Robustness Analysis of the *Escherichia coli* Metabolic Network

Jeremy S. Edwards[†] and Bernhard O. Palsson*

Department of Bioengineering, University of California-San Diego, La Jolla, California 92093-0412

Genomic, biochemical, and strain-specific data can be assembled to define an *in silico* representation of the metabolic network for a select group of single cellular organisms. Flux-balance analysis and phenotypic phase planes derived therefrom have been developed and applied to analyze the metabolic capabilities and characteristics of *Escherichia coli* K-12. These analyses have shown the existence of seven essential reactions in the central metabolic pathways (glycolysis, pentose phosphate pathway, tricarboxylic acid cycle) for the growth in glucose minimal media. The corresponding seven gene products can be grouped into three categories: (1) pentose phosphate pathway genes, (2) three-carbon glycolytic genes, and (3) tricarboxylic acid cycle genes. Here we develop a procedure that calculates the sensitivity of optimal cellular growth to altered flux levels of these essential gene products. The results indicate that the *E. coli* metabolic network is robust with respect to the flux levels of these enzymes. The metabolic flux in the transketolase and the tricarboxylic acid cycle reactions can be reduced to 15% and 19%, respectively, of the optimal value without significantly influencing the optimal growth flux. The metabolic network also exhibited robustness with respect to the ribose-5-phosphate isomerase, and the ribose-5-phosephate isomerase flux was reduced to 28% of the optimal value without significantly effecting the optimal growth flux. The metabolic network exhibited limited robustness to the three-carbon glycolytic fluxes both increased and decreased. The development presented another dimension to the use of FBA to study the capabilities of metabolic networks.

Introduction

Genome sequencing and bioinformatics are beginning to reveal the complete set of molecular components involved in cellular activities. Furthermore, it is also clear that the integrated function of biological systems involves complex interactions among the components that have been identified through bioinformatics and genomics. Importantly, the properties of complex systems cannot be predicted simply on the basis of the complete description of their components, and the emergent properties of biological systems need to be studied (*1, 2*). To understand the complexity inherent in cellular networks, approaches that focus on the systemic properties of the network are required. The focus of such research represents a departure from the classical reductionist approach to the integrated approach (*3*) to understanding the interrelatedness of gene function and the role of each gene in the context of multigenetic cellular functions or *genetic circuits* (*4, 5*).

The engineering approach to analysis and design is to have a mathematical or computer model, e.g., a dynamic simulator, of a cellular process that is based on fundamental physicochemical laws and principles. There has been a long history of mathematical modeling of metabolic systems, which dates back to the to the mid 1960s. With the availability of analogue computers and the knowledge of metabolic regulation, dynamic simulations of simple metabolic and genetic control loops appeared (*6*). The dynamic stability of such control loops became a focus of attention (*7, 8*), given the experimental observations of oscillatory dynamics in yeast glycolysis (*9*).

The systemic nature of metabolic function was apparent, and so was its complexity. However, the availability of enzyme kinetic information was fragmented, and attention turned to developing methods that could shed light on the relative importance of various metabolic events. Methods for sensitivity analysis of metabolic regulation began in the 1960s (*10*) and continued into the 1970s (*11, 12*). The results of these undertakings were biochemical systems theory (BST) and metabolic control analysis (MCA), and some useful results have been obtained using these approaches (*13*).

Establishing complete kinetic models of cellular metabolism became a scientific goal, whose intended use was to elucidate the systemic behavior of metabolic networks. Because of its simplicity, the human red blood cell represented the best opportunity to achieve this goal. Early metabolic models of human red blood cell metabolism appeared in the 1970s (*11*) and continued throughout the 1980s and 1990s (*14−16*). Insights into the functioning of this cell have resulted from these analyses (*11, 17, 18*). Although interesting in their own right, studies of red cell metabolism are not directly useful for organisms of industrial importance.

While the ultimate goal is the development of dynamic models for the complete simulation of metabolic systems, the success of such approaches has been severely hampered by the current lack of kinetic information on the dynamics and regulation of metabolic reactions. However,

* Ph: (619) 534-5668. Fax: (619) 822-0240. Email: palsson@ucsd.edu.

[†] Current address: Department of Genetics, Harvard Medical School, Boston, MA 02115.

10.1021/bp0000712 CCC: $19.00 © 2000 American Chemical Society and American Institute of Chemical Engineers
Published on Web 08/08/2000

in the absence of kinetic information it is still possible to accurately assess the theoretical capabilities and operative modes of metabolic systems using metabolic flux balance analysis (FBA) (*5, 19−23*). FBA is based on the fundamental physicochemical constraints on metabolic networks. FBA only requires information regarding the stoichiometry of metabolic pathways and the metabolic demands; furthermore, FBA can incorporate additional information when it is available. FBA is particularly applicable for post-genomic analysis, because the stoichiometric parameters can be defined from the annotated genome sequence (*21*).

In a previous article, we have examined the capability of *in silico* mutant *E. coli* metabolic networks to support growth and compared the results to the wildtype. By using computer simulations, it was determined that seven metabolic reactions were essential for the aerobic growth of *E. coli* in glucose minimal media (*24*). The remaining reactions were determined to be nonessential, since the metabolic network maintained the capability to bypass simulated metabolic defects, often with little or no effect on the *in silico* maximal biomass yield. In this article we will further examine the essential metabolic reactions by examining the metabolic consequences of reduced metabolic flux carrying capacity in the essential reactions. The results indicate the redundancy and robustness in the function of the respective metabolic reactions in the metabolic network by examining the sensitivity of the objective function to the quantitative flux levels. The sensitivity analysis can provide information regarding the experimental measurements that are likely to provide the most information toward quantitatively describing the metabolic network and can be used for *in silico* experimental design and assessing the value of the *in silico* predictions.

Describing Metabolic Systems

A metabolic network is a collection of enzymatic reactions that serve to biochemically process metabolites within the cell and transport processes that convert extracellular metabolites to intracellular metabolites and vice versa. To quantitatively describe metabolic networks, dynamic mass balances are written for each metabolite in the network, generating a system of ordinary differential equations that describe the transient behavior of metabolite concentrations:

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = \sum_j S_{ij} v_j \tag{1}$$

where v_j corresponds to the jth metabolic flux, X_i represents the ith metabolite, and the stoichiometric coefficient $S_{i,j}$ stands for the number of moles of metabolite i formed (or consumed) in reaction j. Equation 1 is particularly difficult to solve since the metabolic fluxes are often nonlinear functions of the metabolite concentrations, as well as a set of kinetic parameters that are difficult to measure or estimate. The complexity associated with estimating the functional relation between the metabolic fluxes and the metabolite concentrations and the associated kinetic parameters has hampered the quantitative analysis of metabolic networks.

Constraining Metabolic Functions

Given the complexities associated with quantitative analysis of metabolic systems based on kinetic characterization of the components, we have utilized a conceptually different approach to the analysis of metabolic networks. First, we defined fundamental physicochemical constraints to which the metabolic network is constrained. Then, the metabolic capabilities were assessed subject to the imposed constraints. The capabilities are analyzed under the steady state assumption. It should be noted that steady state analysis is applicable to some aspects of metabolism; however, the approach will not be appropriate for studying all cellular processes, such as the cell cycle or signal transduction. Herein, we are interested in metabolic processes and their relation to cellular growth; thus the characteristic time of the processes is about an hour. Metabolic transients within the cell typically occur with time constants on the order of seconds to minutes (*25*); thus under our "window of observation" the metabolic network is essentially in a steady state and the steady-state analysis will be appropriate. The steady-state mass, energy, and redox balance constraints are imposed by simplifying eq 1:

$$\mathbf{S} \cdot \mathbf{v} = 0 \tag{2}$$

where **S** is the stoichiometric matrix and **v** is the flux vector. While the system is closed to the passage of certain metabolites, others are allowed to enter or exit the system via exchange fluxes (or pseudoreactions (*26*)). These fluxes do not represent biochemical conversions or transport processes such as those of internal fluxes but can be thought of as representing the inputs and outputs to the system. For example, the demand on a metabolite for further processing or incorporation into cellular biomass creates an exchange flux on the internal cellular metabolite. Thus, a distinction is made between internal and external metabolites in the system, therefore closing the material balance to all metabolites as indicated by eq 2.

To complete the *in silico* representation of the metabolic network we included the constraints on the individual metabolic reaction fluxes due to reaction thermodynamics and the input/output characteristics of the network. All reversible metabolic reactions were assumed to have the capability to carry any metabolic flux (i.e., $-\infty \leq v_i \leq \infty$; where v_i is the flux in reversible reactions), whereas irreversible metabolic reactions fluxes were restricted to be positive (i.e., $0 \geq v_i \geq \infty$; where v_i is the flux in irreversible reactions). Although constraints on the internal fluxes were defined as infinite, the magnitude of each flux in the optimal solution was examined and compared to measured fluxes (*27, 28*). The reversibility of each reaction in the metabolic network was determined case by case on the basis of the literature and compared to the EcoCyc database (*29*). The metabolic enzymes identified in the complete *E. coli* K12 genome sequence and the online databases (*29−31*) were used to reconstruct the metabolic network (see supplementary information at http://gcrg.ucsd.edu/supplementary_data/BP2000/main.htm). It should be noted that there are instances where the same enzyme can catalyze multiple reactions (e.g., different substrates or cofactors), and this situation was considered by including all reactions catalyzed by an enzyme as a separate column in the stoichiometric matrix. The details of this metabolic reconstruction have been described elsewhere (*24*). Additionally, constraints were placed on the exchange fluxes to indicate the environmental conditions. For example, metabolites not available to the cell are constrained to not enter the cell: $-\infty \leq b_i \leq 0$, where b_i (influx defined as positive) is the exchange flux for a metabolite not available in the simulated environment. It should be noted that all metabolites that have the capability to leave the cell always had unconstrained metabolic fluxes

in the net outward direction, whereas the influx constraints were defined by the simulated environmental conditions. For the analysis herein, the exchange flux for inorganic phosphate, ammonia, carbon dioxide, oxygen, sulfate, potassium, and sodium were unconstrained, whereas the uptake of the carbon source was constrained as specified.

Demands on the Metabolic Network

Under changing substrate/supply conditions metabolic networks are continuously faced with a balanced set of biosynthetic demands (i.e., production of amino acids, nucleotides, phospholipids, as well as energy and redox potential). Effectively this means that the network must generate a balanced set of metabolites that are used to produce biomass. The biosynthetic demands for growth were determined from the biomass composition of *E. coli* (*32, 33*), and a metabolic flux, defined as the growth flux (v_{growth}), utilizes the biosynthetic precursors in the appropriate ratios so as to generate biomass:

$$\sum_{i}^{m} d_i \cdot X_i \xrightarrow{v_{growth}} \text{Biomass}$$

where d_i (mmol · g-dry weight (DW)$^{-1}$) is the *E. coli* biomass composition of metabolite *i*. One gram of biomass is produced per unit flux in the growth flux, v_{growth}, and if the fluxes are represented with a basis of 1 g-DW · h (*22*), the growth flux is equivalent to the growth rate. The biomass composition is not constant but depends on the growth rate and the growth conditions (*33*). However, we have assumed that the biomass composition is constant since it has been shown that the optimal solution is not sensitive to the biomass composition (*34*), and this observation is also true for our system.

In addition to the biosynthetic demands on the metabolic network, we have also imposed maintenance requirements on the metabolic system. The maintenance requirements included were for growth-associated and non-growth-associated maintenance. We imposed a growth-associated maintenance of 23 mmol ATP · g-DW^{-1} and a non-growth-associated maintenance of 5.87 mmol ATP · g-DW^{-1} · h^{-1} (*35*).

Exploring the Metabolic Capabilities

The constraints on the metabolic network define the boundaries within which the metabolic system must operate. The mass, energy, and redox balance constraints are imposed by the linear homogeneous set of equations (eq 2). The nullspace of the stoichiometric matrix, **S**, contains all flux vectors that satisfy the mass, energy, and redox balance constraints (*36*). However, there are additional physicochemical constraints on the metabolic network, such as the thermodynamic constraints and the capacity constraints on the exchange fluxes, which are enforced by linear inequalities. The simultaneous enforcement of all the metabolic constraints defines a region, the *feasible set*, that contains all feasible metabolic flux vectors. The feasible set is not a vector space as is the nullspace, as a result of the linear inequality constraints. Importantly, the feasible set defines the metabolic capabilities of the system. The performance capabilities of any metabolic network reside in the feasible set. In fact, the answer to any question related to the general structure and fitness of the network lies with this region. While the feasible set offers a convenient way of defining metabolic capabilities, the question arises, how do we best explore the specific functions of a metabolic network?

One approach that has been used to explore the relationship between the metabolic genotype and phenotype for a number of organisms is linear optimization (*19, 21, 22, 37*). Linear optimization was used to determine the optimal flux distributions within a network so as to maximize/minimize a particular objective function. A linear programming problem is defined as follows, where a linear objective function is maximized or minimized subject to a series of linear equality and inequality constraints:

$$\text{Maximize/Minimize } Z = c_j v_j$$
$$\text{subject to } S_{ij} v_j = 0,\ \alpha_j \le v_j \le \beta_j \qquad (5)$$

The linear programming formalism is analogous to the system of linear equalities/inequalities that form the constraints on the metabolic network. The objective function, *Z*, is defined by assigning the appropriate values to the **c** vector; herein, the **c** vector was taken as the unit vector in the direction of the growth flux. We used the reduced costs from the linear programming solution to identify alternate optimal solutions. In metabolic engineering applications, the objective function can correspond to a number of diverse objectives, such as maximizing energy or metabolite production (*20*). However, regardless of the objective function the optimal solution will lie within the feasible set that is defined by the physicochemical constraints placed on the system.

The utilization of linear programming to examine metabolic networks defines the optimal flux vector that maximizes (or minimizes) an objective function and satisfies the entire set of constraints. The utilization of design related objectives (such as maximizing the production of an amino acid) can be used to guide genetic engineering of a strain for metabolite overproduction. Herein, we have employed a physiologically realistic objective, the maximization of the growth flux. We have assumed that the cell has evolved the regulatory mechanisms to operate optimally within the feasible set. The feasible set defines the capabilities of the metabolic network, and all metabolic flux vectors within the feasible set satisfy the imposed physicochemical constraints. Therefore, theoretically all flux vectors within the feasible set can be reached by adjusting the enzyme kinetic parameters and gene regulation. The enzyme kinetics and gene regulation constraints on the metabolic system will be referred to as system specific constraints. We assume that the cell has found the optimal set of system specific constraints through the course of evolution, and we attempt to find the same solution using linear programming. The assumption has been experimentally examined under a limited number of conditions, and under defined conditions with a single carbon source, the experimental data is consistent with the optimal utilization of the metabolic network (*27*).

Phenotype Phase Plane Analysis

Flux balance analysis can be used to examine the metabolic network in detail. Optimal solutions to the linear programming problem will then lie on a vertex of the feasible set, which is a polyhedron (*38*). All the metabolic flux vectors (or metabolic phenotypes) attainable from a defined metabolic genotype are mathematically confined to the feasible set. Linear programming was used to search through the feasible set for a solution that maximizes the growth flux. Experimental data for the growth of *E. coli* under nutritionally rich growth conditions (i.e., cell is not starved for phosphate, nitrogen,

etc.) is consistent with the optimal utilization of the metabolic network (*27*); thus, defining the growth flux as the objective function produces physiologically meaningful results. However, the optimal flux distribution is only meaningful when interpreted in terms of the specific environmental conditions. Therefore, phenotype phase planes (*39*) have been developed to define the range of optimal flux vectors and how the optimal flux vector is dependent on the environmental conditions.

The methodology for defining PhPPs has been described (*39*). We will now briefly describe the construction of PhPPs. Two metabolic fluxes can form two axes on an (x,y)-plane (these metabolic fluxes were two unit vectors in R^n). The optimal metabolic flux distribution is calculated for all points in this plane. In other words, the maximum value of the objective function is found as the position of the hyperplanes that bound the feasible set in the respective directions is moved. It has been determined that there are a finite number of fundamentally different optimal metabolic flux distributions (or basis solutions in linear programming terminology) present in such a plane. The demarcations on the phase plane were defined by a shadow price (LP dual variable) analysis (*40*). This procedure leads to the definition of distinct regions, or "phases", in the plane, in which the optimal use of the metabolic network is fundamentally different, corresponding to different optimal phenotypes.

Robustness Analysis

Robustness, defined here with respect to metabolic networks, is a measure of the change in the maximal flux of the objective function (the growth flux was defined as the objective) when the optimal flux through any particular metabolic reaction is changed. The robustness characteristics of the metabolic network were determined by calculating the optimal flux vector so as to maximize the growth flux (with only external flux constraints), this flux was called the *in silico* wildtype flux. Then the flux through the reaction in question was reduced from 100% to 0% of the *in silico* wildtype flux and the objective function was calculated. Additionally, the *in silico* wildtype flux was increased from the wildtype value and the upper bound on increasing the flux level was the maximal allowable flux in the reaction or the flux level for which the objective function was reduced to zero. The calculations were for a simulated aerobic batch culture in glucose minimal media.

The FBA framework was used to address the systemic effect on the metabolic network of increased and decreased (with respect to the *in silico* wildtype) metabolic flux. Herein, we quantified the robustness of the metabolic network to flux changes in the essential enzymatic reactions. The essential enzymes (for growth on glucose minimal media) were previously identified through an *in silico* analysis (*24*). Seven enzymatic reactions in central metabolism (Figure 1) were found to be essential: the transketolase (TKT), ribose-5-phosphate isomerase (RPI), two enzymes (GAP, PGK) in the 3-carbon stage of glycolysis (3CG), and the first three enzymes (GLT, ACN, ICD) of the TCA cycle. Below, the robustness characteristics of the metabolic network with respect to alterations of the flux levels of these essential metabolic reactions will be investigated. We will utilize phenotype phase planes (PhPPs) to define points where the optimal utilization of the metabolic network changes due to capacity constraints on the essential enzymatic reactions.

Transketolase. The transketolase (TKT) catalyzes an essential enzymatic reaction in the pentose phosphate pathway (PPP) (*41*). However, *tkt* mutant strains have been shown to grow on glucose minimal media with low TKT residual activity (3% of wildtype) (*42, 43*). The ability of the metabolic network to support growth with a large reduction in TKT flux was investigated *in silico* by continuously restricting the metabolic flux in the TKT reactions. As the maximum allowable flux through the TKT reactions was reduced from the *in silico* wildtype, it was determined that the ability of the metabolic network to support growth was virtually unchanged for enzymatic fluxes as low as 15% of the *in silico* wildtype (Figure 2). The response to decreased TKT metabolic flux was found to have two qualitatively different regions. The regions were identified in the PhPP (Figure 3).

The PhPP describing the changes in the metabolic pathway utilization as a function of the TKT flux and the glucose uptake rate was calculated (Figure 3). The optimal relation between the glucose uptake rate and the TKT flux was determined from the PhPP (Figure 3). It was determined that there were two qualitatively different regions of metabolic pathway utilization for TKT fluxes lower than optimal, and these regions were defined as A and B (as shown in Figure 3). Furthermore, there were determined to be six qualitatively different regions for TKT fluxes greater than optimal, and these regions were defined as 1−6 (as shown in Figure 3). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all TKT fluxes from zero to the maximum allowable flux (the glucose uptake exchange flux was constrained to 10 mmol g-DW^{-1} h^{-1}) that still permits cellular growth, and the results are shown in Figure 2.

In region A (above 15% of the *in silico* wildtype enzyme flux), the optimal value of the growth flux was hardly changed, and at the demarcation between regions A and B, the growth flux was decreased to 99.2% of the *in silico* wildtype. However, to cope with the decreased TKT metabolic flux carrying capacity, shifts in the metabolic pathway utilization occurred (Figure 4). The redox potential (NADPH) requirement for biosynthetic demands was achieved by a flux redistribution that resulted in the utilization of the transhydrogenase that converted NADH (produced from an increased TCA cycle flux, Figure 4C) into NADPH. The flux diverted from the PPP (Figure 3B) resulted in increased glycolytic fluxes, in particular the pyruvate kinase and the phosphoglucoisomerase fluxes (Figure 3A). In this region, the optimal growth flux was not sensitive to changes in the TKT flux. However, the optimal flux in several metabolic processes were sensitive to the TKT flux in this region (transhydrogenase, PYK, PGI, TCA cycle flux)

In the second region (region B) of reduced TKT flux (enzyme flux less than 15% of the *in silico* wildtype), the metabolic network was limited in the ability to produce the essential biosynthetic precursor, erythrose 4-phosphate. In this region, the optimal growth flux was sensitive to the flux level in the TKT reaction. The metabolic fluxes in this region are not shown in Figure 4 because alternate optimal solutions exist. Cellular growth is solely limited by the availability of a single biosynthetic precursor, the excess glucose can be converted to any of the metabolic byproducts with the same value of the objective function. Furthermore, the excess high-energy phosphate bonds can be eliminated in any futile cycle; thus alternate optimal solutions exist.

The effect on the metabolic network due to TKT fluxes increased beyond the optimal flux for growth was also examined. An increase in metabolic flux may result from the overexpression of the respective gene, and the robustness analysis can be used to identify the constraints on flux changes due to the integrated metabolic network.

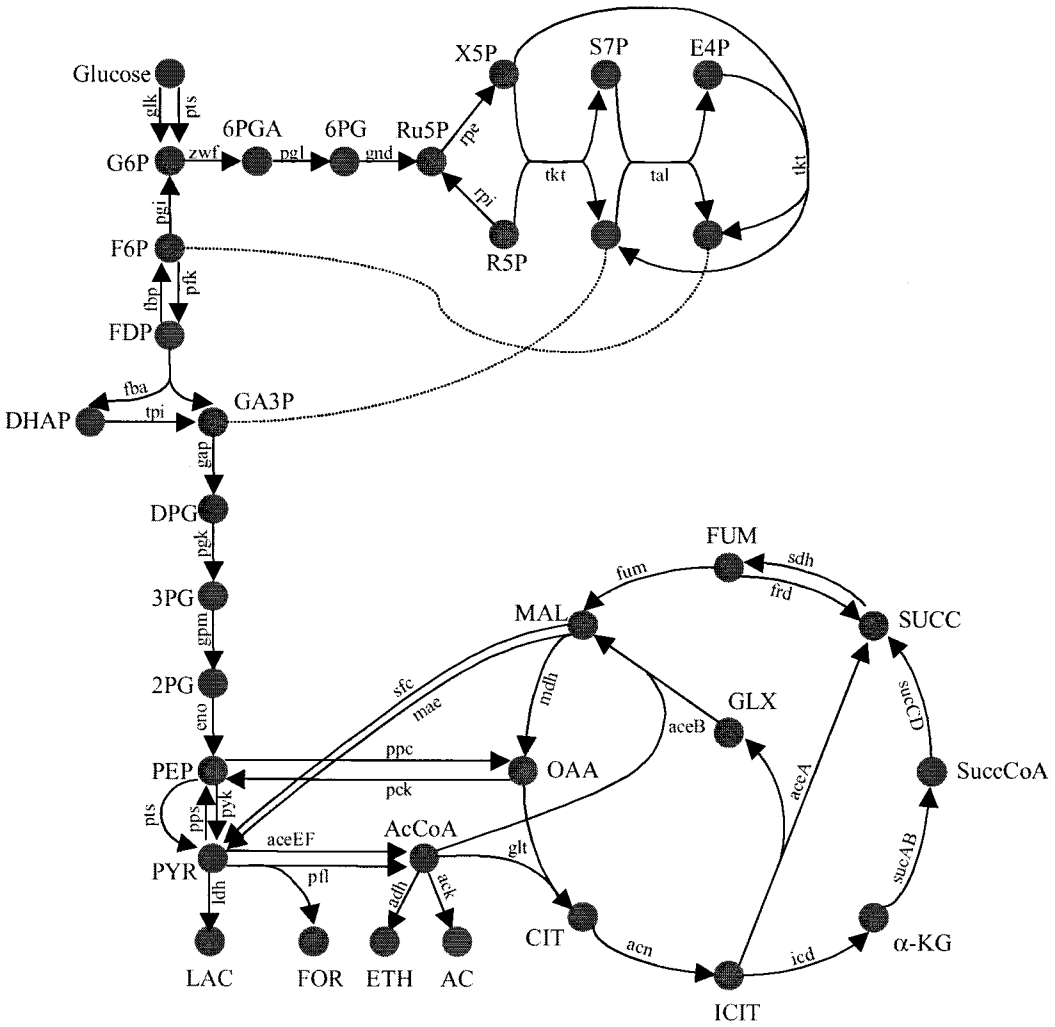

Figure 1. The central metabolic pathway reactions. Reactions: aceA, isocitrate lyase; aceB, malate synthase; aceEF, pyruvate dehydrogenase; ack, acetate kinase; acn, aconitase; adh, acetaldehyde dehydrogenase; eno, enolase; fba, fructose-1,6-bisphosphatate aldolase; fbp, fructose-1,6-bisphosphatase; frd, fumurate reductase; fum, fumarase; gap, glyceraldehyde-3-phosphate dehydrogenase; glk, glucokinase; glt, citrate synthase; gnd, 6-phosphogluconate dehydrogenase; gpm, phosphoglycerate mutase; icd, isocitrate dehydrogenase; ldh, lactate dehydrogenase; mae, malic enzyme; mdh, malate dehydrogenase; pck, phosphoenolpyruvate carboxykinase; pfk, phosphofructokinase; pfl, pyruvate formate lyase; pgi, phosphoglucose isomerase; pgk, phosphoglycerate kinase; pgl, 6-phosphogluconolactonase; ppc, phosphoenolpyruvate carboxylase; pps, phosphoenolpyruvate synthase; pts, phosphotransferase system; pyk, pyruvate kinase; rpe, ribulose phosphate 3-epimerase; rpi, ribose-5-phosphate isomerase; sdh, succinate dehydrogenase; sfc, malic enzyme; sucAB, 2-ketoglutarate dehyrogenase; sucCD, succinyl-CoA synthetase; tal, transaldolase; tkt, transketolase; tpi, triosphosphate isomerase; zwf, glucose 6-phosphate-1-dehydrogenase. Metabolites: 2PG, 2-phosphoglycerate; 3PG, 3-phosphoglycerate; 6PG, D-6-phosphate-gluconate; 6PGA, D-6-phosphate-glucono-δ-lactone; AC, acetate; AcCoA, Acetyl-CoA; α-KG, α -ketoglutarate; CIT, citrate; DHAP, dihydroxyacetone phosphate; DPG, 1,3-bis-phosphoglycerate; E4P, erythrose 4-phosphate; ETH, ethanol; F6P, fructose 6-phosphate; FDP, fructose 1,6-diphosphate; FOR, formate; FUM, fumarate; G6P, glucose 6-phosphate; GA3P, glyceraldehyde 3-phosphate; ICIT, isocitrate; LAC, lactate; MAL, malate; PEP, phosphoenolpyruvate; PYR, pyruvate; R5P, ribose 5-phosphate; Ru5P, ribulose 5-phosphate; S7P, sedo-heptulose; SUCC, succinate; SuccCoA, succinyl CoA; X5P, dihydroxyacetone phosphate.

The optimal growth flux as a function of the flux in the TKT reaction was calculated (Figure 2, insert). The metabolic flux was continuously increased *in silico* from the *in silico* wildtype value to the maximum flux that permits growth. Six qualitatively different patterns of metabolic pathway utilization (numbered 1−6 in Figure 3) were observed when the TKT flux was increased beyond the *in silico* wildtype.

The region 1 of Figure 3 the optimal metabolic flux vector was characterized by an increased PPP flux, an active transhydrogenase reaction, a decreased PYK flux, and a decreased TCA cycle flux. In region 2 the TCA cycle flux was further decreased while the PPP flux was increased, and optimally, the glyoxylate bypass was utilized to replenish the TCA cycle biosynthetic precursors thus reducing the PPC flux (Figure 4). At the demarcation between regions 2 and 3, the TCA cycle was shut off and functioned to produce the biosynthetic precursors, rather than redox potential. Optimally, in region 3, the glucokinase reaction was operative in glucose utilization, and this allowed for a more efficient flow of the metabolites into the PPP due to the increased TKT flux. Regions 4 and 5 were similar with respect to the metabolic pathway utilization, the glyoxylate bypass was no longer utilized in the optimal solution and the PFL reaction optimally carried a small flux. Finally, in region 6, redox potential was overproduced. This region was characterized by alternate optimal solutions to eliminate the excess high-energy phosphate bonds. However, there were no metabolic byproducts produced (other than CO_2); this was because many metabolites were still desirable to the cell (as identified through a shadow price analysis (*40*)). In this region, the optimal oxygen uptake rate was very high (~60 mmol g-DW^{-1} h^{-1}) and it is likely that the maximal TKT flux is much lower due to other constraints on the metabolic network that were not

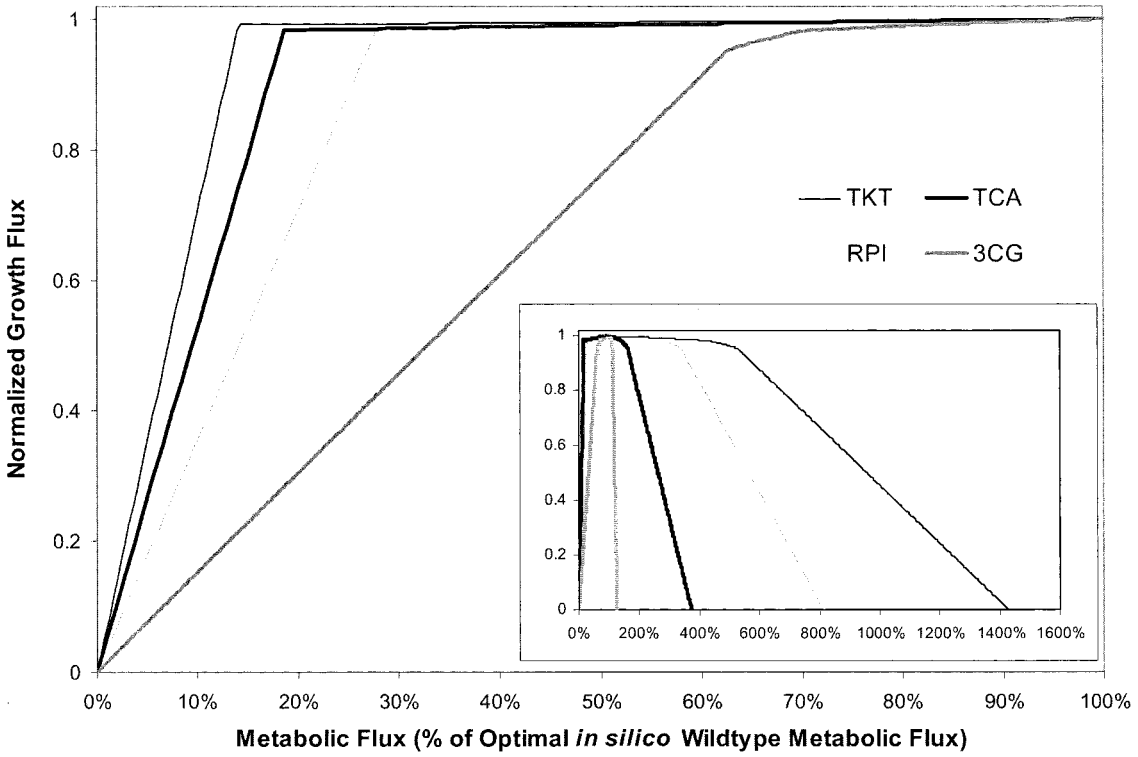

Figure 2. Robustness diagrams. The effect of altered metabolic flux in the essential metabolic reactions on the normalized growth flux is illustrated. The *in silico* wildtype flux is defined as 100%. See the text for a complete discussion.

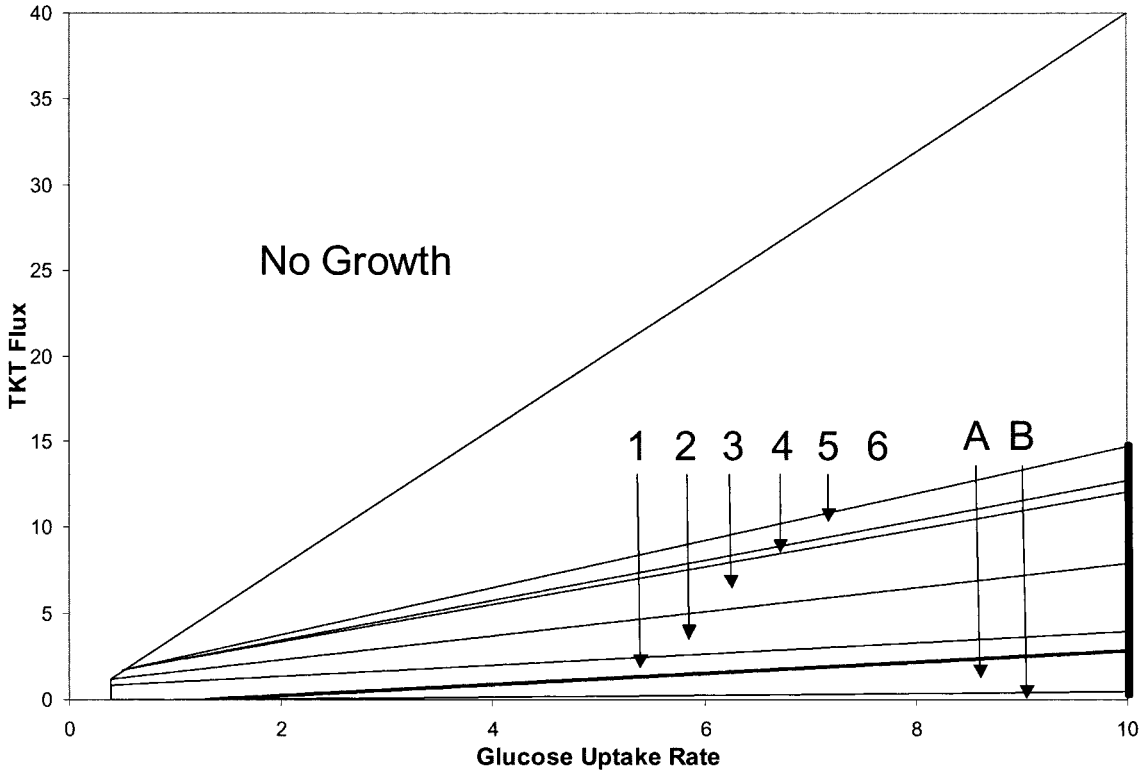

Figure 3. The glucose uptake rate (mmol · g-DW^{-1} · h^{-1})—transketolase flux (substrates converted · g-DW^{-1} · h^{-1}) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to TKT fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify TKT flux reductions below the optimal relation. The metabolic fluxes along the thick vertical line (glucose uptake rate = 10 mmol · g-DW^{-1} · h^{-1}) are shown in Figure 4.

included in the analysis (such as oxygen mass transfer limitations).

Ribose-5-Phosphate Isomerase. The ribose-5-phosphate isomerase reaction (RPI) also catalyzes an essential reaction in the PPP (*41*) for growth in glucose minimal media. Similarly to *tkt* mutants, *rpi* mutants have been shown to grow with enzymatic activity much less than that of the wildtype. For example, Skinner and Cooper have isolated a strain with RPI activity below 10% of the wildtype, and this strain was able to grow (*44*). As the

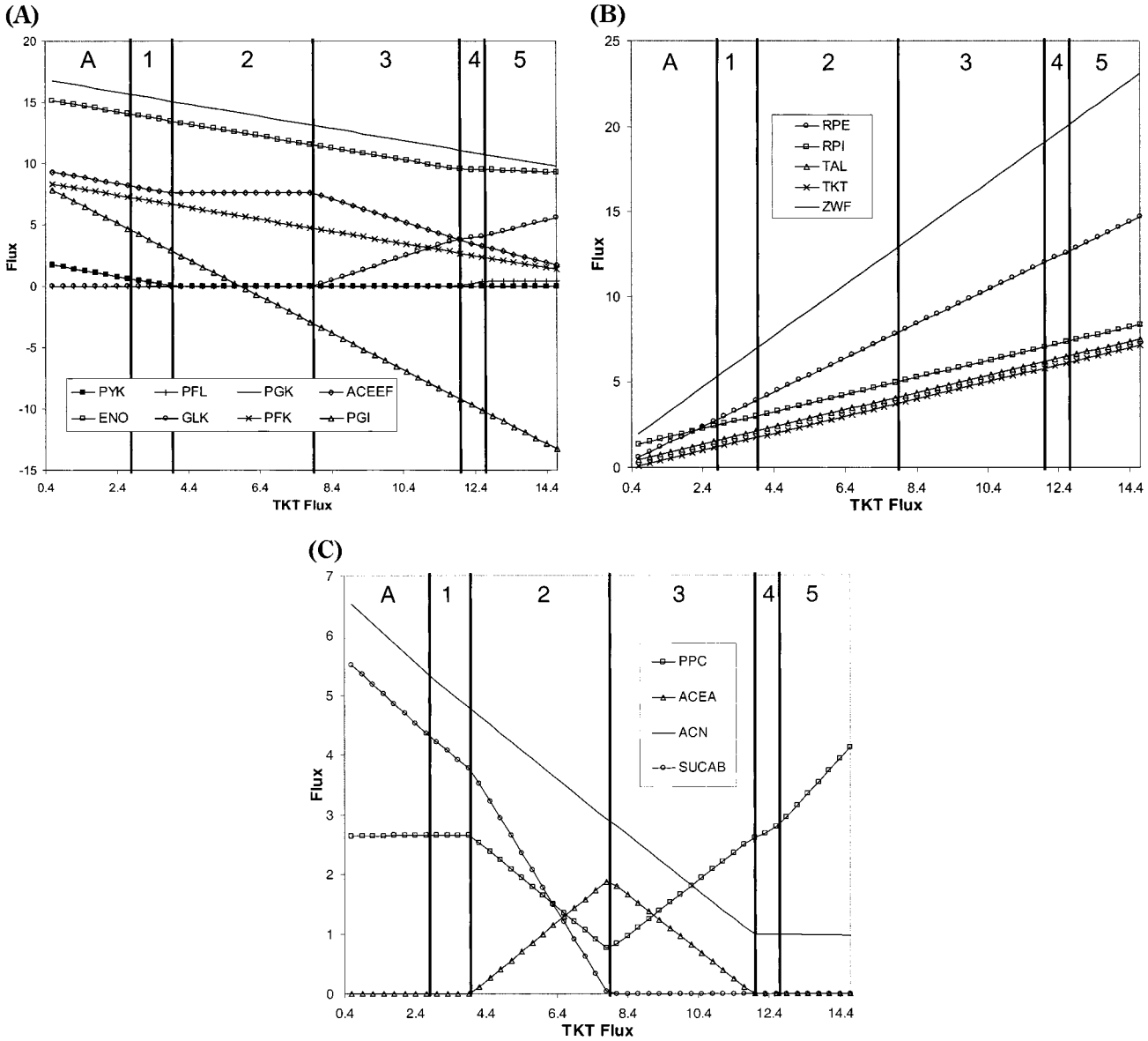

Figure 4. Optimal intracellular fluxes in the central metabolic pathways (substrates converted · g-DW^{-1} · h^{-1}) as a function of the TKT metabolic flux constraint (substrates converted · g-DW^{-1} · h^{-1}). The glucose uptake rate was constrained to 10 mmol · g-DW^{-1} · h^{-1}. (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

maximum allowable RPI flux was reduced from the *in silico* wildtype, it was determined that the ability of the metabolic network to support growth was virtually unchanged for enzymatic fluxes as low as 28% of the *in silico* wildtype (Figure 2). The RPI-glucose uptake rate PhPP was calculated to characterize the effect of altered RPI metabolic fluxes (not shown, similar to Figure 3). The qualitative effect of reduced flux in the RPI reaction from the *in silico* wildtype was investigated with FBA, and the holistic metabolic response to decreased and increased RPI fluxes was similar to the TKT results because the effect on the PPP was similar.

3-Carbon Glycolysis. The 3-carbon glycolytic reactions that the *in silico* analysis predicted to be essential have been shown experimentally to be required for the growth of *E. coli* on a glucose minimal media (GAP, PGK) (*41*). The glycolytic essential reactions were subjected to a robustness analysis to investigate the optimal systemic effect of flux alteration. The ability of the metabolic network to support growth with a reduction in 3CG flux was investigated *in silico* by continuously restricting the 3CG flux. As the allowable flux through the 3CG reactions was reduced from the *in silico* wildtype, it was determined that the sensitivity of the growth flux was increased compared to the other essential reactions. When the 3CG flux was reduced below about 70% of the *in silico* wildtype, the growth flux was sensitive to the 3CG flux (Figure 2). Furthermore, the 3CG fluxes could only be increased to 110% of the *in silico* wildtype before severe limitations in the growth flux were encountered (Figure 2). We have investigated the metabolic response to 3CG flux level alterations by a phenotype phase plane analysis (Figure 5).

The PhPP describing the changes in the metabolic pathway utilization as a function of the 3CG flux and the glucose uptake rate was calculated (Figure 5). The optimal relation between the glucose uptake rate and the 3CG flux was determined from the PhPP (Figure 5). It was determined that there were six qualitatively different regions of metabolic pathway utilization for 3CG fluxes lower than optimal, and these regions were defined as A−F (as shown in Figure 5). Furthermore, there were determined to be two qualitatively different regions for 3CG fluxes greater than optimal, and these regions were

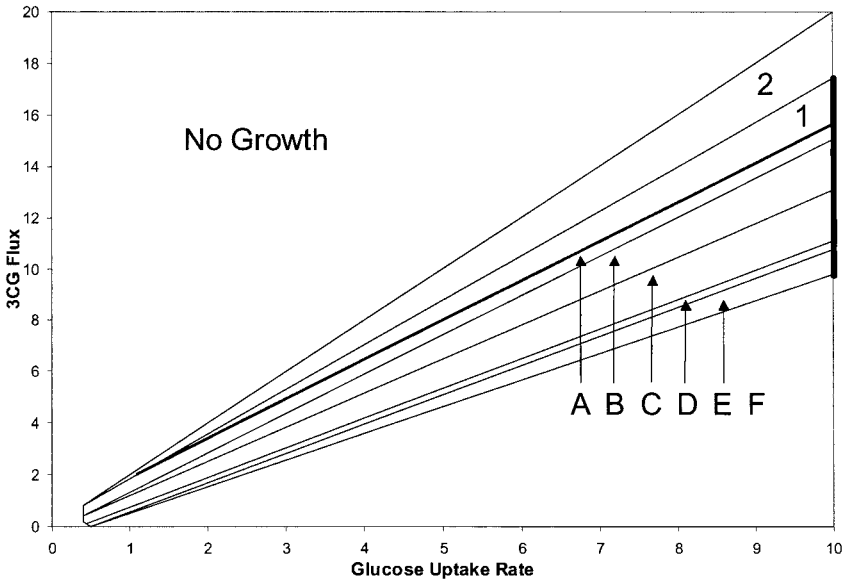

Figure 5. The glucose uptake rate (mmol · g-DW^{-1} · h^{-1})−3-carbon glycolytic flux (substrates converted · g-DW^{-1} · h^{-1}) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to 3CG fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify 3CG flux reductions below the optimal relation. The metabolic fluxes along the thick line (glucose uptake rate = 10 mmol · g-DW^{-1} · h^{-1}) are shown in Figure 6.

defined as 1−2 (as shown in Figure 5). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all 3CG fluxes from zero to the maximum allowable flux (the glucose uptake exchange flux was constrained to 10 mmol g-DW^{-1} h^{-1}) that still permits cellular growth, and the results are shown in Figure 2.

The optimal relation between the glucose uptake and the 3CG flux was calculated (Figure 5). The sensitivity of other optimal fluxes in the metabolic network upon the reduction of the 3CG flux was examined (Figure 6). With 3CG flux reduction just below the optimal value, the optimal metabolic network operation was characterized by region A (Figure 5). In region A, the 3CG flux reduction led to increased PPP fluxes and the transhydrogenase was used (Figure 6B); additionally, the TCA cycle flux was reduced (Figure 6C). The reduced 3CG flux also led to the reduction of the PYK flux, which was optimally completely inactivated at the demarcation between regions A and B (Figure 6A). In region B, the glyoxylate bypass was optimally utilized and the TCA cycle fluxes were further reduced. At the demarcation between region B and C, the TCA cycle no longer operated cyclically but rather served to generate the biosynthetic precursors. In region C, glucokinase was included in the optimal flux vector, the inclusion of the glucokinase decoupled the phosphoenolpyruvate to pyruvate biochemical conversion and the uptake of glucose, thus allowing for the 3CG flux to be decreased and with little effect on the maximal growth flux. The growth flux at the demarcation between region C and D was 98% of the *in silico* wildtype (Figure 2).

Regions D and E were very similar with respect to the optimal metabolic flux vector. In these regions, the pyruvate−formate lyase was optimally active and the glyoxylate bypass was no longer included in the optimal flux vector. Additionally, in regions D and E, the growth flux was more sensitive (compared to regions A−C) to the 3CG flux, and at the demarcation between regions E and F the maximal growth flux was 95% of the *in silico* wildtype.

In the final region of reduced 3CG flux (region F, metabolic flux less than 63% of the *in silico* wildtype), the metabolic network was limited in the ability to produce the essential biosynthetic precursors below the block in the metabolic network. In this region, the optimal growth flux was sensitive to the 3CG flux, and the maximal growth flux linearly decreased to zero as the 3CG flux was reduced to zero from the region E,F boundary. The metabolic fluxes in this region are not shown in Figure 6 because alternate optimal solutions exist. Cellular growth is limited by the availability of the biosynthetic precursors after the metabolic blockage, and the diversion of the flux from glycolysis to the PPP resulted in excess high-energy phosphate bonds and redox potential. The growth flux in region F is dependent upon increased oxygen availability to eliminate the excess redox potential. An additional constraint was imposed on the metabolic network, i.e., the oxygen uptake was constrained below a physiologically realistic value of 20 mmol g-DW^{-1} h^{-1}, and the feasible set did not contain a growth flux for 3CG fluxes below about 40% of the *in silico* wildtype (not shown). Thus, the partial inhibition of the 3CG fluxes can theoretically prevent the growth of *E. coli*; however, growth can be maintained with reduced glucose uptake rates.

The holistic effect of increased 3CG flux on the metabolic networks capability to support cellular growth was assessed with FBA. The optimal growth flux as a function of the flux in the 3CG flux was calculated (Figure 2). The metabolic flux was continuously increased *in silico* from the *in silico* wildtype value to the maximum flux that permits growth, and two qualitatively different metabolic flux vectors (numbered 1 and 2 in Figure 5) were observed.

In region 1 (Figure 5), the optimal metabolic flux vector was characterized by a decreased pentose phosphate pathway (PPP), an active transhydrogenase reaction, an increased PYK flux, and an increased TCA cycle flux. However, region 1 only extends to a 3CG flux of 110% of the *in silico* wildtype (with a glucose uptake of 10 mmol g-DW^{-1} h^{-1}), and region 2 of Figure 5 was characterized by alternate optimal flux distributions. The metabolic network was limited in the ability to produce the essential biosynthetic precursors before the effected reac-

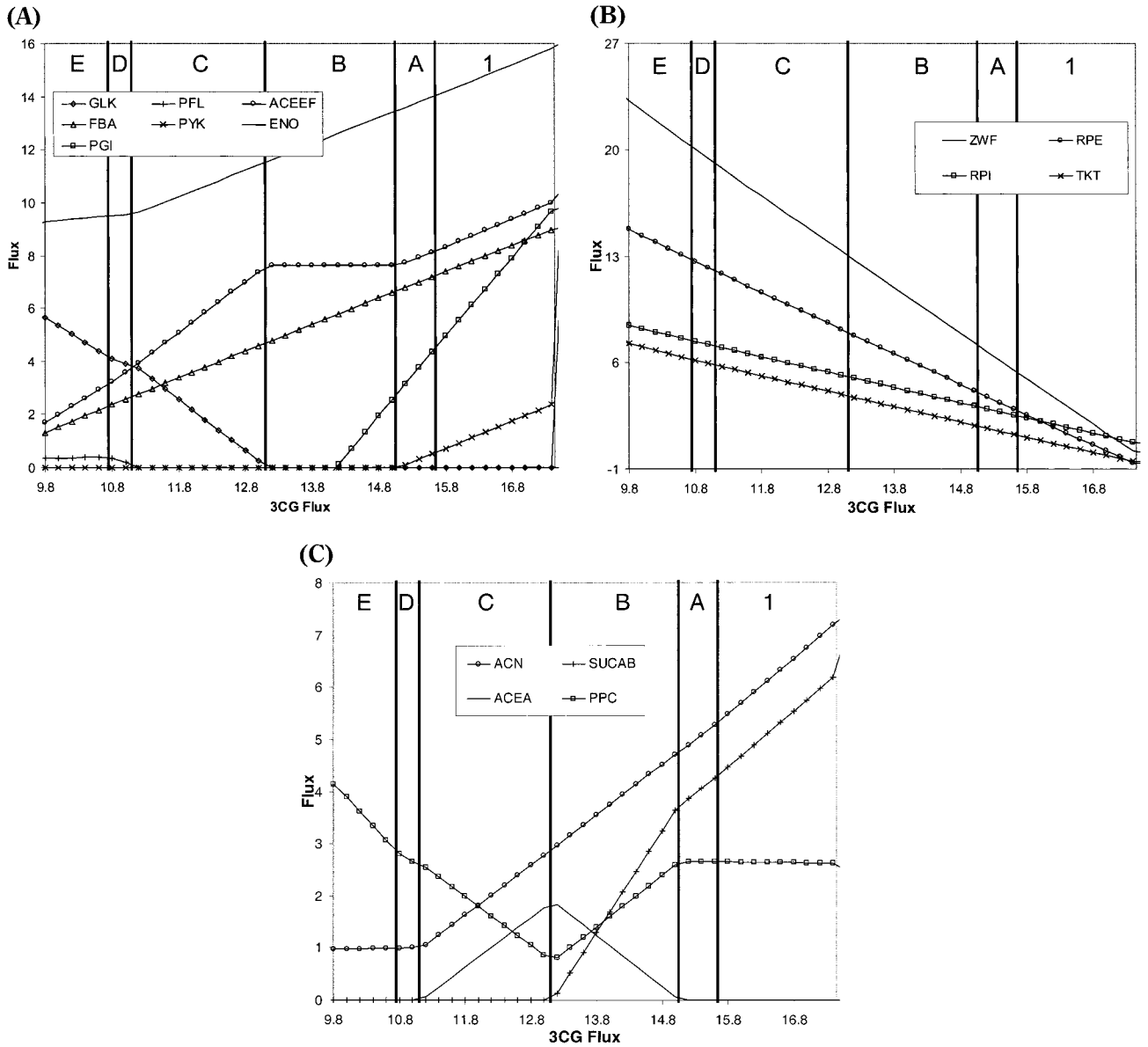

Figure 6. Optimal intracellular fluxes (substrates converted · g-DW^{-1} · h^{-1}) in the central metabolic pathways as a function of the 3CG metabolic flux constraint (substrates converted · g-DW^{-1} · h^{-1}). The glucose uptake rate was constrained to 10 mmol g-DW^{-1} h^{-1}. (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

tions of 3-carbon glycolysis. In this region, the optimal growth flux was sensitive to the 3CG flux. The metabolic fluxes in this region are not shown in Figure 6 because alternate optimal solutions exist. The excess pyruvate produced by the elevated 3CG flux can be converted to any of the metabolic byproducts with the same value of the objective function. Furthermore, the excess high-energy phosphate bonds can be eliminated in any futile cycle, and thus alternate optimal solutions exist.

TCA Cycle. The initial three fluxes of the TCA cycle were determined to be essential. The deletion of any of these enzymatic activities resulted in a glutamate requirement. This requirement has been shown experimentally (*45*). The *in silico* robustness analysis was performed to assess the effect of decreased (and increased) flux entering the TCA cycle (Figures 2, 7, and 8). The ability of the metabolic network to support growth with a reduction in the TCA cycle flux was investigated *in silico* by continuously restricting the citrate synthase flux, which we will refer to as the TCA cycle flux. As the TCA cycle flux constraint was reduced from the *in silico* wildtype, it was determined that the ability of the metabolic network to support growth was not sensitive to the TCA cycle flux above 18% of the *in silico* wildtype. Furthermore, the TCA cycle flux could be increased to about 160% of the *in silico* wildtype before severe limitations in the growth flux were encountered (Figure 2). We have investigated the metabolic response to TCA cycle flux level alterations by a phenotype phase plane analysis (Figure 7).

The PhPP describing the changes in the metabolic pathway utilization as a function of the TCA cycle flux and the glucose uptake rate was calculated (Figure 7). The optimal relation between the glucose uptake rate and the TCA cycle flux was determined from the PhPP (Figure 7). It was determined that there were four qualitatively different regions of optimal metabolic pathway utilization for TCA cycle fluxes lower than optimal, and these regions were defined as A−D (as shown in Figure 7). Furthermore, there were determined to be four qualitatively different regions for TCA cycle fluxes greater than optimal, and these regions were defined as 1−4 (as shown in Figure 7). The maximal growth flux (normalized to the *in silico* wildtype) was calculated for all TCA cycle

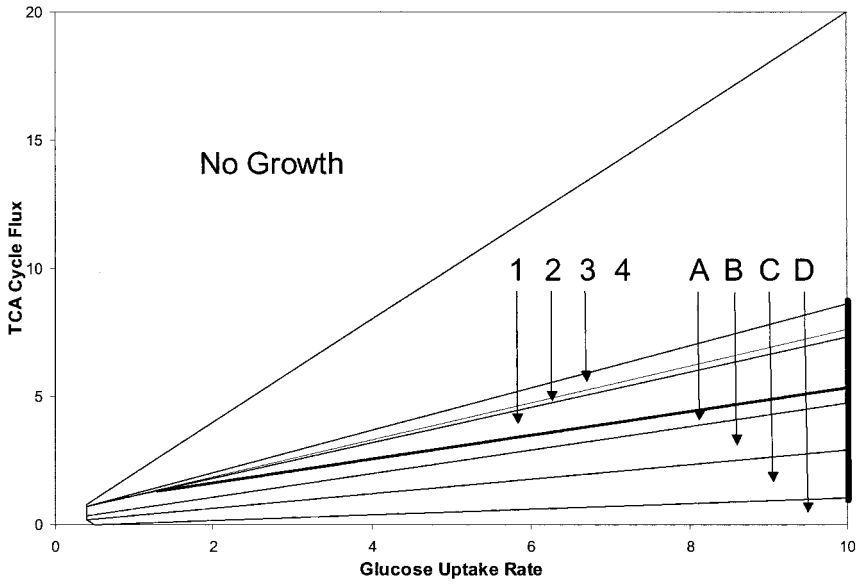

Figure 7. The glucose uptake rate (mmol · g-DW^{-1} · h^{-1})−TCA cycle flux (substrates converted · g-DW^{-1} · h^{-1}) phenotype phase plane. Exchange flux constraints were defined as discussed in the text. The regions are numbered and lettered. The numbered regions correspond to TCA cycle fluxes that are increased relative to the optimal value. The optimal relation is the thick demarcation line. The lettered regions identify TCA cycle flux reductions below the optimal relation. The metabolic fluxes along the thick line (glucose uptake rate = 10 mmol · g-DW^{-1} · h^{-1}) are shown in Figure 8.

fluxes from zero to the maximum allowable flux (The glucose uptake exchange flux was constrained to 10 mmol g-DW^{-1} h^{-1}) that still permits cellular growth, and the results are shown in Figure 2.

The optimal relation between the glucose uptake and the TCA cycle flux was calculated, and the sensitivity of the optimal fluxes in central metabolism to the TCA cycle flux was examined (Figure 8). Region A defines the optimal set of metabolic reactions that are utilized with a reduction of the TCA cycle flux below the optimal value (Figure 7). In region A, the reduction of the TCA cycle flux led to increased PPP fluxes and the transhydrogenase was used; additionally, the glycolytic flux was decreased. The reduced TCA cycle flux led to the reduction of the PYK flux, which was optimally completely inactivated at the demarcation between regions A and B. In region B, the glyoxylate bypass was optimally utilized and the glycolytic fluxes were further reduced. At the demarcation between region B and C, the TCA cycle fluxes were reduced to the point that the TCA cycle no longer operated cyclically but rather served to generate the biosynthetic precursors. In region C, glucokinase was included in the optimal flux vector, and the inclusion of the glucokinase decoupled the phosphoenolpyruvate to pyruvate biochemical conversion and the uptake of glucose. The growth flux at the demarcation between region C and D was 98% of the *in silico* wildtype (Figure 2) and the TCA cycle flux was 18% of the *in silico* wildtype TCA cycle flux.

In the final region of reduced TCA cycle fluxes (metabolic flux less than 18% of the *in silico* wildtype), the metabolic network was limited in the ability to produce α-ketoglutarate, an essential biosynthetic precursor. In this region, the optimal growth flux was sensitive to the TCA cycle flux, and the characteristic behavior was similar to region F of the 3CG-glucose uptake rate PhPP (Figure 5) that was discussed above.

The holistic effect of increased TCA cycle flux on the metabolic networks' capability to support cellular growth was assessed with FBA. The optimal growth flux as a function of the flux in the TCA cycle flux was calculated (Figure 2). The metabolic flux was continuously increased *in silico* from the *in silico* wildtype value to the maximum flux that permits growth, and four qualitatively different metabolic flux vectors (numbered 1−4 in Figure 7) were observed.

The region 1 of Figure 7 the optimal metabolic flux vector was characterized by an active transhydrogenase reaction, an increased PYK flux, an increased glycolytic flux, and a decreased PPP (which was optimally inactivated at the demarcation between regions 1 and 2). Regions 2 and 3 are very similar with respect to the set of metabolic reactions that are optimally utilized, and in these regions, the PPP is optimally inactivated. At the demarcation between regions 3 and 4, the maximal growth flux was about 95% of the *in silico* wildtype and the TCA cycle flux was increased to approximately 160% of the *in silico* wildtype. With TCA cycle flux increases beyond region 3, region 4 is encountered (Figure 7). Region 4 was characterized by alternate optimal flux distributions. The metabolic network was limited in the ability to produce the essential glycolytic and PPP biosynthetic precursors. In this region, the optimal growth flux was sensitive to the TCA cycle flux, and the metabolic fluxes in this region are not shown in Figure 8 because alternate optimal solutions exist.

Discussion

We have illustrated, with the complete *E. coli* metabolic network, how optimal metabolic phenotypes (flux vectors) and shifts in metabolic behavior can be analyzed and interpreted *in silico*. From the fundamental physicochemical constraints on the metabolic network, the feasible set that identifies the capabilities of the metabolic network was identified. Subsequently, a linear optimization routine was utilized to search the feasible set for a flux vector that maximizes a given objective function. Given the complexity associated with developing complete dynamic modeling of cellular processes, the constraining approach, as discussed herein, is a particularly useful alternative approach to metabolic systems analysis. The results presented herein are of fundamental interest for several reasons. First, the ability to define essential genes under various conditions will have many

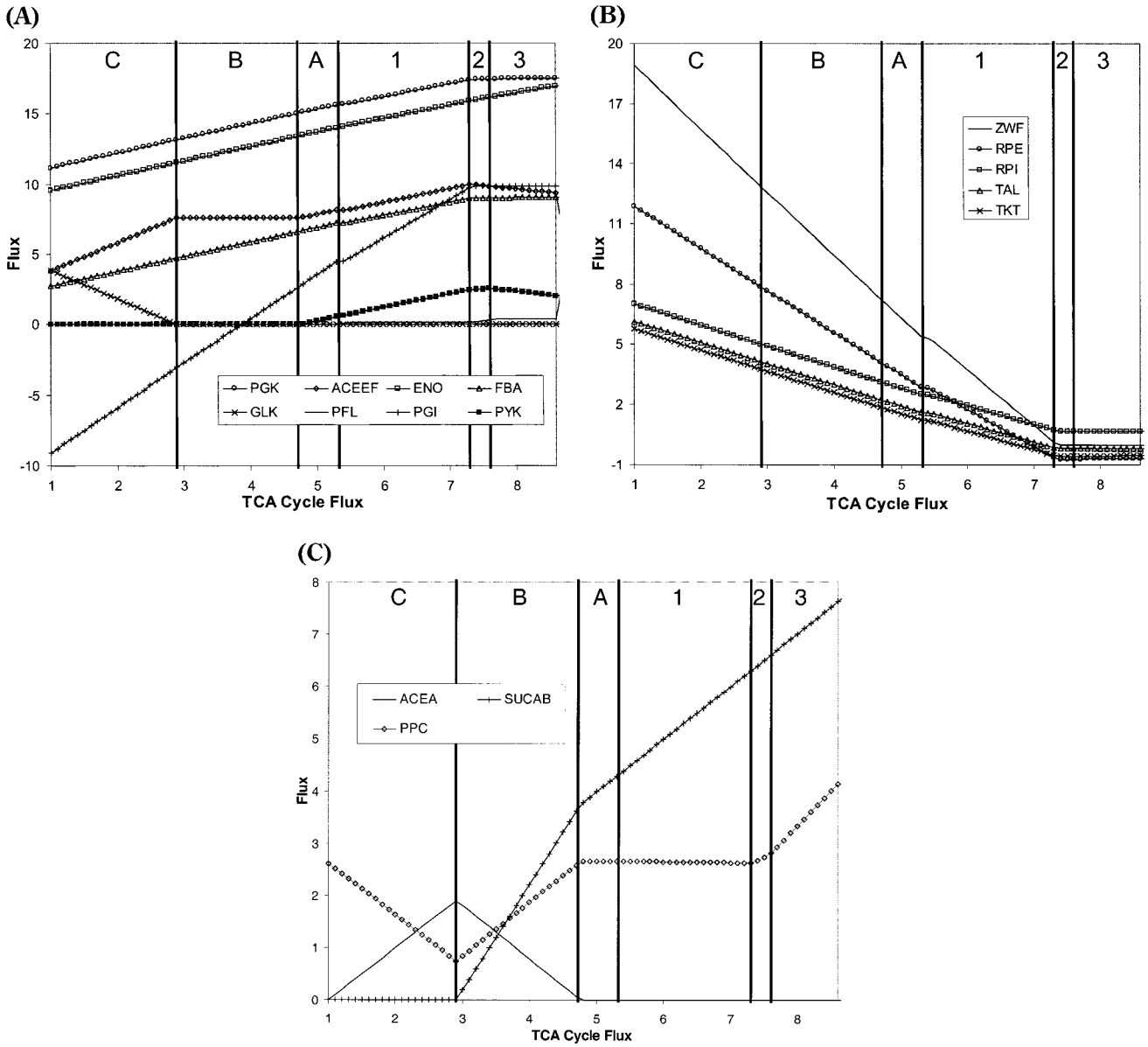

Figure 8. Optimal intracellular fluxes (substrates converted · g-DW^{-1} · h^{-1}) in the central metabolic pathways as a function of the TCA cycle metabolic flux constraint (substrates converted · g-DW^{-1} · h^{-1}). The glucose uptake rate was constrained to 10 mmol g-DW^{-1} h^{-1}. (A) Glycolytic fluxes. (B) Pentose phosphate pathway fluxes. (C) TCA cycle fluxes.

practical applications. Second, the results presented discussed the sensitivity of the objective function to specific fluxes in the metabolic network. Finally, the results demonstrate the potential capabilities of *in silico* analysis of cellular systems. Understanding the relation between individual fluxes and the holistic function of the metabolic network is essential to successfully metabolic engineering a living system, and FBA provides a methodology that can be used to direct the metabolic engineer.

First, we should address the assumptions associated with the utilization of linear optimization to identify the optimal flux vector. The metabolic constraining formalism that we have discussed is based on the fundamental physicochemical constraints that all cells must abide to. Within the set of constraints, the cell will choose a flux vector for which to operate. We have attempted to find the same flux vector by employing linear optimization, and the assumption is that the cell has evolved the regulatory mechanisms to find the optimal solution within the physicochemical constraints to maximize its survival. We have mathematically represented survival as cellular growth. On the basis of comparisons of the linear optimization results and experimental data, the assumption appears to be valid under the tested conditions (*27*). However, currently, the number of situations for which the validity of the assumption has been addressed is limited. Therefore, further experimental validation is in order. It should be noted that, even if the assumption of optimal growth proves correct for wildtype strains, it is not clear if we should expect that an engineered strain will behave in an optimal manner. Therefore, the optimization results may only provide an upper bound on the expected behavior of engineered strains.

The identification of the essential gene products in a metabolic network is of fundamental interest (*46, 47*). The metabolic constraining formalism that was described here provides an efficient method to study the consequences of alterations in the genotype and to gain insight into the genotype-phenotype relation. The study of the removal of individual metabolic enzymes in the central metabolic pathways demonstrated fundamental redundancy properties of the *E. coli* metabolic genotype and the existence of relatively few critical gene products. Seven metabolic reactions were determined to be ess-

ential in the metabolic network. Of these seven reactions, one set of three forms a linear reaction series and another set of two forms another linear reaction series; therefore, the effect of deleting any of the reactions in the linear set is equivalent. Thus, there were basically four different metabolic fluxes that were essential: two in the PPP, the 3-carbon stage of glycolysis, and the first three reactions of the TCA cycle. It was shown that there are regions where the ability of the metabolic network to support growth was not affected by the altered flux levels in the essential enzymatic reactions. This metabolic robustness was due to the capability of the metabolic network to shift the production of redox potential and high-energy phosphate bonds. However, the point where the metabolic network was unable to support growth was quantitatively derived *in silico*. At this point, the metabolic network was limited in the ability to produce biosynthetic precursors. In the biosynthetic precursor limited region, the availability of redox potential and high-energy phosphate bonds did not limit the growth capability of the cell.

The robustness analysis presented herein is essentially a sensitivity analysis. The sensitivity of the objective function to changes in the optimal flux vector was addressed. A constraint was added to the metabolic network, thus altering a single metabolic flux, and a new optimal flux vector was calculated. The relation between the additional constraint and the objective function was examined to investigate the robustness in the system with respect to the essential enzymatic reactions. Furthermore, the relation between the additional constraint and optimal flux vector was also examined. The results of the sensitivity analysis can be used to interpret the optimal metabolic fluxes and their relation to in vivo metabolic fluxes. For example, metabolic fluxes for which the objective function is highly sensitive are likely to obtain in vivo values that are consistent with the optimal values. Furthermore, metabolic fluxes to which the objective is not sensitive may be able to take on flux values in a large range with very little effect on the optimal solution. Additionally, other fluxes in the metabolic network that are sensitive to a given metabolic flux may provide an indication of the accuracy of the optimal predictions. For example, the TCA cycle flux was reduced to approximately 18% of the *in silico* wildtype with little effect on the objective function; however, the PPP fluxes were three times the *in silico* wildtype, fluxes much higher than experimental data indicates.

Understanding the metabolic fluxes and their control is essential to the ability to "design" metabolic networks for the production commodity chemicals (i.e., antibiotics, vitamins, amino acids, etc). Using flux balance analysis the complete range of metabolic phenotypes can be examined under defined environmental and genetic conditions through the use of phenotypic phase planes (PhPP). For the *E. coli* network PhPPs were generated for growth on glucose minimal media spanning the uptake rate of the carbon source and an intracellular metabolic flux. Therefore, the PhPP formalism has provided an efficient methodology for examining the consequence of altered fluxes within the cell. Therefore, bioinformatically based models will undoubtedly have a major impact on the development of metabolic engineering (*4, 5*). Herein, we have investigated the effect of altered metabolic flux levels on the maximal growth flux, thus quantifying the relation between altered flux levels and optimal cellular growth. Furthermore, by examining the entire set of constraints on the metabolic network, constraints on metabolic flux alterations can be identified.

FBA incorporates no information on enzyme kinetics or gene regulation, thus limiting insight into dynamic responses. From flux balance analysis it is possible to realize some of the fundamental constraints that metabolic systems are faced with and define the feasible set that contains all admissible steady-state flux vectors. As in vivo reaction dynamics is further understood, the ability to predict dynamic responses of metabolic networks to environmental and genetic perturbations using dynamic modeling approaches will become more feasible. In general regulatory schemes and reaction dynamics will serve to further constrain metabolic behavior to operate in confined subspaces of the feasible set. Identifying these regions from both the theoretical and experimental side will be a challenge for the future.

A number of experimental technologies have now made the holistic study of biological systems feasible. The ability to assimilate DNA chip-based and protein expression technologies providing genome-scale information with computational methods for metabolic network analysis will become important in advancing the study of metabolic physiology and the practice of metabolic engineering. Currently the interpretation of high-throughput experimental information on systemic behavior is limited by a lack of analysis capabilities. Can systems-based quantitative *in silico* approaches such as flux balance be used to assist in understanding this flood of data? This question will need to be answered as interest builds in the genomics community for quantitative systems analysis.

The analysis of the metabolic phenotype−genotype relation using the bioinformatically based *in silico* metabolic genotype of *E. coli* can serve as a basis for the construction of parallel *in silico* representations of other single-cell organisms. Thus, the results presented are particularly relevant with the current emphasis on genome sequencing. Utilizing the techniques described herein, information can be gained regarding the metabolic physiology of a cell with relatively little experimental biochemical information on the cell of interest. However, this analysis should be considered a single step toward the integrative analysis of bioinformatic databases to predict and understand cellular function based on the underlying genetic content. Continued prediction and experimental verification will be an integral part of the further development of *in silico* strains and their use to represent their in vivo counterparts.

Acknowledgment

Portions of this work are supported by grants from the National Institute of Health (57089-01A1) and the National Science Foundation (9873384).

References and Notes

(1) Bhalla, U. S.; Iyengar, R. Emergent properties of networks of biological signaling pathways. *Science* **1999**, *283*, 381−387.

(2) Weng, G.; Bhalla, U. S.; Iyengar, R. Complexity in biological signaling systems. *Science* **1999**, *284*, 92−96.

(3) Kanehisa, M. Databases of biological information. *Trends Guide Bioinf.* **1998**, 24−26.

(4) Palsson, B. O. What lies beyond bioinformatics? *Nat. Biotechnol.* **1997**, *15*, 3−4.

(5) Edwards, J. S.; Palsson, B. O. How will bioinformatics influence metabolic engineering? *Biotechnol. Bioeng.* **1998**, *58*, 162−169.

(6) Goodwin, B. C. *Oscillatory Organization in Cells, a Dynamic Theory of Cellular Control Processes*; Academic Press: New York, 1963.

(7) Tyson, J. J.; Othmer, H. G. The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* **1978**, *5*, 1−62.
(8) Savageau, M. The behaviour of intact biochemical control systems. *Curr. Top. Cell Regul.* **1972**, *6*, 64−130.
(9) Hess, B.; Boiteux, A. Oscillatory organization in cells, a dynamic theory of cellular control processes. *Hoppe-Seylers Z. Physiol. Chem.* **1968**, *349*, 1567−1574.
(10) Savageau, M. A. Biochemical systems analysis. II. The steady-state solutions for an n- pool system using a power-law approximation. *J. Theor. Biol.* **1969**, *25*, 370−379.
(11) Heinrich, R.; Rapaport, S. M.; Rapaport, T. A. Metabolic regulation and mathematical models. *Prog. Biophys. Mol. Biol.* **1977**, *32*, 1−82.
(12) Kacser, H.; Burns, J. A. The control of flux. *Symp. Soc. Exp. Biol.* **1973**, *27*, 65−104.
(13) Fell, D.: *Understanding the Control of Metabolism*; London: Portland Press: 1996.
(14) Lee, I. D.; Palsson, B. O. A Macintosh software package for simulation of human red blood cell metabolism. *Comput. Methods Prog. Biomed.* **1992**, *38*, 195−226.
(15) Joshi, A.; Palsson, B. O. Metabolic dynamics in the human red cell. Part I-. A comprehensive kinetic model. *J. Theor. Biol.* **1989**, *141*, 515−528.
(16) Mulquiney, P. J.; Kuchel, P. W. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis. *Biochem. J.* **1999**, *342 Pt 3*, 597−604.
(17) Palsson, B. O.; Joshi, A.; Ozturk, S. S. Reducing complexity in metabolic networks: making metabolic meshes manageable. *Fed. Proc.* **1987**, *46*, 2485−2489.
(18) Bailey, J. E. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnol. Prog.* **1998**, *14*, 8−20.
(19) Bonarius, H. P. J.; Schmid, G.; Tramper, J. Flux analysis of underdetermined metabolic networks: The quest for the missing constraints. *Trends Biotechnol.* **1997**, *15*, 308−314.
(20) Edwards, J. S.; Ramakrishna, R.; Schilling, C. H.; Palsson, B. O. Metabolic Flux Balance Analysis. In *Metabolic Engineering*; Lee, S. Y., Papoutsakis, E. T., Eds.; Marcel Dekker: New York, 1999; pp 13−57.
(21) Edwards, J. S.; Palsson, B. O. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* **1999**, *274*, 17410−17416.
(22) Varma, A.; Palsson, B. O. Metabolic Flux Balancing: Basic concepts, Scientific and Practical Use. *Bio/Technology* **1994**, *12*, 994−998.
(23) Sauer, U.; Cameron, D. C.; Bailey, J. E. Metabolic capacity of *Bacillus subtilis* for the production of purine nucleosides, riboflavin, and folic acid. *Biotechnol. Bioeng.* **1998**, *59*, 227−238.
(24) Edwards, J. S.; Palsson, B. O. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5528−5533.
(25) Palsson, B. O.; Lightfoot, E. N. Mathematical modelling of dynamics and control in metabolic networks. Part I. On Michaelis−Menten kinetics. *J. Theor. Biol.* **1984**, *111*, 273−302.
(26) Clarke, B. L. Stoichiometric network analysis. *Cell Biophys.* **1988**, *12*, 237−253.
(27) Varma, A.; Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic byproduct secretion in wildtype *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **1994**, *60*, 3724−3731.
(28) Walsh, K.; Koshland, D. Determination of flux through the branch point of two metabolic cycles. *J. Biol. Chem.* **1984**, *259*, 9646−9654.
(29) Karp, P. D.; Riley, M.; Saier, M.; Paulsen, I. T.; Paley, S. M.; Pellegrini-Toole, A. The EcoCyc and MetaCyc databases. *Nucleic Acids Res.* **2000**, *28*, 56−9.
(30) Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29−34.
(31) Selkov, E., Jr.; Grechkin, Y.; Mikhailova, N.; Selkov, E. MPW: the Metabolic Pathways Database. *Nucleic Acids Res.* **1998**, *26*, 43−45.
(32) Neidhardt, F. C.; Umbarger, H. E. Chemical Composition of *Escherichia coli*. In *Escherichia coli and Salmonella: cellular and molecular biology*, 2nd ed.; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 13−16.
(33) Pramanik, J.; Keasling, J. D. Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.* **1997**, *56*, 398−421.
(34) Varma, A.; Palsson, B. O. Parametric sensitivity of stoichiometric flux balance models applied to wildtype *Escherichia coli* metabolism. *Biotechnol. Bioeng.* **1995**, *45*, 69−79.
(35) Varma, A.; Palsson, B. O. Predictions for Oxygen Supply Control to Enhance Population Stability of Engineered Production Strains. *Biotechnol. Bioeng.* **1994**, *43*, 275−285.
(36) Schilling, C. H.; Palsson, B. O. The underlying pathway structure of biochemical reaction networks. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4193−4198.
(37) Schilling, C. H.; Edwards, J. S.; Palsson, B. O. Towards metabolic phenomics: Analysis of genomic data using flux balances. *Biotechnol. Progr.* **1999**, *15*, 288−95.
(38) Schilling, C. H.; Schuster, S.; Palsson, B. O.; Heinrich, R. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnol. Prog.* **1999**, *15*, 296−303.
(39) Edwards, J. S.; Ramakrishna, R.; Palsson, B. O. *Characterizing phenotypic plasticity: A phase plane analysis*, submitted for publication.
(40) Varma, A.; Palsson, B. O. Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.* **1993**, *165*, 477−502.
(41) Fraenkel, D. G. Glycolysis. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1. pp 189−198.
(42) Josephson, B. L.; Fraenkel, D. G. Sugar metabolism in transketolase mutants of *Escherichia coli*. *J. Bacteriol.* **1974**, *118*, 1082−1089.
(43) Josephson, B. L.; Fraenkel, D. G. Transketolase mutants of *Escherichia coli*. *J. Bacteriol.* **1969**, *100*, 1289−1295.
(44) Skinner, A. J.; Cooper, R. A. Genetic studies on ribose 5-phosphate isomerase mutants of *Escherichia coli* K-12. *J. Bacteriol.* **1974**, *118*, 1183−1185.
(45) Cronan, J. E., Jr.; Laporte, D. Tricarboxylate Acid cycle and glyoxylate bypass. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 189−198.
(46) Arigoni, F.; Talabot, F.; Peitsch, M.; Edgerton, M. D.; Meldrum, E.; Allet, E.; Fish, R.; Jamotte, T.; Curchod, M. L.; Loferer, H. A genome-based approach for the identification of essential bacterial genes. *Nat. Biotechnol.* **1998**, *16*, 851−856.
(47) Akerley, B. J.; Rubin, E. J.; Camilli, A.; Lampe, D. J.; Robertson, H. M.; Mekalanos, J. J. Systematic identification of essential genes by in vitro mariner mutagenesis. *Proc. Natl. Acad. Sci. U.S A.* **1998**, *95*, 8927−8932.

Accepted for publication June 16, 2000.

BP0000712