Research Article

Genomically and biochemically accurate metabolic reconstruction of Methanosarcina barkeri Fusaro, iMG746

Matthew C. Gonnerman[1], Matthew N. Benedict[1], Adam M. Feist[2], William W. Metcalf[3,4], Nathan D. Price[1,4,5§]

[1] Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[2]Department of Bioengineering, University of California, 9500 Gilman Drive #0412, San Diego, La Jolla, CA 92093-0412, United States

[3]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[4]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[5]Institute for Systems Biology, Seattle, WA 98109

[§]Corresponding author. 401 Terry Avenue North, Seattle, WA 98109. nprice@systemsbiology.org

Keywords: Systems biology, Metabolic flux analysis, Metabolic network, Modeling, Microbiology

List of abbreviations

5HBC: 5-hydroxybenzimidazolycolbamide; BLASTP: Basic Local Alignment Search Tool for Proteins; Ech: Energy-conserving hydrogenase; F_{420}: Cofactor F_{420}; FBA: Flux Balance Analysis; Fpo: F_{420} dehydrogenase; Frh: F_{420} reducing hydrogenase; GPR: Gene-Protein-Reaction association; Hdr: Heterodisulfide reductase; iAF692: The previously-published *M. barkeri* model; iMG746: The updated *M. barkeri* model; MH4SPT: N5-methyl-tetrahydrosarcinapterin; Mtr: N5-methyl-tetrahydrosarcinapterin:coenzyme M methyltransferase; Vht: methanophenazine-dependent hydrogenase

Submitted: 26-Nov-2012
Revised: 15-Jan-2013
Accepted: 13-Feb-2013

Abstract

Methanosarcina barkeri is an Archaeon that produces methane anaerobically as the primary byproduct of its metabolism. *M. barkeri* can utilize several substrates for ATP and biomass production including methanol, acetate, methyl amines and a combination of hydrogen and carbon dioxide. In 2006, a metabolic reconstruction of *M. barkeri*, iAF692, was generated based on a draft genome annotation. The iAF692 reconstruction enabled the first large-scale simulations for Archaea. Since the publication of the first metabolic reconstruction of *M. barkeri*, additional genomic, biochemical, and phenotypic data have clarified several metabolic pathways. We have used this newly available data to improve the *M. barkeri* metabolic reconstruction. Modeling simulations using the updated model, iMG746, have led to increased accuracy in predicting gene knockout phenotypes and dynamic simulations of batch growth behavior. We have used the model to examine knockout lethality data and make predictions about metabolic regulation under different growth conditions. Thus, the updated metabolic reconstruction of *M. barkeri* metabolism is a useful tool for predicting cellular behavior, studying the methanogenic lifestyle, guiding experimental studies and making predictions relevant to metabolic engineering applications.

1. Introduction

Methanosarcina barkeri Fusaro is a methanogen that was isolated using a low-salt medium [1] from Lago del Fusaro sediment, where it grows as multicellular aggregates. Members of the genus *Methanosarcina* can utilize at least nine different substrates to produce methane, making them the most metabolically versatile of the methanogenic Archaea [1].

Studies of methanogens provide a means for examining industrial biofuel production and impacts on the environment. Methanogens are currently used in industry through biomethane production and low carbon processing and fixation [2, 3]. Methanogens significantly impact the global carbon cycle by contributing an estimated $5x10^{14}$ g of methane each year to the atmosphere [4]. They also offer an opportunity to examine cross-species relationships [3, 5] with other methanogens or sulfate-reducing organisms. The Methanosarcina are particularly interesting due to 1) their broad substrate specificity relative to other methanogens, and 2) robust techniques having been developed for their genetic manipulation, which have previously been applied both for fundamental studies as well as for strain design (see [6] for review).

Metabolic reconstructions provide structured repositories for links between reactions, metabolites and genes for an organism [7]. These links form a network that can be represented in a mathematical format, and then used to predict growth rates, guide metabolic engineering efforts, examine byproduct yields or perform various other simulations [8]. Metabolic reconstructions can be used as a basis of mathematical models for examining or predicting cellular behavior [9]. Such models have now been built for many bacteria, archaea, and eukaryotes and used for medical and bioengineering applications [9].

A manually curated reconstruction of *M. barkeri*, iAF692, was published in 2006 before the genome was fully annotated [10]. iAF692 was used to perform the first genome-scale metabolic simulations of an Archaeon. iAF692 was used to predict the most likely stoichiometry of the nitrogenase reaction, to examine the number of protons pumped through the Ech hydrogenase reaction, and to analyze the essentiality of genes and reactions in the methanogenic pathways of *M. barkeri*. Since its publication, there have been substantial new findings identifying unique biosynthetic pathways for Archaea and methanogens specifically, motivating us to update the metabolic reconstruction for *M. barkeri* using an established protocol [11] for improving existing reconstructions using newly available information. Specifically, the new additions to the model included: (1) updates to the electron transport chain in the methanogenesis pathway, (2) updates to cofactor and amino acid biosynthesis pathways, (3) a gene-protein-reaction association (GPR) based on a completed annotation, and (4) standard Gibbs free energy information for metabolites and reactions involved in the reconstruction. The revised reconstruction provides improved prediction of biomass generation and byproduct secretion rates, improved accuracy of knockout lethality predictions, and improved growth yield predictions.

2. Materials and Methods

Model Reconstruction procedure

The updated *M. barkeri* str. Fusaro reconstruction is based on iAF692, the genome-scale model previously developed for the same organism [10]. Reactions were added and removed as new information was incorporated into the metabolic reconstruction. Model modifications were done using the COBRA toolbox in MATLAB [12].

Since the publication of iAF692, the genome annotation for *Methanosarcina barkeri* str. Fusaro has been completed [1] and this complete annotation was used in the updated GPRs in iMG746. The draft genome IDs in the iAF692 GPRs were converted to the completed genome IDs using a bidirectional BLASTP [13]. The conversion table is available in the Additional Files.

Sources of new information included the Kyoto Encyclopedia of Genes and Genomes (KEGG) [14], MetaCyc [15], the Model SEED reconstruction [16], UniProt [17] and various literature sources (see supplemental material). The *M. barkeri* gene annotation [1] was used as the primary source of information for the GPR, but the annotation was compared with the genome annotations of *Methanosarcina acetivorans*, *Methanosarcina mazei* and *Methanocaldococcus jannaschii* to ensure consistency. When possible, bidirectional BLASTP against proteins with experimentally validated functions was used to validate gene annotations.

When available, reaction directionality was determined using literature evidence. Other evidence such as basic thermodynamic estimates was used when literature evidence was unavailable. In order to estimate the free energy change for each of the reactions, charged *molfiles* were generated for each compound in the model using ACD/Labs software. The Gibbs free energy of formation for 539 (84%) of metabolites and Gibbs energy of reaction for 615 (83%) of the reactions were computed using the Group Contribution Method (GCM) [18-20]. The remaining metabolites and reactions involve complex ring structures or polymers that are not supported by the group contribution method. The GCM software package provides an estimate of Gibbs free energy of formation and its estimated error at a standard 1 molar concentration, 25°C and a pH of 7. After computing the Gibbs energy of reaction under these conditions, the Gibbs energies were converted to millimolar standard concentrations ΔG_r, and then converted to millimolar standard concentrations using:

$$\Delta G^{o}_{rmM} = \Delta G^{o}_{r} - RTSln(1000). \quad (1)$$

Here, ΔG^{0}_{rmM} is a vector of standard reaction Gibbs free energies with millimolar standards, ΔG^{0}_{r} is a vector of reaction Gibbs energies with molar standards, R is the universal gas constant, T is the temperature (25°C) and S is a stoichiometric matrix with metabolites on rows, reactions on columns and the values being the stoichiometric coefficient of each row in each reaction. The Gibbs free energy of reaction was further refined using previously-published methods to account for proton and charge gradients [21]. Only protons were accounted for since intracellular and extracellular concentrations for other ions are unknown. These final values were used to examine reversibilities and assign them when literature evidence was unavailable.

Flux balance analysis

Model simulations were performed using flux balance analysis (FBA). FBA is an established technique used to predict phenotypes for metabolic reconstructions [8]. Briefly, the sets of possible reaction rates are constrained by physical laws such as conservation of mass and thermodynamic limitations. Flux balance analysis identifies a set of reaction rates that maximizes (or minimizes) an assumed objective subject to these constraints (see supplemental text for a more detailed discussion). The simulations in this paper were performed by assuming maximal biomass production as the cell's objective function. The biomass reaction was updated from the previously-published *M. barkeri* model using data available in the literature and public databases (see results and supplemental tables).

Limiting substrate uptake rates were set based on data available in the literature [22-32]. Other compounds found in the growth media were allowed to enter the system at any rate that was feasible given other constraints on the network. Like the previously reconstructed model of *M. barkeri*, growth and non-growth-associated ATP maintenance parameters were incorporated in the updated model to account for the energy requirements for growth and cell maintenance (respectively). The non-growth associated ATP maintenance cost was increased to 2 mmol/GDW/hr from the iAF692 model, which used an ATP maintenance cost of 1.75. This increase is the result of removing the inefficiencies that had been incorporated into the methanogenesis reactions in iAF692. The relatively low ATP maintenance costs compared to *E. coli* are consistent with experimental evidence for *M. mazei* that maintenance costs are low for methanogens [33] and could be an effect associated with growth on low energy carbon sources. The results of the simulations gave reaction fluxes in units of mmol/GDW/hr and growth rates in units of 1/hr.

The GLPK solver was used for all optimizations using the COBRA toolbox in MATLAB [12]. Simulation results were verified with the LPSolve solver, with identical results. All growth simulations were conducted in high salt minimal media (given in Additional files) with the assumption that *M. barkeri* grows as non-aggregates (reduced generation of cell wall components). Changes in reaction rates discussed in the paper (Figure 1) were verified to hold for all alternative equivalent optima using Flux Variability Analysis (FVA, [34]).

Sensitivity analysis on biomass objective function

The sensitivity analysis was performed in minimal media, an anaerobic environment and methanol limiting conditions. The methanol uptake rate was set between 0 and 30 mmol*(GDW*hr)$^{-1}$ and all other compounds in the minimal media were unconstrained. Protein and RNA percent contributions were examined at ±50% of their baseline contribution in the biomass reaction. DNA and lipid percent contributions were adjusted between 1% and 10% of the weight for one gram dry weight.

Gene-protein-reaction relationships and gene knockout lethality predictions.

The genes were linked together using a binary AND and OR relationship. When a protein functions as part of a complex with known partners, the genes in the complex were assigned an

AND relationship. If the protein complex is unknown, does not exist or the genes are paralogs, the genes were linked with an OR relationship to be conservative with our gene knockout lethality predictions. To simulate a gene deletion, the reaction associated with the gene or genes in question was removed from the model based on the evaluation of a Boolean rule where knocked out genes were FALSE and all others were TRUE. FBA was then performed on the modified model to simulate growth by optimizing the rate of the biomass reaction. The results of simulation were evaluated such that a growth rate above 10^{-4} hr^{-1} was considered nonlethal and anything lower was considered a lethal gene deletion. The simulations were conducted on minimal media, at the estimated substrate uptake rates found in the literature.

3. Results

The updated reconstruction, iMG746, was primarily built using three resources: the previous reconstruction (iAF692), information gathered from online databases, and biological studies, listed in the Additional files (see Methods). An overview of the two models can be seen in Table 1. A detailed spreadsheet containing all reaction formulas, metabolites, thermodynamics and the available gene-protein relationship for reactions is available in the Additional files. Computable versions of the model (in SBML and .mat formats suitable for COBRA toolbox use) are also available in the Additional Files.

Updates to methanogenesis

Methanogenesis is the sole energy-conserving process in *M. barkeri*. In the four distinct methanogenic pathways, the electron transport chain conserves energy by pumping protons and sodium ions across the membrane to generate concentration and charge gradients. The proton gradient created by electron transport is then used to generate ATP for the cell, while the sodium gradient can be used to increase the proton gradient (through a proton/sodium antiporter) or to drive active transport of other molecules. We have carefully curated the methanogenesis pathways in the model and updated the previously published reconstruction as follows: (*i*) Based on the reaction mechanism of the HdrDE heterodisulfide reductase (HDR), two protons are translocated across the membrane (1.8 in iAF692) per methane produced [35, 36]; (*ii*) Reduced F_{420} dehydrogenase (Fpo) is predicted to pump two protons (1.8 in iAF692) for each reduced F_{420} consumed [35]; and (*iii*) Based on the proposed hydrogen cycling mechanism for energy conservation, hydrogen produced from the cytoplasmic F_{420} reducing hydrogenase (Frh) crosses the membrane and is captured by periplasmic methanophenazine-dependent hydrogenase (Vht) creating a net two protons (1.8 in iAF692) transferred across the membrane [35] (Figure 2). Fractional proton pumping had been used in iAF692 to account for inefficiencies in the transport reaction or energy loss through diffusion or leaks. In the updated model, the reactions were modified to pump integral numbers of ions and the inefficiencies were added to the Non-Growth Associated Maintenance (NGAM) instead.

There are two known types of heterodisulfide reductases in methanogens [37] and genes for both are found in *M. barkeri*. The HdrDE type has been shown to be essential in *M. acetivorans* and is involved in the membrane-bound electron transport chain. HdrABC has been shown to be active in *M. acetivorans* [37], where it is thought to couple oxidation of reduced ferredoxin to

reduction of heterodisulfide during methylotrophic methanogenesis. Preliminary studies have suggested that the HdrABC in the *Methanosarcina* links the consumption of reduced ferredoxin to the production of: (i) Coenzyme M, Coenzyme B and reduced F_{420} via a bifurcation method, or (ii) Coenzyme M and Coenzyme B only. A recently published reconstruction of *Methanosarcina acetivorans* was used to predict that HdrABC could be used with either mechanism under certain conditions [38]. In the updated model, we considered the possibility that the cytoplasmic HdrABC could perform a similar function in *M. barkeri.* We have examined both mechanisms by adding the proposed reactions to the updated model and looking for unique phenotypes. Adding either reaction to the iMG746 model did not change knockout predictions for which experimental data is available. However, the model predicts that *M. barkeri mch* mutant could grow on pyruvate by utilizing the HdrABC bifurcation reaction. This phenotype presents an opportunity to test these two possible stoichiometries.

Along with HdrABC, additional metabolic reactions were incorporated into the methanogenesis pathways of iMG746 based on new experimental evidence. 5-hydroxybenzimidazolycolbamide (5HBC), a cofactor required for methyl transfer reactions in *M. barkeri* [39], is not included in iAF692. During the methyltransferase reaction cycle, the cobalt atom in 5HBC has a chance to become oxidized to the inactive Co(II) state. Reactivation of the enzyme requires another protein complex (RamA/M) to reduce the inactive Co(II) to the active Co(I) state [40]. We have added both 5HBC synthesis and Ram to the updated model.

Several reactions affect the growth yields of *M. barkeri* on methanol and hydrogen. As such, additional uncharacterized reaction stoichiometries were examined to determine their effects on growth yields and rates. The studied reactions were Mtr, Ech, and the Na^+/H^+ antiporter in *M. barkeri*. We also examined the effects of different levels of specificity for sodium as opposed to protons by the ATP synthase. Simulation results for all of these studies were compared with experimentally availble growth rates, growth yields, and gene deletion studies (see supplemental text and additional data files). From these simulations, we concluded that the most likely combination of stoichiometries is that Ech pumps 2 protons, Mtr pumps 2 sodiums, the $Na^+\backslash H^+$ antiporter pumps 2 sodiums per proton, and there is no sodium-driven ATP synthesis under normal physiological conditions. These final determinations were incorporated into the model for subsequent analysis.

Updates to cofactor biosynthesis and general metabolism

Since the publication of iAF692, there have been studies characterizing additional metabolic pathways specific for *M. barkeri*, such as the biosynthesis pathway for heme complexes. The genes associated with experimentally studied heme synthesis pathways in other organisms are absent from *M. barkeri*, although it is known to use cytochromes in its electron transport chain [36]. Recent studies offer a possible heme synthesis pathway for *M. barkeri* that involves S-adenosyl-L-methionine as a methyl donor [41, 42], and this putative pathway was included in the iMG746 model. Additional newly-added pathways with recent experimental support in the *Methanosarcina* include a dechlorination mechanism for chlorinated ethane molecules [43, 44], a pathway for synthesis of ribose-5-phosphate that bypasses the pentose phosphate pathway

[45], and a newly-proposed pathway for coenzyme M biosynthesis that has diverged from the pathway found in *Methanocaldococcus* [46].

When appropriate, data for other organisms were used to further refine the metabolic model. The gene annotations were evaluated based on bidirectional BLASTP. If no genes have been identified in the archaeal pathway and the bacterial pathway appeared to be absent, the archaeal pathways were chosen for inclusion in the model. Archaeal-specific pathways were added for coenzyme A, riboflavin, and methanofuran(b) biosynthesis based on biochemical studies in other organisms [47-50].

One reaction of particular interest involves the reduction of SO_3 to H_2S. SO_3 is required for coenzyme M biosynthesis in the *Methanosarcina* [46]. The iAF692 model incorporated SO_3 by adding an exchange from the media. However, it has been shown that *M. barkeri* can grow using H_2S as the only sulfur source and therefore, there must be a mechanism by which *M. barkeri* produces SO_3 from H_2S [51]. *M. barkeri* contains a close homolog to the F_{420}-dependent sulfite reductase from *Methanocaldococcus jannaschii* [52]. Thermodynamics heavily favors the production of H_2S in this reaction under standard conditions, but a thermodynamic sensitivity analysis shows that under certain feasible concentration conditions (i.e. very high H_2S concentrations relative to SO_3), the sulfite reductase reaction could be used as a source of SO_3 for the coenzyme M synthesis. Since only a small amount of coenzyme M production is required for growth, it is possible that this mechanism is used for SO_3 generation, but experimental validation is required to determine if this reaction is indeed the source of the required SO_3.

Construction and analysis of an updated biomass reaction

When generating the biomass reaction for iAF692, specific composition information for *M. barkeri* was utilized when available. Due to a lack of experimental data concerning cell makeup, information on relatives or similar organism was used when *M. barkeri*-specific information was unavailable. Based on information from the literature and MetaCyc [53-56], methanofuran(b) biosynthesis was incorporated into iMG746 and methanofuran(b) was included in the reaction. Additional information provided the components for the cell wall for *M. barkeri* [57]. *M. barkeri* is typically grown in high salt media so it was modeled in the single cell condition without the cell wall component by reducing the cell wall contribution to the biomass objective function.

A sensitivity analysis was previously on the iAF1260 model of *E. coli*, which examined the impact of the exact coefficients of the biomass reaction on the predicted growth rate [58]. We performed a similar sensitivity analysis (see Methods section) on the *M. barkeri* biomass reaction. Changes to DNA, RNA, lipid and protein percent contributions to the biomass objective function had little effect on the predicted growth rate and product secretion rates (see Additional files). Only the growth associated maintenance, non-growth-associated maintenance, and proton and sodium pump stoichiometries had a significant impact on simulation accuracy for the slow-growing organism.

Validation and application studies

To demonstrate improved predictive accuracy with the updated *M. barkeri* reconstruction, the previous iAF692 model and the updated iMG746 model were examined using two simulation studies. The first study compared growth rate and growth yield simulations on three substrates: acetate, methanol, and hydrogen with carbon dioxide. The second study consisted of the predicted effects of gene knockouts on cell viability.

To examine the growth rate and growth yield predictions, we compiled experimental results for growth on the substrates methanol, acetate and hydrogen with carbon dioxide [23, 24, 26-31, 35]. Literature evidence suggests that *M. barkeri* grows faster on methanol than hydrogen or acetate and that *M. barkeri* has the highest growth yield on methanol. We then compared experimental growth rates and yields with simulated results from iAF692 and iMG746 on methanol (see Figure 1), acetate and hydrogen with carbon dioxide. The changes in the methanogenesis pathway incorporated in iMG746 greatly influence the growth yield and the growth rate predictions.

The growth yields and growth rates on acetate simulated using iMG746 and iAF692 agree with each other but seem to overestimate the growth yields. Considering the large variations of growth yields for *M. barkeri* when grown on acetate, both models predict growth yields which are comparable to the experimentally determined values but additional investigation may be needed. One possible reason for the lower growth yield is that *M. barkeri* did not maximally utilize acetate and instead used sub-optimal pathways during the growth experiments.

For growth on methanol, the predicted growth yields and doubling times from iMG746 are comparable to those found experimentally. For growth on H_2/CO_2 for *M. barkeri,* both iMG746 and iAF692 predict doubling times similar to that found experimentally. Since the growth rate for H_2/CO_2 was only available from one study (and no growth yield was explicitly reported), we did not attempt to go further and compare the computed growth yields on that substrate to the simulated growth yields.

To test the lethality of gene deletions, gene knockout phenotypes were predicted using both *M. barkeri* reconstructions. Out of the 58 different knockouts found through literature searches [27, 29, 35, 59, 60], iMG746 was able to correctly predict 56 (96%), an improvement over the 88% accuracy of iAF692 (Table 2).

The first discrepancy between iMG746 knockout lethality predictions and experimental data concerns the lethality of an Ech hydrogenase knockout grown on methanol with hydrogen and carbon dioxide. This inconsistency was also mentioned in the iAF692 reconstruction paper. Ech hydrogenase is required to generate reduced ferroredoxin so that carbon dioxide can be reduced to generate ATP and synthesize biomass components. The *Δech* mutant is capable of growth on methanol, but not growth on hydrogen, carbon dioxide and methanol [60]. Without regulation or additional constraints, a flux balance analysis (FBA) model will never predict lethality as long as a subset of provided substrates is sufficient for growth. It has been proposed that the mutant did not grow due to inhibition or repression of the oxidative branch of methanogenesis [60]. High quantities of hydrogen could cause an increased concentration of

reduced F_{420} via Frh, which would in turn block the flow of electrons from methanol oxidation (which also would require oxidized F_{420}). In addition, high concentrations of CO_2 could cause low concentrations of reduced ferroredoxin via reaction formylmethanofuran(b) dehydrogenase. To test the effects of these changes, the F_{420} reducing hydrogenase reaction (F4RH) was made irreversible in the direction of reduced F_{420} production and formylmethanofuran(b) dehydrogenase (FMFD(b)) was made irreversible in the direction of consumption of CO_2. With these two constraints, the Ech knockout became lethal. Therefore, it was concluded that the presence of hydrogen and CO_2 in a methanol medium could render the *Δech* mutant inviable through thermodynamic constraints on reactions and reduction imbalances within the mutant.

The second discrepancy occurs with a double knockout of Fpo and Frh during growth on acetate [35]. In this knockout, two ways to consume reduced F_{420} are removed from the cell, leaving glutamate synthase as the main reaction consuming reduced F_{420} [61], as seen in Figure 3. Based on simulation results, when grown on acetate, the majority of the reduced F_{420} produced is consumed to synthesize glutamate. Glutamate is predicted to be the primary nitrogen donor in other amino acid synthesis reactions and has the highest free amino acid concentration in the *M. barkeri* cell [61]. It is possible that this reaction has a high turnover rate of glutamate. Simulations show that the growth rate drops linearly with reduction of flux through glutamate synthase, where a knockout of the reaction results in zero growth (results not shown). This implies that kinetic limitations could prevent growth on acetate for the double mutant.

Environment-specific regulation could be another factor resulting in the lethality of the Fpo\Frh double mutant during growth on acetate (Figure 3). *M. barkeri* has annotated enzymes for both NADPH and F_{420}-dependent glutamine synthases. In some bacteria, the NADPH glutamate dehydrogenase is up-regulated if the NH_4 concentration is high, but in lower NH_4 concentration media, F_{420}-dependent glutamate synthase is the preferred enzyme [62]. *M. barkeri* is typically grown in high NH_4 concentration media, which could mean that if this regulation is present, only the NADPH glutamate dehydrogenase is active. To simulate this possible regulation, the F_{420} glutamate synthase reaction was removed leaving only the NADPH reaction available for glutamine biosynthesis in the model. With these changes, the Fpo\Frh double mutant could no longer grow on acetate. Additional experimentation on this mutant could explain these findings.

4. Discussion

Metabolic reconstructions provide a means to probe organisms on a genome scale and guide experimentation. We have provided an update to the metabolic network of *M. barkeri* and examined the differences between our new reconstruction and the previous reconstruction. Flux balance analysis allowed us to examine the effects of the electron transport chain at the network level. Dynamics and growth yield predictions agreed with a lower sodium pump for the Mtr reaction and a larger number of protons pumped for the Ech hydrogenase. The improvements to the electron transport chain directly influenced the accuracy of the growth yield, growth rate and gene deletion predictions.

While not all changes to the metabolic network affected the growth rate predictions, they offered improved insight to the biology of the organism and could aid efforts to refine the knowledge of

methanogen networks to further improve the predictive power of the model. Updating metabolic models is important for maximizing their use in light of currently-available data [63]. Expanding our knowledge of Archaea metabolism through continued efforts in generating and curating genome-scale models will aid future reconstructions for other Archaea, especially other methanogens.

Acknowledgements

MG and MB participated in the reconstruction of the iMG746 model. MG performed validation studies on the model. MG and AF provided detailed information on the reactions in the supplemental. WM provided biological analysis on the simulated predictions of the model. WM and NP conceived the project and NP coordinated research efforts among all the authors. All authors participated in the writing and have read and approved the final manuscript. The authors gratefully acknowledge a critical reading of this manuscript by Julie Bletz. This work was funded by the Energy Biosciences Institute (grant number EBIOO0J08).

Conflict of Interest Statement

The authors declare no commercial or financial conflict of interest.

5. References

[1] Maeder, D. L., Anderson, I., Brettin, T. S., Bruce, D. C*., et al.*, The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol* 2006, *188*, 7922-7931.
[2] Sasaki, K., Sasaki, D., Morita, M., Hirano, S*., et al.*, Efficient treatment of garbage slurry in methanogenic bioreactor packed by fibrous sponge with high porosity. *Appl Microbiol Biotechnol* 2010, *86*, 1573-1583.
[3] Seth-Smith, H., A more convenient truth. *Nat Rev Microbiol* 2007, *5*, 248-250.
[4] Rodgers, J. E., Whitman, W. B., *Microbial production and consumption of greenhouse gases: methane, nitrogen oxides, and halomethanes.*, American Society for Microbiology, Washington, D.C. 1991.
[5] Koschorreck, M., Geller, W., Neu, T., Kleinsteuber, S*., et al.*, Structure and function of the microbial community in an in situ reactor to treat an acidic mine pit lake. *FEMS Microbiol Ecol* 2010, *73*, 385-395.
[6] Kohler, P. R., Metcalf, W. W., Genetic manipulation of Methanosarcina spp. *Front Microbiol* 2012, *3*, 259.
[7] Reed, J. L., Famili, I., Thiele, I., Palsson, B. O., Towards multidimensional genome annotation. *Nat Rev Genet* 2006, *7*, 130-141.
[8] Oberhardt, M. A., Palsson, B. Ø., Papin, J. A., Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 2009, *5*, 320.
[9] Feist, A. M., Herrgard, M. J., Thiele, I., Reed, J. L., Palsson, B. O., Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* 2009, *7*, 129-143.
[10] Feist, A. M., Scholten, J. C. M., Palsson, B. Ø., Brockman, F. J., Ideker, T., Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri. *Mol Syst Biol* 2006, *2*, 2006.0004.
[11] Thiele, I., Palsson, B. Ø., A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* 2010, *5*, 93-121.
[12] Becker, S. A., Feist, A. M., Mo, M. L., Hannum, G*., et al.*, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* 2007, *2*, 727-738.
[13] Camacho, C., Coulouris, G., Avagyan, V., Ma, N*., et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* 2009, *10*, 421.
[14] Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., Hirakawa, M., KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 2010, *38*, D355-360.
[15] Caspi, R., Altman, T., Dale, J. M., Dreher, K*., et al.*, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 2009, *38*, D473-479.
[16] Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M*., et al.*, High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* 2010, *28*, 977-982.
[17] The Uniprot, C., The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* 2010, *38*, D142-148.
[18] Jankowski, M. D., Henry, C. S., Broadbelt, L. J., Hatzimanikatis, V., Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys J* 2008, *95*, 1487-1499.
[19] Finley, S. D., Broadbelt, L. J., Hatzimanikatis, V., Thermodynamic analysis of biodegradation pathways. *Biotechnol Bioeng* 2009, *103*, 532-541.
[20] Henry, C. S., Jankowski, M. D., Broadbelt, L. J., Hatzimanikatis, V., Genome-scale thermodynamic analysis of Escherichia coli metabolism. *Biophys J* 2006, *90*, 1453-1461.
[21] Henry, C. S., Broadbelt, L. J., Hatzimanikatis, V., Thermodynamics-based metabolic flux analysis. *Biophys J* 2007, *92*, 1792-1805.

[22] Liu, J. S., Marison, I. W., von Stockar, U., Microbial growth by a net heat up-take: a calorimetric and thermodynamic study on acetotrophic methanogenesis by Methanosarcina barkeri. *Biotechnol Bioeng* 2001, *75*, 170-180.
[23] Weimer, P. J., Zeikus, J. G., One carbon metabolism in methanogenic bacteria. Cellular characterization and growth of Methanosarcina barkeri. *Arch Microbiol* 1978, *119*, 49-57.
[24] Sowers, K. R., Nelson, M. J., Ferry, J. G., Growth of acetotrophic, methane-producing bacteria in a pH auxostat. *Curr Microbiol* 1984, *11*, 227-229.
[25] Silveira, R. G., Nishio, N., Nagai, S., Growth Characteristics and Corrinoid Production of Methanosarcina barkeri on Methanol-Acetate Medium. *Journal of Fermentation and Bioengineering* 1990.
[26] Hutten, T. J., Bongaerts, H. C., van der Drift, C., Vogels, G. D., Acetate, methanol and carbon dioxide as substrates for growth of Methanosarcina barkeri. *Antonie van Leeuwenhoek* 1980.
[27] Bock, A. K., Schonheit, P., Growth of Methanosarcina barkeri (Fusaro) under nonmethanogenic conditions by the fermentation of pyruvate to acetate: ATP synthesis via the mechanism of substrate level phosphorylation. *J Bacteriol* 1995, *177*, 2002-2007.
[28] Boek, A.-K., Prieger-Kraft, A., Schonheit, P., Pyruvate - a novel substrate for growth and methane formation in Methanosarcina barkeri. *Arch Microbiol* 1994.
[29] Welander, P. V., Metcalf, W. W., Loss of the mtr operon in Methanosarcina blocks growth on methanol, but not methanogenesis, and reveals an unknown methanogenic pathway. *Proc Natl Acad Sci U S A* 2005, *102*, 10664-10669.
[30] Smith, M. R., Mah, R. A., Growth and Methanogenesis by Methanosarcina Strain 227 on Acetate and Methanol. *Appl Environ Microbiol* 1978.
[31] Bomar, M., Knoll, K., Widdel, F., Fixation of molecular nitrogen by Methanosarcina barkeri. *FEMS Microbiol Ecol* 1985.
[32] Ranalli, G., Whitmore, T. N., Lloyd, D., Methanogenesis from methanol in Methanosarcina barkeri studied using membrane inlet mass spectrometry. *FEMS Microbiol Ecol* 1986.
[33] Rajoka, M. I., Tabassum, R., Malik, K. A., Enhanced rate of methanol and acetate uptake for production of methane in batch cultures using Methanosarcina mazei. *Bioresource Technology* 1998.
[34] Mahadevan, R., Schilling, C. H., The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* 2003, *5*, 264-276.
[35] Kulkarni, G., Kridelbaugh, D. M., Guss, A. M., Metcalf, W. W., Hydrogen is a preferred intermediate in the energy-conserving electron transport chain of Methanosarcina barkeri. *Proc Natl Acad Sci U S A* 2009, *106*, 15915-15920.
[36] Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., Hedderich, R., Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol* 2008, *6*, 579-591.
[37] Buan, N. R., Metcalf, W. W., Methanogenesis by Methanosarcina acetivorans involves two structurally and functionally distinct classes of heterodisulfide reductase. *Mol Microbiol* 2010, *75*, 843-853.
[38] Benedict, M. N., Gonnerman, M. C., Metcalf, W. W., Price, N. D., Genome-scale metabolic reconstruction and hypothesis testing in the methanogenic archaeon Methanosarcina acetivorans C2A. *J Bacteriol* 2011, *194*, 855-865.
[39] van der Meijden, P., te Brommelstroet, B. W., Poirot, C. M., van der Drift, C., Vogels, G. D., Purification and properties of methanol:5-hydroxybenzimidazolylcobamide methyltransferase from Methanosarcina barkeri. *J Bacteriol* 1984, *160*, 629-635.
[40] Ferguson, T., Soares, J. A., Lienard, T., Gottschalk, G., Krzycki, J. A., RamA, a protein required for reductive activation of corrinoid-dependent methylamine methyltransferase reactions in methanogenic archaea. *J Biol Chem* 2009, *284*, 2285-2295.

[41] Buchenau, B., Kahnt, J., Heinemann, I. U., Jahn, D., Thauer, R. K., Heme biosynthesis in Methanosarcina barkeri via a pathway involving two methylation reactions. *J Bacteriol* 2006, *188*, 8666-8668.
[42] Storbeck, S., Rolfes, S., Raux-Deery, E., Warren, M. J.*, et al.*, A novel pathway for the biosynthesis of heme in Archaea: genome-based bioinformatic predictions and experimental evidence. *Archaea* 2010, *2010*, 175050.
[43] Holliger, C., Schraa, G., Stams, A. J., Zehnder, A. J., Reductive dechlorination of 1,2-dichloroethane and chloroethane by cell suspensions of methanogenic bacteria. *Biodegradation* 1990, *1*, 253-261.
[44] Holliger, C., Schraa, G., Stupperich, E., Stams, A. J., Zehnder, A. J., Evidence for the involvement of corrinoids and factor F430 in the reductive dechlorination of 1,2-dichloroethane by Methanosarcina barkeri. *J Bacteriol* 1992, *174*, 4427-4434.
[45] Finn, M. W., Tabita, F. R., Synthesis of catalytically active form III ribulose 1,5-bisphosphate carboxylase/oxygenase in archaea. *J Bacteriol* 2003, *185*, 3049-3059.
[46] Graham, D. E., Taylor, S. M., Wolf, R. Z., Namboori, S. C., Convergent evolution of coenzyme M biosynthesis in the Methanosarcinales: cysteate synthase evolved from an ancestral threonine synthase. *Biochem J* 2009, *424*, 467-478.
[47] Yokooji, Y., Tomita, H., Atomi, H., Imanaka, T., Pantoate kinase and phosphopantothenate synthetase, two novel enzymes necessary for CoA biosynthesis in the Archaea. *J Biol Chem* 2009, *284*, 28137-28145.
[48] Grochowski, L. L., Xu, H., Leung, K., White, R. H., Characterization of an Fe(2+)-dependent archaeal-specific GTP cyclohydrolase, MptA, from Methanocaldococcus jannaschii. *Biochemistry* 2007, *46*, 6658-6667.
[49] Grochowski, L. L., Xu, H., White, R. H., An iron(II) dependent formamide hydrolase catalyzes the second step in the archaeal biosynthetic pathway to riboflavin and 7,8-didemethyl-8-hydroxy-5-deazariboflavin. *Biochemistry* 2009, *48*, 4181-4188.
[50] Chatwell, L., Krojer, T., Fidler, A., Romisch, W.*, et al.*, Biosynthesis of riboflavin: structure and properties of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductase of Methanocaldococcus jannaschii. *J Mol Biol* 2006, *359*, 1334-1351.
[51] Mazumder, T. K., Nishio, N., Fukuzaki, S., Nagai, S., Effect of Sulfur-Containing Compounds on Growth of Methanosarcina barkeri in Defined Medium. *Appl Environ Microbiol* 1986, *52*, 617-622.
[52] Johnson, E. F., Mukhopadhyay, B., A new type of sulfite reductase, a novel coenzyme F420-dependent enzyme, from the methanarchaeon Methanocaldococcus jannaschii. *J Biol Chem* 2005, *280*, 38776-38786.
[53] White, R. H., Biosynthesis of the 2-(aminomethyl)-4-(hydroxymethyl)furan subunit of methanofuran. *Biochemistry* 1988, *27*, 4415-4420.
[54] Kezmarsky, N. D., Xu, H., Graham, D. E., White, R. H., Identification and characterization of a L-tyrosine decarboxylase in Methanocaldococcus jannaschii. *Biochim Biophys Acta* 2005, *1722*, 175-182.
[55] DiMarco, A. A., Bobik, T. A., Wolfe, R. S., Unusual coenzymes of methanogenesis. *Annu Rev Biochem* 1990, *59*, 355-394.
[56] Eisenreich, W., Bacher, A., Biosynthesis of methanofuran in Methanobacterium thermoautotrophicum. *J Biol Chem* 1992, *267*, 17574-17580.
[57] Kandler, O., Hippe, H., Lack of peptidoglycan in the cell walls of Methanosarcina barkeri. *Arch Microbiol* 1977, *113*, 57-60.
[58] Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M.*, et al.*, A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 2007, *3*, 121.
[59] Welander, P. V., Metcalf, W. W., Mutagenesis of the C1 oxidation pathway in Methanosarcina barkeri: new insights into the Mtr/Mer bypass pathway. *J Bacteriol* 2008, *190*, 1928-1936.

[60] Meuer, J., Kuettner, H. C., Zhang, J. K., Hedderich, R., Metcalf, W. W., Genetic analysis of the archaeon Methanosarcina barkeri Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proc Natl Acad Sci U S A* 2002, *99*, 5632-5637.
[61] Kenealy, W. R., Thompson, T. E., Schubert, K. R., Zeikus, J. G., Ammonia assimilation and synthesis of alanine, aspartate, and glutamate in Methanosarcina barkeri and Methanobacterium thermoautotrophicum. *J Bacteriol* 1982, *150*, 1357-1365.
[62] Raemakers-Franken, P. C., Brand, R. J., Kortstee, A. J., Van der Drift, C., Vogels, G. D., Ammonia assimilation and glutamate incorporation in coenzyme F420 derivatives of Methanosarcina barkeri. *Antonie Van Leeuwenhoek* 1991, *59*, 243-248.
[63] Feist, A. M., Palsson, B. O., The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol* 2008, *26*, 659-667.

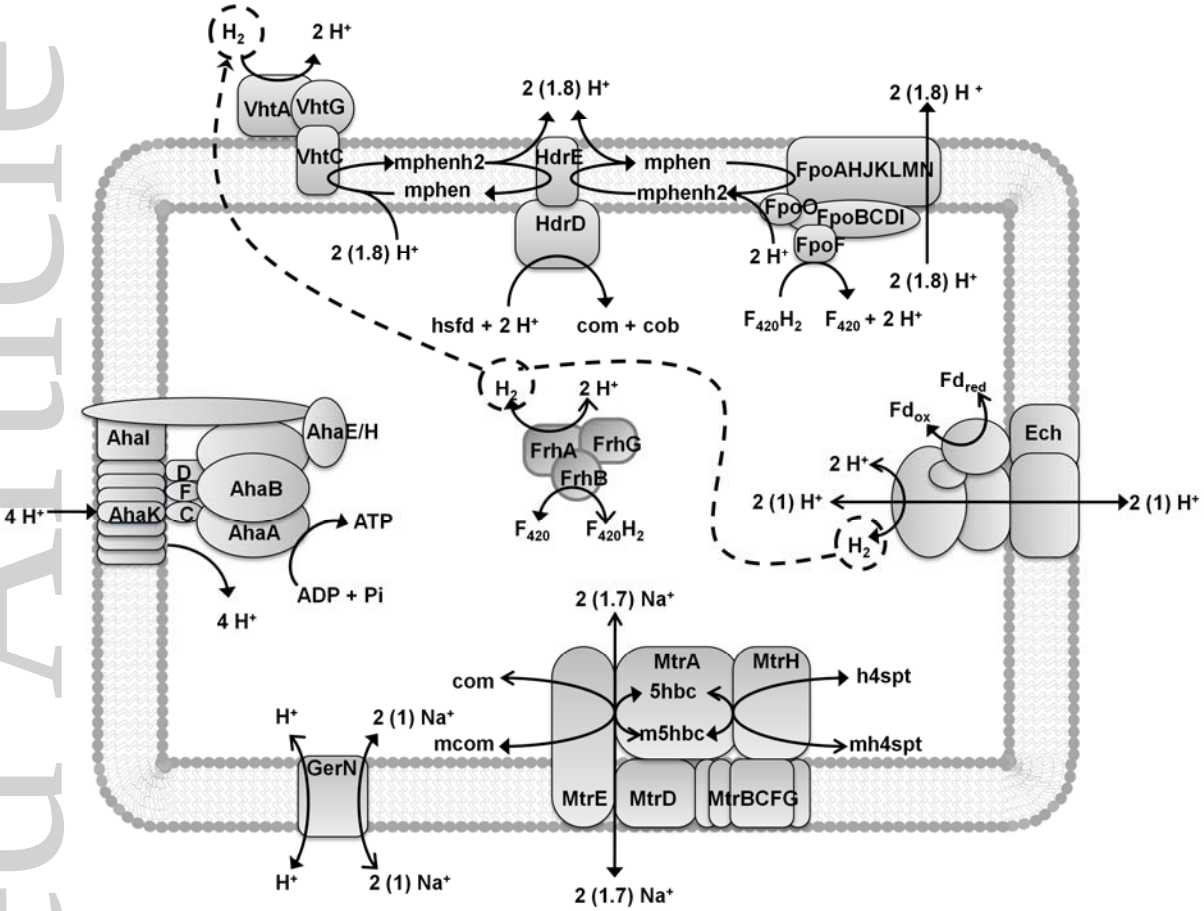

Figure 1: A comparison of simulated doubling times and growth yields from iMG746 and iAF692 compared with available experimental values from the literature. The error bars give a 95% confidence interval (±2*standard error) calculated from the experimental data found in literature. The number of available data points from the literature varied depending on condition as follows: methanol doubling time: 4, methanol yield: 3, acetate doubling time: 4, acetate yield: 3, H2/CO2 doubling time: 4, H2/CO2 yield: 1 (no stderr estimate available).

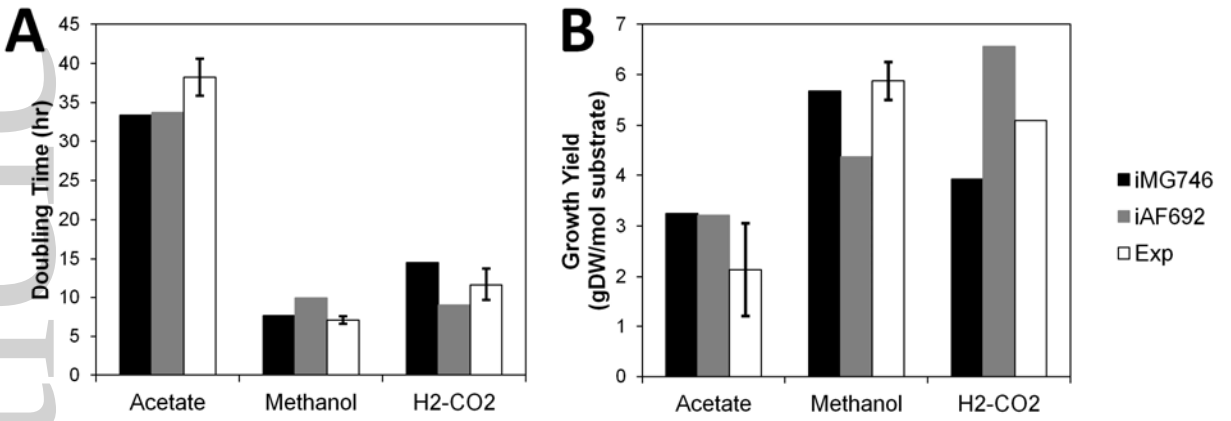

Figure 2: A diagram depicting the current knowledge of the electron transport chain in *M. barkeri* that was reconstructed in the iMG746 model. The numbers in parenthesis are stoichiometric coefficients used in the iAF692 model. Irreversible reactions have only one arrowhead and reversible reactions have two arrowheads.

Abbreviations: 5hbc, 5-Hydroxybenzimidazolylcob(I)amide; ADP, Adenosine diphosphate; ATP, Adenosine triphosphate; cob, coenzyme-b; com, coenzyme-m; F_{420}, coenzyme-ferredoxin-420-2-oxidized; $F_{420}H_2$, coenzyme-ferredoxin-420-2-reduced; Fd_{red}, ferredoxin-reduced-24Fe-4S; Fd_{ox}, ferredoxin-oxidized-24Fe-4S; H^+, proton; H_2, hydrogen; h4spt, tetrahydrosarcinopterin; hsfd, heterodisulfide; Na^+, sodium; mcom, methyl-coenzyme M; mh4spt, Co-Methyl-Co-5-hydroxybenzimidazolylcobamide; mphen, methanophenazine-oxidized; mphenh2, methanophenazine-reduced; Pi, Phosphate; Aha, ATP synthase; Ech, Ech hydrogenase; Fpo, Reduced F_{420} dehydrogenase; Frh, F_{420} reducing hydrogenase; GerN, sodium/hydrogen antiporter; Hdr, heterodisulfide reductase; Mtr, N5-methyl-tetrahydrosarcinapterin : coenzyme M methyltransferase; Vht, methanophenazine-dependent hydrogenase;

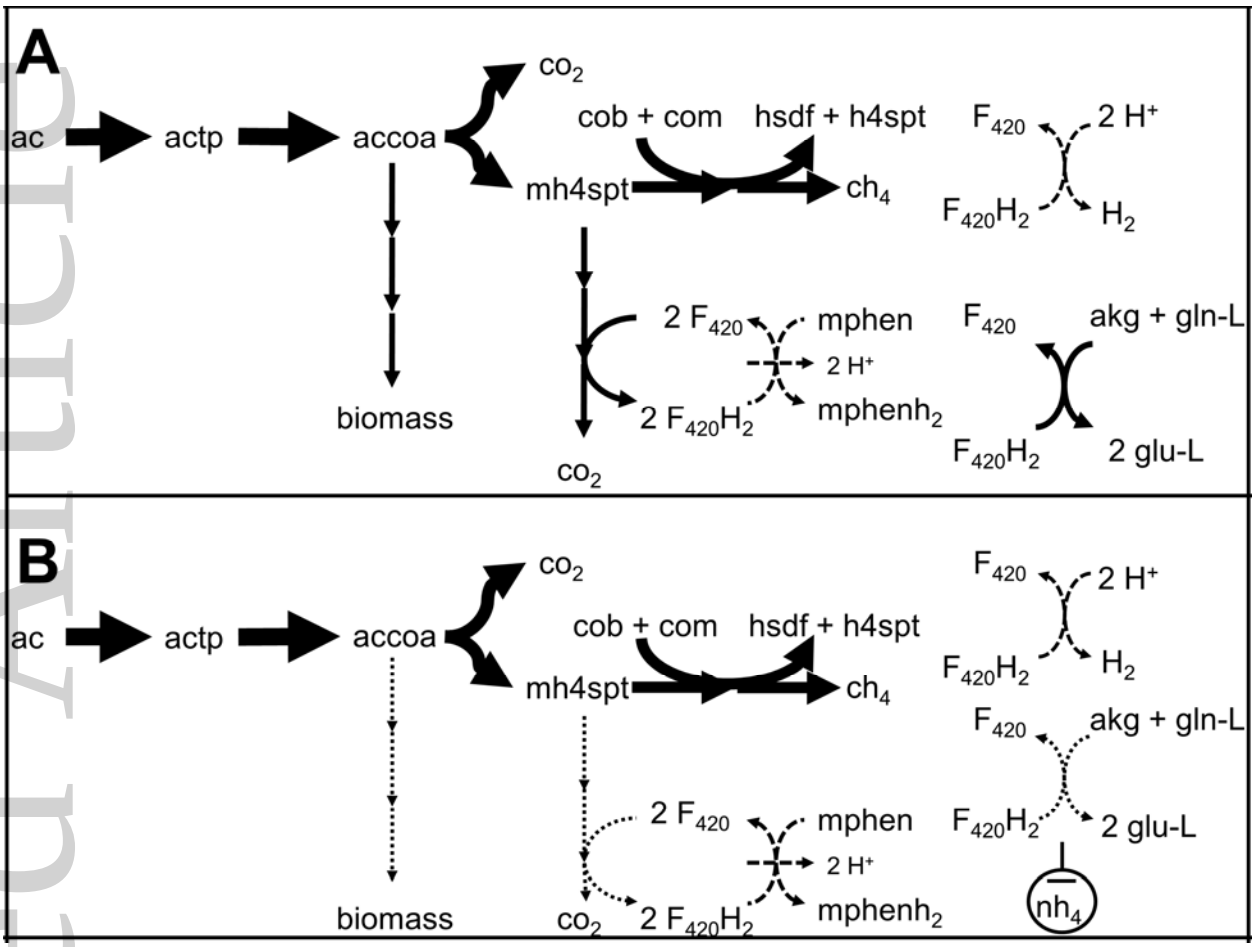

Figure 3: Model prediction pathway and hypothesized regulation an Fpo\Frh knockout mutant in *M. barkeri*. Pathway A represents the model predicted flux results. Arrow thickness indicates relative flux amount. Dashed arrows indicate the reactions used in the wild type model, which are knocked out in the mutant. Pathway B represents the pathway with putative regulation, where dotted arrows indicate reactions predicted to be affected by ammonia regulation on glutamate synthesis.

Tables

Table 1: A comparison between iAF692 and iMG746. The iAF692 reconstruction is based on the draft annotation and the genes associated with their reactions are from the draft annotation.

M. barkeri model comparison		
Model	iAF692	iMG746
Total Number of Reactions	689	815
Metabolic Reactions	619	741
Unique to one model	49	172
Exchange Reactions	70	74
Unique to one model	1	5
Genes	692	746
Gene-protein-reaction associations	509	615
Spontaneous Reactions	0	3
Diffusion Reactions	15	15

Table 2 : A gene deletion study using the updated *M. barkeri* genome-scale reconstruction (iMG746) and comparison to simulation results using the previous reconstruction (iAF692). Gray boxes indicate no experimental data was available. True positive (TP) indicates an agreement with experimentally determined nonlethal knockouts and true negative (TN) indicates an agreement with experimentally determined lethal knockouts. False positive (FP) indicates a disagreement with experiments such that the reconstruction predicts nonlethal knockout, but experiments show cell death. False negative (FN) indicates a disagreement with experiments such that the reconstruction predicts lethal knockout, but experiments show cell growth. In cases where the two models give different results, the first result is from iAF692 and the second from iMG746. P and N: Positive and negative predictions from iMG746 (no data available).

Gene Deletion Study							
Gene(s)	**Reaction(s)**	**MeOH**	**Ac**	**H_2 / CO_2**	**MeOH / Ac**	**MeOH / H_2 / CO_2**	**Pyr**
Mtr	MTSPCMMT	TN	TN	**FP/TN**	TP	TP	P
Ech	ECHH_20	TP	TN	TN	P	**FP**	N
Mch	MTSPC	TN	TN	TN	N	TP	P
Mcr	MCR	P	N	N	P	P	TP
Fpo	F4D	**FN/TP**	TP	TP	P	TP	P
FpoF	F4D	**FN/TP**	TP	TP	P	TP	P
Frh	F4RH	TP	TP	N	P	TP	P
Fpo / Frh	F4D / F4RH	TN	**FP**	TN	P	TP	P
FpoF / Frh	F4D / F4RH	TN	TP	TN	P	TP	P
Mer	F4MTSPR	TN	TN	TN	**FN/TP**	TP	P
Mtd	F4MTSPD	TN	TN	TN	TN	TP	P
Ftr	FMFTSPFT(b)	TN	TN	TN	TN	TP	P
FaeA / Mtr	FAE / MTSPCMMT	TN	TN	**FP/TN**	TP	TP	P
FaeA / Mer	FAE / F4MTSPR	TN	TN	TN	TN	TP	P