J. Microbiol. Biotechnol. (2006), 16(6), 993–998

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Simulation of Dynamic Behavior of Glucose- and Tryptophan-Grown *Escherichia coli* Using Constraint-Based Metabolic Models with a Hierarchical Regulatory Network

LEE, SUNG GUN[1], YU-JIN KIM[1,2], SANG IL HAN[1], YOU-KWAN OH[4], SUNGHOON PARK[1,2], YOUNG HAN KIM[3], AND KYU SUK HWANG[1]*

[1]*Department of Chemical Engineering, College of Engineering, Pusan National University, Pusan 609-735, South Korea*
[2]*Institute for Environmental Technology and Industry, Pusan National University, Pusan 609-735, South Korea*
[3]*Department of Chemical Engineering, Dong-A University, Pusan 604-714, South Korea*
[4]*Biomass Research Team, Korea Institute of Energy Research, Daejeon 305-343, Korea*

Received: December 20, 2005
Accepted: February 26, 2006

Abstract We earlier suggested a hierarchical regulatory network using defined modeling symbols and weights in order to improve the flux balance analysis (FBA) with regulatory events that were represented by if-then rules and Boolean logic. In the present study, the simulation results of the models, which were developed and improved from the previous model by incorporating a hierarchical regulatory network into the FBA, were compared with the experimental outcome of an aerobic batch growth of *E. coli* on glucose and tryptophan. From the experimental result, a diauxic growth curve was observed, reflecting growth resumption, when tryptophan was used as an alternativee after the supply of glucose was exhausted. The model parameters, the initial concentration of substrates (0.92 mM glucose and 1 mM tryptophan), cell density (0.0086 g biomass/l), the maximal uptake rates of substrates (5.4 mmol glucose/g·DCW·h and 1.32 mmol tryptophan/g·DCW·h), and lag time (0.32 h) were derived from the experimental data for more accurate prediction. The simulation results agreed with the experimental outcome of the temporal profiles of cell density and glucose, and tryptophan concentrations.

Key words: Diauxic growth curve, flux balance analysis, hierarchical regulatory network

Metabolites, metabolic circuits, genes, and protein-protein interactions are necessary for the construction of detailed metabolic networks. The recent flood of "-omics" data from genomics, proteomics, transcriptomics, metabolomics, and fluxomics, and the development of programming tools to accurately analyze these data have made available the entire biochemical pathway of microbes. This development has also promoted research on an *in silico* microbe model that can simulate the physiological phenomena generated in the microbe using a PC. There are many approaches to develop the *in silico* models, such as kinetic models, metabolic pathway models, cybernetic models, and stochastic models [10]. Among these, kinetic models and metabolic pathway models are the leading approaches.

The *in silico* model based on kinetics requires parameters in kinetic equations and concentration data that describe enzymatic reactions and regulatory events. With all the kinetic parameters and concentration data known, the microbial behaviors can be predicted. However, it is impossible to conduct all the necessary experiments for obtaining the parameters of kinetic equations [7].

On the other hand, without the detailed kinetic information, flux balance analysis (FBA) can predict cell growth rates, metabolic byproducts, substrate concentrations, and other parameters related to achieving a desired production with the readily available stoichiometry of metabolic pathways, metabolic requirements, and fundamental physicochemical constraints [6, 8, 9, 12, 13]. Occasionally, however, FBA incorrectly predicts cell behavior, because the modeling approach does not use the regulatory gene information, which limits specific metabolic fluxes with extracellular environmental conditions. Recently, some researchers have attempted to incorporate regulatory constraints described by if-then rules and Boolean logic in FBA [1]. Because if-then rules with Boolean logic are a binary system, it is difficult to represent certain situations, such as when true

*Corresponding author
Phone: 82-51-510-2400; Fax: 82-51-512-8563;
E-mail: kshwang@pusan.ac.kr

and false events occur simultaneously, and to represent gene control mechanisms, because of the difficulty of tracking how the rules are applied [4, 5, 14].

In order to resolve these problems, we have recently suggested to use weights for more flexible representation of the degree of connectivity between regulatory elements and hierarchical regulatory networks to describe gene control mechanisms with defined symbols [4]. Here, in order to validate our modeling approach in environmental conditions that were present in one of the case studies in the previous study [4], an experiment was conducted during aerobic batch growth of *E. coli* using glucose and tryptophan. The simulation parameters from this experiment, such as uptake rates, lag time, and maximal growth rates, were determined for more accurate predictions. Finally, the experimental data and simulation results from the proposed modeling method were compared for evaluation.

An *E. coli* K-12 strain, W3110 (KCTC2223), was used for fermentation. The strain is known as being nearly a wild-type, and it is able to grow on a glucose mineral medium. A batch cultivation was performed at 37°C in the M9 medium (Na_2HPO_4 6 g/l, KH_2PO_4 3 g/l, NaCl 0.5 g/l, NH_4Cl 1 g/l, $MgSO_4$ 2 mM, $CaCl_2$ 0.1 mM, $FeCl_3$ 0.01 mM) supplemented with 0.92 mM glucose and 1 mM tryptophan. This low concentration of glucose was used in order to more clearly observe the cell growth on an alternative carbon source in the diauxic growth of *E. coli*. The dissolved oxygen (DO) level was maintained at over 50% of air saturation in a bioreactor (KoBioTech, Korea, Model 5 l) with a working volume of 3 l during aerobic cultivation. Agitation speed was set to 200 rpm with an aeration rate of 1 vvm.

Cell density was measured by a spectrophotometer at 600 nm (Perkin-Elmer, U.S.A.; Model Lambda 20). Glucose was analyzed at 50°C using a high-performance liquid chromatograph (HPLC) (Agilent Technologies, U.S.A.; Model Series 1100) equipped with a Shodex-SH1011 packed column (φ 8 mm×300 mm, Showa Denko K.K., Japan) and a refractive index detector. An aqueous solution of 0.01 N H_2SO_4 was used as an eluent at 0.6 ml/min. Tryptophan was also analyzed by the HPLC on a ZORBAX Eclipse-AAA column (φ 4.6 mm×150 mm, 5 μM; Agilent Technologies, U.S.A.) at 40°C with a UV detector at a wavelength of 280 nm. Two different eluting solutions were used: mobile phase A consisted of 40 mM Na_2HPO_4 (pH 7.8), and mobile phase B consisted of acetonitrile, methanol, and water (45:45:10, v/v/v). The elution was performed by increasing mobile phase B from 0 to 100% at a constant flow rate of 1 ml/min.

LINGO (Lindo System Inc., U.S.A.) was used for linear programming and Matlab (The MathWorks Inc., U.S.A.) for engineering calculations in order to determine the optimal flux balance distribution in the *E. coli* central metabolic pathway, and to quantitatively predict substrate concentration and biomass changes over time.

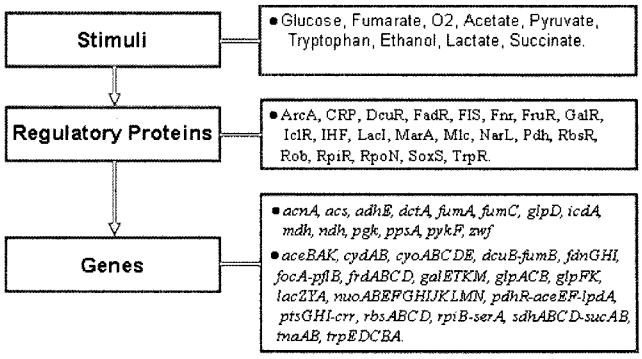


Fig. 1. Stimuli, regulatory proteins, and genes related to the central *Escherichia coli* metabolism modeling symbols.

In this research, the biochemical reactions [11] in the central metabolism of *E. coli* were used as the metabolic pathways for modeling, which is essential for the production of all metabolic byproducts. As shown in Fig. 1, this pathway has 87 genes controlled by 21 regulatory proteins responding to 9 stimuli. A tryptophan degradation pathway (a reaction from tryptophan to phosphoenolpyruate) linking to the central metabolic pathway was created by a one-step reaction with *tnaAB*.

Modeling symbols can be divided into biological and non-biological, as shown in Table 1. The biological symbols

Table 1. Fundamental modeling symbols for a hierarchical regulatory network.

	Function	Symbol
Biological symbols	• Operator	
	• Structural gene	
	• Promoter	
	• Regulatory protein (or transcription factor)	
	• Effector	
	• Gene transcription negative control	
	positive control	
Non-biological symbols	• Weight distributor	
	• And gate	
	• Not gate	
	• Or gate	
	• Symbol state active/inactive	

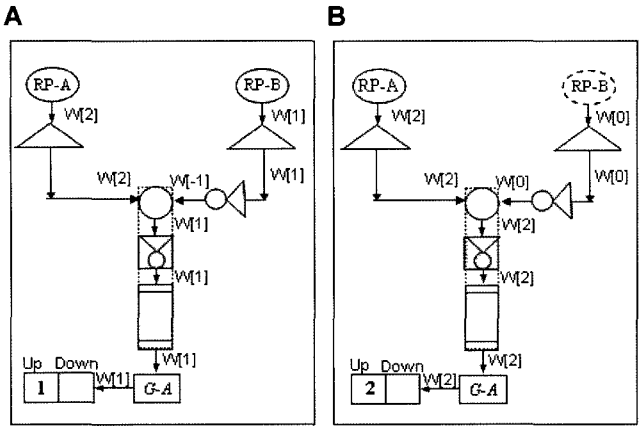


Fig. 2. The level of gene expression in both RP-A and RP-B (**A**) and in the RP-B deletion (**B**) by the weight.
G-A in **A** has "1" as the level of gene expression. *G-A* in **B** has "2" as the level of gene expression. RP-A, activator; RP-B, repressor; *G-A*, gene.

comprised "promoter," "operator," "structural gene," "regulatory protein," and "effector," which were elements in the operon or regulating the operon. When "promoter," "operator," and "structural gene," were assembled into a cluster, they took charge of a gene transcription function. The positive/negative control was determined by active/inactive of "not gate" in the cluster. Non-biological symbols consisted of "not gate," "and gate," and "or gate" to handle weight and weight distribution to indicate operators affected by a stimulus or regulatory protein. Weight was used to determine the active/inactive state of a symbol. In addition, in a situation where two regulatory proteins have influence on one gene (*G-A*) at the same time, the level of up- or downregulation of regulated gene can be represented by the weights. *G-A* is activated by the RP-A in both Figs. 2A and 2B. However, at the RP-B deletion, the *G-A* has higher levels of expression.

A hierarchical regulatory network, constructed by the symbols defined above and consisting of a hierarchy of stimulus, modulon, regulon/operon, gene, and metabolic pathway, allowed us to identify intuitively how genes were controlled by extracellular environmental conditions, as shown in Fig. 4A.

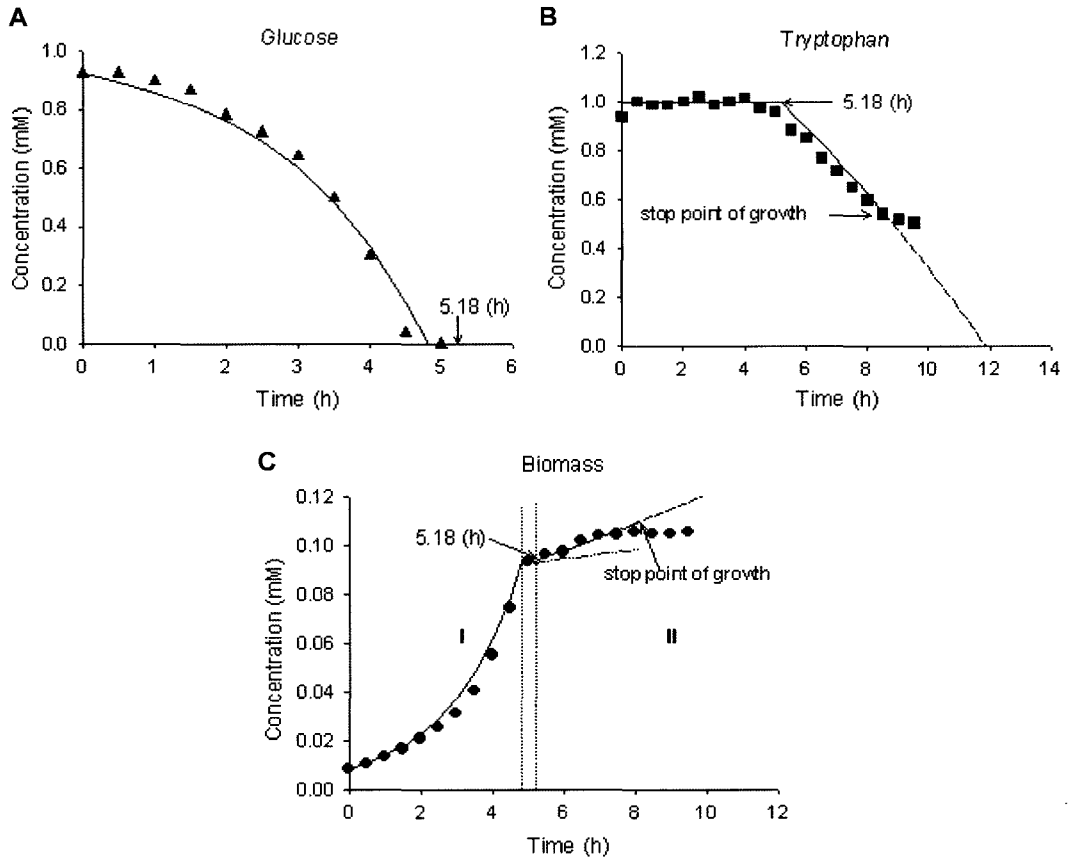


Fig. 3. Three temporal profiles, comparing experimental data obtained by aerobic growth on glucose (I) and tryptophan (II) and the corresponding simulation results.
A. Glucose, experimental (▲) and simulation (—) results. **B.** Tryptophan, experimental (■) and simulation (—) results. The dotted line represents the parts where cell do not grow. **C.** Biomass, experimental (●) and simulation (----) results, and simulation results (—) including the reutilization of byproducts. A dry weight calibration of 0.32 g DCW/optical density at 600 nm was obtained.

A dynamic simulation module has the function of estimating substrate concentration and biomass by the FBA and other equations. This module was composed of two submodules, an FBA and a prediction module. The first step in the execution of the FBA module is to derive a dynamic mass balance for each metabolite in the metabolic pathway constructed from metabolites and fluxes. This mass balance means that the concentration change of a metabolite over time is equal to the difference between the rates of metabolite synthesis and decay [3]. In a steady state, these equations can simply be represented by multiplying the stoichiometric and flux matrices. Here, in order to calculate optimal flux distributions, constraints such as maximum uptake/secretion rate and reversibility/irreversibility of reactions were added to the model. Once a solution space was defined by the constraints, the optimal flux distributions were determined for the optimization of objective function. The prediction module has the function of quantitatively predicting biomass and substrate concentrations in an unsteady state over time. The quasi-steady-state condition on the metabolic network was used to predict dynamic profiles of growth and substrate concentration. The batch time was first divided into small time intervals and given initial parameters from the experiment such as cell density, substrate uptake rate, and aerobic condition, and then the maximal growth rate was obtained from the FBA. With these data, a standard differential equation was repeatedly calculated through an iterative algorithm to predict a concentration change in the next time step. Lastly, the variation of concentration over all the intervals was predicted by integrating small time intervals [1, 4, 13].

The growth pattern for *E. coli* that was aerobically grown on glucose and tryptophan in batch cultures showed a diauxic growth curve. After the glucose culture was selectively consumed by catabolite repression, tryptophan was metabolized as an alternative carbon source. This diauxic growth can be divided into a glucose metabolic phase, a lag phase, and a tryptophan metabolic phase.

As shown in Fig. 3B, tryptophan was not metabolized during the first growth phase, because *tnaAB*, which is used to transfer and degrade tryptophan, was downregulated by glucose. Glucose is always metabolized first in preference to other carbon sources by inhibiting the synthesis of cAMP. This type of glucose effect, known as catabolite repression, is found in many different bacteria. While the concentration of glucose exponentially decreased (0 h–4.83 h), byproducts were secreted. The maximal growth rate was determined to be 0.50 h^{-1}. The initial concentration of glucose, maximal uptake rate of glucose, and initial concentration of biomass were 0.92 mM glucose, 5.4 mmol glucose/g·DCW·h, and 0.0086 g biomass/l, respectively, and these were used as the parameters for the simulation. The lag phase, which temporarily stops the growth of *E. coli*, represents the time required to complete the induction and synthesis of the enzymes, *tnaAB*, which are necessary for utilizing tryptophan. The phase in which *E. coli* did not grow was 0.35 h long, as shown in Fig. 3C. The second growth phase refers to the period when the growth of *E. coli* resumes by metabolizing tryptophan after synthesizing *tnaAB* during the lag time. The initial concentration and maximal uptake rate of tryptophan were 1 mM tryptophan and 1.32 mmol tryptophan/g·DCW·h, respectively.

Determination of these parameters, the initial concentrations, and maximal uptake rates from aerobic batch experiments on glucose and tryptophan are essential for the simulation. If all of these parameters were determined, the behavior of *E. coli* could be predicted through our modeling approach as mentioned above.

A hierarchical regulatory network can simulate the control mechanisms of genes over time by connecting to the dynamic simulation module. When both glucose and tryptophan exist in the extracellular environment, the network can provide information about a catabolite repression mechanism.

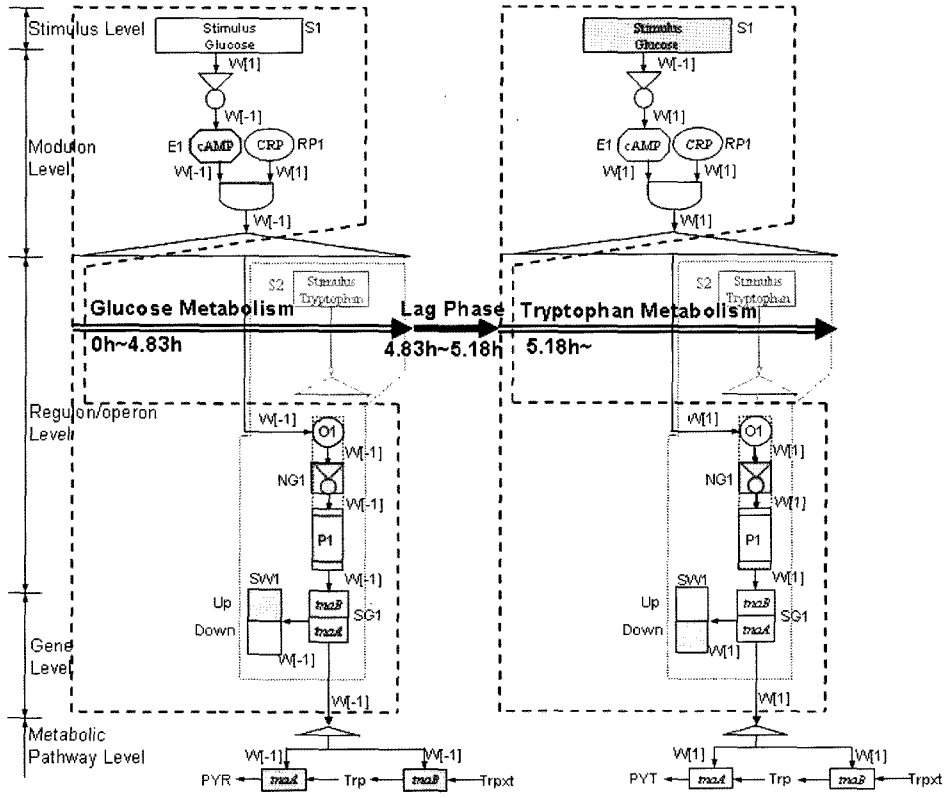
As shown in Fig. 4A, glucose activated S1 at the stimulus level located in the highest part of the hierarchical regulatory network, and the weight became W[1]. By "not gate," it is changed into W[−1] which propagated to the module level through the arc. The cAMP/CRP complex could not be produced, because cAMP is downregulated by W[−1]. This W[−1] was continuously transferred to the regulon/operon level. Because of positive control, W[−1] inactivated genes, *tnaAB*, and did not metabolize tryptophan. This inactive state of *tnaAB* was maintained for 4.83 h, when the concentration of glucose reached zero.

When the glucose was completely utilized, the S1 node became inactive, and the weight, W[-1], was changed to W[1] by "not gate." This weight activated the cAMP/CRP complex and *tnaAB* at the gene level. Eventually, tryptophan was metabolized by an upregulated *tnaAB*. In this way, the hierarchical regulatory network allows us to know the gene regulatory mechanism and the state of each gene over time.

In this study, when both glucose (sugar) and tryptophan (amino acid) were present, the parameters for simulation were determined from the experiment, and the validity of our suggested modeling approach was examined with three temporal profiles showing the experimental data and the corresponding simulation results.

The temporal profiles of cell density, glucose concentration, and tryptophan concentration were in good agreement with the simulation results (Fig. 3). Parameters such as the initial glucose concentration, initial tryptophan concentration, initial cell density, uptake rate of glucose, and uptake rate of tryptophan were used in the simulation. From the results, the growth rate during glucose and tryptophan metabolism was predicted to be 0.5 h^{-1} and 0.02 h^{-1}, respectively. The low growth rate by tryptophan metabolism

A



B

Phase	Time (h)	S1	S2	E1	RP1	O1	NG1	P1	SG1	SW1
Glucose metabolism	0	*A*	*N*	*I*	*A*	*I*	*I*	*I*	*I*	*I*
Lag	4.83	*I*	*N*	*A*	*A*	*A*	*I*	*A*	*A*	*A*
Tryptophan metabolism	5.18	*I*	*N*	*A*	*A*	*A*	*I*	*A*	*A*	*A*

A: Active; *I*: Inactive; *N*: None.

Fig. 4. Dynamic simulation for gene regulation.
A. Weight propagation in glucose and tryptophan. **B.** Active/inactive regulatory elements in a hierarchical regulatory network with time.

means that the precursors required for 1 g of biomass, such as G6P, F6P, R5P, E4P, T3P, 3PG, PEP, PYR, AcCoA, OAA, AKG, and SuccCoA, were produced in small quantities for glucose. It was identified by the FBA that the intracellular fluxes related to biomass precursors were considerably reduced. In particular, the uptake rate in *E. coli* grown on tryptophan (1.32 mmol tryptophan/g·DCW·h) was shown to be 75.5% reduced, compared with that observed in cells grown on glucose (5.4 mmol glucose/g·DCW·h).

As shown in Fig. 3B, the amount of tryptophan was uniformly maintained during glucose metabolism. Generally, tryptophan is produced as a byproduct while glucose is utilized. However, TrpR was upregulated, because an excess of tryptophan was present in the extracellular environment. This regulatory protein keeps *E. coli* from secreting tryptophan to block transcription of the *trp* operon. In addition, while tryptophan was used, we could deduce that the *trp* operon was continuously downregulated from the linear reduction of tryptophan. The lag time in our previous paper was set to 0.5 h, but we obtained more correct simulation results with a lag time of 0.35 h. This lag time can be slightly changed according to type and concentration of substrate in the second growth phase.

The reutilization of byproducts (representatively, acetate) from metabolizing glucose has to be considered for correct

simulation results. When glucose concentrations are above 10 mM, the resulting byproducts are relatively insufficient to support growth after the glucose metabolism phase, because of increased cell density, and therefore, there may be relatively little growth in the second period [13]. Furthermore, when lactose is used in the second growth phase, the effect of byproducts on cell growth may be small, owing to the relatively high growth rate of lactose [2]. However, in the case of low concentrations of glucose and low growth rate of tryptophan, the reutilization of byproducts could be a factor for correctly simulating cell metabolism, because the growth rate ($0.03\ h^{-1}$) by using byproduct was higher than that ($0.02\ h^{-1}$) using tryptophan.

In summary, to date, the transcriptional regulatory constraints have been described by only if-then rules including Boolean logic. In order to solve the problems inherent in this approach, we suggested an improved modeling method that incorporates a hierarchical regulatory network into the FBA. In the present study, an aerobic growth experiment on glucose and tryptophan was carried out to validate our modeling method in the central metabolic pathway of *E. coli*. The parameters for the simulation were determined from the experiment. The simulation results from our modeling method were very similar to the experimental data. The upregulated/downregulated state of the genes in response to the environmental changes was also identified intuitively through the hierarchical regulatory network. The present results support the study on the transcriptional regulatory network and increase the predictive capabilities of *in silico* models.

Acknowledgments

Sung Gun Lee and Yu Jin Kim contributed equally to this work, which was supported by the Brain Korea 21 Project in 2006, and support from the Ministry of Commerce, Industry and Energy of Korea through CANSMC is gratefully acknowledged.

References

1. Covert, W. C., C. H. Schilling, and B. O. Palsson. 2001. Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* **213:** 73–88.
2. Covert, W. C. and B. O. Palsson. 2002. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* **277:** 28058–28064.
3. Edwards, J. S., M. Covert, and B. O. Palsson. 2002. Metabolic modeling of microbes: The flux-balance approach. *Environ. Microbiol.* **4:** 133–140.
4. Lee, S. G., K. S. Hwang, and C. M. Kim. 2005. Dynamic behavior of regulatory elements in the hierarchical regulatory network of various carbon sources-grown *Escherichia coli*. *J. Microbiol. Biotechnol.* **15:** 551–559.
5. Lee, S. G., C. M. Kim, and K. S. Hwang. 2005. Development of a software tool for *in silico* simulation of *Escherichia coli* using a visual programming environment. *J. Biotech.* **119:** 87–92.
6. Lee, S. Y. and E. T. Papoutsakis. 1999. *Metabolic Flux Balance Analysis*, pp. 13–57. Marcel Dekker, New York, U.S.A.
7. Liao, J. C. and M. K. Oh. 1999. Toward predicting metabolic fluxes in metabolically engineered strains. *Metab. Eng.* **1:** 214–223.
8. Park, S. M., H. S. Kang, and D. H. Park. 2004. Metabolic flux shift of *Weissella kimchii* sk10 grown under aerobic conditions. *J. Microbiol. Biotechnol.* **14:** 919–923.
9. Park, S. M., H. S. Kang, D. W. Park, and D. H. Park. 2005. Electrochemical control of metabolic flux of *Weissella kimchii* sk10: Neutral red immobilized in cytoplasmic membrane as electron channel. *J. Microbiol. Biotechnol.* **15:** 80–85.
10. Reed, J. L. and B. O. Palsson. 2003. Thirteen years of building constraint-based *in silico* models of *Escherichia coli*. *J. Bacteriol.* **185:** 2692–2699.
11. Schilling, C. H., J. W. Edwards, D. Letscher, and B. O. Palsson. 2000. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. *Biotechnol. Bioeng.* **71:** 286–306.
12. Varma, A., B. W. Boesch, and B. O. Palsson. 1993. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl. Environ. Microbiol.* **59:** 2465–2473.
13. Varma, A. and B. O. Palsson. 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60:** 3724–3731.
14. Winston, P. H. 1992. *Artificial Intelligence*, pp. 111–137. 3rd Ed. Addison Wesley, Massachusetts, U.S.A.