Supplementary Information

Table of Contents

1. Oxygen Limited Chemostat

Using FBA^{ME}, we have simulated the oxygen limited chemostat growth of *E. coli* by fixing the growth rate at various dilution rates (0.1 – 0.5 h^{-1}). The C* for the cytochromes under microaerobic conditions are calculated using Equations 5 and 10 from the main manuscript and measured Ks values for the cytochromes (Bekker *et al.* 2009). In all conditions studied, Cyd-I is preferred at low oxygen concentrations, while Cyo is preferred at high oxygen concentrations (Figure SI_1). This agrees with the findings of a similar oxygen-limited chemostat experiment with *E. coli* strain MC4100 (Figure SI_1, Tseng et al. 1996). The model predicts that Cyd-II will be utilized at higher dilution rate under microaerobic conditions. Although we do not have the experimental data for the expression of Cyd-II, it is clear that our simulation agrees with experimental and genetic evidence that Cyd-I is the designed for the microaerobic environment, whereas Cyo is designed for the aerobic environment.

2. Model Simplifications and Consequences

There are three major simplifications in the FBA^{ME}. First, we assumed that the S/V ratio of *E. coil* remains constant. It has been observed that *E. coli* cells are bigger under higher growth rates, possibly due to osmotic pressure, leading to a lower S/V ratio. This implies that at higher growth rates, the competition between membrane enzymes is more intense. The direct consequence of this is that respiro-fermentation commences at slightly lower glucose uptake rate than our prediction, as confirmed by experimental observations (Fig 2).

Second, we did not include explicit membrane costs for electron transport proteins other than cytochromes. The membrane-bound electron transport chain also contains succinate dehydrogenase, NADH dehydrogenases, and ATP synthase; unfortunately, the membrane costs of these proteins could not be obtained due to the lack of experimental data and is incorporated implicitly based on the assumption that their membrane costs will be proportional to the flux through the cytochrome oxidases or the electron transport chain. Hence, including the impact of other electron transport proteins is not expected to lead to drastic changes in the model predictions. **However, the high degree of agreement between experimental data and model prediction despite these caveats is perhaps an indication of the fundamental nature of the membrane economics constraint.**

Finally, we have assumed that the transport of fermentation products out of the cell occupies negligible membrane area. The primary reason for this simplification is that we were unable to ascertain the membrane cost of acetate transport. The only known *E. coli* acetate transporter is a permease coded by ActP (EcoCyc). Since this protein was demonstrated to be responsible for the uptake of acetate into the cell (Gimenez 2003), acetate excretion mechanism remains unclear. It is also possible that fermentation

products such as acetate are transported out of the cell through simple diffusion, a process that depends on the pH and the concentration gradient across the membrane (Gimenez 2003, Butler 2004).

Nonetheless, we simulated the glucose-limited chemostat growth of *E. coli* using a FBA^{ME} model with assumed membrane cost for acetate transport. Our simulations suggest that as long as the acetate transport cost is below 7% of the glucose transport cost, the effect of its inclusion is negligible when the GUR is less than the experimentally measured maximum of 10.7 mmol/gdw/h (Fig SI-5). If we assume that acetate transporter for secretion is monomeric (that is the case for ActP), then its membrane cost should be much smaller than the large glucose transport systems (PtsG has 6 subunits per complex, PtsX has 7 subunits per complex) and the very large (22 subunits) ATP synthase required by respiration. For these reasons, neglecting the transport cost of acetate is justified.

3. Model Robustness

In order to analyze the robustness of the FBA^{ME} model with respect to C* values, we have performed 50 simulations with C* values that are randomly generated between 75% and 125% of its original value. The +/- 25% variation is sufficiently large since C* value of Cyo is only 150% times of the C* value of Cyd-I, and many independent experiments have clearly demonstrated that $C^*_{cyo} > C^*_{cyd\text{-}I} > C^*_{cyd\text{-}II}$. (For example, Tseng *et al.* 1996, Bekker *et al*, 2009). The 50 simulations showed that the system is robust against up to 25% perturbation in C* values (Figure SI-2). The general trend behind the utilization of cytochromes remains the same - Cyo for low glucose uptake rate, Cyd-II for high glucose uptake rate, and Cyd-I for microaerobic conditions. The critical glucose uptake rates at which alternative metabolic strategies are adopted show minor variations with varying C* values.

4. Comparison with C-13 Flux Analysis

We compared predictions from FBA and FBA^{ME} with the C-13 flux analysis of glucose-limited chemostat growth performed by Vazquez *et al.* (2008). As illustrated in Figure SI-3, although both FBA^{ME} and FBA predict the repression of the TCA cycle flux at higher growth rate, the predictions from FBA^{ME} are closer to experimental observations. As demonstrated below, FBA with molecular crowding could not predict TCA cycle repression.

5. Simulations using FBA with Molecular Crowding Constraint

Previously, it has been claimed that FBA with molecular crowding constraint (FBAwMC) is capable of predicting the acetate overflow in *E. coli* (Vazquez *et al.* 2008).

The FBAwMC theory suggests that fluxes of all metabolic reactions are constrained by the limited intracellular volume. We evaluated the impact of this constraint by introducing the molecular crowding constraint to the iAF1260 model (Feist 2007). (Note: Vazquez *et al.* used an older model of *E. coli*). We tuned the average crowding coefficient to achieve a growth rate of 0.68. While the FBAwMC was able to predict acetate overflow, it cannot predict growth yield nor rate (Figure SI-4). More importantly, FBAwMC **cannot predict** the acetate overflow if we set the crowding coefficient of the ETC enzymes or the crowding coefficient of all membrane proteins (glucose transporters and the ETC enzymes) to zero. For rod-shaped prokaryotic cells such as *E. coli*, the volume of the membrane is negligible compared to the total volume of the cell; therefore, if the acetate overflow is caused by a volumetric crowding, FBAwMC should be able to predict the respiro-fermentation phenomenon without the inclusion of the membrane proteins into its formulation. The inability of FBAwMC to predict acetate overflow without the inclusion of membrane proteins directly suggests that the available membrane area is the true constraint on cellular metabolism.

6. Construction of FBA^{ME} model

The FBA^{ME} model of *E. coli* is adapted from that of iAF1260 model (Feist 2007) with the following modifications. The iAF1260 model contains multiple version of NADH dehydrogenases (using both menaquinol and ubiquinol) and both Cyo and Cyd-I. It does not contain Cyd-II.

In the FBA^{ME} model, the NADH dehydrogenases and cytochrome oxidases contained in the original iAF1260 are removed, and replaced with reactions NDH-I, NDH-II, CYO, CYD-I, and CYD-II. Only two NADH dehydrogenases are included and are assumed to only use ubiquinol (q8 and q8h2 metabolites in the model), which simplifies the analysis of the ETC system. The glucose exchange reaction, "EX_glc(e)" in the iAF1260 model (representing all glucose transporters) is replaced with a new version of "EX_glc(e)" reaction. The new cytochrome oxidase reactions (CYO, CYD-I, CYD-II) and the new "EX_glc(e)" reaction each contains a pseudo-metabolite called "CMC" which stands for "cytoplasmic membrane cost". This pseudo-metabolite is used to represent the usage of the cytoplasmic membrane area by transmembrane-proteins. A "cytoplasmic membrane demand" reaction (DM_CMC) is also added; the upper bound of this reaction is set to the relative cytoplasmic budget of 1. Together, the CMC pseudo-metabolite and the DM_CMC reaction allow us to implement the membrane budget constraint (Equation SI-1, Equation 11 in main manuscript) using COBRA Toolbox.

$$\sum v_i \, C_i^* \leq 1 \qquad \text{(SI-1)}$$

The resultant model as well as the script used to create it are available for download as supplementary material accompanying this article and can be simulated using COBRA Toolbox.

7. Calculating the Relative Membrane Cost in *E. coli*

Membrane Constraint Binding Requirement during Parameter Estimation
Because the membrane occupancy constraint (SI-1) is the only constraint where the cost parameters appears in our model, this constraint must be binding ($C_i v_i = 1$) for us to acquire unique cost parameters. This implies that only data generated under experimental conditions where the membrane occupancy constraint is the limiting constraint can be used to determine the membrane cost values.

Mathematically, the FBAME can be written as:

maximize growth rate flux
such that

$$\sum Sv = 0 \qquad \text{(R1)}$$

$$\sum v_i \, C_i^* \leq 1 \qquad \text{(R2)}$$

$$v_{glc} \leq v_{glc}^{max} \qquad \text{(R3)}$$

Here, R1 represents the standard pseudo-steady-state assumption used in all FBA models – it provides no information regarding the reaction fluxes (See Feist 2007 for more detail). R2 is the membrane constraint, and R3 is the glucose uptake constraint. R3 is used when simulating the glucose-limited growth in a chemostat, and is not used when simulating batch growth where the glucose uptake rate is not limiting.

When R3 is not used (eg. when we're calibrating for glucose transporter cost), the binding constraint is automatically satisfied in the optimal solution because R2 is the only constraint that limits the flux through the metabolic network. However, when both R2 and R3 are used (eg. when we're calibrating for cytochrome costs), if the glucose availability is very low, the flux through the network may be limited by R3 instead of R2 – in such cases, the binding requirement is not satisfied.

To ensure that the binding requirement is satisfied, we have chosen data generated from batch experiments and from chemostat experiments with sufficiently high glucose uptake rates. Our simulations also confirms that the membrane constraint binding requirement is satisfied for all cases used for parameter estimations.

Relative Cost of C^*_{GUR}
The maximal anaerobic glucose uptake rate of *E. coli* was measured to be 18 mmol/gdw/h (Portnoy 2008) under glucose-excess conditions. We can assume that cytochrome oxidases are not present under anaerobic condition (Tseng *et al.* 1996), and that glucose transporters are the sole occupants of the cytoplasmic membrane (for the purpose of our simplified model). Assuming the cell optimizes its growth rate, we can vary the C^*_{GUR} parameter until the FBA^{ME} model (which contains the membrane budget constraint (Eq SI-1) by definition) correctly predicts the measured glucose uptake rate of 18 mmol/gdw/h. Note: the glucose uptake reaction is left unconstrained to reflect the

glucose-excess condition. At C^*_{GUR}=0.0556 g h/mmol, FBA^{ME} predicted an oxygen uptake rate (OUR) of 18 mmol/gdw/hr and a growth rate of 0.44 hr^{-1}. In this case, the membrane is predicted to be saturated: $\sum v_i C^*_i = 0.0556 \times 18 = 1$.

Relative Cost of C^*_{CYO}

The relative costs of cytochrome oxidases are calculated similarly using experimental data from Bekker *et al.* (2009). In this study, the cytochrome knockouts are grown with a fixed growth rate of 0.15±0.1 hr^{-1} in glucose-limited chemostats. For the Cyo strain (Δcyd1Δcyd2), the measured oxygen uptake rate was 6.4±0.4 mmol/gdw/h, and the measured glucose uptake rate was 2.3 mmol/gdw/h.

To determine the Cyo cost, a glucose uptake constraint of 2.3 mmol/gdw/h is used to reflect the glucose-limiting growth condition of this experiment. Since the C^*_{GUR} is known (from above calculation), and only Cyo exists in this strain, only the C^*_{CYO} parameter is both unknown and relevant. The Cyd-I and Cyd-II reactions are deleted during this process. We can determine the value of C^*_{CYO} by manually varying it until the model correctly predicts both the growth rate and the oxygen uptake rate. At C^*_{CYO} = 0.0658 g h/mmol, FBA^{ME} predicted that the Cyo strain would have a growth rate of 0.15 hr^{-1} and an oxygen uptake rate of 6.7 mmol/gdw/hr. 6.7 is within one standard deviation of the measured average oxygen uptake rate, which is 6.4±0.4 mmol/gdw/h. In this case, the membrane is predicted to be saturated: $\sum v_i C^*_i = 0.0658 \times 6.7 \times 2 + 0.0556 \times 2.3 = 1$.

Relative Cost of $C^*_{CYD\text{-}I}$

For the Cyd-I strain (ΔcyoΔcyd2), the measured oxygen uptake rate was 8.8±1.3 mmol/gdw/h, the measured glucose uptake rate was 3.2 mmol/gdw/h, and the growth rate was again fixed at 0.15±0.1 hr^{-1}.

$C^*_{CYD\text{-}I}$ is similarly determined using FBA^{ME} by constraining the model to the measured glucose uptake rate of 3.2 mmol/gdw/h, and varying the $C^*_{CYD\text{-}I}$ parameter until the model correctly predicts both the oxygen uptake rate and the growth rate. The Cyd-II and Cyo reactions are deleted during this process. Since $C^*_{CYD\text{-}I}$ is usually not expressed under oxygen-rich conditions, the assumption here is that the cell is able to optimally manage $C^*_{CYD\text{-}I}$ expression when both Cyo and Cyd-II are knocked out. At C^*_{CYO} = 0.0427 g h/mmol, FBA^{ME} predicted that the Cyd-I strain would have a growth rate of 0.15 hr^{-1} and an oxygen uptake rate of 9.62 mmol/gdw/hr. 9.62 is within one standard deviation of the measured oxygen uptake rate, which is 8.8±1.3 mmol/gdw/h. In this case, the membrane is predicted to be saturated: $\sum v_i C^*_i = 0.0427 \times 9.62 \times 2 + 0.0556 \times 3.2 = 1$.

Relative Cost of $C^*_{CYD\text{-}II}$

For the Cyd-II strain (ΔcyoΔcyd1), the measured oxygen uptake rate was 11.1±0.6 mmol/gdw/h, the measured glucose uptake rate was 6 mmol/gdw/h, and the growth rate was again fixed at 0.15±0.1 hr^{-1}. Unlike the mutant containing Cyo only, which grew at a

very low glucose uptake rate, the knockout containing Cyd-II only grew at a high glucose uptake rate of 6 mmol/gdw/h and consumes oxygen at a rate of 11 mmol/gdw/h. At this glucose uptake rate, the "membrane economics theory" predicts that the cell membrane of wild-type *E. coli* is shared by Cyo and Cyd-II. Assuming the genetic programming remain functional in the knockout strain (since the cells were not given time to evolve in the experiment), and expression level of Cyd-II is expected to be the same in the knockout strain as in the wild-type growing at a glucose-uptake rate of 6 mmol/gdw/h.

To determine the Cyd-II cost, a glucose uptake constraint of 6 mmol/gdw/h is used to reflect the glucose-limiting growth condition of this experiment. Both Cyo and Cyd-II reactions are left unconstrained, $C^*_{CYD\text{-}II}$ value is manually varied until 11±0.6 mmol/gdw/h of oxygen is processed through the Cyd-II reaction. At $C^*_{CYD\text{-}II}$ = 0.0128 g h/mmol, FBA^{ME} predicted that the wildtype would have an oxygen uptake rate of 14.38 mmol/gdw/h. The flux through Cyo is 5.63 mmol/gdw/h, the flux through Cyd-I is 0, and the flux through Cyd-II is 23.12 mmol/gdw/h (11.55 mmol/gdw/h of oxygen, which falls within 11.1±0.6 mmol/gdw/h). In this case, the membrane is predicted to be saturated: $\sum v_i C^*_i = 5.63\times0.0658 + 23.12\times0.0128 + 6\times0.0556 = 1$

8. Estimation of Membrane Protein Content

We were able to estimate the cellular content of cytochrome oxidases and ATP synthase under several conditions based on experimental measurements and numerical data from BioNumbers database. In this section, BNID refers to the entry ID number in BioNumbers database.

8.1 Number of Polypeptides in *E. coli* Cytoplasmic Membrane

The average cell mass of *E. coli* is 0.5 pg/cell (Philips and Milo, 2009), and about 50% of the cell mass is protein (Feist *et al.* 2007). This means that the average protein content of *E. coli* is 0.25 pg/cell, which is similar to the measured values (BNID: 104878, 104879, 104880). The average size of the polypeptides (a monomeric protein) is 40 kDa in *E. coli* (BNID: 105861) or $6.42\text{x}10^{-20}$ g. This means that the total number of polypeptides in *E. coli* is about 3,900,000. About 18% of all polypeptides in *E. coli* is located in the cytoplasmic membrane (BNID: 100019), therefore, the number of polypeptides in the cytoplasmic membrane is about 700,000.

8.2 Cytochrome Content at Maximum Oxygen Uptake

The maximum oxygen uptake rate (OUR) is 18 mmol/gdw/h in *E. coli* (Feist *et al.* 2007). The kcat of Cyd-II is 70 O_2/enzyme/s, or 252000 O_2/enzyme/h. Assuming only Cyd-II is used at the maximum OUR, then 54000 enzymes are needed. Since Cyd-II is dimeric, the number of polypeptides required is 108000, which is about 15% of all membrane polypetides.

8.3 Cytochrome Content at Microaerobic Condition

Under microaerobic conditions, the expression level of Cyd-I can reach above 250 nmol/g protein (Alexeeva *et al.* 2000). Assume *E. coli* protein content is 0.25pg/cell, each cell contains about 38000 Cyd-I enzymes, or 76000 Cyd-I polypeptides, nearly 11% of the all membrane proteins.

8.4 Glucose Transporter Content

The maximum glucose uptake rate (GUR) is 10.7 mmol/gdw/h in *E. coli* (Feist *et al.* 2007). The Kcat of the PtsG transport system is 756000 Glucose/enzyme/h. Assuming only the PtsG system is used at the maximum GUR, each cell requires 4260 enzymes or 25560 polypeptides (PtsG contains six subunits). This accounts for 4% of all membrane polypeptides. Philips and Milo (2009) has estimated that PtsG accounts for more than 4% of the membrane surface, or 8% of the available membrane surface. The difference between these two estimates suggests that PtsG takes up more membrane area per polypeptide, which is understandable as oligomers are more complex structurally than monomers.

8.3 ATP Synthase Content at Optimal Growths

FBA predicts that the flux through ATP synthase reaction during optimal *E. coli* growth is 52 mmol/gdw/h. The kcat of ATP synthase is 1260000 ATP/enzyme/h. This means each cell requires 4141 ATP synthase. Since each ATP synthase contains 22 polypeptides, about 13% of all membrane polypeptides are ATP synthase.

9. Evidence for Decrease in Respiratory Efficiency

Past theories have attributed the observed tradeoff between the ATP yield and the catabolic rate solely to the utilization of fermentative pathways, resulting in lower ATP per glucose metabolized than oxidative phosphorylation. However, chemostat experiments suggest that the decrease in the respiratory efficiency itself may contribute to the overall yield decrease.

A statistically significant ($p = 0.08$) decrease in biomass yield was observed between $D=0.3\ h^{-1}$ and $D=0.4\ h^{-1}$, with the acetate secretion commencing at $D = 0.4\ h^{-1}$ at a very low level (Figure SI-6). The decrease in biomass yield from $D = 0.3\ h^{-1}$ to $D = 0.4\ h^{-1}$ (0.014 g/mmol based on average yields) is much larger than the yield decrease due to acetate secretion (0.0039 gdw/mmol, based on stoichiometric calculation using FBA), presenting evidence that the reduction in yield is a consequence of decreased respiratory efficiency rather than mere redirection of carbon from biomass to acetate.

FBA^{ME} provides mechanistic insight into the decrease in respiratory efficiency. As the membrane became saturated with glucose transporters at high rate of glucose uptake, the costly Cyo was replaced by the cheaper but less efficient Cyd-II, resulting in an increased ATP production **rate** at the expense ATP **yield**. Therefore, the outcome of the competition for space on the membrane between Cyo and Cyd-II appears to play a pivotal

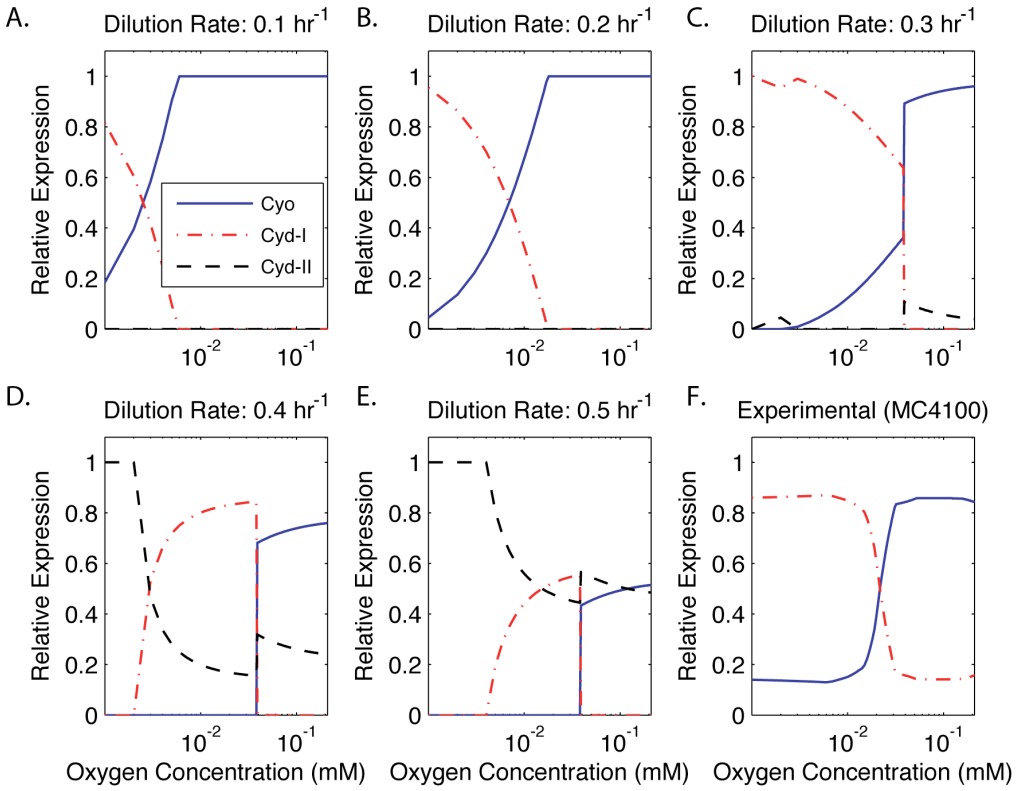
role in deciding the trade-off between the rate and yield of ATP production. This is evident from gene expression data (Figure 3A, B from main manuscript).

10. Model Files

Two versions of the *E. coli* FBA^{ME} model files can be found as a part of the Supplementary Information – a MATLAB version with the suffix ".MAT" and a SBML version with the suffix ".XML". The MATLAB version requires the COBRA Toolbox to load and run. The XML version is generated from the MAT version using COBRA Toolbox's writeCbtoSbml function. COBRA Toolbox can be downloaded from http://opencobra.sourceforge.net/

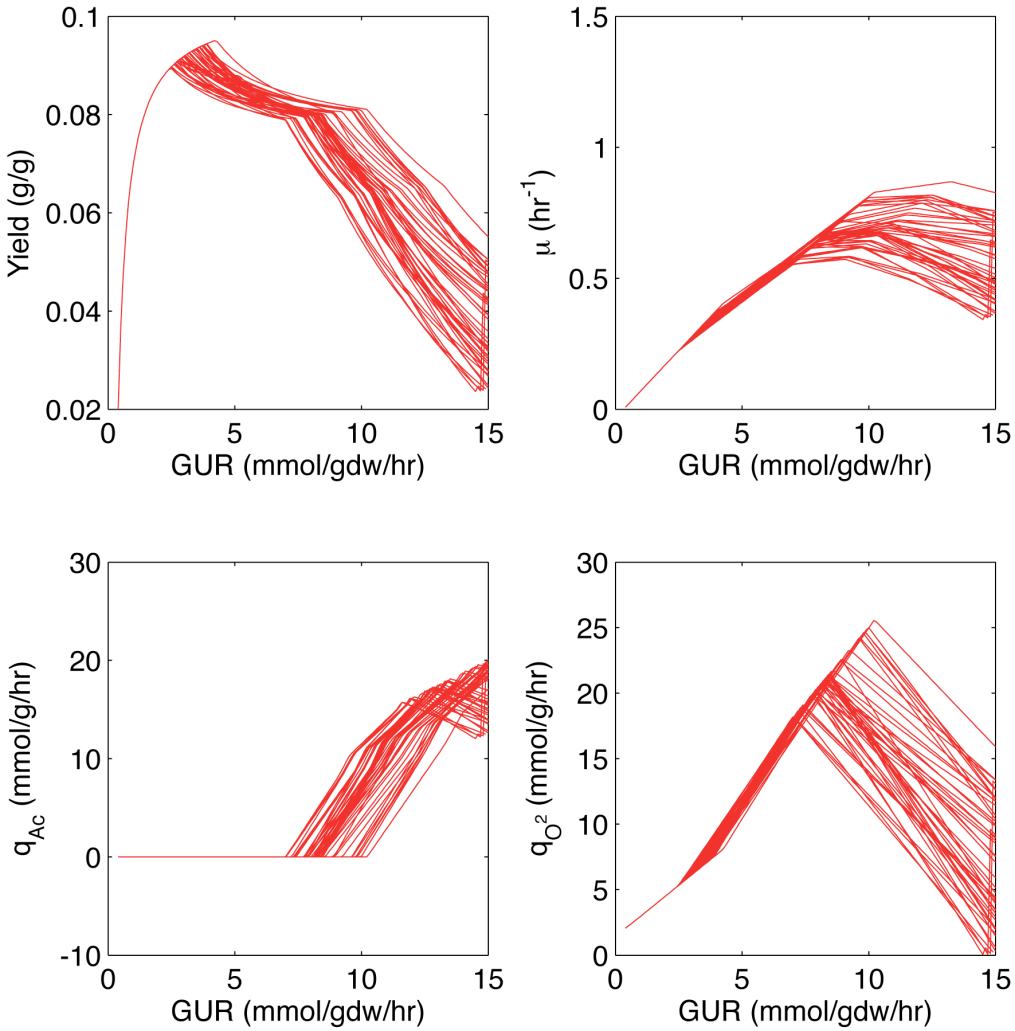
11. Figures

Figure SI-1.



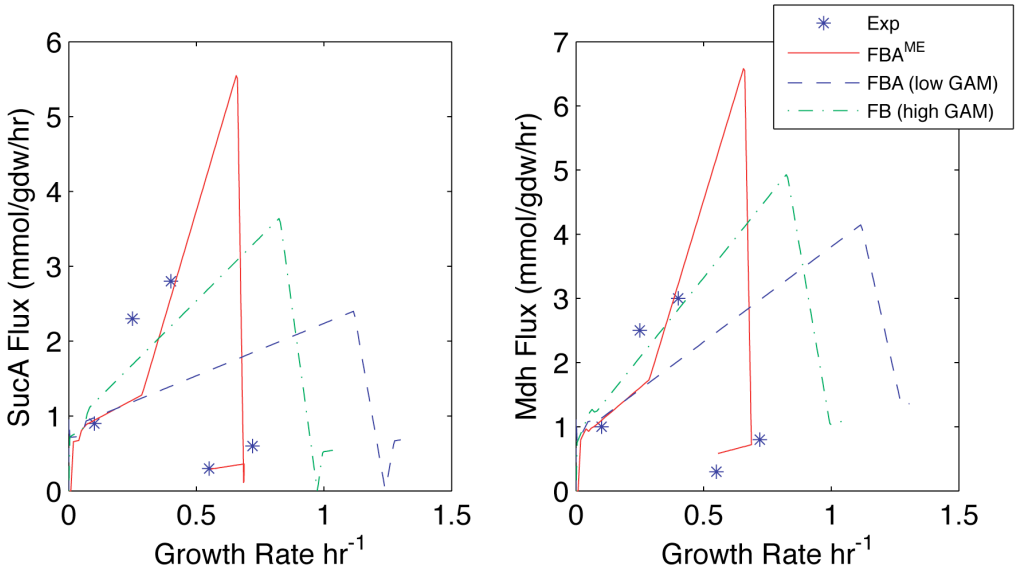
Predicted cytochrome expression of *E. coli* grown in oxygen-limited chemostat at different dilution rates (A-E) compared to the measured cytochrome expression of *E. coli* grown in oxygen-limited chemostat (F). Cyd-II measurement is not available in the experimental data.

Figure SI-2.



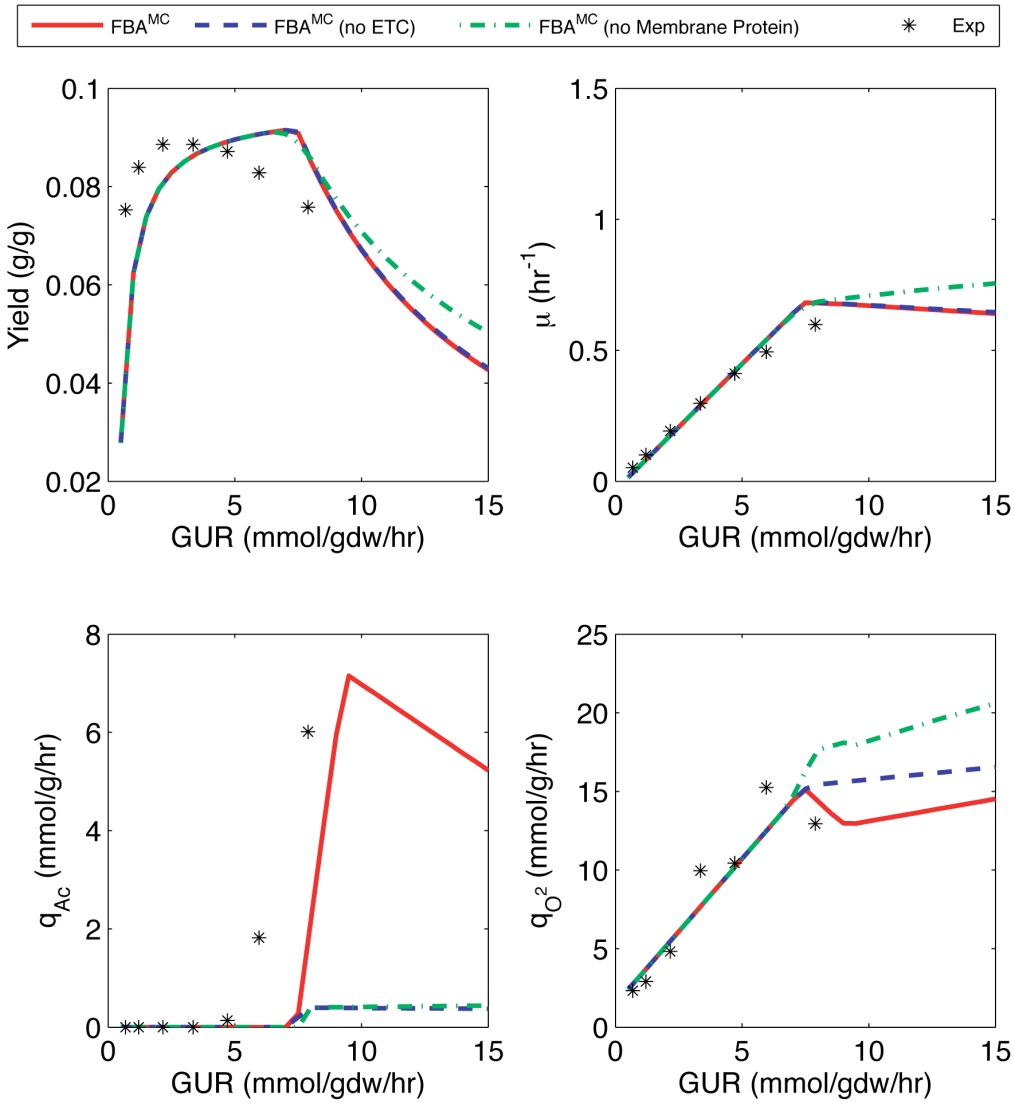
Fifty simulations were performed with C* values that are randomly generated between 75% and 125% of its original value. The result shows that the model is robust against parameter variations.

Figure SI-3.



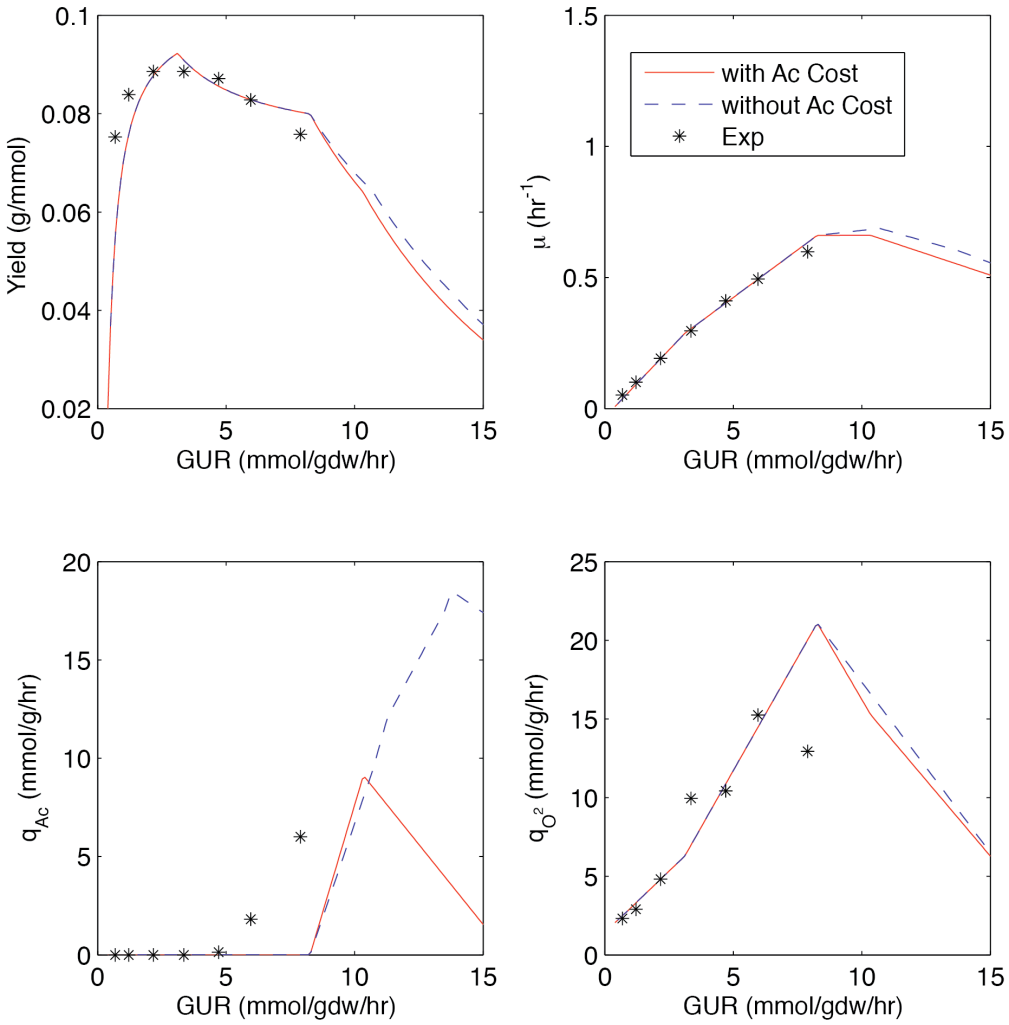
Predicted fluxes through the TCA cycle compared with C-13 flux analysis data.

Figure SI-4.



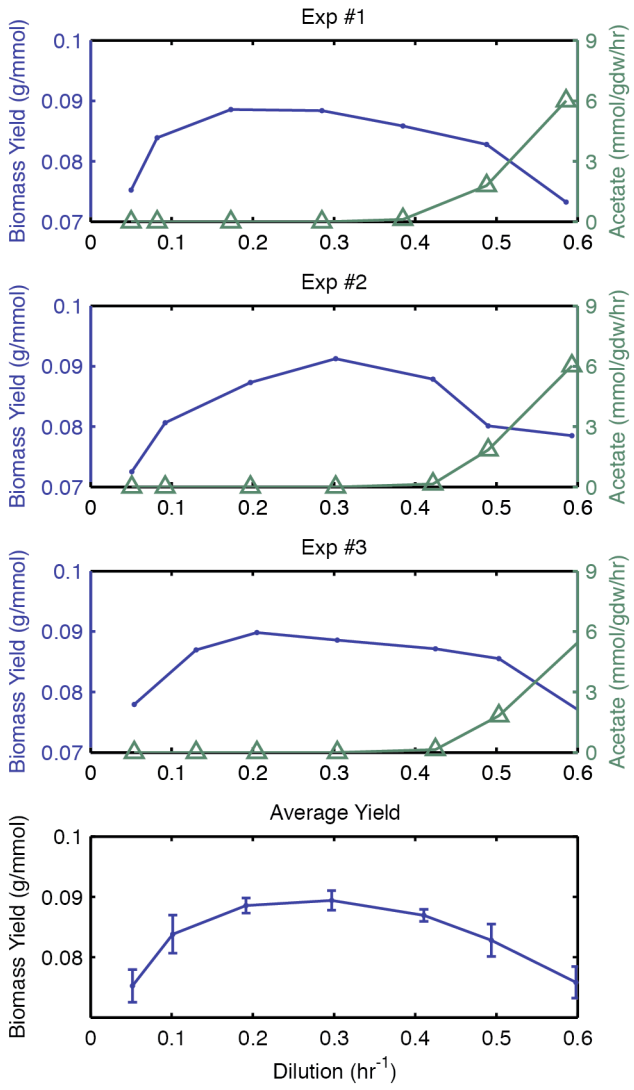
Intracellular volume constraint cannot explain the respire-fermentation phenomenon. FBA framework based on such a constraint (FBAwMC) **cannot predict** the fermentative production of acetate if we set the crowding coefficient of the ETC enzymes or the crowding coefficient of all membrane proteins (transporters and ETC enzymes) to zero. Since the membrane proteins takes up negligible intracellular volume, this indirectly suggests that the available membrane area is the true constraint on cellular metabolism.

Figure SI-5.



The solid red line represents the chemostat simulation with an acetate transport cost that is 7% of the cost of glucose transport. The blue dash line represents the original chemostat simulation without acetate transport cost. The inclusion of the acetate transport cost has no significant effect as long as it is less than 7% of the glucose transport cost.

Figure SI-6.



Triplicate glucose-limited chemostat experiments of *E. coli* growth at various dilution rates (Vemuri et al., 2006).

12. References

Vazquez A, Beg QK, deMenezes MA & Ernst J (2008) Impact of the solvent capacity constraint on *E. coli* metabolism. *BMC Systems Biology* 2: 7.

Feist AM, Henry CS, Reed JL & Krummenacker M (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3: 121.

Portnoy VA, Herrgård M, Palsson B (2008) Aerobic fermentation of D-glucose by an evolved cytochrome oxidase-deficient *Escherichia coli* strain. *Applied and Environmental Microbiology* **74:**7561-9

Bekker M, de Vries S, Ter Beek A, Hellingwerf KJ, de Mattos MJ (2009) Respiration of *Escherichia coli* can be fully uncoupled via the nonelectrogenic terminal cytochrome bd-II oxidase. *J Bacteriol* **191:**5510-5517.

Tseng CP, Albrecht J, Gunsalus RP (1996) Effect of microaerophilic cell growth conditions on expression of the aerobic (cyoABCDE and cydAB) and anaerobic (narGHJI, frdABCD, and dmsABC) respiratory pathway genes in *Escherichia coli*. *J Bacteriol* **178:**1094-1098

Gimenez R, Nunez MF, Badia J, Aguilar Juan, Baldoma L. The Gene yjcG, Cotranscribed with the Gene acs, Encodes an Acetate Permease in *Escherichia coli*. J Bacteriol 185:6448-6455

Philips R, Milo R (2009) A feeling for the numbers in biology. *PNAS* 106: 21465–21471.

Butler T, Lee SG, Wong WW, Fung E, Connor M, Liao JC. (2004) Design of artificial cell–cell communication using gene and metabolic networks. *PNAS* **101**:2299-2304.

Alexeeva S, De Kort B, Sawers G, Hellingwerf KJ, De Mattos MJT (2000) Effects of Limited Aeration and of the ArcAB System on Intermediary Pyruvate Catabolism in *Escherichia coli*. *J Bacteriol* 182: 4934-4940.