J. Microbiol. Biotechnol. (2007), **17**(3), 496–510

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

A Discrete Mathematical Model Applied to Genetic Regulation and Metabolic Networks

ASENJO, J. A.*, P. RAMIREZ, I. RAPAPORT[1], J. ARACENA[†], E. GOLES[1], AND B. A. ANDREWS[1]

Centre for Biochemical Engineering and Biotechnology, Department of Chemical Engineering and Biotechnology
[1]Centre for Mathematical Modelling, Institute for Cell Dynamics and Biotechnology, University of Chile, Beauchef 861, Santiago, Chile
[†]Present address: Department of Mathematical Engineering, University of Concepción, Concepción, Chile

Received: November 7, 2006
Accepted: December 18, 2006

Abstract This paper describes the use of a discrete mathematical model to represent the basic mechanisms of regulation of the bacteria *E. coli* in batch fermentation. The specific phenomena studied were the changes in metabolism and genetic regulation when the bacteria use three different carbon substrates (glucose, glycerol, and acetate). The model correctly predicts the behavior of *E. coli vis-à-vis* substrate mixtures. In a mixture of glucose, glycerol, and acetate, it prefers glucose, then glycerol, and finally acetate. The model included 67 nodes; 28 were genes, 20 enzymes, and 19 regulators/biochemical compounds. The model represents both the genetic regulation and metabolic networks in an integrated form, which is how they function biologically. This is one of the first attempts to include both of these networks in one model. Previously, discrete mathematical models were used only to describe genetic regulation networks. The study of the network dynamics generated 8 (2^3) fixed points, one for each nutrient configuration (substrate mixture) in the medium. The fixed points of the discrete model reflect the phenotypes described. Gene expression and the patterns of the metabolic fluxes generated are described accurately. The activation of the gene regulation network depends basically on the presence of glucose and glycerol. The model predicts the behavior when mixed carbon sources are utilized as well as when there is no carbon source present. Fictitious jokers (Joker1, Joker2, and Repressor SdhC) had to be created to control 12 genes whose regulation mechanism is unknown, since glycerol and glucose do not act directly on the genes. The approach presented in this paper is particularly useful to investigate potential unknown gene regulation mechanisms; such a novel approach can also be used to describe other gene regulation situations such as the comparison between non-recombinant and recombinant yeast strain, producing recombinant proteins, presently under investigation in our group.

*Corresponding author
Phone: 56-2-978-4284; Fax: 56-2-699-1084;
E-mail: juasenjo@ing.jchile.cl

Keywords: Discrete model, metabolic networks, genetic regulation

Metabolic and genetic regulation networks are complex, dynamic systems that are constantly interacting. Changes in concentration levels of components of the metabolic pathways provoke changes in gene expression in the cells. These changes in gene expression result in alterations in enzyme levels and also affect their activity and thus metabolic fluxes. In the last few years, new technologies have been developed that permit the generation of massive amounts of data about genetic regulation and metabolic networks. These technologies include microarrays (gene expression) and metabolic flux analysis (metabolic pathways). Previously, discrete mathematical models had mainly been applied to genetic regulation networks, but most behavior finally depends on the metabolism. Recently, there have been developments in this area; for example, models have been built to find correlations between biochemical effectors and morphogenesis [1] and to investigate the existence of unknown components in a gene regulation network [2–4, 12]. Covert and Palsson [6] have described a constraint-based model that is one of the few that integrates metabolic pathways and gene regulation. There exists sufficient "know how" to attempt the simultaneous simulation of genetic regulation and a metabolic network that will allow a better understanding of the interactions within a cell and also of the data generated by gene expression microarrays and metabolic flux analysis.

This paper describes the use of a discrete mathematical model to represent the mechanisms of regulation of the bacteria *E. coli* in batch fermentation. *E. coli* can use various substrates as energy and metabolite sources, but shows distinct preferences between different substrates. The specific phenomena studied were the changes in

metabolism and genetic regulation when the bacteria use three different carbon substrates (glucose, glycerol, and acetate) and mixtures of the three. Experimental values obtained in microarray experiments by Oh and Liao [16]; and values found by Holms [10] for the metabolic fluxes were used to develop this model.

METHODS

The use, by *E. coli*, of different carbon sources was studied, where an extracellular concentration of 40 mM for each substrate was considered. The metabolic pathways of glycolysis and the citric acid cycle and the genetic regulation associated with these pathways were analyzed. Fig. 1 shows the pathways and enzymes involved. *E. coli* adapts its metabolism to the available carbon source. When presented with a mixture of carbon sources (for example, a medium containing glucose, glycerol, and acetate), *E. coli* will show a distinct preference for one substrate over the others, where glucose is always used before glycerol or acetate and glycerol before acetate (Liao, 2002, personal communication). This is associated with genetic regulation, which forces the cell to metabolize the substrates in the most efficient way. The different directions of the fluxes depend on whether the reactions take place in the direction of the products or the reactants.

Metabolic Fluxes on Glucose

Fig. 2 shows the direction of the fluxes when glucose is used as the carbon source in glycolysis and the TCA cycle. Glucose enters the cell through the PTS system, which is briefly described later. A detailed description of the glycolysis can be found in standard textbooks [15]. Enzyme names can be found in the Nomenclature section. PEP is a very important component in this pathway, as it is the branch point of 2 pathways; one from PEP to OAA and the other from PEO to PYR. The transformation of PEP to OAA is catalyzed by phosphoenolpyruvate carboxylase (Ppc) and, simultaneously, the inverse reaction is catalyzed by phosphoenolpyruvate carboxykinase (PckA). The net flux of these 2 reactions is from PEP to OAA. In the pathway PEP to PYR, the enzymes pyruvate kinases I and II (PykF and PykA, indicated as Pyk in Fig. 2) are responsible for catalysis and the reverse reaction is catalyzed by phosphoenolpyruvate synthetase (Pps). The net flux is from PEP to PYR. When glucose is used as a substrate, the glyoxylate shunt of the TCA cycle is not used.

Acetate is produced during growth on glucose as one of the principal secondary products of fermentation. The first

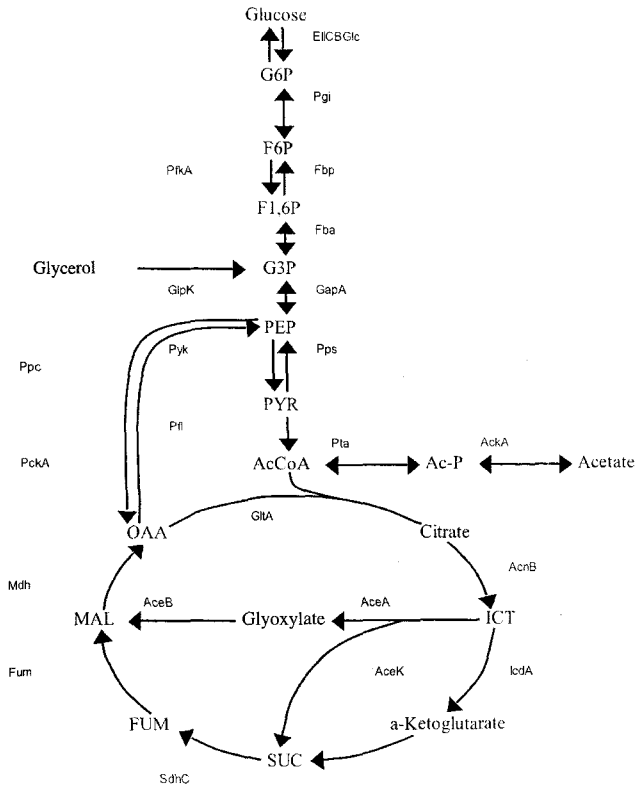

Fig. 1. Metabolic pathways of glycolysis and the citric acid cycle in *E. coli* and the enzymes involved in each step.

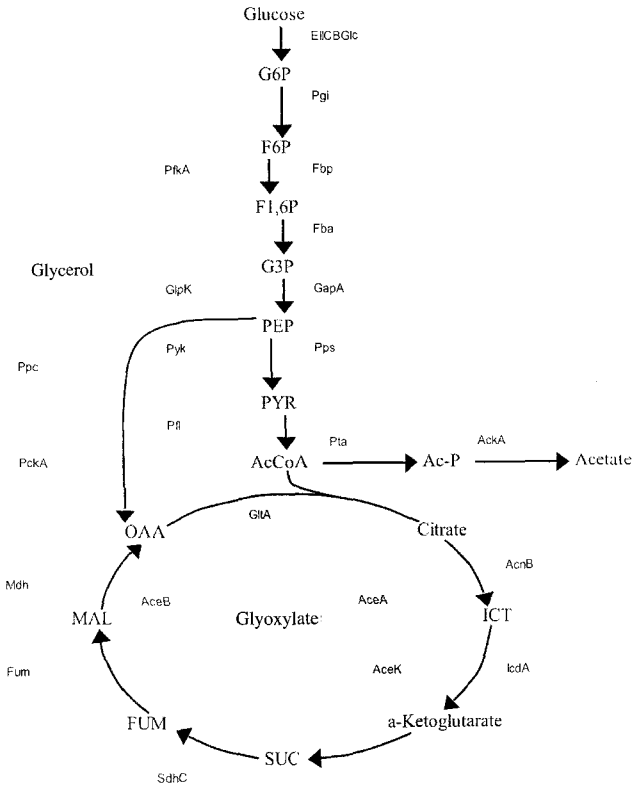

Fig. 2. Direction of metabolic fluxes obtained when *E. coli* uses glucose as a carbon source.

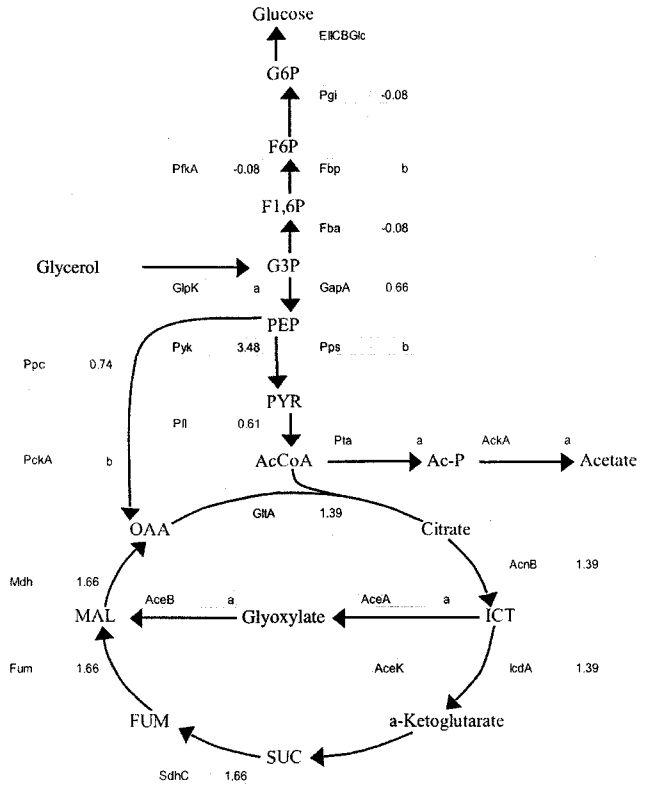

Fig. 3. Direction of the metabolic fluxes obtained when *E. coli* uses glycerol as a carbon source.
The fluxes are expressed as the ratio between growth on glycerol and growth on glucose. These values were taken from Holms [10]; a=unknown flux, b=flux cannot be separated from the reverse reaction catalyzed by another enzyme.

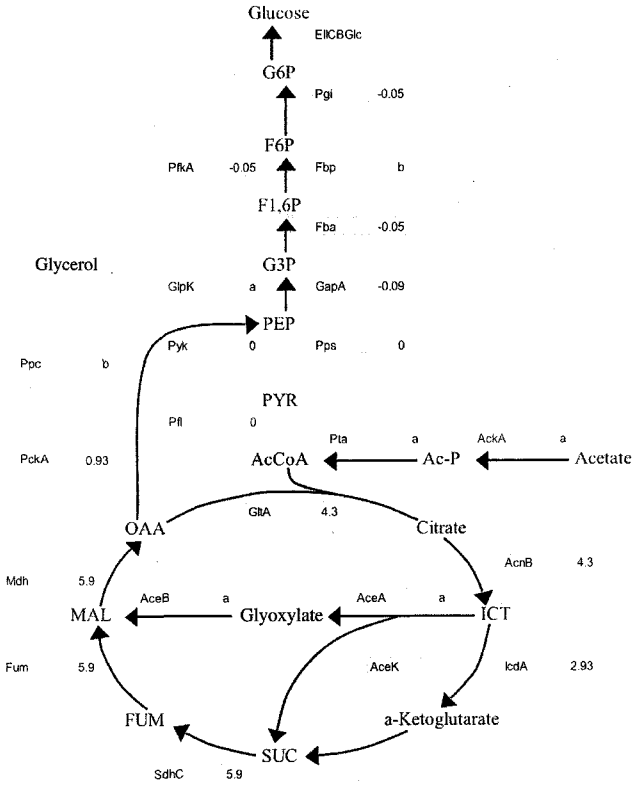

Fig. 4. Direction of the metabolic fluxes obtained when *E. coli* uses acetate as a carbon source.
The fluxes are expressed as the ratio between growth on acetate and growth on glucose. These values were taken from Holms [10]; a=unknown flux, b=flux cannot be separated from the reverse reaction catalyzed by another enzyme.

enzyme in the pathway is acetyl phosphate transferase (Pta), which converts AcCoA to Ac-P, which is then transformed to acetate by the enzyme acetate kinase (AckA).

Metabolic Fluxes on Glycerol

The internal metabolite concentrations and the metabolic fluxes are lower when glycerol is used as a substrate [10], as shown in Fig. 3 (fluxes are shown as the ratio of flux on glycerol compared with flux on glucose). This is because *E. coli* does not metabolize glycerol as easily as glucose. The entry of glycerol into the pathway is via transformation to G3P catalyzed by glycerol kinase (GlpK). The G3P is a substrate for the production of glucose (flux direction is reversed compared with Fig. 2) and glycolysis leading to the TCA cycle. In this case, the glyoxylate shunt is active.

Metabolic Fluxes on Acetate

The biggest changes in flux compared with growth on glucose occur when *E. coli* is using acetate as a carbon source (Fig. 4). The utilization of acetate is more complex than that of glucose or glycerol, as acetate enters in the final stages of glycolysis and the gluconeogenesis pathway is also more complicated. Acetate enters the pathway via Ac-P and AcCoA. The major changes in flux compared with glucose are higher fluxes in the glyoxylate shunt that produces OAA that is used to produce PEP which is then used to produce glucose. In the case of acetate utilization, the net flux is from OAA to PEP [16].

Efficient Adaptation: Regulation of the Amount of Active Enzymes

In *E. coli*, the amount of active enzymes is regulated depending on the substrate being consumed. Enzyme activity is controlled in 2 ways; by modification of the expression of the enzyme and by effectors or inhibitors. The regulation of the amount of mRNA produced is the principal control mechanism, and the quantities produced vary depending on whether glucose, glycerol, or acetate is being used as the carbon source. Oh and Liao [16] studied the expression of mRNA using microarray assays of *E. coli* growing on those 3 substrates.

Fig. 5 shows the transcription levels and metabolic fluxes of genes involved in the metabolism of glycerol compared with those of glucose in glycolysis and the TCA cycle. The flux ratios are close to 1 in most cases, indicating that the metabolism of glycerol is similar to that of glucose, and

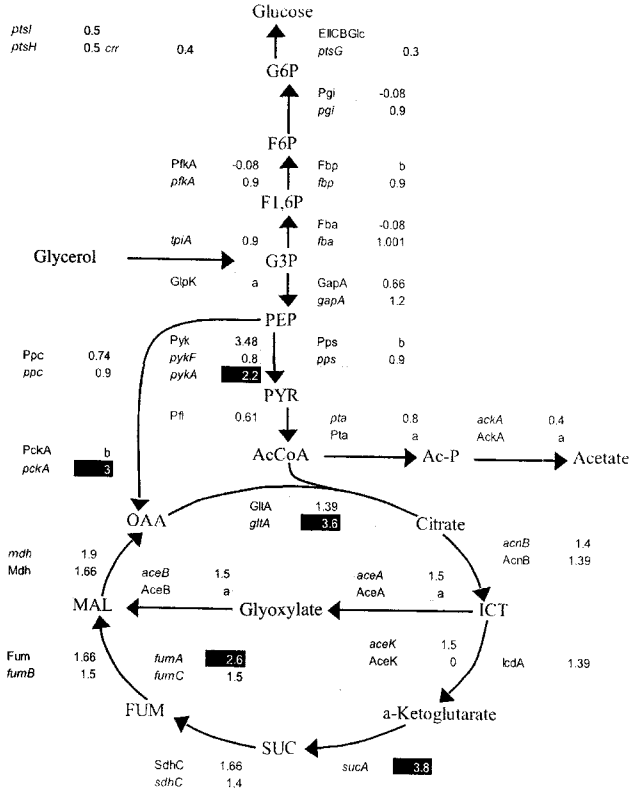

Fig. 5. Glycolysis and the TCA cycle: Transcription levels [16] and metabolic fluxes [10] of the genes involved in glycerol metabolism compared with those involved in glucose metabolism.

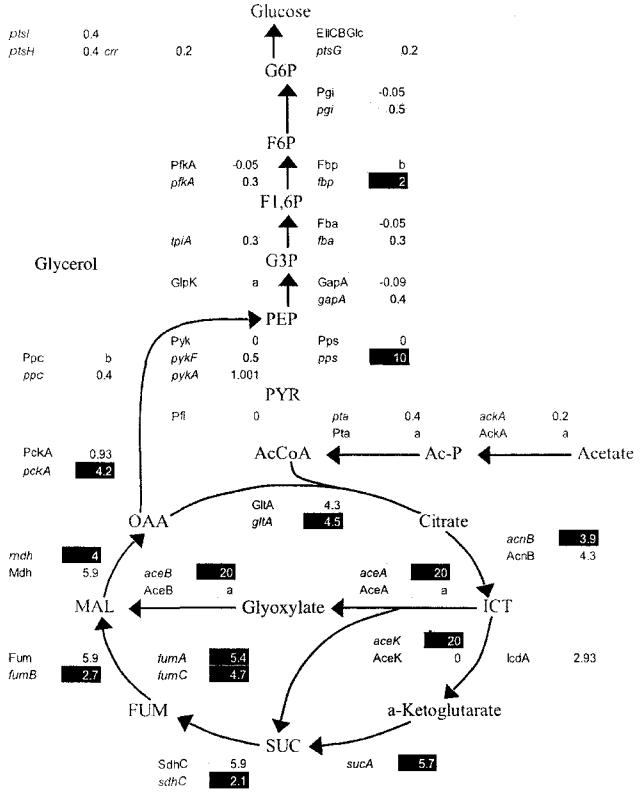

Fig. 6. Glycolysis and the TCA cycle: Transcription levels [16] and metabolic fluxes [10] of the genes involved in the metabolism of acetate compared with those involved in glucose metabolism.

only a few reactions are different. Some enzymes of the TCA cycle are overexpressed to supply the necessary metabolites. The enzymes of the PT (phosphotransferase) system are inactive when the substrate is glycerol, as this system is not used.

Fig. 6 shows the transcription levels and metabolic fluxes of genes involved in the metabolism of acetate compared with those of glucose in glycolysis and the TCA cycle. In this case many of the enzymes associated with glycolysis are underexpressed. Two enzymes of the gluconeogenesis pathway, Fbp and PckA, are overexpressed when acetate is used as a substrate. Enzymes of the TCA cycle and the glyoxylate pathway are also overexpressed.

Mechanisms and Characterization of the Gene Regulation Network

To simulate the behavior shown in Figs. 5 and 6, the mathematical model must take into account the regulatory mechanisms. Mechanisms include components that increase or inhibit enzyme activity, enzymes that enhance or inhibit the activity of other enzymes or messengers, and components that affect the genome to induce or repress gene expression. The first 2 are allosteric regulation mechanisms (change enzyme activity) and the third is a mechanism of gene expression regulation. Of the gene regulation network studied here, some regulate a group of genes or enzymes, and others only single elements. In the PT system, the fructose repressor (FruR) and cyclic AMP (cAMP) regulate the expression of a large number of enzymes. The PT system is very important in the initial stages of glycolysis; it contains 4 enzymes and their expression is controlled by the global repressor Mlc [11, 13, 21].

Cyclic AMP is a repressor that acts when concentrations of extracellular glucose are low. This system has 3 elements; cAMP, CyaA, and a cAMP receptor protein (Crp). CyaA is an enzyme that catalyzes the production of cAMP and is regulated by EIIAGlc and G3P [8]. The level of cAMP varies with different carbon sources; its level is minimal with glucose, intermediate with glycerol, and maximal with acetate. The cAMP receptor protein detects the concentration of cAMP and acts on the genes activating or deactivating their expression. In the metabolic pathways studied here, the proteins regulated by cAMP.Crp are AcnB [7], FumA, FumC [18], and Mdh [17].

Another mechanism of regulation is that dependent on a fructose repressor (FruR). The FruR activates the transcription of various genes that are used in the metabolism of substrates other than glucose, and inhibits

the transcription of genes producing enzymes used in the metabolism of glucose [5, 14]. The activity of FruR is inhibited by F1,6P and F6P, as these compounds are present in high concentrations during the metabolism of glucose, and FruR is inactivated. When glycerol is used as substrate, the activity of FruR is intermediate, and when the substrate is acetate, it is high [14, 19].

Model System

The objective of this research was to develop a discrete mathematical model to represent mechanisms of metabolic regulation in *E. coli*. The phenomenon studied was the changes in the genetic regulation network and the metabolic network when *E. coli* uses different carbon sources in batch fermentation. *E. coli* adapts to its nutrient environment; its metabolic fluxes and quantity of enzymes expressed depends on the available carbon source. In Figs. 2, 3, and 4, the metabolic fluxes associated with the use of each carbon source has been described; glucose, glycerol, and acetate, respectively.

The methodology consists of the model definition and the study of its dynamics. The model definition includes its components: the variables and their activation functions. We have studied the fixed points of the model in order to understand its phenotypes.

Model Definition

The model represents two networks that interact between them: the gene regulation network and the metabolic network. The behavior of both is determined by the carbon source used; these are represented in the system by the biochemical compounds of the metabolic network.

Definition of Variables and States

Carbon Sources. The three carbon sources are represented by the variables $x_{Acetate}$, $x_{Glycerol}$, and $x_{Glucose}$. In all three cases, the value 0 means that the carbon source is not available in the medium. The model incorporates biological knowledge on the global phenomena, which can be summarized as $x_{Glucose}>x_{Glycerol}>x_{Acetate}$. This is why the presence of the carbon source is represented by $x_{Acetate}=1$, $x_{Glycerol}=2$, and $x_{Glucose}=3$ for each case.

Gene Regulation Network. The nodes that are part of the gene regulation network are genes and expression regulators. To each gene (*pgi*, *fba*, *fbp*, etc.), a variable (x_{pgi}, x_{fba}, x_{fbp}, etc.) that reaches the states 1 and −1 has been associated. 1 indicates a high gene expression, whereas −1 indicates a low expression, and the activation function depends on the expression regulators. To the regulators (FruR, cAMP, Crp, etc.), a variable (x_{FruR}, x_{cAMP}, x_{Crp}, etc.) that reaches the values {0, 1, 2, 3} has also been associated. These values reflect the concentration of the regulators when there is no carbon source or when a particular carbon source is being utilized.

Metabolic Network. The nodes that are part of the metabolic network are biochemical compounds and metabolic fluxes. The biochemical compounds (F6P, PEP, G3P, etc.) have a variable associated to them (x_{F6P}, x_{PEP}, x_{G3P}, etc.) that reaches the values {0, 1, 2, 3}. These values, as well as the expression regulators, reflect the concentrations reached with the different carbon sources. The model only included the biochemical compounds that interact with the gene regulation network, F6P, F1, 6P, G3P, and ICT. The metabolic fluxes are related with the enzymes of the metabolic network (Pgi, Fba, Fbp, etc.), and they correspond to the reaction rates of these. The variables associated to the metabolic fluxes (x_{Pgi}, x_{Pba}, x_{Fbp}, etc.) reach the values {−3, −2, −1, 0, 1, 2, 3}. 0 means zero flux, 1 corresponds to the flux reached using acetate, 2 to the flux reached using glycerol, and 3 to the flux reached using glucose. The positive values represent fluxes that follow the sense of reaction defined for the enzyme, and the negative ones those in the opposite sense of reaction.

Integration of Gene Regulation and Metabolic Networks

This integration occurs in two ways:

- gene regulators depend on the concentration of the biochemical compounds, and
- enzymes, represented by the metabolic fluxes, require that the gene expressing them is active.

This integration is shown in the activation functions of the variables, which are defined in the following paragraphs.

Unknown of "Fictitious" Regulators. The literature to date describes many of the interactions between the different compounds, which allowed the mechanistic description of most of the network. However, some specific genes did not have sufficient documentation regarding their specific regulators, which made it necessary to generate unknown or "fictitious" regulators.

Table 1. Fictitious regulators associated with genes whose mechanism of regulation is unknown.

Fictitious regulator	Gene with unknown mechanism of regulation
Joker1	*ackA*
	fba
	fumb
	gltA
	pgi
	ppc
	pta
	tpiA
Joker2	*gapA*
	pykA
RepSdhC	*sdhC*
	sucA

Table 2. Values reached by the fictitious regulators with each carbon source.

Fictitious regulator	Variable	Glucose	Glycerol	Acetate
Joker1	x_{Joker1}	2	1	0
Joker2	x_{Joker2}	0	1	0
RepSdhC	$x_{RepSdhC}$	2	1	0

Three fictitious regulators were generated in order to explain the "undocumented" regulation of some genes in the network; these are Joker1, Joker2, and RepressorSdhC. Regarding the first two, no reference in the literature could be found and they were created for this model. Joker1 and Joker 2 regulate genes whose expression is activated or repressed in the presence of glucose and glycerol. RepressorSdhC represents an unknown "intermediary" that represses SdhC and SucA in the presence of glucose [20]; its activation depends on EIICBGlc. Table 1 shows the genes whose regulation could not be explained with literature data, and the fictitious (unknown) associated regulator. Table 2 shows the values reached by the fictitious regulators with each carbon source.

In the next section, the activation of the unknown or "fictitious" regulators is described as well as those of all the elements of the network. Finally, the logical activation functions are assigned to all the elements of the network.

Activation of Regulators and Biochemical Compounds. These elements are the ones that show which carbon source is being utilized and determine the regulation of the gene and metabolic network. The following list describes how these elements reach the different values.

- EIICBGlc ($x_{EIICBGlc} \in \{0, 3\}$) activated ($x_{EIICBGlc}=3$) means that it is dephosphorylated; this occurs in the presence of glucose ($x_{Glucose}=3$) in the external medium. $x_{EIICBGlc}=0$ in other cases.
- EIIAGlc ($x_{EIIAGlc} \in \{0, 1\}$) active ($x_{EIIAGlc}=1$) means that it is dephosphorylated; this occurs when EIICBGlc is activated. $x_{EIIAGlc}=0$ in other cases.
- Mlc ($x_{Mlc} \in \{-1, 1\}$) is activated ($x_{Mlc}=1$) if it is free; Mlc is inactivated ($x_{Mlc}=-1$) when it is captured by active EIICBGlc.
- G3P ($x_{G3P} \in \{0, 1, 2, 3\}$) depends on the metabolic fluxes that are generated as a product of the reaction: GapA, GlpK, and Fba. The concentration reached by G3P corresponds to the maximum of these fluxes. For example, if $x_{GapA}=-1$, $x_{GlpK}=2$, and $x_{Fba}=3$, then $x_{G3P}=3$.
- Cya ($x_{Cya} \in \{0, 1, 2\}$) reaches a state of maximum activation ($x_{Cya}=2$) when EIIAGlc is phosphorylated and with low concentrations of G3P ($x_{G3P}=0, 1$). Cya is partially activated when EIIAGlc is phosphorylated and G3P is generated in intermediate or high concentrations ($x_{G3P}=2, 3$). Cya is inactive ($x_{Cya}=0$) when EIIAGlc is dephosphorylated.
- Crp ($x_{Crp}=1$) is always active and present in the model.
- cAMP ($x_{cAMP} \in \{0, 1, 2\}$) depends directly on the activity of the enzyme Cya. This means that if $x_{Cya}=0$ then $x_{cAMP}=0$, if $x_{Cya}=1$ then $x_{cAMP}=1$, and if $x_{Cya}=2$ then $x_{cAMP}=2$.
- cAMP.Crp ($x_{cAMP.Crp} \in \{0, 1, 2\}$) is formed depending on the concentration of x_{cAMP}, and reaches the same values.
- F6P and F1,6P ($x_{F6P}, x_{F1,6P} \in \{0, 1, 2, 3\}$) depend on the metabolic fluxes that are generated. In the case of x_{F6P}, the metabolic fluxes that influence its formation are Pgi (x_{Pgi}), Fbp, and PfkA (these last two represented by x_{PfkA}, as they catalyze the same reaction). The concentration reached by F6P corresponds to the maximum flux of these enzymes. In the case of F1,6P, its concentration is the maximum allowed by the fluxes of the enzymes Fba (x_{Fba}), Fbp, and PfkA (x_{PfkA}).
- FruR ($x_{FruR} \in \{0, 1, 2\}$) is totally active ($x_{FruR}=2$) when the concentrations of F6P or F1,6P are low ($x_{F6P}, x_{F1,6P}=0, 1$). It is partially inhibited ($x_{FruR}=1$) at intermediate concentrations ($x_{F6P}, x_{F1,6P}=2$) and totally inhibited ($x_{FruR}=0$) at high concentrations ($x_{F6P}, x_{F1,6P}=3$).
- ICT ($x_{ICT} \in \{0, 1, 2, 3\}$), its concentration depends directly on the enzyme AcnB. When the flux through this enzyme is high, the concentration of ICT will also be high.
- The enzyme AceK ($x_{AceK} \in \{-1, 1\}$) has both kinase and phosphatase activities. AceK is active ($x_{AceK}=1$) when kinase activity predominates, and is inactive in the opposite case ($x_{AceK}=-1$). The activity of AceK is regulated by ICT and G3P: the greater the concentrations of these, the lower is the kinase activity and the greater the phosphatase activity.
- Joker1 ($x_{Joker1} \in \{0, 1, 2\}$) reaches its maximum activity ($x_{Joker1}=2$) when glucose is present in the medium. It reaches an intermediate level when glycerol is present ($x_{Joker1}=1$) and is inactive ($x_{Joker1}=0$) when neither of these carbon sources is present.
- Joker2 ($x_{Joker2} \in \{0, 1\}$) is activated ($x_{Joker1}=1$) when only glycerol is present in the medium, and is inactive ($x_{Joker1}=0$) in other cases.
- RepressorSdhC ($x_{RepressorSdhC} \in \{0, 1, 2\}$) depends on EIICBGlc and on the presence of glycerol. RepressorSdhC reaches its maximum state of activation ($x_{RepressorSdhC}=2$) when EIICBGlc is active. It reaches an intermediate state ($x_{RepressorSdhC}=1$) when glycerol is present in the medium and EIICBGlc is inactive.

Activation of Genes. The following list shows the genes regulated by the mechanisms described and a description of this regulation. All genes reach the states $\{-1, 1\}$, where -1 shows low expression and 1 shows high expression.

Regulated by PTS:

- *ptsI* (x_{ptsI}), *ptsH* (x_{ptsH}), *crr* (x_{crr}), and *ptsG* (x_{ptsG}). Their expression depends on the presence of cAMP.Crp and fundamentally on the concentration of Mlc. This group of genes is repressed when the concentration of free Mlc is high.
- *gapA* (x_{gapA}) is activated when EIICBGlc is dephosphorylated or when there is glycerol in the medium. *gapA* must depend on Joker2 to explain its activation in glycerol.

Regulated by Joker1:

• *ackA* (x_{ackA}) is controlled by Joker1, as its regulation is not sufficiently well documented. *ackA* is activated when glucose is present in the medium, which is equivalent with Joker1 being in its maximum state of activation.

• *fba* (x_{fba}). Its expression is high when the cell uses glucose or glycerol as carbon source, which is when Joker1 reaches high activation states.

• *fumB* (x_{fumB}). This gene is activated when the only carbon source is acetate, which is when Joker1 is inactive.

• *gltA* (x_{gltA}) is inactivated when Joker1 is active at its maximum level, which is when glucose is present in the medium.

• *pgi* (x_{pgi}), *ppc* (x_{ppc}), *pta* (x_{pta}), and *tpiA* (x_{tpiA}) are inhibited when Joker1 is in the minimum level of activation.

Regulated by Joker2:

• *pykA* (x_{pykA}) is only active when *E. coli* uses glycerol as carbon source, when Joker2 is active.

• *gapA* (x_{gapA}); this gene depends on EIICBGlc and Joker2. *gapA* is activated when Joker2 is activated or EIICBGlc is activated.

Regulated by cAmp.Crp:

• *acnB* (x_{acnB}) is regulated by cAMP.Crp; its expression increases when the level of cAMP.Crp are high.

• *fumA* (x_{fumA}) is activated when cAMP.Crp is active.

• *fumC* (x_{fumC}) and *mdh* (x_{mdh}) are activated when the levels of cAMP.Crp are maximum.

Regulated by FruR:

• *fbp* (x_{fbp}) and *pps* (x_{pps}). Their expression is activated when the level of FruR is maximum.

• *aceA*, *aceB*, and *aceK*. The expression of the ace operon, which includes *aceA* (x_{aceA}), *aceB* (x_{aceB}), and *aceK* (x_{aceK}) is activated when FruR is active.

• *pckA* (x_{pckA}) is activated when FruR is active.

• *pfkA* (x_{pfkA}) and *pykF* (x_{pykF}) are inhibited when the level of FruR is maximum.

Regulated by RepressorSdhC:

• *sdhC* (x_{sdhC}) is activated when RepressorSdhC is at the minimum level of activation.

• *sucA* (x_{sucA}) is inhibited when RepressorSdhC is at the minimum level of activation.

Activation of Metabolic Fluxes. The following list shows under which conditions the enzymes of the metabolic network catalyze the reactions and the direction of these fluxes.

• AceA (x_{AceA}). The flux of this enzyme is one way and depends on the flux of the enzyme AcnB and that the gene *aceA* is active.

• AceB (x_{AceB}). Its flux is one way and depends on the flux of AceA and that *aceB* is active.

• AckA (x_{AckA}). Its flux is bidirectional and, therefore, depends on the flux of Pta and on the presence of acetate in the medium ($x_{Acetate}$).

• AcnB (x_{AcnB}). The flux of this enzyme is unidirectional and depends on the flux of GltA.

• Fba (x_{Fba}). Its flux is bidirectional and depends on the fluxes of PfkA, GlpK, and GapA.

• Fbp and PfkA are represented by the same variable (x_{PfkA}) as they catalyze the same reaction, but in opposite directions. Their flux depends on Pgi and Fba.

• FumA, FumB, and FumC are represented by the variable (x_{Fum}), as these three enzymes have the same function. Their flux depends on the flux of SdhC.

• GapA (x_{Gap}) has a bidirectional flux and depends on GlpK, Fba, PckA, and Pps.

• Glpk (x_{GlpK}) has a unidirectional flux and is only active when glycerol is present ($x_{Glycerol}$) in the medium. It is inhibited by high levels of F1,6P and by active EIIAGlc.

• GltA (x_{GltA}) has a unidirectional flux that depends on the flux of Pta and Pfl.

• IcdA (x_{IcdA}) has a unidirectional flux that depends on the flux of GltA and on the activity of AceK.

• Mdh (x_{Mdh}) has a unidirectional flux that depends on the fluxes of AceB and Fumarases (x_{Fum}). Its activity is controlled by the expression level of the gene *mdh*.

• PckA (x_{PckA}) has a unidirectional flux that depends on the fluxes of Mdh and Ppc. Its activity is controlled by the expression level of the gene *pckA*.

• Pfl (x_{Pfl}) has a unidirectional flux that depends on the fluxes of PykF and PykA (x_{Pyk}).

• Pgi (x_{Pgi}) has a unidirectional flux that depends on EIICBGlc being active and on the flux of PfkA.

• Ppc (x_{Ppc}) has a unidirectional flux that depends on the fluxes of GapA and Pps, and the gene *ppc*.

• Pps (x_{Pps}) has a unidirectional flux that depends on the flux of PykF and PykA, and on the gene *pps*.

• Pta (x_{Pta}) has a bidirectional flux and is dependent on the fluxes of Pfl and AckA.

• PykF and PykA, represented by x_{Pyk}, have unidirectional fluxes that depend on the flux of GapA and the expression of one of the genes *pykF* or *pykA*.

• SdhC (x_{SdhC}) has a unidirectional flux that depends on the flux of AceA and the expression of *sdhC*.

Logical Activation Functions. The logical activation functions represent the interactions and the way the activation states were determined for each element of the network; they are described in Table 3. The following functions were used to build the logical functions:

• Max ($x_1, x_2, \ldots, x_n$) gives the maximum of a set of n values.

• Min ($x_1, x_2, \ldots, x_n$) gives the minimum of a set of n values.

• MaxFlux (x_1, x_2) distinguishes which variable has the highest absolute value, when $x_1 \geq 0$ and $x_2 < 0$; it is defined as

$$\text{MaxFlux}(x_1, x_2) = \begin{cases} x_1, \text{ if } |x_1| \geq |x_2| \\ x_2 \text{ in other cases} \end{cases}$$

Table 3. Logical functions of the elements of the model.

Number	Name	Variable	Formula
1	*aceA*	x_{aceA}	$x_{aceA}=sign\_1(x_{FruR}-1.1)$
2	AceA	x_{AceA}	$x_{AceA}=sign0(x_{AcnB} * x_{aceA}-0.1)$
3	aceB	x_{aceB}	$x_{aceB}=sign\_1(x_{FruR}-1.1)$
4	AceB	x_{AceB}	$x_{AceB}=sign0(x_{AceA} * x_{aceB}-0.1)$
5	acek	x_{aceK}	$x_{aceK}=sign\_1(x_{FruR}-1.1)$
6	AceK	x_{AceK}	$x_{AceK}=sign\_1(sign0(x_{aceK}) * sign0(0-x_{ICT} * x_{G3P})-0.1)$
7	Acetate	$x_{Acetate}$	$x_{Acetate}=x_{Acetate}$
8	*ackA*	x_{ackA}	$x_{ackA}=sign\_1(x_{Joker1}-1.1)$
9	AckA	x_{AckA}	$x_{AckA}=MaxFlux(Max(x_{Pta}, 0), Min(-x_{Acetate}, 0))$
10	*acnB*	x_{acnB}	$x_{acnB}=sign\_1(x_{cAMP.Crp}-1.1)$
11	AcnB	x_{AcnB}	$x_{AcnB}=x_{GltA}$
12	cAMP	x_{cAMP}	$x_{cAMP}=x_{Cya}$
13	cAMP.Crp	$x_{cAMP.Crp}$	$x_{cAMP.Crp}=x_{cAMP} * x_{Crp}$
14	Crp	x_{Crp}	$x_{Crp}=1$
15	*crr*	x_{crr}	$x_{crr}=(-x_{Mlc}) * sign0(x_{cAMP.Crp}+0.5)$
16	Cya	x_{Cya}	$x_{Cya}=(1-x_{EIIAGlc}) * (1+sign0(1-x_{G3P}))$
17	EIIAGlc	$x_{EIIAGlc}$	$x_{EIIAGlc}=sign0(x_{EIICBGlc}-0.1)$
18	EIICBGlc	$x_{EIICBGlc}$	$x_{EIICBGlc}=x_{Glucose}$
19	F1,6P	$x_{F1,6P}$	$x_{F1,6P}=Max(x_{PfkA}, -x_{Fba})$
20	F6P	x_{F6P}	$x_{F6P}=Max(x_{Pgi}, -x_{PfkA})$
21	*fba*	x_{fba}	$x_{fba}=sign\_1(x_{Joker1}-0.1)$
22	Fba	x_{Fba}	$x_{Fba}=MaxFlux(Max(x_{PfkA}, 0), Min(-x_{GlpK}, x_{GapA}, 0))$
23	*fbp*	x_{fbp}	$x_{fbp}=sign\_1(x_{FruR}-1.1)$
24	FruR	x_{FruR}	$x_{FruR}=2-MaxFlux(MaxFlux(x_{F6P}-1, x_{F1,6P}-1), 0)$
25	*fuma*	x_{fuma}	$x_{fuma}=sign\_1(x_{cAMP.Crp}-0.1)$
26	*fumb*	x_{fumb}	$x_{fumb}=-sign\_1(x_{Joker1}-0.1)$
27	*fumc*	x_{fumc}	$x_{fumc}=sign\_1(x_{cAMP.Crp}-1.1)$
28	Fum	x_{Fum}	$x_{Fum}=x_{SdhC}$
29	G3P	x_{G3P}	$x_{G3P}=Max(x_{Fba}, x_{GlpK}, -x_{GapA}, 0)$
30	*gapA*	x_{gapA}	$x_{gaPA}=sign\_1(sign0(x_{EIICBGlc}-0.1)+x_{Joker2}-0.1)$
31	GapA	x_{GapA}	$x_{GapA}=MaxFlux(Max(x_{GlpK}, x_{Fba}, 0), Min(-x_{PckA}, -x_{Pps}, 0))$
32	GlpK	x_{GlpK}	$x_{GlpK}=Max(x_{Glycerol}, 0) * (sign0(2-x_{F1,6P}) * sign0(2-x_{EIICBGlc})$
33	*glta*	x_{glta}	$x_{glta}=-sign\_1(x_{Joker1}-1.1)$
34	GltA	x_{GltA}	$x_{GltA}=Max(-x_{Pta}, x_{Pfl}, 0)$
35	Glucose	$x_{Glucose}$	$x_{Glucose}=x_{Glucose}$
36	Glycerol	$x_{Glycerol}$	$x_{Glycerol}=x_{Glycerol}$
37	lcdA	x_{lcdA}	$x_{lcdA}=x_{AcnB} * sign0(-x_{AceK})$
38	ICT	x_{ICT}	$x_{ICT}=x_{AcnB}$
39	Joker1	x_{Joker1}	$x_{Joker1}=sign0(x_{Glucose}-0.1)+sign0(x_{Glucose}+x_{Glycerol}-0.1)$
40	Joker2	x_{Joker2}	$x_{Joker2}=sign0(x_{Glycerol}-x_{Glucose}-0.1)$
41	*mdh*	x_{mdh}	$x_{mdh}=sign\_1(x_{cAMP.Crp}-1.1)$
42	Mdh	x_{Mdh}	$x_{Mdh}=Max(x_{AceB}, x_{Fum}) * sign0(x_{mdh})$
43	Mlc	x_{Mlc}	$x_{Mlc}=-sign\_1(x_{EIICBGlc}-0.1)$
44	*pckA*	x_{pckA}	$x_{pckA}=sign\_1(x_{FruR}-0.1)$
45	PckA	x_{PckA}	$x_{PckA}=MaxFlux(x_{Mdh}, x_{Ppc}, 0) * sign0(x_{pckA})$
46	*pfkA*	x_{pfkA}	$x_{pfkA}=-sign\_1(x_{FruR}-1.1)$
47	PfkA	x_{PfkA}	$x_{PfkA}=MaxFlux(Max(x_{Pgi}, 0), Min(x_{Fba}, 0))$
48	Pfl	x_{Pfl}	$x_{Pfl}=x_{Pyk}$
49	*pgi*	x_{pgi}	$x_{pgi}=sign\_1(x_{Joker1}-0.1)$
50	Pgi	x_{Pgi}	$x_{Pgi}=MaxFlux(x_{EIICBGlc}, Min(x_{PfkA}, 0))$
51	*ppc*	x_{ppc}	$x_{ppc}=sign\_1(x_{Joker1}-0.1)$
52	Ppc	x_{Ppc}	$x_{Ppc}=Max(x_{GapA}, x_{Pps}, 0) * sign0(x_{ppc})$
53	*pps*	x_{pps}	$x_{pps}=sign\_1(x_{FruR}-1.1)$

Table 3. Continued.

Number	Name	Variable	Formula
54	Pps	x_{Pps}	$x_{Pps}=s_{Pyk} * sign0(x_{pps})$
55	*pta*	x_{pta}	$x_{pta}=sign\_1(x_{Joker1}-0.1)$
56	Pta	x_{Pta}	$x_{Pta}=MaxFlux(x_{Pfl}, Min(x_{AckA}, 0))$
57	*ptsG*	x_{ptsG}	$x_{ptsG}=-x_{Mlc} * sign0(x_{cAMP.Crp}+0.5)$
58	*ptsH*	x_{ptsH}	$x_{ptsH}=-x_{Mlc} * sign0(x_{cAMP.Crp}+0.5)$
59	*ptsI*	x_{ptsI}	$x_{ptsI}=-x_{Mlc} * sign0(x_{cAMP.Crp}+0.5)$
60	*pykA*	x_{pykA}	$x_{pykA}=sign\_1(x_{Joker2}-0.1)$
61	Pyk	x_{Pyk}	$x_{Pyk}=Max(x_{GapA}, x_{PckA}, 0) * sign0(x_{pykF}+x_{pykA})$
62	*pykF*	x_{pykF}	$x_{pykF}=-sign\_1(x_{FruR}-1.1)$
63	RepSdhC	$x_{RepSdhC}$	$x_{RepSdhC}=sign0(x_{EIICBGlc}-0.1)+sign0(x_{EIICBGlc}+x_{Glycerol}-0.1)$
64	*sdhC*	x_{sdhC}	$x_{sdhC}=-sign\_1(x_{RepSdhC}-0.1)$
65	SdhC	x_{SdhC}	$x_{SdhC}=x_{AceA} * sign0(x_{sdhC})$
66	*sucA*	x_{sucA}	$x_{sucA}=-sign\_1(x_{RepSdhC}-1.1)$
67	*tpia*	x_{tpia}	$x_{tpia}=sign\_1(x_{Joker1}-0.1)$

- $sign0\ (x)=\begin{cases} 1, \text{if } x \geq 0 \\ 0, \text{in other cases} \end{cases}$
- $sign{-}1\ (x)=\begin{cases} 1, \text{if } x \geq 0 \\ -1, \text{in other cases} \end{cases}$

Description of Phenotypes. The following step for model definition is the description of the phenotypes reached by the cell using the different carbon sources. They correspond to the stable states the system reaches. Tables 4 and 5 show the values of the variables that represent the genes and the metabolic fluxes, respectively. These tables also include the experimental values reached by the enzymes and genes in each phenotype. In the case of the genes expressed, these correspond qualitatively to the values obtained in microarray experiments by Oh and Liao [16]; for the metabolic fluxes, the values found by Holms [10] were used.

Table 6 shows the values of the variables that represent the regulators and biochemical compounds in each phenotype. These values are due to the fact that in glycolysis fluxes, the highest values are obtained on glucose and the lowest

Table 4. Enzymes, associated variable, flux ratios, and value reached with each carbon source (phenotype).

Enzyme	Variable	Glucose		Glycerol		Acetate	
		Flux ratio	Value	Flux ratio	Value	Flux ratio	Value
AceA	x_{AceA}	1	0	a	0	a	1
AceB	x_{AceB}	1	0	a	0	a	1
AckA	x_{AckA}	1	3	a	2	a	−1
AcnB	x_{AcnB}	1	3	1.39	2	4.30	1
Fba	x_{Fba}	1	3	−0.08	−2	−0.05	−1
Fum	x_{Fum}	1	0	1.66	0	5.90	1
GapA	x_{GapA}	1	3	0.66	2	−0.09	−1
GlpK	x_{GlpK}	1	0	a	2	a	0
GltA	x_{GltA}	1	3	1.39	2	4.30	1
IcdA	x_{IcdA}	1	3	1.39	2	2.93	1
Mdh	x_{Mdh}	1	0	1.66	0	5.90	1
PckA	x_{PckA}	1	0	b	2	0.93	1
PfkA	x_{PfkA}	1	3	−0.08	−2	−0.05	−1
Pfl	x_{Pfl}	1	3	0.61	2	0.00	0
Pgi	x_{Pgi}	1	3	−0.08	−2	−0.05	−1
Ppc	x_{Ppc}	1	3	0.74	2	b	0
Pps	x_{Pps}	1	0	b	0	0.00	0
Pta	x_{Pta}	1	3	a	2	a	−1
Pyk	x_{Pyk}	1	3	3.48	2	0.00	0
SdhC	x_{SdhC}	1	0	1.66	0	5.90	1

Table 5. Genes, associated variable, transcription ratios, and states reached with each carbon source.

Gene	Variable	Glucose		Glycerol		Acetate	
		Transcription ratio	State	Transcription ratio	State	Transcription ratio	State
aceA	x_{aceA}	1	−1	1.50	−1	20.00	1
aceB	x_{aceB}	1	−1	1.50	−1	20.00	1
aceK	x_{aceK}	1	−1	1.50	−1	20.00	1
ackA	x_{ackA}	1	1	0.40	−1	0.20	−1
acnB	x_{acnB}	1	−1	1.40	−1	3.90	1
crr	x_{crr}	1	1	0.40	−1	0.20	−1
fba	x_{fba}	1	1	1.00	1	0.30	−1
fbp	x_{fbp}	1	−1	0.90	−1	2.00	1
fuma	x_{fuma}	1	−1	2.60	1	5.40	1
fumb	x_{fumb}	1	−1	1.50	−1	2.70	1
fumc	x_{fumc}	1	−1	1.50	−1	4.70	1
gapA	x_{gapA}	1	1	1.20	1	0.40	−1
gltA	x_{gltA}	1	−1	3.60	1	4.50	1
mdh	x_{mdh}	1	−1	1.90	−1	4.00	1
pckA	x_{pckA}	1	−1	3.00	1	4.20	1
pfkA	x_{pfkA}	1	1	0.90	1	0.30	−1
pgi	x_{pgi}	1	1	0.90	1	0.50	−1
ppc	x_{ppc}	1	1	0.90	1	0.40	−1
pps	x_{pps}	1	−1	0.90	−1	10.00	1
pta	x_{pta}	1	1	0.80	1	0.40	−1
ptsG	x_{ptsG}	1	1	0.30	−1	0.20	−1
ptsH	x_{ptsH}	1	1	0.50	−1	0.40	−1
ptsI	x_{ptsI}	1	1	0.50	−1	0.40	−1
pykA	x_{pykA}	1	−1	2.20	1	1.00	−1
pykF	x_{pykF}	1	1	0.80	1	0.50	−1
sdhC	x_{sdhC}	1	−1	1.40	−1	2.10	1
sucA	x_{sucA}	1	−1	3.80	1	5.70	1
tpiA	x_{tipA}	1	1	0.90	1	0.30	−1

on acetate. TCA cycle fluxes do not follow this order. This has been analyzed in detail in the discussion.

Study of the Model Dynamics

The study of the dynamics of the gene and metabolic regulation network consists of finding the fixed points of the network. They are related to the stable states that the cell will reach and new information can be found from them.

To study the model dynamics, a computational programme was created that included the variables and the activation functions. The programme was built using Visual Basic for Applications. The algorithm used allowed the following:

- a definition of the nutrient configuration in the medium ($x_{Acetate}$, $x_{Glycerol}$, and $x_{Glucose}$)
- generation of random values (within the corresponding bounds) for the rest of the variables of the network.
- generation of the configuration vectors by updating in parallel the activation state for each element of the network using the activation functions.

This process continued until a fixed point or a cycle was reached in the dynamics of the network; the fixed points (stable states) were stored and afterwards compared with the experimental data. For each nutrient configuration in the media, (2^3) 100,000 random vectors were generated and the results obtained are shown in the next section.

MODEL RESULTS

Model Dynamics

The model produced 67 nodes: 28 were genes, 20 enzymes, and 19 regulators/biochemical compounds. The study of the network dynamics generated 8 (2^3) fixed points, one for each nutrient configuration (substrate mixture) in the medium. Some of the fixed points were very similar, so the same values for all the variables were produced in these cases, with the exception of those that represent the nutrients. Glucose had a marked effect on the internal signals in the cell. Hence, the 4 attractors where glucose was present were the same ($x_{Glucose}$=3). The effect of glycerol was milder and hence only 2 attractors were the same when it was present and no glucose existed in the medium. The activation of genes was the same when acetate was present or no carbon source was available; the

Table 6. Biochemical compounds and regulators, variable, and values reached with each carbon source.

Compound or regulator	Variable	Glucose	Glycerol	Acetate
Acetate	$x_{Acetate}$	0	0	1
AceK	x_{AceK}	-1	-1	-1
cAMP	x_{cAMP}	0	1	2
cAMP.Crp	$x_{cAMP.Crp}$	0	1	2
Crp	x_{Crp}	1	1	1
CyaA	x_{CyaA}	0	1	2
EIIAGlc	$x_{EIIAGlc}$	1	0	0
EIICBGlc	$x_{EIICBGlc}$	3	0	0
F1,6P	$x_{F1,6P}$	3	2	1
F6P	x_{F6P}	3	2	1
FruR	x_{FruR}	0	1	2
G3P	x_{G3P}	3	2	1
Glucose	$x_{Glucose}$	3	0	0
Glycerol	$x_{Glycerol}$	0	2	0
ICT	x_{ICT}	3	2	1
Mlc	x_{Mlc}	-1	1	1
Joker1	x_{Joker1}	2	1	0
Joker2	x_{Joker2}	0	1	0
RepressorSdhC	$x_{RepressorSdhC}$	2	1	0

values of other variables differed somewhat. Hence, the results could be grouped as follows:

Glucose Present in the External Media. The four fixed points when glucose was present ($x_{Glucose}$=3) were

1. Acetate Absent, Glycerol Absent, and Glucose Present ($x_{Acetate}$=0, $x_{Glycerol}$=0, $x_{Glucose}$=3)
2. Acetate Present, Glycerol Absent, and Glucose Present ($x_{Acetate}$=1, $x_{Glycerol}$=0, $x_{Glucose}$=3)
3. Acetate Absent, Glycerol Present, and Glucose Present ($x_{Acetate}$=0, $x_{Glycerol}$=2, $x_{Glucose}$=3)
4. Acetate Present, Glycerol Present, and Glucose Present ($x_{Acetate}$=1, $x_{Glycerol}$=2, $x_{Glucose}$=3)

For the rest of the elements, the values of each variable were the same in the four fixed points, as shown in Fig. 7.

This means that *E. coli* does not use nor detects the presence of other carbon sources when glucose is present in the medium. The presence of glucose avoids the utilization of other carbon sources.

Glycerol Present and Glucose Absent in the External Media. The two fixed points for this situation were

5. Acetate Absent, Glycerol Present, and Glucose Absent ($x_{Acetate}$=0, $x_{Glycerol}$=2, $x_{Glucose}$=0)
6. Acetate Present, Glycerol Present and Glucose Absent ($x_{Acetate}$=1, $x_{Glycerol}$=2, $x_{Glucose}$=0)

These two attractors showed similar expression patterns and metabolic fluxes, which means that the presence of glycerol avoids the use of acetate. Fig. 8 shows the configuration of attractors 5 and 6.

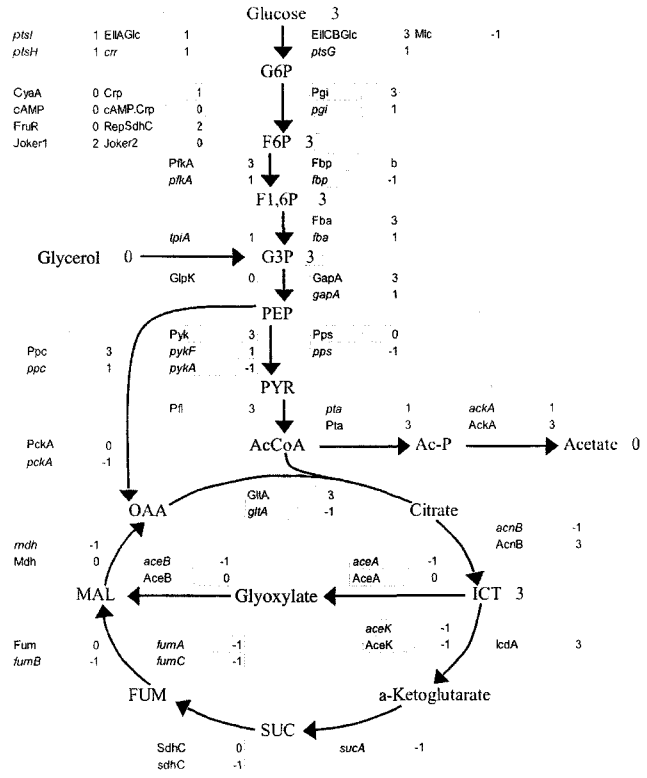

Fig. 7. Fixed points of the model when at least glucose is used as the carbon source.

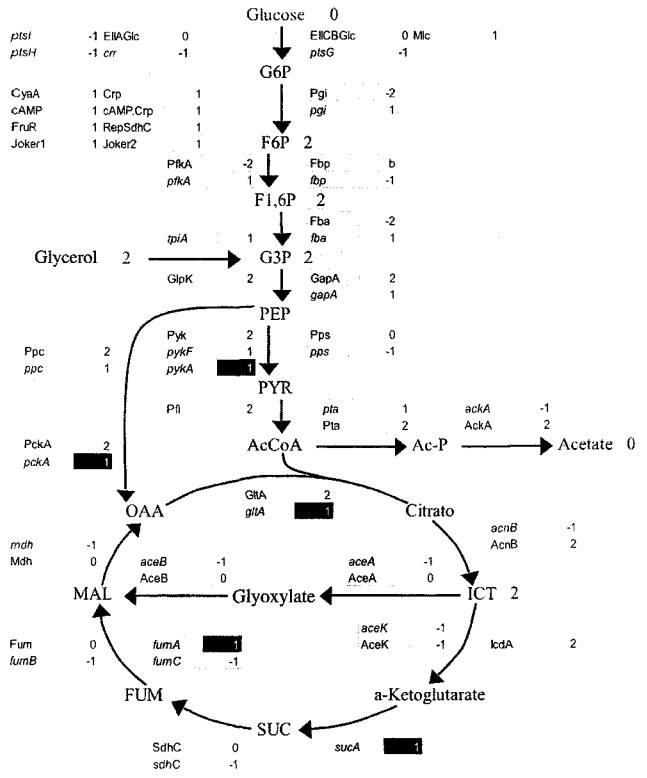

Fig. 8. Fixed points of the model when only glycerol or glycerol and acetate are used as the carbon source.

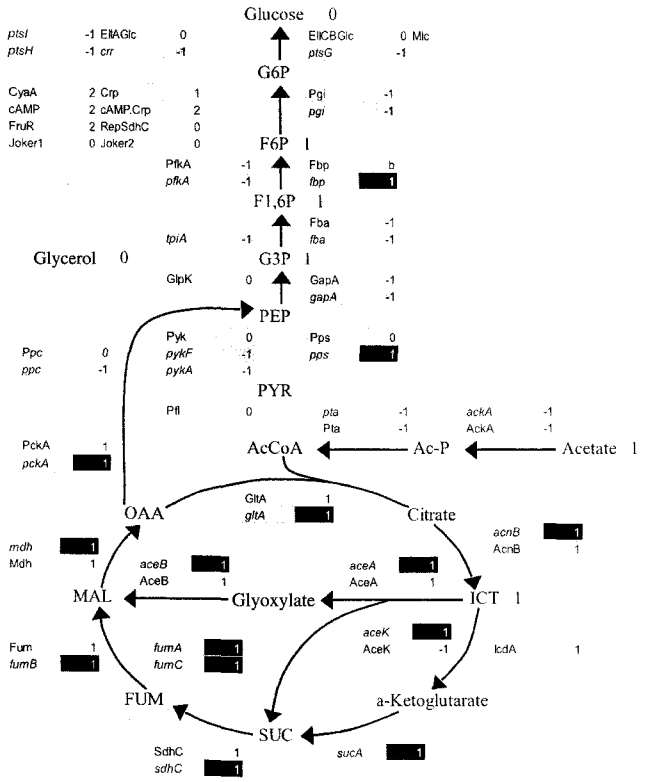

Fig. 9. Fixed points of the model when only acetate is used as the carbon source.

Only Acetate Present in the External Media. There is only one fixed point in this case:

7. Acetate Present, Glycerol Absent, and Glucose Absent ($x_{Acetate}=1$, $x_{Glycerol}=0$, $x_{Glucose}=0$)

Fig. 9 shows the configuration of fixed point 7.

No Carbon Source Present in the External Media. This attractor is the following:

8. Acetate Absent, Glycerol Absent, and Glucose Absent ($x_{Acetate}=0$, $x_{Glycerol}=0$, $x_{Glucose}=0$)

DISCUSSION

The model represents the phenotypes described experimentally and predicts the behavior of *E. coli* metabolizing mixed substrates. Thus, *E. coli* prefers glucose when growing on glucose, glycerol, and acetate, and prefers glycerol when utilizing glycerol and acetate.

The genetic regulatory network is mathematically simple and depends fundamentally on glucose and glycerol. The activation and inhibition of the genes of the network can be fully described only when considering the values of glucose and glycerol. Thus, these would correspond to the reference nodes and all genes would depend on these two nodes. Glucose and glycerol influence the network in a hierarchical fashion. The genes that glucose activates or inhibits are activated and inhibited by glycerol in a weaker fashion. This appears simple and obvious, but it is surprising that biological mechanisms adopt such a simple configuration. Hence, more complex networks could adopt the same regulation. Intermediate regulators only transmit information regarding the presence of glucose and glycerol using, at the most, three values. EIICBGlc uses 2 values, and FruR and cAMP use 3 values. The only gene that differs from this behavior is *pykA*, which is only activated in the presence of glycerol.

Fictitious or Unknown Regulators and Genes of Unknown Regulation

Fictitious regulators Joker1, Joker2, and RepressorSdhC were created to control 12 genes whose mechanism of action was unknown. These regulators indicate the presence of glycerol and glucose in the external medium since these components do not act directly on the genes. As stated previously, in order to regulate these genes, a three-state regulator was sufficient. Joker1 carried out this action and regulated 8 of these genes (*ackA*, *fba*, *fumb*, *gltA*, *pgi*, *ppc*, *pta*, y *tpiA*). Joker1 was the strongest of the fictitious regulators. The rest of the genes (*gapA*, *pykA*, *sdhC* y *sucA*) had to be regulated differently as their regulation was partially known or because they did not follow the same activation patterns. *GapA* is regulated by EIICBGlc indirectly, which is sufficient to describe its behaviour on glycerol, hence *gapA* was regulated by Joker2, which informs that extracellular glycerol is present. *SdhC* and *sucA* are also indirectly regulated by EIICBGlc and were regulated by the RepressorSdhC. This one plays the same role as Joker2 but was named differently as there are indications of the existence of RepressorSdhC and that it regulates simultaneously *sdhC* and *sucA*. *PykA* is activated by Joker2 since this gene is only activated by glycerol.

The similarity of expression patterns of the genes controlled by the fictitious regulators suggests a similar regulation mechanism to FruR or cAMP.Crp. These genes may also be controlled by EIICBGlc and an unknown regulator. Hence, it would be interesting to investigate if the rest of the genes of unknown regulation mechanism depend on EIICBGlc. It will also be necessary to investigate how EIICBGlc acts, since this is a membrane protein and it cannot reach the genes directly. The action of EIIAGlc and another compound is a candidate mechanism to be the intermediary between EIICBGlc and the gene regulated. This mechanism would be analogous to the effect of EIIAGlc on CyaA, where the action of EIIAGlc is regulated by G3P, which allows reaching 3 different values. The gene candidates to be investigated are *ackA*, *fba*, *fumb*, *gltA*, *pgi*, *ppc*, *pta*, and *tpiA*. In particular, it has already been established that *gltA* is not regulated by FruR or by cAMP.Crp.

Gene Activation is Similar When Acetate Only is Present and When There is No Carbon Source

The model also predicts the behavior of *E. coli* when no carbon source is present. The activation of genes in this case is similar to the activation when only acetate is used. Hence, gene expression with acetate corresponds to the basal level. It would be interesting to investigate this experimentally since expression of some genes appears to be very high on acetate (*aceA*, *aceB*, and *aceK*) and it may not be basal. So far, there is no experimental evidence on the possible regulation by acetate of the gene regulation network.

Flux Patterns of the Metabolic Network Depend on All Substrates

Four different flux patterns were obtained and they depend on the three substrates. The mixtures of different substrates do not generate intermediate patterns, and the carbon sources are used hierarchically: first glucose, then glycerol, and finally acetate. The fluxes are zero in the absence of all carbon sources. This situation could not be so if *E. coli* would store a carbon source. If this was the case, metabolic fluxes would exist in this situation and they should be studied in order to complete the model.

Level of Integration Between the Gene and Metabolic Network Regulation

This work stresses the integration between these two networks and on how their regulation works. This integration is necessary since the compounds of the metabolic network control gene regulation. Furthermore, gene activation of the network allows the presence of enzymes and of metabolic fluxes. The first of these interactions (metabolic network acting on gene regulation) was complete since it was not possible to find, and hence include all of the components that were regulating genes.

The second integration (gene regulation network acting over metabolic network) was successful only when gene expression varied substantially and thus modified substantially the fluxes in the metabolic network. This integration established that if one gene was inactive, the corresponding enzyme would not be expressed and, hence, the metabolic flux would be zero. The fluxes that could be regulated by the respective gene were AceA, AceB, AceK, Mdh, PckA, Ppc, PpS, Pyk, and SdhC. When the gene that regulates some of these fluxes was removed, significant changes occurred in the metabolic network. This is not always the case, as for some genes, inactivity did not imply zero expression, but less expression. This allowed for the existence of a corresponding flux even if the gene is inactive.

This model relates qualitatively results of microarrays of gene expression with Metabolic Flux Analysis (MFA). Microarray results can characterize genetic phenotypes and MFA will characterize the phenotypes of the fluxes. In this paper, the experimental results of Oh and Liao [16] for microarrays and those of Holms [10] for MFA were used. Evidently, this analysis and the application of this model are rather universal and can be extended to many different systems, such as the accumulation of recombinant proteins in yeast cells, presently under study in our goup.

Priorities of Carbon Sources and Model Stability

The active states of the model that represented the carbon sources ($x_{Acetate}=1$, $x_{Glycerol}=2$, and $x_{Glucose}=3$) were assigned to establish a metabolic hierarchy. Since, in glycolysis, glucose generates high fluxes, glycerol intermediate ones, and acetate low ones, this hierarchy was essential to establish a model with stable attractors (Liao 2002, personal communication). In each configuration of substrates, the stable state found was the only attractor found. This hierarchy generated appropriate flux patterns and metabolite concentrations that determined a satisfactory activation of the gene regulation network. When a model without this level of hierarchy in the carbon sources was used, the flow patterns, the stable states found, and the metabolite concentrations did not reflect the experimental results at all. Furthermore, the concentration of metabolites could not be differentiated between the different configurations, making it impossible to control the gene regulatory network correctly.

The Values Assigned to the TCA Cycle Fluxes Do Not Correspond with the Ratio of Experimental Fluxes

Our model is sensitive to the value of the fluxes found in glycolysis. However, the values of the TCA cycle fluxes do not affect the regulation network. It is sufficient to have the correct sense of these fluxes.

The model presented in this paper has unique features that have been discussed. On the other hand, the model can be improved and modified to answer different questions. For example, more carbon sources can be added, the accumulation of a recombinant protein, as has recently been done for MFA (Metabolic Flux Analysis) [9], or the effect of including oxygen as a limiting substrate.

The discrete mathematical model developed here is able to successfully simulate the behavior of gene regulation and metabolic networks of glycolysis and TCA cycle of *E. coli*. The model integrates the gene regulation and metabolic networks. This is an important contribution in the application of discrete mathematical models to biological problems since these models had previously been applied only in gene regulation networks.

The model correctly predicts the behavior of *E. coli vis-à-vis* substrate mixtures in batch fermentation. In a mixture of glucose, glycerol, and acetate, it prefers glucose, then glycerol, and finally acetate. The model included 67 nodes; 28 were genes, 20 enzymes, and 19 regulators/ biochemical compounds. The study of the network dynamics

generated 8 fixed points, one for each nutrient configuration (substrate mixture) in the medium. The fixed points of the discrete model reflect the phenotypes described. The gene expression and patterns of the metabolic fluxes generated are described accurately. The activation of the gene regulation network depends basically on the presence of glucose and glycerol. The model predicts the behavior when mixed carbon sources are utilized as well as when there is no carbon source present. The model reached a fixed point in each of the 8 substrate configurations, which means that these fixed points were stable. In a minimal network, glucose and glycerol are the reference nodes and all genes depend on these.

Fictitious jokers (Joker1, Joker2, and RepressorSdhC) had to be created to control 12 genes whose regulation mechanism is unknown, since glycerol and glucose do not act directly on the genes. To regulate these genes, a regulator of three states is sufficient: Joker1 carried out this action and regulated 8 of these genes (*ackA*, *fba*, *fumb*, *gltA*, *pgi*, *ppc*, *pta*, and *tpiA*). The remaining genes (*gapA*, *pykA*, *sdhC*, and *sucA*) had to be regulated in a different form since their regulation is partially known or because they followed a different action pattern. Therefore, *gapA* was regulated by Joker2, and *sdhC* and *sucA* were regulated by RepressorSdhC. *pykA* was activated by Joker2 only in the presence of glycerol.

NOMENCLATURE

Biochemical Compounds

AcCoA: Acetyl Coenzyme A
Ac-P : Acetyl Phosphate
F6P : Fructose-6-Phosphate
F1,6P : Fructose-1,6-Phosphate
FUM : Fumarate
G6P : Glucose-6-Phosphate
G3P : Glycerol-3-Phosphate
ICT : Isocitrate
MAL : Malate
OAA : OxaloAcetate
PEP : Phosphoenol Pyruvate
PYR : Pyruvate
SUC : Succinate

Enzymes

AckA : Acetate kinase
Glta : Citrate synthetase
eno : Enolase
Fbp : Fructose bisphosphate
Fba : Fructose bisphosphate aldolase
Fum : Fumarases (FumA, FumB, and FumC)
Pgi : Glucose phosphate isomerase
GapA : Glyceraldehyde-3-phosphate dehydrogenase A
GlpK : Glycerol kinase
IcdA : Isocitrate dehydrogenase
AceA : Isocitrate lyase
Mdh : Malate dehydrogenase
AceB : Malate synthetase
Pta : Phosphate acetyl transferase
Ppc : Phosphoenolpyruvate carboxylase
PckA : Phosphoenolpyruvate carboxykinase
Pps : Phosphoenolpyruvate synthetase
PfkA and PfkB : Phosphofructokinases I and II
Pgk : Phosphoglyceratekinase
gpmA : Phosphoglycerolmutase 1
Pfl : Pyruvate formate lyases I and II (PflB and PflD)
Pyk : Pyruvate kinases I and II (PykF and PykA)
SdhC : Succinate dehydrogenase
TpiA : Triosephosphate isomerase

Acknowledgments

We would like to thank the Millennium Scientific Initiative (ICM P05-001-F) for financial support and the Centre for Mathematical Modelling (CMM) for scholarship support.

REFERENCES

1. Aracena, J., M. González, A. Zúñiga, M. Méndez, and V. Cambiazo. 2006. Regulatory network for cell shape changes during *Drosophila* ventral furrow formation. *J. Theor. Biol.* **239:** 49–62.
2. Aracena, J. 2001. Discrete Mathematical Models Associated to Biological Systems. Application to genetic regulation networks, Doctoral Thesis in Engineering Sciences, mention Mathematical Modelling, University of Chile.
3. Bledig, S. A., T. M. Ramseier, and M. H. Saier Jr. 1996. Frur mediates catabolite activation of pyruvate kinase (pykF) gene expression in *Escherichia coli. J. Bacteriol.* **178:** 280–283.
4. Charpentier, B., V. Bardey, N. Robas, and C. Branlant. 1998. The EIIGlc protein is involved in glucose-mediated activation of *Escherichia coli* gapA and gapB-pgk transcription. *J. Bacteriol.* **180:** 6476–6483.
5. Cozzone, A. J. 1998. Regulation of acetate metabolism by protein phosphorylation in enteric bacteria. *Annu. Rev. Microbiol.* **52:** 127–164.
6. Covert, M. W. and B. O. Palsson. 2002. Transcriptional regulation in constraints-based models of *Escherichia coli. J. Biol. Chem.* **277:** 28058–28064.
7. Cunningham, L., M. J. Gruer, and J. R. Guest. 1997. Transcriptional regulation of the aconitase genes (acnA and acnB) of *Escherichia coli. Microbiology* **143:** 3795–3805.
8. Eppler, T., P. Postma, A. Schütz, U. Völker, and W. Boos. 2002. Glycerol-3-phosphate-induced catabolite repression in *Escherichia coli. J. Bacteriol.* **184:** 3044–3052.
9. Gonzalez, R., B. A. Andrews, J. Monitor, and J. A. Asenjo. 2003. Metabolic analysis of the synthesis of high levels of

intracellular human SOD in *S. cerevisiae* rhSOD 2060 411 SGA122. *Biotechnol. Bioeng.* **82:** 152–169.
10. Holms, H. 1996. Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol. Rev.* **19:** 85–116.
11. Hurley, J. H., H. R. Faber, D. Worthylake, N. D. Meadow, S. Roseman, D. W. Pettigrew, and S. J. Remington. 1993. Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase. *Science* **259:** 673–677.
12. Nam, T. W., S. H. Cho, D. Shin, J. H. Kim, J. Y. Jeong, J. H. Lee, J. H. Roe, A. Peterkofsky, S. O. Kang, S. Ryu, and Y. J. Seok. 2001. The *Escherichia coli* glucose transporter enzyme IICB(Glc) recruits the global repressor Mlc. *EMBO J.* **20:** 491–498.
13. Jin, J. H., U. S. Jung, J. W. Nam, Y. H. In, S. Y. Lee, D. Lee, and J. Lee. 2005. Construction of comprehensive metabolic network for glycolysis with regulation mechanism and effectors. *J. Microbiol. Biotechnol.* **15:** 161–174.
14. Negre, D., C. Bonod-Bidaud, C. Geourjon, G. Deleage, A. J. Cozzone, and J. C. Cortay. 1996. Definition of a consensus DNA-binding site for the *Escherichia coli* pleiotropic regulatory protein, FruR. *Molec. Microbiol.* **21:** 257–266.
15. Nelson, D., A. Lehninger, and M. Cox. 2000. *Lehninger Principles of Biochemistry.* Worth Publishers.
16. Oh, M. and J. C. Liao. 2000. Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnol. Prog.* **16:** 278–286.
17. Park, S. J., P. A. Cotter, R. P. Gunsalus. 1995. Regulation of malate dehydrogenase (mdh) gene expression in *Escherichia coli* in response to oxygen, carbon, and heme availability. *J. Bacteriol.* **177:** 6652–6656.
18. Park, S. J. and R. P. Gunsalus. 1995. Oxygen, iron, carbon, and superoxide control of the fumarase fumA and fumC genes of *Escherichia coli*: Role of the arcA, fnr, and soxR gene products. *J. Bacteriol.* **177:** 6255–6262.
19. Saier, M. H. Jr. and T. M. Ramseier. 1996. The catabolite repressor/activator (Cra) protein of enteric bacteria. *J. Bacteriol.* **178:** 3411–3417.
20. Takeda, S., A. Matsushika, and T. Mizuno. 1999. Repression of the gene encoding succinate dehydrogenase in response to glucose is mediated by the EIICB(Glc) protein in *Escherichia coli*. *J. Biochem.* **126:** 354–360.
21. Tanaka, Y., K. Kimata, and H. Aiba. 2000. A novel regulatory role of glucose transporter of *Escherichia coli*: Membrane sequestration of a global repressor Mlc. *EMBO J.* **19:** 5344–5352.