Thermodynamic Constraints for Biochemical Networks

Daniel A. Beard[1], Eric Babson[2], Edward Curtis[2], and Hong Qian[1,3]

[1]Department of Bioengineering
[2]Department of Mathematics
[3]Department of Applied Mathematics,
University of Washington, Seattle, WA 98195

June 7, 2008

ABSTRACT

The constraint-based approach to analysis of biochemical systems has emerged as a useful tool for rational metabolic engineering. Flux balance analysis (FBA) is based on the constraint of mass conservation; energy balance analysis (EBA) is based on nonequilibrium thermodynamics. The power of these approaches lies in the fact that the constraints are based on physical laws, and do not make use of unknown parameters. Here, we show that the network structure (i.e., the stoichiometric matrix) alone provides a system of constraints on the fluxes in a biochemical network which are feasible according to both mass balance and the laws of thermodynamics. A realistic example shows that these constraints can be even sufficient for deriving unambiguous, biologically meaningful results. The thermodynamic constraints are obtained by comparing of the sign pattern of the flux vector to the sign patterns of the cycles of the internal cycle space via a strong connection between stoichiometric network theory (SNT) and the mathematical theory of oriented matroids.

1. INTRODUCTION

Theoretical tools for mathematical analysis of biochemical networks traces back to the work done by chemical kineticists in 1970's and 1980's [Clarke, 1980, 1988]. For a given system in the steady state, mass balance restricts the feasible chemical fluxes to the null space of the stoichiometric matrix **S**, [Clarke, 1980, 1988]. Various mathematical approaches based on generating the vectors of the null space of **S** (with certain modifications called the internal representation [Clarke, 1980], elementary modes [Schuster and Hilgetag, 1994], and extreme pathways [Schilling et al. 2000]) have been developed. Mavrovouniotis et al. [1990, 1992a, 1992b, 1992c] have developed a set of algorithms for computing sets of feasible flux pathways for given biochemical objectives. These

approaches to mass balance-based analysis are reviewed elsewhere (see, for example Schilling et al. [1999] and references therein.) Clarke [1980, 1988] combines the null space analysis with the analysis of dynamic stability, representing a viable approach for extending flux balance analysis (FBA) beyond steady-state applications.

In addition to the mass balance constraint, the network structure imposes thermodynamic constraints on the feasible set of chemical fluxes [Beard et al. 2002; Qian et al. 2003]. Katchalsky and colleagues were among the first to raise awareness of network thermodynamics in biophysics [Katzir-Katchalsky and Curran, 1965; Oster et al. 1971]; Prigogine and his colleagues have promoted the cause of nonequilibrium thermodynamics for decades [Nicolis and Prigogine, 1977]. However, none of the classic works on metabolic pathway analysis and constraint-based analysis has explicitly treated nonequilibrium network thermodynamics.

The "data" required for the flux balance analysis are stored in the stoichiometric matrix **S** [Clarke, 1980, 1988; Schuster and Hilgetag, 1994; Schilling et al. 2000, 2001]. The matrix **S** also provides the data needed to determine the thermodynamic constraints [Beard et al. 2002; Qian et al. 2003]. In this paper, we make use of an oriented matroid derived from the nullspace of the internal reaction matrix $\tilde{\mathbf{S}}$. We show that the thermodynamic constraints can be expressed in terms of the "cycles" of the matroid. The main theoretical result of this paper is stated as Theorem 6.1 which gives precise conditions under which a flux vector obeys both the mass balance and the thermodynamic constraints. An application to a realistic metabolic network leading to meaningful biochemical "predictions" is presented.

2. Problem Statement

To demonstrate the need for the thermodynamic constraints, we analyze the network with 28 reactions illustrated in Figure 1, representing energetic metabolism (glycolysis, TCA cycle, oxidative phosphorylation, and ATP consumption). Let **S** be the stoichiometric matrix for this reaction network. If the system is in the steady state, the fluxes satisfy the mass balance equation:

$$\mathbf{S} \cdot J = 0 \tag{2.1}$$

where J is the vector listing the 31 fluxes (28 reactions plus 3 boundary fluxes). Consider a system containing ATP, ADP, Pi, pyruvate, NAD, NADH, and various other metabolites; given an input of one unit of glucose and six units of oxygen, what is the maximal ATP production? In other words, for each glucose molecule consumed, what is the maximal number of times that the internal ATPase reaction can turn over? Flux balance analysis reveals that the production of ATP is unconstrained by steady-state mass balance, because the ATPase reaction participates in a number of internal reaction cycles. For example, consider the sub-system consisting of six species, and four reactions:

$$\begin{aligned} \mathbf{pps}(rev.) &: PEP + AMP + Pi \rightarrow PYR + ATP \\ \mathbf{pyk}(rev.) &: PYR + ATP \rightarrow PEP + ADP \\ \mathbf{ak}(rev.) &: 2\,ADP \rightarrow ATP + AMP \\ \mathbf{ATPase} &: ATP \rightarrow ADP + Pi \end{aligned}$$

These reactions are encoded in a stoichiometric matrix $\mathbf{S_1}$ as follows.

$$\mathbf{S_1} = \begin{array}{c} PEP \\ AMP \\ Pi \\ PYR \\ ATP \\ ADP \end{array} \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 1 \end{bmatrix} \tag{2.2}$$

The null space of $\mathbf{S}_1$ has dimension one, consisting of multiples of the vector $J = [1, 1, 1, 1]^T$. With no further restraint on the system, this cycle could maintain mass balance, while proceeding with arbitrarily large flux.

This unphysical behavior can be avoided in FBA by applying irreversibility constraints to selected fluxes. For example, if either phosphoenolpyruvate synthase (**pps**) or pyruvate kinase (**pyk**) were constrained so that the reverse fluxes cannot occur, then this loop would be disqualified. However, as all reactions are in principle reversible (by the Haldane relationship), a physical theory that allows us to exclude infeasible cycles without bringing in a set of *ad hoc* irreversibility constraints is preferable. Clearly, the

above cycle is not thermodynamically feasible. The problem this paper addresses is how to exclude thermodynamically infeasible cycles.

One way to exclude infeasible cycles is the generalized thermodynamic theory of energy balance analysis (EBA) for stoichiometric matrices, which was developed by Beard et al. [2002] and Qian et al. [2003]. In a related approach, introduced by Mavrovouniotis [1996], knowledge of metabolite concentrations is translated into bounds on reaction potentials, which in turn constrains flux directions. In contrast, the EBA approach requires no prior knowledge of concentrations; EBA is based on the constraint that for a flux vector to be feasible it must be feasible that a potential can exist. The current paper shows that there is a way of imposing the thermodynamic constraints on the fluxes that does not require explicit mention of chemical potential variables. These thermodynamic constraints are obtained by comparing of the sign pattern of the flux vector to the sign patterns of the cycles of the internal cycle space and relies on a strong connection between biochemical networks and oriented matroids.

3. Network Analysis

The central concept introduced by Beard et al. [2002] will be presented first by the very simple example of the network illustrated in Figure 2A. This diagram represents five internal reactions and two external fluxes. The internal reactions are:

$$\begin{aligned} r_1: &\quad A \rightarrow B \\ r_2: &\quad B \rightarrow C \\ r_3: &\quad C \rightarrow A \\ r_4: &\quad C \rightarrow D \\ r_5: &\quad D \rightarrow B \end{aligned}$$

The external fluxes r_6 and r_7 transport A and B into or out of the network. All reactions are considered reversible; the arrows indicate the direction defined as positive. The

stoichiometric matrix for this network is:

$$\mathbf{S} = \begin{matrix} A \\ B \\ C \\ D \end{matrix} \begin{bmatrix} -1 & 0 & +1 & 0 & 0 & +1 & 0 \\ +1 & -1 & 0 & 0 & +1 & 0 & -1 \\ 0 & +1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 & -1 & 0 & 0 \end{bmatrix} \tag{3.1}$$

Suppose J is a vector of fluxes for a network in the steady state. The requirement that $\mathbf{S} \cdot J = 0$ is equivalent to Kirchhoff's current law.

The internal reactions correspond to a matrix $\tilde{\mathbf{S}}$, which is the sub-matrix obtained from **S** by omitting the last two columns (corresponding to external fluxes).

$$\tilde{\mathbf{S}} = \begin{bmatrix} -1 & 0 & +1 & 0 & 0 \\ +1 & -1 & 0 & 0 & +1 \\ 0 & +1 & -1 & -1 & 0 \\ 0 & 0 & 0 & +1 & -1 \end{bmatrix} \tag{3.2}$$

The internal-reaction space is defined to be the null-space of $\tilde{\mathbf{S}}$, and is denoted $\mathcal{N}$. The number of internal reactions is 5, the matrix $\tilde{\mathbf{S}}$ has rank 3, so $\mathcal{N}$ has dimension $5 - 3 = 2$. A basis for $\mathcal{N}$ may be computed from the reduced row echelon factorization of $\tilde{\mathbf{S}}$, [Strang, 1988]. The resulting vectors are the columns of a 2×5 matrix $\mathbf{N_I}$, where

$$\mathbf{N_I} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{3.3}$$

The columns of $\mathbf{N_I}$ form a basis for flux vectors that satisfy mass balance with no transport into or out of the system.

Let $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]$ be a vector of chemical potentials, associated with the species A, B, C and D respectively. The vector $\Delta\mu$ of potential differences is defined by:

$$\Delta\mu = \mu \cdot \mathbf{S} \tag{3.4}$$

Specifically, for each $1 \leq j \leq 5$,

$$\Delta\mu_j = \sum_i \mu_i \cdot \mathbf{S_{ij}} \tag{3.5}$$

$\Delta\mu_j$ is called the (chemical) potential difference associated with the (internal) reaction r_j. Equation (3.4) says that the vector $\Delta\mu$ is a linear combination of the rows of $\tilde{\mathbf{S}}$, that

is $\Delta\mu$ is in the rowspace of $\tilde{\mathbf{S}}$. The statement that each vector in the rowspace of $\tilde{\mathbf{S}}$ is orthogonal to each vector in the nullspace of $\tilde{\mathbf{S}}$ says that

$$\Delta\mu \cdot \mathbf{N_I} = 0 \tag{3.6}$$

This equations is equivalent to Kirchhoff's voltage law. If $\Delta\mu_j \neq 0$, the ratio $-\dfrac{J_j}{\Delta\mu_j}$ is called the *conductance* of the jth reaction.

The flux vectors which satisfy mass balance, with possibly non-zero exchange flux, are the vectors J in the kernel of $\mathbf{S}$. The 4×7 matrix $\mathbf{S}$ has rank 4, so its nullspace has rank 3. A basis for the subspace of fluxes that satisfy mass balance consists of the columns of $\mathbf{N_S}$, where

$$\mathbf{N_S} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{3.7}$$

The columns of $\mathbf{N_S}$ will be called P_1, P_2, P_3. Mass balanced steady state flux vectors have the form

$$J = a_1 P_1 + a_2 P_2 + a_3 P_3 \tag{3.8}$$

with $a_1 \neq 0$.

We proceed to examine the constraint imposed on these fluxes by the Second Law of Thermodynamics, which states that each internal reaction with non-zero flux must dissipate energy. In Beard et al. [2002], this thermodynamic constraint is expressed follows. For a flux vector J to be thermodynamically feasible, there must exist a vector μ, with vector $\Delta\mu$ defined by $\Delta\mu = \mu \cdot \mathbf{S}$, for which the the following inequality is satisfied. For each index $j = 1, 2, \ldots, 5$,

$$\Delta\mu_j \cdot J_j \leq 0 \tag{3.9}$$

Furthermore, $J_j = 0$ if and only if $\Delta\mu_j = 0$. This restriction on J says that the sign pattern (+, 0, or -) in $\tilde{J}$ is the same as the sign pattern in $-\Delta\mu$.

These thermodynamic constraints can be conveniently described by means of an oriented matroid, a mathematical object intimately related to the biochemical cycles and pathways (Oliviera et al. 2001). The oriented matroids of importance for this problem are obtained from the rowspace and the nullspace of $\tilde{\mathbf{S}}$. We include here a brief description of oriented matroids.

4. ORIENTED MATROIDS

A sign pattern α is an n-tuple, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, each of whose entries is taken from the set $\{+, -, 0\}$. For each sign pattern α, α^+ stands for the set of indices for which $\alpha_i = +$, and α^- stands for the set of indices for which $\alpha_i = -$. The signed support of a sign pattern α is the pair (α^+, α^-). If α and β are two sign patterns, we say that $\alpha \subset \beta$ if $\alpha^+ \subset \beta^+$ and $\alpha^- \subset \beta^-$.

Definition 4.1. Two sign patterns α and β are said to be *orthogonal*, if either

(1) The support of α and the support of β have no indices in common, or

(2) There is an index i for which α and β have the same sign, and there is another index j for which α and β have opposite signs.

If two sign patterns α and β are orthogonal, we write $\alpha \perp \beta$.

Euclidean space of n dimensions is denoted $\mathbb{R}^n$. For each vector $v \in \mathbb{R}^n$, let sgn(v) be the sign pattern whose entries are the signs ($+$, 0, or $-$) of the components of v. If follows immediately from 4.1 that two sign patterns α and β are orthogonal if and only if there is a pair of vectors v and w in $\mathbb{R}^n$, with α =sgn(v), β =sgn(w), and v is orthogonal to w in $\mathbb{R}^n$.

If ξ is a subspace $\mathbb{R}^n$, a matroid $\mathcal{M}$ is defined as follows.

(1) A "vector" of $\mathcal{M}$ is a sign pattern of the form sgn(v) for some non-zero vector $v \in \xi$. The set of all vectors of $\mathcal{M}$ is denoted $\mathcal{V}$.

(2) A "cycle" of $\mathcal{M}$ is a member of $\mathcal{V}$ which has minimal support. This means that α is a cycle of $\mathcal{M}$ if $\alpha \in \mathcal{V}$ and there is no $\beta \in \mathcal{V}$ such that the signed support of (β) is a proper subset of the signed support of (α). The set of cycles of $\mathcal{M}$ is denoted $\mathcal{C}$.

The set of cycles of $\mathcal{M}$ has the following three properties, called the circuit axioms for an oriented matroid (see Bjorner et al. [1993], p103; or Ziegler [1994] p162).

(C1) If $\alpha \in \mathcal{C}$ then its negative $-\alpha \in \mathcal{C}$ also.

(C2) If $\alpha, \beta \in \mathcal{C}$, with signed support $(\alpha) \subset$ signed support(β), then $\alpha = \beta$ or $\alpha = -\beta$.

(C3) Suppose $\alpha, \beta \in \mathcal{C}$, with $\alpha \neq -\beta$, and i is an index with $\alpha_i = +$ and $\beta_i = -$. Then there is $\gamma \in \mathcal{C}$, with $\gamma^+ \subset (\alpha^+ \cup \beta^+)$, and $\gamma^- \subset (\alpha^- \cup \beta^-)$, and $\gamma_i = 0$.

These conditions are used to calculate $\mathcal{C}$ from a basis for ξ.

The orthogonal complement of the subspace ξ in $\mathbb{R}^n$ is denoted $\xi^\perp$. The dual matroid $\mathcal{M}^*$ is obtained from $\xi^\perp$ in exactly the same way that $\mathcal{M}$ is obtained from ξ. The "vectors" of $\mathcal{M}^*$ are sometimes called "co-vectors" of $\mathcal{M}$, and are denoted $\mathcal{V}^*$. The "cycles" of $\mathcal{M}^*$ are sometimes called "co-cycles" of $\mathcal{M}$, and are denoted $\mathcal{C}^*$.

The fundamental property which relates the oriented matroid $\mathcal{M}$ to its dual $\mathcal{M}^*$ is the following. (See Bjorner et al. [1993], p 146; or Ziegler [1994], p161.)

Proposition 4.2. *A sign pattern α is in $\mathcal{V}$ if and only if α is orthogonal to every cycle $\gamma \in \mathcal{C}^*$.*

5. Thermodynamic Constraints

We consider a biochemical network that is in the steady-state, with stoichiometric matrix $\mathbf{S}$. Suppose there are n internal reactions, corresponding to the first n columns of $\mathbf{S}$. The remaining columns correspond to external fluxes. Let $\tilde{\mathbf{S}}$ be the submatrix of $\mathbf{S}$ consisting of the first n columns of $\mathbf{S}$.

Suppose J is a vector of fluxes that satisfies mass balance (2.1); that is, J is in the kernel of $\mathbf{S}$. Let $\tilde{J}$ denote the vector in $\mathbb{R}^n$ obtained by projecting J onto the first n components, i.e., the components corresponding to the internal reactions. The thermodynamic constraint (3.9) on J asserts that there is a vector $\Delta\mu$ of the form $\Delta\mu = \mu \cdot \tilde{\mathbf{S}}$, for some vector μ of chemical potentials, such that $\tilde{J}$ has the same sign pattern $-\Delta\mu$. This means that $\Delta\mu$ is in the row-space of $\tilde{\mathbf{S}}$. Let ξ be the row-space of $\tilde{\mathbf{S}}$, and let $\mathcal{M}$ be

the matroid obtained from ξ as described in Section 4. The thermodynamic constraint on J is that the sign pattern $\text{sgn}(\tilde{J})$ must be a member of $\mathcal{V}$, the set of "vectors" of the matroid $\mathcal{M}$.

This condition on $\text{sgn}(\tilde{J})$ can be expressed in terms of the cycles of the dual matroid $\mathcal{M}^*$ as follows. Recall that $\mathcal{M}^*$ is obtained from $\xi^\perp$, where $\xi^\perp$ is the null-space of $\tilde{\mathbf{S}}$. This means that a flux vector J is thermodynamically feasible if and only if

$$\text{sgn}(\tilde{J}) \perp \gamma \tag{5.1}$$

for each cycle in $\mathcal{C}^*$. If the condition on $\tilde{J}$ is relaxed so that J_j is allowed to be 0, even if $\Delta\mu_j \neq 0$, then condition 5.1 takes a slightly different form. To state this condition, $\text{sgn}(\tilde{J})$ and each of the cycles γ in $\mathcal{C}^*$. will be considered to be vectors in $\mathbb{R}^n$ with components 1, 0 or -1, instead of $+$, 0 or $-$. The ordinary inner product in $\mathbb{R}^n$ is denoted $< \ . \ , \ . >$. The (relaxed) condition that J be thermodynamically feasible says that, for each cycle $\gamma \in \mathcal{C}^*$,

$$| < \gamma, \text{sgn}(\tilde{J}) > | \ < \ | < \gamma, \gamma > | \tag{5.2}$$

Continuing the analysis of the example of Figure 2A, it is possible to describe explicitly all the flux vectors J that are biochemically feasible. In order for a vector J to satisfy mass balance it must be of the form:

$$J = a_1 P_1 + a_2 P_2 + a_3 P_3$$

where $\{P_1, P_2, P_3\}$ are the vectors in a basis for $\mathcal{N}$, given as the columns of the matrix $\mathbf{N_S}$. Projecting J on the first 5 components gives

$$\tilde{J} = \begin{bmatrix} a_1 & + & a_2 \\ a_2 & + & a_3 \\ a_2 & & \\ a_3 & & \\ a_3 & & \end{bmatrix} \tag{5.3}$$

A short computation using property (C3), shows that the set of cycles $\mathcal{C}^*$ are the sign patterns of the columns (and their negatives) of the matrix C, where

$$C = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \tag{5.4}$$

The three columns of C will be called C_1, C_2, and C_3 respectively, and the corresponding sign patterns will be called γ_1, γ_2, and γ_3. The columns of C are similar to the elementary modes introduced by Schuster and colleagues [Schuster and Hilgetag, 1994; Schuster et al. 1999; Stelling et al. 2002] in that the elementary modes satisfy a property of minimal support similar to the cycles of $\mathcal{C}^*$.

In order that the sign pattern sgn($\tilde{J}$) be orthogonal to γ_2 (according to Definition 4.1), a_2 and a_3 must have opposite signs, with $|a_2| > |a_3|$. In order that sgn($\tilde{J}$) be orthogonal to γ_1, a_1 and a_2 must have opposite signs, with $|a_1| > |a_2|$; this is sufficient to make sgn($\tilde{J}$) be orthogonal to γ_3 as well. If we take a_3 positive , then a_2 is negative and a_1 is positive, with $a_1 > -a_2 > a_3$. This describes all feasible vectors J which have non-zero flux for each reaction. For example

$$J = 3P_1 - 2P_2 + P_3 = [1, -1, -2, 1, 1, 3, 3]^T$$

is a feasible vector. This argument also shows that, for this example, if J is a feasible vector with each $J_j \neq 0$, then the sign pattern of J must be sgn(J) = $\{+, -, -, +, +, +, +\}$ or its negative. For these fluxes, every feasible chemical potential difference function must have sign pattern $\{-, +, +, -, -\}$ or its negative. One of the (many) such functions corresponding to the feasible vector J is $\Delta\mu = [-3, 2, 1, -1, -1]$, coming from the (chemical) potential function $\mu = [-2, 1, -1, 0]$.

Under the relaxed condition (5.2), the flux vector

$$J = 2P_1 - P_2 + P_3 = [1, 0, -1, 1, 1, 2, 2]^T$$

is feasible. In this situation, it may be appropriate to eliminate the equation r_2, and consider the network with one less reaction.

On the other hand, the flux vector

$$K = 3P_1 - P_2 + 2P_3 = [2, -1, 1, 2, 2, 3, 3]^T$$

satisfies mass balance, but is not thermodynamically feasible, because K is not orthogonal to the cycle γ_3, nor does it satisfy (5.2). K violates the thermodynamic constraint, even though K is a linear combination of the feasible flux vectors.

6. CONCLUSIONS

The previous section has established the following. Suppose $\mathbf{S}$ is the stoichiometric matrix for a biochemical network, and $\tilde{\mathbf{S}}$ is the submatrix corresponding to the internal reactions. Let $\mathcal{M}^*$ be the matroid derived from the nullspace of $\tilde{\mathbf{S}}$, and $\mathcal{C}^*$ be the set of cycles of $\mathcal{M}^*$ as in Section 4.

Theorem 6.1. *A flux vector J is feasible according to steady state mass balance and the thermodynamic constraint if and only if*

(1) $\mathbf{S} \cdot J = 0$ *and*

(2) $\operatorname{sgn}(\tilde{J})$ *is orthogonal to γ for each $\gamma \in \mathcal{C}^*$.*

If J_j is allowed to be 0, even if $\Delta\mu_j \neq 0$, then condition (2) is replaced by (2'):

(2') $|<\gamma, \operatorname{sgn}(\tilde{J})>| \;<\; |<\gamma, \gamma>|$ *for each $\gamma \in \mathcal{C}^*$.*

If a flux vector J satisfies conditions (1) and (2), we will say that J is *strictly feasible*, while if it only satisfies (1) and the relaxed condition (2'), we will say that J is T-feasible, or simply feasible. The set of T-feasible vectors includes the strictly feasible vectors and the vectors on the boundary of the strictly feasible set.

The conditions of Theorem (6.1) serve as a useful thermodynamic constraints because they are violated if an only if a flux vector contains a thermodynamically infeasible cycle. In fact, the second condition (2 or 2') is not limited to steady state fluxes. The thermodynamic feasibility condition applies whether or not the system is in a steady state.

In the example of Figure 2, the flux vectors $J_1 = [1, 0, 0, 0, 0, 1, 1]$ and $J_2 = [0, 1, 1, 0, 0, -1, -1]$ are both T-feasible (but not strictly feasible). The sum $J_3 = [1, 1, 1, 0, 0, 0, 0]$ is not T-feasible. Thus flux vectors constructed from linear combinations of feasible fluxes are not guaranteed to be feasible. Therefore a solution constructed from sums of elementary modes that are known to be feasible is not guaranteed to satisfy the thermodynamic constraints. Some form of thermodynamic constraint must be imposed to guarantee that the fluxes are physically realizable. The advantage of the approach presented here is that these constraints are easy to implement and do not explicitly use chemical potential variables. The disadvantage is that the complete internal cycle space, with all possible sign patterns, must be computed, which can be challenging for large-scale systems.

The method that we have used for generating the cycles involves exhaustively computing linear combinations of null space vectors and compiling the complete set of cycles as defined in Section 4. Since the computational complexity of this search scales exponentially, developing efficient algorithms will be an important goal of future research.

Returning to the original problem of ATP production in energy metabolism (Figure 1), and searching for the flux vector that maximizes ATP production while satisfying the mass balance constraint (Equation (2.1)) and the thermodynamic constraint (Equation (5.1)), we find that at most 20.5 ATP are produced for each glucose molecule consumed. Thus, by applying thermodynamic constraints, there is no need to introduce an *ad hoc* set of irreversibility constraints to obtain biochemically reasonable results in constraints-based analysis of biochemical networks. In fact, even when arbitrary irreversibility constraints are applied, the thermodynamic feasibility of nonequilibrium fluxes can be evaluated only by explicitly considering the network thermodynamics. In this paper, we have demonstrated that the thermodynamic constraints, which arise from the sign patterns of the internal cycle space, can be obtained directly from the stoichiometric matrix $\mathbf{S}$.

REFERENCES

[1] Beard, D. A., Liang, S.-D. & Qian, H. (2002) Energy balance for analysis of complex metabolic networks. Biophys. J. 83:79-83.

[2] Bjorner, A., Las Vergnas, M., Sturmfels, B., White, N., & Ziegler, G. (1993) *Oriented Matroids*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Vol 46.

[3] Clarke, B. L. (1980) Stability of complex reaction networks. in *Adv. Chem. Phys.* edited by I. Prigogine and S.A. Rice. John Wiley, New York. pp. 1-215.

[4] Clarke, B. L. (1988) Stoichiometric network analysis. Cell Biophys. Vol 12:237-253.

[5] Katzir-Katchalsky, A., & Curran, P. F. (1965) Nonequilibrium Thermodynamics in Biophysics. Harvard Univ. Press, Cambridge.

[6] Mavrovouniotis, M. L., Stephanopoulos, G., & Stephanopoulos, G. (1990) Computer-aided synthesis of biochemical pathways. Biotechnology and Bioengineering, 36:119-1132.

[7] Mavrovouniotis, M. L., Stephanopoulos, G., & Stephanopoulos, G. (1992a) Synthesis of biochemical production routes. Computers Chem Eng, 16:605-619.

[8] Mavrovouniotis, M. L., & Stephanopoulos, G. (1992b) Synthesis of reaction mechanisms consisting of reversible and irreversible steps 1. A synthesis approach in the context of simple examples. Ind. Eng. Chem. Res. 31:1625-1637.

[9] Mavrovouniotis, M. L. (1992c) Synthesis of reaction mechanisms consisting of reversible and irreversible steps 2. Formalization and Analysis of the synthesis algorithm. Ind. Eng. Chem. Res. 31:1637-1653.

[10] Mavrovouniotis, M. L. (1996) Duality Theory for Thermodynamic Bottlenecks in Bioreaction Pathways. Chemical Engineering Science, 5:1495-1996.

[11] Nicolis, G., & Prigogine, I. (1977) Self-Organization in Nonequilibrium Systems. Wiley Interscience, New York.

[12] Oliviera, J.S., Bailey, C.G., Jones-Oliviera, J.B. & Dixon, D.A. (2001) An algebraic-combinatorial model for the identification and mapping of biochemical pathways, Bull. Math. Biol. 63:1163-1196.

[13] Oster, G., Perelson, A. & Katchalsky, A. (1971) Network thermodynamics. Nature 234:393-399.

[14] Qian, H., Beard, D. A., & Liang, S.-D. (2003) Stoichiometric network theory for nonequilibrium biochemical systems. Eur. J. Biochem. 270:415-421.

[15] Schilling, C. H., Schuster, S., Palsson, B. O., & Heinrich, R. (1999) Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic era, Biotechnol. Prog. 16:296-303.

[16] 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. 203:229-48.

[17] Schilling, C. H., Edwards, J. S., Letscher, D., & Palsson, B. O. (2001) Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. Biotechnol. Bioeng. 71:286-306.

[18] Schuster, S., & Hilgetag, C. (1994) On elementary flux modes in biochemical reaction systems at steady state. J. Biol. Syst. 2:165-182.

[19] Schuster, S., Dandekar, T., & Fell, D. A. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol. 17:53-60.

[20] Strang, G. (1988) Linear Algebra and its Applications, 3rd Edition. Harcourt Brace Jovanovich, San Diego.

[21] 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.

[22] Ziegler, G. M. (1994) *Lectures on Polytopes*, Graduate Texts in Mathematics No 152, Springer, New York, N.Y.

FIGURE LEGENDS

Figure 1: Energy metabolism network. These reactions are a subset of the Escherichia coli K-12 metabolic network compiled by Palsson et al., and available on the world-wide web at http://gcrg.ucsd.edu/organisms/ecoli.html.

Figure 2: A. Example network structure, with five internal reactions (labeled 1-5) and two transport, or boundary, fluxes (labeled 6 and 7). B. The complete set of distinct internal cycles, pathways for the network in A. C. The complete set of throughput pathways for the network in A.

FIGURE 1

FIGURE 2