

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

The Study of Overflow Metabolism and kinetics in a model of metabolism and gene expression of Escherichia coli

Permalink

<https://escholarship.org/uc/item/66x4v685>

Author

Cheng, Chuankai

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

The Study of Overflow Metabolism and kinetics in a model of metabolism
and gene expression of *Escherichia coli*

A thesis submitted in partial satisfaction of the requirements of the degree
Master of Science

in

Bioengineering

by

Chuankai Cheng

Committee in charge:

Professor Bernhard O. Palsson, Chair
Professor Nathan Lewis
Professor Prashant Mali

2017

Copyright
Chuankai Cheng, 2017
All rights reserved.

The Thesis of Chuankai Cheng is approved and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego
2017

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	v
List of Tables	vii
Abstract of the Thesis	viii
Introduction	1
Chapter 1: Overflow metabolism	2
1.1 Definition	2
1.2 Explanations	3
Chapter 2: Single cell metabolic and gene expression model	5
2.1 Unicellular growth and cell components dilution	6
2.2 The coupling constraint: proteome efficiency	8
2.3 Small ME model with alternative pathway	9
2.4 Genome-scale models of metabolism and macromolecular expression	15
Chapter 3: Results of Simulation	17
3.1 Small ME model	18
3.2 Original genome-Scale model and its adjustment	22
3.3 Parameter sweep on adjusted genome-scale model	27
Chapter 4: Discussions	31
4.1 Function of motility protein	31
4.2 k_{eff} varied by growth rate	32
4.3 Limited Oxygen Uptake	33
Chapter 5: Conclusions	35
References	36

LIST OF FIGURES

Figure 1: Scheme of the phenomenon of overflow metabolism	2
Figure 2: Overflow metabolism based on pathway switch	3
Figure 3: Comparison between the derivation in <i>Basan et al.</i> and optimization upon flux balance analysis	6
Figure 4: Enzymatic/proteomic efficiency	8
Figure 5: Small scale model with alternative pathways	10
Figure 6: Experiment data from <i>Basan et al.</i>	17
Figure 7: Comparison of the results from Small ME model and from <i>Basan et al.</i> ..	18
Figure 8: Acetate lines on multiple k_{eff} s of respiration pathway(left). Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of respiration(right)	19
Figure 9: Acetate lines on multiple k_{eff} s of fermentation pathway(left). Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of fermentation(right)	20
Figure 10: Acetate lines on multiple k_{eff} s of biomass synthesis pathway(left). Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of biomass synthesis(right)	21
Figure 11: The problem of original genome-scale ME model: no fermentation ...	22
Figure 12: The approach of how the model in <i>Basan et al.</i> is mapped in Genome-scale	23
Figure 13: Adjustment of k_{eff} in TCA cycle from the genome-scale model	24
Figure 14: After k_{eff} adjustments: fermentation appears	25
Figure 15: Comparison of the simulation results of pre and post-adjustment, as well as the k_{eff} s in TCA cycle are 95% off	26
Figure 16: Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of TCA cycle, the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines ...	27
Figure 17: Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of reaction ACKr, the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines	28
Figure 18: Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of ATP synthase(left) and cytochrome oxidase(right), the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines	29
Figure 19: Dependence of $S_{ac}, \mu_{threshold}$ and μ_{max} on k_{eff} of ribosome efficiency(left) and global k_{eff} s(right), the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines	29

Figure 20: 50% ATP synthase k_{eff} , 5% TCA1 k_{eff} , 0.5% TCA2 k_{eff} , 0.25% CYT oxidase k_{eff} , 200% ACr k_{eff} , 180% global k_{eff} 30

Figure 21: Good fit for the simulation on the strains which are not able to synthesize motility protein 31

Figure 22: Matching the experiment data on multiple k_{eff} s of cytochrome oxidase 32

Figure 23: A Sigmoid function fit of growth dependence on cytochrome oxidase k_{eff} s, as a fitting function of $k_{eff}(\mu)$ 33

Figure 24: Simulation with oxygen uptake rate bound limited to $14mM/gDW/h$ 34

LIST OF TABLES

Table 1: Comparison between the Small scale model(both <i>Basan et al.</i> and ME) and genome-scale ME model.....	15
Table 2: Adjustment of k_{eff} values in TCA cycle	24

ABSTRACT OF THE THESIS

The Study of Overflow Metabolism and kinetics in a model of metabolism and gene expression of *Escherichia coli*

by

Chuankai Cheng

Master of Science

Universeity of California, San Diego, 2017

Professor Bernhard O. Palsson, Chair

Overflow metabolism can be explained by a pathway switch behavior in the cell based on the tradeoffs between metabolic efficiency and proteomic efficiency(enzymatic kinetics). This thesis simulates this behavior on the both small-scale and genome-scale model. By sweeping the enzyme kinetics parameters on multiple pathways or reactions, the simulation could potentially approach the experimental results, however, would converge to some bounds beyond. Therefore, hypotheses of some extra constraints, such as the limitation of oxygen uptake ,electron transport chain, as well as the function of motility protein are presented.

Introduction

Fermentation, a process of generating less amount of energy(ATP) comparing to respiration, excreting acetate or(and) lactate, is a typical anaerobic behavior of organism for bacteria and nearly all types of organisms[1]. It is easy to explain this behavior in anaerobic condition since the absence of oxygen would cause the respiration being blocked. However, previous studies show that the fermentation process could also be taken place with the presence of oxygen[2]. The seems-to-be wasteful occurrence of fermentation in this circumstances is called overflow metabolism and is considered as the same phenomenon as Warburg effect in cancer cells.

Quantitative biology is a kind of approach that quantifies biological behaviors, such as overflow metabolism in this thesis, then build up some computation models in order to explain the phenomenon.

Cell metabolic network is a model that could be run to simulate the cell behavior such as cellular growth, synthesis of species as well as the excretion of metabolites, etc. This thesis is aiming to take this approach and simulate the behavior of overflow metabolism.

In this thesis, chapter 1 is giving a detailed explanation of overflow metabolism. Chapter 2 is going to focus on how the computational models are built. Chapter 3 will show the results of simulations. By comparing the results of simulations to the experiment data, chapter 4 is going to discuss the differences we met in the comparison and put forward some hypothesis of what caused the difference. Chapter 5 concludes the whole project.

Chapter 1: Overflow metabolism

1.1 Definition

Overflow metabolism is defined as the excretion of fermentation products such as acetate and lactate even in the presence of oxygen. Knowing that fermentation is actually a backup process of respiration which would be less efficient in generating energy that feeds for growth, overflow metabolism is considered as a "seemingly-wasteful" process.

The phenomenon of overflow metabolism exists widely ranging from bacteria (such as *Escherichia coli*) to mammalian cells. It typically occurs when the cell is in high growth rates, such as proliferating cells and cancer cells (known as Warburg effect). Figure 1 is showing the scheme of the phenomenon.

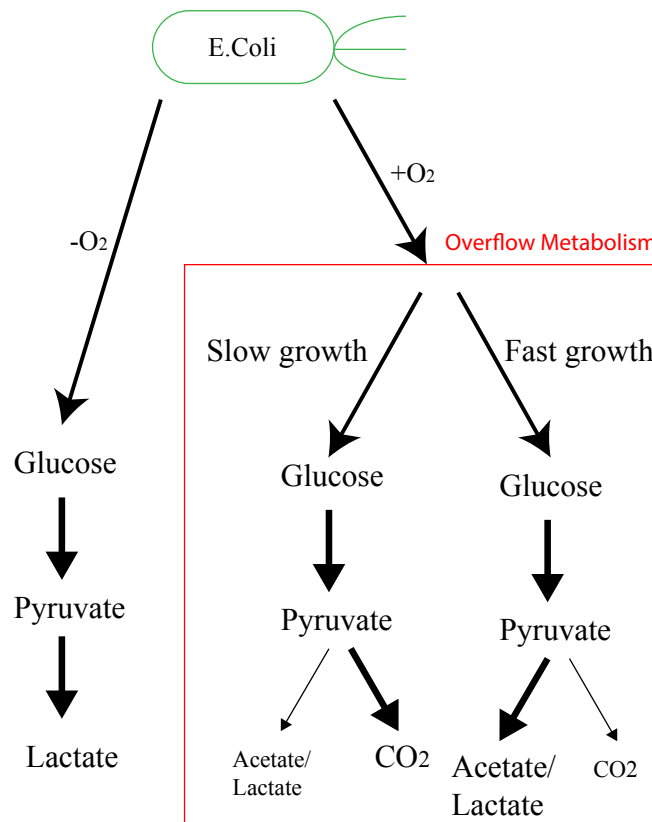


Figure 1: Scheme of the phenomenon of overflow metabolism

1.2 Explanation

The previous study shows that the factor that is trading with energy efficiency could be enzyme efficiency. As shown in figure 2, by uptaking the substrate, the cell can generate energy from either respiration pathway(blue) or fermentation pathway(red), where the energy would feed the biomass synthesis(black). Fermentation pathway would be less efficiency in generating energy(shown as the thinner yellow arrow representing ATP flux) but required less amount of protein for the reaction to be catalyzed(shown as the smaller protein structure as the enzyme for fermentation).

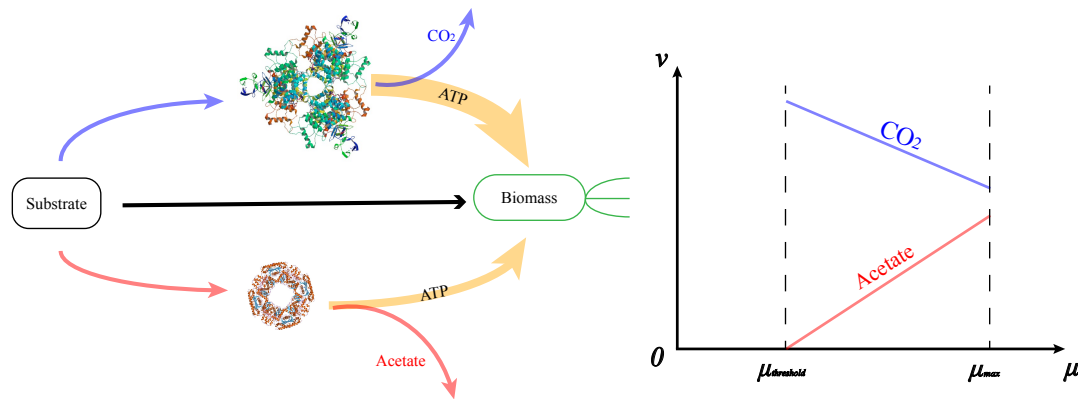


Figure 2: Overflow metabolism based on pathway switch

Therefore, overflow metabolism is indeed a pathway switch behavior that when it occurs, the respiration rate would decrease and fermentation rate would increase due to the higher enzymatic efficiency of fermentation pathway. A toy model created by *Molenaar et al*[3] simulates this behavior based on enzyme kinetics parameters(k_{cat} and K_M in Michaelis-Menten equation[4]).

According to the experiment data gained by *Basan et al*[5], the increase/decrease of fermentation/respiration pathway is actually linearly dependent on growth rate, which is qualitatively shown as the plot in figure 2(right). For the existence of maximum cell growth, the threshold-linear response introduced by *Basan et al* can be extended as the threshold-linear-end-point response:

$$v_{ac}(\mu) = \begin{cases} s_{ac} \cdot (\mu - \mu_{threshold}) & \text{for } \mu_{max} \geq \mu \geq \mu_{threshold} \\ 0 & \text{else} \end{cases} \quad (1)$$

Where s_{ac} is the slope of the line. *Basan et al* also built a computation model whose results would fit their experiment data. Instead of using enzyme kinetics parameters, they used a new variable ε representing "proteomic efficiency" to quantify the enzymatic efficiency of a pathway. In chapter 2, Their model would be discussed more in detail.

Chapter 2: Single cell metabolic and gene expression model

Cell metabolic model is a computation approach that mimics how species in a cell would be involved in the biochemical process, which will typically be presented in a network format[6].

The cell metabolic network would usually treat the species as nodes in a network, while the reactions as edges connecting the nodes with the reaction rate as the weight. For the conservation of mass, the total flux amount that flows into a node should be equal to the total flux that flows out. Considering each reaction is catalyzed by the enzyme, which is a product of gene expression. By balancing the fluxes, genotype and phenotype of the cell would be connected. Therefore, this thesis is using the cell metabolic and gene expression model to figure out the phenotype of overflow metabolism. As introduced in chapter 1, overflow metabolism is caused by a tradeoff between metabolic efficiency and enzymatic efficiency, the gene expression(protein) synthesis must be added into the network.

When the scale of the model is small, the phenotype of the cell can be expressed by exact equations through derivation, like in *Basan et al.* , an exact formula expressing the acetate vs. growth relation is presented. While for more general cases, especially when the model is in genome scale, where the complexity of the metabolic network is high, flux balance analysis would be used to solve the phenotype.

Flux balance analysis[7] is basically a linear optimization approach that maximizes the objective function subject to the vector of the fluxes is the null space of the stoichiometric matrix and other additional constraints.

Figure 3 shows the two approaches to solve a metabolic network. The derivation in *Basan et al.* would take a specific growth rate as an input, and derive the acetate-growth relation. While for a more general case using flux balance analysis, we take a set of constraints and the cell metabolic network interpreted as a stoichiometric matrix as the input, then optimize the objective function(which is maximizing the growth rate). Then scatters

the relation between growth and acetate excretion.

In this thesis, the behavior of (*E.coli*) would be studied in two different scales of metabolic model. Small scale, in which the model is simple and has a huge degree of freedom, mainly simulates the bulk behavior of the cell and can easily fit the experiment data, however, the values we get from the model might be some results of over-fitting and therefore the significance might be obscure. While for the genome-scale model, detail behavior such as the flux of each reaction of the cell can be simulated, but having a limited degree of freedom causing the potential difficulty of fitting the experiment data.

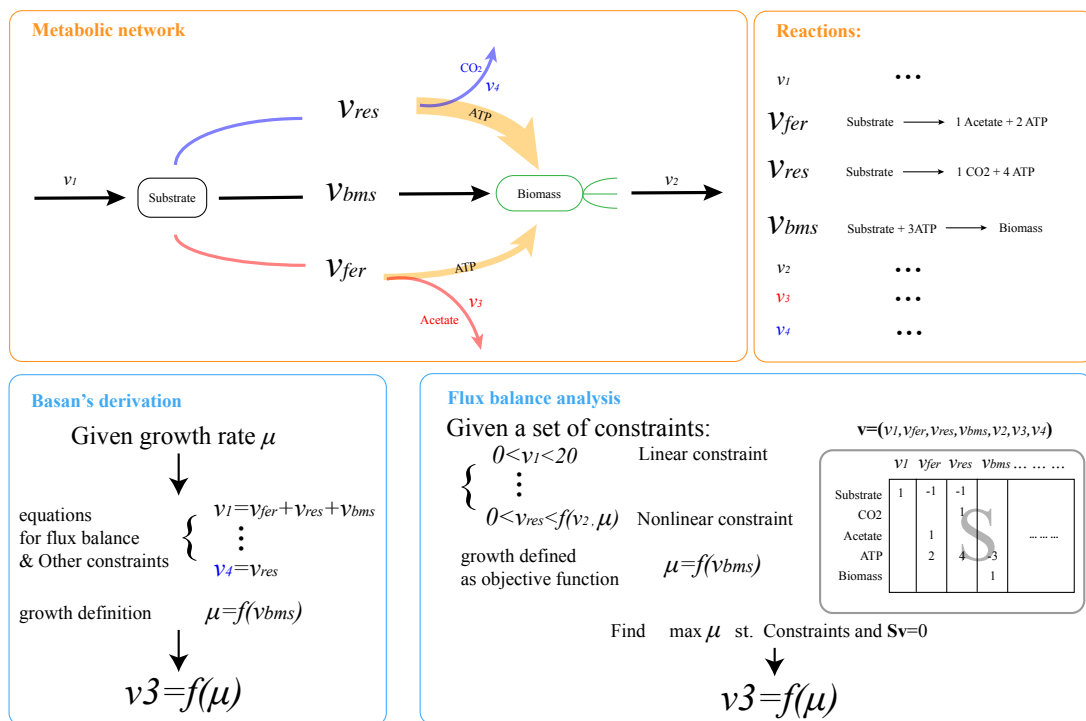


Figure 3: Comparison between the derivation in *Basan et al.* and optimization upon flux balance analysis

2.1 Unicellular growth and cell components dilution

The growth rate means the number of replicated cell per unit of time on a single cell. When the cell gets replicated, it needs to dilute the same amount of cell components as itself.

$$v_{dilution,C} = \mu [C], C \in \text{cell components} \quad (2)$$

In equation (2), $v_{dilution,C}$ is the dilution rate of the corresponding cell component, which represents the rate that the component appearing in the new grown cell. μ is the growth rate, which is the the rate that the cell gets replicated. $[C]$ is the concentration of the components in the original cell.

In computational prospect, the growth rate μ usually has the unit of h^{-1} , while the unit for reaction rate is in $mmol\ gDW^{-1}\ h^{-1}$, and for concentration it is $mmol\ gDW^{-1}$. As $mmol\ gDW^{-1}\ h^{-1} = h^{-1} \times mmol\ gDW^{-1}$, showing that equation (2) holds in units.

2.2 The coupling constraint: proteome efficiency/enzyme kinetics

For the overflow metabolism is caused by a tradeoff between energy efficiency and proteome efficiency, knowing that energy efficiency can be quantified as the stoichiometric coefficient of the amount of energy created per substrate, however, the way of quantifying proteome efficiency would not be shown as a constant stoichiometric coefficient, we, therefore, need extra constraints as the proteome efficiency to model the tradeoff behavior.

This thesis focuses on two kinds of method to provide coupling constraint based on proteome efficiency. *Basan et al.* introduced proteome efficiency based on proteome allocation, considering a kind of protein as a fraction of the entire proteome. *O'brian et al.*[8] derived the coupling constraint based on enzyme kinetics parameters taking account of the dilution rate of the enzyme (figure 4 and eq.6).

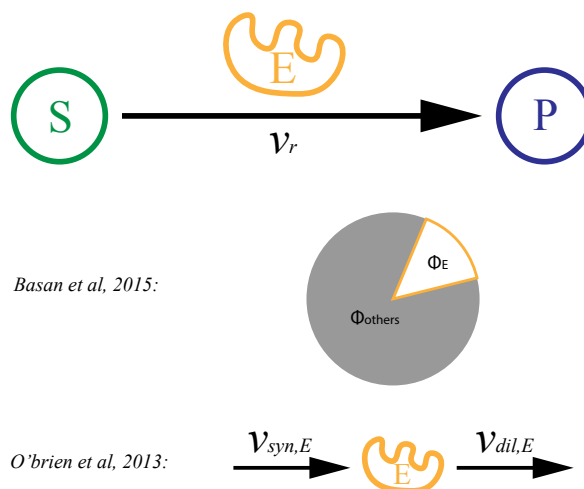


Figure 4: Enzymatic/proteomic efficiency

Basan et al. defined the proteome efficiency with the symbol ε as the coefficient to the corresponding reaction rate multiply with the fraction ϕ of the enzyme upon the entire proteome.

$$v_r = \varepsilon \cdot \phi_E \quad (3)$$

Assuming that the amount of the entire proteome is linearly dependent on the growth rate μ with the coefficient γ . For a particular enzyme E , its dilution rate $v_{dil,E}$ has to be at least as much as the enzyme that required to use, thus:

$$\phi_{E,reaction} \leq \frac{v_{dil,E}}{\gamma\mu} \quad (4)$$

Michaelis-Menten equation:

$$v_{reaction} = \frac{k_{cat}[E][S]}{K_M + [S]} \quad (5)$$

If $K_M = 0$, the concentration of substrate $[S]$ can be canceled out. Considering the actual case $K_M > 0$, there would be a parameter k_{eff} related to k_{cat} (which could be equivalent but might not) that:

$$v_{reaction} \leq k_{eff,E} \cdot [E] = k_{eff,E} \cdot \frac{v_{dil,E}}{\mu} \quad (6)$$

$$k_{eff,E} = \varepsilon_E/\gamma \quad (7)$$

2.3 Small scale model with alternative pathway

Basan et al. built up a small scale with the proteome allocation constraints and solve the growth-acetate relation via derivation. While the model created by *Basan et al.* is in genome-scale and is solved via growth optimization. As we have shown that the parameters providing coupling constraints(k_{eff} vs. ε) are related, how to reconstruct the model defined by *Basan et al.* in a flux balance analysis format using the enzyme kinetics(eq.6) as the proteome constraint? Finding the connection between those two approaches would be helpful to understand the genome-scale model better.

Thus, a small scale metabolic and gene expression model(small scale ME model, abbreviated as "SSME") is built based on the settings from *Basan et al* (figure 5.). After the substrate(carbon from glucose, g6p, lactose, mannitol or other carbon sources)is uptaken, it would contribute to 3 pathways, respiration(shown as "res"), fermentation("fer") and

biomass synthesis pathway (“bms”). The respiration would generate energy(ATP) to feed the synthesis of biomass, and at the mean time, excrete CO₂. The fermentation would also generate energy, which is not as much as the respiration pathway, and excrete acetate. The biomass pathway, as its name, would synthesize biomass, which is a multiplication of the growth rate μ with the coefficient β .

According to Basan et al, the coupling constraint on each reaction(pathway) is based on proteome allocation, or to be more accurate, biomass allocation, which can be explained in three steps:

- The entire biomass($\beta\mu$) which simply represents the entire cell, can be separated into proteome with the coefficient α and other biomass(with the coefficient $1 - \alpha$ such as lipids).
- The proteome can be separated into the proteins serving on the respiration, fermentation and biomass pathways with the coefficient ϕ_{max} , as well as the proteome for other pathways.
- The proteome for respiration(ϕ_r), fermentation(ϕ_f) and biomass($1 - \phi_r - \phi_f$) pathways are divided, serving as the enzymes that catalyze each pathway.

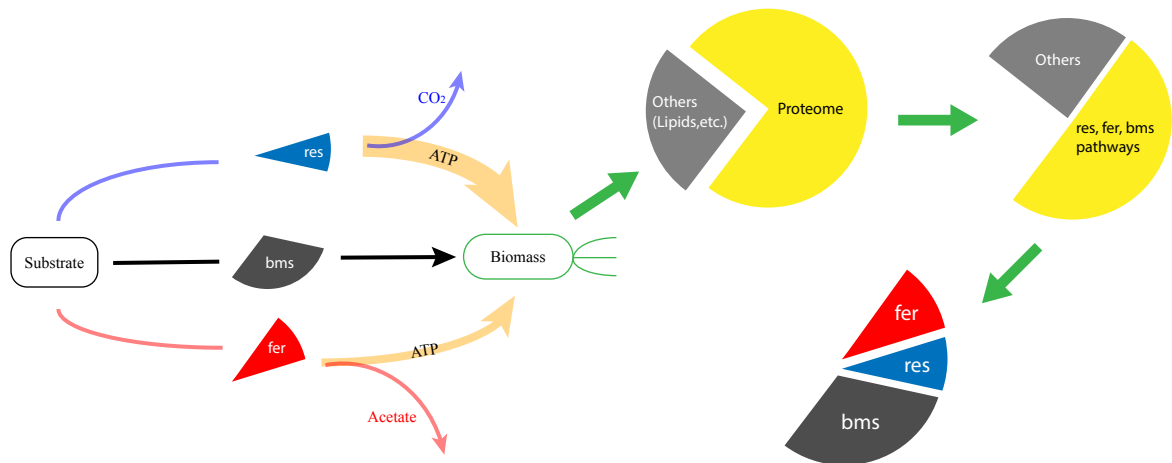


Figure 5: Small scale model with alternative pathways

In *Basan et al.*'s settings, fluxes through respiration and fermentation pathway are correspondingly notated as $J_{E,r}$ and $J_{E,f}$. The total energy that is created through these two path way must meet the energy demand for growth in the rate of μ

$$J_{E,f} + J_{E,r} = J_E(\mu) = \sigma\mu \quad (8)$$

The flux of uptaking substrate(carbon source) is annotated as $J_{C,in}$, while the fluxes that consuming substrate through respiration and fermentation pathways are in terms of $J_{C,f}$ and $J_{C,r}$.

By uptaken through flux $J_{C,in}$ consumed through fluxes $J_{C,f}$ and $J_{C,r}$, the rest of the carbon substrate would be contributed to the biomass flux, meeting the biomass demand for growth.

$$J_{C,bms}(\mu) = J_{C,in} - J_{C,f} - J_{C,r} = \beta\mu \quad (9)$$

Adopting the settings from *O'brien et al.*, the growth energy flux and biomass carbon flux could be considered as the demand reaction that their reaction rates are:

$$\begin{cases} v_{demand,energy} = J_E(\mu) = \sigma\mu \\ v_{demand,biomass} = J_{C,bms}(\mu) = \beta\mu \end{cases} \quad (10)$$

While for the substrate uptake rate bound(annotated as $SURB$), it is considered to be equivalent to the maximum value of the carbon input flux: $J_{C,in} \leq SURB$.

For the respiration and fermentation reactions, the rates would be normalized equaling to the carbon flux of the two pathways.

While according the settings from *Basan et al.*, the amount of energy that is created through either respiration or fermentation pathway is linearly related to the carbon flux with the coefficient e_r and e_f .

$$\begin{cases} v_r = J_{C,r} = J_{E,r}/e_r \\ v_f = J_{C,f} = J_{E,f}/e_f \end{cases} \quad (11)$$

Assuming that the proteome possesses a fixed portion $\alpha(0 \leq \alpha \leq 1)$ of the biomass content.

$$\begin{cases} v_{demand,proteome} = \alpha \cdot v_{demand,biomass}(\mu) = \alpha \cdot \beta\mu \\ v_{demand,otherbiomass} = (1 - \alpha) \cdot \beta\mu \end{cases} \quad (12)$$

While the entire proteome would be divided into growth dependent proteome ($\mu DPro$) possessing the maximum portion $\phi_{max}(0 \leq \phi_{max} \leq 1)$ and growth independent proteome ($\mu IPro$):

$$\begin{cases} v_{demand,\mu DPro} = \phi_{max} \cdot v_{demand,proteome} = \phi_{max} \cdot \alpha \cdot \beta\mu \\ v_{demand,\mu IPro} = (1 - \phi_{max}) \cdot \alpha \cdot \beta\mu \end{cases} \quad (13)$$

The growth dependent proteome is divided into three parts correspondingly contribute to catalyze respiration, fermentation and biomass pathways.

$$\phi_f + \phi_r + \phi_{BM} = 1 \quad (14)$$

Where the proteome part for the biomass would be divided into two parts, a fixed portion annotated by ϕ_0 and a portion whose amount is linearly related to the growth by coefficient b .

$$\phi_{BM} = \phi_0 + b\mu \quad (15)$$

Therefore, the entire proteome demand could be allocated as:

$$\begin{aligned} & v_{demand,proteome} \\ &= v_{demand,\mu DPro} + v_{demand,\mu IPro} \\ &= (\phi_f + \phi_r + \phi_0 + b\mu) \cdot \phi_{max} \cdot \alpha \cdot \beta\mu + (1 - \phi_{max}) \cdot \alpha \cdot \beta\mu \end{aligned} \quad (16)$$

Thus for the part with the fixed portion ϕ_0 in the growth dependent proteome, the dilution rate would be:

$$v_{demand,proteome_0} = \phi_0 \cdot \phi_{max} \cdot \alpha \cdot \beta\mu \quad (17)$$

With the same case, for the dilution of the growth independent proteome:

$$v_{demand,proteome_0} = (1 - \phi_{max}) \cdot \alpha \cdot \beta\mu \quad (18)$$

According to *Basan et al.*'s settings, the portion of the proteome catalyzing respiration/fermentation pathway is linearly related to the energy flux.

$$\begin{cases} \phi_r = \frac{J_{E,r}}{\epsilon_r} = \frac{e_r \cdot v_r}{\epsilon_r} \\ \phi_f = \frac{J_{E,f}}{\epsilon_f} = \frac{e_f \cdot v_f}{\epsilon_f} \end{cases} \quad (19)$$

The amount of components that are demanded for growth should be less than or the same as how much they are diluted into the new grown cell (same as how equation (4) was derived).

$$v_{demand,E} \leq v_{dilution,E}, (E \in \text{catalytic components}) \quad (20)$$

Consequently, for the dilution rate of the proteome catalyzing respiration pathway:

$$\begin{aligned} & v_{dilution,proteome_r} \\ & \geq v_{demand,proteome} \cdot \phi_{max} \cdot \phi_r \\ & = \alpha \cdot \beta \mu \cdot \phi_{max} \cdot \frac{e_r \cdot v_r}{\epsilon_r} \end{aligned} \quad (21)$$

Which could be derived into:

$$\begin{aligned} v_r & \leq \frac{\epsilon_r}{e_r \cdot \alpha \cdot \beta \cdot \phi_{max}} \cdot \frac{v_{dilution,proteome_r}}{\mu} \\ & = k_{eff,r} \times \frac{v_{dilution,proteome_r}}{\mu} \end{aligned} \quad (22)$$

For the fermentation pathway it meets the same case:

$$\begin{aligned} v_f & \leq \frac{\epsilon_f}{e_f \cdot \alpha \cdot \beta \cdot \phi_{max}} \cdot \frac{v_{dilution,proteome_f}}{\mu} \\ & = k_{eff,f} \times \frac{v_{dilution,proteome_f}}{\mu} \end{aligned} \quad (23)$$

And last, for the biomass pathway:

$$\begin{aligned} & v_{dilution,proteome_{bms}} \\ & \geq v_{demand,proteome} \cdot \phi_{max} \cdot (b\mu) \end{aligned} \quad (24)$$

$$\begin{aligned} v_{demand,proteome} & \leq \frac{1}{\phi_{max} \cdot b} \cdot \frac{v_{dilution,proteome_{bms}}}{\mu} \\ & = k_{eff,proteome_{bms}} \times \frac{v_{dilution,proteome_{bms}}}{\mu} \end{aligned} \quad (25)$$

The same as *Basan et al.*'s model, this model with the metabolic and expression format is able to predict the cases of carbon limitation(changing the *SUR*) and translation limitation(changing the value of $k_{eff,proteome_{bms}}$).

In order to make an equivalent effect as the prediction of energy dissipation(amount annotated as w), we could change rate of the energy demand reaction into:

$$v_{demand,energy} = \sigma\mu + w \quad (26)$$

Basan et al. was also able to predict the proteome limitation, by adding a useless protein *LacZ*, taking the portion of ϕ_Z , compressing the portion of growth dependent proteome from ϕ_{max} to $\phi_{max} - \phi_Z$.The way it works in the metabolic and expression format is to keep the k_{eff} values remain unchanged, while adding a new dilution reaction for *LacZ*:

$$\begin{aligned} v_{demand,LacZ} \\ &= v_{demand,proteome} \cdot \phi_Z \\ &= \alpha \cdot \beta\mu \cdot \phi_Z \end{aligned} \quad (27)$$

While the rate of the demand reaction for the portion ϕ_0 of the growth dependent proteome becomes:

$$v_{demand,proteome_0} = \phi_0 \cdot (\phi_{max} - \phi_Z) \quad (28)$$

As a result, eq. 22,23 and 25 are right in the format of the enzyme kinetics constraints(eq.6), which are the same as how the genome-scale model in *O'brien et al.* is formed.

Building up the SSME model could help us connect the settings of different computational model with each other, as well as understand the bulk behavior of the cell. Considering the respiration, fermentation and biomass reactions are actually abstract reactions that are not how the cell exactly works, we need to dig into the genome-scale model to understand more detail behavior. However, the results of the simulation on the SSME model is a good guide to lead us to understand how the genome-scale model works.

Therefore, in the next section, the genome-scale metabolic and macromolecular expression model, will be introduced.

2.4 Genome-scale models of metabolism and macromolecular expression

The genome-scale model is basically an expansion of the toy model introduced in the previous section, where the gene-reaction rule is specified, where the complexity is therefore dramatically increased. As the introduction of the proteome constraint provided by enzyme kinetics in the previous session, each enzyme encoding one or multiple reactions would be having a k_{eff} value. Besides the scale and the detail, another significant different between the genome-scale ME model and the SSME model is that the synthesis(transcription and translation) of proteome will be specified, so as the demand of DNA. In this way, the genome-scale model could potentially simulate the detail behavior of the cell. Table 1 shows the comparison of the two small scale models and the genome-scale ME model.

The k_{eff} values using in the model are determined by a multi-omics framework put forward by *Ebrahim, et al.*[9]. The model that is used in this research is an upgraded version of *O'brien et al's* genome-scale ME model, created by *Lloyd et al.*[10], and is called "iLE1678-ME". For the complexity of the model, binary search is used for solving the optimal growth rate[11].

Table 1: Comparison between the Small scale model(both *Basan et al.* and ME) and genome-scale ME model

	Small Scale		Genome Scale(iLE1678-ME)
	<i>Basan et al.</i>	SSME model	
Growth	$\frac{v_{bms}}{\beta}$ (eq.10)	$\frac{v_{bms}}{\beta}$ & μ plug-in	μ plug-in equations (eq.6)
Growth vs. proteome	Indicated linear	Linear (eq.16)	Not specified
Proteome constraint	Allocation (eq.19)	Turnover rate k_{eff}	k_{eff} (eq.6)
Protein synthesis	Not specified	Not specified	Specified, k_{rna} , k_{ribo} , etc
Energy demand	$v_{dmd,eng} = \sigma\mu$ (eq.26)	$v_{dmd,eng} = \sigma\mu$	GAM & NGAM
Results	Exact formula	Growth optimization	Growth Optimization

By taking account of the biochemistry process in the cell, the three pathways defined in the small scale model, respiration, fermentation and biomass synthesis, can be mapped on the genome scale as:

- Respiration pathway is a process that the cell uptaking oxygen which is completed by the cytochrome oxidase, then generates energy from the TCA cycle.
- Fermentation pathway is relatively straight forward, which is the process that acetyl-CoA turning into acetate.
- Biomass synthesis pathway is a bit obscure. Since every process in the cell can be considered as for the sake of growth. Considering that proteome is a really important portion of the biomass, the process of protein synthesis(transcription and translation) would be an interesting part to dig into.

Furthermore, as energy is an important component of the whole model, the ATP synthesis reaction would also be important to be studied.

Chapter 3: Results of simulation

Before any simulation, it is important to figure out how to validate the results. For cell metabolic model, as we are trying to simulate the overflow metabolism, it would be essential to measure both cellular growth and acetate excretion.

Basan et al. provided rich experiment data, which mainly measures the fluxes using the *E.coli* k-12 wild-type strain NCM3722 and its variations. Knowing that iLE1678-ME is modeling another k-12 wild-type strain MG1655, however, it is similar to NCM3722[13]. Thus, the experiment data can still be a reference for the validation of the results of the iLE1678-ME simulation.

Basan et al. measures using the unit of $mM/OD600nm/hr$. Knowing that $OD600nm$ is equal to $0.44 gDW/L$ in the *E.Coli* k-12 MG1655 strain[13], based on this relation, the experiment data presented with the flux unit $mM/gDW/hr$ and the threshold-linear-end-point fit is shown in figure 6.

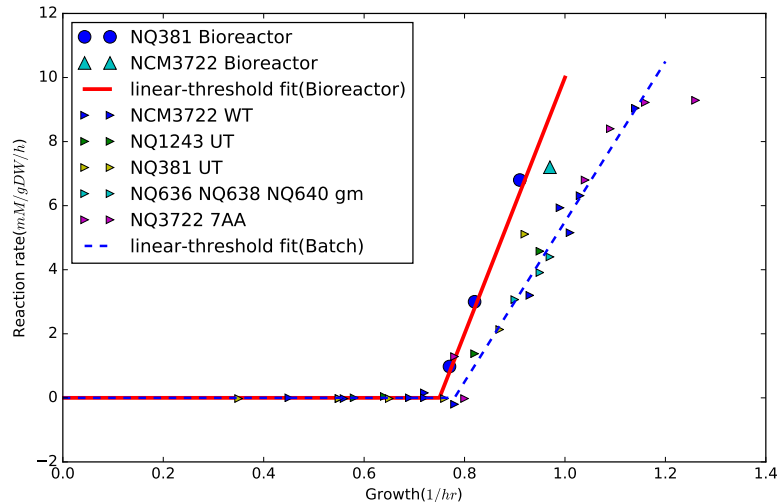


Figure 6: Experiment data from *Basan et al.*

As shown in figure 6, two sets of data are generated in different cell culture, bioreactor (red solid line) and batch culture (blue dash line). While as a result, there are two different linear-threshold-end-point fits for them. As *Basan et al.* is mainly using the result

generated from the bioreactor as their standard case, and this thesis is aiming to find the connection between *Basan et al.*'s model and the genome-scale iLE1678-ME model, the same standard is chosen.

3.1 Small Scale Model

Basan et al. derived the acetate-growth relation $v_{EX_act}(\mu)$ as a threshold-linear response (in the form as eq. 1), shown as the red dots in figure 7(right), and the CO_2 -growth relation as a negative slope line ending when it meets $y = 0$ shown as the blue dots in figure 7(right). In the SSME model, by using the same set of parameter values which fits the experiment data in *Basan et al.*, the exact same $v_{EX_act}(\mu)$ can be simulated, shown as the red solid line in figure 7 right. The CO_2 -growth relation, however, would be slightly different according to the simulation of SSME model, which is a piece-wise function where a region with positive slope and the other with a negative slope. The negative slope region is totally the same as *Basan et al.*. Considering that when the cell is in relatively low growth, respiration is required to generate the energy for biomass and would increase with the rising growth rate, the result of SSME simulation fits this behavior better.

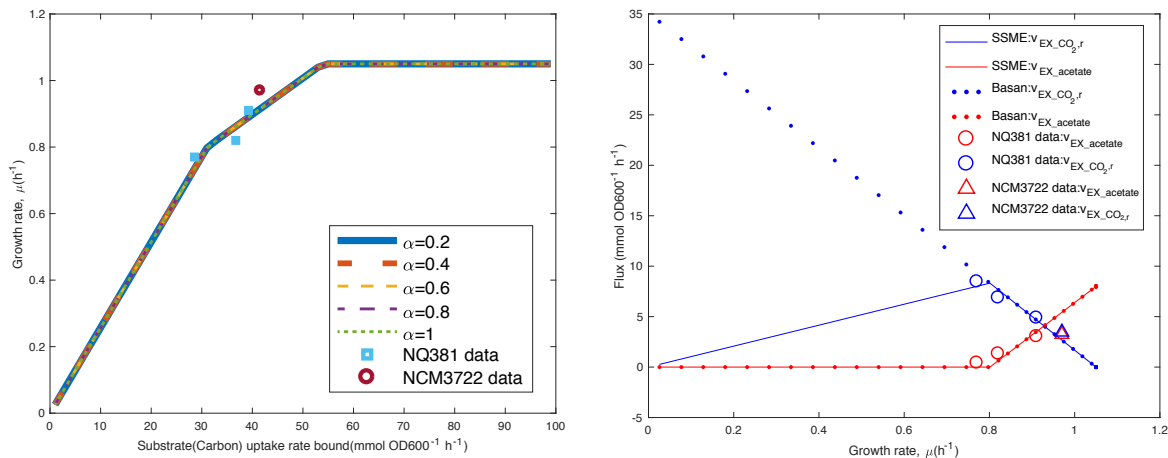


Figure 7: Comparison of the results from small ME model and from *Basan et al.*

The left figure of figure 7 shows the relation between growth and substrate uptake rate bound, with different proteome allocation coefficient α (eq.12). As a result, α does not

affect the results, which might be the reason why *Basan et al.* did not mention this coefficient in the derivation. The growth vs. substrate uptake rate bound is a piecewise function with 3 regions: steep slope region, gradual slope region, and flat region. The steep slope region, which is having a high substrate-growth yield, mainly correspond to respiration functioning. The gradual slope region mainly corresponds to the pathway switching to lower substrate-growth yield. The flat part mainly mean the metabolism is saturated and the growth has reached its maximum that providing more nutrients will contribute nothing.

The k_{eff} of the respiration pathway would provide an effect on the threshold while not changing the max growth (figure 8). As the model shows the intuition that the max growth is when the pathway has entirely switched with only fermentation, the respiration pathway would not be involved here. The slope, therefore, would change as the threshold changes, which rises as the k_{eff} increases.

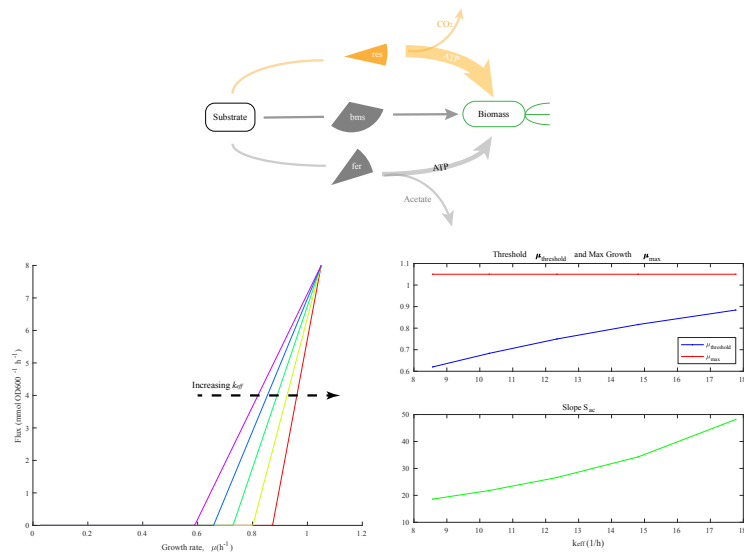


Figure 8: Acetate lines on multiple k_{eff} s of respiration pathway(left), Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of respiration(right)

On the contrary, the k_{eff} of the fermentation pathway would not affect the threshold but change the max growth rate (figure 9). And the slope would drop as the k_{eff} increases.

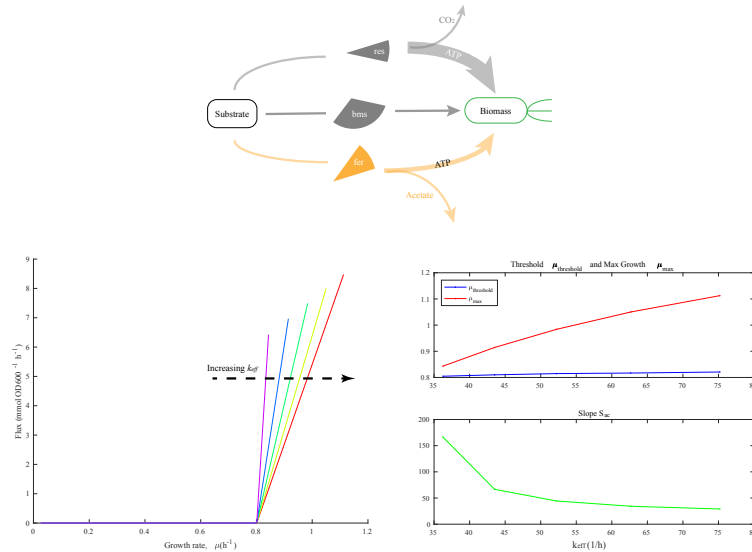


Figure 9: Acetate lines on multiple k_{eff} s of fermentation pathway(left), Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of fermentation(right)

The k_{eff} of biomass pathway would affect both the threshold and max growth. As a result, both the threshold and max growth would rise when increasing the k_{eff} , while the threshold would drop. The rationale of this effect can be that as the protein synthesis is part of the biomass, enzymes catalyzing respiration and fermentation are both more efficient to be synthesized, which potentially increases both their capacity.

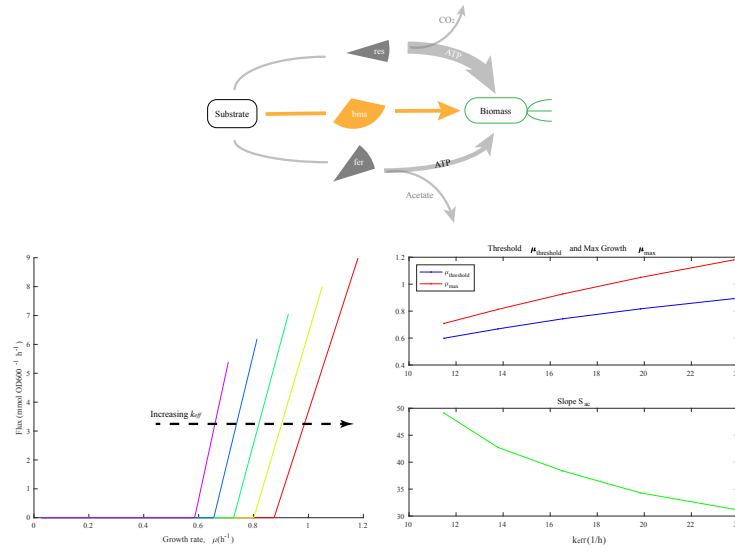


Figure 10: Acetate lines on multiple k_{eff} s of biomass synthesis pathway(left), Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of biomass synthesis(right)

3.2 Original genome-Scale model and its adjustment

By simulating the original iLE1678-ME model with free glucose uptake, which means the growth rate would be optimized to its maximum, there's only little acetate excretion, $0.66mM/gDW/hr$. To make things worse, by checking the metabolic pathway map (figure 11.), acetate is actually not synthesized by the fermentation pathway. Phosphotransacetylase and acetate kinase are not carrying any fluxes. Thus, some proper adjustment to the model in order to generate the pathway switch behavior is necessary.

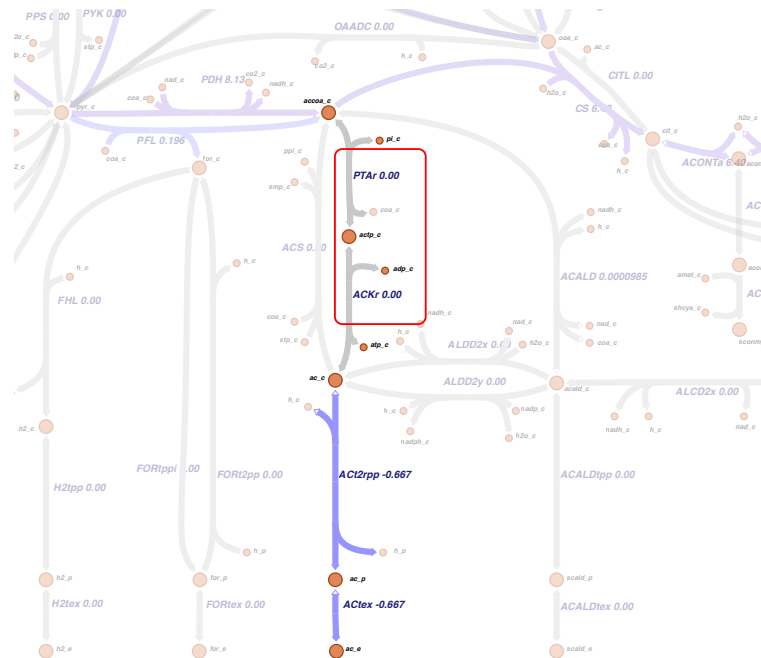


Figure 11: The problem of original genome-scale ME model: no fermentation

The reason of not having fermentation in high growth rate can be a result from $\mu_{threshold} \leq \mu_{max}$. In this case, we need to figure out a way to keep μ_{max} unchanged while decreasing $\mu_{threshold}$. The simulation of the small scale ME model (SSME) tells us that we can achieve this goal by decreasing the k_{eff} in respiration pathway, or in another word, figure out the correct k_{eff} values. *Basan et al.* provided a method to quantify their proteome efficiency ε . Considering the linear relation between ε and k_{eff} shown in eq.7. Even though the value of γ is unknown, as far as it is a constant and we are planning to scale the k_{eff} , replacing

k_{eff} with ε can be an initial guess of the adjustment.

Basan et al. measured the protein abundance of the gene expressions encoding the reactions in 5 different pathways (glycolysis, lactose degradation, TCA cycle, oxidative phosphorylation and acetate synthesis) in 4 different growth rates. While each reaction is separated into portions of respiration, fermentation, and biomass based on the stoichiometry in *Neidhardt et al.* [14]. Afterward, *Basan et al.* summed up all the protein abundance corresponding to respiration and fermentation pathways, then calculate ε_r , ε_f by getting the slope of the linear fit. The process of how *Basan et al.* mapping their 3-pathway model into the genome-scale is shown in figure 12, where only glycolysis, acetate synthesis, and TCA cycle are shown.

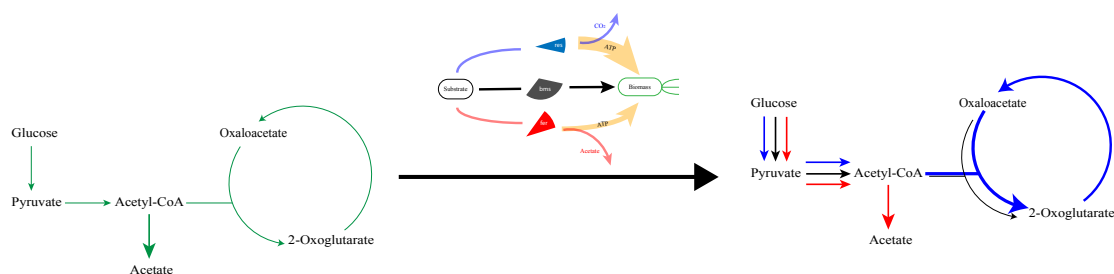


Figure 12: The approach of how the model in *Basan et al.* is mapped in Genome-Scale

On the vision from the 3-pathway model, since every reaction in genome-scale can potentially involve in multiple pathways, manipulating k_{eff} on a wrong reaction could change the proteome efficiency of both respiration and fermentation at the same time. Thus, as our goal is to change the k_{eff} of only respiration pathway, we need to manipulate the k_{eff} of the reactions which only involve in or mostly involve in respiration. TCA cycle is the only pathway that meets this requirement. The other pathway, for example, even though oxidative phosphorylation is the one that consumes oxygen, which seems to be closely related to respiration, as it is essential for generating energy, which is the key factor for growth, it would potentially involve hugely in biomass synthesis.

Table 2 shows how the ε is calculated from 4 different ATP fluxes, which corresponding to 4 different growth rates. The value of ε is the slope of the linear fit of the 4 protein

fraction percentages upon 4 ATP fluxes.

Table 2: Adjustment of k_{eff} values in TCA cycle

Reaction	Ori. k_{eff}	Gene-Reaction rule	Prot. % on different ATP fluxes				Slope (ε)
			37.22	30.22	21.46	15.22	
ACONTa	1598.02	acnA or acnB	0.92	0.84	0.66	0.57	59.8
ACONTb	1603.47	acnA or acnB	0.92	0.84	0.66	0.57	59.8
AKGDH	3000.00	sucA&sucB	1.34	1.36	1.2	1.07	66.34
CS	493.21	gltA	0.88	0.8	0.61	0.48	52.89
FUM	354.28	fumB or fumC or fumA	0.24	0.21	0.17	0.13	201.05
ICDHyr	3000.00	icd	1.55	1.55	1.31	1.39	65.38
MDH	95.90	mdh	0.45	0.45	0.41	0.39	307.74
SUCDi	75.44	sdhA&sdhB&sdhC&sdhD	0.49	0.45	0.42	0.35	158.90
SUCOAS	346.56	sucC&sucD	0.88	0.84	0.66	0.52	56.60

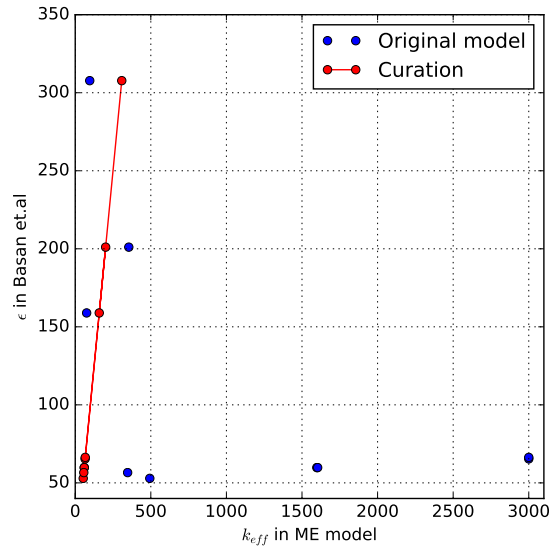


Figure 13: Adjustment of k_{eff} in TCA cycle from the genome-scale model

After getting the ε values of each reaction in the TCA cycle, the update of the k_{eff} is simply replacing them with ε values. As shown in figure 13, the original k_{eff} (blue dots) are replaced by ε (red dots).

The simulation with the new k_{eff} values results in the appearance of fermentation in maximum growth rate (figure 14).

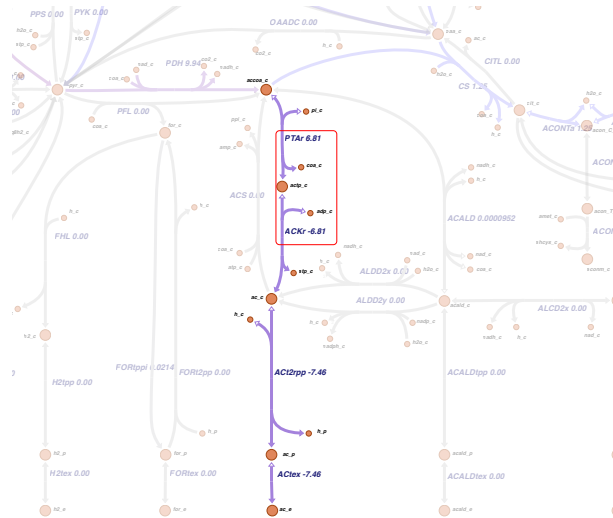


Figure 14: After k_{eff} adjustments, fermentation appears

Moreover, if we run the full simulation by varying the glucose uptake rate bound and plot the fermentation/respiration-growth relation (figure 15), we can tell that after the k_{eff} adjustments, pathway switch behavior appears, which is the same mechanism as the explanation of overflow metabolism.

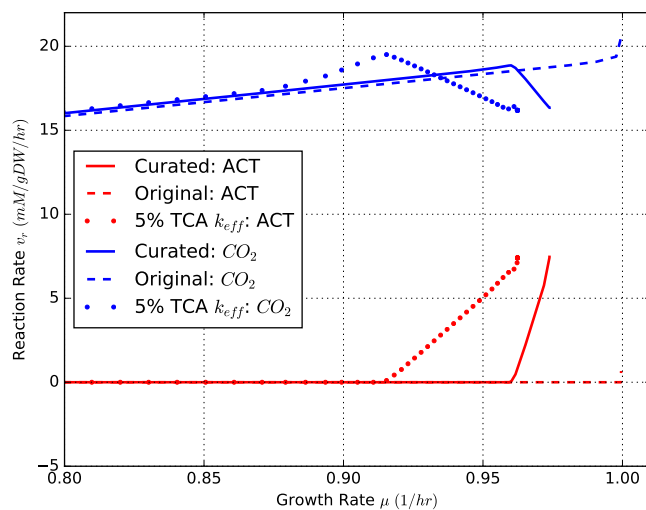


Figure 15: Comparison of the simulation results of pre and post-adjustment, as well as the k_{eff} s in TCA cycle are 95% off

In fact, besides the adjustment process introduced above, another simple method to recreate the pathway switch behavior is to scale down the k_{eff} of the entire TCA cycle (dot lines in figure 15). However, the result curation process validates the result of the derivation of the connection between the model created by *Basan et al.* and SSME. In addition, the post-adjustment result actually fits a set of experiment data provided by *Basan et al.* well (figure 21 in chapter 4.1). Therefore, the post-adjustment k_{eff} s would be considered as the standard case for the k_{eff} variation analysis.

3.3 Parameter sweep on adjusted genome-scale model

According to the stoichiometry in *Neidhardt et al.* used by *Basan et al.*, the TCA cycle can be separated into two steps. The first step (figure 16 left), ending with 2-Oxoglutarate, a precursor of glutamate, is mainly a respiration process but having a biomass portion. The glutamate is essential to growth, which is a component of the biomass, shown in the reconstruction of genome-scale *E.coli* model iAF1260[15]. The second step (figure 16 right), end with Oxaloacetate, is pure respiration process.

As a result (figure 16 left), for the first step of TCA cycle, the behavior of varying the k_{eff} is similar to a combination of varying the k_{eff} s of both respiration and fermentation. As there are both rising and dropping trends (appearing in different regions) of the slope, corresponding to the behavior in biomass and respiration. In addition, varying the k_{eff} would affect both $\mu_{threshold}$ and μ_{max} , which is seemingly the same behavior as the biomass pathway.

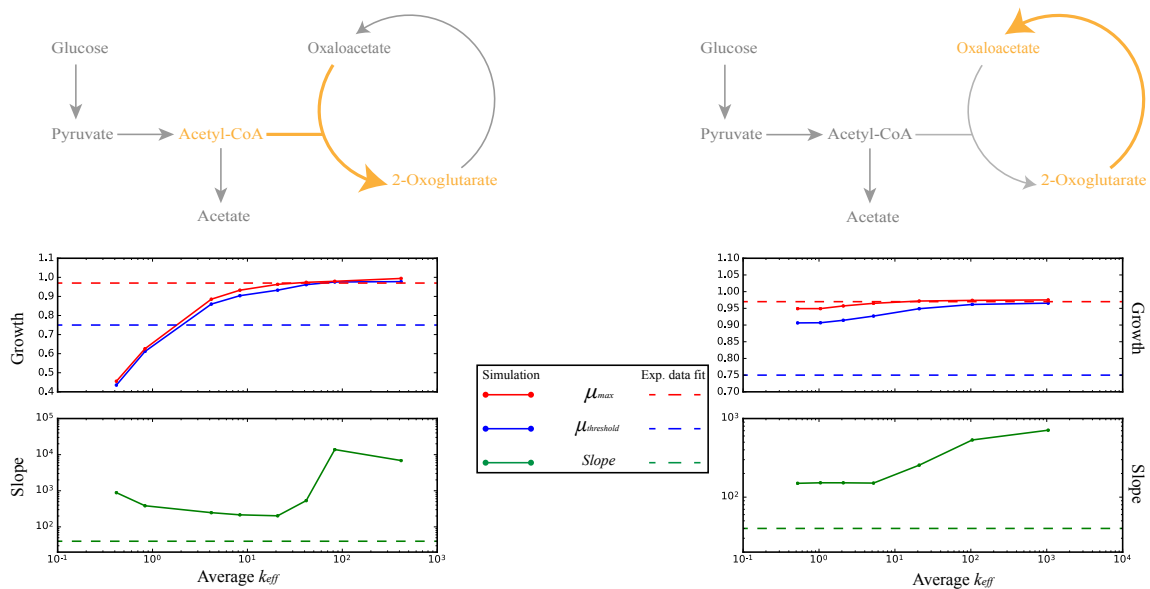


Figure 16: Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of TCA cycle, the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines

The second step of TCA cycle (figure 16 right) behaves exactly the same as how the respiration pathway works, that μ_{max} is stable while $\mu_{threshold}$ would be varying with the

variation of k_{eff} s. The slope would increase with the k_{eff} increases, which is the same as the respiration pathway. However, due to the complexity of the genome-scale model, when the k_{eff} shrinks to a really small value, the effect of the variation converges.

The process of acetate kinase (figure 17), which is an important step of Acetyl-CoA turning into acetate, is obviously a fermentation process. Therefore, as the result, it is not surprising that vary in the k_{eff} of acetate kinase would have the exact same effect as for how the fermentation pathway behaves. The only problem is similar to the second step of TCA cycle, for which the effect converges.

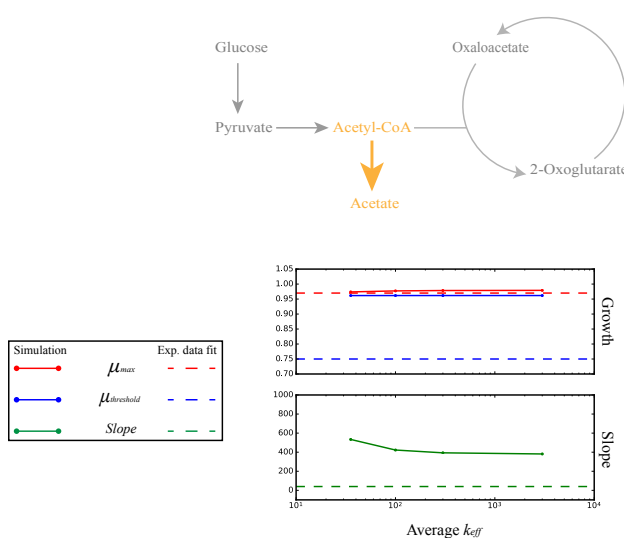


Figure 17: Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of reaction ACKr, the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines

Oxidative phosphorylation (i.e. ATP synthase and cytochrome oxidase, as shown in figure 18), as it uptakes oxygen, is obviously a respiration process. However, as it is crucial for generating energy, which is a key demand for growth, it is significantly involved in biomass pathway too.

As shown in figure 18, the result of varying the k_{eff} values of ATP synthase or cytochrome oxidase can be interpreted as an effect on both biomass pathway and respiration pathway. Both μ_{max} and $\mu_{threshold}$ are increasing with the rise of k_{eff} , which is the same as

biomass pathway. The slope behaves the same as respiration pathway that keeps increasing when the k_{eff} increases.

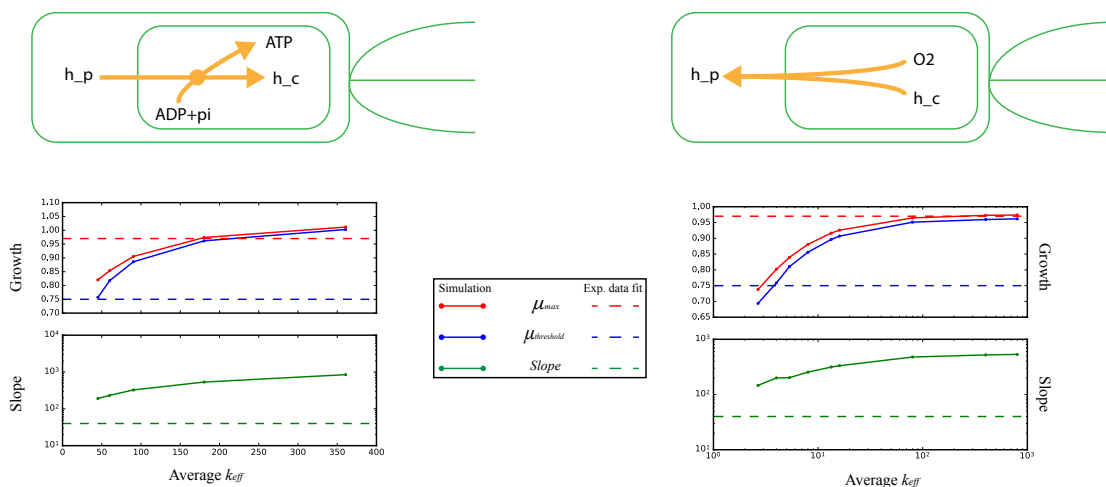


Figure 18: Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on k_{eff} of ATP synthase(left) and cytochrome oxidase(right), the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines

Some other parameters are tested where results are shown in figure 19. The translation efficiency noted as m_{rr} in iLE1678-ME, as expectation, provide the same effect as the k_{eff} of biomass pathway, as the biomass pathway would also take charge of protein synthesis. The global k_{eff} , which is simply scaling up or down the enzyme kinetics of all the metabolic reactions regardless of the protein synthesis, provide the capability to translate the acetate line.

As what is shown in the result, varying k_{eff} of a single pathway could either converge beyond the linear fit of the experiment data(dash horizontal lines in the plots from figure 16 to figure 19) or affecting the μ_{max} , $\mu_{threshold}$ and the slope at the same time.

By trying to make a combination of the k_{eff} variation, the result is shown in figure 20, where the slope as dropped significantly comparing to the standard case but still had some distance to fit the experiment data.

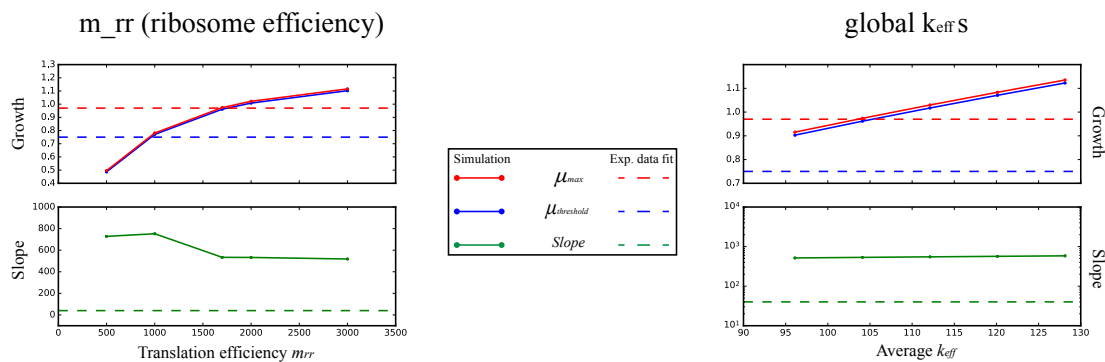


Figure 19: Dependence of S_{ac} , $\mu_{threshold}$ and μ_{max} on ribosome efficiency(left) and global k_{effs} (right), the parameters that fit the experiment data in figure 6 is shown as the dash horizontal lines

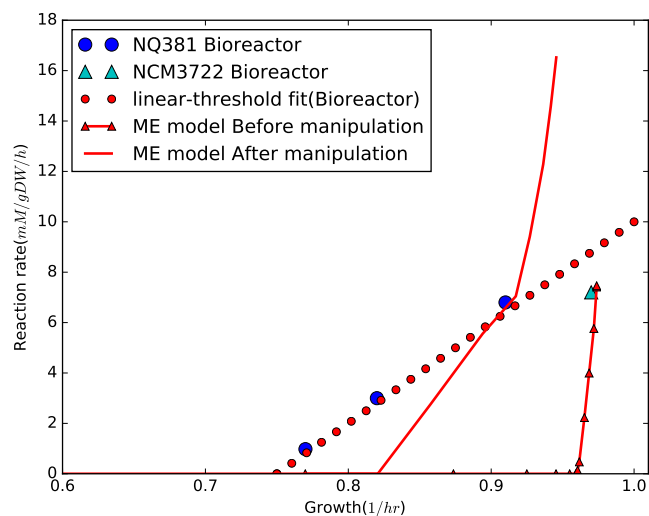


Figure 20: 50% ATP synthase k_{eff} , 5% TCA1 k_{eff} , 0.5% TCA2 k_{eff} , 0.25% CYT oxidase k_{eff} , 200% ACkr k_{eff} , 180% global k_{eff}

Chapter 4: Discussion

The result of varying k_{eff} s of different pathways can be approached to fit the experiment data but would converge to some bounds beyond. Thus, some extra constraint would need to be added to the genome-scale model. Three potential constraints are introduced in this chapter.

4.1 The function of motility protein

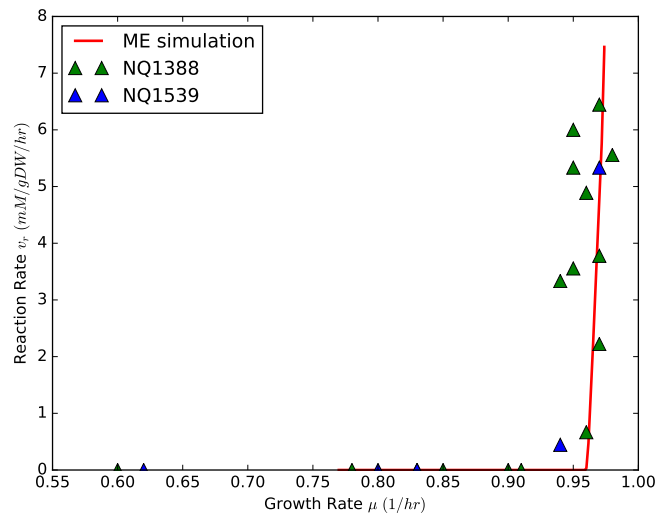


Figure 21: Good fit for the simulation on the strains which are not able to synthesize motility protein

Basan et al. provided the data (figure 21) with two special strains: NQ1388 and NQ1539, where their gene encoding for flagella assembly (flhD for NQ1388, and fliA for NQ1539) are knocked out, which means they are incapable of synthesizing motility protein. The experiment data can be fit by a linear-threshold response with a steep slope and high $\mu_{threshold}$. The explanation of the steeper slope response is that as the cells are cultivated in well-stirred culture, motility protein is useless. Knocking out the gene for flagella assembly would create more free proteome portion for respiration, thus the $\mu_{threshold}$ gets delayed.

The flagella synthesis pathway is not included in iLE1678-ME. Moreover, the simulation result of the standard case (post k_{eff} adjustment introduced in chapter 3.2) fits the experiment data of those two strains well, which validates the process of k_{eff} adjustment

and shows the necessity for adding the flagella assembly pathway to fit the NCM3722 and NQ381 data.

4.2 k_{eff} varied by growth rate

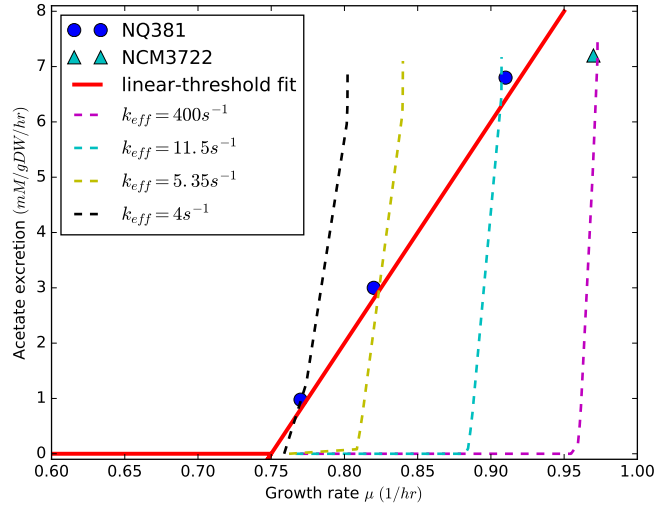


Figure 22: Matching the experiment data on multiple k_{eff} s of cytochrome oxidase

As introduced in figure 12 (chapter 3.2), when fitting the experiment data, *Basan et al.* treated each particular metabolic reaction as a combination of respiration, fermentation and biomass pathway with the coefficient from stoichiometry. This indicates that if every single metabolic reaction has a pathway switch behavior, which means the change of k_{eff} when changing the growth rate, the experiment data is guaranteed to get a good fit.

Therefore, a hypothesis that instead of a constant value k_{eff} , the enzyme kinetics of a single metabolic reaction is a function of the growth $k_{eff}(\mu)$. Figure 22 shows how the variation of k_{eff} for cytochrome oxidase varying with growth rate can fit the experiment data. Figure 23 shows a sigmoid fit of the k_{eff} samples where the fitting function is

$$k_{eff}(\mu) = \frac{398}{-70(\mu-0.935)} + 2.$$

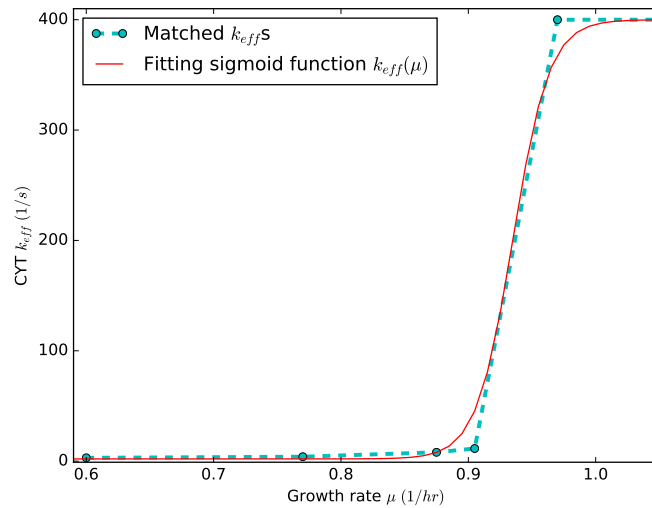


Figure 23: A Sigmoid function fit of growth dependence on cytochrome oxidase $k_{eff}s$, as a fitting function $k_{eff}(\mu)$

4.3 Limited oxygen uptake

When simulating iLE1678-ME, there is no bound for oxygen uptake. However, in an actual case of cell culturing, limitation of oxygen uptake might actually exist. Figure 24 shows a case of running the simulation of the adjusted iLE1678-ME model with a limitation of oxygen uptake (magenta line). As a result, it has a better fit on the experiment data, but loose linearity. There's a drop of acetate excretion rate when reaching the max growth where another pathway switch appears: pyruvate starts to synthesize lactate.

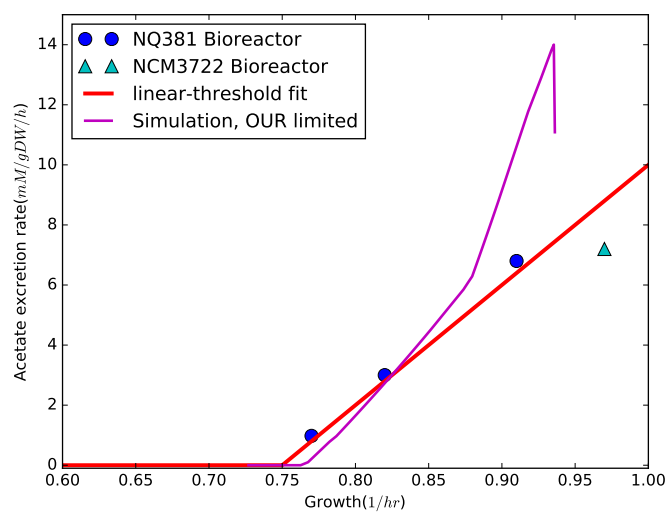


Figure 24: Simulation with oxygen uptake rate upper bound limited to 14 mM/gDW/h

Chapter 5: Conclusions

Overflow metabolism is a pathway switch behavior caused by the tradeoff of energy efficiency and proteome efficiency. The energy efficiency is quantified as the stoichiometric coefficient of the metabolic pathway, while proteome efficiency is quantified as enzyme kinetics based on gene expression. Thus, overflow metabolism cannot be replicated by the traditional metabolic model[15] without taking account of the enzyme kinetics. As a model combining cell metabolic network and gene expressions, the metabolic and gene expression model(ME model)[10] can potentially model overflow metabolism and in reverse, overflow metabolism can be a good test case to validate the model.

Basan et al. measured the behavior of overflow metabolism and fit it as a linear-threshold response. Besides, a small scale metabolic and proteome allocation model is created to fit the experiment data. This thesis derived the proteome allocation part of the model into the format of genome-scale ME model (*O'brien et al.*), named as small scale ME model(SSME).As a result, 3 models in 2 different scales and 2 different formats are presented.

The model created by *Basan et al.* is having a good fit on the experiment data and in a small scale, the latest ME model iLE1678-ME(*Lloyd*) simulates the detailed metabolic flux but need adjustment on the parameters to fit the experiment data. The SSME model, using the same metabolic pathway defined by *Basan et al.* and able to replicate the exact same results, while using the ME format, would be helpful to be as the reference guide to the adjustment and manipulation of the parameters.

As a result, it turns out to be hard to fit the standard experiment data using the genome-scale ME model and therefore, 3 hypotheses are presented, which could be interesting topics for future study.

References

- [1] De Mey, Marjan, Sofie De Maeseneire, Wim Soetaert, and Erick Vandamme. "Minimizing acetate formation in *E. coli* fermentations." *Journal of industrial microbiology & biotechnology* 34, no. 11 (2007): 689-700.
- [2] Vander Heiden, Matthew G., Lewis C. Cantley, and Craig B. Thompson. "Understanding the Warburg effect: the metabolic requirements of cell proliferation." *science* 324, no. 5930 (2009): 1029-1033.
- [3] Molenaar, Douwe, Rogier Van Berlo, Dick De Ridder, and Bas Teusink. "Shifts in growth strategies reflect tradeoffs in cellular economics." *Molecular systems biology* 5, no. 1 (2009): 323.
- [4] Michaelis, Leonor, and Maud L. Menten. "Die kinetik der invertinwirkung." *Biochem. z* 49.333-369 (1913): 352.
- [5] Basan, Markus, Sheng Hui, Hiroyuki Okano, Zhongge Zhang, Yang Shen, James R. Williamson, and Terence Hwa. "Overflow metabolism in *E. coli* results from efficient proteome allocation." *Nature* 528, no. 7580 (2015): 99.
- [6] Francke, Christof, Roland J. Siezen, and Bas Teusink. "Reconstructing the metabolic network of a bacterium from its genome." *Trends in microbiology* 13, no. 11 (2005): 550-558.
- [7] Orth, Jeffrey D., Ines Thiele, and Bernhard Ø. Palsson. "What is flux balance analysis?." *Nature Biotechnology* 28.3 (2010): 245-248.
- [8] O'brien, Edward J., Joshua A. Lerman, Roger L. Chang, Daniel R. Hyduke, and Bernhard Ø. Palsson. "Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction." *Molecular systems biology* 9, no. 1 (2013): 693.
- [9] Ebrahim, Ali, Elizabeth Brunk, Justin Tan, Edward J. O'brien, Donghyuk Kim, Richard Szubin, Joshua A. Lerman, Anna Lechner, Anand Sastry, Aarash Bordbar, Adam M. Feist, and Bernhard Ø. Palsson. "Multi-omic data integration enables discovery of hidden biological regularities." *Nature Communications* 7 (2016).
- [10] Colton J Lloyd, Ali Ebrahim, Laurence Yang, View ORCID ProfileZachary Andrew King, Edward Catoi, Edward J O'Brien, Joanne K Liu, Bernhard O Palsson,"COBRAme: A Computational Framework for Building and Manipulating Models of Metabolism and Gene Expression", doi:10.1101/106559.
- [11] Ma, Ding, Laurence Yang, Ronan MT Fleming, Ines Thiele, Bernhard O. Palsson, and Michael A. Saunders. "Reliable and efficient solution of genome-scale models of Metabolism and macromolecular Expression." *Scientific reports* 7 (2017).
- [12] Lyons, Eric, Michael Freeling, Sydney Kustu, and William Inwood. "Using genomic sequencing for classical genetics in *E. coli* K12." *PloS one* 6, no. 2 (2011): e16717.
- [13] Conrad, Tom M., Michael Frazier, Andrew R. Joyce, Byung-Kwan Cho, Eric M. Knight, Nathan E. Lewis, Robert Landick, and Bernhard Ø. Palsson. "RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media." *Proceedings of the National Academy of Sciences* 107, no. 47 (2010): 20500-20505.
- [14] Neidhardt, Frederick Carl, John L. Ingraham, and Moselio Schaechter. *Physiology of the bacterial cell: a molecular approach*. Vol. 20. Sunderland: Sinauer, 1990.

- [15] Orth, Jeffrey D., Tom M. Conrad, Jessica Na, Joshua A. Lerman, Hojung Nam, Adam M. Feist, and Bernhard Ø. Palsson. "A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011." *Molecular systems biology* 7, no. 1 (2011): 535.