COMPARATIVE DETERMINATION OF BIOMASS COMPOSITION IN DIFFERENTIALLY ACTIVE METABOLIC STATES

HSUAN-CHAO CHIU[1] hcchiu@bu.edu
DANIEL SEGRÈ[1,2] dsegre@bu.edu

[1] *Graduate Program in Bioinformatics, Boston University, Boston, MA, 02215, USA*
[2] *Departments of Biology and Biomedical Engineering, Boston University, Boston, MA, 02215, USA*

Flux Balance Analysis (FBA) has been successfully applied to facilitate the understanding of cellular metabolism in model organisms. Standard formulations of FBA can be applied to large systems, but the accuracy of predictions may vary significantly depending on environmental conditions, genetic perturbations, or complex unknown regulatory constraints. Here we present an FBA-based approach to infer the biomass compositions that best describe multiple physiological states of a cell. Specifically, we seek to use experimental data (such as flux measurements, or mRNA expression levels) to infer best matching stoichiometrically balanced fluxes and metabolite sinks. Our algorithm is designed to provide predictions based on the comparative analysis of two metabolic states (e.g. wild-type and knockout, or two different time points), so as to be independent from possible arbitrary scaling factors. We test our algorithm using experimental data for metabolic fluxes in wild type and gene deletion strains of *E. coli.* In addition to demonstrating the capacity of our approach to correctly identify known exchange fluxes and biomass compositions, we analyze *E. coli* central carbon metabolism to show the changes of metabolic objectives and potential compensation for reducing power due to single enzyme gene deletion in pentose phosphate pathway.

Keywords: flux balance analysis; systems biology; data integration; metabolic objectives

1. Introduction

An important goal of systems biology is to reconstruct and simulate biological networks to facilitate the understanding of complex cellular metabolism. Constraint based approaches have been applied to characterize the cellular flux distribution and predict metabolic phenotypes for cells grown in different conditions. One of the most prominent constraint based approaches, Flux Balance Analysis (FBA), relies on a steady state approximation and optimization algorithms to predict metabolic fluxes at cellular level [15]. The steady state approximation translates into a set of constraints on the fluxes, namely that the net sum of all fluxes producing or consuming each metabolite has to be zero. FBA determines these steady state fluxes by searching the space of feasible solutions, a polyhedral space defined by multiple constraints, for a choice of fluxes that minimizes/maximizes an objective function associated with a biological task. For instance, for a unicellular organism, one may ask what is the solution that maximizes an appropriately defined growth (or biomass production) flux, reflecting selection for fast-growth during evolution [15]. In addition to maximizing growth, van Gulik and Heijnen suggested maximization of ATP yield, based on the assumption that evolution drives

maximal energy efficiency [14]. Bonarius *et al*. suggested minimization of overall intracellular flux, reflecting the hypothesis that organisms are evolved to maximize enzymatic efficiency [1]. Several works have proposed methods to identify objective functions from experimental data. Knorr *et al*. proposed a Bayesian-based probability ranking method to evaluate multiple objective functions [7]. Schuetz *et al*. have measured fluxes and evaluated different objectives with a Euclidean metric approach [11]. Among all the objectives studied by Schuetz *et al*., nonlinear maximization of the ATP yield best described unlimited growth on glucose in oxygen or nitrate respiring batch cultures while linear maximization of the overall ATP or biomass yields achieved the highest accuracy under nutrient limited continuous cultures [11]. Although FBA optimal growth seems to work well in several cases, it has been shown to be sometimes insufficient for predicting perturbed metabolic states, such as the one found in gene deletion knockout strains. A better way to determine mutant fluxes is to use Minimization Of Metabolic Adjustment (MOMA) [12], which assumes that the mutants would stay as close to wild type flux distribution as possible. One lesson learned from MOMA is that metabolic networks perturbed from a simple average behavior may be better described by objective functions different than standard growth rate maximization.

One can imagine, in general, that a living system may switch its objective when facing a physiological change. For example, the diauxic shift in yeast, which is the switching from anaerobic growth to aerobic respiration upon depletion of glucose, is known to be correlated with widespread changes in the expression of genes involved in carbon metabolism, protein synthesis, and carbohydrate storage [3, 6]. Understanding the physiology of such a natural progress is still an open challenge.

Lacking knowledge of objectives for perturbed cells and changes of objectives under different metabolic states limits the capacity to correctly describe metabolic networks using FBA methods. An alternative way to study metabolism is to infer metabolic flux objectives from available data. Comparative analyses of biomass compositions in different physiological states, either between wild type and mutants or throughout naturally occurring physiological transitions, could provide insight helpful towards understanding the design of metabolic networks. Previously Burgard and colleagues proposed ObjFind and BOSS to identify putative objective functions from flux measurements. Specifically, these methods identify the coefficients of importance responsible for flux distributions in *E. coli* and yeast [2, 4]. Uygun *et al*. proposed a multilayer optimization framework to discover the major fluxes of metabolic objective that account for the flux distribution in a mammalian cell [13]. However, these methods rely on flux measurements and cannot take advantage of other high throughput data.

Here we present an FBA-based approach to infer the biomass compositions that best describe multiple physiological states of a cell. Our method is designed to incorporate high throughput data for comparatively determining metabolic objectives in two physiological states. As a first step, we analyze here flux data from *E. coli* central carbon metabolism pathways [5] to demonstrate our method for predicting metabolic objectives.

2. Method

2.1. *Flux Balance Analysis*

FBA describes the cellular level reaction rates (fluxes) under a steady state approximation, thereby imposing linear mass balance constraints. All the nutrients taken from the extracellular environment would be consumed to produce biomass or other byproducts and taken out from the system without intracellular metabolite accumulation. The steady state equation responsible for mass balance can be written as follows:

$$dx/dt = Sv = 0 \tag{1}$$

where x is the vector of metabolites, v is the vector of reaction fluxes and S is the stoichiometric matrix of the network. S is an m by n matrix where m is the number of metabolites and n is the number of reactions. The value S_{ij} in S is the stoichiometric coefficient for metabolite i in reaction j. Additional constraints such as lower and upper bound for specific enzymatic reactions or nutrient uptake rates may also be imposed as $LB_j \leq v_j \leq UB_j$, for reaction v_j.

FBA determines a specific flux prediction by maximizing/minimizing a linear objective function associated with a biological task. A typical FBA objective used in microbial systems is the maximization of biomass production [15] based on the assumption that unicellular organisms have been selected to reach maximum growth performance during evolution. Biomass production is approximated by a growth flux v_{growth}, which is defined as follows:

$$\sum_i c_i X_i \xrightarrow{v_{growth}} 1\ Biomass \tag{2}$$

where c is the vector of biomass coefficients, whose component c_i indicates the proportion of metabolite X_i required for the formation of a unit of biomass.

The linear programming statement for maximizing growth in FBA could be formulated as:

$$\begin{aligned} &\max \quad v_{growth} \\ &\text{s.t.} \quad Sv = 0 \\ &\qquad LB_j \leq v_j \leq UB_j \end{aligned} \tag{3}$$

2.2. *Objective inference*

We extend the conventional FBA formulation to concurrently infer metabolic objectives in two different metabolic states of a system. Here we limit our search to maximization of biomass production as an objective function, but we allow the biomass composition to assume in principle any vector of coefficients. For instance, the two states could be the wild type and a given mutant. The goal is then to infer the corresponding c^1 and c^2 vectors of biomass coefficients best representing the metabolic objectives for the two corresponding physiological states.

To reverse engineer the objectives, we implement a linear optimization procedure to identify the FBA objectives maximally compatible with given vectors E^1 *and* E^2 encoding reference experimental data:

$$\begin{aligned} &\min \sum_{\substack{j \\ E_j^1 * E_j^2 \neq 0}} \left| \frac{v_j^1}{E_j^1} - \frac{v_j^2}{E_j^2} \right| \\ &\text{s.t. } S' \cdot v' = 0, \text{ where } S' = \begin{bmatrix} S & \Phi \\ \Phi & S \end{bmatrix}, \; v' = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \\ &LB_j^1 \le v_j^1 \le UB_j^1 \\ &LB_j^2 \le v_j^2 \le UB_j^2 \\ &\sum_j |v_j| \ge v_{\min} \end{aligned} \tag{4}$$

where v is the vector of fluxes to be determined, Φ is a zero-containing matrix with the same dimensions of S. In this optimization problem the overall flux activity (the sum of the absolute values of all fluxes) is imposed to be above a threshold v_{min} (e.g. 25% of the flux activity obtained with regular FBA). Biomass production reactions for the first and second state are disabled from the stoichiometric matrix and a sink reaction for each biomass component is added. Each single biomass component originally flowing into biomass is exported separately and the inferred fluxes will correspond to the biomass coefficients for the corresponding metabolic state. Our optimization method tries to optimize biomass coefficients simultaneously for two metabolic states, hence allowing us to take advantage of the fact that certain data could provide only relative changes between reaction activities in the two states. Here we limit the optimization to intracellular fluxes.

To test our objective function inference approach and demonstrate its performance, we apply our method to experimental flux measurements in *E. coli* central carbon metabolism pathways, taken from the paper published by Ishii *et al*. [5]. In their flux measurements, wild type strain of *E. coli* K-12 and 24 single gene deletion mutants of glycolysis and pentose phosphate pathway were grown in glucose-limited chemostat cultures. The mutant cells were grown at fixed dilution rate of 0.2 hours^{-1}, and wild-type cells cultured at the same specific growth rate were used as a reference sample. They also cultured wild type cells in different dilution rates (0.1, 0.4, 0.5, and 0.7 hours^{-1}) for comparison. In this work, we apply an *E. coli* central carbon metabolism FBA model [9] to study these data. The biomass production reaction in this model is a sink for the linear combination of several metabolites that are precursors of amino acids, nucleotides or lipids:

0.205 g6p + 0.361 e4p + 1.496 3pg + 1.787 oaa + 1.079 akg + 2.833 pyr +
0.898 r5p + 0.519 pep + 0.129 g3p + 0.071 f6p + 18.225 nadph + 3.748 accoa +
3.547 nad + 55.703 atp + 55.703 h2o
→
18.225 nadp + 3.748 coa + 3.547 nadh + 55.703 adp + 55.703 pi + 41.025 h (5)

The fact that we are dealing with a small model and there are a lot of sink reactions for the metabolites results in many alternative optima for the optimization in Eq. (4). Therefore we further use Minimization Of Metabolic Adjustment (MOMA) [12] to find the most probable steady state solution for v^l by exploring the solution space we get from optimizing Eq. (4). Coefficients for biomass precursors listed above are inferred after the primary and secondary optimization process.

3. Results

Performing gene deletions is a commonly used approach to study how an organism responds to perturbations. FBA and MOMA have been used for generating predictions of these metabolic responses. MOMA, in particular, has been shown to be more accurate for predicting mutant fluxes than FBA. However, there are cases in which neither FBA nor MOMA objectives seem to capture well enough the true metabolic state (Fig. 1). Hence this is a good test case for our algorithm, in search for biomass composition coefficients that would be compatible with experimental data.

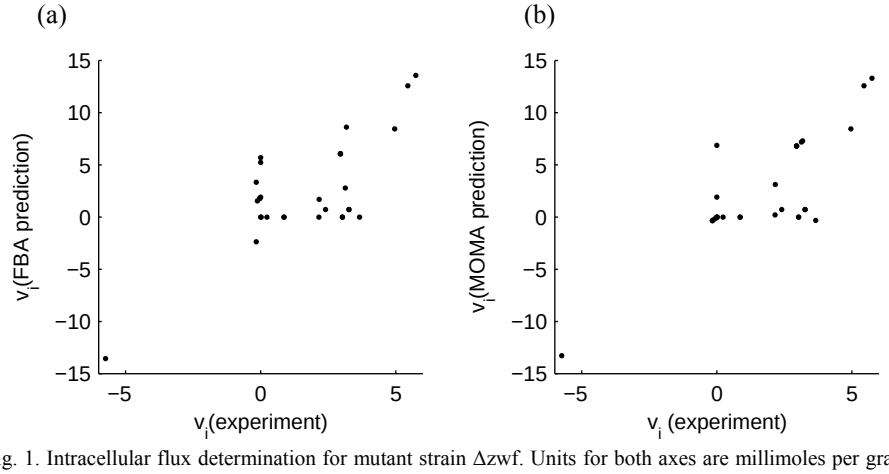

Fig. 1. Intracellular flux determination for mutant strain Δzwf. Units for both axes are millimoles per gram dry weight per hour (mmol/gDW/h). (a) FBA flux predictions for the mutant do not correlate well with experimental measurements. (b) MOMA predictions are expected to better correlate with experimental fluxes. However, in this case, even MOMA predicted mutant fluxes are not satisfactory enough for inferring biomass coefficients.

We applied our method to flux measurements in *E. coli* central carbon metabolism pathways [5] to infer the metabolic objectives in wild type and mutant strains. The reference (Ref) strain we use is the average of the four replicates available experimentally. Fig. 2 shows the correlation of predicted and experimental exchange rates (which were not part of the input of the above inference algorithm). Our predictions for glucose uptake rates agree with all wild type and mutant measurements studied here. In general, predicted oxygen uptake rates match well with experiments except for those of the GR03 wild type strain (See Table 1). The less accurate predictions for oxygen or CO_2 production rates may be caused by reactions consuming or producing these compounds that are not in central carbon metabolism pathways. For instance,

Ubiquinone-8 biosynthesis requires oxygen. Therefore, under-predicted oxygen uptake rates will result in a corresponding under-prediction of Ubiquinone-8 related reactions, such as NADH dehydrogenase or succinate dehydrogenase. In addition, predictions may also be affected by inaccurate flux measurements. For example, three CO_2–associated reactions have large standard deviations (larger than 0.5*mean, see Table S5B in [5]) in the wild type replicates, possibly due to experimental difficulties or resulted from the fitting procedure for flux corrections to achieve isotopomeric steady state.

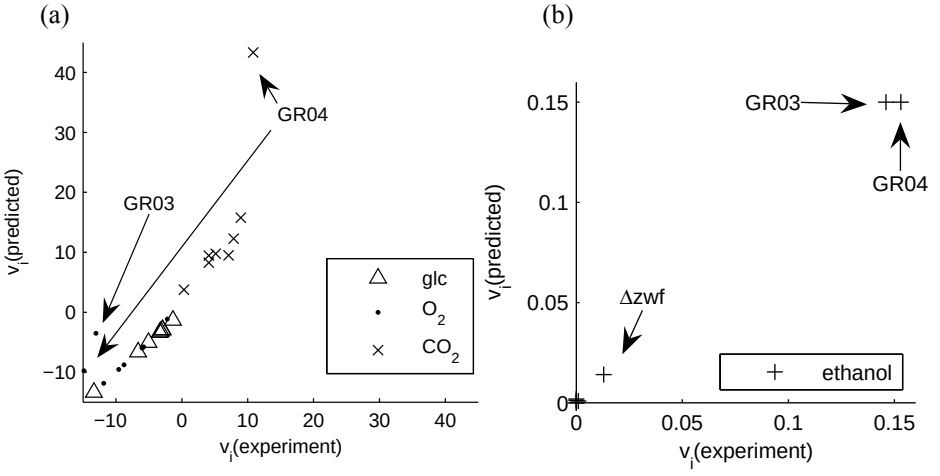

Fig. 2. Predicted uptake and secretion rates in wild type and three mutant strains Δzwf, Δpgl and Δgnd (mutants for pentose phosphate pathway single gene deletion). Unit for flux is millimoles per gram dry weight per hour (mmol/gDW/h). Negative values refer to uptake rates. (a) Glucose uptake rates in all cultures are predicted quite well. Some oxygen uptake rates are under-predicted in wild type strains under high dilution rates. The wild type strain with largest dilution rate (GR04) has a large deviation for CO_2 predictions. (b) All significant ethanol production rates are correctly predicted.

3.1. *Conserved biomass coefficients across different glucose supply rates*

For the wild type strains grown in different dilution rates, we implement our algorithm relative to the reference strain (Ref) mentioned above. In Fig. 3, the predicted production rates of the ten biomass precursors defined in the FBA model are plotted against the corresponding biomass coefficients. A linear correlation is observed across all dilution rates, ranging from an almost glucose-starved state to a nearly unlimited glucose supply. In addition, the slope of the line determined by the aligned data points roughly reflects the growth rate of each wild type strain. For instance, the slope observed in Fig. 3d is roughly 1.8 times the one in Fig. 3b, which matches the fold change of growth rates (0.7 vs. 0.4 h^{-1}) between these two experiments. These results suggest that *E. coli* grown in glucose supply cultures apply robust metabolic objectives for biomass precursors, in agreement with the FBA optimal growth assumption for central carbon metabolism, regardless of glucose supply rate.

3.2. *Influence of single gene deletion in pentose phosphate pathway*

The pentose phosphate pathway is responsible for generating NADPH and nucleotides. A perturbation in the pentose phosphate pathway could change the levels of NADPH and nucleotides and may result in less efficient growth. To study the changes of metabolic objectives caused by the deletion of pentose phosphate pathway genes, we applied our algorithm to three single gene deletion mutants Δzwf (glucose 6-phosphate-1-dehydrogenase), Δpgl (6-phosphogluconolactonase) and Δgnd (glucose 6-phosphate dehydrogenase), relative to the Ref state. Our goal was to see whether considerable changes of biomass coefficients or fluxes rerouting could be detected.

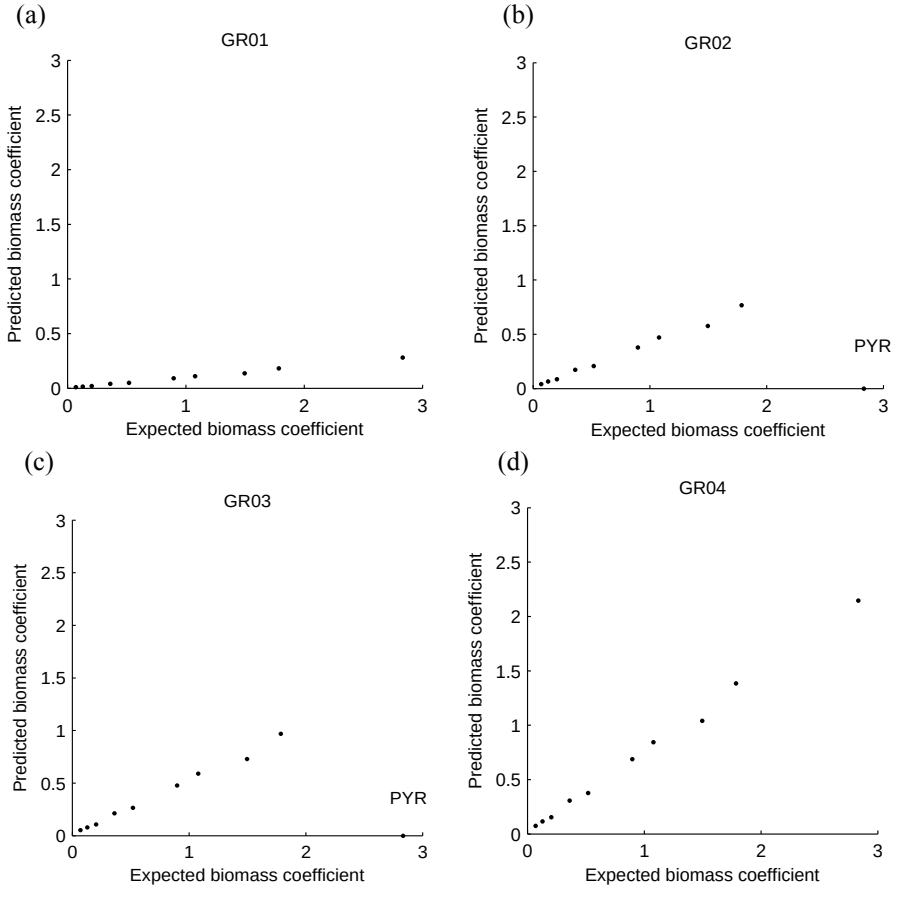

Fig. 3. Predicted production rates for biomass precursors in wild type strain under different dilution rates. Units for both axes are millimoles per gram dry weight per hour (mmol/gDW/h). The y coordinate for each data point represents the predicted flux production rate for the corresponding biomass component, and x coordinate is the biomass coefficient taken from the FBA model (see Eq. (5)). *E. coli* is cultured at dilution rate of $0.1h^{-1}$(a), $0.4h^{-1}$ (b), $0.5\ h^{-1}$ (c), $0.7h^{-1}$ (d) respectively [5]. The slope of the data line roughly reflects the growth rate for each experiment. Pyruvate coefficients in GR02 and in GR03 are erroneously predicted to be zero. This might be related to the less accurate prediction of CO_2 production rate, since several reactions consuming or producing pyruvate generate CO_2.

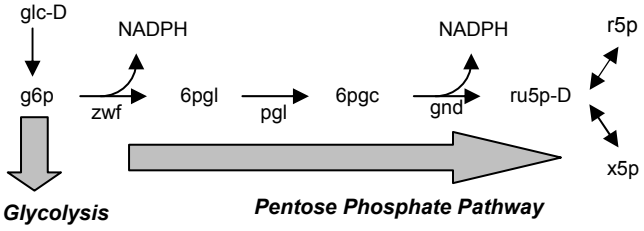

Fig. 4. Map of the initial reactions in the pentose phosphate pathway. zwf and gnd are responsible for NADPH production to generate reducing power for growth.

Fig. 4 illustrates the reactions being knocked out from the pentose phosphate pathway in our computational study. Detailed predictions for biomass components and several key fluxes are shown in Table 1. Note that all mutants were grown in chemostat cultures at the same dilution rate ($0.2h^{-1}$) as the wild type strain (column Ref in Table 1). Hence they can all be considered to grow at the same rate. The predicted production rates for the different biomass precursors can therefore be directly compared between different mutants, and relative to the corresponding biomass composition coefficients used in FBA calculations, appropriately normalized. Our results show that most biomass coefficients for biomass precursors change proportionally to the coefficients themselves across different strains. However, individual deviations from this trend can be seen. Fig. 5 shows the predicted production rates for biomass precursors in wild type and mutants. Amino acids and nucleotide precursors (e4p bar to pep bar) tend to be over-produced and under-produced in Δpgl and in Δgnd respectively, compared to the Ref strain. Meanwhile, the measured dry weight for Δpgl and Δgnd show the same trend as the production rates for these biomass precursors. One explanation for the deviations is that these mutants may not grow at exactly the same rate due to possible experimental error, since the dry weight matches the under/over production trend. Another interpretation would be that these mutants reprogram their fluxes differently in response to gene deletion. However, more investigation would be required to draw a clear conclusion.

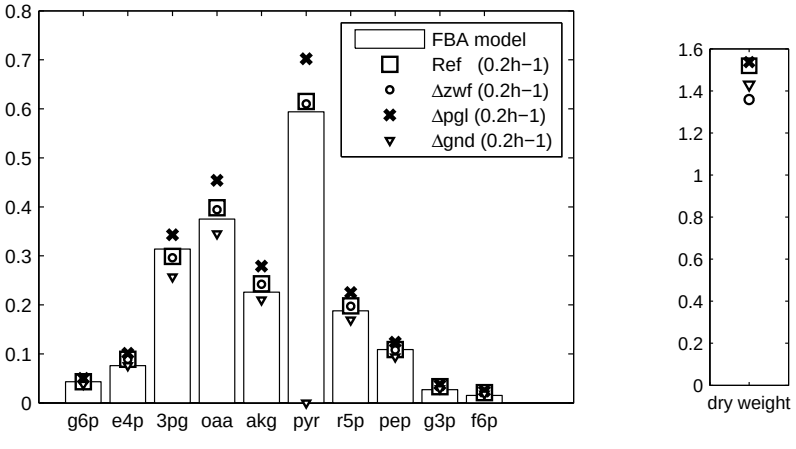

Fig. 5. Left panel is the predicted production rates for biomass precursors (mmol/gDW/h) in wild type and mutants under the same dilution rate ($0.2h^{-1}$). Right panel is the measured dry weight (g/L) for these strains.

NADPH serves as the electron donor in reductive biosynthesis. Gene deletions in the pentose phosphate pathway perturb NADPH levels and further cause oxidative damage to the mutants [8]. One possible response for these mutants is to reroute their fluxes and generate NADPH from NADP in other pathways. Our results suggest that these mutants may use another strategy for replenishing NADPH level. As shown in Table 1, all three mutants are predicted to have higher PntAB transhydrogenase activity, suggesting that mutants may replenish NADPH level by converting NADH into NADPH. This prediction supports the previous suggestion that PntAB transhydrogenase plays an important role for generating NADPH in *E. coli* [10]. The predicted PntAB flux ratio for Δzwf/Ref (1.55) agrees with previously reported Δzwf/wild-type mRNA ratio (about 1.7) [10].

In contrast to the predicted robust metabolic requirements for biomass precursors, the cofactor requirements vary. NAD and NADPH requirements show considerable increase per unit of biomass production in Δzwf and Δpgl. It is not clear how to biologically interpret the increase of redox requirements in these mutants. One possible explanation is the mutants result in redox imbalance and their regulatory networks react consequently, causing unusual ways to direct the network operation for central carbon metabolism.

In all cases analyzed here, ATP coefficients are predicted to be zero. This is because ATP synthase is present in the FBA model, and we have no information about the proton and phosphate uptake rate, Hence, the reverse ATP synthase flux is indistinguishable from the sink flux of ATP in biomass. Therefore the amount of ATP synthase reaction actually contains the ATP biomass production (but in the opposite direction). When we block (set to zero) ATP synthase in the model, the ATP biomass coefficient results equal to the absolute value of the ATP synthase flux listed in Table 1. However, this is not enough to draw a conclusion at this point on the actual ATP biomass coefficient for each strain. This issue could be examined in detail in the future, with more experimental information.

4. Discussion

We proposed an FBA-based approach to infer the biomass compositions that best describe multiple physiological states of a cell. Our results show that *E. coli* maintains robust biomass coefficients for biomass precursors in central carbon metabolism pathways under glucose supply medium, ranging from an almost glucose-starved state to a nearly unlimited glucose supply. This result suggests that *E. coli* operates its central carbon metabolism pathways with the same biomass objective, in agreement with optimal growth criteria under glucose supply medium. One should keep in mind that this result might be partially biased by the fact that experimental inference of fluxes requires fitting to a stoichiometric model that usually involves a biomass production flux as well. Our predictions for mutants indicate that there is an increase usage for the PntAB transhydrogenase flux, suggesting another potential strategy for the mutants to

compensate the less efficient NADPH production caused by single gene deletion in the pentose phosphate pathway.

Some of our flux predictions cannot be fully understood, partly due to the use of an incomplete model (as opposed to a genome-scale one) and partly due to potential experimental errors in the flux measurements. For instance, if we had information about Ubiquinone-8 associated fluxes, we could correct the missing information in the current model and improve the accuracy of oxygen uptake rate prediction. On the other hand, it would be difficult to apply a genome-scale *E. coli* FBA model in our study, since the experimental data is limited to central carbon metabolism pathways. At present, large scale flux measurements are still unavailable, due to experimental difficulties. One way to overcome this limitation would be to take advantage of other types of high throughput data. Our method is designed to incorporate not only flux measurements, but also other high throughput data as the reference vector E in Eq. (4), such as mRNA expression or protein levels for two distinct physiological states.

In ongoing work, we are applying our method to time series data such as gene expression along the cell cycle, to provide insights into the physiology of cellular growth. This will allow us to learn more about how living organisms organize their biomass requirements and manage energy or redox balance during their life cycle. The method should provide insights into how a cell allocates its metabolic resources in a time-dependent and condition-specific manner, and can be extended to integrate multiple data sources with FBA models, to shed new light on the system-level organization of metabolic networks.

Biomass component					Biomass coefficients				
Biomass Precursors	*FBA model	Ref ($0.2h^{-1}$)	Δzwf ($0.2h^{-1}$)	Δpgl ($0.2h^{-1}$)	Δgnd ($0.2h^{-1}$)	*GR01 ($0.1h^{-1}$)	*GR02 ($0.4h^{-1}$)	*GR03 ($0.5h^{-1}$)	*GR04 ($0.7h^{-1}$)
g6p	0.043	0.043	0.044	0.050	0.038	0.040	0.043	0.042	0.044
e4p	0.076	0.089	0.089	0.101	0.076	0.080	0.086	0.085	0.088
3pg	0.314	0.299	0.296	0.343	0.257	0.272	0.288	0.292	0.297
oaa	0.375	$0.398	0.394	0.454	0.345	0.364	0.384	0.387	0.396
akg	0.226	0.243	0.242	0.279	0.210	0.222	0.235	0.236	0.241
pyr	0.594	0.615	0.610	0.702	0.000	0.562	0.000	0.000	0.613
r5p	0.188	0.198	0.197	0.225	0.169	0.182	0.190	0.191	0.196
pep	0.109	$0.109	0.108	0.124	0.094	0.098	0.104	0.106	0.107
g3p	0.027	0.033	0.032	0.037	0.029	0.030	0.033	0.032	0.033
f6p	0.015	0.021	0.022	0.024	0.018	0.018	0.020	0.021	0.021
Cofactors									
#atp	11.684	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
nad	0.744	23.631	40.600	63.507	25.496	0.000	23.516	0.000	0.000
nadph	3.823	31.252	45.218	61.827	34.733	12.992	28.807	12.322	21.787
accoa	0.786	0.825	0.817	0.941	0.000	0.752	0.000	0.000	0.000

Specific reac	Ref	Δzwf	Δpgl	Δgnd	*GR01	*GR02	*GR03	*GR04
NADPH->NADH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
£NADH->NADPH	27.713	£42.818	£57.991	£32.026	9.538	25.244	9.618	15.141
©NADH->NAD	9.279	14.506	21.057	9.053	2.526	8.043	©1.146	©2.406
#ADP->ATP (ATP synthase)	-5.906	-7.114	-6.839	-8.288	-5.030	-5.695	-6.455	-9.590
EX_ac	0.000	0.000	0.000	1.245	0.000	1.389	1.597	1.371
EX_akg	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EX_co2	8.251	9.747	9.482	9.456	7.512	6.128	6.311	12.380
EX_etoh	0.000	0.013	0.000	0.000	0.000	0.000	0.058	0.044
EX_for	0.000	0.000	0.000	0.532	0.000	0.594	0.600	0.000
EX_fum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EX_glc	-2.934	-3.178	-3.361	-2.922	-2.676	-2.525	-2.653	-3.813
EX_h2o	4.252	8.711	15.300	2.103	-2.172	2.939	-3.947	-6.669
EX_h	9.924	6.903	0.955	12.474	15.098	8.902	16.164	25.450
EX_lac_D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EX_o2	-5.792	-8.765	-11.867	-6.006	-2.250	-4.785	©-1.408	©-2.786
EX_pi	-0.792	-0.788	-0.904	-0.681	-0.722	-0.763	-0.769	-0.787
EX_succ	0.000	0.000	0.000	5.2E-6	0.000	0.000	0.000	0.000

Table 1. Predicted biomass production and important fluxes (mmol/gDW/h). Negative values refer to uptake fluxes. *Normalized to the same scale with Ref column for comparison. £PntAB transhydrogenase activities increase in all three mutants. #The ATP biomass coefficient would be the absolute value of ATP synthase fluxes if we block ATP synthase reaction from the model. $One flux pair (Ref and Δzwf) fails to predict the correct value for oaa and pep (results not shown). This is due to erroneous prediction for a single reaction, PPC (phosphoenolpyruvate carboxylase), which converts pep and co2 into oaa. The deviation of PPC fluxes in two predictions of Ref biomass (Ref vs. Δzwf and Ref vs. Δgnd) matches the deviation of co2 production rates. In addition, the flux measurement for PPC has large standard deviation [5]. ©The NADH dehydrogenase flux seems to be under-predicted in GR03 and GR04 due to the unprecise oxygen uptake prediction.

Acknowledgements

The authors would like to thank Evan Snitkin, Niels Klitgord and William Riehl for discussion and critical reading of the manuscript. Linear Programming calculations were performed using the software Xpress, kindly provided by Dash Optimization under free academic license. This work is supported by research grants from the US National Institute of Health (5012846-00) and the US Department of Energy (DE-FG02-07ER64388 and DE-FG02-07ER64483).

References

[1] Bonarius, H.P.J., Hatzimanikatis, V., Meesters, K.P.H., *et al*., Metabolic flux analysis of hybridoma cells in different culture media using mass balances, *Biotechnol Bioeng*, 50(3):299-318, 1996.
[2] Burgard, A.P. and Maranas, C.D., Optimization-based framework for inferring and testing hypothesized metabolic objective functions, *Biotechnol Bioeng*, 82(6):670-7, 2003.
[3] DeRisi, J.L., Iyer, V.R. and Brown, P.O., Exploring the metabolic and genetic control of gene expression on a genomic scale, *Science*, 278(5338):680-6, 1997.
[4] Gianchandani, E.P., Oberhardt, M.A., Burgard, A.P., *et al*., Predicting biological system objectives de novo from internal state measurements, *BMC Bioinformatics*, 9(43, 2008.
[5] Ishii, N., Nakahigashi, K., Baba, T., *et al*., Multiple high-throughput analyses monitor the response of *E. coli* to perturbations, *Science*, 316(5824):593-7, 2007.
[6] Johnston, M. and Carlson, M., *The Molecular Biology of the Yeast Saccharomyces: Gene Expression*, 1992.
[7] Knorr, A.L., Jain, R. and Srivastava, R., Bayesian-based selection of metabolic objective functions, *Bioinformatics*, 23(3):351-7, 2007.
[8] Minard, K.I. and McAlister-Henn, L., Antioxidant function of cytosolic sources of NADPH in yeast, *Free Radic Biol Med*, 31(6):832-43, 2001.
[9] Palsson, B.O., *Systems Biology: Properties of Reconstructed Networks*, Cambridge University Press, 2006.
[10] Sauer, U., Canonaco, F., Heri, S., *et al*., The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*, *J Biol Chem*, 279(8):6613-9, 2004.
[11] Schuetz, R., Kuepfer, L. and Sauer, U., Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli, *Mol Syst Biol*, 3(119, 2007.
[12] Segre, D., Vitkup, D. and Church, G.M., Analysis of optimality in natural and perturbed metabolic networks, *Proc Natl Acad Sci U S A*, 99(23):15112-7, 2002.
[13] Uygun, K., Matthew, H.W. and Huang, Y., Investigation of metabolic objectives in cultured hepatocytes, *Biotechnol Bioeng*, 97(3):622-37, 2007.
[14] van Gulik, W.M. and Heijnen, J.J., A metabolic network stoichiometry analysis of microbial growth and product formation, *Biotechnol Bioeng*, 48(6):681-698, 1995.
[15] Varma, A., Boesch, B.W. and Palsson, B.O., Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates, *Appl Environ Microbiol*, 59(8):2465-73, 1993.