Supporting Materials for “Thermodynamics and H₂ Transfer in a Methanogenic, Syntrophic Community”

Joshua J. Hamilton¹, Montserrat Calixto Contreras¹, and Jennifer L. Reed¹,§

¹Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA
§Corresponding author

Email addresses:
JJH: jjhamilton2@wisc.edu
MCC: calixto@wisc.edu
JLR: reed@engr.wisc.edu

Table of Contents

1 Supporting Results .................................................................................................................. 1
  1.1 Testing and Parameterizing the iMhu428 Metabolic Model .................................................. 1
  1.2 Testing and Parameterizing the iSfu648 Metabolic Model .................................................. 2
      1.2.1 Elimination of Thermodynamically Infeasible Closed Cycles ..................................... 2
      1.2.2 ATP Production Mechanisms in S. fumaroxidans ..................................................... 2
      1.2.3 Validation of ATP Production Mechanisms ............................................................... 3
      1.2.4 Thermodynamics of H₂ Production .............................................................................. 4
2 Supporting Discussion ........................................................................................................... 6
  2.1 Testing and Parameterizing of the iSfu648 Metabolic Model ............................................. 6
3 Supporting Methods ............................................................................................................... 7
  3.1 Reconstruction and Parameterization of the iMhu428 Metabolic Model .............................. 7
  3.2 Reconstruction and Parameterization of the iSfu648 Metabolic Model .............................. 8
      3.2.1 Curation of Draft iSfu648 Model ................................................................................ 8
      3.2.2 Manual Curation of Energy Conservation Mechanisms ............................................ 9
      3.2.3 Manual Curation of Carbon Transfer Mechanisms .................................................. 11
      3.2.4 Experimental Justification for Theoretical Flux Distributions .................................. 11
      3.2.5 Regulatory Constraints ............................................................................................... 12
      3.2.6 Development and Gapfilling of Biomass Equation .................................................... 13
      3.2.7 Gapfilling of Blocked Reactions ................................................................................ 14
      3.2.8 Conversion to a Thermodynamic Model .................................................................... 14
      3.2.9 iSfu648 Model Parameterization ................................................................................ 14
  3.3 Metabolite Structures and Molecular Formulas .................................................................. 15
  3.4 Thermodynamics-Based Metabolic Flux Analysis (TMFA) ............................................... 15
  3.5 TMFA and Thermodynamics of H₂ Production ................................................................... 17
  3.6 Thermodynamic Lumping ................................................................................................. 18
  3.7 Linear Programming Approximation of TMFA (TMFA-LP) ................................................. 19
  3.8 Parsimonious TMFA (pTMFA) .......................................................................................... 19
  3.9 Community Formulation .................................................................................................... 20
  3.10 Community Thermodynamic Variability Analysis ............................................................ 22
  3.11 Minimal Probabilistic Sets (MPS) .................................................................................. 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.11.1</td>
<td>MPS for Validating $i$Mhu428: ATP Gain</td>
<td>24</td>
</tr>
<tr>
<td>3.11.2</td>
<td>MPS for Validating $i$Mhu428: Biomass Growth</td>
<td>25</td>
</tr>
<tr>
<td>3.11.3</td>
<td>MPS for Validating $i$Sfu648: Closed-Network ATP Cycles</td>
<td>25</td>
</tr>
<tr>
<td>3.11.4</td>
<td>MPS for Validating $i$Sfu648: ATP Gain</td>
<td>26</td>
</tr>
<tr>
<td>3.12</td>
<td>Simulation Conditions</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>Supporting Figures</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Fig S1</td>
<td>27</td>
</tr>
<tr>
<td>4.2</td>
<td>Fig S2</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>Supporting References</td>
<td>29</td>
</tr>
</tbody>
</table>
1 Supporting Results

1.1 Testing and Parameterizing the iMhu428 Metabolic Model

Methanogens can be classified into two groups based on the presence or absence of cytochromes in their energy transfer mechanisms [7]. To date, GEMs have been constructed for two methanogens, *Methanosarcina barkeri* [63,64] and *Methanosarcina acetivorans* [41,65], both of which use cytochromes. To the best of our knowledge, there are no published GEMs for methanogens lacking cytochromes, such as *M. hungatei*.

After the iMhu428 reconstruction was converted to a thermodynamic model, TMFA predicted a no-growth phenotype, due to the inability of the iMhu428 model to oxidize H₂S to SO₃²⁻ via sulfite reductase, a necessary reaction for biomass production. The estimated Δ_rG' for this reaction was between 55.6 kJ/mol and 284 kJ/mol, indicating that the reaction could not proceed in the direction required for growth. However, replacing the coenzyme F₄₂₀ with a generic ferredoxin enabled sulfite reductase to proceed in the required direction. Because sulfur metabolism in methanogens remains poorly understood [66], this difference in sulfur metabolism between *M. acetivorans* and *M. hungatei* seems plausible.

The iMhu428 model predicted that *M. hungatei* is able to generate 0.5 mol ATP per mol of CO₂ converted to CH₄. However, the iMhu428 model predicted other, higher-yielding, ATP-generating mechanisms outside the methanogenesis pathway, despite experimental evidence suggesting methanogenesis is the sole source of ATP. The enumeration of all such ATP-generating cycles [67] proved computationally intractable, so a previously proposed probability-based approach [68] was used to qualitatively constrain reaction directions in the iMhu428 model. This approach calculates the probability that a reaction’s standard transformed Gibbs free energy of reaction (Δ_rG°') is negative. If the probability is greater (less) than 70% (30%), the reaction is constrained to the forward (reverse) direction. Constraining all 497 such reactions eliminated these higher ATP yielding mechanisms, but also resulted in a no-growth phenotype.

To overcome this problem, a new optimization approach was developed (see Minimal Probabilistic Sets, or MPS). MPS minimizes the number of qualitative reaction direction constraints needed to reduce the maximum ATP gain to 0.5 mol ATP/mol CO₂, while still allowing for growth. A set of 41 reactions had to be constrained to a single direction (listed in Table S2 in S1 Dataset of the Supporting Material). Many of the reactions identified by MPS may be thermodynamically unidirectional *in vivo* due to intracellular metabolite concentrations related to the cellular energy charge [69], but these constraints are not captured in the iMhu428 model. These additional reaction direction constraints were used in all subsequent analyses with the iMhu428 model.

It has also been observed that biomass and energy generation are independent processes in *M. hungatei*, with CO₂ being the sole source of ATP (using methanogenesis), and acetate being the sole source of carbon for biomass [55]. In contrast, the iMhu428 model predicts that *M. hungatei* can produce some biomass using CO₂ alone, and MPS could not identify any qualitative reaction direction...
constraints which would prevent CO\textsubscript{2} from being used to generate biomass while still allowing biomass production from acetate (see Supporting Methods for additional details). Nonetheless, during growth on CO\textsubscript{2} and acetate, over 90\% of the CO\textsubscript{2} consumed is used for methanogenesis.

1.2 Testing and Parameterizing the \textit{iSfu648 Metabolic Model}

1.2.1 Elimination of Thermodynamically Infeasible Closed Cycles

A thermodynamically consistent metabolic network should contain no thermodynamically infeasible closed cycles [29]. Such cycles may become problematic if they enable undesirable behaviors, such as free ATP production. The thermodynamic model for \textit{iSfu648} allowed ATP producing cycles when no nutrients were provided. As with the \textit{iMhu428} reconstruction, MPS was used to minimize the number of qualitative reaction direction constraints needed to eliminate these cycles, while maintaining biomass growth. The resulting qualitative reaction direction constraints are given in Table S2 in S1 Dataset of the Supporting Material.

1.2.2 ATP Production Mechanisms in \textit{S. fumaroxidans}

Experimental studies have elucidated five growth modes for \textit{S. fumaroxidans}: four in monoculture and one in coculture with \textit{M. hungatei} [36,48,52]. In monoculture, \textit{S. fumaroxidans} can convert fumarate to succinate and CO\textsubscript{2}. The bacterium can also grow on fumarate with formate or H\textsubscript{2} as an electron donor, producing succinate as the only by-product. \textit{S. fumaroxidans} can also oxidize propionate to succinate using fumarate as an electron acceptor. A final monoculture growth mode has been observed in which \textit{S. fumaroxidans} couples propionate oxidation to sulfate reduction [70]. However, the sulfate reduction pathway in \textit{S. fumaroxidans} has not been fully characterized [48]. In coculture with \textit{M. hungatei}, \textit{S. fumaroxidans} grows on propionate and produces acetate, H\textsubscript{2} and CO\textsubscript{2}/formate; however, \textit{S. fumaroxidans} can not grow on propionate alone in monoculture. This work examines the most commonly studied growth modes: monoculture growth on fumarate (Fig 2A), monoculture growth on fumarate plus propionate (Fig 2B), and coculture growth on propionate (Fig 2C).

During monoculture growth on fumarate alone (Fig 2A), one mole of fumarate gets fully oxidized to CO\textsubscript{2}, while six moles of fumarate get reduced to succinate [48,52]:

\[
7 \text{ fumarate} \rightarrow 6 \text{ succinate} + 4 \text{ CO}_2
\]  
(S1)

The oxidation of one fumarate to CO\textsubscript{2} generates one ATP and five reducing equivalents (three NADH and two pairs of reduced ferredoxin) [48,52], while the reduction of additional fumarate to succinate by fumarate reductase (\textit{FRD}) produces menaquinone [48]. The Rnf complex (\textit{RNF}) reoxidizes NADH via ferredoxin reduction, while the ferredoxin-oxidizing hydrogenase (\textit{frH\textsubscript{2}ase}) transfers electrons from ferredoxin to protons, generating H\textsubscript{2} and oxidized ferredoxin. The cytosolic hydrogenase (\textit{cytH\textsubscript{2}ase}) then transfers electrons from H\textsubscript{2} to menaquinone, generating menaquinol. An additional pair of electrons (in the form of H\textsubscript{2}) is provided to \textit{cytH\textsubscript{2}ase} via formate hydrogen lyase (\textit{FHL}), the final step in fumarate oxidation to CO\textsubscript{2}. Reduction of fumarate to succinate recycles menaquinol back into menaquinone and generates the proton motive force (PMF) responsible for driving the \textit{RNF} reaction and producing ATP.
During monoculture growth on fumarate plus propionate (Fig 2B), one mole of propionate gets oxidized to succinate, while one mole of fumarate gets oxidized to acetate and CO$_2$. Two additional moles of fumarate get reduced to succinate [36,48]:

$$\text{propionate + 3 fumarate } \rightarrow \text{acetate + CO}_2 + 3 \text{ succinate} \quad \text{(S2)}$$

The oxidation of fumarate to acetate and CO$_2$ produces one NADH and one pair of reduced ferredoxin. The confurcating hydrogenase ($cH_2$ase) couples NADH and ferredoxin oxidation with H$_2$ production. The cytosolic hydrogenase ($cytH_2$ase) then transfers electrons from H$_2$ to menaquinone, generating menaquinol and driving fumarate reductase ($FRD$). Oxidation of propionate to succinate produces one ATP, while $FRD$ generates the PMF necessary for additional ATP production.

During coculture growth on propionate (Fig 2C), propionate gets oxidized to acetate and CO$_2$ via the methylmalonyl-CoA pathway [48,52]:

$$\text{propionate } \rightarrow \text{acetate + CO}_2 + 3 \text{ H}_2 \quad \text{(S3)}$$

A single ATP is generated during the oxidation of propionate to succinate. This ATP is hydrolyzed to generate the PMF necessary for the endergonic oxidation of succinate to fumarate ($SDH$). This oxidation generates menaquinol; the cytosolic hydrogenase ($cytH_2$ase) then transfers electrons from menaquinol to two protons, generating H$_2$. The further oxidation of fumarate to acetate and CO$_2$ produces one NADH and one pair of reduced ferredoxin. The confurcating hydrogenase ($cH_2$ase) couples NADH and ferredoxin re-oxidation with H$_2$ production. Unlike in the monoculture growth modes, the H$_2$ is not consumed by the cytosolic hydrogenase and must diffuse outside the cell. It has been proposed that the net production of H$_2$ by S. fumaroxidans is only thermodynamically favorable at the low H$_2$ concentrations maintained by methanogens, thereby explaining why S. fumaroxidans only produces H$_2$ during coculture growth.

S. fumaroxidans exhibits considerable flexibility in its ATP production mechanisms during syntrophic growth ([8,9,48,51]), and can produce formate instead of CO$_2$ (Fig 3) yielding an overall transformation of:

$$\text{propionate } \rightarrow \text{acetate + formate + 2 H}_2 \quad \text{(S4)}$$

In one mechanism (Fig 3A), activity of the cytosolic formate hydrogenase ($cytFDH$) substitutes for the activity of the $cytH_2$ase. CO$_2$ produced by the oxidation of propionate serves as a substrate for $cytFDH$, which transfers electrons from menaquinol to CO$_2$, generating formate. In a second mechanism (Fig 3B), the confurcating formate dehydrogenase ($cFDH$) substitutes for the $cH_2$ase. Here, the $cFDH$ couples NADH and ferredoxin re-oxidation with the conversion of CO$_2$ (from propionate oxidation) to formate.

### 1.2.3 Validation of ATP Production Mechanisms

Experimental evidence suggests that the carbon and electron transport mechanisms described in Table 2 and Table S3 (in S1 Dataset in the Supporting Material) provide the sole source for ATP production in S. fumaroxidans, either by substrate-level phosphorylation or through establishment of a proton gradient used by ATP synthase [5]. However, the $i$Sfu648 model includes additional pathways for ATP synthesis with higher ATP yields. To identify these pathways, the $i$Sfu648 model was first used to identify the highest amount of ATP that could be produced using the experimentally known mechanisms. The $i$Sfu648 model was then analyzed to find and eliminate other higher ATP yielding pathways.
As with the iMhu428 reconstruction, mechanisms with higher ATP yields were identified, and MPS was used to eliminate them. For each growth condition, MPS identified the smallest number of qualitative reaction direction constraints needed to maintain biomass growth while ensuring the appropriate maximal ATP gain. The resulting qualitative reaction direction constraints are given in Table S2 in S1 Dataset of the Supporting Material. Many of the reactions identified by MPS may be thermodynamically unidirectional \textit{in vivo} due to intracellular metabolite concentrations related to the cellular energy charge [69], but these constraints are not captured in the iSfu648 model. These additional reaction direction constraints were used in all subsequent analyses with the iSfu648 model.

1.2.4 Thermodynamics of H\textsubscript{2} Production

Experimental studies of \textit{S. fumaroxidans} have shown that H\textsubscript{2} is not produced during growth in monoculture [36,52,53]. However, our simulations reveal that H\textsubscript{2} production remains thermodynamically feasible under monoculture conditions. For example, during monoculture growth on fumarate, the iSfu648 model predicts that H\textsubscript{2} can be produced via the following mechanism:

\[
\text{fumarate} \rightarrow 4 \text{CO}_2 + 6 \text{H}_2 \quad (S5)
\]

In this scenario, H\textsubscript{2} molecules produced by the ferredoxin-oxidizing hydrogenase are exported outside the cell, instead of serving as substrates for the cytosolic hydrogenase. As a result, no proton motive force is generated by fumarate reductase, and the net ATP yield is zero. This suggests that thermodynamic considerations alone cannot explain the absence of H\textsubscript{2} production during monoculture growth, but that the observed flux distribution may instead be driven by demands for energy generation.

To determine if thermodynamics could play a role in explaining the absence of H\textsubscript{2} production in monoculture conditions, we sought to identify a constraint on extracellular metabolite concentrations that would eliminate H\textsubscript{2} production [34] (e.g., if extracellular CO\textsubscript{2} and H\textsubscript{2} concentrations are too high, H\textsubscript{2} production would become thermodynamically unfavorable). Such a constraint was identified, but the iSfu648 model predicted a shift from CO\textsubscript{2} and H\textsubscript{2} to formate production:

\[
\text{fumarate} \rightarrow 3 \text{CO}_2 + 5 \text{H}_2 + \text{formate} \quad (S6)
\]

Once again, a constraint on extracellular CO\textsubscript{2}, H\textsubscript{2}, and formate concentration was identified to render this extracellular flux distribution thermodynamically unfeasible, and the iSfu648 model predicted fumarate reduction to succinate:

\[
\text{fumarate} \rightarrow 2 \text{CO}_2 + 0.5 \text{succinate} + 2.5 \text{H}_2 \quad (S7)
\]

However, no constraint could be found which would fully eliminate H\textsubscript{2} production.

Similar results were obtained during monoculture growth during simultaneous utilization of fumarate and propionate. The iSfu648 model predicted that H\textsubscript{2} could be produced via the following mechanism:

\[
\text{propionate} + \text{fumarate} \rightarrow \text{CO}_2 + \text{acetate} + \text{succinate} + 2 \text{H}_2 \quad (S8)
\]

In this scenario, propionate is oxidized to succinate, while fumarate is oxidized to CO\textsubscript{2} and acetate, producing NADH and reduced ferredoxin. The confucrating hydrogenase reoxidizes ferredoxin and NAD\textsubscript{H}, producing H\textsubscript{2}. The H\textsubscript{2} is exported outside the cell, instead of participating in fumarate reduction as a substrate for the cytosolic hydrogenase. The net ATP yield is one mole ATP per mole propionate, less than the maximum ATP yield of 2 1/3 ATP per mole propionate, which occurs when fumarate is
reduced. Again, this suggests that the observed flux distribution may be driven by demands for energy generation instead of thermodynamics.

Furthermore, no constraint on metabolite concentrations could be identified which would make this H$_2$-production mechanism thermodynamically infeasible. Thus, if the absence of H$_2$ production under monoculture conditions is governed by thermodynamics, current thermodynamic models cannot capture this phenomenon. As a result, simulations performed on monoculture included an explicit constraint blocking net H$_2$ production.
2 Supporting Discussion

2.1 Testing and Parameterizing of the iSfu648 Metabolic Model

The iSfu648 model also has some important limitations. In particular, the iSfu648 reconstruction does not distinguish between cytosolic, periplasmic, and membrane-bound versions of electron transport complexes (such as H₂ase). For example, the fumarate reductase and succinate dehydrogenase complexes are membrane-bound, and menaquinol should transfer electrons to a membrane-bound or periplasmic enzyme complex. However, the iSfu648 reconstruction does not differentiate between differently-localized versions of the same complex. As a consequence, the iSfu648 model’s predicted electron transport mechanisms may be simpler than occur in vivo.

In addition, by using the standard transformed Gibbs free energy of reaction (\(\Delta r G^0\)) as a basis for thermodynamic calculations, the iSfu648 model contained too much network flexibility, as it did not account for thermodynamic interconnectivity arising from shared metabolites. Thermodynamic interconnectivity can be captured using the standard transformed Gibbs free energy of molecules or groups directly when modeling thermodynamics, as well as through lumping constraints. Enumerating all such constraints which could be applied to a network is computationally infeasible, but it may be possible to identify important linear combinations of reactions which would benefit from a lumping constraint.
3 Supporting Methods

3.1 Reconstruction and Parameterization of the iMhu428 Metabolic Model

The iMhu428 reconstruction of *M. hungatei* was built from the iMB745 reconstruction of *M. acetivorans* [41], the newest methanogen reconstruction available at the time this work began. A preliminary draft reconstruction was built using the RAVEN Toolbox [42], but, the reconstruction contained less than 200 genes, or 27% of the genes in the iMB745 reconstruction (S2 Dataset in the Supporting Material). Instead of performing extensive gapfilling, all 756 reactions from the iMB745 *M. acetivorans* reconstruction were copied into the iMhu428 reconstruction, with modifications to reflect key metabolic features of *M. hungatei*.

Because *M. hungatei* can only utilize acetate and CO₂ [55], reactions which enabled growth on CO and methylated carbon sources were removed. In addition, the oxidative arm of the TCA cycle was removed, and the reductive arm was added [71]. The methanogenesis pathway was also replaced [7].

The iMhu428 reconstruction was then converted to a thermodynamic model, as described in “Metabolite Structures and Molecular Formulas” and “Thermodynamics-Based Metabolic Flux Analysis.” On defined minimal media (Table S1 in S1 Dataset in the Supporting Material), TMFA predicted a no-growth phenotype, due to the inability of the iMhu428 model to oxidize H₂S to SO₄²⁻ via sulfite reductase, a necessary reaction for biomass production. The estimated transformed Gibbs free energy of reaction (ΔrG') for this reaction was between 55.6 kJ/mol and 284 kJ/mol, indicating that the reaction could not proceed in the forward direction required for growth. However, replacing the coenzyme F₄₂₀ with a generic ferredoxin enabled sulfite reductase to proceed in the required direction, though the group contribution method could not provide an estimate of the standard transformed Gibbs free energy of reaction (ΔrG°'). Because sulfur metabolism in methanogens remains poorly understood [66], this difference in cofactor utilization between *M. acetivorans* and *M. hungatei* seems possible.

Substrate uptake rates (SURs), and growth- (GAM) and non-growth-associated (NGAM) maintenance requirements for the iMhu428 model of *M. hungatei* were estimated using experimental data. NGAM represents the amount of energy spent to maintain the cell (i.e., maintenance energy), while GAM represents energy spent on growth-related functions (e.g., protein synthesis). In genome-scale models, these ATP requirements are usually expressed using ATP hydrolysis reactions: in the case of NGAM, the ATP hydrolysis reaction is constrained to some lower bound, and in the case of GAM, an ATP hydrolysis term is added to the biomass equation.

The value of the NGAM parameter was found by maximizing ATP hydrolysis at the reported maintenance cost as measured in terms of CO₂ uptake [53]. The value of the GAM parameter is typically obtained by plotting a series of substrate uptake and growth rates taken from chemostat experiments [72], but no such data were available for *M. hungatei*. However, studies have reported the growth rate of *M. hungatei* on CO₂ [51], and its yield per mol methane (Yₓ/P) [72]. From these data, the following iterative procedure was used to simultaneously estimate SURCO₂ and GAM:
1. Select an initial estimate for the GAM.
2. Identify the \( \text{SUR}_{\text{CO}_2} \) for which the observed growth rate and model-predicted maximum growth rates agree.
3. Calculate the \textit{in silico} \( \text{CH}_4 \) production rate at the model-predicted maximum growth rate and compute \( Y_{X/P} \).
4. Adjust the value of the GAM upwards or downwards until \( Y_{X/P} \) at the model-predicted maximum growth agrees with the experimental measurement. Changing the GAM will change the model-predicted maximum growth. Thus,
5. Repeat steps 2 to 4 until the model-predicted maximum growth and corresponding \( Y_{X/P} \) agree with the experimental measurement.

An identical procedure was followed to compute the \( \text{SUR} \) for growth on formate, assuming the same NGAM and GAM as for growth on \( \text{CO}_2 \). For the \( i\text{Mhu}428 \) model, the NGAM was estimated to be 0.6 mmol ATP/gDW/day, GAM was estimated to be 47 mmol ATP/gDW, \( \text{SUR}_{\text{CO}_2} \) was estimated to be 75.7 mmol/gDW/day, and \( \text{SUR}_{\text{formate}} \) was estimated to be 955 mmol/gDW/day.

The final \( i\text{Mhu}428 \) reconstruction contains 806 reactions, 273 genes (associated with 285 reactions), 706 metabolites. Of the 273 genes, 196 were added based on sequence homology, and 77 were added manually. The \( i\text{Mhu}428 \) reconstruction is available in S2 Dataset and S3 Dataset, in Excel and SBML formats, respectively.

### 3.2 Reconstruction and Parameterization of the \( i\text{Sfu}648 \) Metabolic Model

The \( i\text{Sfu}648 \) reconstruction of \( S. \text{fumaroxidans} \) was built using the RAVEN Toolbox [42], which uses sequence orthology to construct draft reconstructions from the proteins and reactions in the KEGG database. Briefly, the RAVEN Toolbox uses protein homology to identify the KEGG Orthology (KO) ID, which best matches each gene. The reactions and genes corresponding to that KO ID are then imported into the \( i\text{Sfu}648 \) reconstruction. Sequence homology was computed using default parameters: e-value < \( 10^{-30} \), alignment length > 200 nucleotides, sequence similarity > 40%. The resulting draft reconstruction was manually refined following recommendations given in a recent review [73]. While the \( i\text{Sfu}648 \) reconstruction was built using the RAVEN Toolbox, curation and refinement occurred in the GAMS modeling environment.

#### 3.2.1 Curation of Draft \( i\text{Sfu}648 \) Model

First, non-metabolic reactions were removed from the draft reconstruction (primarily those involved in tRNA charging). Next, all generic metabolites (e.g., acceptor) were identified and replaced with specific metabolites (e.g., NAD). Then, all reactions in the \( i\text{Sfu}648 \) reconstruction were mass- and charge-balanced, with a generic ferredoxin molecule used for charge balancing if atomic balancing was insufficient to do so.

In addition, all gene-protein-reaction (GPR) relationships were evaluated. Instead of providing detailed GPRs, the RAVEN Toolbox generates lists of genes associated with each reaction, and requires users to define the detailed GPR structure themselves.

First, gene annotations were examined to ensure that they matched the function of the associated reactions. This was particularly problematic in the case of one-to-many relationships, in which a single protein can carry out multiple reactions. For each gene which matches a particular KO group, the RAVEN Toolbox associates all
reactions in that group with the gene. Many such reactions were removed, due to a low likelihood of physiological relevance. In addition, 80 of the reactions contained in the final iSfu648 reconstruction had their GPRs adjusted through the removal of one or more genes, and 16 had their GPRs replaced entirely (approximately 11% of all reactions). The draft reconstruction contained a number of hypothetical genes, and both these genes and their associated reactions were removed from the iSfu648 reconstruction.

The logical structures of the GPRs provided by the RAVEN Toolbox were also determined. In addition to simple associations, in which a single gene encodes a single enzyme, GPRs can take the form of isozymes, in which multiple genes encode distinct proteins carrying out the same function, and multimeric protein complexes, in which multiple genes encoding distinct protein subunits come together to form an active enzyme. Genes annotated as separate subunits of a complex were given an ‘AND’ relationship, while genes with no such annotation were given an ‘OR’ relationship. Of the reactions retained in the final iSfu648 reconstruction, approximately 32% of them required manual curation of the logical relationships among the subunits and isozymes.

All told, the final iSfu648 reconstruction retained 717 of 859 reactions (83%) and 556 of 720 genes (77%) from the draft reconstruction, and around 50% of the GPRs required some level of curation.

Finally, the RAVEN Toolbox does not provide reaction direction information, instead assuming reactions are bidirectional in the absence of any specifying information. Reaction directions for the iSfu648 reconstruction were computed using group contribution, as described in the TMFA-LP section.

3.2.2 Manual Curation of Energy Conservation Mechanisms
S. fumaroxidans contains a wide variety of hydrogenases, dehydrogenases, and other enzyme complexes involved in electron transport [8,9,48–51]. A brief description of each reaction and justification for each annotation are given below. These reactions were incorporated into the final iSfu648 reconstruction, replacing draft reconstruction content where it disagreed with experimental evidence.

3.2.2.1 Hydrogenases and Formate Dehydrogenases
S. fumaroxidans contains a large number of gene clusters encoding hydrogenases and formate dehydrogenases, which we annotate as follows:

The gene cluster Sfum_0844-0846 encodes for a bifurcating [FeFe]-hydrogenase [8]. In Thermotoga maritima, this complex has been shown to simultaneously use electrons from NADH and reduced ferredoxin to produce H₂ [74], a process known as confurcation. A similar function has been proposed for a formate dehydrogenase (Sfum_2703-2707), and a [NiFe]-hydrogenase (Sfum_2713-2716) [8]. Experimental evidence supports these annotations [51], with the [NiFe]-hydrogenase having formate dehydrogenase activity. The reactions associated with these three complexes are described as “confurcating” H₂ases and FDHases in our model.

The gene clusters Sfum_0030-0031 and Sfum_1273-1275 are annotated as cytosolic formate hydrogenases, and the clusters Sfum_2952-2953 and Sfum_2220-2222 as cytosolic hydrogenases [8,48]. It has been proposed that these complexes are involved in succinate oxidation, where they receive electrons from b- and c-type cytochromes [8,48]. The cytochromes would in turn receive electrons from the
menaquinone pool, as has been observed in Desulfovibrio vulgaris [75]. However, group contribution methods cannot compute free energies for cytochromes and cytochrome-containing reactions, so we have chosen to lump the two steps into a single reaction, in which electrons are transferred directly from menaquinol. We describe these reactions as “cytochrome mediated” to emphasize the role cytochromes play in this process.

Finally, the genome of S. fumaroxidans contains clusters encoding for two cytoplasmic [NiFeSe]-hydrogenases (Sfum_3535-3537 and Sfum_3954-3956), and a cytoplasmic formate dehydrogenase (Sfum_3509-3511) [8,51]. Experimental evidence suggests these complexes are induced to reoxidize ferredoxin [51], so we have given these reactions ferredoxin reducing H2ases and FDHase activity, respectively.

### 3.2.2.2 Formate-Hydrogen Lyase
The gene cluster Sfum_1791-1796 shows similar to those coding for cytoplasmic-oriented membrane proteins which interconvert hydrogen and CO2 to formate [51], collectively referred to as formate-hydrogen lyase activity. Specifically, Sfum_1795-1796 show similarity to fhlF, and Sfum_1791 shows similarity to fhlH. We have assigned the entire cluster a formate-hydrogen lyase activity.

### 3.2.2.3 RNF Complex
The gene cluster Sfum_2694-2699 encodes for an RNF cluster [8,51], which has been propose to reoxidize NADH by ferredoxin reduction via the use of a proton motive force [8,51]. This proposal is consistent with the biochemistry of the RNF complex in a variety of prokaryotes [76].

### 3.2.2.4 Succinate Dehydrogenase
The gene cluster sdhABC (Sfum_1998-2000) codes for two cytoplasmic subunits (sdhA and sdhB) and a trans-membrane subunit containing heme (sdhC) which are similar to the succinate dehydrogenase subunits of Bacillus subtilis [48]. In addition, S. fumaroxidans contains menaquinones [37]. Oxidation of succinate to fumarate via menaquinone is an endergonic process ($\Delta G^0 \approx 60$ kJ/mol by our estimate). In B. subtilis and sulfate-reducing bacteria [77,78], the energy of a proton gradient is used to move electrons energetically “uphill” via a reverse electron transfer mechanism involving heme groups.

### 3.2.2.5 Fumarate Reductase
The gene cluster frdABEF (Sfum_4092-4095) encodes a fumarate reductase [48] which is believed to function similarly to that of Wolinella succinogenes [8,79]. In this bacterium, the transmembrane proton gradient is generated by hydrogen or formate oxidation coupled to fumarate reduction via cytochromes and a menaquinone loop [80]. While the proton translocation is believed to occur via the hydrogenases and formate dehydrogenases, the energy comes from fumarate reduction. Furthermore, our thermodynamic model predicts proton translocation via the H2ase or FDH to be impossible. For these reasons, we have modeled proton translocation as part of the fumarate reductase reaction.
3.2.2.6 NADH Dehydrogenase
The *S. fumaroxidans* genome contains two NADH dehydrogenase gene clusters (Sfum_0199-0209 and Sfum_1935-1943). We assigned the former to a NADH dehydrogenase using menaquinone without proton transport, and the latter to a NADH dehydrogenase using menaquinone with proton transport.

3.2.2.7 ATP Synthase
The endergonic oxidation of succinate to fumarate requires a proton gradient to function. Thermodynamic calculations suggest that about 0.67 mol ATP has to be invested to make this reaction energetically possible at conditions maintained by methanogens in a syntrophic community [5]. Because the succinate dehydrogenase reaction in the iSfu648 model pumps two protons, the ATP synthase reaction is assumed to pump three protons.

3.2.3 Manual Curation of Carbon Transfer Mechanisms
*S. fumaroxidans* metabolizes carbon in five well-defined ways, as described in Table 1. The steps of each pathway have been determined by $^{13}$C-NMR experiments [36,52], and a recent genomic survey identified the genes associated with each enzyme [48]. These pathways were incorporated into the final iSfu648 reconstruction, replacing draft reconstruction content where it disagreed with experimental evidence.

3.2.4 Experimental Justification for Theoretical Flux Distributions
Experimental studies have elucidated six growth modes for *S. fumaroxidans*: five in monoculture and one in coculture with *M. hungatei* (Table 2) [36,48,52]. Brief descriptions of the three growth modes examined in this work are presented below.

3.2.4.1 Monoculture Growth on Fumarate
$^{13}$C labeling studies reveal that one mole of fumarate gets fully oxidized to CO$_2$ for every six moles of fumarate which get reduced to succinate [48,52]:

$$7 \text{ fumarate} \rightarrow 6 \text{ succinate} + 4 \text{ CO}_2$$  \hspace{3cm} (S9)

Enzymatic assays revealed that oxidation of fumarate to CO$_2$ generates six reducing equivalents (NADH or reduced ferredoxin) [48,52], while the reduction of fumarate to succinate is driven by menaquinol [48]. Based on the results of gene expression assays [51], we propose the following mechanism for the transfer of electrons from NADH and reduced ferredoxin to menaquinol. High levels of NADH from fumarate oxidation induce RNF, which reoxidizes NADH by ferredoxin reduction. In turn, ferredoxin induces expression of the ferredoxin-oxidizing formate dehydrogenase and hydrogenase, which transfer electrons from ferredoxin to protons, generating H$_2$. The cytosolic hydrogenase then transfers electrons from H$_2$ to menaquinone, generating menaquinol. Reduction of fumarate to succinate generates the proton motive force (PMF) responsible for driving RNF and producing ATP.

3.2.4.2 Monoculture Growth on Fumarate and Propionate
$^{13}$C labeling studies reveal that one mole of propionate gets oxidized to succinate, while one mole of fumarate gets oxidized to acetate and CO$_2$. An additional two moles of fumarate get reduced to succinate [36,48]:

$$\text{propionate} + 3 \text{ fumarate} \rightarrow \text{acetate} + \text{CO}_2 + 3 \text{ succinate}$$  \hspace{3cm} (S10)
Based on the results of gene expression assays [51], the following mechanism for energy conservation has been proposed. The oxidation of fumarate to acetate and CO₂ produces one mole NADH and one mole reduced ferredoxin. The confurcating hydrogenase couples NADH and ferredoxin oxidation with H₂ production. The cytosolic hydrogenase then transfers electrons from H₂ to menaquinone, generating menaquinol. Reduction of fumarate to succinate generates the proton motive force (PMF) responsible for ATP production.

3.2.4.3 Coculture Growth on Propionate

¹³C labeling studies reveal that propionate gets oxidized to fumarate and CO₂ via the methylmalonyl-CoA pathway [48,52].

\[ \text{propionate} \rightarrow \text{acetate} + \text{CO}_2 + 2 \text{H}_2 \quad (S11) \]

Based on the results of gene expression assays [51], the following mechanism for energy conservation has been proposed. A single ATP molecule is generated during the oxidation of propionate to succinate. This ATP molecule is hydrolyzed to generate the PMF necessary for the endergonic oxidation of succinate to fumarate. This oxidation generates menaquinol; the cytosolic hydrogenase then transfers electrons from menaquinol to two protons, generating H₂. The oxidation of fumarate to acetate and CO₂ produces one mole NADH and one mole reduced ferredoxin. The confurcating hydrogenase couples NADH and ferredoxin re-oxidation with H₂ production.

3.2.5 Regulatory Constraints

The expression levels of the enzymes described in 3.2.2 have been shown to vary across conditions [51]. We developed the following regulatory constraints to ensure the proper energy conservation mechanisms were active under each condition. These constraints are justified below, and are summarized in Table S3 in S1 Dataset of the Supporting Materials.

3.2.5.1 Confurcating Hydrogenases and Formate Dehydrogenases

Gene expression evidence indicates these reactions are expressed at statistically significant higher levels during coculture growth, and in monoculture growth with fumarate and propionate. Thus, we allow these reactions to carry flux only under those conditions.

3.2.5.2 RNF Complex and Ferredoxin-Oxidizing Hydrogenases and Formate Dehydrogenases

Experimental evidence suggests that high NADH levels may induce transcription of genes encoding for the RNF complex, which couples NADH oxidation to ferredoxin reduction [51]. The ferredoxin-oxidizing hydrogenases and formate dehydrogenases then reoxidize the ferredoxin [51]. Because growth on fumarate in monoculture generates significantly more NADH than other growth modes, we allow these reactions to carry flux only under that condition.

3.2.5.3 Cytosolic Hydrogenases and Formate Dehydrogenases

In the conceptual model of energy conservation in S. fumaroxidans, the cytosolic hydrogenases (Sfum_0030-0031 and Sfum_1273-1275) and formate dehydrogenases (Sfum_2952-2953 and Sfum_2220-2222) are involved in succinate oxidation and fumarate reduction, where they receive electrons from the menaquinone pool. We
introduced flux coupling constraints to ensure that electrons from succinate oxidation and fumarate reduction are transferred strictly to the cytosolic hydrogenases and formate dehydrogenases.

3.2.5.4 NADH dehydrogenases.
It is unknown what, if any, role the two NADH dehydrogenase gene clusters play in energy conservation. To the best of our knowledge, no conceptual models of energy conservation in S. fumaroxidans predict a role for the NADH dehydrogenases, and their expression levels were not measured in the single gene expression study we were able to find [51]. In light of this, we chose to constrain the two NADH dehydrogenase reactions to carry zero flux across all conditions.

3.2.5.5 Fumarate Reductase and Succinate Dehydrogenase
Fumarate reductase is predicted to be active only during growth in monoculture, and to exclusively reduce fumarate to succinate. Succinate dehydrogenase is predicted to be active only during growth in coculture, and to exclusively oxidize succinate to fumarate. Thus, these two reactions were constrained as appropriate during each simulation condition.

3.2.6 Development and Gapfilling of Biomass Equation
The RAVEN Toolbox does not provide users with a biomass equation; they must instead construct one manually. Unfortunately, the dry cell weight biomass composition of S. fumaroxidans has not yet been described. In this work, a template biomass equation (proposed by the developers of the Model SEED [81]) was used as a scaffold to construct the biomass equation based on data from related organisms.

The weight fractions of major macromolecule classes (proteins, DNA, etc) were taken from Geobacter sulfurreducens, a deltaproteobacterium which has been extensively studied and modeled [13,82]. (S. fumaroxidans is also a deltaproteobacterium). Mole fractions of individual macromolecules were obtained as follows:

- amino acids, from the iAF1260 model of E. coli [83];
- DNA and RNA, from the S. fumaroxidans genome sequence [48];
- lipids and carbohydrates, from G. sulfurreducens [84];
- cell wall, cofactors, and small molecules, from the Model SEED template.

Additional details can be found in Table S5 of S1 Dataset in the Supporting Materials.

The iSfu648 model was unable to sustain flux through this biomass equation under any known growth condition. Individual biomass precursors which the iSfu648 model could not synthesize were identified, and SMILEY was used to identify those reaction additions which would enable biomass growth [85]. Manual curation was used to identify reactions necessary for lipid, cell wall, and carbohydrate biosynthesis. Where possible, those solutions for which genomic evidence could be found were selected. In the case of many cofactors, no genomic evidence was found for any solution. These cofactors were eliminated from the biomass equation.

SMILEY is a mixed-integer programming approach for model refinement, which calculates the minimum number of reactions from a universal reaction database which must be added to a reconstruction to enable cellular growth. For this study, the universal reaction database was a manually-curated subset of version 57 of the KEGG
database, available through the BioWebDB Consortium (http://www.biowebdb.org). During curation, all reactions which met any of the following criteria were removed:

- contained an elongation;
- contained the same metabolite on both sides;
- contained compounds in the KEGG glycan database;
- contained a compound with an R group;
- contained a compound without a formula.

For reactions containing a generic acceptor molecule, two separate versions of the reaction were created, one using NAD as an acceptor, and one using NADP. Additionally, any reaction which could not be balanced through the addition of protons or ferredoxins was removed. Reaction directions for the gapfilling database were computed using group contribution, as described in the TMFA-LP section of Methods. Metabolite formulas and molecular structure files (molfiles) were obtained directly from KEGG. The curated database contains over 6000 reactions and represents approximately 70% of the original KEGG version 57 database. The curated database is provided in S6 Dataset and S7 Dataset, in Excel and SBML formats, respectively.

### 3.2.7 Gapfilling of Blocked Reactions

Flux Variability Analysis (FVA) [86] was used to identify reactions which were blocked under one or more growth conditions. SMILEY was used to identify gapfilling solutions for these blocked reactions, and those reactions with genomic evidence were added to the iSfu648 model.

### 3.2.8 Conversion to a Thermodynamic Model

The iSfu648 reconstruction was then converted to a thermodynamic model, as described in “Metabolite Structures and Molecular Formulas” and “Thermodynamics-Based Metabolic Flux Analysis.” On defined minimal media (Table S1 in S1 Dataset in the Supporting Material), TMFA predicted a no-growth phenotype, most likely because biosynthetic routes which were feasible under TMFA-LP may not be feasible under TMFA [34]. Once again, SMILEY identified those biomass precursors which the iSfu648 model could not synthesize, and solutions which would enable biomass growth under TMFA were added to the iSfu648 reconstruction.

The final iSfu648 reconstruction contains 874 reactions, 648 genes (associated with 770 reactions), 893 metabolites. Considering the entire content of the final iSfu648 reconstruction, reactions and genes contained in the draft reconstruction make up 82% and 86% of the total reaction and gene content, respectively. The iSfu648 reconstruction is available in S4 Dataset and S5 Dataset, in Excel and SBML formats, respectively.

### 3.2.9 iSfu648 Model Parameterization

Substrate uptake rates (SURs), and growth- (GAM) and non-growth-associated (NGAM) maintenance requirements for the iSfu648 model of *S. fumaroxidans* were estimated using experimental data.

The value of the NGAM parameter was found by maximizing ATP hydrolysis at the reported maintenance cost as measured in terms of propionate uptake [53].
The value of the GAM parameter is typically obtained by plotting a series of substrate uptake and growth rates taken from chemostat experiments [72], but no such data were available for S. fumaxoridans. However, studies have reported the growth rate of S. fumaxoridans on a variety of substrates [51], and its yield per mol propionate (Y_{X/S}) [72]. From these data, the following iterative procedure was used to simultaneously estimate SUR_{propionate} and GAM:

1. Select an initial estimate for the GAM.
2. Identify the SUR_{propionate} for which the observed growth rate and model-predicted maximum growth rates agree.
3. Adjust the value of the GAM upwards or downwards until Y_{X/P} at the model-predicted maximum growth agrees with the experimental measurement. Changing the GAM will change the model-predicted maximum growth. Thus,
4. Repeat steps 2 to 3 until the model-predicted maximum growth and corresponding Y_{X/P} agree with the experimental measurement.

Step 2 was used to compute the SUR for growth on fumarate, assuming the same NGAM and GAM as for growth on propionate. For the iSfu648 model, the NGAM was estimated to be 3.36 mmol ATP/gDW/day, GAM was estimated to be 22.8 mmol ATP/gDW, SUR_{propionate} was estimated to be 37.7 mmol/gDW/day, and SUR_{fumarate} was estimated to be 27.6 mmol/gDW/day.

### 3.3 Metabolite Structures and Molecular Formulas

To estimate Gibbs free energies, molecular structure files (molfiles) for metabolites in the reconstructions were obtained from KEGG (iMhu428 and iSfu648) or manually constructed (iMhu428). For the iMhu428 reconstruction, molfiles were obtained for 94% of the metabolites, enabling the calculation of the standard transformed Gibbs free energy of reaction (∆rG'0) for 83% of the reactions. For the iSfu648 reconstruction, molfiles were obtained for 100% of the metabolites, enabling the calculation of ∆rG'0 for 84% of the reactions.

All metabolites were then converted to their predominant ionic species (pseudoisomer) at biochemical standard state: pH 7, zero ionic strength, and temperature 298K. The major pseudoisomer form of each molecule was determined using pKa estimation software (Marvin pKa plug-in, version 5.11.4, ChemAxon, Budapest, Hungary). Finally, all reactions in the reconstructions were mass- and charge-balanced using the new metabolite formulas. Supplemental Files “S1 File” and “S2 File” contain molfiles for all compounds in the iMhu428 and iSfu648 reconstructions, respectively.

### 3.4 Thermodynamics-Based Metabolic Flux Analysis (TMFA)

Flux-balance analysis (FBA) [27] is a constraint-based technique for predicting the state of a metabolic network consistent with physiochemical principles. FBA identifies a flux distribution which maximizes cellular growth (or some other objective function), subject to steady-state mass-balance and enzyme capacity constraints.

Specifically, given a stoichiometric matrix S and a set of reactions J, FBA seeks a steady-state flux distribution v maximizing the flux through the biomass reaction (v_{BM}).
while also satisfying mass-balance and enzyme capacity constraints for individual reactions, $j$:

$$\begin{align*}
\text{max} & \quad v_{BM} \\
\text{s.t.} & \quad S \cdot v = 0 \\
& \quad v_{\text{min}} \leq v_j \leq v_{\text{max}} \quad \forall j \in J 
\end{align*}$$

Enzyme capacities, $v_{\text{min}}$ and $v_{\text{max}}$, should be set on the basis of available evidence, such as thermodynamic irreversibility. Fluxes are typically constrained to $v_{\text{max}} = 1000$ mmol/gDW/day and $v_{\text{min}} = -1000$ mmol/gDW/day, except for measured fluxes (e.g., carbon uptake rates). Uptake rates used in the simulations are given in Table S1 in S1 Dataset of the Supporting Material.

Thermodynamics-Based Metabolic Flux Analysis (TMFA, [33,34]) extends FBA via the introduction of thermodynamic constraints. TMFA assumes no a priori evidence for enzyme directionality, instead enforcing the second law of thermodynamics, which states that the transformed Gibbs free energy of reaction ($\Delta_r G'$) and flux ($v$) have opposite signs:

$$v_j \cdot \Delta_r G_j' < 0$$

This nonlinear constraint is converted to a mixed-integer constraint as described previously [34].

The $\Delta_r G'$ of a reaction is in turn a function of the standard transformed Gibbs free energy of reaction ($\Delta_r G^0$) and the concentrations ($x_i$) of those metabolites participating in the reaction:

$$\Delta_r G_j' = \Delta_r G_j^0 + RT \sum_i S_{i,j} \ln(x_i)$$

where $R$ is the gas constant, and $T$ is the temperature. In the absence of specific information, metabolite concentrations are constrained to global bounds of 0.01 mM to 20 mM. Specific metabolite concentrations were imposed for water, protons, and compounds in the growth media, as indicated in Table S2 in S1 Dataset of the Supporting Material.

Due to a paucity of experimental data, group contribution methods [61,87–89] are used to provide estimates and uncertainties of $\Delta_r G^0$ for reactions. Estimates ($\Delta_r G_{j,\text{est}}^0$) and uncertainties ($SE_{\Delta_r G_{j,\text{est}}^0}$) of $\Delta_r G^0$ for the reactions in the reconstructions were obtained using a software implementation of the latest group contribution method for biological systems, which the authors refer to as component contribution [61]. The implementation is available via the von Bertalanffy 2.0 add-on to the COBRA Toolbox [62].

The GCM method returns estimates of $\Delta_r G^0$ for each reaction, given the pKa values for all compounds in the reconstruction, and information on temperature ($T$), pH, ionic strength ($I$) and electrical potential ($\Phi$) in each cellular compartment. pKa values were calculated using pKa estimation software (Marvin pKa plug-in, version 5.11.4, ChemAxon, Budapest, Hungary) from molecular structure files (molfiles).
Thermodynamic parameters (T, pH, I, and Φ), and $\Delta_r G_{j,est}^{0}$ and $SE_{\Delta_r G_{j,est}^{0}}$ values for each reconstruction are given in S2 Dataset (iMhu428) and S4 Dataset (iSfu648).

The standard transformed Gibbs free energy ($\Delta_r G^{0}$) of each reaction was allowed to vary within its 95% confidence interval, as determined by the standard error (SE) reported by the GCM software:

$$\Delta_r G_{j,est}^{0} - 2SE_{\Delta_r G_{j,est}^{0}} \leq \Delta_r G_j^{0} \leq \Delta_r G_{j,est}^{0} + 2SE_{\Delta_r G_{j,est}^{0}}$$

(S17)

If the GCM method is unable to obtain an estimate of $\Delta_r G^{0}$ for a particular reaction, $\Delta_r G^{0}$ was allowed to vary freely.

The GCM method estimated $\Delta_r G^{0}$ for approximately 82% of the reactions in the iMhu428 reconstruction, and approximately 84% of the reactions in the iSfu648 reconstruction. For some reactions for which $\Delta_r G^{0}$ could not be estimated, a lumping approach was used to constrain $\Delta_r G^{0}$, as described in “Thermodynamic Lumping.”

Aggregating the above constraints gives the final formulation for TMFA:

$$\max \quad v_{BM}$$

s.t.  
FBA constraints, (S13) and (S14) \quad \forall j \in J \quad \text{(TMFA)}

$\Delta_r G^{0}$ constraints, (S16) and (S17) \quad \forall j \in J

consistency constraints, (S15) \quad \forall j \in J

3.5 TMFA and Thermodynamics of H₂ Production

When evaluating the thermodynamic feasibility of H₂ production, simulations with the iSfu648 model were performed using estimates of Gibbs free energies of formation ($\Delta_j G_{i,est}^{0}$) as a basis for thermodynamic calculations. This switch from $\Delta_j G_{i,est}^{0}$ to $\Delta_j G_{i,est}^{0}$ as basis for thermodynamic calculations is necessary to capture the thermodynamic interconnectivity (arising from shared metabolites) between the individual reactions which contribute to the overall metabolic transformation. For these simulations, the formulation in the previous section was extended as follows:

First, the standard transformed Gibbs free energy of reaction $\Delta_j G^{0}$ was computed as a function of the standard Gibbs free energy of formation $\Delta_j G^{0}$ of its constituent metabolites, plus adjustments ($\Delta_{pH} G_r^{0}$ and $\Delta_{\Phi} G_r^{0}$) due to changes in pH gradient ($\Delta pH$) and electrochemical potential ($\Delta \Phi$) across the cell membrane:

$$\Delta_r G_j^{0} = \sum_i S_{j,i} \Delta_j G_i^{0} + \Delta_{pH} G_r^{0} + \Delta_{\Phi} G_r^{0}$$

(S18)

Values for $\Delta_{pH} G_r^{0}$ and $\Delta_{\Phi} G_r^{0}$ were computed by the GCM software, as described previously [33,34,61]. Estimates ($\Delta_j G_{i,est}^{0}$) and uncertainties ($SE_{\Delta_j G_{i,est}^{0}}$) of $\Delta_j G^{0}$ for the metabolites in the reconstructions were obtained using the GCM software, and $\Delta_j G_{i,est}^{0}$ and $SE_{\Delta_j G_{i,est}^{0}}$ values are given in S4 Dataset.
The standard transformed Gibbs free energy ($\Delta_f G^0$) of each metabolite was allowed to vary within its 95% confidence interval, as determined by the standard error (SE) reported by the GCM software:

$$\Delta_f G_{i,est}^0 - 2SE_{\Delta_f G_{i,est}^0} \leq \Delta_f G^0_i \leq \Delta_f G_{i,est}^0 + 2SE_{\Delta_f G_{i,est}^0}$$  \hspace{1cm} (S19)

If the GCM method is unable to obtain an estimate of $\Delta_f G^0$ for a particular reaction, $\Delta_f G^0$ was allowed to vary freely.

Aggregating these constraints with the formulation from the previous section gives the final formulation:

$$\max v_{BM}$$

s.t.  \hspace{0.5cm} FBA constraints, (S13) and (S14) \hspace{1cm} \forall j \in J

$$\Delta_r G^0_i$$ constraints, (S16) and (S17) \hspace{1cm} \forall j \in J \hspace{0.5cm} \text{(TMFA-dGf)}

$$\Delta_r G^0_i$$ constraints, (S19) and (S23) \hspace{1cm} \forall j \in J

consistency constraints, (S15) \hspace{1cm} \forall j \in J

After each H$_2$-producing mode was identified, the overall transformation was introduced as a new reaction to the iSfu648 model. To block flux through the reaction, a constraint was sought which would ensure $\Delta_r G > 0$ for the reaction, thereby eliminating the erroneous H$_2$-producing phenotype. (We have previously referred to such constraints as phenotype-correction constraints [34].) That is,

$$\Delta_r G_j^0 + RT \sum_i S_{i,j} \ln(x_i) > 0$$  \hspace{1cm} (S20)

Then, as long as

$$\max RT \sum_i S_{i,j} \ln(x_i) > \min \Delta_r G_j^0$$

adding a new constraint of the form

$$RT \sum_i S_{i,j} \ln(x_i) > -\min \Delta_r G_j^0$$  \hspace{1cm} (S21)

will ensure H$_2$ production via the identified mechanism is infeasible. Values of $\max RT \sum_i S_{i,j} \ln(x_i)$ and $\min \Delta_r G_j^0$ were found by solving TMFA-dGf with the objective function of $\max RT \sum_i S_{i,j} \ln(x_i)$ or $\min \Delta_r G_j^0$ as appropriate.

If such a constraint could be identified, it was added to the iSfu648 model and new H$_2$-producing reactions were sought. The process was repeated until no additional H$_2$-producing reactions could be found, or a phenotype-correction constraint could not be identified.

### 3.6 Thermodynamic Lumping

Due to limitations in the GCM [61], there were some reactions for which $\Delta_r G^0$ could not be estimated. Among these reactions were the formate dehydrogenase and confuracting hydrogenase involved in H$_2$ production during syntophic growth. To understand why these reactions were predicted to be active under monoculture
conditions, a previously described lumping approach [34] was used to constrain the free energies of these reactions.

In this approach, reactions with unknown $\Delta r G^0$ are linearly combined into lumped representations for which $\Delta r G^0$ can be calculated. The set of lumped reactions is called $J_L$, and introduced into the $S$ matrix as a subset of the set $J$. Constraints are derived which ensure both the lumped reactions and their constituent reactions remained thermodynamically consistent, irrespective of the value of $\Delta r G^0$ of the unknown reactions.

For lumped reactions, $\Delta r G'$ is calculated as described in (S16), with (S17) modified to allow $\Delta r G^0$ to vary freely. The transformed Gibbs free energy ($\Delta r G'$) of the lumped reactions is calculated from $\Delta r G'$ of their constituent reactions. To do this, parameters are defined such that $\alpha_{j,l} = 1$ if reaction $j$ (one of the reactions with unknown $\Delta r G^0$) combines in the forward direction to make up lumped reaction $l$, and defined $\alpha_{j,l} = -1$ if reaction $j$ combines in the reverse direction. The transformed Gibbs free energy ($\Delta r G'$) of the lumped reaction is then computed:

$$\Delta r G'_l = \sum_j \alpha_{j,l} \Delta r G'_j \quad \forall l \in J_L$$

(S23)

This constraint ensures that the thermodynamics of the lumped reaction and its constituents are internally consistent, irrespective of the value of $\Delta r G^0$ of the unknown reactions. The complete list of lumped reactions $J_L$, and parameters $\alpha$ can be found in S4 Dataset in the Supporting Material.

### 3.7 Linear Programming Approximation of TMFA (TMFA-LP)

Because TMFA is a mixed-integer program, its use may be impractical under certain conditions. For example, TMFA cannot be used as the inner problem in a bilevel optimization, and extending TMFA to large networks (as in gapfilling) is computationally intensive. As a consequence, during model refinement TMFA-LP, rather than TMFA, is used as described previously [34]. In this approach, extreme values for $\Delta r G'$ are computed as given by (S16), assuming $\Delta r G^0$ falls within the 95% confidence interval given by (S17). If the predicted $\Delta r G'$ range for a reaction was entirely negative, the reaction is assumed to proceed only in the forward direction (i.e., $v_{\text{min}} = 0$ mmol/gDW/day); if the predicted $\Delta r G'$ range is entirely positive, the reaction is assumed to proceed only in the reverse direction (i.e., $v_{\text{max}} = 0$ mmol/gDW/day). Otherwise, the reaction is allowed proceed in both directions (i.e., $v_{\text{min}} = -1000$ mmol/gDW/day, $v_{\text{min}} = 1000$ mmol/gDW/day).

### 3.8 Parsimonious TMFA (pTMFA)

While FBA often assumes selective pressure for the fastest growing strains, other selective forces may shape an organism’s phenotype [90]. pFBA [54] is a constraint-based approach which assumes selective pressure not only for the fastest growing strains, but also for the lowest total flux through the network (a proxy for minimizing the
total mass of enzymes required to sustain optimal growth through the network). pTMFA uses the same assumptions as pFBA while implementing the thermodynamic constraints of TMFA.

pTMFA was implemented as a two-stage optimization process. In Stage 1, TMFA is solved as described previously, where growth rate is maximized. In the second stage, the total flux through the network is minimized, by decomposing each flux $v$ into its non-negative forward and reverse components ($v^+$ and $v^-$), and summing over the flux components for all reactions in the network:

$$
\min \sum_j v^+_j + v^-_j \\
\text{s.t. } v = v^+ - v^- \\
\quad v^+ \geq 0 \\
\quad v^- \geq 0 \\
\tag{S24}
$$

To ensure that the pTMFA solution is also optimal with respect to growth, the growth rate is fixed to the Stage 1 solution:

$$
v_{BM} = \max (\text{TMFA})
\tag{S27}
$$

Aggregating the new constraints with those from TMFA gives pTMFA:

$$
\min \sum_j v^+_j + v^-_j \\
\text{s.t. } \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \\
\quad \Delta_r G^\text{c} \text{ constraints, (S16) and (S17)} \quad \forall j \in J \\
\quad \text{consistency constraints, (S15)} \quad \forall j \in J \\
\quad \text{flux decomposition constraints, (S25) and (S26)} \quad \forall j \in J \\
\quad \text{optimal biomass given by TMFA, (S26)} \\
\tag{pTMFA}
$$

### 3.9 Community Formulation

For the coculture simulations, a community model which simulates growth in a continuous stirred-tank reactor (CSTR) was developed. The formulation takes into account the biomass concentrations $X_n$ of each species, $n$. Conceptually, the formulation attempts to minimize the species-weighted total flux through the metabolic networks, subject to TMFA constraints for each species, and a mass balance around the entire reactor:

$$
\min \sum_n \left( X_n \sum_j (v^+_j + v^-_j) \right) \\
\text{s.t. } \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \quad \forall n \in N \\
\quad \Delta_r G^\text{c} \text{ constraints, (S16) and (S17)} \quad \forall j \in J \quad \forall n \in N \\
\quad \text{consistency constraints, (S15)} \quad \forall j \in J \quad \forall n \in N \\
\quad \text{reactor balance, (S28)} \quad \forall n \in N \\
\quad v_{BM,n} = D
\tag{CSTR}
$$

As described in Results and illustrated in Fig 3B, both species exchange metabolites with a shared pool of metabolites $I$. Each metabolite $i_n$ is exported to (or imported from) the shared pool to species, $n$, via an exchange reaction $j_n$ at flux $v_{j,n}$ (with
units of mmol/gDW/day). Likewise, each metabolite $i$ has a net in (or out) flow from the reactor, denoted by rate $F_i$ (note the units for $F$ are mmol/L/day). Thus, each metabolite can flow into/out of the shared pool via $|N|+1$ reactions. Such tuples of reactions are indicated by the set $J_{n=1,...,|N|,i}$, with the set $J_{\text{Shared}}$ containing all such tuples. The mass balance around the reactor can be then written:

$$\sum_{n} X_n v_{j,n} = F_i \quad \forall (j_{n=1,...,|N|,i}) \in J_{\text{Shared}}$$

(S28)

Realizing that the reactor operates at a fixed dilution rate ($D$) [91], the final constraint can be derived from a mass-balance around each species:

$$v_{BM,n} = D \quad \forall n \in N$$

As presented, this formulation is general and can be applied to a community with any number of species.

The use of thermodynamic feasibility constraints necessitates the use of concentrations as variables in the coculture formulation. Additional constraints enforce this consistency of metabolite concentrations. A set of tuples $(i_{n=1,...,|N|,i},)$ called $I_{\text{Shared}}$ indicates those tuples of in silico metabolites corresponding to the same physical metabolite (e.g., $\text{CO}_2$ has an extracellular version for each strain, and a version in the shared pool, for a total of three versions). A constraint forces each version of a metabolite to have the same concentration, thus ensuring that a concentration constraint on metabolite $i$ in the shared pool constrains the metabolite concentration in each model as well:

$$\ln(x_{i,n=1}) = \ldots = \ln(x_{i,n=|N|}) = \ln(x_i) \quad \forall (i_{n=1,...,|N|,i}) \in I_{\text{Shared}}$$

(S30)

Failing to do so may result in inconsistent thermodynamic predictions arising from a single metabolite having multiple concentrations.

Finally, metabolites known to be exported to (or imported from) the shared space by each species via diffusion are identified. For each such metabolite, a pair $(i_n, T_n)$ containing the extra- ($\hat{i}$) and intracellular ($\bar{I}$) versions of that metabolite gets defined, as do sets $I_{n,\text{Out}}$ and $I_{n,\text{In}}$ containing those pairs which diffuse out and in, respectively. Constraints on these metabolite concentrations ensure consistency with diffusion:

$$\ln(x_{i,n}) < \ln(x_{\hat{i},n}) \quad \forall (i_n, T_n) \in I_{n,\text{Out}}$$

$$\ln(x_{i,n}) > \ln(x_{\bar{i},n}) \quad \forall (i_n, T_n) \in I_{n,\text{In}}$$

(S31)

The expanded coculture model formulation includes those constraints necessary for thermodynamic consistency across models.

$$\min \sum_{n} \left( X_n \sum_{j,n} (v_{j,n}^+ + v_{j,n}^-) \right)$$

s.t. FBA constraints, (S13) and (S14) \hspace{1cm} \forall j \in J \hspace{1cm} \forall n \in N

\Delta G' constraints, (S16) and (S17) \hspace{1cm} \forall j \in J \hspace{1cm} \forall n \in N

consistency constraints, (S15) \hspace{1cm} \forall j \in J \hspace{1cm} \forall n \in N

concentration consistency constraints, (S30) \hspace{1cm} \forall n \in N

diffusion constraints, (S31) \hspace{1cm} \forall n \in N

reactor balance, (S28) \hspace{1cm} v_{BM,n} = D \hspace{1cm} \forall n \in N

(CSTR)

21
Definitions for the sets $l_{in}, l_{ln}, l_{Out}$, and concentrations for metabolites in the shared pool are given in Table S2 of S1 Dataset of the Supporting Material.

This formulation is a mixed-integer non-linear program (MINLP): the thermodynamic consistency constraints (S4) are integer constraints, and the reactor balance constraint (S12) is nonlinear. To avoid solving the MINLP, the dilution rate $D$ and biomass concentrations of $X_n$ for all species were fixed, converting the MINLP into an easily-solvable MIP.

To explore the behavior of the community under a variety of operating conditions, the reactor dilution rate and the relative ratio of $M. \text{hungatei}$ to $S. \text{fumaroxidans}$ were systematically changed, while allowing unlimited propionate uptake by the reactor ($F_{\text{propionate}}$). Simulations were performed at dilution rates between 0 to 0.2 days$^{-1}$ at an interval of 0.01 days$^{-1}$. The ratio of $M. \text{hungatei}$ to $S. \text{fumaroxidans}$ was varied from 0 to 1.6, with simulations performed at intervals of 0.1. The maximal dilution rate of 0.2 days$^{-1}$ corresponds to the maximal growth rate of the slower-growing $S. \text{fumaroxidans}$.

### 3.10 Community Thermodynamic Variability Analysis

We performed community thermodynamic variability analysis (cTVA) to identify thermodynamically feasible metabolite concentration ranges for each metabolite in the coculture. Simulations were performed at dilution rates between 0 to 0.2 days$^{-1}$ at an interval of 0.01 days$^{-1}$. The ratio of $M. \text{hungatei}$ to $S. \text{fumaroxidans}$ was varied from 0 to 1.6, with simulations performed at intervals of 0.1. Feasible metabolite concentration ranges for the coculture were obtained by maximizing (or minimizing) individual metabolite concentrations, subject to the constraints of the CSTR formulation. In addition, the species-weighted total flux through the metabolic network was constrained to the minimal value predicted by the CSTR formulation. That is, for each metabolite, the following optimization was performed:

$$\max (\min) \ln(x_{i,n})$$

s.t.  

- FBA constraints, (S13) and (S14) $\forall j \in J$ $\forall n \in N$
- $\Delta_j \Gamma$-consistency constraints, (S16) and (S17) $\forall j \in J$ $\forall n \in N$
- Concentration consistency constraints, (S15) $\forall j \in J$ $\forall n \in N$
- Diffusion consistency constraints, (S30) $\forall n \in N$
- Reactor balance, (S28)
- $D = \hat{D}$ $\forall n \in N$
- $\frac{X_{S.\text{fum.}}} {X_{M.\text{hun.}}} = \frac{\hat{X}_{S.\text{fum.}}} {\hat{X}_{M.\text{hun.}}}$
- $\sum_{n}(x_n\sum_{j,n}(v^+_j + v^-_j))$ given by solving (CSTR) for $\hat{D}, \hat{X}_{S.\text{fum.}}, \hat{X}_{M.\text{hun.}}$

### 3.11 Minimal Probabilistic Sets (MPS)

Under some conditions, thermodynamic feasibility constraints predict that a solution space contains undesired flexibility (e.g., the maximal predicted rate of ATP hydrolysis...
exceeds the experimental observation). A previous study of thermodynamic constraints used probability to qualitatively constrain reaction directions and reduce network flexibility [68]. This approach calculated the probability that a reaction’s $\Delta_r G^0$ was negative. If the probability was greater (less) than 70% (30%), the reaction was constrained to the forward (reverse) direction. Otherwise, the reaction is assumed to be bidirectional.

This probabilistic approach was applied to reduce network flexibility in *M. hungatei*, but was found to reduce network flexibility too much, and phenotypes which were previously correct become incorrect (e.g., the model no longer predicted growth). To prevent this, an optimization procedure was developed to identify the smallest number of probabilistic (qualitative reaction direction) constraints needed to correct one or more phenotypes, while preserving one or more other phenotypes. The required formulation, Minimal Probabilistic Sets (MPS) maximizes network flexibility (as measured by the number of unconstrained reactions), with the probabilistic (qualitative reaction direction) constraints identified by subtraction. The formulation is defined as follows:

The sets $J_{\text{probFwd}}$ and $J_{\text{probRev}}$ consist of those reactions $j$ in $J$ which probability suggests should be constrained to the forward or reverse directions, respectively, while the set $J_{\text{Bidir}}$ contains those reactions assumed to be directional. The binary variables $y$ and $z$ take a value of 0 if a reaction is to be constrained to the forward or reverse direction, respectively, and take a value of 1 otherwise. The appropriate constraints are:

$$y_j \in \{0,1\} \quad \forall j \in J_{\text{probFwd}}$$

$$z_j = 1 \quad \forall j \in J_{\text{probFwd}}$$

$$y_j \in \{0,1\} \quad \forall j \in J_{\text{probRev}}$$

$$z_j = 1 \quad \forall j \in J_{\text{probRev}}$$

$$y_j = z_j = 1 \quad \forall j \in J_{\text{Bidir}}$$

The objective is to maximize the network flexibility (as measured by the number of unconstrained reactions), while maintaining consistency with each $n$ of $N$ phenotypes:

$$\text{max} \sum_{j,n} y_{j,n} + \sum_{j,n} z_{j,n}$$

subject to

- phenotype constraints $\forall n \in N$ (S36)
- computed cellular phenotype $\forall n \in N$ (S37)
- $y_{j,n} = y_{j,\tilde{n}} \quad \forall (n,\tilde{n}) \in N$ (S38)
- $z_{j,n} = z_{j,\tilde{n}} \quad \forall (n,\tilde{n}) \in N$ (S39)

The final constraint (S38) ensures the same qualitative reaction direction constraints are applied to each computed cellular phenotype (S37).

Those reactions which require a qualitative reaction direction constraint can be easily identified, as the $y$ or $z$ variable associated with that reaction will have a value of 0. The desired phenotypes (S36) can be imposed as a variety of constraints, e.g.,

$$v_{BM} \geq \varepsilon$$

$$v_{ATP} = \gamma$$

and may enforce phenotypes to be corrected (e.g., ATP gain) or maintained (e.g., cellular growth). Individual cellular phenotypes (S37) are computed by an inner problem,
which maximizes or minimizes some cellular objective, $v_{obj}$, subject to mass-balance and qualitative reaction direction constraints under a particular media condition:

$$\max \quad v_{obj} \quad \text{(S40)}$$

$$\text{s.t.} \quad S \cdot v = 0 \quad \text{(S41)}$$

$$v_{\min} \leq v_j \leq v_{\max} \quad \forall j \in J \quad \text{(S42)}$$

media constraints

$$v_j \geq 0 \quad \forall j \in y_j = 0 \quad \text{(S44)}$$

$$v_j \leq 0 \quad \forall j \in z_j = 0 \quad \text{(S45)}$$

The inner problem enforces the qualitative reaction direction (probabilistic) constraints (S44) and (S45) identified by the outer problem. Because thermodynamic feasibility constraints cannot be enforced directly in the inner problem, reactions are constrained using their TMFA-LP directions instead.

The final outer problem constraint (S38) ensures that the same qualitative reaction direction constraints are enforced on each inner problem. Aggregating all of the above constraints gives the final formulation for MPS:

$$\max \quad \sum_{j,n} y_{j,n} + \sum_{j,n} z_{j,n} \quad \text{(MPS)}$$

s.t. computed cellular phenotype(s), (S40) to (S45) $\forall n \in N$

phenotype constraint(s), (S39) $\forall n \in N$

allowed probabilistic constraints, (S32) to (S34) $\forall n \in N$

consistency of probabilistic constraints, (S38) $\forall n \in N$

For purposes of implementation, this bilevel problem is converted to a single-level problem via well-established techniques which rely on duality theory [92,93].

The formulation enables any number of phenotypes to be predicted by the inner problems, provided that a constraint for each is given in the outer problem. MPS was able to correct a number of phenotypes in the iMhu428 and iSfu648 models.

### 3.11.1 MPS for Validating iMhu428: ATP Gain

Experimental evidence suggests that *M. hungatei* is able to generate 0.5 mol ATP per mol of CO$_2$ converted to CH$_4$ [7]. A methanogenesis pathway which produces this ATP yield could be identified, but the iMhu428 model predicted other, higher-yielding, ATP-generating mechanisms outside the methanogenesis pathway. MPS was used to identify qualitative reaction direction constraints which would eliminate these extra mechanisms, while simultaneously ensuring biomass growth:

$$\max \quad \sum_{j} y_j + \sum_{j} z_j \quad \text{(MPS)}$$

s.t. $\max$ ATP gain

s.t. FBA constraints, (S13) and (S14) $\forall j \in J$

carbon uptake via CO$_2$ and acetate

enforce probabilistic constraints, (S44) and (S45) $\forall j \in J$

ATP gain = 0.5

$v_{BM} \geq 0.01$

allowed probabilistic constraints, (S32) to (S34) $\forall n \in N$
MPS identified a set of 38 reactions to be constrained to a single direction (listed in Table S2 in S1 Dataset of the Supporting Material).

### 3.11.2 MPS for Validating \(\text{iMhu}428\): Biomass Growth

It has been observed that biomass and energy generation are independent and non-interacting in \(M.\ hungatei\), with methanogenesis via \(\text{CO}_2\) being the sole source of ATP, and acetate being the sole source of carbon in the biomass [55]. In contrast to experimental evidence, the \(\text{iMhu}428\) model predicts that \(M.\ hungatei\) can produce some biomass from \(\text{CO}_2\) alone. MPS was used to resolve the discrepancy with two inner problems: the first inner problem tries to enforce a no growth phenotype on \(\text{CO}_2\) alone, while the second inner problem tries to maintain growth on acetate alone:

\[
\begin{align*}
\text{max} & \quad \sum_{j,n} y_{j,n} + \sum_{j,n} z_{j,n} \\
\text{s.t.} & \quad \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \\
& \quad \text{carbon uptake via } \text{CO}_2 \\
& \quad \text{enforce probabilistic constraints, (S44) and (S45)} \quad \forall j \in J \\
& \quad v_{BM} = 0 \\
\end{align*}
\]

\[
\begin{align*}
\text{max} & \quad v_{BM} \\
\text{s.t.} & \quad \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \\
& \quad \text{carbon uptake via acetate} \\
& \quad \text{enforce probabilistic constraints, (S44) and (S45)} \quad \forall j \in J \\
& \quad v_{BM} \geq 0.01 \\
& \quad \text{allowed probabilistic constraints, (S32) to (S34)} \quad \forall n \in N \\
& \quad \text{consistency of probabilistic constraints, (S38)} \quad \forall n \in N \\
\end{align*}
\]

The problem proved infeasible, meaning that with the current \(\text{iMhu}428\) model topology ATP generation and biomass growth cannot be fully uncoupled.

### 3.11.3 MPS for Validating \(\text{iSfu}648\): Closed-Network ATP Cycles

Thermodynamic feasibility constraints dictate that solutions containing closed cycles (e.g., \(A \rightarrow B \rightarrow C \rightarrow A\)) should be infeasible. In the absence of complete thermodynamic information, such cycles remain possible. MPS was used to identify qualitative reaction direction constraints which would eliminate closed-cycles which synthesize ATP, while simultaneously ensuring biomass growth:

\[
\begin{align*}
\text{max} & \quad \sum_{j,n} y_{j,n} + \sum_{j,n} z_{j,n} \\
\text{s.t.} & \quad \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \\
& \quad \text{no uptake fluxes} \\
& \quad \text{enforce probabilistic constraints, (S44) and (S45)} \quad \forall j \in J \\
& \quad v_{ATPM} = 0 \\
\end{align*}
\]

\[
\begin{align*}
\text{max} & \quad v_{BM} \\
\end{align*}
\]
\[ s.t. \quad FBA \text{ constraints, (S13) and (S14)} \quad \forall j \in J \]
\[ \text{growth in monoculture (or coculture)} \]
\[ \text{enforce probabilistic constraints, (S44) and (S45)} \forall j \in J \]
\[ v_{BM} \geq 0.01 \]
\[ \text{allowed probabilistic constraints, (S32) to (S34)} \quad \forall n \in N \]
\[ \text{consistency of probabilistic constraints, (S38)} \quad \forall n \in N \]

MPS identified a set of two reactions to be constrained to a single direction during monoculture growth, and one direction to be constrained during coculture growth (listed in Table S2 in S1 Dataset of the Supporting Material).

MPS was used also to identify a single set qualitative reaction direction constraints which would ensure growth in both monoculture and coculture simultaneously (formulation not shown). That problem was infeasible, indicating that distinct sets of qualitative reaction direction (probabilistic) constraints are required under the two conditions.

### 3.11.4 MPS for Validating \( i \)Sfu648: ATP Gain

*In-silico* simulations identified the theoretical maximum ATP yield for each growth mode of *S. fumaroxidans*. However, the \( i \)Sfu648 model enabled other mechanisms of ATP generation with a higher yield. MPS was used to identify qualitative reaction direction constraints which would eliminate these extra mechanisms, while simultaneously ensuring biomass growth:

\[
\begin{align*}
\text{max} & \quad \sum_j y_j + \sum_j z_j \\
\text{s.t.} & \quad \text{max ATP gain} \\
\text{s.t.} & \quad \text{FBA constraints, (S13) and (S14)} \quad \forall j \in J \\
& \quad \text{appropriate growth mode uptake constraints} \\
& \quad \text{enforce probabilistic constraints, (S44) and (S45)} \forall j \in J \\
& \quad v_{BM} \geq 0.01 \\
& \quad \text{allowed probabilistic constraints, (S32) to (S34)} \quad \forall n \in N \\
& \quad \text{consistency of probabilistic constraints, (S38)} \quad \forall n \in N
\end{align*}
\]

(MPS)

MPS identified a set of 38 reactions to be constrained to a single direction (listed in Table S2 in S1 Dataset of the Supporting Material).

### 3.12 Simulation Conditions

Simulations were performed for both individual models (\( i \)Sfu648 and \( i \)Mhu428) and the coculture model. Uptake fluxes for carbon and other nutrients are given in Table S1 in S1 Dataset of the Supporting Material. All simulations were performed using CPLEX 12 (IBM, Armonk, NY) accessed via the General Algebraic Modeling System, Version 23.9.5 (GAMS, GAMS Development Corporation, Washington, DC).
Growth rate and product stoichiometry of *S. fumaroxidans* in monoculture on fumarate with propionate. The fumarate to propionate uptake ratio was increased from one to three (the experimentally observed value) at intervals of 0.1. Simulations maximized growth rate, then minimized total enzyme cost (pTMFA). The experimentally observed product yields are three mol succinate per mol propionate; one mol acetate per mol propionate and one mol CO$_2$ per mol propionate.
Metabolite exchange in coculture simulations. (A) Allowed carbon exchange fluxes. Propionate is fed into the system, which *S. fumaroxidans* converts into acetate, and CO\(_2\) and H\(_2\) or formate. These products are secreted into the media, where they can be consumed by *M. hungatei*. *M. hungatei* converts acetate to biomass, and CO\(_2\), H\(_2\), and formate to CH\(_4\). *M. hungatei* can also optionally interconvert excess CO\(_2\) and formate via a formate-hydrogen lyase. (B) Mass balance around the coculture system. In this system, both species grow at the same rate (equal to the dilution rate) and are present in the reactor at biomass concentration \(X\). They exchange metabolites (circles) into a shared environment with flux \(v\). The rate \(F\) represents the net flux into / out of the system, and is related to \(v\) and \(X\) for each species via the given equation. The medium components (cells and metabolites) are subject to mass balance constraints, which relate each species’ biomass concentration \(X\), individual species uptake/secretion fluxes \(v\), and the net flux into or out of the reactor \(F\).
5  Supporting References


