Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 225 (2003) 185–194
Journal of Theoretical Biology
www.elsevier.com/locate/jtbi

Network-based analysis of metabolic regulation in the human red blood cell

Nathan D. Price[1], Jennifer L. Reed[1], Jason A. Papin[1], Sharon J. Wiback, Bernhard O. Palsson*

Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, 92093-0412 USA

Received 10 December 2002; received in revised form 20 May 2003; accepted 3 June 2003

Abstract

Reconstruction of cell-scale metabolic networks is now possible. A description of allowable metabolic network functions can be obtained using extreme pathways, which are the convex basis vectors of the solution space containing all steady state flux distributions. However, only a portion of these allowable network functions are physiologically possible due to kinetic and regulatory constraints. Methods are now needed that enable us to take a defined metabolic network and deduce candidate regulatory structures that control the selection of these physiologically relevant states. One such approach is the singular value decomposition (SVD) of extreme pathway matrices (**P**), which allows for the characterization of steady state solution spaces. Eigenpathways, which are the left singular vectors from the SVD of **P**, can be described and categorized by their biochemical function. SVD of **P** for the human red blood cell showed that the first five eigenpathways, out of a total of 23, effectively characterize all the relevant physiological states of red blood cell metabolism calculated with a detailed kinetic model. Thus, with five degrees of freedom the magnitude and nature of the regulatory needs are defined. Additionally, the dominant features of these first five eigenpathways described key metabolic splits that are indeed regulated in the human red blood cell. The extreme pathway matrix is derived directly from network topology and only knowledge of V_{max} values is needed to reach these conclusions. Thus, we have implemented a network-based analysis of regulation that complements the study of individual regulatory events. This topological approach may provide candidate regulatory structures for metabolic networks with known stoichiometry but poorly characterized regulation.

Keywords: Extreme pathways; Singular value decomposition; SVD; Constraint-based models; Elementary modes

1. Introduction

It is now possible to construct cell-scale networks describing various cellular functions (Covert et al., 2001). A major challenge we are now faced with is to develop in silico modeling strategies that generate meaningful network-level analyses. Various modeling philosophies and approaches have been pursued, including deterministic kinetic models (Joshi and Palsson, 1989; Tomita et al., 1999; Werner and Heinrich, 1985), stochastic models (Arkin et al., 1998; Elowitz et al., 2002; McAdams and Arkin, 1997), and constraint-based models (Edwards et al., 2002; Palsson, 2000; Price et al., 2003a; Schuster et al., 2000). Constraint-based models use governing constraints, such as mass balance and maximum reaction rates, to restrict potential cellular behavior. This modeling approach has been applied to metabolism and results in a solution space in which all feasible steady state flux distributions reside (Covert et al., 2001; Palsson, 2000; Papin et al., 2003). The edges of this space, which form a convex cone, can be calculated and are called the extreme pathways (Fig. 1a). All valid steady state flux distributions through the metabolic network are nonnegative linear combinations of these extreme pathways. If all kinetic parameters are known, then kinetic models can be used to calculate a single solution that lies within the convex cone formed by the extreme pathways (Fig. 1a). If the kinetic constants are not fully known, then partial

*Corresponding author. Tel.: +1-858-534-5668; fax: +1-858-822-3120.
E-mail address: bpalsson@ucsd.edu (B.O. Palsson).
[1] These authors contributed equally to this work.

0022-5193/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0022-5193(03)00237-6

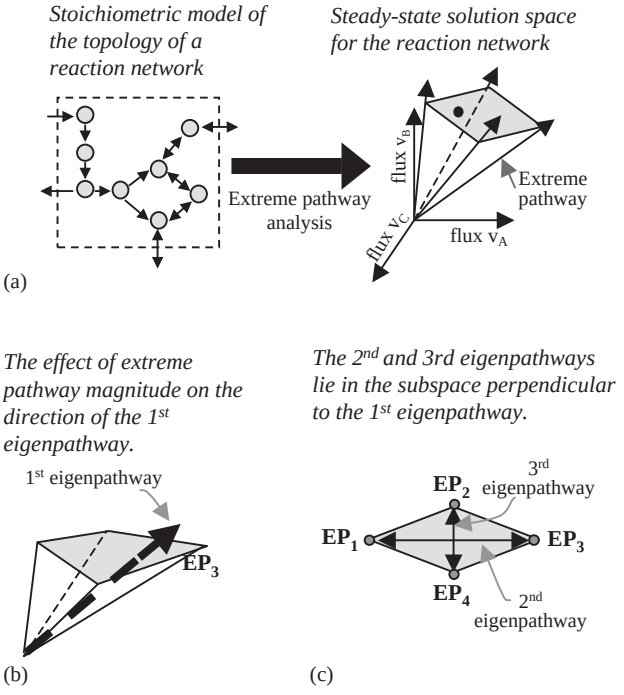

Fig. 1. Conceptual framework for application of singular value decomposition to extreme pathway analysis. Extreme pathway analysis generates a set of vectors from network stoichiometry that defines a convex solution space (panel a). This solution space circumscribes all possible phenotypes of the biochemical reaction network. A complete kinetic model can be used to calculate a single solution (shown as a black dot) at which the network operates. All steady state solutions of the kinetic model will lie within the convex solution space. In panel b, the effect of varying extreme pathway magnitude (in this case EP_3) on the calculated first eigenpathway is shown. The first eigenpathway represents the principal direction of the cone; it is pulled towards the longer (or higher flux capacity) extreme pathway (EP_3). In panel c, we can see how the second and third eigenpathways characterize the convex cone. In the plane perpendicular to the first eigenpathway, the second and third eigenpathways characterize the mutually orthogonal directions. Movement along the second eigenpathway allows for the most control of the space characterized by the first eigenpathway.

kinetic and flux information can be used to approximate the location of physiologically meaningful solutions. Thus, to characterize which regions of the steady state solution space are of interest, there is a need to develop methods that bridge the kinetic and constraint-based approaches. Furthermore, we need to find candidate regulatory strategies that direct solutions to these regions.

The human red blood cell provides an excellent model system for the development of biologically relevant computational methodologies. A kinetic whole-cell simulator based on a previously formulated model (Joshi and Palsson, 1989) has been developed for human red blood cell metabolism (Jamshidi et al., 2001). The extreme pathways for the red blood cell metabolic network are known, along with the maximum allowable flux through many of the reactions (Wiback and Palsson, 2002). Thus, both the constraint-based and kinetic models are available to describe the functionalities of the complete human red blood cell metabolic network. The kinetic regulation of metabolism in the red blood cell is also well known. For these reasons, the red blood cell network provides a good test system for evaluating analytical methods.

Regulation can be thought of as the process by which a cell "chooses" its state within the constrained solution space in which it must operate. Study of metabolic regulation has focused on the identification of large numbers of individual regulatory events, their characteristics, and the molecules that participate in them (Davidson, 2001; Ptashne and Gann, 2002). Now with reconstructed metabolic networks available we can approach the study of regulatory issues from a network-based perspective, focusing on the regulatory needs of the network as a whole. Understanding these systemic regulatory needs can provide a holistic view of the organization of individual regulatory components and how they are used to satisfy cellular objectives.

Mathematical approaches have been used to study cellular regulation. Elementary modes (Schuster et al., 2000), closely related to the concept of extreme pathways, have recently been used to make predictions regarding gene expression levels under changing environmental conditions (Stelling et al., 2002). Here we apply singular value decomposition of extreme pathway matrices to study metabolic regulation from a network-based perspective.

2. Materials and methods

2.1. Extreme pathways in human red blood cell metabolism

Extreme pathways define the edges of a high dimensional cone that circumscribes all possible flux distributions in a metabolic network (Fig. 1a) (Schilling et al., 2000). They are calculated from a stoichiometric matrix, **S**, where the columns contain the stoichiometric coefficients of the metabolic reactions and rows correspond to the metabolites. Reversible internal reactions are decoupled into two separate reactions for the forward and reverse directions. The application of mass balance (Eq. (1)) and reaction directionality (Eq. (2)) constraints defines a set of solutions to the metabolic network. The constraints are mathematically represented as

$$\mathbf{S} \cdot \mathbf{v} = 0, \tag{1}$$

$$v_i \geqslant 0, \quad \forall i, \tag{2}$$

where $\mathbf{v}$ is the flux vector containing the individual reaction fluxes, v_i. The set of extreme pathways is the convex basis of **S**, defined by Eqs. (1) and (2). An

algorithm for calculating extreme pathways has been previously published (Schilling et al., 2000). The extreme pathways for the red blood cell metabolic network (shown in Fig. 3) were previously calculated and analysed (Wiback and Palsson, 2002). There are 55 extreme pathways for the red blood cell metabolic network; however, 19 of these extreme pathways are excluded since they are internal loops that must have zero net flux based on thermodynamic principles (Beard et al., 2002; Price et al., 2002).

2.2. Normalization of the extreme pathways

The extreme pathways were scaled to the maximum allowable flux (V_{max}) through the most limiting reaction in each extreme pathway, thus forming a closed solution space. The length of the pathways affects the results obtained from SVD. Scaling by the kinetic parameter V_{max} normalizes each pathway to its functional capacity, as opposed to normalizing each pathway to unit length, which is mathematically convenient but less physiologically relevant. V_{max} values for the reactions in the red blood cell metabolic network have been previously published (Joshi and Palsson, 1990; Wiback and Palsson, 2002). The maximum glucose uptake flux was set to 1.5 mM/h in accordance with the maximum observed value in the human red blood cell model (Jamshidi et al., 2001; Joshi and Palsson, 1990). All extreme pathways studied herein contained at least one reaction with a defined V_{max} (Wiback and Palsson, 2002).

2.3. Singular value decomposition (SVD)

The extreme pathway matrix, $\mathbf{P}$, is formed using the extreme pathway vectors, $\mathbf{P}_i$, as its columns. Each row corresponds to a flux through a given reaction in the metabolic network. $\mathbf{P}$ can be decomposed into three matrices ($\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}^T$) using SVD:

$$\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T. \tag{3}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ contain a basis set for each of the four fundamental subspaces of the matrix, $\mathbf{P}$ (Lay, 1997). The column space of $\mathbf{P}$ contains any steady state flux vector, $\mathbf{v}$, through the metabolic network. The columns of $\mathbf{U}$ and $\mathbf{V}$ were ordered based upon monotonically decreasing singular values (described below). While this is not a mathematical necessity, results from SVD are typically arranged in this manner. The column space is described by the first r columns of $\mathbf{U}$, where r is the rank of $\mathbf{P}$. The first r columns of $\mathbf{U}$ are a set of unique orthonormal vectors, called the left singular vectors. The left singular vectors from the SVD of $\mathbf{P}$ are the eigenvectors of $\mathbf{P}\cdot\mathbf{P}^T$ and are referred to herein as the eigenpathways.

The elements of the diagonal matrix $\mathbf{\Sigma}$ are called the singular values and are ordered so that they decrease in magnitude down the diagonal. The singular values, σ_i, indicate the relative contribution of the columns of $\mathbf{U}$ and $\mathbf{V}$ in reconstructing $\mathbf{P}$ (Lay, 1997). These singular values are the eigenvalues of $\mathbf{P}\cdot\mathbf{P}^T$ (or, equivalently, of $\mathbf{P}^T\cdot\mathbf{P}$) (Lay, 1997). Binary representations of the $\mathbf{P}\cdot\mathbf{P}^T$ and $\mathbf{P}^T\cdot\mathbf{P}$ matrices have been studied in (Papin et al., 2002). The first eigenpathway corresponds to the largest singular value and is thus the vector that most contributes to the reconstruction of $\mathbf{P}$. Each subsequent eigenpathway also contains the next largest contribution to the reconstruction of $\mathbf{P}$, subject to being orthogonal to all the other eigenpathways. If $\mathbf{P}$ were projected into m dimensions, the projection with the highest information content would be a projection onto the first m eigenpathways.

The fractional contribution of each eigenpathway to the reconstruction of the space is the magnitude of the corresponding singular value divided by the sum of all singular values. Thus, the fractional contribution is a measure of how much a particular eigenpathway contributes to the reconstruction of the space and not necessarily how much it is used in reconstructing a particular vector in the space. The cumulative fractional contribution is defined as the sum of the first n fractional contributions, where n varies from 1 to r. Note that in this paper both the fractional contribution and the cumulative fractional contribution are reported as a percent.

2.4. Decomposition of steady state flux vectors into the eigenpathways

Each steady state flux vector, $\mathbf{v}$, which lies within the solution space, can be written as a linear combination of the eigenpathways resulting from the SVD of the extreme pathway matrix as

$$\mathbf{v} = \sum_{i=1}^{r} w_i \cdot \mathbf{u}_i, \tag{4}$$

where $\mathbf{u}_i$ is the ith column vector of $\mathbf{U}$ (or the ith eigenpathway), w_i is the weight for that eigenpathway, and r is the rank of $\mathbf{P}$. The inner product of $\mathbf{v}$ with each one of the eigenpathways gives the corresponding weights

$$\mathbf{v}\cdot\mathbf{u}_j = \left[\sum_{i=1}^{r} w_i \cdot \mathbf{u}_i\right]\cdot\mathbf{u}_j = w_j \tag{5}$$

since the columns of $\mathbf{U}$ are orthonormal (i.e. $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ if $i \neq j$ and $\mathbf{u}_i\cdot\mathbf{u}_j = 1$ if $i = j$).

2.5. Relative error of the reconstructed solution

The weighting needed on each eigenpathway in order to reconstruct the vector, $\mathbf{v}$, can be calculated from Eq. (5). The weighted addition of $k(\leqslant r)$ eigenpathways

brings the recomposed vector, $\mathbf{v}_k$, closer to $\mathbf{v}$ as k increases. When $k = r$, then $\mathbf{v}_r = \mathbf{v}$. The relative error (RE) was calculated to evaluate how much each of the eigenpathways contributed to reconstructing the original steady state flux vector $\mathbf{v}$:

$$\mathbf{E(k)} = \mathbf{v} - \mathbf{v}_k = \mathbf{v} - \sum_{i=1}^{k} w_i \cdot \mathbf{u}_i, \tag{6}$$

$$RE(k) = \frac{\mathbf{E(k)}^{\mathrm{T}} \cdot \mathbf{E(k)}}{\mathbf{v}^{\mathrm{T}} \cdot \mathbf{v}}, \tag{7}$$

where **E(k)** is the difference between the actual vector, $\mathbf{v}$, and the reconstructed vector, $\mathbf{v}_k$, made up by the weighted linear combination of the first k eigenpathways.

2.6. Calculating steady state flux vectors using the kinetic human red blood cell metabolic model

Steady state flux distributions ($\mathbf{v}$) for red blood cell metabolism were obtained using an in silico kinetic model (Jamshidi et al., 2001). The unstressed (no metabolic load applied) physiological steady state was calculated. In addition, the flux distributions corresponding to three types of metabolic loads were calculated to mimic maximal environmental stresses associated with ATP (energy load), NADPH (oxidative load), and NADH (methemoglobin load). These loads correspond to the primary metabolic demands placed upon the red blood cell based upon its physiologic role in the human body. One other main metabolic function of the red blood cell is control of the oxygen–hemoglobin binding curve by regulating the levels of 2,3-diphosphoglycerate (23DPG). The effect of a 23DPG load was studied, but because of low demand flux for 2,3-DPG, the results based on including that load were very similar to the unloaded state (results not shown). Loads can be simulated in silico by assigning increased fluxes to non-specific conversions of cofactors (which are zero in the unstressed state) (Jamshidi et al., 2001; Wiback and Palsson, 2002). The maximum environmental stress sustainable by red blood cell metabolism was defined as the highest load the cell could sustain based on results from the kinetic model.

3. Conceptual framework

3.1. Applying SVD to extreme pathway matrices

The SVD of the extreme pathway matrix, **P**, has been used to characterize the steady state solution space circumscribed by the extreme pathways (Price et al., 2003b). Since the first eigenpathway corresponds to the greatest contribution to the reconstruction of P, it is the single vector that best characterizes the solution space (see Fig. 1b). By decoupling all reversible exchange fluxes in the extreme pathways, all elements of **P** are non-negative, ensuring that the first eigenpathway is a valid biochemical pathway (Price et al., 2003b); it satisfies mass balance and reaction directionality constraints and is directed through the "center" of the cone. The first eigenpathway does not necessarily lie in the geometric center of the cone; rather, it may be "pulled" towards the edges that have either a higher density of extreme pathways or extreme pathways with higher magnitude (see Fig. 1b). Eigenpathways subsequent to the first characterize key tradeoffs in the metabolic network (see Fig. 1c). For example, moving from EP_1 to EP_2 within the solution space requires the addition of vectors (described by the second and third eigenpathways) to the first eigenpathway. A detailed description of SVD of extreme pathway matrices has been previously published (Price et al., 2003b).

4. Results

4.1. SVD of the human red blood cell extreme pathway matrix

The rank of the V_{max}-limited red blood cell extreme pathway matrix, **P**, was 23. The first eigenpathway has a fractional contribution of 47% (Fig. 2). Combined, the first five eigenpathways have a cumulative fractional contribution of 86%, while the cumulative fractional contribution of first nine eigenpathways is 96%. When the extreme pathways were normalized to a unit length

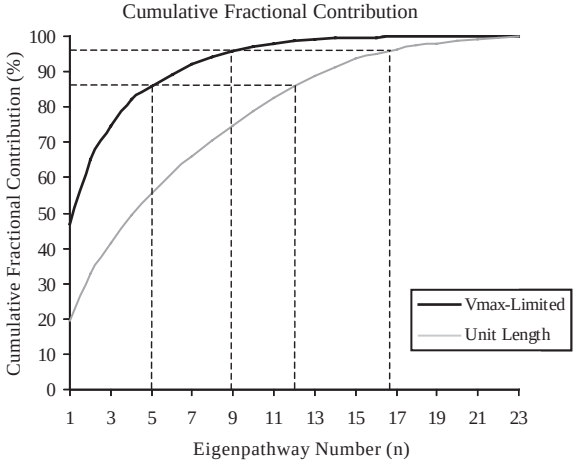

Fig. 2. Cumulative fractional contribution resulting from the SVD of two pathway matrices: V_{max}-limited pathways and pathways of unit length. The cumulative fractional contribution is the sum of the first n fractional singular values (as a percent). The cumulative fractional contribution from the V_{max}-limited pathway matrix is shown in black while the cumulative fractional contribution for the unit pathway matrix is shown in gray. Two cutoff values are also shown, indicating how many eigenpathways are needed for the cumulative fractional contribution to be 86% and 96%.

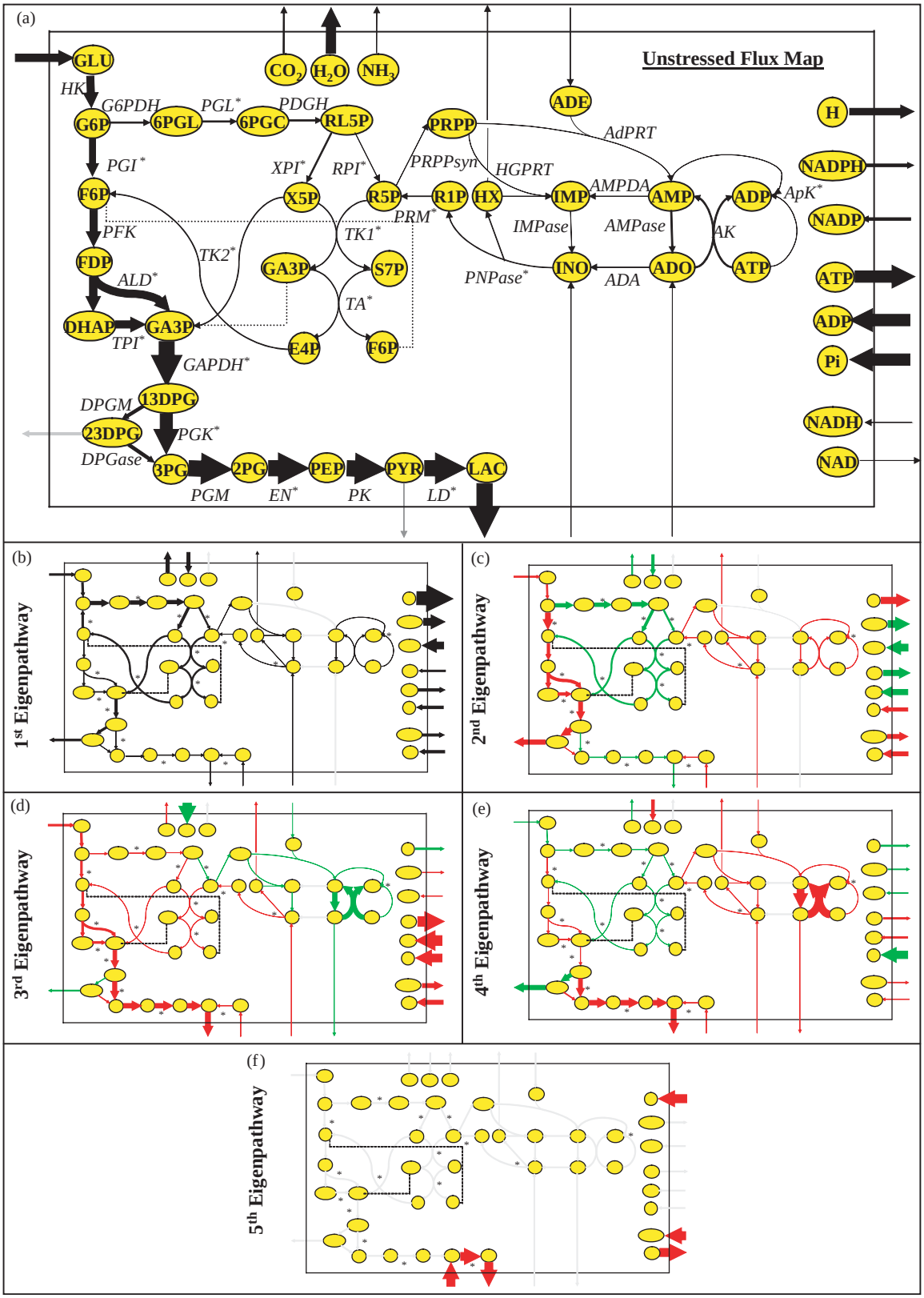

Fig. 3. The singular value decomposition of the extreme pathways of the human red blood cell metabolic network. The unstressed flux map (a) and the first–fifth eigenpathways of SVD of the human red blood cell extreme pathways (b–f) are shown. The first eigenpathway is a valid pathway (i.e. mass balance and thermodynamic constraints are satisfied). The arrow widths are proportionate to the flux values. Red (positive) and green (negative) corresponds to a change with respect to the first eigenpathway. Zero fluxes are in gray. The indicated reactions are catalysed by the enzymes whose names and abbreviations are listed in Table 2. Enzymes marked with an asterisk (*) catalyse reversible reactions. See Table 1 for metabolite abbreviations.

rather than the limiting V_{max}, more eigenpathways were needed to describe the same cumulative fractional contribution in the network (Fig. 2). With the extreme pathways normalized to unit length, the first eigenpathway had a fractional contribution of 20%, while the cumulative fractional contribution of the first five and nine eigenpathways were 56% and 75%, respectively. This difference demonstrates the importance of the V_{max} parameters in shaping the solution space and describing physiologically meaningful states. For the remainder of this paper the V_{max}-limited **P** was used.

The first five eigenpathways resulting from the SVD of **P** are shown on the metabolic maps in Fig. 3(b–f). A list of metabolites and their abbreviations as used in Fig. 3 can be found in Table 1. The associated enzymes and full reaction list that make up the network are found in Table 2. The first eigenpathway shows low flux values through the adenosine reactions, higher fluxes through the glycolytic reactions, with an exit through the Rapoport–Leubering (R/L) shunt, and the highest flux levels through the pentose phosphate pathway. Movement along one of the subsequent eigenpathways in the positive direction (the first eigenpathway can only be used in a positive direction) corresponds to increasing the fluxes shown in red and decreasing those shown in green. Since the eigenpathways are required to be orthogonal, they specifically describe the directions in the cone that are independent from each other. The subsequent eigenpathways can be interpreted biochemically as follows:

- The second eigenpathway describes the flux split between glycolysis and the pentose phosphate reactions. If the contribution of this eigenpathway were added to the first eigenpathway it would lead to decreased flux through the pentose phosphate reactions and reduced production of NADPH. The increased glycolytic flux exits through the R/L shunt leading to decreased ATP production since ATP is used in upper glycolysis and not recovered in lower glycolysis. The production of NADH increases.
- The third eigenpathway uses the glycolytic reactions from glucose to pyruvate with production of ATP and NADH. It results in a lower dissipation of ATP as a consequence of a reduced use of the coupled actions of AMPase and AK. As a result this eigenpathway has significant ATP production.
- The fourth eigenpathway describes the flux split between lower glycolysis and the R/L shunt. It further describes an increase in ATP dissipation via the AMPase-AK cycle leading to little net production of ATP, and counteracts the effects of the third eigenpathway with respect to ATP dissipation.
- The fifth eigenpathway is actually one of the extreme pathways (Wiback and Palsson, 2002). It describes the uptake of pyruvate and conversion to lactate, consuming NADH. It is thus important in balancing NADH redox metabolism.

As discussed below the first five eigenpathways account for most of the red blood cell's physiological states.

Table 1
Metabolite abbreviations used in Fig. 3

Abbr.	Metabolite
13DPG	1,3-diphosphoglycerate
23DPG	2,3-diphosphoglycerate
2PG	2-phosphoglycerate
3PG	3-phosphoglycerate
6PGC	Phosphogluconate
6PGL	Phosphogluco-lactone
ADE	Adenine
ADO	Adenosine
DHAP	Dihydroxyacetone phosphate
E4P	Erythrose-4-phosphate
F6P	Fructose-6-phosphate
FDP	Fructose-1,6-bisphosphate
G6P	Glucose-6-phosphate
GA3P	Glyceraldehyde-3-phosphate
GLU	Glucose
HX	Hypoxanthine
IMP	Inosine monophosphate
INO	Inosine
LAC	Lactate
PEP	Phosphenolpyruvate
Pi	Inorganic phosphate
PRPP	5-phosphoribosyl-1-pyrophosphate
PYR	Pyruvate
R1P	Ribose-1-phosphate
R5P	Ribose-5-phosphate
RL5P	Ribulose-5-phosphate
S7P	Sedoheptulose-7-phosphate
X5P	Xylulose-5-phosphate

4.2. *Reconstructing physiological steady states from eigenpathways*

The unstressed state of the red blood cell metabolic network (defined as the steady state flux distribution without a metabolic load) was calculated using a full kinetic model (Jamshidi et al., 2001; Joshi and Palsson, 1989), and is shown on the metabolic map (Fig. 3a). The values of the maximum sustainable loads for NADPH, ATP, and NADH were also calculated using the kinetic model as well as their corresponding flux distributions. An NADPH load simulates the red blood cell's response to oxidative free radicals. The maximum NADPH load is 2.5 mM/h, a very high value relative to glucose input at 1.5 mM/h. The ATP load simulates conditions of increased energy loads, such as when the cell is in a hyperosmotic environment. The maximum ATP load is 0.37 mM/h. A 1.7 mM/h NADH load, important for methemoglobin reduction in the human red blood cell,

Table 2
Reaction abbreviations used in Fig. 3. The corresponding reaction stoichiometries are also included

Abbr.	Enzyme name	Metabolic reaction
ADA	Adenosine deaminase	$ADO + H_2O \rightarrow INO + NH_3$
AdPRT	Adenine phosphoribosyl transferase	$PRPP + ADE + H_2O \rightarrow AMP + 2\ Pi$
AK	Adenosine kinase	$ADO + ATP \rightarrow ADP + AMP$
ALD	Aldolase	$FDP \leftrightarrow GA3P + DHAP$
AMPase	Adenosine monophosphate phosphohydrolase	$AMP + H_2O \rightarrow ADO + Pi + H$
AMPDA	Adenosine monophosphate deaminase	$AMP + H_2O \rightarrow IMP + NH_3$
ApK	Adenylate Kinase	$2\ ADP \leftrightarrow ATP + AMP$
DPGase	Diphosphoglycerate phosphatase	$23DPG + H_2O \rightarrow 3PG + Pi$
DPGM	Diphosphoglyceromutase	$13DPG \rightarrow 23DPG + H$
EN	Enolase	$2PG \leftrightarrow PEP + H_2O$
G6PDH	Glucose-6-phosphate dehydrogenase	$G6P + NADP \rightarrow 6PGL + NADPH + H$
GAPDH	Glyceraldehyde phosphate dehydrogenase	$GA3P + NAD + Pi \leftrightarrow 13DPG + NADH + H$
HGPRT	Hypoxanthine guanine phosphoryl transferase	$PRPP + HX + H_2O \rightarrow IMP + 2\ Pi$
HK	Hexokinase	$GLU + ATP \rightarrow G6P + ADP + H$
IMPase	Inosine monophosphatase	$IMP + H_2O \rightarrow INO + Pi + H$
LD	Lactate dehydrogenase	$PYR + NADH + H \leftrightarrow LAC + NAD$
PDGH	6-phosphoglyconate dehydrogenase	$6PGC + NADP \rightarrow RL5P + NADPH + CO_2$
PFK	Phosphofructokinase	$F6P + ATP \rightarrow FDP + ADP + H$
PGI	Phosphoglucoisomerase	$G6P \leftrightarrow F6P$
PGK	Phosphoglycerate kinase	$13DPG + ADP \leftrightarrow 3PG + ATP$
PGL	6-phosphoglyconolactonase	$6PGL + H_2O \leftrightarrow 6PGC + H$
PGM	Phosphoglyceromutase	$3PG \leftrightarrow 2PG$
PK	Pyruvate kinase	$PEP + ADP + H \rightarrow PYR + ATP$
PNPase	Purine nucleoside phosphorylase	$INO + Pi \leftrightarrow HX + R1P$
PRM	Phosphoribomutase	$R1P \leftrightarrow R5P$
PRPPsyn	Phosphoribosyl pyrophosphate synthetase	$R5P + ATP \rightarrow PRPP + AMP$
RPI	Ribose-5-phosphate isomerase	$RL5P \leftrightarrow R5P$
TA	Transaldolase	$GA3P + S7P \leftrightarrow E4P + F6P$
TK1	Transketolase	$X5P + R5P \leftrightarrow S7P + E4P$
TK2	Transketolase	$X5P + E4P \leftrightarrow F6P + GA3P$
TPI	Triose phosphate isomerase	$DHAP \leftrightarrow GA3P$
XPI	Xylulose-5-phosphate epimerase	$RL5P \leftrightarrow X5P$

was also applied. These stressed flux vectors represent extreme physiological states of the red blood cell, and help identify the limits of the kinetically attainable solutions within the stoichiometrically defined solution space.

These four steady state solutions were uniquely decomposed into the 23 eigenpathways (Fig. 4a). The first five eigenpathways are weighted to different degrees in reconstructing the flux distributions for the four metabolic states. In addition, "fine tuning" occurs with the addition of eigenpathways 6–11 to the reconstructed solution. Eigenpathways 12–23 are essentially insignificant in reconstructing the four solutions to the red blood cell kinetic model. The weighted addition of each subsequent eigenpathway to each reconstructed steady state solution ($\mathbf{v}_k$) reduced the relative error (Fig. 4b). After the inclusion of the first five eigenpathways, the relative error (RE(5)) of all the reconstructed kinetic solutions ranged from 0.005 to 0.018. Thus, in all four cases, the first five eigenpathways essentially reconstructed each of the individual steady state solutions (Fig. 4b). Therefore, the physiologically relevant region of the steady state solution space can be described using only five vectors and thus is only five dimensional. Consequently, there are effectively only five degrees of freedom in the problem of regulating red blood cell metabolism.

An inspection of the first five eigenpathways (Fig. 3b–f) demonstrates how together they reconstruct the unstressed steady state solution (Fig. 3a). Relative to the first eigenpathway (Fig. 3b), adding the second eigenpathway (Fig. 3c) increases the flux through the first half of glycolysis, decreases the flux through the pentose phosphate reactions, and decreases NADPH production, all of which moves the reconstructed solution significantly towards the unstressed steady state (Fig. 3a). Adding the third eigenpathway (Fig. 3d) increases the flux through all of glycolysis, particularly through lower glycolysis. The addition of the fourth eigenpathway (Fig. 3e) appropriately decreases the amount of 23DPG that is produced and instead sends that flux through lower glycolysis. Finally, the addition of the fifth eigenpathway increases the flux from pyruvate to lactate, which leads essentially to the steady state solution where lactate is the primary output of glycolysis. Thus, the significant features of the unstressed steady state can be reconstructed with only the first five eigenpathways.

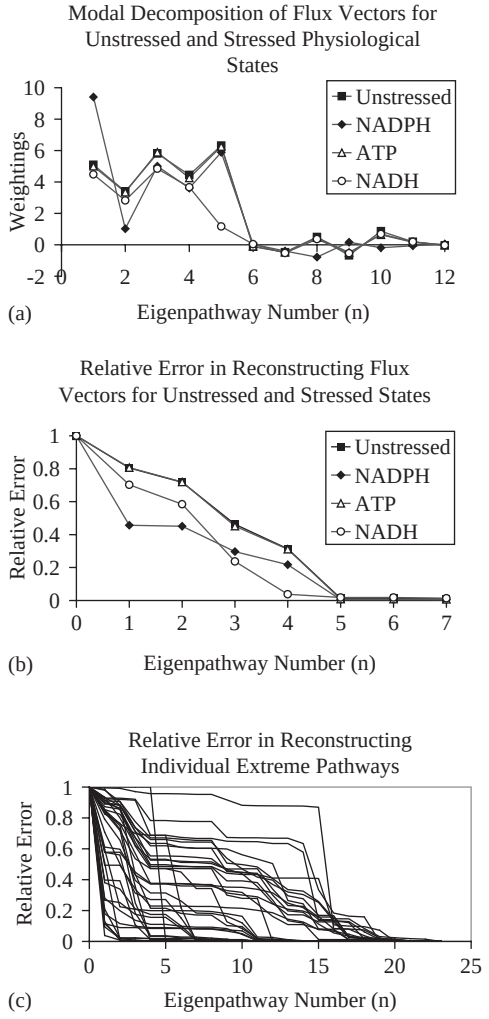

Fig. 4. Reconstructing possible steady state solutions using the eigenpathways calculated by SVD. The decompositions into the eigenpathways for the unstressed flux vector and indicated stressed flux vectors are shown (a). The relative distance from the reconstructed vector (a weighted combination of eigenpathways up to and including that eigenpathway) to the actual flux vector is shown (b). By the fifth eigenpathway the physiological flux vectors have been nearly completely reconstructed. The relative errors from the reconstruction of all 39 extreme pathways are shown (c).

The weightings for the first five eigenpathways vary when the cell is under stressed conditions. Increases in the NADPH load resulted in a substantial increase of the weighting on the first eigenpathway, thereby increasing the flux through the pentose phosphate reactions and thus elevating the production of NADPH. The weightings on the second, third, fourth, and fifth eigenpathways decrease with the application of higher NADPH loads largely because NADPH production approaches its optimum as the flux distribution approaches that of the first eigenpathway. The reduction in the weighting of the second eigenpathway, however, is the most dramatic. The application of increasing ATP loads resulted in little change in the values of the weightings on all of the first five eigenpathways. The ATP load is handled in the red blood cell by a decrease in an ATP-consuming futile cycle, with the generated ATP used instead to satisfy the load imposed upon the cell. Thus, the usage of an ATP-dissipating futile cycle in the unstressed state acts to dampen the effects of changing ATP loads, allowing the red blood cell to respond to changing ATP loads with little change in the overall flux distribution in the cell. The application of the NADH loads resulted in a significant decrease of all the eigenpathway weightings mainly because the magnitude of the steady state flux vector decreases. The weighting on the fifth eigenpathway decreased most dramatically since it consumes NADH when utilized in the positive direction and thus is needed less.

4.3. Reconstructing individual extreme pathways

Decomposition of the extreme pathway vectors into the eigenpathways shows that the most important eigenpathway of each reconstruction is often not one of the first five eigenpathways (Fig. 4c). Therefore, many regions of the allowable solution space, as defined by the extreme pathways, are poorly characterized by only the first five eigenpathways; however, all physiologically relevant states examined in this study were well captured with the first five eigenpathways. Thus, many of the extreme pathways are physiologically irrelevant and can be identified using SVD of **P**, if the approximate location of physiological solutions is known.

5. Discussion

The study of metabolic regulation has historically focused on the identification and characterization of individual regulatory events. Now that we can reconstruct full metabolic reaction networks we can address the need for regulation from a network-based perspective. This study focused on interpreting regulation from a network-based perspective using singular value decomposition of the extreme pathway matrix for human red blood cell metabolism. Each of the first five eigenpathways obtained by SVD was interpreted biochemically in the context of red blood cell metabolism.

The first five eigenpathways effectively characterized all the relevant physiological states of the red blood cell. Using the calculated eigenpathways, physiologically relevant solutions to the full red blood cell kinetic model were reconstructed. The relative error using only the first five eigenpathways was within 1.8%. Thus, the first five eigenpathways can be used to essentially recapture each of the physiologically relevant kinetic solutions. However, individual extreme pathways could not be reconstructed to such a high degree with only the first five eigenpathways. Thus, the first five

eigenpathways represented the space relevant to solutions of the full kinetic model better than they did to the space as a whole, even though they were calculated from a description of the entire solution space. This fact suggests that developing constraint-based methods that take into account kinetics and metabolomics (Palsson, 2002) will result in defining a much reduced solution space. This new solution space will be a more constrained region within the space circumscribed by the topologically based extreme pathways. The reduced solution space will allow for more physiologically relevant analyses, such as an evaluation of the effects of SNPs on red blood cell network function (Jamshidi et al., 2002).

The singular value decomposition of **P** demonstrated that the regulatory problem in red blood cell metabolism is essentially five dimensional. Hence, controlling these five eigenpathways would allow for control of the system. The question now becomes: does the human red blood cell in fact control these eigenpathways? While no evidence exists that the red blood cell precisely controls each of the eigenpathways as a whole, by examining known regulatory loops (Joshi and Palsson, 1990), a case can be made that the human red blood cell does in fact control the dominant features of these eigenpathways (reaction splits with large changes in flux levels). For example, the dominant feature of the second eigenpathway is the flux split between glycolysis and the pentose phosphate pathway and this split is known to be regulated in the red blood cell (Kirkman and Gaetani, 1986). The third eigenpathway describes the glycolytic pathway from glucose to pyruvate with production of ATP and NADH, and this pathway is known to be regulated in the red blood cell through the regulation of hexokinase (HK) and pyruvate kinase (PK) (Gerber et al., 1974; Holzhütter et al., 1985). The dominant feature of the fourth eigenpathway is the flux split between lower glycolysis and the R/L shunt. This split is known to be regulated in the red blood cell, primarily through regulation of diphosphoglycerate mutase (DPGM) (Heinrich et al., 1977). The fifth eigenpathway describes uptake of pyruvate and conversion to lactate (dissipating one NADH) and is simply controlled through mass-action kinetics of lactate and pyruvate exchange (Halestrap, 1976). Thus, the dominant features of the first five eigenpathways are indeed controlled in the red blood cell. Control of the dominant features of the first five eigenpathways, along with knowledge of the cofactor loads, is sufficient to control red blood cell metabolism.

Taken together, this study presents a network-based approach to infer a regulatory network and to define the degrees of freedom of the regulatory problem. The method used calculates the modalities needed for the metabolic network to navigate its solution space and thus could be used to infer candidate regulatory loops of metabolic systems for which the regulation is largely unknown. Further, based upon their contribution to the steady state solution space, these regulatory loops can potentially be ordered in terms of their importance to the reduction of the solution space to tightly confine the physiological states of interest. Network-based approaches to studying regulation, such as the one offered herein, complement component-based studies and provide a potential framework to better understand the interaction of regulatory components needed to achieve the regulatory demands of the cell.

Acknowledgements

The authors would like to thank Iman Famili and Dr. Radhakrishnan Mahadevan for insightful discussions, as well as Natalie Duarte for helpful feedback on this manuscript. We would like to acknowledge the support of the NIH (GM 57089) and the Whitaker Foundation (Graduate Research Fellowship to JP).

References

Arkin, A., Ross, J., McAdams, H.H., 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. Genetics 149, 1633–1648.

Beard, D.A., Liang, S.D., Qian, H., 2002. Energy balance for analysis of complex metabolic networks. Biophys. J. 83, 79–86.

Covert, M.W., Schilling, C.H., Famili, I., Edwards, J.S., Goryanin, I.I., Selkov, E., Palsson, B.O., 2001. Metabolic modeling of microbial strains in silico. Trends Biochem. Sci. 26, 179–186.

Davidson, E.H., 2001. Genomic Regulatory Systems: Development and Evolution. Academic Press, San Diego.

Edwards, J.S., Covert, M., Palsson, B., 2002. Metabolic modelling of microbes: the flux-balance approach. Environ. Microbiol. 4, 133–140.

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. Science 297, 1183–1186.

Gerber, G., Preissler, H., Heinrich, R., Rapoport, S.M., 1974. Hexokinase of human erythrocytes. Purification, kinetic model and its application to the conditions in the cell. Eur. J. Biochem. 45, 39–52.

Halestrap, A.P., 1976. Transport of pyruvate NAD lactate into human erythrocytes. Evidence for the involvement of the chloride carrier and a chloride-independent carrier. Biochem. J. 156, 193–207.

Heinrich, R., Rapaport, S.M., Rapaport, T.A., 1977. Metabolic regulation and mathematical models. Prog. Biophys. Mol. Biol. 32, 1–82.

Holzhütter, H.G., Jacobasch, G., Bisdorff, A., 1985. Mathematical modelling of metabolic pathways affected by an enzyme deficiency. A mathematical model of glycolysis in normal and pyruvate-kinase-deficient red blood cells. Eur. J. Biochem. 149, 101–111.

Jamshidi, N., Edwards, J.S., Fahland, T., Church, G.M., Palsson, B.O., 2001. Dynamic simulation of the human red blood cell metabolic network. Bioinformaics 17, 286–287.

Jamshidi, N., Wiback, S.J., Palsson, B.O., 2002. In silico model-driven assessment of the effects of single nucleotide polymorphisms (SNPs) on human red blood cell metabolism. Genome Res. 12, 1687–1692.

Joshi, A., Palsson, B.O., 1989. Metabolic dynamics in the human red cell. Part I—A comprehensive kinetic model. J. Theor. Biol. 141, 515–528.

Joshi, A., Palsson, B.O., 1990. Metabolic dynamics in the human red cell. Part III—metabolic reaction rates. J. Theor. Biol. 142, 41–68.

Kirkman, H.N., Gaetani, G.F., 1986. Regulation of glucose-6-phosphate dehydrogenase in human erythrocytes. J. Biol. Chem. 261, 4033–4038.

Lay, D.C., 1997. Linear Algebra and Its Applications, 2nd Edition. Addison Wesley Longman Inc, Reading, MA.

McAdams, H.H., Arkin, A., 1997. Stochastic mechanisms in gene expression. Proc. Natl Acad. Sci. 94, 814–819.

Palsson, B.O., 2000. The challenges of in silico biology. Nat. Biotechnol. 18, 1147–1150.

Palsson, B.O., 2002. In silico biology through "omics". Nat. Biotechnol. 20, 649–650.

Papin, J.A., Price, N.D., Palsson, B.O., 2002. Extreme pathway lengths and reaction participation in genome-scale metabolic networks. Genome Res. 12, 1889–1900.

Papin, J.A., Price, N.D., Wiback, S.J., Fell, D., Palsson, B., 2003. Metabolic pathways in the post-genome era. Trends Biochem. Sci. 28, 250–258.

Price, N.D., Famili, I., Beard, D.A., Palsson, B.O., 2002. Extreme pathways and Kirchhoff's second law. Biophys. J. 83, 2879–2882.

Price, N.D., Papin, J.A., Schilling, C.H., Palsson, B., 2003a. Genome-scale microbial in silico models: the constraints-based approach. Trends Biotechnol. 21, 162–169.

Price, N.D., Reed, J.L., Papin, J.A., Famili, I., Palsson, B.O., 2003b. Analysis of metabolic capabilities using singular value decomposition of extreme pathway matrices. Biophys. J. 84, 794–804.

Ptashne, M., Gann, A., 2002. Genes and Signals. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

Schilling, C.H., Letscher, D., Palsson, B.O., 2000. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. J. Theor. Biol. 203, 229–248.

Schuster, S., Fell, D.A., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nature Biotechnol. 18, 326–332.

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E.D., 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420, 190–193.

Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T.S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J.C., Hutchison, C.A.I., 1999. E-CELL: software environment for whole-cell simulation. Bioinformatics 15, 72–84.

Werner, A., Heinrich, R., 1985. A kinetic model for the interaction of energy metabolism and osmotic states of human erythrocytes. Analysis of the stationary "in vivo" state and of time dependent variations under blood preservation conditions. Biomed. Biochim. Acta 44, 185–212.

Wiback, S.J., Palsson, B.O., 2002. Extreme pathway analysis of human red blood cell metabolism. Biophys. J. 83, 808–818.