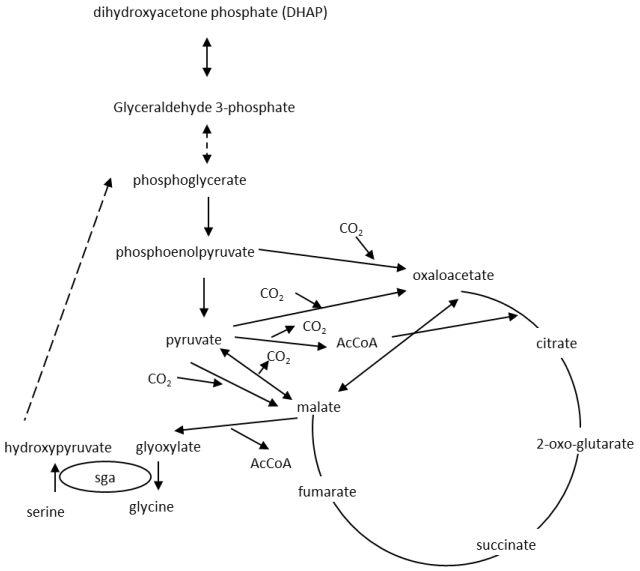


Figure 1.3 Metabolic pathways after DHAP and pyruvate node, multiple routes could convert C3 compounds to oxaloacetate and malate node.

1.3.2 Metabolic engineer for lipid production

Methylomicrobium buryatense 5G has potential to be industrial workhorse methanotroph to convert methane or nature gas into liquid fuels. The native methanotroph has packed inner membranes, which are mainly fatty acids. It has whole genome sequenced and annotated by JGI[28], enabling genomic manipulation for metabolic engineering as well as genomic model construction, which could further be used to design and guide rational metabolic engineering. Lidstrom lab has developed toolboxes for genetic manipulation[27], also has transcriptome data available under various culture conditions of 5G, trying to looking for genes correlated to high inner membrane contents. However, knowledge on

metabolic flux distribution through central carbon metabolism is still lacking. This would be a critical piece of information to be obtained for metabolic engineering.

1.4 Metabolic flux analysis

Flux balance analysis and ^{13}C labeled metabolic flux analysis are two major fields in metabolic flux analysis[29]. Both of them are essential for providing insights and prediction to the process of systematic metabolic engineering. Flux balance analysis uses a stoichiometric model developed from annotated genome (called genome scale metabolic model). It is a constraint-based approach (mass balance and stoichiometry are the major constraints) to calculate flux distribution in a pre-defined metabolic model under steady state consideration. The optimization needs to achieve a specific objective function, for instance, maximization of cell growth or certain compound production. It is particularly useful for phenotype prediction of mutants as well as metabolic model probing in the process of metabolic design. However, the ability of describing accurate flux distribution for a specific phenotype is very limited for FBA. In order to get quantitatively description of *in vivo* flux, stationary ^{13}C flux analysis or non-stationary ^{13}C flux analysis need to be conducted.

1.4.1 Stationary ^{13}C flux analysis

Stationary ^{13}C flux analysis is a powerful tool, which combined both experimental and computational approaches to understand the metabolic pathways in a living organism. It is based on stoichiometric reaction model and extracellular consumption and secretion, along with ^{13}C labeling information to calculate the *in vivo* reaction rate[30–32]. This technique needs biological system to be at both metabolic steady state as well as isotopical steady state. Multiple techniques could be applied to measure the isotopomer distribution of biomass-derived amino acids and intracellular metabolites. NMR and MS are the major players. In this study, only MS is used due to the advantage of high sensitivity.

The computational approach has been optimized based on Elementary Metabolite Unite (EMU) framework[33], which significantly reduced the computation time for flux calculation. Several software packages have been developed based on EMU framework[34–36]. And Influx_s was used in this study due to its excellent performance on numerical stability and computation time[36].It generates both flux

maps with absolute values as well as flux distribution normalized to specific substrate uptake rate, offering insights into cell behavior at the metabolic activity level.

Stationary ^{13}C metabolic flux analysis has usually applied to well-studied biological system where the central carbon metabolism model is well-established. Mostly the analysis has also been done on systems with multi-carbon input. The carbon rearrangement from the substrate carries the flux information to labeling patterns of intracellular metabolites or proteingenic amino acids, which could later inferred back to flux distribution. Due to the limitation of stationary ^{13}C metabolic flux analysis, in order to decipher C1 metabolism where carbon input is sole entry of single carbon compound, non-stationary ^{13}C flux analysis need to be deployed as well.

1.4.2 Non-stationary ^{13}C flux analysis

Non-stationary ^{13}C flux analysis also named as kinetic flux analysis, distinguished from stationary ^{13}C flux analysis mainly on the following two aspects. First of all, metabolic steady state is the only prerequisite for non-stationary ^{13}C flux analysis. That leads to shorter experiment duration. Secondly, non-stationary ^{13}C flux analysis could be applied to global level (i.e. the entire central carbon metabolism) for local point without systematical isotopomer mapping and balancing (i.e branch points, where flux ratio is needed). Several approaches including software packages have been developed for local flux ratio quantification[37], local flux quantification[38] as well as global flux estimation based on non-stationary flux analysis[39].

1.5 Goals and dissertation layout

In order to facilitate and speed up the biotechnology application on methylotrophs, basic questions about metabolism need to be answered. With the advance of current genome sequence techniques and other 'omics' technique, it is much easier nowadays to conduct investigation around specific methylotroph bacteria. The goal for my study is taking metabolomics and fluxomics approaches to understand C1 metabolism at the level of metabolites through two strains of methylothophs: *Methylobacterium extorquens* AM1 and *Methylomicrobium buryatense* 5G.

This thesis includes 4 more chapters. In Chapter 2, I will present the study of trade-off between growth rates and yield using *M. extorquens* AM1 though stationary ^{13}C flux analysis (this chapter has been

published on BMC microbiology, 2016). In Chapter 3, the central metabolism of *M. buryatense* 5G growing on methane was investigated using both stationary ^{13}C flux analysis as well as mutant analysis. In this study, evidences of oxidative TCA cycle operating were provided (manuscript submitted to Metabolic Engineering Journal, currently under review). In Chapter 4, the central metabolism of *M. buryatense* 5G growing on methanol was investigated using stationary ^{13}C flux analysis, RNA-seq as well as targeted metabolomics. Significant metabolism rearrangement was found in the system when substrate shifted from methane to methanol (manuscript in preparation). Chapter 5 summarized all the major findings and layout the future directions.

Chapter 2: Difference in C3-C4 metabolism underlies tradeoff between growth rate and biomass yield in Methylobacterium extorquens AM1

Abstract

Background: Two variants of *Methylobacterium extorquens* AM1 demonstrated a trade-off between growth rate and biomass yield. In addition, growth rate and biomass yield were also affected by supplementation of growth medium with different amounts of cobalt. The metabolism change relating to these growth phenomena as well as the trade-off was investigated in this study. ^{13}C metabolic flux analysis was used to generate a detailed central carbon metabolic flux map with both absolute and normalized flux value.

Results: The major differences between the two variants occurred at the formate node as well as within C3-C4 inter-conversion pathways. Higher relative flux through formyltetrahydrofolate ligase, phosphoenolpyruvate carboxylase, and malic enzyme led to higher biomass yield, while higher relative flux through pyruvate kinase and pyruvate dehydrogenase led to higher growth rate. These results were then tested by phenotypic studies on three mutants (null *pyk,* null *pck* mutant and null *dme* mutant) in both variants, which agreed with the model prediction.

Conclusions: In this study, ^{13}C metabolic flux analysis for two strain variants of *M. extorquens* AM1 successfully identified metabolic pathways contributing to the trade-off between cell growth and biomass yield. Phenotypic analysis of mutants deficient in corresponding genes supported the conclusion that C3-C4 inter-conversion strategies were the major response to the trade-off.

Background

Methylobacterium extorquens AM1 is a facultative α-proteobacterial methylotroph, that has been studied intensively over 50 years [40]. The availability of the genome sequence for *M. extorquens* AM1 [18], intensive developments of genetic tools [41, 42], and well-studied biochemistry and physiology have made the organism a model system for C1 metabolism. With the development of transcriptomics, proteomics, metabolomics and fluxomics, studies on C1 metabolism in *M. extorquens* AM1 have recently

been carried out using systems approaches [43][44][45]. C1 metabolism involves multiple C1-specific metabolic pathways, including the tetrahydromethanopterin-dependent oxidation pathway, the serine cycle, and the ethylmalonyl-CoA pathway as shown in Figure 2.1.

Recent studies have shown that cobalt is an important trace metal for methylotrophic growth in *M. extorquens* AM1. Cobalt is needed for vitamin B_{12} production used as cofactor for two enzymes involved in methylotrophy growth, methylmalonyl-CoA mutase (Mcm) and ethylmalonyl-CoA mutase (Ecm) in the ethylmalonyl-CoA pathway, and plays an important role in strain fitness [8, 9,11]. Three research groups have published optimized media recipes, including optimization of cobalt levels [46, 47][48][49]. However, the effect of cobalt on the overall central carbon metabolism in *M. extorquens* AM1 remains unknown.

Strain integrity can become compromised when the same strain is transferred between labs using different storage procedures, as illustrated recently for *M. extorquens* AM1 [50]. Phenotypic divergence was observed between an archival strain and a modern strain in terms of growth rate and fitness across various culture conditions [50]. The literature shows that two other strains have diverged in Mary Lidstrom's lab and Julia Vorholt's lab, after these strains were separated for 14 years. Different growth rates were reported from previous studies for both strains [17][51], which could be ascribed to a combination of culturing environment and unintended domestication of the *M. extorquens* AM1 strain, but the basis for this difference is not known.

It has been well-documented that a trade-off exists between rate and yield for heterotrophic organisms in which growth rate is predicted to be limited by ATP [52, 53]. However, it was not known whether such a tradeoff occurs in the *M. extorquens* AM1 strain variants. In *M. extorquens* AM1, the cell growth is predicted to be limited by reducing power instead of ATP [13], making the metabolic basis for such tradeoffs unclear. The availability of two strain variants with differences in growth rate and possibly in biomass yield offers an opportunity to decipher system-wide metabolic responses in *M. extorquens* AM1, including the possible trade-off between growth rate and biomass yield.

^{13}C metabolic flux analysis is a powerful tool, which combines both experimental and computational approaches to quantitatively understand the metabolic pathways in a living organism. It is based on a stoichiometric reaction model and extracellular consumption and secretion, along with ^{13}C labeling information to calculate *in vivo* reaction rates [30–32, 35, 54]. It generates both flux maps with absolute

values as well as flux distribution normalized to specific substrate uptake rate, offering insights into cell behavior at the metabolic activity level. This technique provides an approach to address the metabolic changes underlying the physiological differences observed in the two *M. extorquens* strain variants noted above.

In this study, it was established that a tradeoff exists between growth rate and biomass yield in these two *M. extorquens* AM1 strain variants, and ^{13}C metabolic flux analysis (MFA) was applied to both strain variants in order to assess the metabolic differences underlying the trade-off. In addition, MFA was applied to the same strain with different cobalt supplements to assess the effect of cobalt on overall central carbon metabolism and its contribution to the trade-off.

Results

2.1 Tradeoff between growth rate and biomass yield exists in *M. extorquens* AM1 for two strain variants as well as two cobalt levels

Observations of growth rate differences prompted a more in-depth study of two *M. extorquens* AM1 strains, a parent strain from the laboratory of Mary Lidstrom (LL strain), and a strain that originated from the Lidstrom laboratory, but was carried in the laboratory of Julia Vorholt (VL strain) since 2001. Both genomes of the two variants were sequenced in this study. A small number of genomic differences were detected between three different *M. extorquens* strains (VL, LL and Chris Marx modern strain[55]) (see supporting material table S2_1). However, these differences did not provide an obvious solution to the trade-off or suggest how the divergence in genome sequence affects growth rate. Growth experiments were conducted for the two strains with methanol as the sole carbon substrate. A minimal medium from the Lidstrom laboratory (called hypho medium, HY) was used to be consistent with previous work done in the Lidstrom laboratory. In addition, since levels of cobalt in the medium also affect growth rate of *M. extorquens* AM1 strains [8, 9] but the effect on biomass yield is unknown, this factor was also studied. Three different levels of cobalt were tested for growth rate experiments (data not shown). For both strains, the most dramatic differences were between 1.35 μM and 6.31 μM. These two concentrations were then chosen to test the effect of cobalt supplementation.

Biomass yield on methanol, which was defined as g cell dry weight generated per mol methanol consumption, was monitored by measuring methanol concentration in supernatant and biomass concentration during log phase (at least two time points were taken during log phase). Excretion products were tested in culture supernatant and were found to be negligible. A negative correlation between growth rate and biomass yield was observed, with both strain difference and cobalt level playing a role in the trade-off. As shown in Figure 2.2, the VL strain grew 22% faster than the LL strain in the presence of 1.35μM cobalt (normal HY medium), and similarly, 21% faster in medium with 6.31μM cobalt (termed HYC medium). However, biomass yield shows the opposite trend. The biomass yield of was 28% lower than that of the LL strain in both HY and HYC. A similar trade-off was also observed for the same strain growing in medium with different cobalt levels. The LL strain grew 26% faster but with 32% lower biomass yield in HYC medium than in HY medium, while the VL strain grew 25% faster with 28% lower biomass yield in HYC than HY.

In summary, within the four conditions tested in this study, LL in HY medium has the slowest growth rate but highest biomass yield while VL in HYC medium has the fastest growth rate but lowest biomass yield. Since both strain difference and cobalt level were taken into account in this set of comparisons, analysis of the metabolic response to the trade-off includes both factors.

2.2 Flux distribution reveals different metabolic strategies in C_3-C_4 interconversions and the formate branch node

In order to assess metabolic differences between the two strains, detailed central carbon flux maps were generated for each strain to provide a quantitative description for the two growth conditions using ^{13}C metabolic flux analysis (^{13}C-MFA). The published flux balance model and biomass composition for the VL strain [13] was used as a basis for a model developed in Influx_s. The specific methanol uptake rate of cells was measured in exponential phase, and an analysis of extracellular metabolites for cells grown in the same way showed negligible amounts. The model consists of 114 reactions including 9 reversible reactions and 2 scrambling reactions. A total of 149 MS measurements were used for model fitting with 12 calculated label measurements. The simulated flux data were then normalized to a specific methanol uptake rate to generate the flux distribution map for each condition. The flux distribution comparison shown in Figure 2.3 indicates the major differences were in the formate branch point and in C3-C4

interconversion pathways. As shown in the central carbon metabolism map, formate could either be oxidized to CO_2 by formate dehydrogenase (FDH) or be converted into formyl-H_4F by formyltetrahydrofolate ligase (FTHFL). As expected by the growth rate difference, the absolute methanol uptake rate and flux through formate to both branches was higher for the VL strain than the LL strain (see supporting file 2_3). However, the flux distribution across the branch point, which is normalized to the methanol uptake rate, shows a strain-specific difference in keeping with the higher biomass yield for the LL strain. For the LL strain, FTHFL carries 36% more flux entering the assimilation pathway than the VL strain in HY, 28% more in HYC.

Likewise, two routes exist for conversion of phosphoenolpyruvate (PEP) into acetyl-CoA (AcCoA). In the first route, the classic serine cycle involves conversion of PEP into oxaloacetate (OAA) through a carboxylation pathway involving PEP carboxylase (PEPCL), and then conversion of OAA into malate (MAL) by malate dehydrogenase (MDH), which is then further converted into glyoxylate and AcCoA by malyl-coA lyase (MCL). A reaction also exists to decarboxylate OAA back to PEP by phosphoenolpyruvate carboxykinase (PEPCK). An alternative route utilizes PEP to synthesize pyruvate through pyruvate kinase (PK), and then pyruvate is decarboxylated by pyruvate dehydrogenase (PYRDH) to produce AcCoA. Pyruvate could also be produced from decarboxylation of malate through malic enzyme (ME). The flux distribution shows that PEPCL had a higher normalized flux in the LL strain than the VL strain by 91% in HY and by 55% in HYC. For each strain grown at different cobalt concentrations, the normalized PEPCL flux decreased by 53% at the higher cobalt concentration in HYC for the LL strain and 42% for the VL strain. PEPCK is active for the LL strain in HY, but is negligible for the LL strain in HYC. For the VL strain, PEPCK is negligible regardless of the level of cobalt. For the malic enzyme pathway (ME), the LL strain has 97% higher flux than the VL strain in HY, but the difference became smaller in HYC, with the higher cobalt concentration. PK in the LL strain is 49% lower than VL in HY with the lower cobalt level, 25% in HYC. A similar trend was observed for PYRDH for both conditions but to different extents, 67% and 61% respectively.

2.3 Phenotypic differences of mutants

Based on the flux analysis results, three genes were targeted for mutation (*dme* encoding ME*, pck* encoding PEPCK and *pyk* encoding PK), to assess mutant phenotypes in the two strains. It is known that PEPCL, MCL, and MDH are required for growth on methanol [56] so those mutants were not generated. From the metabolic flux distribution result, the following phenotypes were predicted. The growth rate was associated with the PYK pathway encoded by *pyk*. Conditions with faster growth have higher flux through PYK. The *pyk* mutant should have higher negative impact on growth rate for the VL strain under both medium conditions, while for the LL strain the impact should be minor. The growth rate experiments on *pyk* mutants agree with this prediction. As shown in Table 2.1, for HY medium, LL:Δ*pyk* shows similar growth as LL, while VL:Δ*pyk* has 56% slower growth rate than VL. For HYC medium, LL:Δ*pyk* shows a 13% growth defect compared to LL under the same condition, while VL:Δ*pyk* has a 36% slower growth rate than VL under the same condition. The metabolic flux distributions can only predict the impact of this mutation on growth rate, but it is also of interest to assess impact on yield. The biomass yield of LL:Δ*pyk* and VL:Δ*pyk* showed minor changes in HY medium, while in HYC medium, biomass yield increased slightly for both strains. This suggests that in the Δ*pyk* mutant, the route of the standard serine cycle with PEPCL, MDH and ME carries more flux than in the original strain, which can lead to higher biomass yield. Secondly, metabolic flux map results also indicate that the non-essential pathways associated with biomass yield were ME encoded by *dme* and PEPCK encoded by *pck*. The higher biomass yield conditions have higher carbon flux through both ME and PEPCK. This predicts that both of these mutants should have a more negative impact on conditions with higher biomass yield. The experimental results on biomass yield for these two mutants agree with the prediction shown in Table 2.1. As shown in Table 2.1, the *dme* mutation had a negative effect on biomass yield in the LL strain in HY medium (54% decrease) compared to wild type LL. A smaller decrease in biomass yield (28%) occurred in VL:Δ*dme*. Likewise, the *pck* mutation also had a negative impact on biomass yield in the LL strain (46% decrease compared to wild type LL), while for VL:Δ*pck* in HY, it decreased by only 18%. The effect on biomass yield for both strains was less when the cells were grown with the higher level of cobalt. For both VL:Δ*dme* and VL:Δ*pck*, the biomass yield increased. Noticeably, the *dme* mutant in VL and LL strain had the same biomass yield in HYC when extra cobalt was supplied to the medium, while the difference was more

pronounced in HY medium. The Δ*pck* and Δ*dme* strains did not have large impacts on growth rate, but showed different extents of growth defects (Table 2.1). These two pathways involve NADH and ATP consumption, respectively. Therefore, these two mutants are expected to alter consumption of ATP and NADH, which might be expected to affect growth rate.

Discussion

This work generated four detailed flux distribution maps for two strain variants of *M. extorquens* AM1 (the LL and VL strains) grown with different cobalt levels. ^{13}C MFA has previously been applied to the VL strain during growth on methanol at a higher cobalt level [13]. However, it is not possible to compare the flux maps directly, as in the previous study the culture was grown with 5% CO_2 purging the culture system. The influence of 5% CO_2 on metabolism in *M. extorquens* AM1 is unknown, but given the number of carboxylation and decarboxylation reactions involved in growth on one-carbon compounds, it might alter fluxes. In this study we used air (0.5% CO_2) and also a slightly different medium recipe, to allow direct comparisons with previous metabolic studies of this bacterium carried out in this medium with 0.5% CO_2.

We have shown that the VL strain grew faster with lower biomass yield than the LL strain, regardless of the cobalt supplementation. From these results we conclude that a tradeoff exists between growth rate and biomass yield that is strain-specific. This suggests that these two strains are an example of unintended domestication, which has been reported previously for *M. extorquens* AM1 [12]. In addition, the cobalt level also affects growth rate and biomass yield in a manner that is not strain-specific. Improved growth rate and increased absolute fluxes through the EMC cycle in response to higher cobalt are in keeping with previous studies [57][49] suggesting that the EMC pathway could be a potential bottleneck for methanol assimilation under cobalt limitation conditions.

The occurrence of natural strain variants with growth rate/biomass yield tradeoffs along with the cobalt affect provided the opportunity to determine features of metabolic network response underlying these major physiological attributes of growth rate and biomass yield. ^{13}C-flux labeling revealed that FTHFL, the pathway involving the standard serine cycle (PEPCL, MDH, and MCL), and PEPCK and ME all had higher relative flux in conditions when biomass yield was higher, while the alternate pathway to AcCoA

involving PK and PYRDH had higher relative flux in conditions when growth rate was faster. These differences were predicted to affect NADH and ATP usage. PEPCK consumes one ATP, while PK and PYRDH produce one ATP and one NADH, respectively. MDH and MCL consume NADH and ATP, respectively. These results suggest that higher biomass yield is achieved under conditions that result in relatively higher relative fluxes through the standard serine cycle-based carbon assimilation pathways, even though this metabolic scheme consumes more NADH and ATP than the one involving conversion of pyruvate to AcCoA, at the expense of growth rate. Likewise, higher growth rate is achieved by diverting carbon through pathways with higher NADH and ATP production, with concomitant decrease in biomass yield.

There must be a genetic explanation for the trade-off. However, genome sequence results did not provide a clear clue, since none of the genomic changes is predicted to directly impact any of the C3/C4 reactions that show altered flux. One of the possible explanations is that the mutants in the gene cluster with unknown functions lead to shifted flux through C3/C4 enzymes due to the change in small molecule pools, which in turn affects yield versus growth rate.

The results of mutant phenotypes are consistent with these conclusions regarding metabolic network response, since loss of the biomass yield-related enzymes resulted in an impact on biomass yield, and loss of a growth-rate related enzyme resulted in an impact on growth rate. It is more difficult to predict the impact of the *pyk* mutant on yield, and of the *dme* and *pck* mutants on growth rate. C3-C4 interconversion pathways in this bacterium are flexible due to the fact that some of the enzymes are reversible (such as MDH, ME) as well as the fact that different enzymes are present that catalyze forward and reverse reactions (such as PK, PEP synthase, PEPCL and PEPCK). In the two strain variants, two C3/C4 pathways dominate but with different ratios, as shown in Figure 2.3, but we do not know how the network will change in the mutants. Therefore, although we can predict that removing the optimal variant would decrease growth rate or yield, respectively, the mutation may or may not also affect the other parameter, depending on how the metabolic network rearranges.

The cobalt concentration used in the growth medium also affected the fluxes and mutant phenotypes. In both strains, growth in higher cobalt resulted in higher growth rates and a correspondingly higher absolute flux through the assimilatory pathways. Since the EMC pathway contains two steps requiring B_{12}-utilizing

mutases (ethylmalonyl-CoA mutase and methylmalonyl-CoA mutase), it is expected that higher availability of cobalt would affect flux through the EMC shown by others[46][49]. The increased absolute flux through the EMC suggests that a cobalt-related component of one or both of these enzymes might limit flux under the lower cobalt growth condition. However, although the absolute flux increased through all of the central assimilatory pathways with increased cobalt, the flux distribution (normalized to the methanol uptake rate) did not. The flux distribution through the EMC was essentially the same for both the high and low cobalt conditions. This result suggests that the assimilatory flux is well balanced with both the methanol uptake rate and the formate branch point distribution, and that the metabolic network has the capacity to increase flux to this level. The significant changes in the flux distribution that occur in response to cobalt are not in the EMC pathway, but instead, are in the nexus between the serine cycle, the partial TCA cycle, the pentose-phosphate pathway, and the EMC pathway. It seems likely that this redundancy in C3-C4 interconversions at this metabolic intersection provides robustness for the metabolic network. The mechanism(s) for these changes are not yet known, but could occur at the transcriptional or post-transcriptional level.

The flux distribution maps presented here provide information that is potentially useful in guiding strain development for biotechnical applications. For applications in which faster strain growth is favored to create optimal economic conditions for a product (for instance, when a biomass product is the target), reactions identified in this study associated with faster growth are targets for manipulation. Similarly, if instead higher biomass yield is favored, for instance, with an excreted product, the reactions associated with higher biomass yield could be manipulation targets. It should also be noted that as shown in this study, a single trace mineral change in the medium changes the flux distribution, pointing out the importance of flux analysis in assessing metabolic dynamics and growth conditions.

Conclusion

Growth rate and biomass yield measurements for two culture conditions with two strain variants suggest trade-offs exist between growth rate and biomass yield. ^{13}C flux analysis was successfully applied to identify metabolic differences that could contribute to the trade-off. Pathways with different activities were identified, which were used to generate a hypothesis of the metabolic changes responding to the trade-

off, focused on C3-C4 inter-conversion strategies. The hypothesis was tested using mutants, and their phenotypes supported the conclusion that the flux difference mainly involves C3-C4 interconversion strategies.

List of abbreviations

AM1: *Methylobacterium extorquens* AM1; HY: hypho medium with 1.35μM cobalt; HYC: Hypho medium with 6.31μM

Methods

2.4 Chemicals and medium composition

^{13}C methanol of 99% purity was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). All other chemicals including metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO). Phusion DNA polymerase, dNTP, buffer, ligases, OneTaq Quick-Load 2X Master Mix and the Gibson Assembly Master Mix kits used in this study were from New England Biolabs (Ipswich, MA). Primers, the sequences of which are shown in Table 2.1, were obtained from Invitrogen (Grand Island, NY) and IDT (Coralville, IA). Acetonitrile and water used as UPLC solvents were UPLC-MS grade. Hypho minimal medium as previously described [58] was used for ^{13}C flux analysis cell culture and growth rate determination, involving two different concentrations of cobalt (1.35μM and 6.31μM respectively).

2.5 Bacterial strains, plasmid and growth condition

Strains used in this study are listed in Table 2.1. *Escherichia coli* strains Top 10 and S17-1 were cultivated at 37ºC in Luria-Bertani medium. Two wild type *M. extorquens* AM1 strains were used, in which VL is from the Julia Vorholt lab and LL is from the Mary Lidstrom lab. Plasmid pCM184 was used as template for backbone and *cre-kan* amplification. A Gibson assembly kit (New England Biolabs, Ipswich, MA) was used for plasmid construction to create knock out strains with a kanamycin marker FYF1, FYF2, FYF3 and FYF4, (primer sequences shown in Table 2.3). Insertion mutants of *pyk* were generated for both LL and VL variants as described earlier[59] because of our inability to obtain clean knockout using the method described in this study. 125mM methanol was used as the sole carbon source for cell culture and growth curves. For ^{13}C flux analysis, ^{13}C methanol was introduced at the stage of 3 ml seed culture in

tubes, and then inoculated into flasks with screw caps in a dilution ratio of 1:200 in triplicates. Cell culture was then quenched using fast filtration and liquid nitrogen after the cell reaches OD_{600} around 0.6-0.8 (7 generations). For normal growth curve cell culture, methanol was added as carbon source in 3 ml seed culture, and then inoculated into flasks with screw caps with the same dilution ratio.

2.6 Methanol measurements in supernatant

Methanol consumption rates were determined by taking time course samples for supernatant methanol measurements. At least 2 points were taken during exponential phase growth. 750 ul of the supernatant sample were centrifuge filtered through 0.22μm centrifuge filters (Costar® Spin-X® centrifuge tube filters, cellulose acetate membrane, pore size 0.22 μm, non-sterile, Sigma-Aldrich, (St. Louis, MO) for 2 minutes at 14000 rpm. GC-FID (Aglient, Santa Clara, CA) was used for methanol detection with a 6890 Gas Chromatograph equipped with flame ionization detector (FID). Data were collected and converted into matlab input files with LabVIEW 2010. Data analysis was later done in Matlab. SLB-IL60 column (Supelco, Bellefonte, PA, USA) with a 30 m x 0.25 mm inner diameter (i.d.) x 0.2 μ m thickness was installed for gas chromatograms. 1μl of supernatant was injected and separated in the column and then entering into FID detector with a split ratio 50:1. Oven temperatures were programmed as follows. An initial temperature of 120 °C was held for 1 minute, and then ramped at 35 °C /min until a final temperature of 200 °C was achieved. The FID was operated at 220 °C. All measurements were done in triplicate. With each run, a new calibration curve was then generated using the same instrument setting.

2.7 Metabolite extraction and biomass hydrolysis for mass isotopomer distribution measurements

Both proteinogenic amino acids and intracellular metabolites were used for isotopomer measurements in order to obtain good coverage of central carbon metabolism under methylotrophic growth. 20 ml of cell culture were collected and quenched using fast filtration using 0.2 μm nylon membrane filters (0.2 μm, 47mm from PALL life science, Port Washington, NY) and liquid nitrogen. Cell pellets on 0.22 μm filters were lyophilized using FreeZone 4.5 Liter benchtop freezedry system (Labconco, Kansas) for 8-12 hours to remove medium residue. The filter was incubated in 20ml of boiling water for 10 minutes, and then placed on ice for at least 20 minutes for protein precipitation. The whole broth was then centrifuged at 4°C, 5000rpm for 10 minutes. The clear supernatant was transferred into a clean 50 ml falcon tube, and

then frozen using liquid nitrogen. The remaining tubes containing protein and cell debris were centrifuged again, at 4°C, 5000rpm for 20 minutes, and supernatants were removed carefully. Frozen samples were then lyophilized twice, and reconstituted in 100µl nanopure water. The 100µl of sample was distributed into 2 vials. One 50µl sample was used for LC/MS/MS measurements, and the other was dried using a Speedvac (Labconco, Kansas), and derivatized with TBDMS and methoxyamine hydrochloride O-Methylhydroxylamin-hydrochlorid (Sigma-Aldrich, St. Louis, MO) using a previously described method [31]. Cell pellets were then hydrolyzed using 1ml of 6N HCl at 105°C for 22-24 hours. The remaining HCl was later evaporated with nitrogen flow. Dried samples were reconstituted in 500µl nanopure water, and ash and cell debris were then filtered out of the sample using centrifuge tube filter (spin-x costar, 0.22µm, Sigma-Aldrich). 50µl of the reconstituted sample was then dried and derivatized using TBDMS as noted above.

Both TBDMS-derivatized proteinogenic amino acids and intracellular metabolites were detected using an Agilent (Santa Clara, CA) GCMS with an Agilent J&W HP-5ms Ultra Inert GC Column (30 m×0.25 mm×0.25 µm, catalog number: 19091s-433). Each sample was run in duplicate on the GCMS with the same temperature program, and also as one more run with a slower oven temperature ramping rate. The oven temperature was programmed as follows: for regular runs, the oven temperature was programmed as follows:, initial temperature at 100 °C , hold for 4 minutes, then ramp to 300°C with ramp rate at 5 °C /min, then hold for 5 minutes, with total run time 49 minutes. In order to further separate compounds with similar retention times a slower run was developed. For the slower run, the ramp rate slows down to 2.5°C/min with total run 89 minutes. MS source temperature was set at 230 °C and MS quad temperature set as150°C.

Sugar phosphate was separated by a Zic-pHilic column (SeQuant, PEEK 150mm length ×2.1mm metal free, with 5µm polymeric film thickness) using LC-MS/MS (Xevo, Waters, Milford, MA). Mobile phase A is LC grade water with 2mM formic acid and 4mM ammonium, mobile phase B is acetonitrile with 0.1% formic acid. Column temperature was set at 30°C. The LC condition starts with 0.15ml/min flow rate with initial gradient A=15%, hold for 2 minutes, and then increased to 40% over the next 2 minutes. At 7 minute, A=50%, at 10 min, A=80%, at 11 minutes, A is set to 90%, held for 0.5 minutes, and then switched to 15% at 11.50 minute to re-equilibrate the column for another 4.5 minutes.

Waters Xevo G2-S Q-Tof equipped with Waters Acquity UPLC I class (Milford, MA) was used for some of the proteinogenic amino acids using HILIC-BEH-amide (Waters Acquity, part number 186004802, with dimension 2.1mm×150mm, 1.7um). Mobile phase A is water with 20 mM ammonium formate and 0.1% formic acid, mobile phase B is acetonitrile with 0.1% formic acid. LC solvent gradient was set as follow, starting at 10% A for 1 minute, at 2 minute, A=35%, at 3 minute, A=40%, at 6 minute, A=50%, at 7.5 minutes, A=80%, holding for another 2 minutes. Then the column is re-equilibrated at 9.51 minutes by running 10% A through the column until 5.5 minutes

2.8 Central carbon metabolism model construction and ^{13}C flux analysis

The central carbon metabolism pathway model used for ^{13}C flux analysis is shown in Figure 2.1. It was constructed based on a previous model [51], and details are shown in supporting file 2_1. The pathways included are: oxidation of methanol to CO_2 via formaldehyde and formate, conversion of formate to methylene-H_4F using the H_4F pathway; assimilation of methylene H_4F via the serine cycle; the TCA cycle; the ethylmalonyl-CoA pathway; gluconeogenesis; and anaplerotic reactions [reactions catalyzed by pyruvate kinase, PEP carboxylase and malic enzyme] overlapping with the serine cycle in key intermediate metabolites nodes (such as 2-O-Phosphono-D-glyceric acid (2PGA) and phosphoenolpyruvate (PEP). Some of the C3 and C4 metabolites were involved in more than one pathway in this interconnected set of cycles. For instance, 1) 2PGA is the branch node for the serine cycle and gluconeogenesis, 2) PEP and pyruvate are both in the serine cycle and anaplerotic pathways, 3) Malate and OAA were both in the serine cycle, TCA cycle and anaplerotic pathway. This complexity and redundancy offers the cell alternatives for the connections between the serine cycle and EMC pathway under environmental perturbation and gene manipulation.

Methanol uptake rates and biomass yields were measured for individual conditions. Extracellular metabolite secretion was measured using 1H NMR with a previously described protocol [11]. For *in silico* simulation, Influx_s was used [36].

Tables

Table 2.1 Growth rate and biomass yield for Δ*dme*, Δ*pck,* and Δ*pyk* mutants of both strain variants, base cases included, marked as *

Strain + medium condition	Biomass yield (g cdw/mol methanol consumed)	Absolute change in biomass yield (g cdw/mol methanol consumed)	% of original strain biomass yield under same condition	Growth rate (per hr)	Absolute change in growth rate (per hr)	% of original strain growth under same condition
LL + HY	8.90 ± 0.59	*	*	0.09 ± 0.01	*	*
VL + HY	6.44 ± 0.90	*	*	0.13 ± 0.01	*	*
LL + HYC	6.09 ± 0.04	*	*	0.13 ± 0.01	*	*
VL + HYC	4.38 ± 0.31	*	*	0.15 ± 0.01	*	*
LL: Δ*pck* + HY	4.84 ± 0.98	-4.07	54 % ± 11%	0.10 ± 0.01	0.01	111% ± 12%
VL: Δ*pck* + HY	5.26 ± 0.52	-1.17	82% ± 8%	0.10 ± 0.01	-0.03	77% ± 8%
LL: Δ*pck* + HYC	4.76 ± 1.67	-1.33	78% ± 27%	0.09 ± 0.01	-0.04	75% ± 10%
VL: Δ*pck* + HYC	5.06 ± 0.49	0.68	115% ± 11%	0.09 ± 0.01	-0.06	60% ± 8%
LL: Δ*dme* + HY	4.07 ± 2.06	-4.83	46% ± 23%	0.09 ± 0.01	0	100% ± 10%
VL: Δ*dme* + HY	4.62 ± 1.02	-1.81	72% ± 16%	0.09 ± 0.01	-0.04	69% ± 8%
LL: Δ*dme* + HYC	5.92 ± 0.37	-0.16	97% ± 6%	0.07 ± 0.01	-0.06	58% ± 10%
VL: Δ*dme* + HYC	6.01 ± 0.76	1.63	137% ± 17%	0.10 ± 0.01	-0.05	67% ± 7%
LL: Δ*pyk* + HY	7.77 ± 0.45	-1.13	87% ± 6%	0.10 ± 0.01	0.01	108% ± 6%
VL: Δ*pyk* + HY	6.30 ± 0.17	-0.14	98% ± 4%	0.06 ± 0.01	-0.07	44.9% ± 6%
LL: Δ*pyk* + HYC	7.60 ± 0.56	1.51	125% ± 8%	0.11 ± 0.01	-0.02	86% ± 10%
VL: Δ*pyk* + HYC	6.63 ± 1.14	2.25	151% ± 17%	0.10 ± 0.01	-0.05	63% ± 10%

Table 2.2 Strains and plasmid

Strain	Plasmid/Genotype *	Reference
LL	Rif derivative	M.Lidstrom
VL	Rif derivative	J.Vorholt
FYF1	LL::Δ*dme*	This study
FYF2	VL:Δ*dme*	This study
FYF3	LL::Δ*pck*	This study
FYF4	VL::Δ*pck*	This study
Plasmids	Description	Reference
pCM184	Ap^r, Kn^r,Tc^r; pCM182 with *kan* from pCM183; allelic exchange vector	41

Table 2.3 gene sequences for primers used for PCR amplification in this study

Primer name	Primer description	Sequence
YFP1	dme upstream flank forward	cgcacatttccccgaaaagtgccacctgacgtctagatctatccaccgcatcaccgtctc
YFP2	dme upstream flank backward	tggcataacttcgtataatgtatgctatacgaagttatgggcactgttgaggcatctgtt
YFP3	dme downstream flank forward	ctcagcgataacttcgtatagcatacattatacgaagttagcattcagttccgggcgttg
YFP4	dme downstream flank backward	gcgtggaaagcctggtcggctggatcctctagtgagctcgcccttcttctccttggtcca
YFP5	pck upstream flank forward	cgcacatttccccgaaaagtgccacctgacgtctagatctgatcggacacgtcctccacc
YFP6	pck upstream flank backward	acgtggcataacttcgtataatgtatgctatacgaagttatggggtgttcctccctcacg
YFP7	pck downstream flank forward	gacctcagcgataacttcgtatagcatacattatacgaagttagcggtcacagcccctcc
YFP8	pck downstream flank backward	cgtggaaagcctggtcggctggatcctctagtgagctccgagccgctctacaatctcgac
YFP9	pCM184 cre-km forward	ataacttcgtatagcatacattatacgaagttatgccacgttgtgtctcaaaatctctg
YFP10	pCM184 km-cre backward	taacttcgtataatgtatgctatacgaagttatcgctgaggtctgcctcgtg
YFP11	pCM184 backbone forward	gagctcactagaggatccagccg
YFP12	pCM184 Backbone backward	agatctagacgtcaggtggcacttttc

Figures

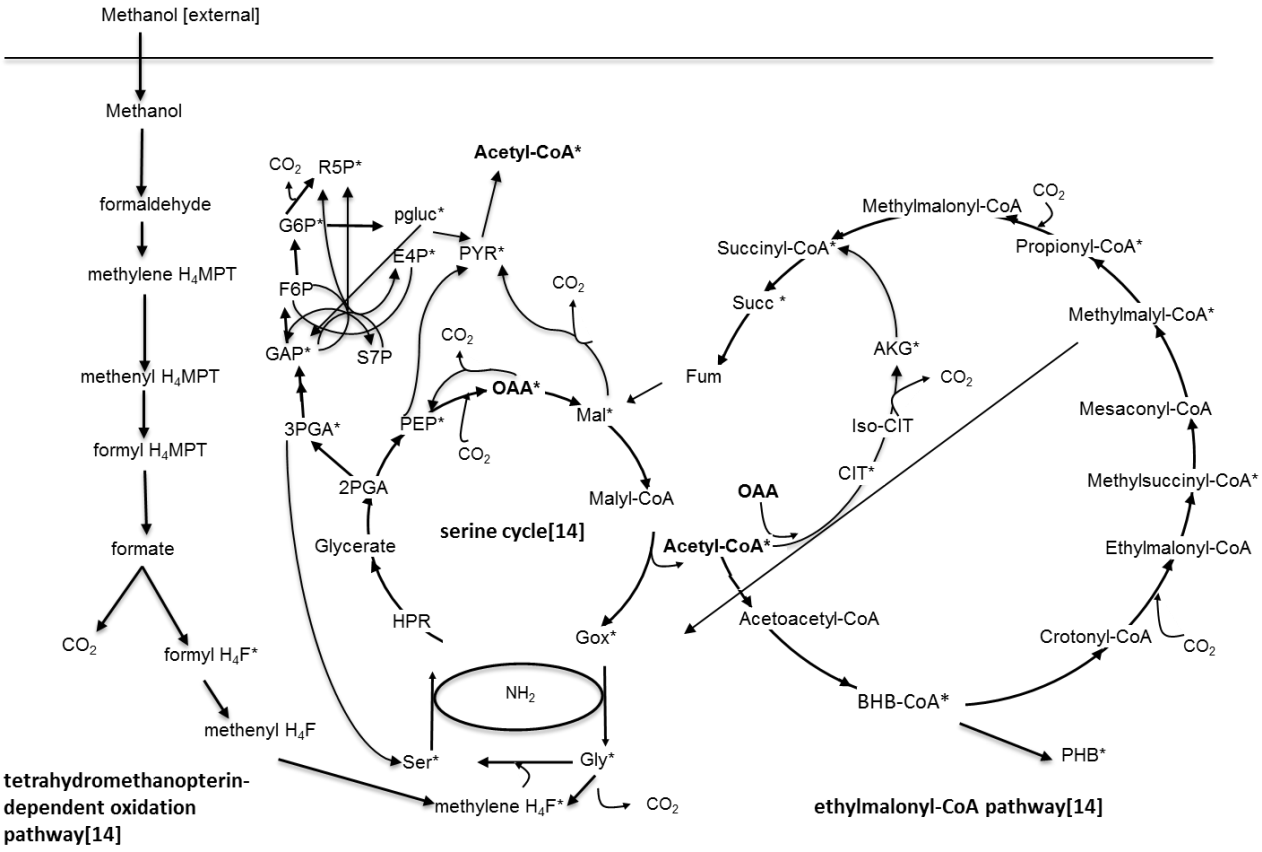
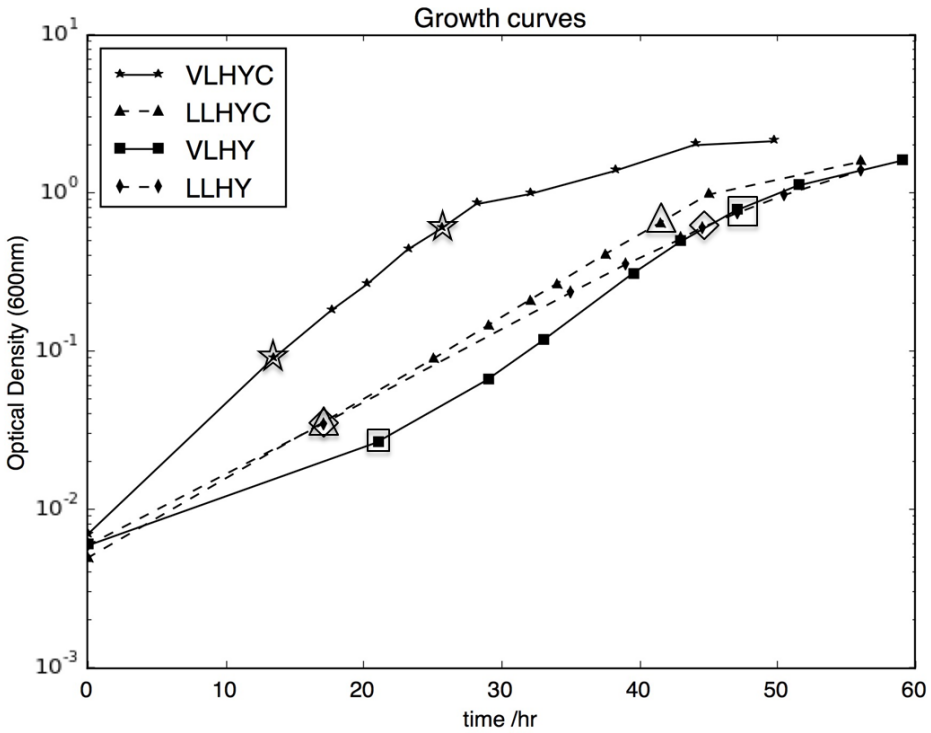


Figure 2.1. Central carbon metabolism model for *M. extorquens* AM1 methylotrophic growth. Metabolites with * are precursors for biomass. Metabolites in bold are branch points. The model includes 114 reactions with 9 reversible reactions and 2 scramble reactions. 71 intracellular metabolites are included in the model. Methanol is oxidized to formate via the H_4MPT pathway. Part of the formate pool is converted to CO_2 by formate dehydrogenase. The other part is converted into methylene H_4F via the H_4F pathway, entering the serine cycle. The serine cycle is the main assimilation pathway, with the EMC pathway regenerating glyoxylate in an anaplerotic function. Portions of the TCA cycle, gluconeogenesis, and the pentose-phosphate cycle were also operating to provide intracellular metabolites for biomass.



Strains+ conditions	Growth rate per hr	p value		Biomass yield g.cdw/mol	p value	
LLHY	0.092±0.006	1.23E-08	5.38E-06	8.90 ± 0.59	0.016	0.008
VLHY	0.132±0.008			6.44 ± 0.90		
LLHYC	0.125±0.009	1.08E-06	4.65E-04	6.09 ± 0.04	0.005	0.02
VLHYC	0.151±0.008			4.38 ± 0.31		

Figure 2.2. Representative growth curves for two strain variants with different concentration of cobalt supplement. Time points for supernatant methanol concentration measurements were highlighted by shapes corresponding to each condition. Growth rate and biomass yield were reported for each condition with p values from student t-test with n>=3.

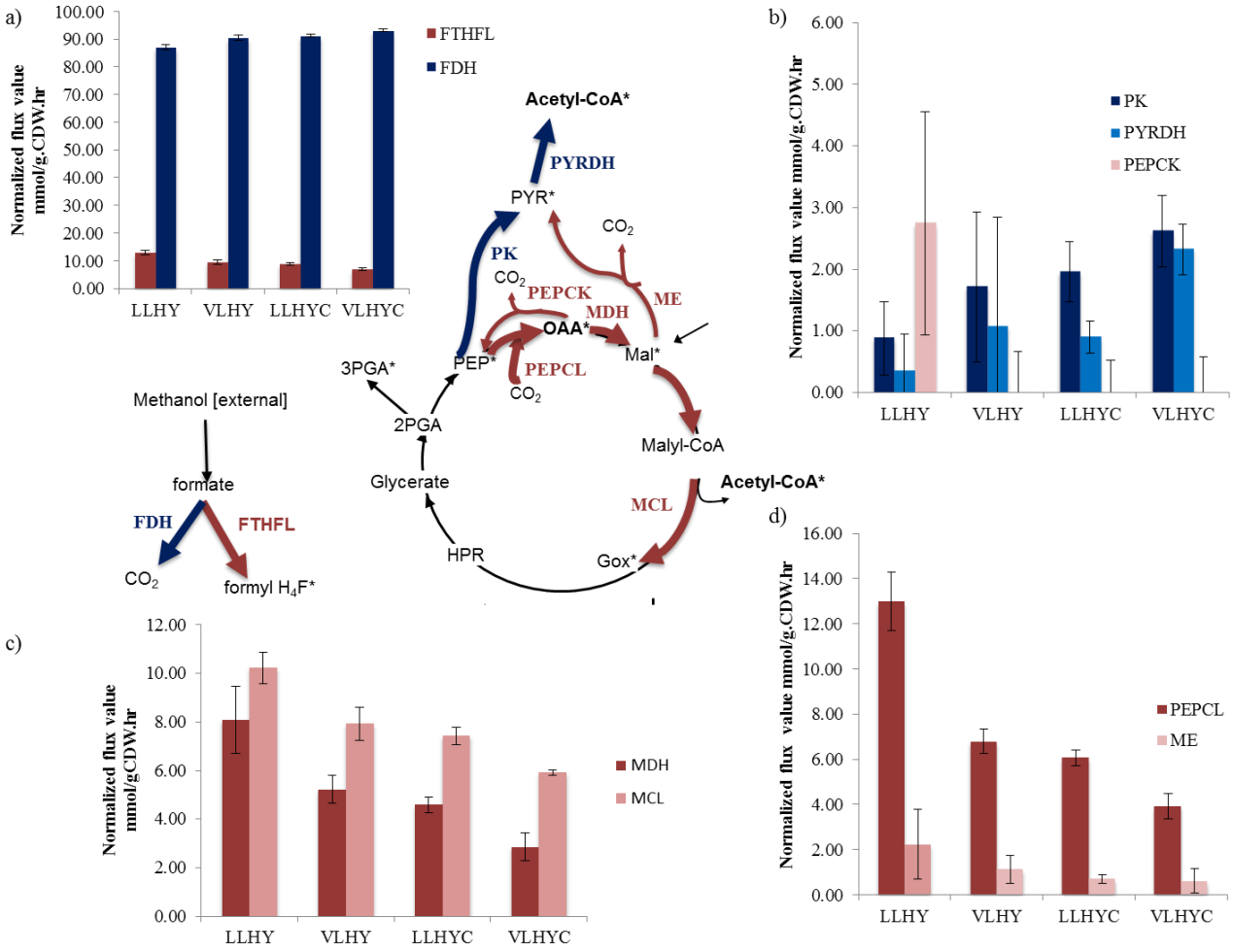


Figure 2.3. Flux distribution map showing flux values normalized to the methanol uptake rate. The major differences are highlighted in colors. In green, the increased fluxes correlated with higher growth rate; in red and pink, the increased fluxes correlated with higher biomass yield. a) flux partition at the formate node; b) fluxes from PEP to AcCoA using PK and PYRDH pathway, and OAA to PEP with PEPCK. c) flux through MDH and MCL. d) flux through PEPCL and ME. Data represent the averages from 3 biological replicates ± standard deviations.

Supplementary materials

Table S2_1 Genomic differences between different *M. extorquens* AM1 strains. Genome of the Marx laboratory strain has been published [55]

Supporting file 2_1: central carbon metabolism model for *M. extorquens* AM1 and list of metabolites with full name

Supporting file 2_2: three replicates of isotopomer labeling pattern of metabolites used in the simulation

Supporting file 2_3: detailed table for simulated flux with absolute values for all four conditions with 3 replicates each

Supporting file 2_4: detailed table for simulated flux distribution with relative values to specific methanol uptake rate for all four conditions with 3 replicates each

Chapter 3: Oxidative TCA cycle operating during methanotrophic growth of the Type I methanotroph Methylomicrobium buryatense 5GB1

Abstract

Methanotrophs are a group of bacteria that use methane as sole carbon and energy source. Type I methanotrophs are gamma-proteobacterial methanotrophs using the ribulose monophosphate cycle (RuMP) cycle for methane assimilation. In order to facilitate metabolic engineering in the industrially promising Type I methanotroph *Methylomicrobium buryatense* 5GB1, flux analysis of cellular metabolism is needed and ^{13}C tracer analysis is a foundational tool for such work. This biological system has a single-carbon input and a special network topology that together pose challenges to the current well-established methodology for ^{13}C tracer analysis using a multi-carbon input such as glucose, and to date, no ^{13}C tracer analysis of flux in a Type I methanotroph has been reported. In this study, we showed that by monitoring labeling patterns of several key intermediate metabolites in core metabolism, it is possible to quantitate the relative flux ratios for important branch points, such as the malate node. In addition, it is possible to assess the operation of the TCA cycle, which has been thought to be incomplete in Type I methanotrophs. Surprisingly, our analysis provides direct evidence of a complete, oxidative TCA cycle operating in *M. buryatense* 5GB1 using methane as the carbon and energy substrate, contributing about 45% of the total flux for *de novo* malate production. Combined with mutant analysis, this method was able to identify *fumA* (METBUDRAFT_1453/MBURv2__60244) as the primary fumarase involved in the oxidative TCA cycle, among 2 predicted fumarases, supported by ^{13}C tracer analysis on both *fumA* and *fumC* single knockouts. Interrupting the oxidative TCA cycle leads to a severe growth defect, suggesting that the oxidative TCA cycle functions to not only provide precursors for *de novo* biomass synthesis, but also to provide reducing power to the system. This information provides new opportunities for metabolic engineering of *M. buryatense* for the production of industrially relevant products.

Introduction

Methane is the second most prevalent greenhouse gas, and is also a cheap, abundant and renewable carbon source[10, 60]. Converting methane to valuable products at an industrial scale using methanotrophs will require iterations of strain optimization for higher yield and efficiency. *M. buryatense* 5GB1 is a robust and fast-growing Type I (gamma-proteobacterial) methanotroph, which is also highly resistant to contamination due to the high pH and salt in optimal culture conditions[61]. *M. buryatense* 5GB1 is a good candidate for the industrial scale application of methane bioconversion to valuable chemicals[4], with a series of developed genetic manipulation tools[62, 63] as well as bioreactor performance datasets[64] available. The whole genome of *M. buryatense* 5GB1 was also annotated and published[28]. However, in order to facilitate the process of metabolic engineering of *M. buryatense* 5GB1, quantitative characterization of cellular metabolism using ^{13}C tracer analysis needs to be developed. Stable isotope tracer experiments along with flux estimation and analysis is a key component of the metabolic engineering toolbox[65–67]. This approach has not yet been developed and used in Type I methanotrophs to quantitatively describe the metabolic phenotype of cells. Steady state ^{13}C flux analysis is well-established in other organisms with multi-carbon inputs[30–32]. However, the details of the network topology and the single carbon input of methanotrophs results in only partial flux elucidation from steady state isotope tracer experiments. The predicted core metabolic pathways operating in *M. buryatense* 5GB1 summarized in Figure 3.1 have been identified through both genome sequence as well as transcriptome analysis[68]. Methane is first oxidized to methanol through pMMO (particulate methane monooxygenase), and then converted into formaldehyde via methanol dehydrogenase. Formaldehyde is the first branch point, where it can either enter assimilation using the ribulose monophosphate cycle (RuMP cycle) or be further oxidized to formate and CO_2 to produce NADH. There are two variants of the RuMP cycle in *M. buryatense* 5GB1, the Embden-Meyerhof-Parnas (EMP) variant as well as the Entner–Doudoroff (ED) variant, which convert formaldehyde into C3 intermediate metabolites, such as pyruvate and phosphoenolpyruvate (PEP). Downstream of the pyruvate node, a series of multi-carbon compound inter-conversions lead to carbon precursors for biosynthesis, such as oxaloacetate (OAA), 2-oxoglutarate and acetyl-CoA (AcCoA).

One essential foundation of steady state ^{13}C tracer analysis is differentiated carbon rearrangements resulting from different metabolic pathways. However, this condition is not met in the *M. buryatense* 5GB1 metabolic network topology. The net outcome of the RuMP cycle is condensing 3 molecules of formaldehyde into 1 molecule of C3 compounds such as pyruvate. In this case, all intermediate metabolites will become fully labeled with the ^{13}C input. The two pathway variants both lead to fully labeled product, which makes the flux partition between EMP and ED variants unresolvable by steady state ^{13}C tracer analysis. However, carboxylation and decarboxylation steps downstream of the pyruvate and PEP nodes result in pathways that lead to differentiated labeling patterns. Several carboxylation reactions and decarboxylation reactions are involved in anaplerotic reactions as well as the TCA cycle. The carbon skeleton is also rearranged through malyl-CoA lyase reactions. Table 3.1 summarizes all the carbon atom transitions for the sub-network downstream of the PEP node. These transitions create the opportunity to determine fluxes through the steady state isotopomer labeling patterns. However, several parallel reactions (R_pyc, R_me, and R_mdh) with identical atom transitions exist in the network, which places more challenges on resolving the flux between those reactions. Here we use mutant analysis coupled with ^{13}C tracer analysis to solve this challenge.

One important question this approach can address is whether *M. buryatense* 5GB1 operates an incomplete TCA cycle when growing on methane. An incomplete TCA cycle has been proposed for Type I methanotrophs including *M. buryatense*, based on the inability to detect 2-oxoglutarate dehydrogenase enzyme activity as well as radioactive tracer studies[69–72]. From these results, it has been assumed that in Type I methanotrophs, the major function of the TCA cycle is to provide precursors for biomass[1]. However, no direct evidence of intermediate metabolite labeling patterns has been provided to support this conclusion. For metabolic modeling, it is important to determine flux through the branches of the TCA cycle and whether the TCA cycle is incomplete, as a baseline for accurate predictions.

In this study, we applied steady state ^{13}C metabolic flux analysis to *M. buryatense* 5GB1, focusing on reactions downstream of the pyruvate and phosphoenolpyruvate (PEP) nodes. First, we obtained the surprising evidence that the TCA cycle is complete during growth on methane, operating in the full oxidative mode, and we quantitatively determined the relative flux ratio of the TCA cycle. Second, we were able to disrupt the TCA cycle by mutating genes encoding fumarase, and then assessed phenotype

and labeling patterns. Third, using the same approach, the flux partition through the partial serine cycle present in *M. buryatense* 5GB1 was determined.

Results

3.1 Steady state labeling patterns of key metabolites

In order to initiate flux studies in *M. buryatense* 5GB1, labeling patterns of key metabolites were determined for cultures grown on methane. Cultures were exposed to ^{13}C methane for 5-6 generations, when the culture was assumed to reach both metabolic steady state as well as isotopic steady state. Mass isotopomer distributions (MIDs) of the metabolites listed in Table 3.2 were monitored using LC-MS/MS. As expected, the intermediate metabolites in the RuMP cycle were fully labeled, as shown for selected metabolites in Figure 3.1a (full list of labeled metabolites shown in Table S3_1). These results provided validation of isotopic steady state. Among the metabolites of the RuMP cycle, 6-phosphogluconate (6PG) had an unexpected peak at M+4, which was further identified as interference from the sample using Q-TOF (MS spectra of Q-tof data shown in supplemental materials Figure S3_1). C4 metabolites downstream of the pyruvate node such as malate and citrate were partially labeled in a specific pattern of adjacent peaks of two isotopomers with 1 mass shift, illustrated in Figure 3.1. This pattern was identified as a carboxylation signature pattern shown in Figure 3.1, consistent with a mixed labeled CO_2 pool (which contains a mixture of M0 and M1 with one mass shift).

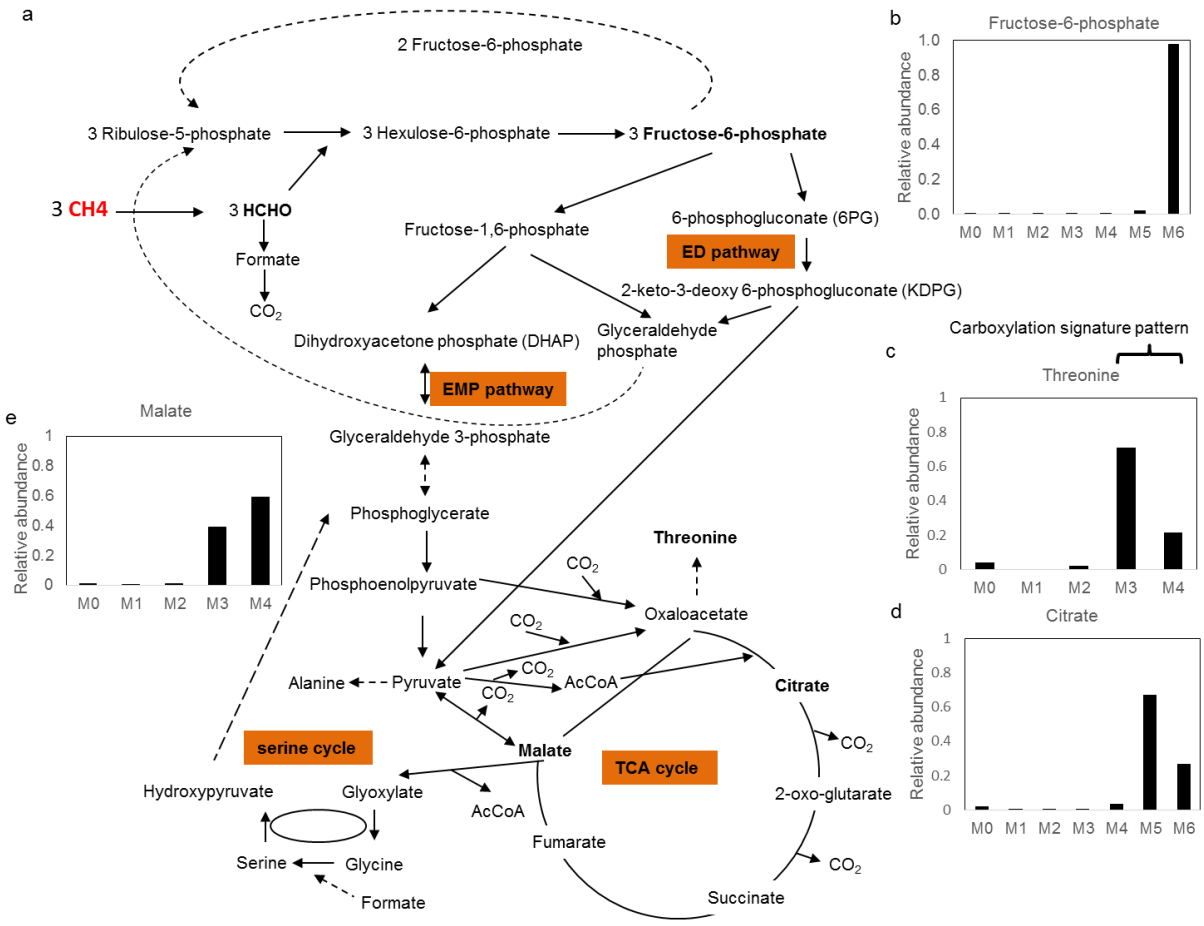


Figure 3.1 Central carbon metabolism pathways and steady state labeling patterns of key metabolites of M. buryatense 5GB1 with methanotrophic growth. Carboxylation signature is also shown in mass isotopomer distribution (MID) of malate, threonine and citric acid. 3 molecules of methane are converted into 1 molecule of a C3 compound through ribulose monophosphate (RuMP) cycle assimilation. Metabolite highlighted in red is the ^{13}C tracer substrate, metabolites in bold are measured using LC-MS/MS for MIDs. In WT, the MIDs of malate (Figure e) and OAA (Figure c) suggest a complete TCA cycle while the citrate (Figure d) labeling pattern suggests minor activity through malyl-CoA lyase and the serine cycle. Metabolites in the RuMP cycle are fully labeled, shown representatively using F6P (Figure b)

Pyruvate could not be detected in the cell extracts, likely due to a low pool and/or inefficient ionization. In addition, oxaloacetate (OAA) could not be detected due to its instability[73]. Therefore, intracellular threonine was measured to reflect labeling of OAA, and intracellular alanine was measured to reflect the labeling of pyruvate. Based on the alanine labeling pattern, pyruvate is fully labeled. The threonine and

citrate labeling patterns were similar to each other, both containing the carboxylation signature as shown in Figure 3.1b.

The mass isotopomer distribution (MID) of acetyl-CoA could be calculated through the ^{13}C enrichment of OAA and citrate. The consistency of mass isotopomer distribution (MID) between OAA and citrate suggested that acetyl-CoA is fully labeled with 2 mass shifts in the carbon backbone, which is an indication of negligible activity of malyl-CoA lyase. If acetyl-CoA were generated from malate via malyl-CoA, it would result in an acetyl-CoA pool that was partially labeled, changing the labeling pattern for citrate. This result predicts that the partial serine cycle present in *M. buryatense* 5GB1 (Figure 3.1) has minor flux during growth on methane. This prediction was tested by mutant analysis (see section 3.3 below). A significant difference in the MID of malate (shown in Figure 3.1c) compared to OAA (shown in Figure 3.1b) was observed, suggesting the operation of a full oxidative TCA cycle contributing to malate *de novo* production. This suggestion was also tested by mutant analysis (see section 3.2 below).

In the culture medium for *M. buryatense* 5GB1, carbonate buffer is provided to maintain pH. With ^{13}C methane and ^{12}C carbonate buffer experiments, the intracellular pool of CO_2 is partially labeled, due to the fact that both $^{13}CO_2$ from methane (^{13}C labeled) oxidation and extracellular $^{12}CO_2$ (from ^{12}C carbonate buffer) will contribute to the pool. The partially labeled CO_2 incorporates into central metabolites through carboxylation steps. In order to estimate the contribution from methane oxidation and CO_2 from carbonate as well as to discover carboxylation steps in central metabolism, another experiment with ^{13}C carbonate buffer and ^{12}C methane was conducted. In this case, the CO_2 pool was 48± 4% labeled (from methane oxidation), and 52± 4% unlabeled (from carbonate). Meanwhile, both malate and citric acid showed a carboxylation signature pattern that agrees with the ^{13}C methane and ^{12}C carbonate experiment. Similarly, the results showed a complete oxidative TCA cycle, again consistent with the ^{13}C methane experiment. The full list of labeled metabolites is in Table S3_2. The metabolites in the RuMP cycle were not labeled, confirming that no carboxylation steps occur in the RuMP cycle or contribute significantly to these intermediates.

3.2. Complete oxidative TCA cycle in *M. buryatense* 5GB1 grown on methane and TCA cycle interruption by mutation

As noted above, the labeling difference between OAA and malate suggested that part of the malate pool was generated via the operation of a complete TCA cycle. Malate showed higher M+4 than threonine, which suggested a TCA cycle contribution through fumarate via fumarase. Citrate, on the other hand, which is synthesized from a condensation reaction between OAA and acetyl-CoA (AcCoA) showed a consistent labeling pattern with threonine, allowing the OAA labeling pattern to be calculated using either citrate or threonine. The flux distribution of the TCA cycle to malate was then determined quantitatively using the method described in the Method section. Based on the calculation, TCA cycle fluxes contribute 45% of the flux to malate.

As a further test that the observed labeling pattern was the result of a complete TCA cycle, we also interrupted the complete TCA cycle by mutating fumarase, which converts fumarate to malate. Two copies of genes encoding fumarase are predicted in the genome (*fumC*: METBUDRAFT_2764, *fumA*: METBUDRAFT_1453). Both single knockouts were made successfully with the method described previously[62]. Steady state ^{13}C tracer analysis was then performed on both constructs. As shown in Figure 3.2a, the *fumC* null mutant does not show significant differences in labeling of malate and threonine (which reflects OAA) compared to the wild type (WT). The inconsistency between malate and OAA labeling persists in the *fumC* mutant, again suggesting a complete TCA cycle with around 45% flux contribution to malate. No growth defect is observed in the *fumC* mutant. However, the *fumA* null mutant showed a different labeling pattern of malate compared to the WT (Figure 3.2a). Consistent labeling of malate and threonine was also observed, which suggested the TCA cycle was indeed interrupted in this construct. The relative flux contribution of the TCA cycle to malate is close to 0 in the *fumA* mutant based on the relative flux ratio calculation described in the Method section, as shown in Figure 3.2b. This suggests that under these growth conditions, *fumA* is the dominant gene in *M. buryatense* 5GB1 encoding fumarase. The *fumA* mutant has a severe growth defect, with a doubling time on methane of 17 hours (compared to 3 hours for the WT), further indicating the importance of a complete TCA cycle in growth of *M. buryatense* 5GB1 on methane.

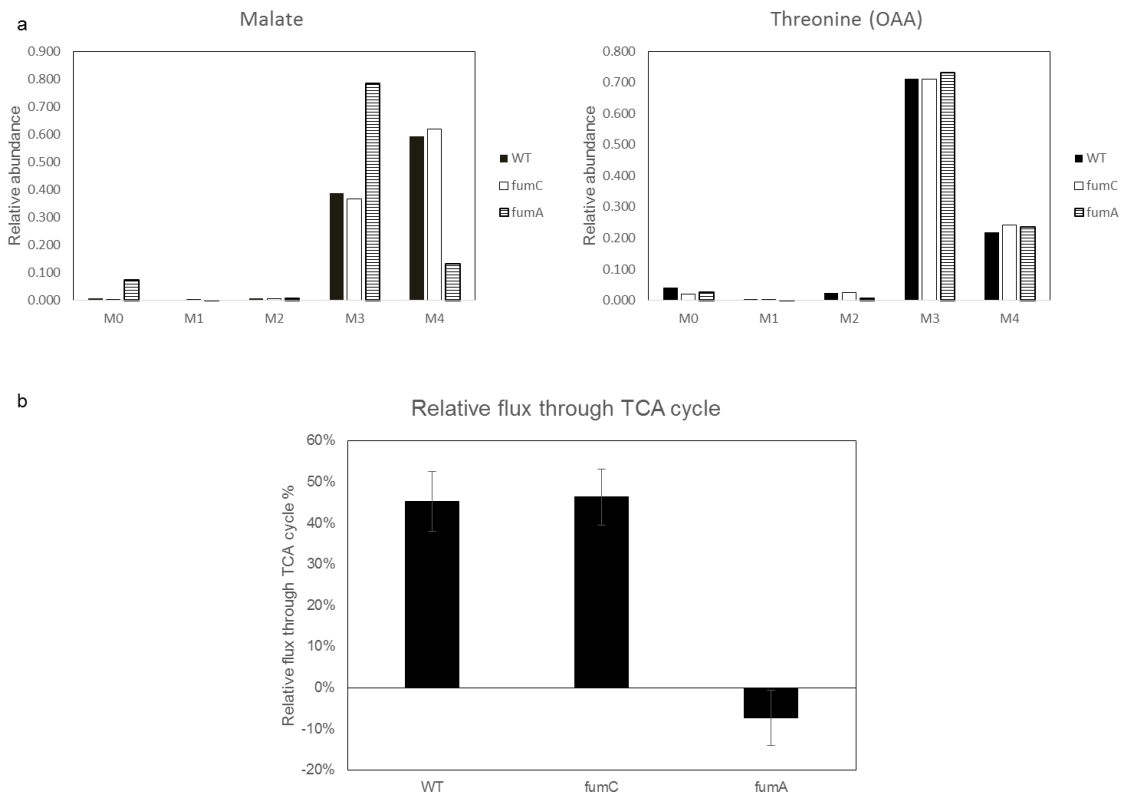


Figure 3.2. Labeling pattern of key metabolites (Figure 3.2a) and relative flux through TCA cycle in WT, ΔfumA mutant and ΔfumC mutant (Figure 3.2b). ΔfumA mutant showed higher M+3 of malate than WT and ΔfumC mutant, OAA showed same MID for all. The consistent MID between malate and OAA in the ΔfumA mutant suggested interrupted TCA cycle. Relative flux through TCA cycle was then quantitated, ΔfumA mutant showed interrupted TCA cycle flux.

In order for *M. buryatense* 5GB1 to operate a complete TCA cycle, it must have the ability to convert 2-oxoglutarate to succinylCoA or succinate, and the absence of detectable 2-oxoglutarate dehydrogenase activity has been used as part of the evidence for an incomplete TCA cycle in Type I methanotrophs[70, 71] . The genome sequence predicts three paths for such a conversion, including a classic 2-oxoglutarate dehydrogenase complex, a 2-oxoglutarate ferredoxin oxidoreductase, and a pathway through succinate semialdehyde (Figure 3.3a). Transcriptomic data show that all sets of genes are expressed at levels similar to other TCA cycle genes, such as citrate synthase[68]. A mutation was generated in the 2-oxoglutarate dehydrogenase complex E3 subunit (METBUDRAFT_0801) to remove the 2-oxoglutarate dehydrogenase activity, and unlike the *fumA* mutant, this mutant did not show a growth

phenotype. Also unlike the *fumA* mutant, the labeling pattern of malate and citrate in this strain still shows the inconsistency of heavily labeled isotopomers, as shown in Figure 3.4b, and supplemental material Table S3_1. The flux ratio of the TCA cycle was determined using the method described above, and found to be 51%. These results strongly suggest the TCA cycle is still functioning as an oxidative cycle in this mutant, further supporting our results suggesting that other pathways can substitute for the functionality of 2-oxoglutarate dehydrogenase in this bacterium.

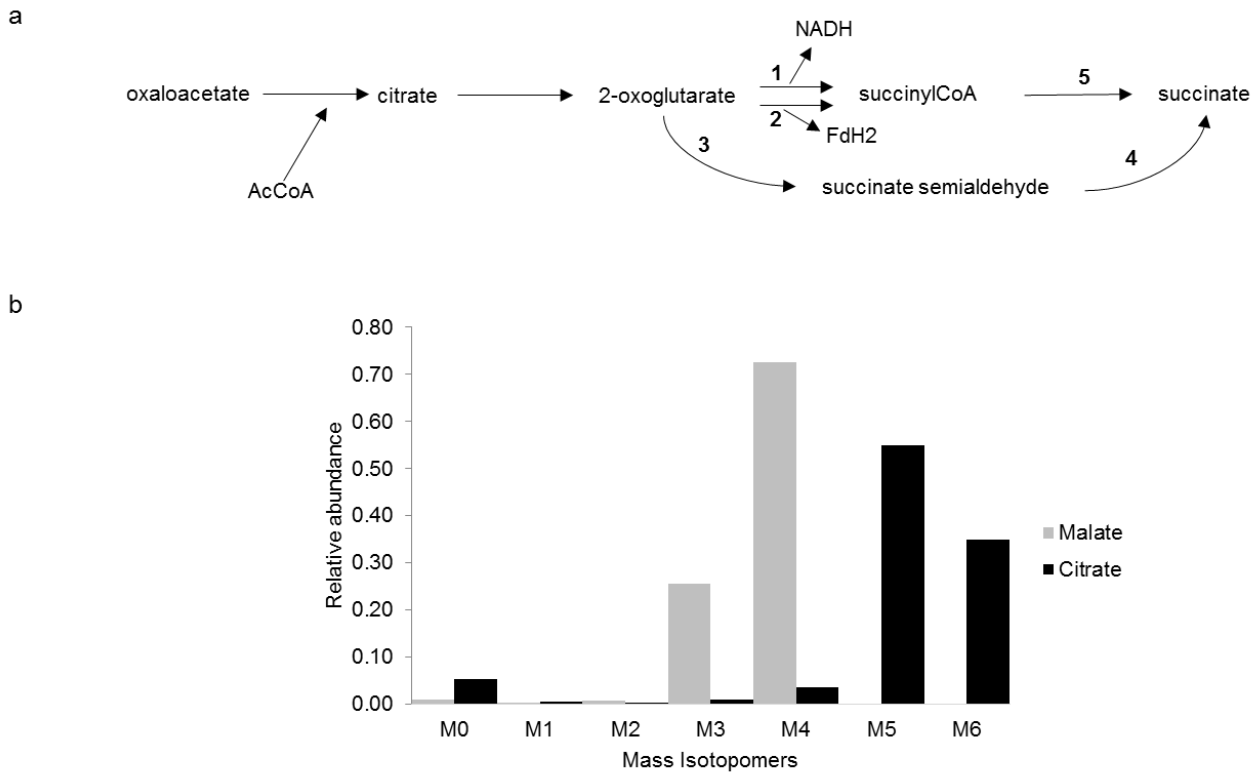


Figure 3.3. Three ways of converting 2-oxoglutarate to succinate in M. buryatense 5GB1 (Figure 3.3a). 1. Convert 2-oxo-glutarate into succinyl-CoA with production of NADH using the 2-oxoglutarate dehydrogenase complex. 2. Catalyze the same reaction generating FdH2, with 2-oxoglutarate ferredoxin oxidoreductase. 3. Convert 2-oxoglutarate into succinate semialdehyde with 2-oxo-glutarate decarboxylase, and convert succinate semialdehyde to succinate with succinate-semialdehyde dehydrogenase. Numbers on arrows are designation of genes and enzymes show as follows: 1: 2-oxoglutarate dehydrogenase complex METBUDRAFT_0802 (E2 subunit), METBUDRAFT_0803 (E1 subunit, decarboxylase), METBUDRAFT_0801 (E3 subunit). 2: 2-oxoglutarate ferredoxin oxidoreductase METBUDRAFT_3369 (alpha-Sub Unit),

METBUDRAFT_3370 (beta-sub unit) 3:2-oxoglutarate decarboxylase METBUDRAFT_3235. 4: succinate-semialdehyde dehydrogenase METBUDRAFT_0798, METBUDRAFT_3202. 5: succinyl-CoA synthetase complex METBUDRAFT_2196 (succinyl-CoAsynthetase NAD(P)-binding, alpha subunit), METBUDRAFT_2197 (succinyl-CoA synthetase, beta subunit), METBUDRAFT_2277 (succinyl-CoA synthetase, NAD(P)-binding, alpha subunit), METBUDRAFT_2278 (succinyl-CoA synthetase, beta subunit). In the Δlpd strain, MIDs of citrate and malate were similar to WT, which suggested that TCA cycle flux is still contributing to malate production.

3.3. *sga* mutants suggest negligible flux through serine cycle.

M. buryatense 5GB1 contains a set of genes predicted to encode a partial serine cycle that converts formate plus CO_2 to AcCoA (Figure 3.1a). Serine glyoxylate aminotransferase (*sga*: METBUDRAFT_2200/MBURv2_130302) catalyzes the entry step of the serine cycle, converting glyoxylate into glycine, and serine into hydroxypyruvate. This gene is expressed at a level similar to other genes in core metabolism (such as pyruvate dehydrogenase) in transcriptomic data[68]. As noted above, steady state ^{13}C analysis of the WT suggested minor flux through this partial serine cycle. Therefore, to confirm this hypothesis, *sga* was also mutated in order to assess flux through the serine cycle and stationary ^{13}C tracer analysis was conducted on this mutant with ^{13}C methane. The labeling patterns for the key metabolites in this condition were consistent with the WT (Figure 3.4a; supplemental material Table S3_1). The flux ratio at the malate node was then calculated and found to also not be different from WT (Figure 3.4b). In addition, this mutant does not have a growth defect. These results allow us to conclude that serine cycle flux through *sga* is negligible under methane growth condition.

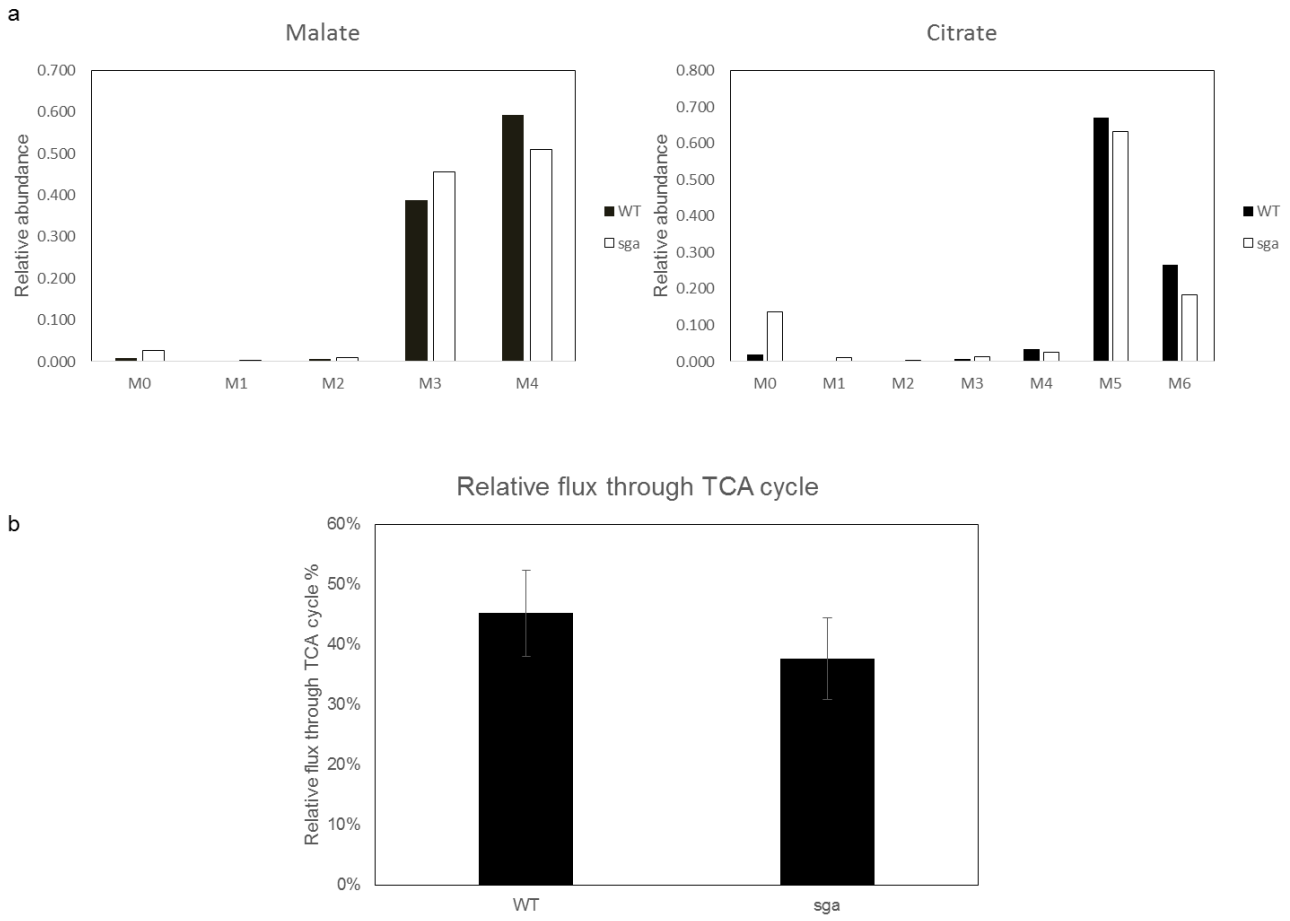


Figure 3.4 MIDs of key metabolites, relative flux to malate node (Figure 3a), and relative flux through TCA cycle (Figure 3b) in Δsga mutant. No difference in MID between WT and Δsga mutant, slight difference in TCA cycle contribution. This validates minor flux through serine cycle in WT.

Discussion

In order to carry out robust metabolic engineering in methanotrophs, it is important to define which metabolic pathways are active in these organisms. We report here an approach to steady state ^{13}C analysis in a Type I methanotroph, as a step towards a greater understanding of the functional metabolic pathways in methanotrophs. As predicted by theoretical simulation, the flux distribution for the two variants of the RuMP cycle (ED and EMP) could not be distinguished using this technique. Kinetic ^{13}C labeling analysis[37, 39] will be required for these determinations. Fortunately, the presence of a partially labeled CO_2 pool as well as carbon rearrangements in C3-C4 interconversion pathways allows flux

through these pathways to be distinguished. The metabolic network includes multiple predicted routes for interconversion of C3-C4 compounds, most of which are reversible reactions. These fluxes could be further distinguished through steady state flux analysis coupled with mutant analysis.

One unresolved question of central metabolism in Type I methanotrophs is whether the TCA cycle is complete or not, since it is commonly reported that no 2-oxoglutarate dehydrogenase activity is detected[1, 69–72]. In this study, we discovered the surprising result that in *M. buryatense* 5GB1, the TCA cycle operates as an oxidative cycle and confirmed the labeling results with a fumarase mutant. We also identified that *fumA* is the likely the dominant fumarase enzyme under these growth conditions. In addition, the genome sequence predicts three separate pathways for converting 2-oxoglutarate to succinylCoA or succinate, including a classic 2-oxoglutarate dehydrogenase. All of these pathways offer the possibility of a complete TCA cycle, and would be predicted to provide network robustness. The role of each pathway is not known, but a mutant in the 2-oxoglutarate dehydrogenase E3 subunit (METBUDRAFT_0801/MBURv2_50047) did not show a growth phenotype and the labeling pattern was similar to WT, suggesting that at least one of the other pathways is functional. Since the oxidative TCA cycle is predicted to generate 45% of the malate pool and the fumarase A mutant that disrupts the oxidative TCA cycle had a severe growth defect, our results suggest that the TCA cycle functions both in the generation of precursors for biomass synthesis and in the generation of reducing power.

From previously published genome-scale modeling of *M. buryatense* 5GB1 under methane growth, a partial serine cycle was predicted to contribute to *de novo* acetyl-CoA synthesis with Δpdh and Δxfp knockouts[68], but carries no flux in WT. This prediction matches with our ^{13}C tracer analysis results of WT as well as the *Δsga* mutant.

The presence of a complete TCA cycle in this methanotroph is an important new finding, which opens up new opportunities for metabolic engineering. Pyruvate and AcCoA are key precursors for valuable chemical products that might be generated from methane using *M. buryatense* 5GB1 through metabolic engineering[74], so this increased understanding of how pyruvate and AcCoA are metabolized is important for such manipulations. In addition, since the metabolic network is more similar to *E. coli* than previously thought, it suggests the possibility of ‘drop-in’ modules designed for production of valuable

compounds in *E. coli*. Our results also provide additional constraints for incorporation into FBA models using the method previously described[75]. Our demonstration of an approach for steady state flux analysis in a Type I methanotroph together with the new insights gained into the central metabolism in such a bacterium set the stage for more robust metabolic engineering approaches for conversion of methane to value-added products.

Conclusion

This study initiated flux estimates in *M. buryatense* 5GB1, an industrially promising candidate for bioconversion of methane into value-added chemicals, providing informative insights into how core metabolism operates under tested environmental conditions. The addition of flux estimates into the metabolic engineering toolbox for this bacterium will accelerate the metabolic engineering process. The surprising finding of an oxidative TCA cycle operating in this Type I methanotroph opens more possibilities for metabolic engineering targets to improve process characteristics.

Materials and Methods

3.4. Cell culture

NMS2 medium as previous described[28, 76] was used for both liquid and plate culture. The following antibiotics were used for phenotypic characterization for knockout strains: kanamycin (Km), 50 ug/ml; zeocin (Zeo), 30ug/ml.

For growth curve experiments, liquid pre-cultures were grown in 25 ml tubes with an atmosphere of 25% methane in air at 30C, shaken at 200 rpm, sealed with rubber stoppers and aluminum seals (Wheaton, Millville,NJ,USA). Pre-cultures were then inoculated in new 25 ml tubes with starting OD_{600} = 0.01 for growth rate experiments.

For isotope labeling experiments, precultures were grown in 25 ml tubes with an atmosphere of 25% methane (^{12}C methane or ^{13}C methane for strains in which a growth defect was observed) for 18 hours. The precultures were then inoculated into 50 ml fresh medium in 250 ml serum bottle (with starting

OD_{600}=0.01). Cell were harvested at 0.3 ~ 0.6 when the culture was under both isotopic steady state and metabolic steady state. All strains used in this study are listed in Table 3.3.

3.5. Genetic manipulation of 5GB1 using electroporation

All gene knockouts are constructed with assembled PCR products which were then electroporated into *M. buryatense* 5GB1 using the protocol described previously[62]. Primers for knockout constructs are listed in Table S3_3 in the supplemental material. Briefly, a construct containing the zeocin resistance gene flanked by FLP-FRT sites was assembled in the middle of flanks with 800-1000 base pair from the gene targeted for deletion using PCR. The Δ*sga* strain was constructed with kanamycin resistance gene flanks in the middle using the same procedure. The PCR mixture after assembly was then transferred into *M. buryatense* 5GB1 using electroporation. After electroporation, the cells were recovered in NMS2 with 80 rpm shaking overnight, and plated out on an agar plate containing zeocin or kanamycin. Colonies were screened using diagnostic colony PCR, and later confirmed using sequence (Genewiz).

3.6. MIDs of metabolites measurements using LC/MS-MS

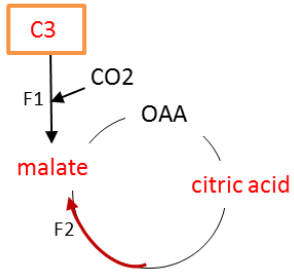
Sample collection and extraction were performed using the method described previously[77]. Briefly, cell cultures were quenched using fast filtration using 0.22 μm nylon membrane (0.22 μm, 47 mm from PALL life science, Port Washington, NY) once it reaches the desired stage, and saved in 50 ml falcon tubes submerged in liquid nitrogen. The collected samples were lyophilized for 12 hours to get rid of extra medium on filter. A hot water extraction protocol was used to extract intracellular metabolites. 20 ml of boiling water was added to 50 ml falcon tubes, the tubes were incubated in a boiling water bath for 20 minutes, and then placed on ice for 30 minutes. The cell extracts were centrifuged at 4°C, 5000rpm for 30 minutes to remove cell debris. The supernatant was transferred into new tubes and fast frozen using liquid nitrogen. The cell lysates were lyophilized and concentrated into 50 ul water. The reconstituted samples were centrifuged with filtering (spin-x costar, 0.22 μm, Sigma-Aldrich), then kept at -20°C until LC-MS/MS analysis. A Waters Xevo mass spectrometry triple quad (Xevo, Waters, Milford, MA) with UPLC system equipped with a Zic-pHILIC column (SeQuant, PEEK 150 mm length × 2.1 mm metal free, with 5 μm polymeric film thickness, EMD Millipore) was used for detection of MID of metabolites. Multiple

reaction monitors (MRM) were set up for each metabolite of interest. For each metabolite, ^{12}C chemical standards were used to set up the mass channel for unlabeled isotopomers. The predicted mass fragments were then used to predict the MRM for labeled isotopomers for each metabolite. This information is included in the supplemental material Table S3_4.

3.7. Mathematically determine relative flux distribution at malate node.

Malate will have a differentiated labeling pattern, depending on the pathways for synthesis. In the current understanding of central carbon metabolism, malate could be synthesized by three different paths illustrated in Figure 3.5. First, it could be produced via malic enzyme from pyruvate and CO_2. Second, it could be produced by malate dehydrogenase from OAA. Third, it could be produced through the TCA cycle from fumarate via fumarase. Route 1 and 2 lead to identical malate labeling patterns. Route 3 will produce fully labeled malate. The equations shown in Figure 3.5 were then developed to quantitate the relative flux at the malate node between route 3 and the other 2 routes with ^{13}C methane (Figure 3.5a) or ^{13}C carbonate as tracer (Figure 3.5b). The flux of carboxylation summarizes all the flux contributions to malate through CO_2 incorporation. The calculation was done within in-house developed scripts using python.

a. With ^{13}C methane, ^{12}C carbonate

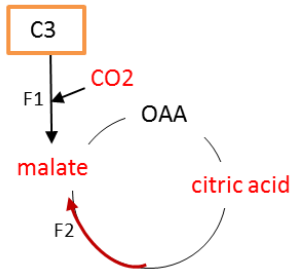


$$F1*M0_{(CO2)}+F2*0=M3_{(malate)}$$
$$F1*M1_{(CO2)}+F2*1=M4_{(malate)}$$

Matrix A | Matrix F | Matrix M

$$\begin{pmatrix} M0_{(CO2)} & 0 \\ M1_{(CO2)} & 1 \end{pmatrix} \begin{pmatrix} F1 \\ F2 \end{pmatrix} = \begin{pmatrix} M3_{(malate)} \\ M4_{(malate)} \end{pmatrix}$$

b. With ^{13}C carbonate, ^{12}C methane



$$F1*M0_{(CO2)}+F2*1=M0_{(malate)}$$
$$F1*M1_{(CO2)}+F2*0=M1_{(malate)}$$

$$\begin{pmatrix} M0_{(CO2)} & 1 \\ M1_{(CO2)} & 0 \end{pmatrix} \begin{pmatrix} F1 \\ F2 \end{pmatrix} = \begin{pmatrix} M0_{(malate)} \\ M1_{(malate)} \end{pmatrix}$$

Matrix F = Matrix A^{-1} Matrix M

Figure 3.5. Flux ratio quantification methodology. Figure 3.5a shows the equation for the condition with ^{13}C methane, ^{12}C carbonate, and Figure 3.5b shows the equation for the condition with ^{12}C methane, ^{13}C carbonate. Summary equation: Matrix A*Matrix F = Matrix M. The Matrix F is then further calculated using in-house script implemented in python.

Tables

Table 3.2 Carbon atom transition in sub-networks of core metabolism in M. buryatense 5GB1

Reaction name	Stoichiometry	Atom transition
R_pk	PEP -> PYR	ABC -> ABC
R_pyc	PYR + CO2 -> OAA	ABC + D -> ABCD
R_me	PYR + CO2 -> MAL	ABC + D -> ABCD
R_mdh	OAA -> MAL	ABCD -> ABCD
R_mcl	MAL -> GOX + ACCOA	ABCD -> AB + DC
R_pdh	PYR -> ACCOA + CO2	ABC -> BC + A
R_sga	GOX + SER -> HPR +GLY	AB + CDE -> CDE + AB
R_TCA1	ACCOA + OAA -> CIT	AB + CDEF -> FEDBAC
R_TCA2	CIT -> AKG + CO2	ABCDEF -> ABCDE + F
R_TCA3	AKG -> SUCC +CO2	ABCDE -> BCDE + A
R_TCA4	SUCC -> FUM	ABCD -> ABCD
R_fum	FUM -> MAL	ABCD -> ABCD

Table 3.3 List of metabolites monitored using LC-MS/MS

Metabolites name	Ionization mode
Succinate	negative
Malate	negative
Phosphoenolpyruvate (PEP)	negative
Citrate	negative
Ribulose-5-phosphate (R5P)	negative
Fructose 6 phosphate (F6P)	negative
Glucose 6 phosphate (G6P)	negative
6-phosphoglucnate (6PG)	negative
Sedoheptulose 7-phosphate (S7P)	negative
Fructose 1,6-biphosphate (FBP)	negative
Alanine	positive
Threonine	positive

Table 4.*3* Strains used in this study

Strain	Description	Reference or source
M. buryatense 5GB1C	Variant of *M. buryatense* 5GB1 intentionally cured of native plasmid by knocking out repA (METBUDRAFT_0076) to repB (METBUDRAFT_0086) native plasmid locus.	[63]
M. buryatense Δsga	*M. buryatense* 5GB1C (METBUDRAFT_2200)	this study
M. buryatense ΔfumA	*M. buryatense* 5GB1C (METBUDRAFT_1453)	this study
M. buryatense Δ*fumC*	*M. buryatense* 5GB1C (METBUDRAFT_2764)	this study
M. buryatense Δ*lpd*	*M. buryatense* 5GB1C (METBUDRAFT_0801)	this study

Supplemental materials

Table S3_1. Sample information and isotopomer labeling pattern of key intermediate metabolites with ^{13}C methane and ^{12}C carbohydrate buffer as substrate.

Table S3_2. Sample information and isotopomer labeling pattern of key intermediate metabolites with ^{12}C methane and ^{13}C carbohydrate buffer.

Table S3_3. Primers for knockout construction.

Table S3_4. Multiple reaction monitors (MRM) for key metabolites measured in this study.

Figure S3_1. Q-TOF MS spectra of 6PG showing that M+4 peak is an interference. The mz of 279.0176 is 50 ppm away from the theoretical mz of M+4 of 6PG based on the mass difference between ^{12}C and ^{13}C is 1.003355. The High resolution Q-TOF was calibrated and tuned at accuracy of 30 ppm, showing that the mz of 279.0176 is not the M+4 of 6PG.

Chapter 4: Core metabolism rearrangements in the methanotroph *Methylomicrobium buryatense* 5GB1 grown on methanol

Abstract

Methylomicrobium buryatense 5GB1 is an obligate methylotroph, which can grow on either methane or methanol with similar growth rates. Although the basic understanding of *M. buryatense* 5GB1 metabolism under methane growth is moving forward rapidly, little is known about metabolic flux during growth on methanol. In this study, both targeted metabolomics as well as ^{13}C tracer analysis have been undertaken to understand the core carbon metabolism in *M. buryatense* 5GB1 during methanol growth, to provide baseline metabolic performance characterization for future metabolic engineering procedures. Targeted metabolomics analysis were performed on both methane and methanol cultures to identify metabolic nodes with altered fluxes. Several key metabolites showed large differences in pool size. Notably, 2-keto-3-deoxy-6-phosphogluconate (KDPG) showed much larger pools under methanol growth, suggesting the Entner-Doudoroff (ED) pathway was active. Intermediates in other parts of metabolism also showed pool size changes under methanol growth. A systematic shift of active core metabolism is proposed to explain the pool changes. The ED Pathway was also manipulated by swapping the native promoter with a tunable promoter system to assess pool size changes correlated with pathway activity. In order to distinguish flux partition differences in the C3-C4 node, ^{13}C tracer analysis was also applied to methanol culture. These flux ratios were then incorporated into an FBA model for better prediction of global metabolism change. Taken together, these results show that the metabolism on methanol is substantially different from that on methane, with a shift towards flux through the ED pathway and serine cycle, and an incomplete TCA cycle balanced with increased flux through formate dehydrogenase.

Introduction

Methylomicrobium buryatense 5GB1 is an obligate type I methanotroph, which is an industrial-promising candidate for converting methane into valuable chemicals due to advantages of robust growth and high resistance to contamination[60, 74, 78]. The basic understanding of metabolism for *M. buryatense* 5GB1

growing on methane is moving forward rapidly[27, 64, 68, 79]. This obligate methylotroph can also grow on methanol with a growth rate similar to that on methane[4]. The expanded scope of usable substrates for *M. buryatense* 5GB1 provides more opportunities for metabolic engineering this strain with C1 substrates compared to many methanotrophs, which grow poorly if at all on methanol. Similarly to methane-based biotechnology, methanol-based biotechnology is also an exciting and fast-moving area[14, 23]. Currently, in methanol-based biotechnology, several categories of methylotrophs are used as model organisms. One example is *Methylobacterium extorquens* AM1[23], which use the serine cycle as the assimilation pathway for C1 units. This model organism has been thoroughly explored using both metabolomics and metabolic flux analysis[16, 44, 80]. In the category of methylotrophs with the ribulose monophosphate (RuMP) cycle for formaldehyde assimilation, *Bacillus methanolicus* MGA3 is facultative methylotroph with core metabolism that has been well-investigated[81, 82].

A full suite of genetic manipulation techniques [62, 63] as well as transferable metabolic engineering toolkits are available for *M. buryatense* 5GB1. Therefore, we propose that *M. buryatense* 5GB1 might be a candidate for methanol-based biotechnology, if similar information about metabolism were available. Little is currently known about growth on methanol. In a set of bioreactor experiments, it was determined that compared to methane-grown cells, methanol-grown cells had a high glycogen content (43% of cell dry weight) and a high level of formate excretion (12 mM $gCDW^{-1}$)[4]. These results suggested that the metabolic network during growth on methanol might differ from that during growth on methane.

The known core metabolism network of *M. buryatense* 5GB1 during methane growth could support growth on methanol (Fig. 4.1). Methanol is converted into formaldehyde, which is further oxidized to formate and then CO_2 to provide reducing equivalents as NADH. The formaldehyde could also be assimilated via the RuMP cycle. Two variants of the RuMP cycle are possible, the EMP variant and the ED variant. In the EMP variant, fructose-6-phosphate (F6P) is converted to fructose 1, 6-biphosphate (FBP) and split into C3 sugar phosphate via fructose-bisphosphate aldolase. In the ED pathway, 6-phosphogluconate (6PG) is converted into 2-keto-3-deoxy 6-phosphogluconate (KDPG), and then enters the node of pyruvate directly via 2-keto-3-deoxy-6-phosphogluconate aldolase. It has recently been shown that this methanotroph operates a complete TCA cycle during growth on methane, and operates a set of C3/C4 interconversions[83].

In order to determine the metabolic network involved in growth on methanol, both metabolic profiling and ^{13}C tracer analysis approaches have been used. These approaches provide a direct view of cellular behavior resulting from the integrative effect on gene expression and regulation, which provides critical information for metabolic engineering. In this study, targeted metabolite profiling was performed for core metabolites under both methane and methanol growth to discover and quantitate which part of the network is responding to the substrate change. Meanwhile, ^{13}C tracer analysis was also conducted to determine whether the organism changes strategies for interconversion of C2-C3-C4 compounds downstream of the pyruvate node.

Constraint-based analysis of genome-scale metabolic network (COBRA analysis[84, 85]) is flux balance analysis (FBA) based on stoichiometric equations describing a genome scale model network, which predicts flux distribution to achieve optimization of a specified objective function. It is a complementary approach to ^{13}C-based flux analysis, since the network covered in FBA analysis is much broader. However, the assumption that the network operates to optimize around a single, specified objective goal as well as the fact that multiple possibilities could lead to same value of the objective function, often result in a mismatch between the predicted flux distribution and the experimentally validated ^{13}C MFA, leading to less accurate prediction. In this study, the data from the ^{13}C analysis were incorporated into a genome scale model as additional constraints, to better predict flux distribution of central metabolism during methanol growth, providing the basis for future rational strain design.

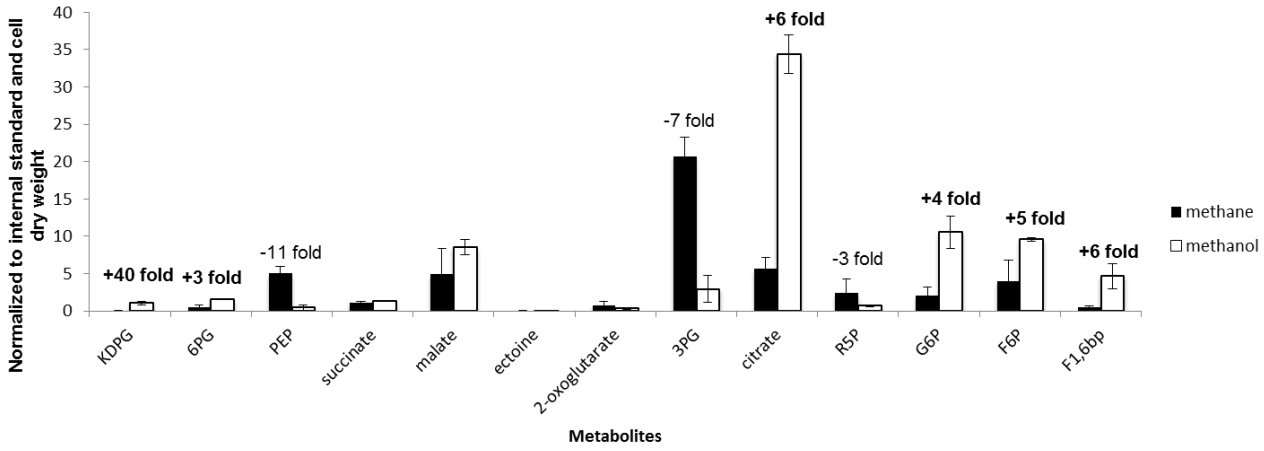
Results

4.1. Targeted metabolomics analysis showed systematical flux distribution shifts under methanol growth

Changes in the intracellular level of metabolites indicate metabolite nodes with altered flux. Targeted metabolomics analysis was performed for both methane-grown and methanol-grown cultures from the exponential growth phase to locate the metabolite nodes that respond to the two substrates. Intermediate metabolites were extracted and quantitated under both conditions. The levels of metabolites were normalized to both internal standard ($^{13}C_4$ succinate) as well as cell biomass. As seen in Figure 4.2a, the following metabolites showed pool size increase in methanol cultures compared to methane cultures of at

least 3-fold: G6P, F6P, F1, 6BP, citrate, KDPG, 6PG, which are intermediates of the EMP pathway, ED pathway, and TCA cycle. Other metabolites showed pool size decreases of at least 3-fold: PEP, 3PG and R5P, which are intermediates of the EMP pathway and the RuMP cycle. Pools of malate, succinate and 2- oxo-glutarate did not change significantly, which are also intermediates in the TCA cycle.

a



b

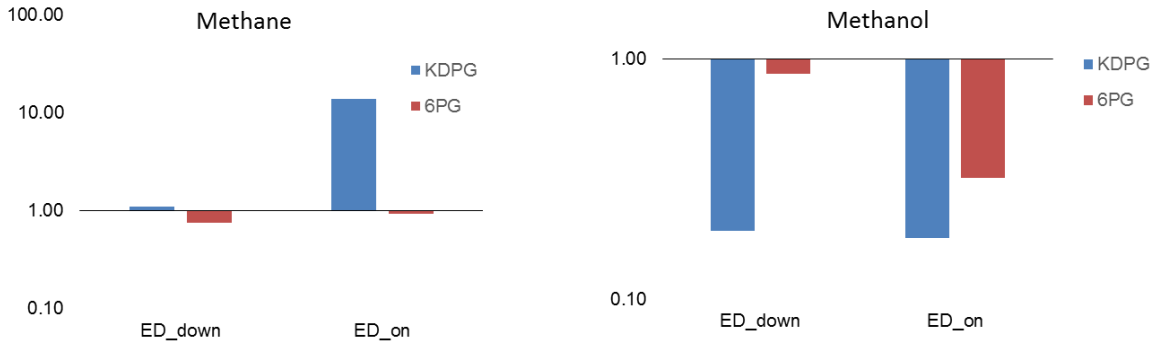


Figure 4.2 Targeted metabolite profile for M. buryatense 5GB1. a) Pools of the following metabolites changed, indicating altered flux through those metabolite nodes. Metabolites in the EMP and ED pathway, such as G6P, F6P, FBP, 6PG and KDPG increased under methanol growth, as did citrate, an intermediate of the TCA cycle. R5P, PEP, and 3PG decreased. C4 metabolites such as malate and succinate did not change. b) KDPG and 6PG pool changes for various conditions for WT and strain with tunable promoter for *edd*. Values were in triplicate

These results showed that the pools for most of the metabolites in core carbon metabolism changed under methanol growth, which indicates a systematic shift of flux distribution under methanol growth involving the EMP pathway, the ED pathway, C3-C4 metabolite interconversions and the TCA cycle. The increased levels of 6PG and KDPG suggested a shift towards the ED pathway. In order to test this hypothesis, attempts were made to generate null mutants in the first two genes of the ED pathway, *edd* (METBUDRAFT*1903/MBURv2*130008) encoding 6-phosphogluconate dehydratase and *eda* (METBUDRAFT*1904/MBURv2*130009), encoding 2-keto-3-deoxy 6-phosphogluconate aldolase. No deletion mutants were obtained. Instead, the native *edd* promoter was replaced in the chromosome with a tunable promoter system described previously[78]. The promoter system showed leaky expression when repressed by TetR protein, and is fully induced with anhydro-tetracycline above the level of 0.3 ug/ml (unpublished data), and this construct was successful. The key metabolites of the ED pathway were monitored in this strain with either the 'on' state or 'off' state under both methane and methanol growth. Shown in Figure 4.2b, under methanol growth, the KDPG pool decreased and 6PG level did not change significantly when the promoter was ‘off’ (with leaky expression), compared to the native promoter. When the promoter was tuned up by the inducer, both 6PG and KDPG levels dropped compared to the native promoter. Under methane growth, when turning the promoter ‘off’ (with leaky expression), no significant change in either the KDPG or 6PG level was observed compared to the native promoter, while tuning the promoter up to the ‘on’ state resulted in increased KDPG level but little change for the 6PG level. The increase of KDPG in the 'on' state under methane growth suggested that the construction functioned as designed, to increase flux through KDPG during growth on methane.

4.2. Gene expression of core metabolism response to substrate shifts

RNA-seq data was obtained and analyzed in order to identify gene expression response to the change of substrate from methane to methanol. RNA samples were taken from bioreactor runs with either methane or methanol as sole carbon source as described previously[64]. Table 4.1 shows the genes in core metabolism with fold change and p values (cut-off value =0.05) comparing methane and methanol culture. The gene expression fold changes were in the range of 0.3 - 2.8. In methanol culture compared to methane, genes of the RuMP cycle (such as hexulose phosphate synthase), ED pathway (such as 6PG dehydratase), and serine cycle (such as serine-glyoxylate aminotransferase) tended to be upregulated,

while those of the TCA cycle, such as fumarate hydratase, tended to be down-regulated. Most genes in C1 metabolism did not change, while Fdh expression was down-regulated.

4.3. Supernatant metabolite profile showed more formate secretion than methane culture

Previous bioreactor studies had shown high excretion of formate during methanol growth compared to methane[4], and so vial cultures of *M. buratense* 5GB1 growth with either methanol or methane were collected during exponential growth stage, and the supernatants were then analyzed using NMR to obtain a full profile for extracellular metabolites. As shown in Table 4.2, in methanol cultures 16.45 ± 1.02μM (g cdw^{-1}) of formate was secreted to the supernatant compared to 1.92 ± 0.51μM (g cdw^{-1}) for the methane culture. Much smaller amounts of acetate and lactate were excreted, and in this case, the amounts were similar to the methane culture. Methanol concentration in the supernatant was also monitored and the specific product yields for both biomass and substrate were then calculated. During growth on methanol, 11% ± 4% of the total methanol consumed is excreted as formate, while less than 0.1% of the methanol consumed is excreted as acetate plus lactate.

Table 4.2 Extracellular product yield based on biomass and methanol.

Product yield vs. biomass (mmol [g cdw^{-1}])	methanol	methane	SD methanol	SD methane
Formate	16.45	1.92	1.02	0.51
Acetate	0.18	0.13	0.13	0.01
Lactate	0.01	0.01	0.00	0.00
Product yield vs. methanol (mmol [mmol methanol $consumed^{-1}$])	methanol	methane	SD methanol	SD methane
Formate	0.11	NA	0.04	NA
Acetate	1.00E-03	NA	3.66E-04	NA
Lactate	5.92E-05	NA	2.92E-05	NA

Three biological replicates for methanol culture, two biological replicates for methane culture. SD, standard deviation. NA, not applicable

4.4. ^{13}C tracer analysis elucidated relative flux ratio downstream of PEP node.

As described previously[83], ^{13}C analysis cannot resolve relative flux contributions of the RuMP cycle, EMP pathway, or ED pathway, but can resolve relative flux contributions downstream of the PEP and pyruvate node due to carboxylation reactions. In this study, the same procedure was applied to WT with

^{13}C methanol as the tracer substrate. Figure 4.1 shows the MIDs of 2 key metabolites (full list in supplemental material Table S4_1). It clearly shows that malate and threonine (indicator of OAA) have a consistent labeling pattern, which suggested that the major flux to *de novo* production of OAA and malate was through carboxylation reactions, which on the other hand suggested a minor flux contribution from the oxidative TCA cycle to *de novo* malate production. The CO_2 pool was also determined with 76% unlabeled (diffusion from extracellular sources) and 24% labeled (from ^{13}C methanol oxidation through formate dehydrogenase). Meanwhile, citrate has about 20% of total label as M4, suggesting that AcCoA has two sources, from pyruvate (fully labeled) through pyruvate dehydrogenase and from malate, through malyl-CoA lyase (Figure 4.1a). The relative flux contribution to *de novo* production of AcCoA was then determined using the method described in Figure 4.3a. As shown in Figure 4.3b, the relative flux contribution to Acetyl-CoA in WT is calculated to be 30% from pyruvate, presumably from pyruvate dehydrogenase (Pdh) and 70% from malyl-CoA lyase (Mcl).

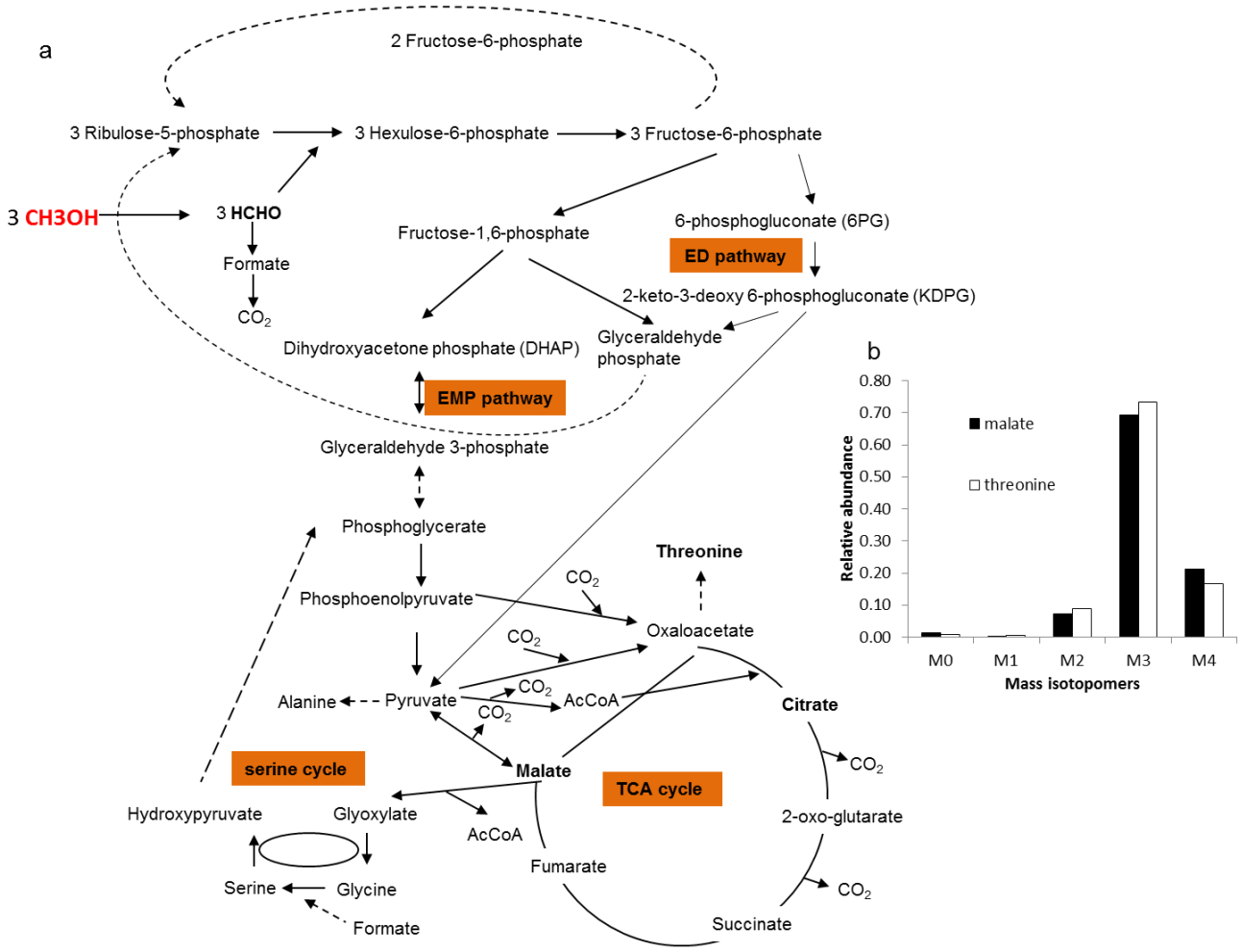
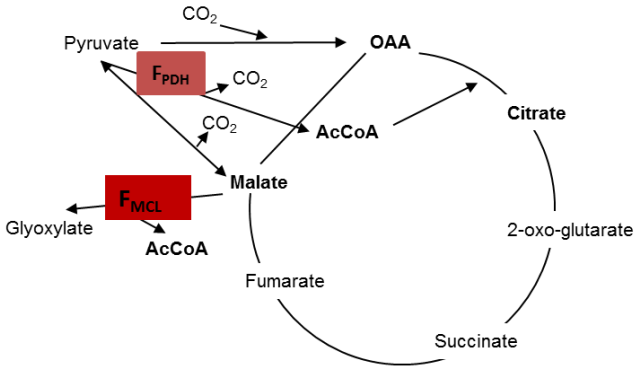


Figure 4.1 Core metabolism for methanol growth with MIDs for key metabolites. a) Methanol is oxidized to formaldehyde, and either enters into the RuMP cycle or is further oxidized to formate and then CO_2. Sugar phosphates with 6-carbon backbones are converted into sugar phosphates with 3-carbon backbones via either the EMP pathway or the EDD pathway. Downstream of the PEP and pyruvate node, multiple routes exist for the interconversion of C3-C4 compounds. Both pyruvate dehydrogenase as well as malyl-CoA lyase could contribute to *de novo* AcCoA synthesis. AcCoA enters the TCA cycle to provide precursors for biomass synthesis. b) MIDs of malate and threonine indicate negligible flux through fumarase to malate. Values were in triplicate.

The relative contribution to AcCoA was also determined in the construct with the tunable ED pathway in the 'off' state (Figure 4.3b). In this case, the contribution from Pdh was calculated to increase by 20%

while the contribution from Mcl decreased by 20%.

a. Culture with ^{13}C methanol



OAA_{M+3} * $AcCoA_{M+1}$=$Citrate_{M+4}$
OAA_{M+3} * $AcCoA_{M+2}$ + OAA_{M+4} * $AcCoA_{M+1}$=$Citrate_{M+5}$
F_{PDH}/F_{MCL}= $AcCoA_{M+1}$/ $AcCoA_{M+2}$
F_{pdh} + F_{mcl}* Mal_{M4} = $AcCoA_{M+2}$
F_{mcl}* Mal_{M3}=$AcCoA_{M+1}$
CO_2 pool: 76% unlabeled, 24 % labeled

b. Relative flux through Mcl and Pdh

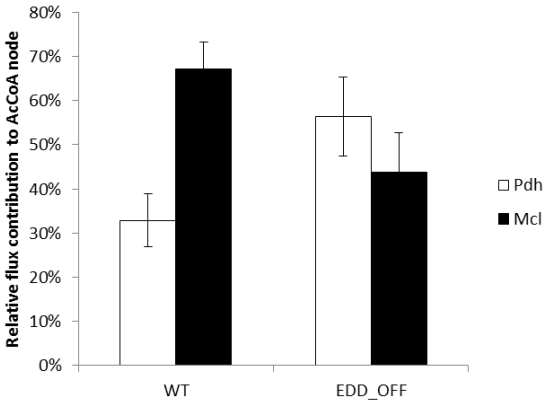


Figure 4.3 Quantitation method and result of relative flux ratio between FPDH and FMCL. a) Equations describing the labeling pattern of intermediate metabolites as well as relative flux from F_{PDH} and F_{MCL} to AcCoA. b) Relative flux through Pdh and Mcl in both WT and the construct with the tunable EDD pathway ('off') state.

4.5. Modeling growth on methanol

The experimental results of growth rate on methanol were compared to simulated results from a genome scale model, based on the original model published previously[68], with modifications to reflect methanol metabolism. First, the methanol transport reaction was added to the model, and also the electron transport summary equation was replaced with individual reactions including those for methanol-cyt*c* electron transfer (for the methanol dehydrogenase). The updated model files are included as supplemental material.

In the original model, the flux ratio of EMP and EDD is constrained by 3:1, based on previous results for methane growth[11]. This constraint is removed for methanol metabolism, since the relative ratio cannot

be determined without kinetic flux profiling. Secreted products, such as extracellular polysaccharide (EPS), formate, acetate and lactate were decoupled from the biomass equation for convenience in changing the values. In addition, the glycogen content was decoupled from the biomass equation, since the glycogen content changed under methanol growth. We assumed that the biomass composition other than glycogen was similar to methane growth, which leads to unchanged stoichiometry of precursors included in the biomass equation.

Specific methanol uptake was measured in vial cultures as 31 mmol (±11) methanol (g cdw^{-1} hr^{-1}). All of the flux balance analysis done in this study was optimized for maximum biomass growth. With the control model, substrate uptake rate, measured extracellular product rate, glycogen content[64] as well as EPS content (similar as methane condition) were constrained based on the experimental values. Several cases were simulated in the model with COBRApy, the results are summarized in Table 4.3. The optimal growth rate under this condition is 0.475 hr^{-1}, higher than the measured growth rate in vials (0.17-0.2 hr^{-1}). The formate dehydrogenase flux is predicted to be 0 under the optimal growth condition. Robustness analysis was then performed for formate dehydrogenase flux (FDH) to biomass flux as shown in Figure 4.4a. A negative correlation was found between FDH and biomass flux with two slopes. Based on the experimental data for growth rate (with 25% variation), the lower bound of FDH was set to 15 mmol ($gcdw^{-1}hr^{-1}$). As expected, in this model, the biomass flux dropped to 0.257 hr^{-1}. The model was then further constrained based on flux ratios determined through ^{13}C tracer analysis. Growth rate dropped to 0.25 hr^{-1} with an incomplete TCA cycle (upper bound of flux through fumarase was constrained =0). And with the ratio constraint of malyl-CoA lyase and pyruvate dehydrogenase (MCLA1/PDH=3), the growth rate dropped to 0.239 (in total 7% decrease in biomass flux with additional constraints from ^{13}C tracer analysis). The impact of the flux ratio between EMP and ED pathways on growth rate was also evaluated based on the model with ^{13}C results integrated. As shown in Figure 4.5a, there is limited impact on growth rate unless the EMP/ED ratio is smaller than 0.25. One of the extreme scenarios was simulated in which fructose-bisphosphate aldolase was turned off, forcing the system to exclusively use the ED pathway for assimilation, which dropped the biomass flux to 0.182 hr^{-1}. The flux map generated from the last scenario is shown in Figure 4.5b. These models also suggest a decrease in absolute flux through the TCA cycle.

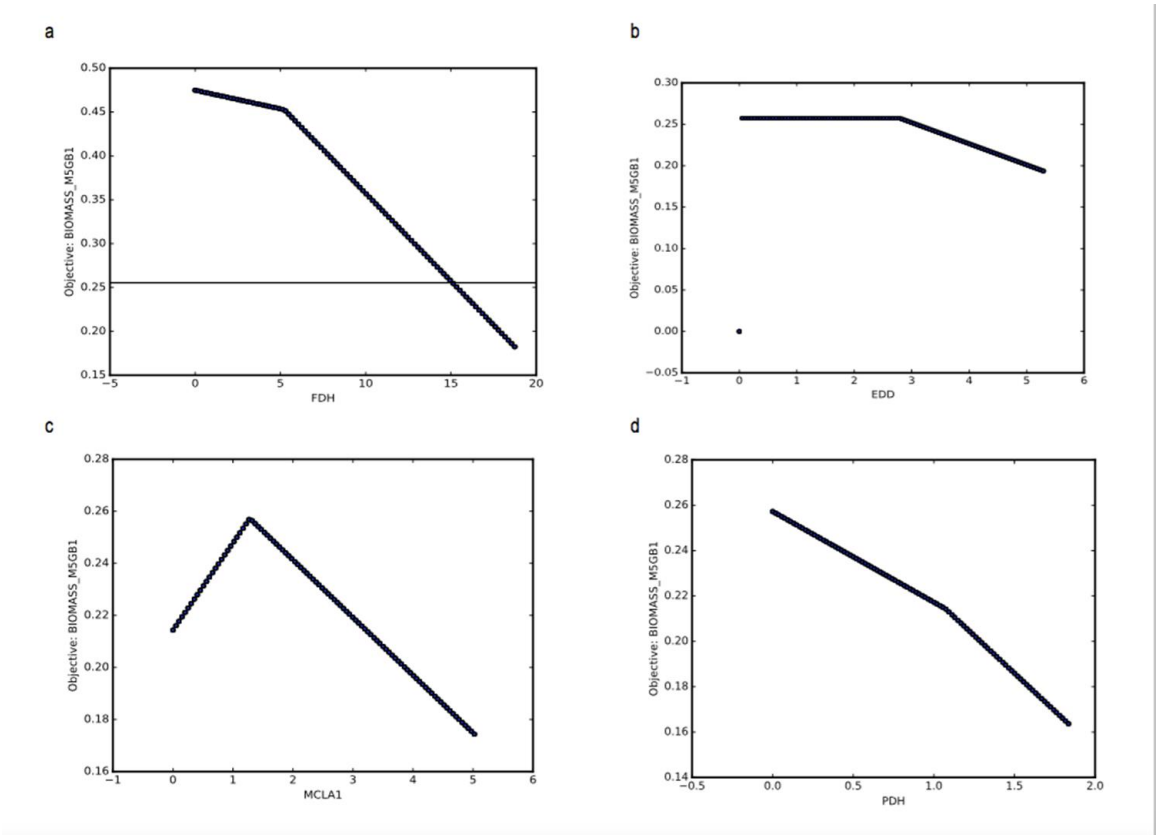


Figure 4.4 Robustness analysis of FDH, ED, MCLA1, and PDH on biomass flux. a) FDH flux correlates to biomass flux negatively with two distinct slopes. When the FDH flux > 15 mmol/gcdw.hr, the growth rate will fall below the region of 0.25 hr-1. b) ED flux correlates to biomass flux negatively once the value exceeds 3 mmol/g.cdw.hr. c) MCLA1 (malyl-CoA lyase) has an optimal point for biomass flux at 1.5 mmol/g.cdw.hr. d) PDH correlates to biomass flux negatively with two distinct slopes.

Table 4.3 Result summary for FBA analysis

Model	Predicted growth rate	Note
Control	0.475	methanol uptake rate = 31 mmol /$g^{-1}hr^{-1}$, extracellular product secretion included, glycogen content was set at 42% of total cell biomass, EPS was set at 10% of total cell biomass
FDH_constrained	0.257	based on model 'control', lower bound of FDH flux was set to 15 mm $g^{-1}hr^{-1}$
Fum_constrained	0.25	based on model 'FDH_constrained', lower_bound of fumarase flux was set =0
Ratio_construmed	0.239	Based on model 'Fum_constrained', flux ratio of MCL1A/PDH was set = 3:1 based on ^{13}C result.
ED pathway only	0.182	EMP pathway turned off.

In addition, we also evaluated how the flux through ED pathway, PDH, and MCLA1 is predicted to impact the cell growth using robustness analysis. As shown in Figure 4.4b, flux through the ED pathway does not impact the optimal growth rate until the value is higher than 2.8 mmol (g $cdw^{-1}hr^{-1}$). Beyond that point, the growth rate correlates negatively with higher flux through the ED pathway. As shown in Figure 4.4c, for flux through the MCL pathway (MCLA1 reaction), there was an optimum point for biomass growth. Less or greater than the optimum point, the growth rate is predicted to be negatively impacted. On the other hand, as shown in Figure 4.4d, PDH correlates negatively to biomass fluxes with two slopes. The upper limit for PDH flux is around 1.8 mmol $gcdw^{-1}$ hr^{-1}. Beyond that point the model does not generate a feasible

solution.

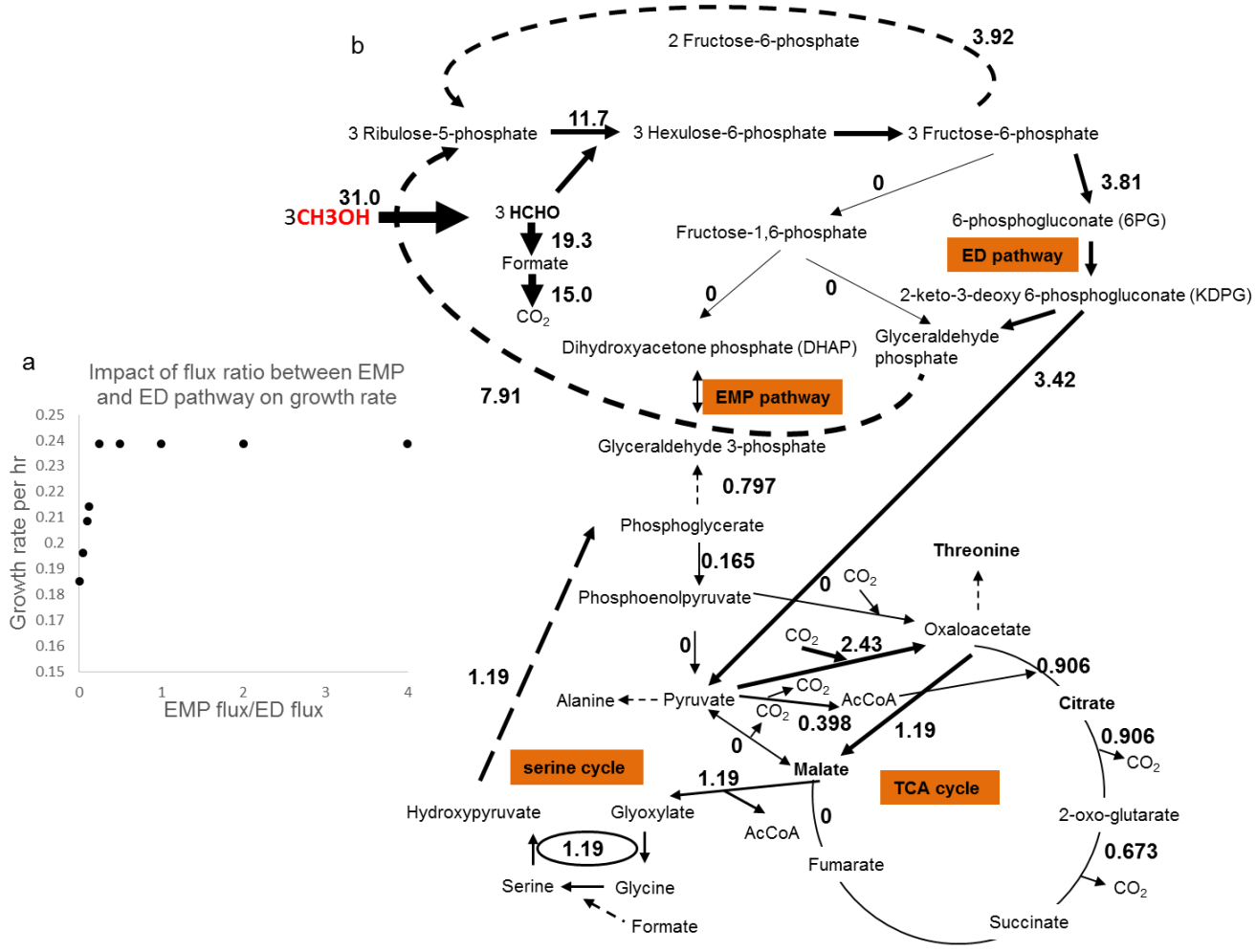


Figure 4.5 a) The sensitivity analysis of EMP flux/ED flux ratio to growth rate. This suggests that the ratio will only positively impact growth rate when the ratio is smaller than 0.25. b) Flux distribution predicted in COBRApy with the scenario of EMP pathway shut-off.

Discussion

In this study, we used a combination of metabolite pool measurements, ^{13}C flux labeling, engineered pathway flux, transcriptomics, and flux modeling to show that the metabolic network operating during the growth of *M. buryatense* 5GB1 on methanol is substantially different from that operating during growth of this strain on methane. Significant shifts occur in the pathway used for converting C6 compounds to C3 compounds, from the EMP pathway to the ED pathway. In addition, the TCA cycle shifts from operating in

a complete, oxidative mode, to an incomplete branched mode. Finally, substantial redirection of carbon flux occurs to glycogen and to excreted formate. These results are unexpected, because a metabolic pathway diagram approach to methanol growth (Figure 4.1a) suggests that it should be highly similar to methane growth, and initial metabolic models we used resulted in similar predictions. Our results provide more insights into the changes that occurred.

The high use of the ED pathway caused us to manipulate the ED pathway activity by swapping the native promoter with a tunable promoter system. The 'on' state under methanol growth showed the pools of both KDPG and 6PG changed compared to WT, which could be explained if *edd* (METBUDRAFT*1903/MBURv2*130008) and *eda* (METBUDRAFT*1904/MBURv2*130009) were sharing the same promoter. In addition, both pools decreased, suggested that the *eda* expression is higher than native expression of *eda* comparing to WT. Similarly, when turning down the promoter (with leaky expression), in methanol culture, the level of KDPG decreased while 6PG level keeps similar level with WT. This matches with expectation that the 6PG dehydrogenase (encoded by *edd*) was turned down. While in methane culture, by turning down the promoter, both KDPG and 6PG doesn't change much, which suggested low ED pathway activity in WT under methane growth, comparable to the construction with leaky expression.

To sum up, these results suggest that in WT, the ED pathway carried more flux in growth on methanol comparing to growth on methane. In order to evaluate the impact of the ED pathway to flux through later pathways, ^{13}C analysis was also conducted on the strain with 'off' state. In this case, that manipulation resulted in altered downstream flux distribution through the serine cycle as well as pyruvate dehydrogenase. This suggested that one of the control points for the network response to methanol as a substrate is the ED pathway activity. Meanwhile, the sensitivity analysis of the flux ratio between EMP/ED to biomass flux (growth rate) suggested that the biomass flux does not response until the EMP/ED ratio is lower than 0.25, and our results support the concept that the ED pathway dominates during methanol growth. It was predicted that the EMP pathway was not essential to cell growth, but will cause a growth defect on methanol when it is completely shut down.

The flux distribution among the C3-C4 interconversions, especially relative flux through both pyruvate dehydrogenase and malyl-CoA lyase was significantly different for methanol growth compared to

methane growth. Sensitivity analysis on both showed that growth rate responds to changes in fluxes of both pathways. We propose that these two pathways are also control points of the network for the two substrate shift.

Our results also suggest that the main function of the TCA cycle under methanol growth is to provide precursors for *de novo* biosynthesis, not to provide NADH, as in methane growth. Instead, during methanol growth, it appears that formate dehydrogenase (FDH) takes over the role of producing NADH. The predicted high FDH activity matches with the extracellular formate secretion, suggesting that the activity of FDH reaches a maximum that is lower than the formate production rate.

In general, the RNA-seq transcriptomics data supported the other experimental findings with regard to trends. However, the changes in gene expression level were subtle but statistically significant, while the relative changes in fluxes were much larger. This suggests that the changes in metabolites and fluxes could be caused as much if not more by small molecule regulation than by gene expression level. Of the targeted metabolites, some changed by over 5-fold, including KDPG, citrate, F6P, and F1,6BP (increased on methanol) and 3PG and PEP (decreased on methanol). These are candidates for small molecule regulation.

As shown in another study[86], we found that integrating ^{13}C results into the genome scale model provides additional constraints to the model that redefine the solution space for better prediction as well as redefining the design space for future strain engineering. The critical information about flux distribution in the network can only be measured locally, but with this approach, we were able to obtain a better prediction about the flux through genome-scale model. A recent effort has been made to develop a method to constrain genome-scale models with ^{13}C labeling data[75] and eliminate the need to assume an objective function for optimization. However, this approach is difficult to apply to a C1 network for the same reasons that make the well-established ^{13}C MFA method difficult to apply, namely the full labeling of many key intermediates.

In summary, a systematic shift of metabolism was observed when switching the substrate from methane to methanol, which suggested significant flexibility of core central metabolism in this obligate methylotroph. The presence of both EMP and ED pathways provides flexibility to the whole network in response to environmental perturbations (such as shift in substrate type), genetic manipulations or both.

This study is an important step forward to understanding the core metabolism in Type I methanotrophs for methanol-based biotechnology.

Conclusion

In this study, we demonstrated a systematic shift of metabolism in *M. buryatense* 5GB1 in cells grown on methanol compared to methane towards the ED pathway, the serine cycle, and an incomplete TCA cycle. The combination of flux balance analysis of genome-scale model and flux ratio from ^{13}C data changed the solution space for better prediction of cell behavior.

Method

4.6. Cell culture for ^{13}C labeling experiment, phenotypic characterization, and medium composition

For growth curve experiments, liquid pre-cultures were grown in 25 ml tubes with NMS2 medium with 0.2% methanol in 30°C, 200 rpm, sealed with rubber stoppers and aluminum seals (Wheaton, Millville, NJ, USA). Pre-cultures were then inoculated in new 25 ml tubes with starting OD_{600} = 0.01 for growth rate experiments as well as ^{13}C analysis. NMS2 medium as described in[28] was used for liquid culture. For strain construction, zeocin (Zeo), was used at 30μg/ml.

For ^{13}C labeling experiment, precultures were grown in 25 ml tubes with 0.2% methanol for 18 hours. The precultures were then inoculated into 50 ml fresh medium in 250 ml serum bottles (with starting OD_{600}=0.01). Cell were harvested at 0.3 ~ 0.6 when the culture was under both isotopic steady state and metabolic steady state. All strains used in this study are listed in supplemental Table S4_2.

4.7. Genetic manipulation of *M. buryatense* 5GB1 with Ptac-tetR-PtetA system

Primers for the tunable promoter system constructs are listed in Table S4_3 in the supplemental material. The tunable promoter system was constructed by swapping the predicted promoter region (in total 570bp) in *edd* (METBUDRAFT*1903 /MBURv2*130008) region with the PtetA promoter[87]. The construction was implemented with assembled PCR products which were then electroporated into unmarked strain FC35 (with Ptac-tetR integrated into the METBUDRAFT_*1431-1432/MBURv2*_60220-60221 site) using the

protocol described previously[62]. The promoter is repressed without anhydrotetracycline (aTc) and is fully induced by adding aTc to the medium above a concentration of 0.3 µg/ml (unpublished data).

4.8. ^{13}C labeling pattern of metabolite measurements and targeted metabolomics analysis using LC/MS-MS

Cell quenching procedure with fast filtration and extracted using hot water extraction method was carried out as previously described[77]. Briefly, cell cultures were quenched using fast filtration, and saved in 50 ml falcon tubes submerged in liquid nitrogen. The collected samples were lyophilized for 12 hours to get rid of extra medium. Hot water extraction protocol was used to extract intracellular metabolites as described previously[77]. Briefly, 20 ml of boiling water was added to 50 ml falcon tubes, and then the tube was placed into a hot water bath for 20 minutes. Then the tubes were placed on ice for 30 minutes. The cell extract was centrifuged at 4C, 5000rpm for 30 minutes to remove cell biomass. The supernatant was transferred into new tubes, and fast frozen using liquid nitrogen. The cell lysates were lyophilized and concentrated into 50 µl of water. The reconstituted samples were centrifuged with filter (spin-x costar, 0.22 µm, Sigma-Aldrich), and kept at -20°C until LCMS analysis for both ^{13}C labeling pattern measurement as well as targeted metabolomics determinations. Waters Xevo Mass spectrometry with UPLC system was used for detection of labeling pattern of metabolites. Multiple reaction monitors (MRM) were set up for each metabolite of interest. This information is also included in the supplemental material Table S4_4. Similarly, for targeted metabolomics, internal standards of succinate- ^{13}C4 (Sigma-Aldrich, St.Louis, MO) and L-serine-^{13}C3 ^{15}N (Sigma-Aldrich, St.Louis, MO) were spiked into the cell extract for relative quantification of metabolites pools. The relative abundance of metabolites was determined by normalizing to the specific peak area to peak area of internal standards and cell dry weight. Extracellular metabolite secretion was measured using 1H NMR with a previously described protocol[11].

4.9. Mathematically determine relative flux distribution in acetyl-CoA node.

Acetyl-CoA will have a differentiated labeling pattern depending on pathways for synthesis. Based on current understanding of central carbon metabolism, acetyl-CoA could be synthesized by two different ways: first, it could be produced from pyruvate through pyruvate dehydrogenase. Secondly, it could be produced from malyl-CoA through malyl-CoA lyase. Route 1 leads to full labeled AcCoA. Route 2 will

produce 1 carbon labeled AcCoA. MIDs of AcCoA are calculated from citrate as well as threonine (OAA) since AcCoA could not be measured directly due to the low pool. The following equation was then developed to quantitate the relative flux in the AcCoA node between route 1 and route 2. The following equation is developed to determine the labeling pool of AcCoA and flux ratio between those 2 reactions.

$OAA_{M+3} * AcCoA_{M+1} = Citrate_{M+4}$

$OAA_{M+3} * AcCoA_{M+2} + OAA_{M+4} * AcCoA_{M+1} = Citrate_{M+5}$

$F_{PDH}/F_{MCL} = AcCoA_{M+1}/ AcCoA_{M+2}$

$F_{pdh} + F_{mcl} * Mal_{M4} = AcCoA_{M+2}$

$F_{mcl} * Mal_{M3} = AcCoA_{M+1}$

4.10. RNA-seq data analysis

The RNA-seq data analysis was implemented using the method described previously[68].

4.11. Flux balance analysis and robustness analysis

Genome-scale model simulation was done in COBRApy[88] with an updated model. The updated model is shown in Supplemental file 4_1. Robustness analysis as well as adding flux ratio constraints to the model was done with in-house developed python scripts, both scripts are shown in Supplemental file 4_2.

Tables

Table 4.1 Fold change of gene expressions in core metabolism comparing methanol to methane condition with p value less than 0.05

Locus tag	Gene	Product	Fold change	Pathway involved	P value
MBURv2_210062		lactate/malate dehydrogenase	0.36	C3-C4 interconversion	0.00
MBURv2_130865	fumC	fumarate hydratase (fumarase C), aerobic class II	0.42	TCA cycle	0.01
MBURv2_20327	pykA	pyruvate kinase	0.45	EMP pathway	0.04
MBURv2_210058	ald	alanine dehydrogenase	0.47	alanine biosynthesis	0.00
MBURv2_50079	fdh1B	Tungsten-containing NAD-dependent formate dehydrogenase, beta subunit	0.45	C1 pathway	0.00
MBURv2_50078	fdh1A	Tungsten-containing NAD-dependent formate dehydrogenase, alpha subunit	0.55	C1 pathway	0.03
MBURv2_130931	fdhD	Protein FdhD homolog	0.58	C1 pathway	0.02
MBURv2_130613	sps	sucrose-phosphate synthase	0.60	sucrose biosynthesis	0.00
MBURv2_130610	ams	amylosucrase	0.63	sucrose biosynthesis	0.01
MBURv2_210335		6-phosphogluconate dehydrogenase NAD-binding	0.67	RuMP cycle	0.00
MBURv2_210177	glgA	glycogen synthase	0.68	glycogen biosynthesis	0.003
MBURv2_210178	glgB	1,4-alpha-glucan branching enzyme	0.48	glycogen biosynthesis	0.0007
MBURv2_130310	mdh	malate dehydrogenase	0.69	C3-C4 interconversion	0.01
MBURv2_160480	fda	fructose-1,6-bisphosphate aldolase, class II	0.70	EMP pathway	0.03

Locus tag	Gene	Product	Fold change	Pathway involved	P value
MBURv2_80101	sdhA	succinate dehydrogenase, flavoprotein subunit	0.76	TCA cycle	0.01
MBURv2_20302	pps	phosphoenolpyruvate synthase	0.79	C3-C4 interconversion	0.02
MBURv2_80100	sdhB	succinate dehydrogenase, FeS subunit	0.83	TCA cycle	0.01
MBURv2_160358	icd	Isocitrate dehydrogenase [NADP]	0.86	TCA cycle	0.04
MBURv2_80063	tkt	transketolase domain protein	1.09	RuMP cycle	0.02
MBURv2_160221	zwf	glucose-6-phosphate dehydrogenase	1.11	RuMP cycle	0.03
MBURv2_160313	rmpB	3-hexulose-6-phosphate isomerase	1.23	RuMP cycle	0.00
MBURv2_160305	rmpB	3-hexulose-6-phosphate isomerase	1.24	RuMP cycle	0.00
MBURv2_160244	rpe	D-ribulose-5-phosphate 3-epimerase	1.25	RuMP cycle	0.03
MBURv2_210071	oadB	putative oxaloacetate decarboxylase beta chain	1.37	C3-C4 interconversion	0.04
MBURv2_130012	pdhB	pyruvate dehydrogenase E2 component; dihydrolipoamide acetyltransferase	1.41	pyruvate metabolism	0.04
MBURv2_130389	sucC	succinyl-CoA synthetase, beta subunit	1.46	TCA cycle	0.01
MBURv2_130008	edd	6-phosphogluconate dehydratase	1.49	ED pathway	0.03
MBURv2_130009	eda	multifunctional 2-keto-3-deoxygluconate 6-phosphate aldolase and 2-keto-4-hydroxyglutarate aldolase and oxaloacetate decarboxylase	1.30	ED pathway	0.05
MBURv2_130011	aceE	pyruvate dehydrogenase, decarboxylase component E1, thiamin-binding	1.51	pyruvate metabolism	0.01

Locus tag	Gene	Product	Fold change	Pathway involved	P value
MBURv2_130313	glyA	serine hydroxymethyltransferase	1.59	serine cycle	0.00
MBURv2_130401	pfp	pyrophosphate--fructose 6-phosphate 1-phosphotransferase	1.68	EMP pathway	0.00
MBURv2_210131	pgk	phosphoglycerate kinase	1.68	EMP pathway	0.01
MBURv2_20405	eno	enolase	1.87	EMP pathway	0.01
MBURv2_160304	rmpA	3-hexulose-6-phosphate synthase	1.91	RuMP cycle	0.00
MBURv2_130302	sgaA	serine--glyoxylate aminotransferase	1.93	serine cycle	0.01
MBURv2_160312	rmpA	3-hexulose-6-phosphate synthase	1.94	RuMP cycle	0.00
MBURv2_160308	rmpA	3-hexulose-6-phosphate synthase	1.94	RuMP cycle	0.00
MBURv2_130299	sucC	succinyl-CoA synthetase, beta subunit	2.05	TCA cycle	0.01
MBURv2_50413	gpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	2.82	EMP pathway	0.00

Supplemental materials

Table S4_1. MIDs for intermediate metabolites of *M. buryatense* 5GB1 grown on methanol

Table S4_2. Strains used in this study

Table S4_3. Sequence of primers used in this study

Table S4_4. Multiple reaction monitoring (MRM) for intermediate metabolites

Supplemental file 4_1: updated genome-scale model for methanol growth in SBML format

Supplemental file 4_2: python scripts used (1. constrain the model with flux ratios, 2. Robustness analysis) in this study

Chapter 5: Summary and future directions

Summary

In the first project, I have successfully applied steady state flux analysis on a methanol utilizing bacterium, *Methylobacterium extorquens* AM1 to investigate the trade-off between growth rate and biomass yield found in two laboratory variants of this bacterium. This analysis has identified the source of the trade-off, involving changes in the routes taken for C3-C4 inter-conversion, which was validated by mutant analysis. These results were surprising, because some of these reactions have been previously thought to be side reactions, without major importance to core parameters such as growth rate and yield. Only by combining the approaches used here, was it possible to dissect the importance of these reactions.

In the second project, I have established an analytical method for measuring intracellular metabolite pools in *Methylomicrobium buryatense* 5GB1 as well as measuring isotopomer distribution among intracellular metabolites, which lays the foundation of ^{13}C metabolic flux analysis. I showed that a complete oxidative TCA cycle operates during growth on methane, in direct contrast to the literature of several decades. This finding significantly changes our understanding of the metabolic network during growth on methane, with regard to flux through specific pathways and the source of NADH. I also found a systematic metabolism shift occurred in *M. buryatense* 5GB1 growth on methanol compared to growth on methane. This was also very surprising, as it has been assumed that growth on methanol is similar to growth on methane. Taken together, these new results have provided major new insights into methylotrophic metabolic networks and how they function. Again, by combining approaches I was able to dissect roles and directly measure fluxes, uncovering the misunderstandings of past work, which was carried out before metabolic flux analysis and metabolomics were possible.

Another important result of this study is that the special network topology as well as C1 input of the network brought limitations as well as challenges for implementation of metabolic flux analysis. To address these challenges, I have used multiple strategies (such as mutant analysis

coupled with ^{13}C tracer analysis, as well as genome scale model coupled with ^{13}C flux ratio) as complementary approaches to metabolic flux analysis, which has led to a better understanding of the network, and has led to more accurate predictions.

To sum up, several cases are shown in this dissertation using the approaches of metabolic flux analysis as well as metabolomics to investigate cell behavior at the level of 'end products' in central dogma (metabolites as well as fluxes). These cases have demonstrated the usefulness of these two approaches in terms of measuring intracellular metabolism, which could further provide critical information for understanding cell physiology as well as guide strain optimization process in the design – build – test iterations in metabolic engineering.

Future directions

The overall impact of this work is for both methods development (metabolite and ^{13}C flux analysis and FBA analysis) and for uncovering fundamental information regarding flux through metabolic pathways in two different bacteria that use one-carbon compounds. My dissertation will lay the groundwork for future metabolic engineering of these industrially-promising methylotrophs, by delineating how carbon is partitioned across multiple branch points within the metabolic networks of these bacteria, under different growth conditions.

To date, most successful engineered strains for valuable chemicals or fuel production which could be economically feasible are in sugar-based systems, which will always suffer from the dilemma of competing food with energy or chemicals. With increasing interest in using C1 substrates (either methane or methanol) for valuable chemical production via methylotrophs, the basic understanding for C1 model organism is a crucial step in the metabolic engineering process. For methane-based biotechnology, *M. buryatense* 5GB1 is both a model system for basic research on methanotrophy as well as an industrial-promising candidate for bio-conversion of methane into valuable chemicals. With the increasing understanding of central carbon metabolism as well as 'omics' data, it is possible to expand the scope of valuable products that could be over-produced by *M. buryatense* 5GB1. Based on the surprising new finding of a

complete TCA cycle, it is possible that the current network is capable of accepting drop-in modules from other biological systems designed for *E. coli*, for instance. One other direction we are following right now is the 'pseudofermentation process' of conversion of C1 substrates to carboxylic acids under O_2 limitation. With the help of the updated genome-scale model and baseline performance data from bioreactors, we can quickly move forward in our understanding of how central metabolism operates under this 'pseudofermentation process'.

One key approach that needs to be developed in the C1 system regarding the RuMP cycle is kinetic flux labeling profiling. This is the only solution for quantitating the flux via the core RuMP pathway, and between the EMP and ED pathways. As shown in the later study of *M. extorquens* 5GB1 with the genome-scale model, the sensitivity analysis of the flux ratio between EMP/ED actually has two stages, in one of which it correlates positively with biomass flux, and beyond that it will not affect the biomass flux. Once the ratio of EMP and ED fluxes are quantitated through kinetic flux profiling, this will provide additional constraints on the model, which will lead to a better description of the WT flux distribution in the genome-scale model. In the design-build-test cycle, metabolic flux analysis as well as targeted metabolomics plays important roles in the test step, which further guides the direction of design in a rational way. The cases presented here are the first steps towards the goal of metabolic engineering the strains for valuable chemical production. The baseline of WT is extremely important in the sense of both understanding the model organisms as well as the design capacity of each organism. As we move forward, the power of the tools used here as well as the baseline knowledge will become essential to further progress in this important area.

Bibliography

1. *The Biochemistry of Methylotrophs - Anthony,C*. 1983.

2. Chistoserdova L, Lidstrom ME: **Aerobic Methylotrophic Prokaryotes**. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013.

3. Chistoserdova L, Kalyuzhnaya MG, Lidstrom ME: **The expanding world of methylotrophic**

metabolism. *Annu Rev Microbiol* 2009, **63**:477–99.

4. Kalyuzhnaya M, Puri AW, Lidstrom ME: **Metabolic engineering in methanotrophic bacteria.** *Metab Eng* 2015, **29**:142–152.

5. Lan EI, Liao JC: **Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources.** *Bioresour Technol* 2013, **135**:339–49.

6. Gronenberg LS, Marcheschi RJ, Liao JC: **Next generation biofuel engineering in prokaryotes.** *Curr Opin Chem Biol* 2013, **17**:462–71.

7. Nielsen J, Fussenegger M, Keasling J, Lee SY, Liao JC, Prather K, Palsson B: **Engineering synergy in biotechnology.** *Nat Chem Biol* 2014, **10**:319–22.

8. Kalscheuer R, Stölting T, Steinbüchel A: **Microdiesel: Escherichia coli engineered for fuel production.** *Microbiology* 2006, **152**(Pt 9):2529–36.

9. Wackett LP: **Biomass to fuels via microbial transformations.** *Curr Opin Chem Biol* 2008, **12**:187–93.

10. Haynes CA, Gonzalez R: **Rethinking biological activation of methane and conversion to liquid fuels.** *Nat Chem Biol* 2014, **10**:331–9.

11. Kalyuzhnaya MG, Yang S, Rozova ON, Smalley NE, Clubb J, Lamb a, Gowda G a N, Raftery D, Fu Y, Bringel F, Vuilleumier S, Beck D a C, Trotsenko Y a, Khmelenina VN, Lidstrom ME: **Highly efficient methane biocatalysis revealed in a methanotrophic bacterium.** *Nat Commun* 2013, **4**(May):2785.

12. Conrado RJ, Gonzalez R: **Chemistry. Envisioning the bioconversion of methane to liquid fuels.** *Science* 2014, **343**:621–3.

13. Cameron DC: **Methane as a Feedstock for Bio-based Processes An Industrial and Venture Perspective**. 2012.

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

15. Semrau JD: **Bioremediation via Methanotrophy: Overview of Recent Findings and Suggestions for Future Research.** *Front Microbiol* 2011, **2**:209.

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

17. Crowther GJ, Kosály G, Lidstrom ME: **Formate as the main branch point for methylotrophic metabolism in Methylobacterium extorquens AM1.** *J Bacteriol* 2008, **190**:5057–62.

18. Chistoserdova L, Chen S-W, Lapidus A, Lidstrom ME: **Methylotrophy in Methylobacterium extorquens AM1 from a genomic point of view.** *J Bacteriol* 2003, **185**:2980–7.

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

20. Marx CJ, Lidstrom ME: **Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria.** *Biotechniques* 2002, **33**:1062–1067.

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

22. Yang S, Sadilek M, Synovec RE, Lidstrom ME: **Liquid chromatography-tandem quadrupole mass spectrometry and comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry measurement of targeted metabolites of Methylobacterium extorquens AM1 grown on two different carbon sources.** *J Chromatogr A* 2009, **1216**:3280–9.

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

24. Helling RB: **Speed versus efficiency in microbial growth and the role of parallel**

pathways. *J Bacteriol* 2002, **184**:1041–5.

25. Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S: **Experimental tests for an evolutionary trade-off between growth rate and yield in E. coli.** *Am Nat* 2006, **168**:242–51.

26. Skovran E, Crowther GJ, Guo X, Yang S, Lidstrom ME: **A systems biology approach uncovers cellular strategies used by Methylobacterium extorquens AM1 during the switch from multi- to single-carbon growth.** *PLoS One* 2010, **5**:e14091.

27. Puri AW, Owen S, Chu F, Chavkin T, Beck D a. C, Kalyuzhnaya MG, Lidstrom ME: **Genetic Tools for the Industrially Promising Methanotroph Methylomicrobium buryatense**. *Appl Environ Microbiol* 2015, **81**:1775–1781.

28. Khmelenina VN, Beck DAC, Munk C, Davenport K, Daligault H, Erkkila T, Goodwin L, Gu W, Lo C-C, Scholz M, Teshima H, Xu Y, Chain P, Bringel F, Vuilleumier S, Dispirito A, Dunfield P, Jetten MSM, Klotz MG, Knief C, Murrell JC, Op den Camp HJM, Sakai Y, Semrau J, Svenning M, Stein LY, Trotsenko YA, Kalyuzhnaya MG: **Draft Genome Sequence of Methylomicrobium buryatense Strain 5G, a Haloalkaline-Tolerant Methanotrophic Bacterium.** *Genome Announc* 2013, **1**:e00053–13.

29. Chen X, Alonso AP, Allen DK, Reed JL, Shachar-Hill Y: **Synergy between 13C-metabolic flux analysis and flux balance analysis for understanding metabolic adaption to anaerobiosis in E. coli**. *Metab Eng* 2011, **13**:38–48.

30. Wiechert W: **13C metabolic flux analysis.** *Metab Eng* 2001, **3**:195–206.

31. Wittmann C: **Fluxome analysis using GC-MS.** *Microb Cell Fact* 2007, **6**:6.

32. Zamboni N, Fendt S-M, Rühl M, Sauer U: **(13)C-based metabolic flux analysis.** *Nat Protoc* 2009, **4**:878–92.

33. Antoniewicz MR, Kelleher JK, Stephanopoulos G: **Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions.** *Metab Eng* 2007, **9**:68–86.

34. Quek L-E, Wittmann C, Nielsen LK, Krömer JO: **OpenFLUX: efficient modelling software for 13C-based metabolic flux analysis.** *Microb Cell Fact* 2009, **8**:25.

35. Wiechert W, Möllney M, Petersen S, de Graaf AA: **A universal framework for 13C metabolic flux analysis.** *Metab Eng* 2001, **3**:265–83.

36. Sokol S, Millard P, Portais J-C: **influx_s: increasing numerical stability and precision for metabolic flux analysis in isotope labelling experiments.** *Bioinformatics* 2012, **28**:687–93.

37. Hörl M, Schnidder J, Sauer U, Zamboni N: **Non-stationary (13)C-metabolic flux ratio analysis.** *Biotechnol Bioeng* 2013, **110**:3164–76.

38. Yuan J, Bennett BD, Rabinowitz JD: **Kinetic flux profiling for quantitation of cellular metabolic fluxes.** *Nat Protoc* 2008, **3**:1328–40.

39. Young JD: **INCA: a computational platform for isotopically non-stationary metabolic flux analysis.** *Bioinformatics* 2014, **30**:1333–5.

40. Large PJ, Peel D, Quayle JR: **Microbial growth on C1 compounds. II. Synthesis of cell constituents by methanol- and formate-grown Pseudomonas AM 1, and methanol-grown Hyphomicrobium vulgare.** *Biochem J* 1961, **81**:470–80.

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

42. Chubiz LM, Purswani J, Carroll SM, Marx CJ: **A novel pair of inducible expression vectors for use in Methylobacterium extorquens.** *BMC Res Notes* 2013, **6**:183.

43. Skovran E, Crowther GJ, Guo X, Yang S, Lidstrom ME: **A systems biology approach uncovers cellular strategies used by Methylobacterium extorquens AM1 during the switch from multi- to single-carbon growth.** *PLoS One* 2010, **5**:e14091.

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

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

46. Kiefer P, Buchhaupt M, Christen P, Kaup B, Schrader J, Vorholt JA: **Metabolite profiling uncovers plasmid-induced cobalt limitation under methylotrophic growth conditions.** *PLoS One* 2009, **4**:e7831.

47. Orita I, Nishikawa K, Nakamura S, Fukui T: **Biosynthesis of polyhydroxyalkanoate**

copolymers from methanol by Methylobacterium extorquens AM1 and the engineered strains under cobalt-deficient conditions. *Appl Microbiol Biotechnol* 2014, **98**:3715–25.

48. Delaney NF, Kaczmarek ME, Ward LM, Swanson PK, Lee M-C, Marx CJ: **Development of an optimized medium, strain and high-throughput culturing methods for Methylobacterium extorquens.** *PLoS One* 2013, **8**:e62957.

49. Chou HH, Berthet J, Marx CJ: **Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation**. *PLoS Genet* 2009, **5**.

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

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

52. Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S: **Experimental tests for an evolutionary trade-off between growth rate and yield in E. coli.** *Am Nat* 2006, **168**:242–51.

53. Russell JB: **Energetics of bacterial growth : balance of anabolic and catabolic reactions . Energetics of Bacterial Growth : Balance of Anabolic and Catabolic Reactions**. 1995, **59**.

54. Wiechert W, de Graaf AA: **In vivo stationary flux analysis by 13C labeling experiments.** *Adv Biochem Eng Biotechnol* 1996, **54**:109–54.

55. Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ: **Diminishing returns epistasis among beneficial mutations decelerates adaptation.** *Science* 2011, **332**:1190–2.

56. Van Dien SJ: **Reconstruction of C3 and C4 metabolism in Methylobacterium extorquens AM1 using transposon mutagenesis**. *Microbiology* 2003, **149**:601–609.

57. Kiefer P, Buchhaupt M, Christen P, Kaup B, Schrader J, Vorholt JA: **Metabolite profiling uncovers plasmid-induced cobalt limitation under methylotrophic growth conditions.** *PLoS One* 2009, **4**:e7831.

58. Okubo Y, Skovran E, Guo X, Sivam D, Lidstrom ME: **Implementation of microarrays for Methylobacterium extorquens AM1.** *OMICS* 2007, **11**:325–40.

59. Chistoserdova L, Lidstrom ME: **Identification and mutation of a gene required for glycerate kinase activity from a facultative methylotroph, Methylobacterium extorquens AM1**. *J Bacteriol* 1997, **179**:4946–4948.

60. Fei Q, Guarnieri MT, Tao L, Laurens LML, Dowe N, Pienkos PT: **Bioconversion of natural gas to liquid fuel: opportunities and challenges.** *Biotechnol Adv* 2014, **32**:596–614.

61. Khmelenina VN, Kalyuzhnaya MG, Starostina NG, Suzina NE, Trotsenko YA: **Isolation and Characterization of Halotolerant Alkaliphilic Methanotrophic Bacteria from Tuva Soda Lakes**. *Curr Microbiol* 1997, **35**:257–261.

62. Yan X, Chu F, Puri AW, Fu Y, Lidstrom ME: **Electroporation-Based Genetic Manipulation in Type I Methanotrophs**. *Appl Environ Microbiol* 2016, **82**:2062–2069.

63. Puri AW, Owen S, Chu F, Chavkin T, Beck DAC, Kalyuzhnaya MG, Lidstrom ME: **Genetic tools for the industrially promising methanotroph Methylomicrobium buryatense.** *Appl Environ Microbiol* 2015, **81**:1775–81.

64. Gilman A, Laurens LM, Puri AW, Chu F, Pienkos PT, Lidstrom ME: **Bioreactor performance parameters for an industrially-promising methanotroph Methylomicrobium buryatense 5GB1.** *Microb Cell Fact* 2015, **14**:182.

65. Klapa M, Stephanopoulos G: *Metabolic Flux Analysis*. 2000.

66. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, DeBerardinis RJ, Feron O, Frezza C, Ghesquiere B, Gottlieb E, Hiller K, Jones RG, Kamphorst JJ, Kibbey RG, Kimmelman AC, Locasale JW, Lunt SY, Maddocks OD, Malloy C, Metallo CM, Meuillet EJ, Munger J, Nöh K, Rabinowitz JD, Ralser M, Sauer U, Stephanopoulos G, St-Pierre J, Tennant DA, et al.: **A roadmap for interpreting 13C metabolite labeling patterns from cells**. *Curr Opin Biotechnol* 2015, **34**:189–201.

67. Toya Y, Shimizu H: **Flux analysis and metabolomics for systematic metabolic engineering of microorganisms.** *Biotechnol Adv* 2013, **31**:818–26.

68. De la Torre A, Metivier A, Chu F, Laurens LML, Beck DAC, Pienkos PT, Lidstrom ME, Kalyuzhnaya MG: **Genome-scale metabolic reconstructions and theoretical investigation of methane conversion in Methylomicrobium buryatense strain 5G(B1).** *Microb Cell Fact* 2015,

14:188.

69. Davey JF, Whittenbury R, Wilkinson JF: **The distribution in the methylobacteria of some key enzymes concerned with intermediary metabolism**. *Arch Mikrobiol* 1972, **87**:359–366.

70. Hazeu W, Batenburg-van der Vegte WH, de Bruyn JC: **Some characteristics of Methylococcus mobilis sp. nov.** *Arch Microbiol* 1980, **124-124**:211–220.

71. Patel R, Hoare L, Hoare DS, Taylor BF: **Incomplete tricarboxylic acid cycle in a type I methylotroph, Methylococcus capsulatus.** *J Bacteriol* 1975, **123**:382–4.

72. Kaluzhnaya M, Khmelenina V, Eshinimaev B, Suzina N, Nikitin D, Solonin A, Lin J-L, Mcdonald I, Murrell C, Trotsenko Y: **Taxonomic characterization of new alkaliphilic and alkalitolerant methanotrophs from soda lakes of the Southeastern Transbaikal region and description of Methylomicrobium buryatense sp.nov.** *Syst Appl Microbiol* 2001, **24**:166–176.

73. Zhang R, Watson DG, Wang L, Westrop GD, Coombs GH, Zhang T: **Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites**. *J Chromatogr A* 2014, **1362**:168–179.

74. Kalyuzhnaya M, Puri AW, Lidstrom ME: **Metabolic engineering in methanotrophic bacteria.** *Metab Eng* 2015, **29**:142–52.

75. Martín HG, Kumar VS, Weaver D, Ghosh A, Chubukov V, Mukhopadhyay A, Arkin A, Keasling JD: **A Method to Constrain Genome-Scale Models with 13C Labeling Data.** *PLoS Comput Biol* 2015, **11**:e1004363.

76. Chu F, Lidstrom ME: **XoxF Acts as the Predominant Methanol Dehydrogenase in the Type I Methanotroph Methylomicrobium buryatense**. *J Bacteriol* 2016, **198**:1317–1325.

77. Fu Y, Beck DAC, Lidstrom ME: **Difference in C3-C4 metabolism underlies tradeoff between growth rate and biomass yield in Methylobacterium extorquens AM1.** *BMC Microbiol* 2016, **16**:156.

78. Henard CA, Smith H, Dowe N, Kalyuzhnaya MG, Pienkos PT, Guarnieri MT: **Bioconversion of methane to lactate by an obligate methanotrophic bacterium.** *Sci Rep* 2016,

6(February):21585.

79. Chu F, Lidstrom ME: **XoxF Acts as the Predominant Methanol Dehydrogenase in the Type I Methanotroph Methylomicrobium buryatense.** *J Bacteriol* 2016, **198**:1317–25.

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

81. Schendel FJ, Bremmon CE, Flickinger MC, Guettler M, Hanson RS: **L-lysine production at 50 degrees C by mutants of a newly isolated and characterized methylotrophic Bacillus sp.** *Appl Environ Microbiol* 1990, **56**:963–70.

82. Müller JEN, Meyer F, Litsanov B, Kiefer P, Vorholt JA: **Core pathways operating during methylotrophy of Bacillus methanolicus MGA3 and induction of a bacillithiol-dependent detoxification pathway upon formaldehyde stress.** *Mol Microbiol* 2015, **98**:1089–100.

83. Fu Y, Li Y, Lidstrom M: **The oxidative TCA cycle operates during methanotrophic growth of the Type I methanotroph Methylomicrobium buryatense 5GB1,***Metabolic Engineering, 2016 (under review)*

84. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox**. *Nat Protoc* 2007, **2**:727–738.

85. Hyduke D, Schellenberger J, Que R, Fleming R, Thiele I, Orth J, Feist A, Zielinski D, Bordbar A, Lewis N, Rahmanian S, Kang J, Palsson B: **COBRA Toolbox 2.0**. *Protocol Exchange* 2011.

86. Hay JO, Shi H, Heinzel N, Hebbelmann I, Rolletschek H, Schwender J: **Integration of a constraint-based metabolic model of Brassica napus developing seeds with (13)C-metabolic flux analysis.** *Front Plant Sci* 2014, **5**(December):724.

87. Henard CA, Smith H, Dowe N, Kalyuzhnaya MG, Pienkos PT, Guarnieri MT, Anderson B, Fei Q, Conrado RJ, Gonzalez R, Haynes CA, Gonzalez R, Chistoserdova L, Kalyuzhnaya MG, Lidstrom ME, Strong PJ, Xie S, Clarke WP, Bothe H, Harrison DE, Hamer G, Khosravi-Darani K, Mokhtari Z-B, Amai T, Tanaka K, Ye RW, Tao L, Sharpe PL, Ivanova EG, Fedorov DN, et al.: **Bioconversion of methane to lactate by an obligate methanotrophic bacterium**. *Sci Rep*

2016, **6**:21585.

88. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR: **COBRApy: COnstraints-Based Reconstruction and Analysis for Python**. .