CHAPTER FOUR

Application of Metabolic Flux Analysis in Metabolic Engineering

Sang Yup Lee,[*,†,‡,§] Jong Myoung Park,[*,†] *and* Tae Yong Kim[*,‡]

Contents

1. Introduction 68
   1.1. Systems metabolic engineering and metabolic flux analysis 68
   1.2. ^{13}C-based flux analysis 68
   1.3. Constraints-based flux analysis 70
2. General Structure of Constraints-Based Flux Analysis 70
3. Algorithms of Metabolic Flux Analysis 73
   3.1. Flux balance analysis 75
   3.2. Identifying gene targets for engineering strain development: Gene knockout 77
   3.3. Identifying gene targets for engineering strain development: Up- or downregulation of Genes 79
   3.4. Identifying gene targets for engineering strain development: Foreign genes insertion 82
   3.5. Identifying gene targets for engineering strain development: Metabolite essentiality 82
   3.6. Accurately describing cellular physiology: Incorporation of experimental data and physiological properties into the *in silico* model 83
4. Concluding Remarks 85
Acknowledgments 86
References 87

Abstract

Metabolic flux analysis (MFA) is an important analytical technique to quantify intracellular metabolic fluxes as a consequence of all catalytic and transcriptional interactions. In systems metabolic engineering, MFA has played important

* Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 Program), KAIST, Daejeon, Republic of Korea
† BioProcess Engineering Research Center, Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, KAIST, Daejeon, Republic of Korea
‡ Bioinformatics Research Center, KAIST, Daejeon, Republic of Korea
§ Department of Bio and Brain Engineering, KAIST, Daejeon, Republic of Korea

Methods in Enzymology, Volume 498
ISSN 0076-6879, DOI: 10.1016/B978-0-12-385120-8.00004-8

role to understand cellular physiology under particular conditions and predict its metabolic capability after genetic or environmental perturbations. Two methods using optimization procedure, ^{13}C-based flux analysis and constraints-based flux analysis, have been used generally on the basis of stoichiometry of metabolic reactions and mass balances around intracellular metabolites under pseudo-steady state assumption. Practically, MFA has been applied to generate new knowledge on the biological system, analyze cellular physiology system-wide, and consequently design metabolic engineering strategies at a systems-level. In this chapter, we study the basic principle of MFA (more particularly constraints-based flux analysis), inspect the characteristics of several *in silico* algorithms developed for system-wide analysis of cellular metabolic fluxes, and discuss their applications.

1. Introduction

1.1. Systems metabolic engineering and metabolic flux analysis

Systems metabolic engineering has been provided as a new paradigm for generating new knowledge on biological systems and for systematically designing novel strategies to develop improved strain (Joyce and Palsson, 2006; Lee *et al.*, 2005b, 2007; Park and Lee, 2008; Park *et al.*, 2007). In systems metabolic engineering, metabolic flux analysis (MFA) has played an important role in understanding cellular physiology and predicting its metabolic capability under specified environmental or genetic conditions (Orth *et al.*, 2010; Raman and Chandra, 2009; Sauer, 2006; Zamboni and Sauer, 2009). MFA is an analytical and powerful technique using optimization procedure to quantify intracellular metabolic fluxes from all known catalytic and transcriptional interactions. MFA is based on the stoichiometry of the metabolic reactions and the mass balances around intracellular metabolites under pseudo-steady state assumption. Two methods have been used to study the metabolic flux in a biological system: ^{13}C-based flux analysis and constraints-based flux analysis.

1.2. ^{13}C-based flux analysis

The ^{13}C-based flux analysis utilizes an isotope-labeled carbon substrate and allows the determination of intracellular fluxes in metabolic networks by analyzing ^{13}C enrichment patterns of metabolites with nuclear magnetic resonance (NMR) or gas chromatography–mass spectrometry (GC–MS) (Sauer, 2006; Zamboni and Sauer, 2009). The ^{13}C-labeled substrates are fed to growing cells until the isotope-labeled carbon is distributed throughout the metabolic network. The measured ^{13}C-isotope pattern data and

additional physiological data during cultivation, including exchange fluxes (uptake rate and production rate) determined from time courses of extracellular metabolite concentrations and biomass composition data, are simultaneously integrated with computational analysis. The intracellular fluxes are then estimated by fitting iteratively the simulated fluxes in stoichiometric models to the measured data. The difference between simulated and measured labeled pattern is minimized (Sauer, 2004, 2006; Wiechert, 2001).

Typically, ^{13}C-based flux analysis has been used to understand the physiological status of a cell by quantifying intracellular fluxes under a particular condition (Al Zaid Siddiquee *et al.*, 2004; Li *et al.*, 2006; Peng *et al.*, 2004; Schmidt *et al.*, 1999; Zhao *et al.*, 2004a,b). ^{13}C-based flux analysis also has been used to discover and quantify the *in vivo* operation of unusual pathways within complex metabolic networks and to elucidate the pathways in less-characterized species (Rabinowitz, 2007; Risso *et al.*, 2008; Sauer, 2006; Tang *et al.*, 2007). Other applications that ^{13}C-based flux analysis has been utilized for are the elucidation of mechanisms in network-wide balancing of intracellular components, for example, energy and redox balancing, and the demonstration of the role of unfamiliar pathways in the metabolic network (Fuhrer and Sauer, 2009; Peyraud *et al.*, 2009; Zamboni and Sauer, 2009). Combining ^{13}C-based flux analysis and other data pertaining to the network being investigated has enabled characterization of condition-dependent regulatory circuits that ultimately govern the metabolic phenotype (Ishii *et al.*, 2007; Nanchen *et al.*, 2008; Tang *et al.*, 2009; Tannler *et al.*, 2008; Zamboni and Sauer, 2009).

Experimental fluxes based on ^{13}C-based flux analysis were used to predict cellular physiology using the genome-scale metabolic model, with relatively high accuracy, by constraining the flux solution space (Herrgard *et al.*, 2006a; Kim and Lee, 2006; Sauer, 2006) and evaluation of the model predictions (Park *et al.*, 2010; Segre *et al.*, 2002). This has allowed members of the biotechnology community to utilize ^{13}C-based flux analysis for metabolic engineering (e.g., isoprenoid production in *Escherichia coli*; Kizer *et al.*, 2008), drug development (e.g., dihydrofolate reductase inhibitor in *E. coli*; Kwon *et al.*, 2008), and the identification of functional side effects of drugs (Schneider *et al.*, 2009) because of the extensive perspectives on cellular energetics and network-wide balancing provided by ^{13}C-based flux analysis (Zamboni and Sauer, 2009). In practice, despite relatively accurate estimation of intracellular fluxes, ^{13}C-based flux analysis typically focuses on small-scale metabolic network (i.e., central metabolism) rather than the entire, or genome-scale, metabolic network because of difficulties in experimentation and subsequent computational calculations required for large-scale metabolic models, limiting its applications for large-scale analysis (Kim *et al.*, 2008a; Sauer, 2006).

1.3. Constraints-based flux analysis

Constraints-based flux analysis is a general mathematical method using optimization-based simulation techniques to analyze cellular metabolism under a specified environmental or genetic condition and predict metabolic capability when the specified conditions are perturbed (Park *et al.*, 2009). To implement constraints-based flux analysis, a stoichiometric model needs to be first constructed based on genomic information, databases, and literatures. As the genomes of increasing number of organisms have been completely sequenced, *in silico* (means "performed on computer or via computer simulation.") genome-scale metabolic models have been constructed for several organisms in the domains of bacteria, archaea, and eukarya to use them for exploring their metabolic characteristics at a systems-level (Duarte *et al.*, 2007; Durot *et al.*, 2009; Feist and Palsson, 2008; Joyce and Palsson, 2006; Kim *et al.*, 2008a) (Fig. 4.1).

Reconstruction of the *in silico* genome-scale metabolic model begins with utilizing the genome annotation to generate a collection of metabolic reactions and the stoichiometric coefficients of the metabolites, giving a set of linear mass balance equations for cellular metabolites describing the cellular metabolism (Davidsen *et al.*, 2010; Kanehisa *et al.*, 2010). This collection of equations forms the foundation of the metabolic network. Gaps in the metabolic network, due to insufficient data or characterization in the genome annotation, are filled in, and the errors are corrected based on knowledge from literature, databases, and experiments. The *in silico* genome-scale metabolic model is then validated by comparing simulation results with actual experimental data. If the simulation results differ greatly from experimental observations, the metabolic model should be refined iteratively until the discrepancies are resolved (Fig. 4.1). After the *in silico* genome-scale metabolic model has been validated by iterative processes, constrains-based flux analysis can be utilized using appropriate objective functions (e.g., maximization of cell growth rate) and constraints that restrict the solution space of the model to exclude incorrect or infeasible metabolic states. In this chapter, we focus on the methods of constraints-based flux analysis using the genome-scale metabolic model and its applications in systems metabolic engineering for strain improvement.

2. General Structure of Constraints-Based Flux Analysis

In silico algorithms developed for the genome-scale metabolic models, to date, are based on optimization techniques with various constraints applied for improving the accuracy of the simulation (Park *et al.*, 2009). Before inspecting each *in silico* algorithm, a general structure of constraints-based flux analysis,

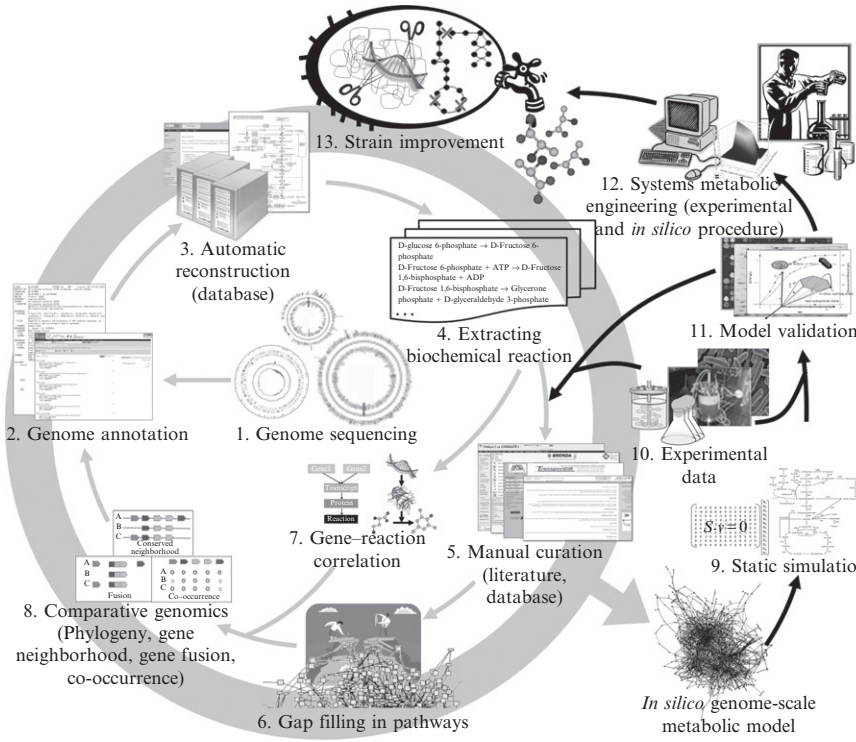

Figure 4.1 Procedure for the reconstruction of *in silico* genome-scale metabolic model and its application to metabolic engineering. (A) Automatic reconstruction of metabolic network based on genome sequence and annotation data (1-2-3-4). (B) Manual curation and fine-tuning of the metabolic network using literatures, databases, gene/reaction correlation, and comparative genomics to fill gaps and correct errors in the pathways (4-7-8-2 or 5-6-8-2). (C) Validation of the metabolic model in comparison with experimental data (9-10-11). The biomass composition determined by the experiments is applied to the model. If the simulation results do not correspond with experimental data, the model needs to be refined further by an iterative process until the differences between predictions and experiments are resolved. (D) Systems metabolic engineering for strain improvement by combining experimental and *in silico* procedures (12-13). The gray arrows indicate the procedures for the construction of the model. The black arrows indicate the procedure for the simulation and validation of the model and its application.

based on optimization, is worth understanding. The general structure of constraints-based flux analysis is the basic principle in constructing *in silico* algorithms or metabolic models. It is able to cover all classes of mathematical optimization methods. The structure of constraints-based flux analysis consists of two parts: objective functions and the constraints to metabolic fluxes in the metabolic model (Lee and Papoutsakis, 1999; Stephanopoulos *et al.*, 1998).

Constraints are the conditions that must be satisfied while solving for the optimal solution to the metabolic network by maximizing/minimizing the

objective function(s). Constraints can be in the form of either equality or inequality statements. A general form of the constraints is as follows:

$$\alpha y_j \leq v_j \leq \beta y_j, y_j \in \{0, 1\} \tag{4.1}$$

where v_j is a continuous variable, y_j is a discrete variable having a binary value of 0 or 1, and α and β are constants that represent the upper and lower limits, respectively.

Constraints-based flux analysis solves for an optimal solution to the metabolic network by maximizing or minimizing an objective function(s) subject to the constraints defined for the independent variables. More than one objective function can be selected and solved for. Solving the system of equations defining the metabolic network proceeds in the following manner:

1. Determine the decision (or control) variables
2. Formulate all objectives representing the purpose of decision maker
3. Formulate constraints
4. Maximize or minimize the objective function(s) subject to the constraints

Objective function:

$$\text{Maximize/Minimize} \quad Z(x) = \left(c_1 x_1^{m_1} + c_2 x_2^{m_2} + \ldots + c_n x_n^{m_n}\right)^k, \quad \text{for all } n \tag{4.2}$$

Constraints:

$$\text{Subject to} \quad \frac{dX_i}{dt} = S_{ij} v, \quad \alpha_j \leq v_j \leq \beta_j \tag{4.3}$$

The objective function $Z(x)$ is a mathematical expression of the goals for the system desired by the user, where c, m, and k are constants and x_n is a variable that can be designated as a vector representing the fluxes of metabolic reactions, the number of significant metabolic flux changes represented by a binary variable, or any other characteristic of interest. The type of simulation is determined by the form of objective function $Z(x)$. The system of equations is linear or nonlinear according to the values of m and k and consequently determines the method used to solve the system (i.e., linear programming (LP) or nonlinear programming). The number of objective functions also determines what type of problem the system becomes, whether it is a single, or multiple objective function system.

In silico genome-scale metabolic model is composed of metabolic reactions that define the stoichiometric conversion of substrate metabolites into various intracellular metabolites that are precursors to different components important for cellular function. Mass balances can be set up as Eq. (4.3),

where the difference between consumption rate and production rate of a specific metabolite is equal to the change rate of metabolite concentration. The subscripts i and j represent the indices of metabolites and reactions that the metabolite participates in, respectively. X denotes the vector representing the concentrations of metabolites. The stoichiometric matrix S is an $m \times n$ matrix where m is the total number of metabolites and n is the total number of reactions in the metabolic network that is being described. v is a vector of the fluxes for the reactions that consume and produce the metabolites. The stoichiometric coefficients are negative if the metabolite i is a substrate of the reaction and positive if the metabolite i is a product of the reaction. The fluxes in v are subject to lower and upper bounds, α and β, respectively (Fig. 4.2).

To simplify the process of solving this system of equations, the pseudo-steady state assumption is applied to eliminate the time derivative from Eq. (4.3), reducing it to a system of linear equations in the form of Eq. (4.4) (Lee and Papoutsakis, 1999; Stephanopoulos *et al.*, 1998) (Fig. 4.2A). This pseudo-steady state assumption is based on the observation that the changes in intracellular concentrations of the metabolites are infinitesimally small compared to the overall timescale of cellular functions, such as cell division.

$$S_{ij}v_j = 0, \quad \alpha_j \leq v_j \leq \beta_j \tag{4.4}$$

3. Algorithms of Metabolic Flux Analysis

Constraints-based flux analysis of the *in silico* genome-scale metabolic model allows us to investigate the metabolic status of the cell under specified conditions and to rapidly predict and evaluate phenotypes that would result from genetic and/or environmental perturbations to the cell. This has been used with great success in improving the metabolic capability for the overproduction of the desired product (Kim *et al.*, 2008b; Park *et al.*, 2009). Based on the general structure of constraints-based flux analysis, several *in silico* algorithms have been developed to tailor the objective functions or constraints according to the desired goals such that the cellular physiology can be accurately described for a specific condition and identify targets to engineer for strain improvement (Fig. 4.3). Recently, constraints-based flux analysis was used for identifying metabolic engineering targets for the overproduction of industrially important products, including petroleum-alternative biochemicals, amino acids, biopolymers, and biofuels (Kim *et al.*, 2008a,b; Park *et al.*, 2009; Raman and Chandra, 2009), and for identifying drug targets in pathogens (Hu *et al.*, 2007; Jamshidi and Palsson, 2007; Kim *et al.*, 2010; Lee *et al.*, 2009; Yeh *et al.*, 2004).

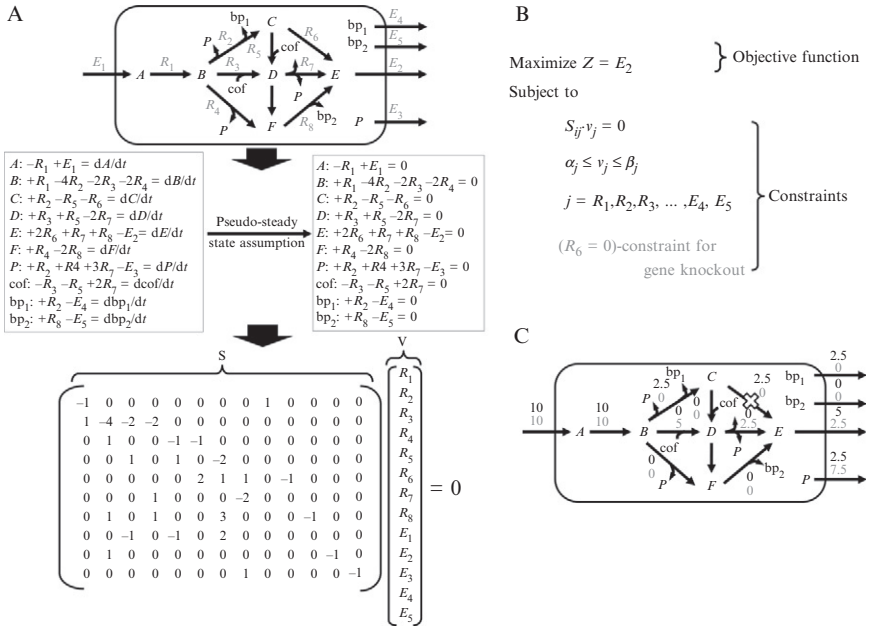

Figure 4.2 Construction of a metabolic model expressed by a stoichiometric matrix and its simulation using constraints-based flux analysis. (A) An example metabolic model consisting of 13 reactions and 15 metabolites (10 internal and 5 external metabolites). Mass balances for each metabolite are set up as differential equations, where the difference between consumption rate and production rate of a specific metabolite is equal to the change rate of metabolite concentration. Based on the pseudo-steady state assumption, the time derivative can be eliminated, giving a set of linear equations. The stoichiometric coefficients for substrate and product of a reaction are negative and positive, respectively. The stoichiometric matrix S is an $m \times n$ matrix where m is the number of metabolites, and n is the number of reactions. v is a vector representing the fluxes of reactions that consume and produce the metabolites. In this model, internal reactions are represented by R, and reactions related with external metabolites are represented by E. The subscripts i and j represent the indices of metabolites and reactions. (B) Optimization based on constraints-based flux analysis is formulated with an objective function(s) subject to mass balances and additional constraints. v is subject to lower and upper bound constraints, represented as α and β, respectively. In this example, to investigate the metabolic capability after gene knockout, the metabolic flux of reaction R_6 is constrained to zero while maximizing an objective function E_2. (C) The distribution of metabolic fluxes calculated by constraints-based flux analysis for wild type and knockout mutant (i.e., $R_6 = 0$) strains is shown. Upper black and lower gray values indicate the flux values of wild type and knockout mutant, respectively. The deletion of reaction R_6 in this exampled mutant increases the production rate of metabolite P but decreases the flux rate of objective reaction E_2, compared with those of wild-type strain. The units of fluxes are mmol/gDCW/h.

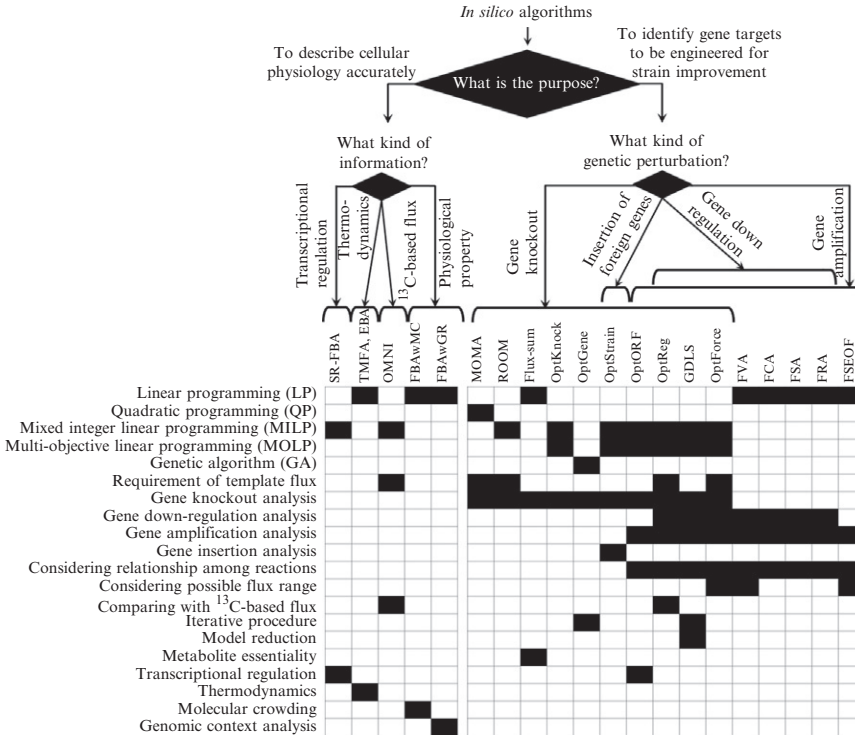

Figure 4.3 **Flowchart for the simulations based on constraints-based flux analysis of several representative *in silico* algorithms. The black box denotes a particular property that corresponds to an *in silico* algorithm.**

3.1. Flux balance analysis

Flux balance analysis (FBA) is a widely used and basic approach of constraints-based flux analyses. FBA quantifies the intracellular flux distribution of a metabolic network, as represented in Eq. (4.4), by optimizing a linear objective function with LP. Additional constraints can be applied to represent perturbations that can be made to the system and thereby allowing the user to predict the change of metabolic fluxes in response to that perturbation (Fig. 4.2B and C). For example, the change in physiology to gene knockout can be investigated by constraining the metabolic flux of the corresponding reaction to the gene that is to be knocked out to zero. Another example is to apply inequality constraints to the flux values (e.g., $v_j \geq$ or $\leq$ desired level of a flux) to represent intervention of gene expression (i.e., down- or upregulation of gene expression). To eliminate unrealistic metabolic fluxes, the flux solution space of *in silico* metabolic model is

restricted by utilizing constraints determined from experimentally measured fluxes or physiological data. This limits the solution space to fluxes that are realistically within the cell's capacity. These constraints are applied as inequality or equality constraints (e.g., $v_j \geq, \leq$, or $=$ measured value).

Generally, the metabolic reaction representing biomass formation, which is based on experimental measurements of biomass composition under various cultivation conditions, has been used as the objective function in FBA of the metabolic network. The selection of the biomass formation reaction as the objective function is based on the assumption that the cell seeks to maximize cellular growth to ensure survival (Orth *et al.*, 2010; Raman and Chandra, 2009; Schuetz *et al.*, 2007; Smallbone and Simeonidis, 2009; Varma *et al.*, 1993). The maximization of growth rate in FBA is useful in predicting the essentiality of gene/reaction and the robustness of a cell under specific genetic or environmental conditions (Edwards and Palsson, 2000; Kauffman *et al.*, 2003). The essentiality of a reaction and robustness of a cell can be explored by observing the change in the objective value for the biomass formation in response to variations in the flux of a particular reaction (Edwards and Palsson, 2000; Kauffman *et al.*, 2003; Orth *et al.*, 2010). For example, if the objective value for the biomass formation is zero when the flux of a particular reaction is constrained to zero to simulate gene deletion, then the relevant gene or reaction is determined to be essential. For the development of novel drugs to kill a pathogenic microorganism, gene/reaction essentiality analysis on the metabolic network of that microorganism can provide useful information for identifying drug targets. To identify drug targets, essential genes or reactions identified through FBA are further characterized by sequence analyses and structural studies (Hu *et al.*, 2007; Jamshidi and Palsson, 2007; Raman and Chandra, 2009; Yeh *et al.*, 2004).

Additionally, by applying constraints to external metabolites, such as substrate uptake rate, oxygen presence to simulate aerobic/anaerobic condition, and biochemical secretion rate, FBA also can be used for quantifying the cellular growth rate under different environmental conditions (Edwards *et al.*, 2001; Oberhardt *et al.*, 2009), investigating byproduct secretion under increasingly anaerobic conditions (Varma *et al.*, 1993), evaluating carbon source utilization capacity (Oberhardt *et al.*, 2009; Orth *et al.*, 2010; Varma *et al.*, 1993), and identifying the optimal growth media composition (Song *et al.*, 2008). The capability of carbon source utilization and the optimal media composition for growth can be examined through varying the constraints related to carbon source uptake and media composition and observing the effects on the growth rate (i.e., viable, nonviable, or maximal growth rate).

However, the objective function for FBA is not restricted to only biomass formation and other objective functions have been utilized to investigate other characteristics of the metabolic network, including maximization of ATP or reducing power (Ramakrishna *et al.*, 2001; Schuetz *et al.*, 2007), and maximization of a particular biochemical production

(Hong *et al.*, 2003; Kauffman *et al.*, 2003). This allows FBA the flexibility to investigate a wide range of target phenotypes. In investigating the production capability of a desired biochemical and identifying alternative metabolic pathways that lead to the production of desired biochemical, the *in silico* theoretical maximum yield is evaluated (i.e., maximizing the production rate of the target biochemical in FBA; Hong *et al.*, 2003). FBA can also be used to calculate the yields of important cofactors, such as ATP, NADH, or NADPH (Orth *et al.*, 2010; Varma *et al.*, 1993).

FBA has also been used for the refinement of *in silico* metabolic models by filling gaps in the metabolic network due to incomplete information in the genome annotation and the databases. Gaps in the metabolic network appear where the predicted results are inconsistent with experimental data. Analysis of the results generated using FBA can identify the missing reactions that are not annotated in the genome and are required in the metabolic network to reconcile the disagreements between predictions and experiments. By filling in these gaps for the model refinement, the genome annotation is concurrently updated (Oberhardt *et al.*, 2008, 2009; Raman and Chandra, 2009; Reed *et al.*, 2006).

FBA and FBA-based approaches have been utilized in metabolic engineering to identify gene targets with the goal of improving the production yield of a desired biochemical. This is accomplished by selecting targets which increase the availability of metabolic precursors and cofactor balancing by redirecting the metabolic fluxes through fluxes that generate the desired biochemical (Park *et al.*, 2009; Raman and Chandra, 2009). Gene knockout approaches, such as OptKnock and its derivatives, have been widely used to identify target genes that will block competing fluxes and funnel the flux toward the overproduction of biochemicals (Burgard *et al.*, 2003; Lee *et al.*, 2005a; Pharkya *et al.*, 2003). In addition to gene knockout approaches, other FBA-based approaches have been developed to analyze features of metabolic network and the relationship among different reactions to each other in the metabolic network. These approaches include flux variability analysis (FVA), flux coupling analysis (FCA), flux sensitivity analysis (FSA), and flux response analysis (FRA) (Burgard *et al.*, 2004; Jung *et al.*, 2010; Lee *et al.*, 2007; Mahadevan and Schilling, 2003; Price *et al.*, 2004) and will be discussed later in detail (Fig. 4.3). By considering the relationship of the metabolic reaction to the desired biochemical, regulatory targets can be identified to improve the production of the desired biochemical.

3.2. Identifying gene targets for engineering strain development: Gene knockout

In metabolic engineering, gene knockout is the most common and important tool that generates strategies leading to the overproduction of the desired biochemical by redirecting metabolic fluxes and redesigning the

metabolic pathways of the host strain (Fig. 4.2B and C). However, there is a problem in trusting strategies generated from knockout simulations to accurately reflect *in vivo* knockout phenotypes. The problem is that biological systems do not instantly attain the optimal phenotype that is displayed from FBA. The cell requires an adjustment to the perturbation that is introduced to its metabolic network. To account for this adjustment to the metabolic network, algorithms describing the physiological characteristics of a cell after gene knockout perturbations were developed: minimization of metabolic adjustment (MOMA) and regulatory on/off minimization (ROOM; Segre *et al.*, 2002; Shlomi *et al.*, 2005) (Fig. 4.3). These algorithms require a template flux distribution, which calculates the flux distribution of mutant. Typically, the template flux distribution is the flux distribution of wild-type strain or base strain for the next stage of engineering. MOMA assumes that the metabolic fluxes of the metabolic network in the mutant go through a minimal flux redistribution in relation to the wild type. Therefore, the objective function in MOMA finds a unique flux distribution for the mutant network that is closest to a given template flux distribution using Euclidian norm with quadratic programming (QP; Segre *et al.*, 2002). ROOM also utilizes a different objective function which looks for a flux distribution that minimizes the number of significant flux changes from the template flux distribution using mixed integer linear programming (MILP; Shlomi *et al.*, 2005). Comparing the two algorithms reveals that MOMA fluctuates most of metabolic flux values in the metabolic network in relation to the template flux distribution. However, ROOM minimizes the number of flux changes compared with the template flux distribution.

MOMA was utilized to identify gene knockout targets to develop strains capable of enhanced production of lycopene (Alper *et al.*, 2005; Choi *et al.*, 2010), L-valine (Park *et al.*, 2007), and polylactic acid (Jung *et al.*, 2010) in *E. coli* and sesquiterpene in *Saccharomyces cerevisiae* (Asadollahi *et al.*, 2009). Particularly, sequential and iterative optimization approach using MOMA, whereby single gene knockouts are investigated in the genetic background of mutants identified from previous iterations, was used to identify knockout target genes for the overproduction of lycopene and L-valine in *E. coli* (Alper *et al.*, 2005; Park *et al.*, 2007). For the production of sesquiterpene in *S. cerevisiae*, the effects of gene knockouts were evaluated using MOMA as objective function and OptGene as simulation framework (Asadollahi *et al.*, 2009). ROOM was utilized to show improved flux predictions in pyruvate kinase (*pyk*) knockout *E. coli* and good performances for the prediction of gene essentiality in *S. cerevisiae* compared with either FBA or MOMA (Shlomi *et al.*, 2005).

Although the enhanced production of the desired biochemical through genetic modifications is the desired outcome, increasing the production rate of the desired biochemical often negatively affects the cellular growth rate,

and *vice versa*. To resolve this dilemma, the bi-level optimization framework using MILP, OptKnock, was developed (Burgard *et al.*, 2003) (Fig. 4.3). OptKnock allows the user to find a set of gene knockout targets that increase the fluxes toward the production of the desired biochemical, while biomass precursors are simultaneously generated to maintain a sufficient level of growth. OptKnock has been utilized to suggest gene knockout strategies for the production of amino acids (Pharkya *et al.*, 2003), lactic acid (Fong *et al.*, 2005; Hua *et al.*, 2006), succinic acid (Burgard *et al.*, 2003), and 1,3-propanediol (Burgard *et al.*, 2003) in *E. coli*. Based on the OptKnock framework, OptGene was developed to identify target knockout genes for optimization of a desired biochemical production using genetic algorithm, instead of MILP, to reduce computational time (Patil *et al.*, 2005) (Fig. 4.3). A population of several genotypes is initiated by assigning an on/off status for each gene, and each individual genotype is then scored for their fitness by using FBA, MOMA, ROOM, or any other algorithm. After scoring their fitness, the best individual is selected for the generation of a new population by applying random genetic modifications, crossovers, and mutations. This cycle of evolution is repeated until the performance of the mutant achieves a satisfactory performance. OptGene has suggested potential gene knockout targets for the improved production of vanillin, glycerol, succinic acid, and sesquiterpene in *S. cerevisiae* (Asadollahi *et al.*, 2009; Patil *et al.*, 2005). The other algorithms, such as OptStrain (Pharkya *et al.*, 2004), OptReg (Pharkya and Maranas, 2006), OptORF (Kim and Reed, 2010), and OptForce (Ranganathan *et al.*, 2010) using the OptKnock framework as a starting point and a heuristic algorithm called genetic design through local search (GDLS) (Lun *et al.*, 2009) to reduce computational burden, can also be applied to predict gene knockout targets (Fig. 4.3).

3.3. Identifying gene targets for engineering strain development: Up- or downregulation of Genes

Increasing or decreasing gene expression levels to increase the production of the target biochemical has been widely recognized in the community of metabolic engineering (Jensen and Hammer, 1998; Koffas *et al.*, 2003). Determining whether a gene should be upregulated or downregulated is based on the relationship among metabolic reactions and the response of reactions according to varying the flux of a specific reaction (e.g., production of desired biochemical). Several different algorithms based on LP have been developed to investigate the relationship between the metabolic reactions in metabolic network to the characteristic of interest (Burgard *et al.*, 2004; Jung *et al.*, 2010; Lee *et al.*, 2007; Mahadevan and Schilling, 2003; Price *et al.*, 2004) (Fig. 4.3). FVA investigates the possible flux ranges of reactions (i.e., flux solution space of the metabolic reactions) by examining the maximal and minimal fluxes for each reaction (Bushell *et al.*, 2006;

Khannapho *et al.*, 2008; Puchalka *et al.*, 2008). FVA was utilized to identify inactive or infeasible reactions and classify the reactions according to their simulated behaviors in the *in silico* genome-scale metabolic models by considering a minimal and maximal flux values (Faria *et al.*, 2010; Feist *et al.*, 2007; Lun *et al.*, 2009; Teusink *et al.*, 2006). FVA can also analyze the changes of flux ranges of reactions after the flux of a metabolic reaction is forced to up or down and was applied to identify gene targets to be engineered for the production of biochemicals, such as succinic acid, 1-butanol, and lycopene in *E. coli* (Choi *et al.*, 2010; Ranganathan and Maranas, 2010; Ranganathan *et al.*, 2010). FCA examines the correlations for every pair of metabolic fluxes in the metabolic network (Bundy *et al.*, 2007; Burgard *et al.*, 2004; Puchalka *et al.*, 2008). FCA was used to analyze the coupled reaction sets *in silico* genome-scale metabolic models of *E. coli*, *Pseudomonas putida*, *Helicobacter pylori*, *S. cerevisiae*, and *Homo sapiens* (Burgard *et al.*, 2004; Duarte *et al.*, 2007; Pal *et al.*, 2005; Puchalka *et al.*, 2008). FSA explores the change in the objective function flux in response to the flux changes of other metabolic reactions (Delgado and Liao, 1997; Price *et al.*, 2004). FSA suggested metabolic engineering strategies for improving the production of biochemicals, such as acetate, phenylalanine, and erythromycin precursors in *E. coli* (Delgado and Liao, 1997; Gonzalez-Lergier *et al.*, 2006; Wahl *et al.*, 2004) and was applied to estimate the usefulness of a metabolite toward increasing the growth rate (Finley *et al.*, 2010; Grafahrend-Belau *et al.*, 2009). FRA examines the response of the flux values for target reactions (i.e., desired biochemical production rate and cell growth rate) to the variation in the fluxes of other metabolic reactions. FRA was applied to identify metabolic engineering strategies to increase the production of L-threonine, malic acid, and polylactic acid in *E. coli* (Jung *et al.*, 2010; Lee *et al.*, 2007; Moon *et al.*, 2008). FRA was employed to identify targets, in conjunction with MOMA, to develop PLA-overproducing *E. coli* strain (Jung *et al.*, 2010).

The method to identify gene amplification targets called flux scanning based on enforced objective flux (FSEOF) scans the changes of all the metabolic fluxes in response to the enhancement of the flux toward the desired biochemical (Choi *et al.*, 2010) (Fig. 4.3). FSEOF selects the reactions, as amplification targets, representing fluxes that increase when the flux toward the production of desired biochemical is forced to increase. This method was validated by identifying amplification targets that improved the production of lycopene in *E. coli* (Choi *et al.*, 2010).

To consider simultaneous applications of multiple up- or downregulations and elimination of genes, OptReg and OptForce, derivatives of OptKnock, were developed (Pharkya and Maranas, 2006; Ranganathan *et al.*, 2010) (Fig. 4.3). OptReg requires the determination of initial steady-state fluxes for all metabolic reactions (Pharkya and Maranas, 2006). The fluxes of metabolic reactions are defined as repressed or activated when the fluxes are sufficiently higher or lower compared to the

corresponding initial steady-state fluxes. OptReg suggested metabolic engineering strategies for the production of ethanol and succinic acid in *E. coli* (Pharkya and Maranas, 2006).

OptForce identifies possible engineering interventions by comparing the maximal ranges of flux variability for all metabolic reactions in the wild-type metabolic network, with those in a hypothetical overproducing networks (Ranganathan *et al.*, 2010). The fluxes of overproduction targets in the overproducing networks are forced to maintain the desired limits the fluxes can achieve as additional constraints (i.e., $v^{\text{target}} > v^{\text{desired}}$). By doing so, OptForce identifies the sets that must be changed to achieve the prespecified overproducing networks (i.e., MUST sets) for the overproduction of desired biochemical and classifies metabolic reactions into MUST sets including reactions whose flux value must increase, decrease, and become zero: MUST^{U}, MUST^{L}, and MUST^{X}, respectively. Based on these sets, OptForce subsequently extracts a minimal set of fluxes (i.e., FORCE set) that must be modified to obtain the desired phenotype. OptForce was employed to examine the metabolic network for the increased production of succinic acid and 1-butanol in *E. coli* (Ranganathan and Maranas, 2010; Ranganathan *et al.*, 2010).

OptORF, a bi-level optimization method, was developed to design optimal gene knockout and amplification strategies for strain improvement by integrating transcriptional regulatory networks and metabolic networks (Kim and Reed, 2010) (Fig. 4.3). OptORF identifies the targets for modification that maximize biochemical production along with maximizing cellular growth rate using transcriptional regulatory constraints based on Boolean logics (e.g., AND, OR, TRUE, and FALSE) to allow for transcriptional regulation in the metabolic network. OptORF was implemented for producing ethanol and higher alcohol (e.g., isobutanol) in *E. coli* (Kim and Reed, 2010).

Developing multiple genetic manipulation strategies is necessary in maximizing the capabilities of the metabolic network in the production of the desired biochemical and to achieve improved cellular performance. However, as the scale of metabolic model and the number of genetic manipulations employed increase, the calculation time and required computational resources needed to formulate these strategies increase exponentially. To resolve these difficulties, GDLS was developed and involves a process of first, model reduction and then an iterative cycle of sequential gene knockout simulation (Lun *et al.*, 2009) (Fig. 4.3). GDLS framework starts with the reduction of the metabolic model into a smaller model that is equivalent in performance to the original model. This step removes dead-end reactions and linked reactions that are not essential in the functioning of the metabolic model. Then, GDLS randomly selects an initial set of genetic manipulations and yields a recombinant network. Subsequently, GDLS searches for the best additional genetic manipulation that improves the

phenotype using MILP. The best perturbed network selected is used as the start point in the next round of the search. This search cycle continues until no further improvements are found within the allowed range of genetic manipulations. This GDLS framework can operate any other optimization algorithms, such as OptKnock or OptReg, in each cycle to find target genes for manipulation. The GDLS algorithm was applied to the production of acetate and succinic acid in *E. coli*, and its performance was compared to the global search method used by OptKnock (Lun *et al.*, 2009).

3.4. Identifying gene targets for engineering strain development: Foreign genes insertion

To confer nonnative functionality into a host organism to achieve a desired phenotype, the insertion of foreign genes is considered. The OptStrain framework was developed for examining the insertion of foreign genes for strain improvement (Pharkya *et al.*, 2004) (Fig. 4.3). OptStain first identifies a pathway that can achieve the maximum *in silico* yield of desired biochemical using a universal reaction database, which includes all elementally balanced metabolic reactions. Optknock subsequently redesigns a stoichiometrically balanced metabolic pathway that contains the minimum number of nonnative reactions from the universal database, incorporates the nonnative reactions into the host's metabolic model, and finally applies the OptKnock framework to optimize the phenotype of the newly designed strain. The OptStrain framework was validated through the designing of strategies for the production of hydrogen and vanillin in *E. coli* (Pharkya *et al.*, 2004).

3.5. Identifying gene targets for engineering strain development: Metabolite essentiality

Biological systems maintain phenotypic stability against diverse genetic and environmental perturbations due to redundant or alternative pathways. This is an inherent property of metabolic networks called as robustness (Kitano, 2004, 2007a,b; Stelling *et al.*, 2004; Xu *et al.*, 2009). The study on the robustness of an organism has generally depended on the identification of genes or reactions essential for cell growth as a reaction-centric viewpoint. However, the reaction-centric approach has met with difficulties, where the only limited number of genes or reactions has been identified as essential ones that destroy the cellular robustness. Thus, a metabolite-centric approach, flux-sum, for analyzing the robustness of metabolic network was reported (Chung and Lee, 2009; Kim *et al.*, 2007) (Fig. 4.3). The flux-sum is defined to be one half of the summation of all consumption and generation fluxes around a particular metabolite under pseudo-steady state. Flux-sum analysis elucidates the essentiality of metabolites in the

network through observing the behavior of the flux-sum to perturbations to the metabolic network. For example, essential metabolites are capable of maintaining a steady flux-sum against diverse perturbations by redistributing metabolic fluxes so that the flux-sum remains steady. Hence, the breakdown of the flux-sum around essential metabolites can have negative effects on cellular robustness and cell growth or survival. This metabolite-centric approach provides unique insights into cellular robustness and relevant fragility. Flux-sum has been applied to several applications where robustness or the disruption of robustness is examined, such as identifying drug target candidates in pathogens. Using the metabolic network of the pathogen *Acinetobacter baumannii* AYE, the flux-sum approach was used to find the most effective drug targets in killing the pathogen by targeting metabolites that are essential in the robustness of its metabolic network (Kim *et al.*, 2010).

3.6. Accurately describing cellular physiology: Incorporation of experimental data and physiological properties into the *in silico* model

The results simulated from *in silico* genome-scale model usually do not agree with the experimental data because the information used to reconstruct the model is incomplete. Also, the discrepancies between predictions and experimental data can be caused by the broad flux solution space of *in silico* genome-scale model, which represents all physiologically feasible states and is much larger in comparison with the biologically feasible flux solution space of a real organism (Kim *et al.*, 2008a; Palsson *et al.*, 2003; Park *et al.*, 2009). Thus, there have been several efforts to reduce the flux solution space of the *in silico* genome-scale model to reduce the differences between prediction and experiment.

In some cases, the inaccurate prediction results can be caused by the dissimilarity of active reaction sets between *in vivo* system and *in silico* model. The flux solution space of *in silico* genome-scale model can be reduced by eliminating unrealistic reactions under a given condition. The algorithm, optimal metabolic network identification (OMNI), determines the active reactions in the *in silico* genome-scale model through comparison with experimentally measured fluxes from ^{13}C-based flux analysis (Herrgard *et al.*, 2006a) (Fig. 4.3). OMNI efficiently identifies the set of reactions, by minimizing the discrepancies between experimental data and *in silico* predictions using MILP, that need to be included in the metabolic model and finds bottleneck reactions that need to be excluded from the metabolic model.

Another strategy in reducing the flux solution space of *in silico* genome-scale model is to supply additional and more detailed information and procedures regarding the metabolic network in the form of additional constraints for the model and simulations (Fig. 4.3). Experimentally measured flux data typically obtained from ^{13}C-based flux analysis and

fermentation data can be utilized as constraints during the simulation of *in silico* genome-scale model (Blank *et al.*, 2005; Fischer and Sauer, 2005; Kim and Lee, 2006) (Fig. 4.3). Using ^{13}C-based flux analysis, the *in vivo* fluxes calculated by using isotope-labeled substrates can serve as realistic constraints by limiting the flux values of corresponding intracellular reactions.

In vivo systems are controlled by complex regulatory mechanism that responds to various environmental changes, such as temperature, pH, oxygenic condition, or genetic perturbations. Thus, attempts were made to integrate regulatory information into the *in silico* metabolic models using Boolean logics to describe the regulatory mechanism (Barrett *et al.*, 2005; Covert *et al.*, 2004, 2008; Herrgard *et al.*, 2006b; Shlomi *et al.*, 2007) (Fig. 4.3). Steady-state regulatory flux balance analysis (SR-FBA) was developed by combining *in silico* genome-scale model with a transcriptional regulation network that represents GPR relationship through Boolean logics. SR-FBA can express regulatory effects, using MILP, of environmental or genetic perturbations by operating on/off conditions of gene, protein, and reaction as binary variables.

Conventional FBA determines flux distributions depending only on mass balances of metabolites but does not consider the thermodynamics of reactions. Accordingly, this results in several reactions showing thermodynamically infeasible fluxes. To resolve this issue, FBA extensions, such as energy balance analysis (EBA) (Beard *et al.*, 2004; Yang *et al.*, 2005), thermodynamics-based metabolic flux analysis (TMFA) (Henry *et al.*, 2007), and so on (Feist *et al.*, 2007; Kummel *et al.*, 2006), have been carried out (Fig. 4.3). The feasibility, directionality, and reversibility of metabolic reactions in the model can be determined by calculating the Gibbs free energy of the metabolic reactions based on the laws of thermodynamics. Metabolic reactions that are found to violate the laws of thermodynamics can be modified or excluded from the *in silico* metabolic model.

Intracellular cytoplasm is occupied by macromolecules, many of which are enzymes (Beg *et al.*, 2007). The cytoplasmic enzymes are restricted within the available cytoplasmic space. Thus, the concentration of cytoplasmic enzymes cannot increase further without drastic effects on protein structures, biochemical reaction kinetics, and dynamics of transport within the limited cytoplasmic space of a cell. Consequently, the competitions among enzymes in the limited cytoplasmic space might affect the attainable flux values of each reaction. To incorporate this cellular physiological property into FBA, FBA with molecular crowding (FBAwMC), representing physical and spatial constraints, was applied to predict the growth rate of *E. coli* wild type and mutant strains and to examine the dynamic patterns of substrate utilization (i.e., the sequence and mode of substrate uptake) of the *E. coli* cell in mixed-substrate media (Beg *et al.*, 2007) (Fig. 4.3).

Intracellular proteins, including enzymes responsible for catalyzing metabolic reactions, interact directly through physical binding and may also

interact indirectly through utilizing a substrate together during enzymatic actions, regulating each other transcriptionally, or forming larger multi-protein assemblies (von Mering *et al.*, 2003, 2007). Functional associations among proteins can be analyzed by genomic context of the genes in the form of conserved neighborhood (i.e., the degree of proximity), gene fusion (i.e., events of forming a hybrid gene), and co-occurrence (i.e., presence or absence across organisms; Jensen *et al.*, 2009). To incorporate this cellular physiological property into FBA, the constraints regarding the grouping of functionally and physically related reactions in the metabolic network were developed by considering genomic context and flux-converging patterns (Park *et al.*, 2010) (Fig. 4.3). Based on genomic context analysis, functionally related reactions are organized together. Followed by genomic context analysis, reactions in each group are further clustered by flux-converging pattern analysis that considers the carbon number of metabolites in reactions and the flux patterns converged from a carbon source in metabolic network. Based on the assumption that the functionally related reactions in the same group show similar expression patterns by similar regulation under several conditions, FBA with grouping reaction constraints (FBAwGR) was applied to describe the changes of fluxes under several different genotypic (*pykF*, *zwf*, *ppc*, and *sucA* knockout mutants) and environmental (i.e., carbon source shift from glucose to acetate) conditions in *E. coli* (Park *et al.*, 2010).

Mutualisms, in which two or more organisms interact with one another and each individual obtains fitness benefits, may significantly influence the community structure and stability of ecosystems. To describe the mutualistic interactions between two different organisms (i.e., sulfate-reducing bacteria and methanogens; *Desulfovibrio vulgaris* and *Methanococcus maripaludis*), a method using a multiobjective system was developed by designing a system of three compartments (Stolyar *et al.*, 2007). Two metabolic models representing two different species were constructed and contained in separate compartments. The third compartment contained exchange reactions connecting the two organisms through the transfer of metabolites between the two species.

4. Concluding Remarks

In systems metabolic engineering, MFA plays an important role for the generation of new biological knowledge on the cellular system, system-wide analysis of cellular physiology, and in developing metabolic engineering strategies at the systems-level. Practically, MFA by means of several *in silico* algorithms and constraints has been applied for the understanding of metabolic characteristics of a cell and for the design of metabolic strategies identifying target genes to be engineered for strain improvement. When performing simulations using *in silico* methods, determining a suitable

algorithm for application is important (Fig. 4.3). The simulation starts with the consideration of the desired purposes, whether it is to accurately describe cellular physiology or to identify target genes for strain improvement.

In describing cellular physiology, the flux solution space of an *in silico* genome-scale model that represents all biologically feasible metabolic states for a given condition is examined. However, the flux solution space of the model is broader than the physiologically feasible flux solution space of the real cell because of various levels of cellular mechanisms, such as cellular regulation, signaling, and homeostasis, that are not considered due to incomplete information regarding the metabolic network. Thus, additional constraints, including experimental flux data, Boolean logics representing transcriptional regulation, thermodynamics, and physiological data, during simulation can improve the flux solution space of the metabolic model to represent that of a real cell. In improving cellular performance, metabolic engineering approaches, using gene knockout, gene amplification, gene down-regulation, and introduction of foreign genes, have been considered. In identifying gene targets to be engineered for strain improvement, several *in silico* algorithms can be applied and have been discussed above.

There are still several important issues that need to be advanced to complete *in silico* cell that realizes a real cell successfully. The first thing to be done is to reconstruct thoroughly the *in silico* genome-scale metabolic model with accurate information. In the aspect of describing cellular physiology, other new constraints and improved algorithms need to be developed to incorporate other physiological properties into the *in silico* genome-scale metabolic model. Additionally, innovative and progressive *in silico* algorithms beyond static approach based on pseudo-steady state assumptions need to be developed to describe the dynamic behaviors of a cell. Still, the operation of an individual algorithm has limitations. Since several mechanisms in a real cell operate synchronously, integration of several algorithms or constraints might be advisable for the description of complex cellular physiology. Accordingly, *in silico* strategies to integrate metabolic, regulatory, and signal transduction mechanisms, such as integrated dynamic FBA (idFBA) (Lee *et al.*, 2008) and integrated FBA (iFBA) (Covert *et al.*, 2008), and combining dynamic kinetics with *in silico* genome-scale metabolic model (Yugi *et al.*, 2005) have been proposed. In conclusion, to make great progress on these issues, we rather make constant efforts on developing the advanced models and *in silico* algorithms for their applications in biological and biotechnological studies.

ACKNOWLEDGMENTS

This work was supported by the Korean Systems Biology Research Project (20100002164) of the Ministry of Education, Science and Technology (MEST). Further support by the World Class University Program (R32-2009-000-10142-0) through the National Research Foundation of Korea funded by the MEST is appreciated.

REFERENCES

Al Zaid Siddiquee, K., Arauzo-Bravo, M. J., and Shimizu, K. (2004). Metabolic flux analysis of *pykF* gene knockout *Escherichia coli* based on ^{13}C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations. *Appl. Microbiol. Biotechnol.* **63,** 407–417.

Alper, H., Jin, Y. S., Moxley, J. F., and Stephanopoulos, G. (2005). Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.* **7,** 155–164.

Asadollahi, M. A., Maury, J., Patil, K. R., Schalk, M., Clark, A., and Nielsen, J. (2009). Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through *in silico* driven metabolic engineering. *Metab. Eng.* **11,** 328–334.

Barrett, C. L., Herring, C. D., Reed, J. L., and Palsson, B. O. (2005). The global transcriptional regulatory network for metabolism in *Escherichia coli* exhibits few dominant functional states. *Proc. Natl. Acad. Sci. USA* **102,** 19103–19108.

Beard, D. A., Babson, E., Curtis, E., and Qian, H. (2004). Thermodynamic constraints for biochemical networks. *J. Theor. Biol.* **228,** 327–333.

Beg, Q. K., Vazquez, A., Ernst, J., de Menezes, M. A., Bar-Joseph, Z., Barabasi, A. L., and Oltvai, Z. N. (2007). Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl. Acad. Sci. USA* **104,** 12663–12668.

Blank, L. M., Kuepfer, L., and Sauer, U. (2005). Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* **6,** R49.

Bundy, J. G., Papp, B., Harmston, R., Browne, R. A., Clayson, E. M., Burton, N., Reece, R. J., Oliver, S. G., and Brindle, K. M. (2007). Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling. *Genome Res.* **17,** 510–519.

Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84,** 647–657.

Burgard, A. P., Nikolaev, E. V., Schilling, C. H., and Maranas, C. D. (2004). Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14,** 301–312.

Bushell, M. E., Sequeira, S. I., Khannapho, C., Zhao, H., Chater, K. F., Butler, M. J., Kierzek, A. M., and Avignone-Rossa, C. A. (2006). The use of genome scale metabolic flux variability analysis for process feed formulation based on an investigation of the effects of the *zwf* mutation on antibiotic production in *Streptomyces coelicolor*. *Enzyme Microb. Technol.* **39,** 1347–1353.

Choi, H. S., Lee, S. Y., Kim, T. Y., and Woo, H. M. (2010). *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.* **76,** 3097–3105.

Chung, B. K., and Lee, D. Y. (2009). Flux-sum analysis: A metabolite-centric approach for understanding the metabolic network. *BMC Syst. Biol.* **3,** 117.

Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J., and Palsson, B. O. (2004). Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429,** 92–96.

Covert, M. W., Xiao, N., Chen, T. J., and Karr, J. R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* **24,** 2044–2050.

Davidsen, T., Beck, E., Ganapathy, A., Montgomery, R., Zafar, N., Yang, Q., Madupu, R., Goetz, P., Galinsky, K., White, O., and Sutton, G. (2010). The comprehensive microbial resource. *Nucleic Acids Res.* **38,** D340–D345.

Delgado, J., and Liao, J. C. (1997). Inverse flux analysis for reduction of acetate excretion in *Escherichia coli*. *Biotechnol. Prog.* **13,** 361–367.

Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R., and Palsson, B. O. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. USA* **104,** 1777–1782.

Durot, M., Bourguignon, P. Y., and Schachter, V. (2009). Genome-scale models of bacterial metabolism: Reconstruction and applications. *FEMS Microbiol. Rev.* **33,** 164–190.

Edwards, J. S., and Palsson, B. O. (2000). The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* **97,** 5528–5533.

Edwards, J. S., Ibarra, R. U., and Palsson, B. O. (2001). *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19,** 125–130.

Faria, J. P., Focha, M., Stevens, R. L., and Henry, C. S. (2010). Analysis of the effect of reversibility constraints on the predictions of genome-scale metabolic models. *In* "Advances in Bioinformatics," (M. P. Rocha, F. F. Riverola, H. Shatkay, and J. M. Corchado, eds.), pp. 209–215. Springer, Berlin.

Feist, A. M., and Palsson, B. O. (2008). The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nat. Biotechnol.* **26,** 659–667.

Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson, B. O. (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.

Finley, S. D., Broadbelt, L. J., and Hatzimanikatis, V. (2010). In silico feasibility of novel biodegradation pathways for 1, 2, 4-trichlorobenzene. *BMC Syst. Biol.* **4,** 7.

Fischer, E., and Sauer, U. (2005). Large-scale *in vivo* flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat. Genet.* **37,** 636–640.

Fong, S. S., Burgard, A. P., Herring, C. D., Knight, E. M., Blattner, F. R., Maranas, C. D., and Palsson, B. O. (2005). *In silico* design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* **91,** 643–648.

Fuhrer, T., and Sauer, U. (2009). Different biochemical mechanisms ensure network-wide balancing of reducing equivalents in microbial metabolism. *J. Bacteriol.* **191,** 2112–2121.

Gonzalez-Lergier, J., Broadbelt, L. J., and Hatzimanikatis, V. (2006). Analysis of the maximum theoretical yield for the synthesis of erythromycin precursors in *Escherichia coli*. *Biotechnol. Bioeng.* **95,** 638–644.

Grafahrend-Belau, E., Schreiber, F., Koschutzki, D., and Junker, B. H. (2009). Flux balance analysis of barley seeds: A computational approach to study systemic properties of central metabolism. *Plant Physiol.* **149,** 585–598.

Henry, C. S., Broadbelt, L. J., and Hatzimanikatis, V. (2007). Thermodynamics-based metabolic flux analysis. *Biophys. J.* **92,** 1792–1805.

Herrgard, M. J., Fong, S. S., and Palsson, B. O. (2006a). Identification of genome-scale metabolic network models using experimentally measured flux profiles. *PLoS Comput. Biol.* **2,** e72.

Herrgard, M. J., Lee, B. S., Portnoy, V., and Palsson, B. O. (2006b). Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. *Genome Res.* **16,** 627–635.

Hong, S. H., Moon, S. Y., and Lee, S. Y. (2003). Prediction of maximum yields of metabolites and optimal pathways for their production by metabolic flux analysis. *J. Microbiol. Biotechnol.* **13,** 571–577.

Hu, W., Sillaots, S., Lemieux, S., Davison, J., Kauffman, S., Breton, A., Linteau, A., Xin, C., Bowman, J., Becker, J., Jiang, B., and Roemer, T. (2007). Essential gene identification and drug target prioritization in *Aspergillus fumigatus*. *PLoS Pathog.* **3,** e24.

Hua, Q., Joyce, A. R., Fong, S. S., and Palsson, B. O. (2006). Metabolic analysis of adaptive evolution for *in silico*-designed lactate-producing strains. *Biotechnol. Bioeng.* **95,** 992–1002.

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., Ho, P. Y., Kakazu, Y., *et al.* (2007). Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **316,** 593–597.

Jamshidi, N., and Palsson, B. O. (2007). Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the *in silico* strain iNJ661 and proposing alternative drug targets. *BMC Syst. Biol.* **1,** 26.

Jensen, P. R., and Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* **64,** 82–87.

Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., and von Mering, C. (2009). STRING 8—A global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37,** D412–D416.

Joyce, A. R., and Palsson, B. O. (2006). The model organism as a system: Integrating 'omics' data sets. *Nat. Rev. Mol. Cell Biol.* **7,** 198–210.

Jung, Y. K., Kim, T. Y., Park, S. J., and Lee, S. Y. (2010). Metabolic engineering of *Escherichia coli* for the production of polylactic acid and its copolymers. *Biotechnol. Bioeng.* **105,** 161–171.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010). KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38,** D355–D360.

Kauffman, K. J., Prakash, P., and Edwards, J. S. (2003). Advances in flux balance analysis. *Curr. Opin. Biotechnol.* **14,** 491–496.

Khannapho, C., Zhao, H., Bonde, B. K., Kierzek, A. M., Avignone-Rossa, C. A., and Bushell, M. E. (2008). Selection of objective function in genome scale flux balance analysis for process feed development in antibiotic production. *Metab. Eng.* **10,** 227–233.

Kim, T. Y., and Lee, S. Y. (2006). Accurate metabolic flux analysis through data reconciliation of isotope balance-based data. *J. Microbiol. Biotechnol.* **16,** 1139–1143.

Kim, J., and Reed, J. L. (2010). OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst. Biol.* **4,** 53.

Kim, P. J., Lee, D. Y., Kim, T. Y., Lee, K. H., Jeong, H., Lee, S. Y., and Park, S. (2007). Metabolite essentiality elucidates robustness of *Escherichia coli* metabolism. *Proc. Natl. Acad. Sci. USA* **104,** 13638–13642.

Kim, H. U., Kim, T. Y., and Lee, S. Y. (2008a). Metabolic flux analysis and metabolic engineering of microorganisms. *Mol. Biosyst.* **4,** 113–120.

Kim, T. Y., Sohn, S. B., Kim, H. U., and Lee, S. Y. (2008b). Strategies for systems-level metabolic engineering. *Biotechnol. J.* **3,** 612–623.

Kim, H. U., Kim, T. Y., and Lee, S. Y. (2010). Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen *Acinetobacter baumannii* AYE. *Mol. Biosyst.* **6,** 339–348.

Kitano, H. (2004). Biological robustness. *Nat. Rev. Genet.* **5,** 826–837.

Kitano, H. (2007a). A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* **6,** 202–210.

Kitano, H. (2007b). Towards a theory of biological robustness. *Mol. Syst. Biol.* **3,** 137.

Kizer, L., Pitera, D. J., Pfleger, B. F., and Keasling, J. D. (2008). Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl. Environ. Microbiol.* **74,** 3229–3241.

Koffas, M. A., Jung, G. Y., and Stephanopoulos, G. (2003). Engineering metabolism and product formation in *Corynebacterium glutamicum* by coordinated gene overexpression. *Metab. Eng.* **5,** 32–41.

Kummel, A., Panke, S., and Heinemann, M. (2006). Systematic assignment of thermodynamic constraints in metabolic network models. *BMC Bioinform.* **7,** 512.

Kwon, Y. K., Lu, W., Melamud, E., Khanam, N., Bognar, A., and Rabinowitz, J. D. (2008). A domino effect in antifolate drug action in *Escherichia coli*. *Nat. Chem. Biol.* **4,** 602–608.

Lee, S. Y., and Papoutsakis, E. T. (1999). Metabolic Engineering. Marcel Dekker, New York.

Lee, S. J., Lee, D. Y., Kim, T. Y., Kim, B. H., Lee, J., and Lee, S. Y. (2005a). Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and *in silico* gene knockout simulation. *Appl. Environ. Microbiol.* **71,** 7880–7887.

Lee, S. Y., Lee, D. Y., and Kim, T. Y. (2005b). Systems biotechnology for strain improvement. *Trends Biotechnol.* **23,** 349–358.

Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., and Lee, S. Y. (2007). Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* **3,** 149.

Lee, J. M., Gianchandani, E. P., Eddy, J. A., and Papin, J. A. (2008). Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput. Biol.* **4,** e1000086.

Lee, S. Y., Kim, H. U., Park, J. H., Park, J. M., and Kim, T. Y. (2009). Metabolic engineering of microorganisms: General strategies and drug production. *Drug Discov. Today* **14,** 78–88.

Li, M., Ho, P. Y., Yao, S., and Shimizu, K. (2006). Effect of *sucA* or *sucC* gene knockout on the metabolism in *Escherichia coli* based on gene expressions, enzyme activities, intracellular metabolite concentrations and metabolic fluxes by ^{13}C-labeling experiments. *Biochem. Eng. J.* **30,** 286–296.

Lun, D. S., Rockwell, G., Guido, N. J., Baym, M., Kelner, J. A., Berger, B., Galagan, J. E., and Church, G. M. (2009). Large-scale identification of genetic design strategies using local search. *Mol. Syst. Biol.* **5,** 296.

Mahadevan, R., and Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **5,** 264–276.

Moon, S. Y., Hong, S. H., Kim, T. Y., and Lee, S. Y. (2008). Metabolic engineering of *Escherichia coli* for the production of malic acid. *Biochem. Eng. J.* **40,** 312–320.

Nanchen, A., Schicker, A., Revelles, O., and Sauer, U. (2008). Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. *J. Bacteriol.* **190,** 2323–2330.

Oberhardt, M. A., Puchalka, J., Fryer, K. E., Martins dos Santos, V. A., and Papin, J. A. (2008). Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **190,** 2790–2803.

Oberhardt, M. A., Chavali, A. K., and Papin, J. A. (2009). Flux balance analysis: Interrogating genome-scale metabolic networks. *Methods Mol. Biol.* **500,** 61–80.

Orth, J. D., Thiele, I., and Palsson, B. O. (2010). What is flux balance analysis? *Nat. Biotechnol.* **28,** 245–248.

Pal, C., Papp, B., and Lercher, M. J. (2005). Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37,** 1372–1375.

Palsson, B. O., Price, N. D., and Papin, J. A. (2003). Development of network-based pathway definitions: The need to analyze real metabolic networks. *Trends Biotechnol.* **21,** 195–198.

Park, J. H., and Lee, S. Y. (2008). Towards systems metabolic engineering of microorganisms for amino acid production. *Curr. Opin. Biotechnol.* **19,** 454–460.

Park, J. H., Lee, K. H., Kim, T. Y., and Lee, S. Y. (2007). Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc. Natl. Acad. Sci. USA* **104,** 7797–7802.

Park, J. M., Kim, T. Y., and Lee, S. Y. (2009). Constraints-based genome-scale metabolic simulation for systems metabolic engineering. *Biotechnol. Adv.* **27,** 978–988.

Park, J. M., Kim, T. Y., and Lee, S. Y. (2010). Prediction of metabolic fluxes by incorporating genomic context and flux-converging pattern analyses. *Proc. Natl. Acad. Sci. USA* **107,** 14931–14936.

Patil, K. R., Rocha, I., Forster, J., and Nielsen, J. (2005). Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinform.* **6,** 308.

Peng, L., Arauzo-Bravo, M. J., and Shimizu, K. (2004). Metabolic flux analysis for a *ppc* mutant *Escherichia coli* based on ^{13}C-labelling experiments together with enzyme activity assays and intracellular metabolite measurements. *FEMS Microbiol. Lett.* **235,** 17–23.

Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., and Vorholt, J. A. (2009). Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. *Proc. Natl. Acad. Sci. USA* **106,** 4846–4851.

Pharkya, P., and Maranas, C. D. (2006). An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab. Eng.* **8,** 1–13.

Pharkya, P., Burgard, A. P., and Maranas, C. D. (2003). Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol. Bioeng.* **84,** 887–899.

Pharkya, P., Burgard, A. P., and Maranas, C. D. (2004). OptStrain: A computational framework for redesign of microbial production systems. *Genome Res.* **14,** 2367–2376.

Price, N. D., Reed, J. L., and Palsson, B. O. (2004). Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2,** 886–897.

Puchalka, J., Oberhardt, M. A., Godinho, M., Bielecka, A., Regenhardt, D., Timmis, K. N., Papin, J. A., and Martins dos Santos, V. A. (2008). Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. *PLoS Comput. Biol.* **4,** e1000210.

Rabinowitz, J. D. (2007). Cellular metabolomics of *Escherchia coli*. *Expert Rev. Proteomics* **4,** 187–198.

Ramakrishna, R., Edwards, J. S., McCulloch, A., and Palsson, B. O. (2001). Flux-balance analysis of mitochondrial energy metabolism: Consequences of systemic stoichiometric constraints. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **280,** R695–R704.

Raman, K., and Chandra, N. (2009). Flux balance analysis of biological systems: Applications and challenges. *Brief. Bioinform.* **10,** 435–449.

Ranganathan, S., and Maranas, C. D. (2010). Microbial 1-butanol production: Identification of non-native production routes and *in silico* engineering interventions. *Biotechnol. J.* **5,** 716–725.

Ranganathan, S., Suthers, P. F., and Maranas, C. D. (2010). OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput. Biol.* **6,** e1000744.

Reed, J. L., Patel, T. R., Chen, K. H., Joyce, A. R., Applebee, M. K., Herring, C. D., Bui, O. T., Knight, E. M., Fong, S. S., and Palsson, B. O. (2006). Systems approach to refining genome annotation. *Proc. Natl. Acad. Sci. USA* **103,** 17480–17484.

Risso, C., Van Dien, S. J., Orloff, A., Lovley, D. R., and Coppi, M. V. (2008). Elucidation of an alternate isoleucine biosynthesis pathway in *Geobacter sulfurreducens*. *J. Bacteriol.* **190,** 2266–2274.

Sauer, U. (2004). High-throughput phenomics: Experimental methods for mapping fluxomes. *Curr. Opin. Biotechnol.* **15,** 58–63.

Sauer, U. (2006). Metabolic networks in motion: ^{13}C-based flux analysis. *Mol. Syst. Biol.* **2,** 62.

Schmidt, K., Nielsen, J., and Villadsen, J. (1999). Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models. *J. Biotechnol.* **71,** 175–189.

Schneider, K., Kromer, J. O., Wittmann, C., Alves-Rodrigues, I., Meyerhans, A., Diez, J., and Heinzle, E. (2009). Metabolite profiling studies in *Saccharomyces cerevisiae*: An assisting tool to prioritize host targets for antiviral drug screening. *Microb. Cell Fact.* **8,** 12.

Schuetz, R., Kuepfer, L., and Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3,** 119.

Segre, D., Vitkup, D., and Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA* **99,** 15112–15117.

Shlomi, T., Berkman, O., and Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc. Natl. Acad. Sci. USA* **102,** 7695–7700.

Shlomi, T., Eisenberg, Y., Sharan, R., and Ruppin, E. (2007). A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* **3,** 101.

Smallbone, K., and Simeonidis, E. (2009). Flux balance analysis: A geometric perspective. *J. Theor. Biol.* **258,** 311–315.

Song, H., Kim, T. Y., Choi, B. K., Choi, S. J., Nielsen, L. K., Chang, H. N., and Lee, S. Y. (2008). Development of chemically defined medium for *Mannheimia succiniciproducens* based on its genome sequence. *Appl. Microbiol. Biotechnol.* **79,** 263–272.

Stelling, J., Sauer, U., Szallasi, Z., Doyle, 3rd, F. J., and Doyle, J. (2004). Robustness of cellular functions. *Cell* **118,** 675–685.

Stephanopoulos, G. N., Aristidou, A. A., and Nielsen, J. (1998). Metabolic Engineering. Academic Press, San Diego.

Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., and Stahl, D. A. (2007). Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* **3,** 92.

Tang, Y. J., Chakraborty, R., Martin, H. G., Chu, J., Hazen, T. C., and Keasling, J. D. (2007). Flux analysis of central metabolic pathways in *Geobacter metallireducens* during reduction of soluble Fe(III)-nitrilotriacetic acid. *Appl. Environ. Microbiol.* **73,** 3859–3864.

Tang, Y. J., Martin, H. G., Dehal, P. S., Deutschbauer, A., Llora, X., Meadows, A., Arkin, A., and Keasling, J. D. (2009). Metabolic flux analysis of *Shewanella* spp. reveals evolutionary robustness in central carbon metabolism. *Biotechnol. Bioeng.* **102,** 1161–1169.

Tannler, S., Fischer, E., Le Coq, D., Doan, T., Jamet, E., Sauer, U., and Aymerich, S. (2008). CcpN controls central carbon fluxes in *Bacillus subtilis*. *J. Bacteriol.* **190,** 6178–6187.

Teusink, B., Wiersma, A., Molenaar, D., Francke, C., de Vos, W. M., Siezen, R. J., and Smid, E. J. (2006). Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale metabolic model. *J. Biol. Chem.* **281,** 40041–40048.

Varma, A., Boesch, B. W., and Palsson, B. O. (1993). Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl. Environ. Microbiol.* **59,** 2465–2473.

von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., and Snel, B. (2003). STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res.* **31,** 258–261.

von Mering, C., Jensen, L. J., Kuhn, M., Chaffron, S., Doerks, T., Kruger, B., Snel, B., and Bork, P. (2007). STRING 7–recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res.* **35,** D358–D362.

Wahl, A., El Massaoudi, M., Schipper, D., Wiechert, W., and Takors, R. (2004). Serial ^{13}C-based flux analysis of an L-phenylalanine-producing *E. coli* strain using the sensor reactor. *Biotechnol. Prog.* **20,** 706–714.

Wiechert, W. (2001). ^{13}C metabolic flux analysis. *Metab. Eng.* **3,** 195–206.

Xu, Z., Sun, X., and Yu, S. (2009). Genome-scale analysis to the impact of gene deletion on the metabolism of *E. coli*: Constraint-based simulation approach. *BMC Bioinform.* **10** (Suppl. 1), S62.

Yang, F., Qian, H., and Beard, D. A. (2005). *Ab initio* prediction of thermodynamically feasible reaction directions from biochemical network stoichiometry. *Metab. Eng.* **7,** 251–259.

Yeh, I., Hanekamp, T., Tsoka, S., Karp, P. D., and Altman, R. B. (2004). Computational analysis of *Plasmodium falciparum* metabolism: Organizing genomic information to facilitate drug discovery. *Genome Res.* **14,** 917–924.

Yugi, K., Nakayama, Y., Kinoshita, A., and Tomita, M. (2005). Hybrid dynamic/static method for large-scale simulation of metabolism. *Theor. Biol. Med. Model.* **2,** 42.

Zamboni, N., and Sauer, U. (2009). Novel biological insights through metabolomics and ^{13}C-flux analysis. *Curr. Opin. Microbiol.* **12,** 553–558.

Zhao, J., Baba, T., Mori, H., and Shimizu, K. (2004a). Effect of *zwf* gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. *Metab. Eng.* **6,** 164–174.

Zhao, J., Baba, T., Mori, H., and Shimizu, K. (2004b). Global metabolic response of *Escherichia coli* to *gnd* or *zwf* gene-knockout, based on ^{13}C-labeling experiments and the measurement of enzyme activities. *Appl. Microbiol. Biotechnol.* **64,** 91–98.