A Genome-Scale Metabolic Model for *Methylococcus capsulatus* (Bath) Suggests Reduced Efficiency Electron Transfer to the Particulate Methane Monooxygenase

Lieven, Christian; Petersen, Leander A. H.; Jørgensen, Sten Bay; Gernaey, Krist V.; Herrgard, Markus J.; Sonnenschein, Nikolaus

*Published in:*  
Frontiers in Microbiology

*Link to article, DOI:*  
10.3389/fmicb.2018.02947

*Publication date:*  
2018

*Document Version*  
Publisher's PDF, also known as Version of record

*Link back to DTU Orbit*

*Citation (APA):*  
https://doi.org/10.3389/fmicb.2018.02947

General rights  
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
A Genome-Scale Metabolic Model for *Methylococcus capsulatus* (Bath) Suggests Reduced Efficiency Electron Transfer to the Particulate Methane Monoxygenase

Christian Lieven1*, Leander A. H. Petersen2,3, Sten Bay Jørgensen2, Krist V. Gernaey2, Markus J. Herrgard1 and Nikolaus Sonnenschein1

1 The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark, 2 Unibio A/S, Kongens Lyngby, Denmark, 3 Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

Background: Genome-scale metabolic models allow researchers to calculate yields, to predict consumption and production rates, and to study the effect of genetic modifications *in silico*, without running resource-intensive experiments. While these models have become an invaluable tool for optimizing industrial production hosts like *Escherichia coli* and *S. cerevisiae*, few such models exist for one-carbon (C1) metabolizers.

Results: Here, we present a genome-scale metabolic model for *Methylococcus capsulatus* (Bath), a well-studied obligate methanotroph, which has been used as a production strain of single cell protein (SCP). The model was manually curated, and spans a total of 879 metabolites connected via 913 reactions. The inclusion of 730 genes and comprehensive annotations, make this model not only a useful tool for modeling metabolic physiology, but also a centralized knowledge base for *M. capsulatus* (Bath). With it, we determined that oxidation of methane by the particulate methane monoxygenase could be driven both through direct coupling or uphill electron transfer, both operating at reduced efficiency, as either scenario matches well with experimental data and observations from literature.

Conclusion: The metabolic model will serve the ongoing fundamental research of C1 metabolism, and pave the way for rational strain design strategies toward improved SCP production processes in *M. capsulatus*.

Keywords: COBRA, genome-scale metabolic reconstruction, C1 metabolism, single cell protein, methanotrophy, constraint-based reconstruction and analysis

INTRODUCTION

The Gram-negative, obligate-aerobe *Methylococcus capsulatus* (Bath) is a methane oxidizing, gamma-proteobacterium. Since its initial isolation by Foster and Davis (1966), the organism has been subject to a wide array of studies. The global role of *M. capsulatus* as a participant in the carbon cycle has been elucidated (Hanson and Hanson, 1996; Jiang et al., 2010) as well as its effects on
human (Indrelid et al., 2017) and animal health and disease (Kleivelend et al., 2013). Specifically the latter studies have been triggered by a considerable commercial interest in *M. capsulatus* (Bath) as the primary microbe used for the production of Single Cell Protein (SCP) as animal feed starting in the 70 s (Överland et al., 2010). Now that hydraulic fracturing has made natural gas a cheap and abundant feedstock (Ritala et al., 2017), the application of *M. capsulatus* (Bath) for this purpose is being explored again (Nunes et al., 2016; Petersen et al., 2017). Another portion of studies, however, has focused on uncovering the biochemical and genetic basis of the organism’s unique metabolism (Anthony, 1983). Yet, the greatest interest has been the role and function of the initial enzyme in methanotrophy, methane monooxygenase (Ross and Rosenzweig, 2017), responsible for oxidation of methane to methanol.

*Methylococcus capsulatus* (Bath) is able to express two distinct types of methane monoxygenases: a soluble form of methane monoxygenase (sMMO) and a particulate, or membrane-bound form (pMMO). The expression of these enzymes is strongly influenced by the extracellular concentration of copper; when *M. capsulatus* (Bath) is grown in the presence of low levels of copper the sMMO is active, while the pMMO is predominantly active at high levels. Both enzymes require an external electron donor to break a C-H bond in methane. While the electron donor for the sMMO is NADH (Colby and Dalton, 1978, 1979; Lund et al., 1985; Blazyk and Lippard, 2002), the native reductant to the pMMO has not yet been elucidated due to difficulties to purify the enzyme and assay its activity in vitro (Ross and Rosenzweig, 2017). Three hypotheses regarding the mode of electron transfer to the pMMO have been suggested previously:

1. In the redox-arm mode (Dawson and Jones, 1981), the methanol dehydrogenase (MDH) passes electrons via cytochrome c555 (cL) (Anthony, 1992) and cytochrome c553 (cH) (DiSpirito et al., 2004) to either a CBD- or AA3-type terminal oxidase (Larsen and Karlsen, 2016) and thus contributes to building up a proton motive force (PMF) and the synthesis of ATP (Figure 1). The electrons required for the oxidation of methane are provided through ubiquinone (Q8H2). Reactions downstream of MDH responsible for oxidizing formaldehyde to CO2 feed electrons into the Q8 pool. This includes dye-linked formaldehyde dehydrogenase (DL-FALDH), formate dehydrogenase (FDH) and any NAD reducing reactions which ultimately contribute electrons through the NADH dehydrogenases (NDH-1, NDH-2, NQR). See Figure 1. Although no binding site has been identified, there is support for pMMO reduction by endogenous quinols (Shiemke et al., 2004).

2. With the direct coupling mode, the MDH is able to directly pass electrons to the pMMO (Wolfe and Higgins, 1979; Leak and Dalton, 1983; Culpepper and Rosenzweig, 2014). Here, cytochrome c555 is the electron donor to the pMMO instead of ubiquinol. This mode is supported by results from cryoelectron microscopy which indicates that the pMMO and the MDH form a structural enzyme complex (Myronova et al., 2006).

3. The uphill electron transfer mode supposes that the electrons from cytochrome c553 can reach the ubiquinol-pool facilitated by the PMF at the level of the ubiquinol-cytochrome-c reductase. This mode was proposed by Leak and Dalton (1986b) as it could explain the observed reduced efficiency.

A genome-scale metabolic model (GEM) not only represents a knowledge base that combines the available physiological, genetic and biochemical information of a single organism (Thiele and Palsson, 2010), but also provides a testbed for rapid prototyping of a given hypothesis (Benedit et al., 2012). Hence, we present the first manually curated GEM for *M. capsulatus* Bath, with the intent of supplying the basis for hypothesis-driven metabolic discovery and clearing the way for future efforts in metabolic engineering (Kim et al., 2015). Using the GEM, we investigate the nature of electron transfer in *M. capsulatus* (Bath) by comparing the model’s predictions against experimental data from Leak and Dalton (1986a). Furthermore, we compare its predictions to those of the model of *Methylobacterium buryatense* 5G(B1) (de la Torre et al., 2015) and explain notable differences by considering the proposed electron transfer modes.

**RESULTS AND DISCUSSION**

**Reconstruction**

The presented genome-scale metabolic reconstruction of *M. capsulatus* (Bath) termed iMcBath is based on BMID000000141026, an automatic draft published as part of the Path2Models project (Büchel et al., 2013). The whole genome sequence of *M. capsulatus* (Bath) (GenBank AE017282.2; Ward et al., 2004) was used to aid the curation process and to supply annotations (see section “Materials and Methods”). This metabolic reconstruction consists of 843 enzymatic reactions that interconvert and transport 759 unique metabolites. The total number of reactions, including exchange, demand and the biomass reactions, is 913. The model attributes reactions and metabolites to four distinct compartments: The cytosol, the periplasm, an intramembranous compartment and the medium, which is referred to as extracellular in the model. Gene-Protein-Reaction rules (GPR) associated with 730 unique genes support 86.97% of the included reactions with, leaving 119 reactions without a GPR. The GPRs include representation of 43 enzyme complexes.

The model and all scripts used during reconstruction are made available publicly on Github at https://github.com/ChristianLieven/memote-m-capsulatus. The model is available in the community-standard SBML format (Level 3 Version 2 with FBC; Olivier and Bergmann, 2015) and a JSON format that is native to cobrapy (Ebrahim et al., 2013).

**Biomass Reaction**

In stoichiometric models biological growth is represented as the flux through a special demand reaction. The so-called biomass reaction functions as a drain of metabolites, which are either highly reduced non-polymeric macromolecules such as lipids and steroids or precursors to typical biopolymers such as nucleic acids, proteins or long-chain carbohydrates. The stoichiometry of an individual precursor was calculated from the principal composition of *M. capsulatus* (Bath) as reported
The growth-associated maintenance (GAM) and the non-growth associated maintenance (NGAM) requirements for *M. capsulatus* are yet to be determined experimentally. Therefore, a GAM value of 23.087 mmol ATP gDW⁻¹ h⁻¹ was estimated according to Thiele and Palsson (2010) based on the data for *Escherichia coli* published by Neidhardt et al. (1990). The value for GAM is expected to increase with the growth rate of the cells (Varma et al., 1993). Like de la Torre et al. (2015) had done for *M. buryatense* 5G(B1), we assumed the NGAM of *M. capsulatus* (Bath) to be similar to that of *E. coli* thus setting it to 8.39 mmol ATP gDW⁻¹ h⁻¹ (Feist et al., 2010).

**Metabolism**

To study which of the three modes of electron transfer is active in *M. capsulatus* (Bath), they were each implemented in the model. The implementation is illustrated in Figure 2. To include the **redox-arm** we implemented the reaction representing the particulate methane monooxygenase, in the model termed as PMMOiipp, utilizing Q8H2 as a cofactor. Accordingly, a variant of the pMMO reaction was added to account for a possible direct coupling to the MDH. In this variant reaction, termed PMMOSDipp, the cofactor is cytochrome c555, represented as the metabolite focyc555_p in the model. To enable an **uphill electron transfer**, the reaction representing the ubiquinol-cytochrome-c reductase (CYOR_q8ppi in the model), was constrained to be reversible while keeping PMMOiipp active.

Following the path of carbon through metabolism downstream from the MDH, the model includes both the reaction for a ubiquinone-dependent formaldehyde dehydrogenase (Zahn et al., 2001), termed FALDDHiipp, and an NAD-dependent version, termed ALDD1 (Figure 3). Despite the initial evidence for the latter reaction (Tate and Dalton, 1999) having been dispelled by Adeosun et al. (2004), it was added to allow further investigation into the presence of a putative enzyme of that function. An additional pathway for formaldehyde oxidation represented in both genome and model is the tetrahydromethanopterin(THMPT)-linked pathway (Vorholt, 2002).

Formaldehyde assimilation in *M. capsulatus* (Bath) occurs primarily through the ribulose monophosphate (RuMP)-pathway. As outlined by Anthony (1983), the RuMP-pathway has four hypothetical variants. Based on the annotated genome published by Ward et al. (2004), we identified
Three different formaldehyde oxidation pathways are represented in the model. Black text denotes the common name, while faint gray text denotes the reaction ID in the metabolic model.

All four variants of the RuMP pathway are represented in the metabolic model. If there is a common enzyme name the black text denotes the common name, while faint gray text denotes the reaction ID in the metabolic model. Otherwise the black text is also the reaction ID in the model.

Not only both C6 cleavage pathways depending either on the 2-keto, 3-deoxy, 6-phosphogluconate (KDPG) aldolase (EDA) or the fructose bisphosphate aldolase (FBA), but also the transaldolase (TA) involved in the rearrangement phase that regenerates ribulose 5-phosphate. The alternative to a transaldolase-driven rearrangement step is the use of a sedoheptulose bisphosphatase, which was not included in the initial annotation. Strøm et al. (1974) could not detect specific activity using cell-free preparations. Yet, we decided to add a corresponding reaction for two reasons. First, the FBA has been characterized to reversibly catalyze sedoheptulose bisphosphate cleavage (Rozova et al., 2010), which is reflected by the reaction FBA3 in the model. Second, the pyrophosphate-dependent 6-phosphofructokinase (PFK_3_ppi) was reported to have higher affinity and activity to the reversible phosphorylation of sedoheptulose phosphate than compared to fructose 6-phosphate (Reshetnikov et al., 2008). Thus, all of the resulting four combinations that make up the RuMP-pathway are represented in this metabolic model (Figure 4).
The genome of *M. capsulatus* (Bath) encodes genes for a complete Calvin Benson Bassham (CBB) cycle (Taylor, 1977; Taylor et al., 1980; Baxter et al., 2002) and a partial Serine pathway for formaldehyde assimilation (Ward et al., 2004). It was argued by Taylor et al. (1981) that *M. capsulatus* (Bath) can metabolize glycolate [a product of the oxygenase activity of the ribulose bisphosphate carboxylase (RBPC)] via this pathway. Both Taylor and Ward, further suggested the presence of unique key enzymes to complete the Serine pathway, such as hydroxymethyl transferase, hydroxypyruvate reductase and malate-CoA lyase. However, since the gene annotation did not reflect this and the RuMP pathway is reportedly the main pathway for formaldehyde assimilation (Kelly et al., 2005), these putative reactions were not included.

All genes of the TCA cycle were identified in the genome sequence and all associated reactions were included accordingly (Ward et al., 2004). Because no activity of the 2-oxoglutarate dehydrogenase has so far been measured *in vivo* (Wood et al., 2004; Kelly et al., 2005), the associated reactions have been constrained to be blocked (both lower and upper bounds were set to zero). This way they can easily be activated if needed. For instance, if a growth condition is discovered where activity for this enzyme can be detected.

Based on reactions already present in BMID000000141026, the information in the genome annotation, and the measured biomass composition, we curated the biosynthetic pathways of all proteinogenic amino acids, nucleotides, fatty acids, phospholipids, panthothenate, coenzyme A, NAD, FAD, FMN, riboflavin, thiamine, myo-inositol, heme, folate, cobalamine, glutathione, squalene, lanosterol, peptidoglycan. Since no corresponding genes could be identified, reactions catalyzing the biosynthesis of lipopolysaccharide (LPS) were adopted from iJO1366 (Orth et al., 2011) under the assumption that the biosynthesis steps among gram-negative bacteria require amounts of ATP comparable to *E. coli*.

*Methylococcus capsulatus* (Bath) is able to metabolize the nitrogen sources ammonium (NH$_4$) and nitrate (NO$_3$) in a variety of ways. When the extracellular concentration of NH$_4$ is high (>1 mM), alanine dehydrogenase (ADH) is the primary pathway for nitrogen assimilation into biomass, when it is low (<1 mM) assimilation is carried out via the glutamine synthetase/glutamine synthase (GS/GOGAT) pathway (Murrell and Dalton, 1983). In addition to assimilation, the two monooxygenases are able to oxidize ammonium to hydroxylamine (Colby and Dalton, 1978; Bédard and Knowles, 1989), which is then further oxidized by specific enzymes first to nitrite (Bergmann et al., 1998; Figure 5), and even to dinitrogen oxide (Campbell et al., 2011). NO$_3$ is reduced to NH$_4$ via nitrite and ultimately assimilated via GS/GOGAT. Furthermore, it has been shown that *M. capsulatus* (Bath) is able to fix atmospheric nitrogen (N$_2$) (Murrell et al., 1983). The nitrogenase gene cluster has been identified (Oakley and Murrell, 1991) and annotated accordingly (Ward et al., 2004), and the corresponding reactions have been included in the model. As the enzyme has not yet been specifically characterized, the nitrogenase reaction was adapted from iAF987 (Feist et al., 2014). A schema showing these reactions side-by-side is displayed in Figure 6.

![Figure 5 | Oxidation of ammonia to nitrite.](image)

A metabolic map of all the reactions in the model was built using Escher (King et al., 2015) and is available as Supplementary Figure 1.

**Prediction of Transport Reactions**

As a last step we added transport reactions that were not in the draft reconstruction. Inferring membrane transport reactions from the genome sequence is difficult, as usually the precise 3D structure of transport proteins dictates which metabolite classes can be transported (Mishra et al., 2014). Even if the substrates are known, the energy requirements of transport are often undefined. Working on protein sequence matches using PsortB 3.0 (Yu et al., 2010), combined with BLAST (Altschul et al., 1990) matches against TransportDB (Elbourne et al., 2017) and the Transporter Classification Database (TCDB) (Saier et al., 2006), we were able to identify 56 additional transport reactions. We have limited the number of transporters to be added, focusing specifically on transporters with known mechanisms and transport of metabolites already included in the model. A list of putative transport-associated genes that we did not consider is available at https://github.com/ChristianLieven/memote-m-capsulatus. Ward et al. (2004) hypothesized that some of these genes may facilitate the uptake of sugars for growth. Kelly et al. (2005) argue that these genes instead may allow a...
function similar to *Nitrosomas europaea*, which is able to grow on fructose as a carbon-source with energy from ammonia oxidation (Hommes et al., 2003; Kelly et al., 2005). Thus, this list is a good starting point to study potential alternate carbon source use by *M. capsulatus* (Bath).

**Parameter Fitting to Determine the Mode of Electron Transfer**

To determine which combination of the three aforementioned electron transfer modes is active in *M. capsulatus* (Bath), we constrained the model based on experiments conducted by Leak and Dalton (1986a). Since the three modes relate to how the pMMO receives electrons, we focused on the data generated by growing *M. capsulatus* (Bath) in high-copper medium, which is the condition in which pMMO is predominantly active. We used the average of carbon and oxygen-limited measurements as a reference. Having constrained the model, we compared the Leak and Dalton’s measurements for the ratio of oxygen consumption to methane consumption ($O_2/CH_4$) to the predictions of the model (Figure 7). We considered the $O_2/CH_4$ ratio to be a key metric for the respiratory chain in *M. capsulatus* (Bath), as it is a function of the mode of electron transfer to the pMMO. The central carbon metabolism was left unconstrained.

Under the assumption that the mode of electron transfer to pMMO would be independent of the source of nitrogen, we compared the $O_2/CH_4$ ratios of the model constrained to employ one of the three modes of electron transfer exclusively to the corresponding reference values of *M. capsulatus* (Bath) grown on either NO$_3$ or NH$_4$. However, neither of the modes adequately represented the measured $O_2/CH_4$ ratios of about 1.43 and 1.6, respectively. Although Leak and Dalton had proposed that the reverse or uphill electron transfer is the most probable mode (Leak and Dalton, 1986b), the model predictions allowing for an unbounded uphill transfer did not support this (Figure 7E), as the efficiency was almost comparable to predictions using direct coupling.

We altered the efficiency of the three modes to determine whether the fit could be improved. For the redox arm, we gradually decreased the mol protons required for the synthesis of 1 mol ATP, thereby improving the efficiency. This did not change the $O_2/CH_4$ ratio (see Figure 7A). de la Torre et al. (2015) constructed a GEM for *M. buryatense* 5G(B1) to investigate growth yields and energy requirements in different conditions. Similar to the results presented herein, they found that the redox-arm mode correlated least with their experimental data for *M. buryatense*.

We decreased the efficiency of the direct coupling mode (PMMODCipp) by forcing a portion of the flux through the regular pMMO reaction (PMMOipp) using ratio constraints (Figure 7B). To achieve a ratio representing the measured value of 1.43 required a large decrease in efficiency with almost 85% of the incoming carbon routed through the regular pMMO reaction (see Supplementary Table 2). Lastly, we iteratively constrained the lower bound of the reaction associated with the ubiquinol-cytochrome-c reductase (CYOR_q8ppi), to reduce the efficiency of the uphill-electron transfer (Figure 7C).

Leak and Dalton (1986b) developed mathematical models for each mode based on previous calculations (Harder and Van Dijken, 1976; Anthony, 1978). Their models lead them to conclude that both direct coupling and uphill electron transfer operating at reduced efficiency may account for the observed high ratios of $O_2/CH_4$, which agrees with our predictions. Since experimental
results on the nature of protein–protein interactions between pMMO and MDH seem to be inconclusive (Myronova et al., 2006; Culpepper and Rosenzweig, 2014; Ross and Rosenzweig, 2017), neither of the modes can be ruled out entirely.

Parameter Fitting to Determine the Impact of NH₄ Oxidation

Leak and Dalton pointed out, that the unexpectedly high O₂/CH₄ ratio of 1.6 was the product of latent NH₄ oxidation rather than assimilation leading to elevated levels of NO₂ which detected at medium concentrations of around 200 µM (Leak and Dalton, 1986a). They were uncertain whether this increase in the ratio of O₂/CH₄ could be attributed to the energetic burden of oxidizing NH₄ or an uncoupling effect of NO₂.

To investigate this effect, we introduced a ratio constraint (see Methods) coupling the uptake of NH₄ to the excretion of NO₂ and explored a number of values for this ratio (Figure 7C). According to the simulations, the energy spent oxidizing about 50% of incoming NH₄ to NO₂ is sufficient to account for the observed, high O₂/CH₄ ratio of 1.6. Although this shows that the loss of energy could be significant enough to account for the increased ratio, this does not exclude a potential combined effect because of energy decoupling.
Regardless it shows that calculations using the metabolic model can accurately reflect the in vivo behavior of *M. capsulatus* (Bath).

**Validation of the Model**

Specific growth rates of 0.25 and 0.37 h\(^{-1}\) were measured by Joergensen and Degen (1987) for *M. capsulatus* (Bath) on copper-free and copper-rich NO\(_3\) mineral salt medium, respectively. Simulations of iMcBath with an active sMMO reaction yielded 0.19 h\(^{-1}\) on NO\(_3\) and 0.26 h\(^{-1}\) on NH\(_4\) as the nitrogen source, as well as 0.20 h\(^{-1}\) for growth driven by the pMMO on NO\(_3\) and 0.30 h\(^{-1}\) on NH\(_4\). The predicted growth rates scale together with the uptake rates for methane and the respective nitrogen sources. The uptake rate for methane in iMcBath was constrained to 18.46 mmol g\(^{-1}\) h\(^{-1}\), a value adopted from de la Torre et al. (2015) due to the lack of a specific measurement for *M. capsulatus* (Bath). Uptake rates for NO\(_3\) and NH\(_4\) have been left unconstrained for the same reasons. An adjustment of the scale will be possible as soon as these data become available.

The model qualitatively agrees with a report from Patel and Hoare (1971), who determined that *M. capsulatus* (Bath) is capable of using certain amino acids as sources for nitrogen (Table 1). Although the transport reactions for these amino acids could not be inferred, the addition of corresponding boundary reactions sufficed to predict growth using FBA. However, the predicted growth rates are not comparable as it seems that in the model some of the carbon is assimilated as the amino acids are ultimately converted into pyruvate. On L-glutamate and L-glutamine as N-sources the model further predicts the production of alpha-ketoglutarate, and on L-valine that of alpha-ketoisovalerate. This is in accordance with findings by Patel and Hoare (1971).

### Table 1 | Model validation against experimental conditions from literature.

<table>
<thead>
<tr>
<th>Growth condition</th>
<th>Predicted growth rate [h(^{-1})]</th>
<th>Measured growth rate [h(^{-1})]</th>
</tr>
</thead>
<tbody>
<tr>
<td>sMMO – NO(_3)</td>
<td>0.18</td>
<td>0.12(^1)</td>
</tr>
<tr>
<td>sMMO – NH(_4)</td>
<td>0.23</td>
<td>0.09(^1)</td>
</tr>
<tr>
<td>sMMO – L-Alanine</td>
<td>1.0</td>
<td>0.08(^1)</td>
</tr>
<tr>
<td>sMMO – L-Aspartate</td>
<td>1.0</td>
<td>0.11(^1)</td>
</tr>
<tr>
<td>sMMO – L-Asparagine</td>
<td>1.0</td>
<td>0.11(^1)</td>
</tr>
<tr>
<td>sMMO – L-Cysteine</td>
<td>1.0</td>
<td>0.09(^1)</td>
</tr>
<tr>
<td>sMMO – L-Glutamate</td>
<td>0.47</td>
<td>0.11(^1)</td>
</tr>
<tr>
<td>sMMO – L-Glutamine</td>
<td>0.47</td>
<td>0.11(^1)</td>
</tr>
<tr>
<td>sMMO – L-Valine</td>
<td>0.46</td>
<td>0.07(^1)</td>
</tr>
</tbody>
</table>

\(^1\) Pateland Hoare (1971) converted from Doubling Time [h\(^{-1}\)].\(^2\) Baxter et al. (2002).

To further validate iMcBath, we simulated autotrophic growth on carbon dioxide and molecular hydrogen (H\(_2\)) as observed previously by Baxter et al. (2002). Indeed, growth was possible using carbon dioxide as the C-source and H\(_2\) as the source of energy (Table 1). Replacement of H\(_2\) with formate also allowed the production of biomass, although formate is both assimilated and oxidized. This is because the model allows for the consumption of cytosolic carbon dioxide created as the end-product of formate oxidation. In addition to the ribulose-bisphosphate carboxylase converting CO\(_2\) to 3-phosphoglycerate, this mode of operation relies on the pyrophosphate-dependent 6-phosphofructokinase (PFK\(_3\)-ppi) to reversibly convert sedoheptulose-1,7-bisphosphate to sedoheptulose-7-phosphate and fructose bisphosphate to fructose-6-phosphate thus providing these intermediates to the RuMP cycle (Reshetnikov et al., 2008). Growth just on carbon dioxide or molecular hydrogen exclusively could not be simulated as reported by the authors.

Flux balance analysis confirms the results from Stanley and Dalton (1982) predicting latent CO\(_2\) fixation even during growth on methane.

### Comparison With Other Models

We compared iMcBath, the preceding automated draft reconstruction BMID000000141026, and a GEM of the gram-negative, gamma-proteobacterium *M. buryatense* strain 5G(B1) (de la Torre et al., 2015) to illustrate how much the model has progressed from the draft through manual curation and how the mode of electron transfer affects growth parameters within the group of gamma-proteobacteria (see Table 2).

Unsurprisingly, the automated draft generally performs quite poorly in comparison with the curated models. It’s not possible to produce biomass from methane, even in rich media conditions, which is indicative of gaps in the pathways leading toward individual biomass components. Moreover, ATP can be produced without the consumption of substrate, which in turn means that key energy reactions are not constrained thermodynamically to have the correct reversibility. In the draft, 51% of the reactions are not supported by Gene-Protein-Reaction rules (GPR), while in iMb5G(B1) this percentage is only 32% and in iMcBath the percentage of these reactions is under 20%. GPRs allow modelers to distinguish between isozymes and enzyme complexes by connecting gene IDs either through boolean “OR” or “AND” rules. In iMcBath, 25 complexes in total were curated and formulated as GPR. Neither the draft model nor iMb5G(B1) make this distinction.

In the automated draft, the oxidation of methane was only possible through a reaction representing the sMMO (MNXR6057). In iMcBath, this was corrected by also implementing reactions that represent the pMMO, one with ubiquinone as the electron donor (PMMOipp), and another that receives electrons from the MDH (PMMODCipp). *M. capsulatus* (Bath), like *M. buryatense* 5G(B1), expresses both the soluble and the particulate monooxygenase depending on the availability of copper in the medium. In iMb5G(B1), however, only the particulate monooxygenase is represented by the reaction “pMMO.” The ability of *M. capsulatus* (Bath). \[\text{CO}_2 + \text{H}_2 \rightarrow \text{NO}_3 \]
M. capsulatus on nitrate. While the draft model for adding the respective transport reactions iMcB5G(B1) only grows except for urea, it couldn’t grow at all, which again is likely contains exchange reactions for all the tested nitrogen sources is necessary to verify this in vivo former of which is then assimilated. Further experimental work and broken down into ammonia and carbon dioxide, the consequence of this, urea can be taken up into the cytosol when curating the cobalamine biosynthesis pathway. As a possible in the model since we identified MCA1662 only be possible on low concentrations. Growth on urea is

1986a; Nyerges and Stein, 2009), hence 
in vivo growth may only be possible on low concentrations. Growth on urea is possible in the model since we identified MCA1662 as an ATP-driven urea transporter and gap-filled a urease reaction when curating the cobalamine biosynthesis pathway. As a consequence of this, urea can be taken up into the cytosol and broken down into ammonia and carbon dioxide, the former of which is then assimilated. Further experimental work is necessary to verify this in vivo. M. buryatense 5G(B1) is reported to grow on nitrate, ammonia and urea, yet without adding the respective transport reactions iMb5G(B1) only grows on nitrate. While the draft model for M. capsulatus (Bath) contains exchange reactions for all the tested nitrogen sources except for urea, it couldn’t grow at all, which again is likely because of gaps in the pathways leading to biomass precursor metabolites.

The difference between M. capsulatus (Bath) and M. buryatense 5G(B1) becomes apparent when comparing the growth energetics of iMcBath and iMb5G(B1). iMb5G(B1) produces more mmol gDW$^{-1}$ h$^{-1}$ ATP than iMcBath. Because of this the hypothetical growth-rates predicted by iMb5G(B1) are higher than those of iMcBath regardless of the respective nitrogen source. This difference is likely a direct consequence of the mode of electron transfer to the monooxygenase and thus the efficiency of the methane oxidation reaction in total. When comparing the ratio of the uptake rate of oxygen and the uptake rate of methane for the two models, we can see that the resulting values in iMb5G(B1) are lower than in iMcBath. It was recently reported, that instead of the reverse-electron transfer and redox-arm mode active in M. capsulatus (Bath), a mixture of direct coupling from pMMO to MDH and reverse electron transfer seems to be the most likely mode in M. buryatense 5G(B1) (de la Torre et al., 2015).

**CONCLUSION**

iMcBath is the first, manually curated, GEM for M. capsulatus (Bath). With iMcBath, we combine biochemical knowledge of half a century of research on M. capsulatus (Bath) into a single powerful resource, providing the basis for targeted metabolic engineering, process design and optimization, omic-data contextualization and comparative analyses. We applied the metabolic model to study the complex electron transfer chain of M. capsulatus, by analyzing the three modes of electron transfer that had been proposed previously (Leak and Dalton, 1986b). We did so by corresponding each mode with the flux through a reaction in the model, and consequently comparing the predicted O$_2$/CH$_4$ ratios to experimentally measured values by Leak and Dalton (1986a). Simulation and experiment agreed either when the model was constrained to employ the uphill electron transfer at reduced efficiency or direct coupling at strongly reduced efficiency for M. capsulatus (Bath) grown in silico on NO$_3$ as the source of nitrogen. Further experimental validation is required as neither mode can be ruled out exclusively. The experimentally observed effect of NH$_4$ oxidation to NO$_2$ could be replicated by considering the energy burden alone.

Future applications of the metabolic model could include hypothesis testing of the regulation of the MMO in other growth conditions (Kelly et al., 2005), studying the effects of the predicted hydrogenases on the energy metabolism of M. capsulatus (Bath) (Ward et al., 2004), and exploring venues of metabolic engineering for an improved production of metabolites (Kalyuzhnaya, 2016). Improved validation of the model requires more experimental data, which is surprisingly sparse. Specific uptake and growth rates in various conditions, as well as gene deletion studies are paramount to improving iMcBath. The authors would like to reach out to any experts on M. capsulatus to get involved in maintaining and improving iMcBath. The model and additional data is hosted publicly on Github$^1$.

$^1$https://github.com/ChristianLieven/memote-m-capsulatus

**TABLE 2 | Model Comparison.**

<table>
<thead>
<tr>
<th>Structure</th>
<th>iMcBath</th>
<th>BMID000000141026</th>
<th>iMb5G(B1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane oxidation</td>
<td>SMMO1,</td>
<td>MNXR6057</td>
<td>pMMO</td>
</tr>
<tr>
<td></td>
<td>PMMDClpp,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PMMDClpp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth on N-sources</td>
<td>No Growth</td>
<td>NO$_3$</td>
<td>NO$_3$</td>
</tr>
<tr>
<td></td>
<td>Urea, NO$_2$, NO$_3$, NH$_4$, N$_2$, Ala, Asn, Asp, Cys, Gln, Glt, Val</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The presented reconstruction iMcBath, the automated draft BMID000000141026 and iMb5G(B1), a genome-scale reconstruction of Methylomicrobium buryatense strain 5G(B1).
contributions are welcome regardless of whether they come in the form of pull requests, issues, comments, or suggestions.

MATERIALS AND METHODS

Model Curation
Starting reconstruction on the basis of an automated draft required additional effort to be able to use it for calculations. The automated draft used two sets of ID namespaces, BiGG (King et al., 2016) and MetaNetX (MNX) (Ganter et al., 2013). Hence, the first step in the curation efforts consisted of unifying the namespace by mapping all metabolite and reaction identifiers from MNX into the human-readable BiGG namespace. Individual metabolites and reactions with unmappable IDs, that could not be identified in the BiGG database and for which there was little evidence in the form of GPR rules, were removed from the model. Several metabolite formulas contained polymer units, and many reactions lacked EC numbers. Using the MNX spreadsheet exports “chem_prop.tsv” and “reac_prop.tsv” from version 1.0 (Bernard et al., 2014) and 2.0 (Moretti et al., 2016) most of these issues were resolved. Due to said malformed and missing metabolite formulae, many reactions were not mass-charge-balanced. We used the “check_mass_balance” function in cobrapy (Ebrahim et al., 2013) to identify and balance 99.4% of the reactions in the model. The remaining five reactions, belonging to fatty acid biosynthesis, involve a lumped metabolite that represents the average fatty acid content of M. capsulatus (Bath).

After mapping the reaction and metabolite identifiers from the MetaNetX namespace to the BiGG namespace, we proceeded with the curation efforts as follows: First, we chose a subsystem of interest, then we picked a pathway and using information from either the genome sequence, published articles, the metacyc or uniprot databases, and lastly, we enriched each enzymatic step with as much information as possible. Information that was added included for instance: GPR, reaction localization, EC numbers, a confidence score, possible cofactors and inhibitors and cross references to other databases such as KEGG, BIGG and MetaNetX. For each metabolite involved in these reactions, we defined the stoichiometry, charge and elemental formula, either based on the corresponding entries in the BiGG database or on clues from literature.

If reactions from a pathway were present in the draft, we checked their constraints and directionality, consulting the corresponding reactions in MetaCyc (Caspi et al., 2014). This was necessary as many of the irreversible reactions in the draft reconstruction seemed to have been “inverted” when compared to the corresponding reactions in the reference databases, which made flux through them impossible in normal growth conditions.

The energy metabolism and methane oxidation were curated first. Except for the reaction representing the sMMO, all reactions were newly constructed, as they were absent in the draft. Then, in order to achieve sensible FBA solutions for growth on methane, the central carbon metabolism, amino acid and nucleotide biosynthesis pathways were manually curated. Simultaneous to the manual curation a metabolic pathway map was created in Escher, which helped us to maintain a visual checklist of curated pathways.

The automated draft contained a rudimentary, generic biomass reaction, which only accounted for the production of proteins, DNA and RNA, but not for the biosynthesis of a cell wall and cell membrane, the maintenance of a cofactor pool, the turnover of trace elements or the energetic costs associated with growth. After calculating a more specific biomass composition for M. capsulatus (Bath), further pathway curation was necessary to achieve growth in silico. This included the biosynthesis pathways of fatty acids (even, odd and cyclopropane varieties), phospholipids, coenzyme A, Ubiquinol, Lanosterol, FAD, FMN, Riboflavin, NAD and NADP, Heme, Glutathione, Cobalamin, Thiamine, Myo-Inositol, and Lipopolysaccharides. The corresponding biosynthetic pathways were gap-filled and manually curated. To identify the appropriate pathways, MetaCyc pathway maps from the M. capsulatus (Bath) specific database were used to compare the reactions that were present in the draft reconstruction.

To account for the reported differences in ammonia assimilation of M. capsulatus (Bath) when grown in the presence of excess ammonia versus the growth on either atmospheric nitrogen or nitrate, we curated the nitrogen metabolism including the oxidation of ammonia to nitrite via hydroxylamine, the reduction of nitrate and nitrite, the glutamine synthetase/ glutamate synthase reactions and the ADH. Reversible degradation reactions producing ammonia that would potentially bypass the characterized ammonia assimilation pathways were constrained to be irreversible accordingly.

After we had enriched the annotations already in the draft with annotations from the metabolic models iJO1366 (Orth et al., 2011), iRC1080 (Chang et al., 2011), iJN678 (Nogales et al., 2012), and iHN637 (Nagarajan et al., 2013), they were converted into a MIRIAM-compliant format. As a final step, we manually added transport reactions to reflect the uptake energetics of cofactors.

Throughout the reconstruction process, we iteratively tested and validated the reconstruction. For instance, we checked the mass and charge balance of each reaction, attempting to manually balance those that weren’t balanced. In the majority of cases metabolites were missing formula or charge definitions. In order to remove energy generating cycles, problematic reactions were manually constrained to be irreversible. Validation was carried out against growth data (Leak and Dalton, 1986a), which was also the point of reference for the parameter fitting.

Biomass Composition
For the principal components of biomass, measurements were made available through the website of our collaborators Unibio (2018). This included specifically the content of RNA (6.7%), DNA (2.3%), crude fat (9.1%), and glucose (4.5%) as a percentage of the cell dry weight. We did not use the percentage values for crude protein (72.9%) and N-free extracts (7.6%) as these measurements are fairly inaccurate relying on very generalized factors. The percentage value of Ash 550 (8.6%) measurements

2https://biocyc.org/organism-summary?object=MCAP243233
was inconsistent with the sum of its individual components (4.6%) and was hence excluded.

On the homepage of Unibio, we were also able to find g/kg measurements of all amino acids except for glutamine and asparagine, trace elements and vitamins, which could directly be converted into mmol/g DW. However, we omitted some of data: The stoichiometries for Selenium, Cadmium, Arsenic, Lead, and Mercury were not included in the biomass reaction as their values were negligibly small. Beta-Carotene (Vitamin A) and Gamma-Tocopherol (Vitamin E) were omitted because no genes were found supporting their biosynthesis, in addition to both being reported below the detection limit (Overland et al., 2010).

For the lack of better measurements, and assuming that M. buryatense and M. capsulatus (Bath) are more similar than M. capsulatus (Bath) and E. coli, the stoichiometries of glutamine and asparagine, intracellular metabolites such as ribulose-5-phosphate, organic cofactors such as coenzyme A, and cell wall components such as LPS were introduced from de la Torre et al. (2015).

Using the GC content calculated from the genome sequence (Ward et al., 2004) and the percentage measurements from Unibio for RNA and DNA, we were able to calculate the stoichiometries of all individual nucleobases.

Unibio’s measurements that 94% of the crude fat portion were fatty acids conflicted with previously published results, which indicated that in fact phospholipids are likely to be the main lipid component in M. capsulatus (Bath) (Makula, 1978; Müller et al., 2004). Thus, we assumed 94% of the crude fat to be phospholipids. This meant that 6% of the crude fat was composed of fatty acids, the distributions of which were again provided by Unibio. However, in order to calculate the stoichiometry of each fatty acid we recalculated the distribution to exclude the unidentified part. Makula had also measured the composition of the phospholipid pool itself (Makula, 1978), from which we calculated the corresponding stoichiometries for phosphatidylethanolamine, phosphatidylcholine, phosphatidylglycerol, and cardiolipin.

Bird et al. (1971) had reported the percentage of squalene and sterols of dry weight, which we converted into mmol/g DW without further corrections. Since the genes for hopanoid synthesis were identified we included diplopterol in the biomass reaction (Ward et al., 2004; Nakano et al., 2007). For a lack of more detailed measurements we estimated a similar contribution to the overall composition as squalene. We specifically used lanosterol to represent the general class of sterols in the biomass reaction, since we had only been able to reconstruct the synthesis pathway of lanosterol and since lanosterol is a key precursor metabolite in the biosynthesis of all other sterols.

The growth associated maintenance (GAM) energy requirements were calculated in accordance with protocol procedures outline in Figure 13 of Thiele and Palsson (2010). The final value of 23.087 mmol ATP g<sub>DW</sub>⁻¹ h⁻¹ is the sum of the amount of ATP required to synthesize the proportional content of proteins, DNA and RNA per one gram of cell dry weight (DW). For each component <i>i</i> we obtained the synthesis energy coefficient <i>C</i> [mmol ATP/ mmol component] from Thiele and Palsson (2010), who calculated them based on Tables 5 and 6 of Chapter 3 in Neidhardt et al. (1990). With <i>x</i> denoting the amount of a specific component <i>i</i> per gram DW [mmol component/ g<sub>DW</sub>], the following formula determines the GAM requirement:

\[ \sum_{i=1}^{n} (x_i C_i) \quad (1) \]

Specifically, the coefficients used herein are C<sub>P</sub> = 4.324, C<sub>D</sub> = 1.365, C<sub>R</sub> = 0.406 and x<sub>P</sub> = 5.312, x<sub>D</sub> = 0.047, x<sub>R</sub> = 0.135.

**Transport Reactions**

The identification of transport reactions involved the two databases for transport proteins, the TCDB (Saier et al., 2016) and the Transport Database (TransportDB) (Elbourne et al., 2017), and two computational tools, PSORTb v3.0 (Yu et al., 2010) and BLASTp (Altschul et al., 1990). We employed the following semi-automatic way of inferring the putative function of transport proteins in M. capsulatus (Bath).

Using the protein sequences in AE0172822_protein.faa (Ward et al., 2004), we determined the subcellular location of each protein using PSORTb v3.0. We filtered the results and focused only on proteins with a final score larger than 7.5, which the authors of PSORTb consider to be meaningful. We combined this list with the M. capsulatus (Bath) specific entries from the TransportDB, which allowed us to use the PSORT-scores as an additional measure of confidence. At this point, 242 putative transport proteins were identified. From this list we then selected all proteins which were predicted to transport metabolites and were already included in the model, which reduced the number to 133. Since for many of the entries, the exact mechanism of transport is unknown, we combined the previously selected transporters with the results from running BLAST against the TCDB. The e-value and bitscore from BLAST provided yet another measure to confidently assess the correctness of the automatic TransportDB predictions, and the Transporter Classification-Numbers (TC-numbers) allowed us to gather information on the mechanism of transport. This led to a list of 97 transport proteins with high confidence, which was filtered once more as follows.

We checked the general description for a given specific TC-number, and then considered the BLAST result to read about a given transporter in more detail, especially with regards to finding the specific substrates. When we were able to identify the corresponding transport reaction in the BiGG database, we incorporated only the simplest, smallest set of reactions. In cases of conflicts between our own BLAST results and the automatic TransportDB transporter annotation, we preferentially trusted the BLAST results. Thus we ultimately added 75 transporter-encoding genes connected via GPR to 56 distinct transport reactions.

**Applied Constraints**

To identify which mode of electron transfer is active in M. capsulatus (Bath), we fit the solutions of flux balance analysis using iMcBath to measured values (Leak and Dalton, 1986a). The authors had experimentally determined the O<sub>2</sub>/CH<sub>4</sub> ratio and the growth yield of M. capsulatus (Bath) in several conditions. They varied the nitrogen source using KNO<sub>3</sub>, NH<sub>4</sub>Cl, and both
simultaneously; the concentration of copper in the medium, which directly affects the activity of either sMMO or pMMO; and whether the culture was oxygen or carbon limited.

The baseline constraints applied to the model allowed only the uptake of CH$_4$, O$_2$, SO$_4$, PO$_4$, either NH$_4$ or NO$_3$, and an array of mineral salts which are required by the biomass equation. The methane uptake rate was limited to allow at most 18.46 mmol gDW$^{-1}$ h$^{-1}$ (adopted from de la Torre et al., 2015), all remaining uptake rates were left unconstrained. We accounted for the differential expression of ADH (model reaction ID: ALAD_L) in the presence of excess NH$_4$ in the medium by blocking the GS/GOGAT (model reaction ID: GLNS). In the presence of NO$_3$ or N$_2$ we allowed flux through both reactions (Murrell and Dalton, 1983). Maximization of flux through the biomass equation was set as the objective for parsimonious flux balance analysis with which all calculations were done.

Since each electron transfer mode, respectively, is represented by the flux through one specific reaction in the model (model reaction IDs: redox-arm = PMMOipp, direct coupling = PMMODCipp, and uphill electron transfer = CYOR_q8ppi), we were able to investigate each simply by allowing flux through the corresponding reaction.

Several studies have shown that NH$_4$ is a co-metabolic substrate to the methane monoxygenases in *M. capsulatus* (Bath) leading to the production of hydroxylamine first and nitrite later (Hutton and Zobell, 1953; Bédard and Knowles, 1989; Nyerges and Stein, 2009). Hence, when simulating the growth on NH$_4$ we assumed that varying ratios $r$ of the nitrogen taken up would eventually be converted into nitrite (Figure 7).

\[ v_{EX_{NH4}} + rv_{EX_{nle}} = 0 \]  

The program code and data for all analyses presented in this publication are available as Model 15 Calculations.ipynb at https://github.com/ChristianLieven/memote-m-capsulatus.

### Stoichiometric Modeling

The reactome of an organism can be represented mathematically as a stoichiometric matrix $S$. Each row of $S$ corresponds to a unique compound, while each column corresponds to a metabolic structure. Hence the structure of the matrix for an organism with $m$ compounds and $n$ reactions equals $m \times n$. The values in each row denote the stoichiometric coefficients of that metabolite in all reactions. The coefficients are either negative for compounds that are educts, positive for those that are products, or zero for those that are not involved in a given metabolic reaction.

The vector $v$ of length $n$ contains as values the turnover rates of each reaction. These rates are commonly referred to as fluxes and are given the unit mmol gDW$^{-1}$ h$^{-1}$. Vector $v$ is also referred to as the flux vector.

### Availability of Data and Materials

The metabolic model, scripts and corresponding datasets generated and/or analyzed during the current study are available in the GitHub repository.
REFERENCES


Unibio (2018). Unibio Homepage. Available at: www.unibio.dk


Conflict of Interest Statement: Unibio is a collaborator in the “Environmentally Friendly Protein Production (EFPro2)” project. CL, MH, and NS are publicly funded through the Novo Nordisk Foundation and the Innovation Fund Denmark. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lieven, Petersen, Jørgensen, Gernaey, Herrgard and Sonnenschein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.