In ***Encyclopedia of Genetics, Genomics, Proteomics, and Bioinformatics***

Volume 3: Proteomics, M. Dunn, Editor
Section 8: Systems Biology, R.L. Winslow, Editor

Constraint-Based Modeling of Metabolomic Systems

Daniel A. Beard and Hong Qian

John Wiley & Sons (2005)

http://www.wiley.com/legacy/wileychi/ggpb/id1-proteomics.htm

Specialist Review

99. Constraint-based modeling of metabolomic systems

Daniel A. Beard
Medical College of Wisconsin, Milwaukee, WI, US

Hong Qian
University of Washington, Seattle, WA, US

1. Introduction

Since the time of Newton, a central dogma of scientific thinking has been that physical systems are best understood by representation in terms of the smallest possible subsystem (i.e., model) that captures the important mechanistic interactions. The influence of gravity in maintaining the earth's orbit about the sun is satisfactorily explained by analyzing the equations of motion representing a universe consisting of two massive bodies. Still, a complete mathematical analysis of the three-body problem remains out of reach, and most quantitative analyses of practical engineering problems, such as the analysis of the material properties of metal composites and the building of aircraft, are based on a combination of mechanistic models with intermediate-level empirical descriptions in quantitative terms. This approach is typical of engineering solutions to complex problems in physical sciences.

Living biological systems consist of not two, or even two hundred interacting components. Analysis, prediction, and rational manipulation of cellular function requires a mechanistic understanding of physical systems of unimaginable complexity. These are truly *complex systems*. Furthermore, each biological system is geared toward a unique function (or functions), a concept usually lacking in the physical sciences but central to engineering (Hopfield, 1994).

Two other disciplines that have tackled complexities similar to those encountered in Systems Biology are economics and ecology. It is arguable that mathematical thinking, language, and modeling are well-accepted components of these two fields. In addition, in both fields there is a clear distinction between the so-called micro- and macroscopic views. The microscopic view focuses on the small-scale mechanistic components of a system, while the macroscopic view treats the systems as a whole. As in biology, both economics and ecology have yet to establish a quantitative framework connecting the two viewpoints.

While it may seem natural to follow the path paved by Newton in constructing a differential equation–based representation for biological systems, we should not treat uncritically the assumption that a useful representation of this form always exists. Popular mythology aside, Newton stood on no one's shoulders. Similarly, we do not require Systems Biology to stand on his alone.

Clearly, certain aspects of cellular physiology are successfully quantified via a traditional kinetic differential equation–based theory. A classical example is the electrophysiology of excitable cells, a field pioneered by Hodgkin and Huxley (1952). On the basis of their foundational description, models of ever-increasing complexity have been introduced, leading toward more complete computational models of cellular electrophysiology (Noble, 2004). However, it is not immediately clear that differential equations will provide a universal framework for integrating these models with cellular mechanics, metabolism, genetics, and other phenomena. It seems more likely that no such universal mathematical framework will exist for Systems Biology.

In the past several decades, progress in cell biology has been driven mainly by the philosophies of and methodology from biochemistry and molecular genetics. Biochemistry emphasizes individual reactions and their chemical nature: what are the reactants and products; which enzyme is involved; and what is the protein structure and enzyme mechanism? Molecular genetics, on the other hand, focuses on the function and/or malfunction of molecular components and on information processing in living cells. The traditional modeling approach in biochemistry is differential equation–based enzyme kinetics. Modeling studies of one or a few enzymatic reactions at a time have been successfully applied to in vitro kinetics. It remains to be demonstrated, however, that this approach can be scaled up to an in vivo system of hundreds of reactions and species with thousands of parameters (Teusink● *et al*., 2003). More importantly, it is not clear that this is the ideal approach for integrating biochemistry with molecular genetics.

●[AQ2]

The constraint-based modeling (CBM) approach circumvents several of these difficulties. It facilitates integration of experimental data of disparate types and from disparate sources while increasing the accuracy in its prediction. It does not require a priori knowledge (or assumptions) regarding all of the mechanisms and parameters for a given system. However, when a priori knowledge exists, such as data on enzyme kinetics or measurements of in situ concentrations and fluxes, this knowledge can be introduced in the form of constraints, on equal footing with molecular genetic observations on the topology and information flow in a biochemical network. While the differential equation–based modeling provides predictions with "high information content and high false rate", the CBM approach, by design, provides a low level of false prediction, but can often give little information in its prediction. Thus, we believe that these two approaches will ultimately prove to be complementary.

It is the aim of this chapter to review this alternative, constraint-based approach to modeling biochemical systems (Reed and Palsson, 2003). We shall focus on modeling of cellular metabolic networks, provide a comprehensive description of the mathematical and theoretical basis for CBM of metabolic systems. The material has been organized with the objective of a convenient exposition, rather than as an

historical account. Wherever possible, we point to detailed reviews on specialized topics.

2. Biochemical variables and constraints on fluxes

For metabolic systems, the variables of interest are the concentrations of biochemical species (c_i), the fluxes of the biochemical reactions (J_j), and the activities of the associated enzymes (X_j). Furthermore, metabolic fluxes are responsible for maintaining the homeostatic state of the cell, a condition that translates into the assumption that the metabolic network functions in or near a steady state, that is, all of the concentrations are treated as constant in time.

2.1. The mass-balance constraint

In a metabolic steady state, the biochemical fluxes are balanced to maintain constant concentrations of all internal metabolic species. If the stoichiometry of a system made up of M species and N fluxes is known, then the stoichiometric numbers can be systematically tabulated in an $M \times N$ matrix, known as *the stoichiometric matrix* $\widetilde{\mathbf{S}}$ (Clarke, 1980). The S_{ij} entries of the stoichiometric matrix are determined by the stoichiometric numbers appearing in the reactions in the network. For example, if the jth reaction has the form:

$$S_{1j}^{L} A_1 + \ldots + S_{Mj}^{L} A_M \rightleftharpoons S_{1j}^{R} A_1 + \ldots + S_{Mj}^{R} A_M \tag{1}$$

where A_i represents the ith species, then the stoichiometric matrix has the form $S_{ij} = S_{ij}^{R} - S_{ij}^{L}$.

The fundamental law of conservation of mass dictates that the vector of Steady state fluxes, $\mathbf{J}$, satisfies

$$\widetilde{\mathbf{S}}\mathbf{J} = \mathbf{b} \tag{2}$$

where $\mathbf{b}$ is the vector of boundary fluxes that transport material into and out of the system. If the values of the boundary fluxes are not known, equation (2) can be written as $\mathbf{SJ} = \mathbf{0}$ in which the boundary fluxes have been incorporated into the $M \times (N + N')$ matrix $\mathbf{S}$, where N' is the number of boundary fluxes (Qian *et al*., 2003). As an example, consider a simple network of three unimolecular reactions:

$$A \overset{J_1}{\rightleftharpoons} B, \quad B \overset{J_2}{\rightleftharpoons} C, \quad C \overset{J_3}{\rightleftharpoons} A \tag{3}$$

where all reactions are treated as reversible; left-to-right flux is the direction defined as positive. The network of equation (3) is represented by the stoichiometric matrix:

$$\widetilde{\mathbf{S}} = \begin{matrix} A \\ B \\ C \end{matrix} \begin{bmatrix} -1 & 0 & +1 \\ +1 & -1 & 0 \\ 0 & +1 & -1 \end{bmatrix} \tag{4}$$

Next, consider that species A is transported into the system at rate b_A and species B is transported out at rate b_B. Then the mass-balance equations $\widetilde{\mathbf{S}}\mathbf{J} = \mathbf{b}$ can be expressed:

$$\begin{bmatrix} -1 & 0 & +1 \\ +1 & -1 & 0 \\ 0 & +1 & -1 \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix} = \begin{bmatrix} -b_A \\ +b_B \\ 0 \end{bmatrix}. \tag{5}$$

Algebraic analysis of this equation reveals that mass-balanced solutions exist if and only if $b_A = b_B$. Equation (5) can be simplified to $J_2 = J_3 = J_1 - b_A$. Thus, mass balance does not provide unique values for the internal reaction fluxes. In fact, for this example, solutions exist for

$$J_1 \in (-\infty, +\infty), \quad J_2 \in (-\infty, +\infty), \quad J_3 \in (-\infty, +\infty) \tag{6}$$

Equation (6) illustrates the fact that often the mass-balance constraint poses an underdetermined problem; typically, it is necessary to identify additional constraints and/or to formulate a model objective function to arrive at meaningful estimates for biochemical fluxes.

2.2. *Thermodynamic constraints*

In addition to the stoichiometric mass-balance constraint, constraints on reaction fluxes and species concentration arise from nonequilibrium steady state biochemical thermodynamics (Hill, 1977). For a biochemical reaction at a constant temperature T, such as

$$A + B \rightleftharpoons C + D \tag{7}$$

there exists a forward flux J_+ and a backward flux J_-, with net flux $J = J_{+-}J_-$. The concentrations of the reactants and products in this reaction are related to the chemical potential of the reaction via (Qian, 2001)

$$\Delta\mu = \mu_C + \mu_D - \mu_A - \mu_B = k_B \mathrm{T} \ln(J_-/J_+) \tag{8}$$

where k_B is the Boltzmann constant. We see from equation (8) that J and $\Delta\mu$ always have opposite signs, and are both zero only when a reaction is in equilibrium.

For a system of reactions, let $\boldsymbol{\Delta\mu}$ be the vector that contains potential differences for all the reactions. The magnitude of the product $-\mathbf{J}\boldsymbol{\Delta\mu}$ is the amount of heat the reactions dissipate per unit time. The negative sign reflects the Second Law of Thermodynamics according to Lord Kelvin: One cannot convert 100% of heat energy into useful work.

The fundamental law of conservation of energy (First Law of Thermodynamics) dictates that the heat dissipated by the reactions is supplied by constant feeding

of the *open* biochemical system. We have shown that this statement can be mathematically expressed as

$$\boldsymbol{\mu}\mathbf{J} = \boldsymbol{\Delta\mu}\widetilde{\mathbf{S}}\mathbf{J} = \boldsymbol{\mu}\mathbf{b} \tag{9}$$

where the right-hand side of equation (9) is the amount of energy supplied to the system through boundary fluxes (Qian *et al*., 2003). If the system is closed by removing the boundary fluxes, Equation (9) becomes

$$\boldsymbol{\Delta\mu}\mathbf{K} = \mathbf{0} \tag{10}$$

where the matrix $\mathbf{K}$ contains a basis for the right null space of $\widetilde{\mathbf{S}}$. This relation reflects *energy balance*, a generalization of Kirchhoff's loop law and the Tellegen's theorem in electrical circuits (Beard *et al*., 2002).

For the example of equation (3), equation (10) is expressed as:

$$[\Delta\mu_1\ \Delta\mu_2\ \Delta\mu_3]\begin{bmatrix}1\\1\\1\end{bmatrix} = 0 \tag{11}$$

Here, the matrix $\widetilde{\mathbf{S}}$ of equation (4), has a one-dimensional null space, for which the vector $[1\ 1\ 1]^T$ is a basis. Equation (11) corresponds to summing the reaction potentials about the closed loop formed by the reactions in Equation (3).

The requirement that J and $\Delta\mu$ for each reaction in a network have opposite signs provides a stringent constraint on the vectors $\mathbf{J}$ and $\boldsymbol{\Delta\mu}$. Each vector $\mathbf{J}$ that satisfies mass balance $\widetilde{\mathbf{S}}\mathbf{J} = \mathbf{b}$ must also satisfy the following condition in order to be thermodynamically feasible. The system of inequalities

$$\sum_i \mu_i \widetilde{S}_{ij} J_j \leq 0,\ \ (j = 1, 2, \ldots, N) \tag{12}$$

must have at least one nonzero solution for $\{\mu_i\}$ (Qian *et al*., 2003; Beard *et al*., 2004).

Under this constraint, the bounds on the fluxes of equation (5) are narrowed from those of equation (6) to:

$$J_1 \in (0, b_A), \quad J_2 \in (0, b_A), \quad J_3 \in (0, b_A) \tag{13}$$

Thus, in general, the thermodynamic constraint narrows the feasible flux space, but not necessarily to a unique solution. Knowledge of the boundary fluxes translates into constraints on the reaction directions. Thus, the feasible reaction directions are a function of an open system's (i.e., a cell's) interaction with its environment.

Prior to the introduction to the generalized constraint-based thermodynamic theory outlined above, applications of constraint-based modeling of metabolic systems relied on an empirical set of irreversibilities that were obtained from

prior observations of reaction directions in physiological settings. By treating certain reactions as implicitly unidirectional, biologically reasonable results can often be obtained without considering the system thermodynamics as outlined above. Since the system-level thermodynamic constraint is inherently nonlinear, in current and future application of constraint-based modeling, it may be practical to implement unidirectional (irreversibility) constraints on specific reactions in large-scale network models.

3. Biochemical concentrations and enzyme activities

Combining mass balance with the thermodynamic constraint, one can determine the feasible ranges for the biochemical concentrations and the levels of enzyme activities.

3.1. Feasible biochemical concentrations from potentials

Introducing the chemical potential and the energy balance equation provides a solid physical chemistry foundation for the CBM approach to metabolic systems analysis. Proper treatment of the network thermodynamics not only improves the accuracy of the predictions on the steady state fluxes, but can also be used to make predictions on the steady state concentrations of metabolites. To see this, we substitute the well-known relation between the chemical potential μ and the concentration c of a biochemical species, $\mu = \mu_o + k_B \mathrm{T} \ln c$ into the equation (12):

$$\sum_i \ln c_i \widetilde{S}_{ij} J_j \leq -\Delta\mu_j^o \tag{14}$$

where $\Delta\mu_j^o$ is the standard state chemical potential difference for reaction j, which may be obtained from a standard chemical reference source. There is also a concerted effort to establish a central database for this information at the National Institute of Standard and Technology (NIST). If a solution for μ_i exists for equation (12), then there exists a set of corresponding concentrations c_i. In fact, equation (14) provides a feasible space for the metabolites concentrations as a convex cone in the log-concentration space. If the set of feasible concentrations is empty, then the vector $\mathbf{J} = \{J_j\}$ is thermodynamically infeasible.

3.2. Biochemical conductance and enzyme activity

From the traditional biochemical kinetics standpoint, both steady state biochemical concentrations and reaction fluxes are predictable from known enzyme reactions with appropriate rate constants and initial conditions. In steady state, the fluxes are computable from the concentrations of the reactants and products. However, a realistic challenge we confront is that our current understanding of the reaction mechanisms and measurements of rate constants are significantly deficient. From

the standpoint of CBM, however, the ratio between the J_k and $\Delta\mu_k$ of a particular reaction is analogous to the conductance, which can be shown to be proportional to the enzyme activity of the corresponding reaction. We emphasize that the magnitudes of both J_k and $\Delta\mu_k$ are functions of the reaction networks topology. Therefore, each one alone will not be sufficiently informative of the level of enzyme activity (i.e., the level activity due to gene expression or posttranslational modification).

3.3. Conserved metabolic pools

In addition to the constraint on concentrations imposed by equation (14), a reaction network's stoichiometry imposes a set of constraints on certain conserved concentration pools (Alberty, 1991). These constraints follow from the equation for the kinetic evolution of the metabolite concentration vector:

$$\mathbf{P}\mathrm{d}\mathbf{c}/\mathrm{d}t = \mathbf{S}\mathbf{J} \tag{15}$$

where **P** is a diagonal matrix, with diagonal entries corresponding to the partition coefficients, or fractional intracellular spaces, associated with each metabolite in the system. In equation (15), columns corresponding to the boundary fluxes have been grouped into the matrix **S**. Here, the vector **J** includes both internal reaction fluxes and boundary fluxes. The left null space of the matrix $\mathbf{P}^{-1}\mathbf{S}$ may be computed and bases for this space stored in a matrix **L**, such that:

$$\mathbf{L}\mathrm{d}\mathbf{c}/\mathrm{d}t = \mathbf{L}\mathbf{P}^{-1}\mathbf{S}\mathbf{J} = \mathbf{0} \tag{16}$$

It follows from equation (16) the product **Lc** remains constant and defines a number of conserved pools of metabolic concentrations. For example, if we were to consider the glycolytic series as an isolated system, with no net flux of phosphate-containing metabolites into or out of the system, then as phosphate is shuttled among the various metabolites, the total amount of phosphate in the system is conserved.

3.4. Incorporating metabolic control analysis

One of the theoretical frameworks in quantitative analysis of metabolic networks is metabolic control analysis (Westerhoff and van Dam, 1987; Heinrich and Schuster, 1996). In metabolic control analysis, the enzyme elasticity coefficients provide empirical constraints between the metabolites concentrations and the reaction fluxes. These constraints can be considered in concert with the interdependencies in the **J** and **c** spaces that are imposed by the network stoichiometry. If the coefficients $\epsilon_k^i = (c_k/J_i)\partial J_i/\partial c_k$ are known, then these values bind the fluxes and concentrations to a hyperplane in the **J**, **c** space. In addition, the constraint-based approach finds applications (Beard *et al*., 2003) in dynamic metabolic control

analysis, which derives flux control coefficients from linear relaxation kinetics in response to perturbations of a steady state (Reder, 1988; Liao and Delgado, 1993).

4. Optimization theory and objective functions

Constraint-based metabolic analysis is intimately tied to the mathematical field of optimization. Readers may find an accessible introduction to optimization theory, which represents a mature field of modern applied mathematics, in Strang's *Introduction to Applied Mathematics* (Strang, 1986). Optimization theory has been the major mathematical engine behind bioinformatics and genomic analysis (Waterman, 1995). Constraint-based approaches have also been very successful in biological structural modeling ranging from distance geometry calculations for protein structure prediction from NMR to the recent structural determination of large macromolecular complexes (Alber *et al*., 2004). For the purposes of this chapter, it is sufficient to be familiar with the following basic concepts of optimization theory.

Mathematical optimization deals with determining values for a set of unknown variables $x_1, x_2, \ldots, x_n$, which best satisfy (optimize) some mathematical objective quantified by a scalar function of the unknown variables, $F(x_1, x_2, \ldots, x_n)$. The function F is termed the *objective function*; bounds on the variables, along with mathematical dependencies between them, are termed *constraints*. Constraint-based analysis of metabolic systems requires definition of the constraints acting on biochemical variables (fluxes, concentrations, enzyme activities) and determining appropriate objective functions useful in determining the behavior of metabolic systems.

Therefore, in CBM, the objective functions play a crucial role. A given objective function can be thought of as a mathematical formulation of a working hypothesis for the function of particular cell or cellular system. These objective functions should not be considered to be as theoretically sound as the physiochemical constraints; but they may be informative and biologically relevant. They can serve as concrete statements about biological functions and powerful tools for quantitative predictions, which must be checked against experimental measurements. One of the surprising discoveries in constraint-based modeling is how well certain simple objective functions have described biological function (see below).

That a cell functions precisely following some rule of optimality (such as optimal energetic efficiency for optimal cell growth rate) is, of course, highly suspect. There may be an evolutionary argument in favor of certain objective functions, but the ultimate justifications lie in the correctness of its predictions. In this sense, the constraint-based optimization approach provides a convenient way to efficiently generate quantitative predictions of biological hypotheses formulated in terms of objective functions. The value of this approach is in facilitating the systematic prediction–experimental verification–hypothesis modification cycle, ideally leading to new discoveries.

The aim of metabolic engineering is different from that of the traditional biological research (Bailey, 1991; Stephanopoulos, 1994). In metabolic engineering, one is more interested in the "capacity" and optimal behavior of a "biological

hardware" (Edwards and Palsson, 2000; van Dien and Lidstrom, 2002) rather than its natural function per se. For this reason, the metabolic engineering community has been a major proponent of the constraint-based optimization approach, while the cell biologists, whose traditional inclination follow reductionism, view the approach with certain healthy skepticism. Hopefully, with the introduction of thermodynamics, and the establishment of physiochemical basis of the CBM, the two communities will join force in furthering the research on Systems Biology of cells (Hartwell *et al*., 1999).

5. Applications of constraint-based modeling

For bacterial cells, growth rate (rate of biomass production) has been a widely used objective function. This objective is constructed as a net flux out of the cell of the components of biomass (amino acids, nucleotides, etc.) in their proper stoichiometric ratios, which translates into a linear function of the reaction fluxes. On the basis of this elegant paradigm, predictions from FBA of the fate of the *E. coli* MG1655 cell following deletions of specific genes for central metabolic enzymes have been remarkably accurate (Edwards and Palsson, 2000). When combined with energy balance analysis (EBA), it has been shown (Beard *et al*., 2002) that cells with nonessential genes deleted can redirect the metabolic fluxes under relatively constant enzyme activity levels, with few changes due to gene expression regulation and/or posttranslational regulation. The FBA/EBA combined approach predicts which enzyme must be upregulated, which must be down regulated, and which reactions must be reversed given, perturbations to the genotype and/or cellular environment. Using this combined approach, a clear relation is established between the enzyme regulation and constraint-based analysis of metabolism.

Different objective functions can be used in studying other biological systems and problems. When addressing cellular metabolic pathway regulation, robustness has proven to be a useful concept. Robustness can be defined as minimal changes in metabolic fluxes, in steady state concentration, or in enzymes activities following perturbations. For example, minimization of flux adjustment has been used to model the metabolic response of *E. coli* JM101 with pyruvate kinase knockout (Segrè *et al*., 2002). In this study, it is assumed that the cell acts to maintain its wild-type flux pattern in response to the challenge imposed by a gene knockout. However, a minimal change in the flux pattern may require an unrealistic level of metabolic control. We have shown (Beard *et al*., 2004) that the objective of minimal changes in enzyme activities predicts the key regulatory sites in switching between glycogenic and gluconeogenic operating modes in hepatocytes. This approach facilitates inverse analyses, where the regulatory system is treated as a black box and control mechanisms are identified from measurements of the inputs and outputs to the system.

6. Summary

In summary, the constraint-based approach to the analysis of metabolic systems is based on a set of constraints that are imposed by the fundamental laws of mass

conservation and thermodynamics. These laws impose mathematical dependencies on the feasible fluxes, potentials, and concentrations in a given metabolic system. Biological hypotheses may be formulated as mathematical objectives that metabolic systems optimize under the imposed constraints. These hypotheses (e.g., optimal growth or optimally efficient control) are accepted, rejected, or revised on the basis of comparisons between the model predictions and experimental measurements.

Further Reading

Schilling CH, Letscher D and Palsson BO (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal Theoretical Biology*, **203**, 229–248.

Schilling CH, Schuster S, Palsson BO and Heinrich R (1999) Metabolic pathway analysis: basic concepts and scientific applications in post-genomic era. *Biotechnology Progress*, **15**, 296–303.

Schuster S, Dandekar T and Fell DA (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, **17**, 53–60.

Schuster S and Hilgetag C (1994) On elementary flux modes in biochemical reaction systems at steady-state. *Journal of Biological Systems*, **2**, 165–182.

Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, *et al*. (2000) Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *European Journal of Biochemistry*, **267**, 5313–5329.

References

Alber F, Eswar N and Sali A (2004) Structure determination of macromolecular complexes by experiment and computation. In *Practical Bioinformatics*, Bujnicki J (Ed.).●

●[AQ3]

Alberty RA (1991) Equilibrium compositions of solutions of biochemical species and heat of biochemical reactions. *Proceedings of the National Academy of Science United States of America*, **88**, 3268–3271.

Bailey JE (1991) Toward a science of metabolic engineering. *Science*, **252**, 1668-1674.

Beard DA, Liang SD and Qian H (2002) Energy balance for analysis of complex metabolic networks. *Biophysical Journal*, **83**, 79–86.

Beard DA, Qian H and Bassingthwaighte JB (2003) Stoichiometric foundation of large-scale biochemical system analysis. In *Modeling in Molecular Biology*, Ciobanu GM (Ed.), Springer: New York.

Beard DA, Babson E, Curtis E and Qian H (2004) Thermodynamic constraints for biochemical networks. *Journal of Theoretical Biology*, **228**, 327–333.

Clarke BL (1980) Stability of complex reaction networks. *Advances in Chemical Physics*, **43**, 1–215.

Edwards JS and Palsson BO (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capacities. *Proceedings of National Academy of Science United States of America*, **97**, 5528–5533.

Noble D (2004) Modeling the heart. *Physiology*, **19**, 191–197.

Hartwell LH, Hopfield JJ, Leibler S and Murray AW (1999) From molecular to modular cell biology. *Nature*, **402**, C47–C52.

Heinrich R and Schuster S (1996) *The Regulation of Cellular Systems*, Chapman & Hall: New York.

Hill TL (1977) *Free Energy Transduction in Biology: The Steady-State Kinetic and Thermodynamic Formalism*, Academic Press: New York.

Hodgkin AL and Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, **117**, 500–544.

Hopfield JJ (1994) Physics, computation, and why biology looks so different. *Journal of Theoretical Biology*, **171**, 53–60.

Liao JC and Delgado J (1993) Advances in metabolic control analysis. *Biotechnology Progress*, **9**, 221–233.

Qian H (2001) Nonequilibrium steady-state circulation and heat dissipation functional. *Physical Review E*, **64**, 022101.

Qian H, Beard DA and Liang SD (2003) Stoichiometric network theory for nonequilibrium biochemical systems. *European Journal of Biochemistry*, **270**, 415–421.

Reed JL and Palsson BO (2003) Thirteen years of building constraint-based in silico models of *Escherichia coli*. *Journal of Bacteriology*, **185**, 2692–2699.

Reder C (1988) Metabolic control theory: a structural approach. *Journal of Theoretical Biology*, **135**, 175–201.

Segrè D, Vitkup D and Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of National Academy of Science United States of America*, **99**, 15112–15117.

Stephanopoulos G (1994) Metabolic engineering. *Current Opinion in Biotechnology*, **5**, 196–200.

Strang G (1986) *Introduction to Applied Mathematics*, Wellesley-Cambridge Press: Wellesley.

van Dien SJ and Lidstrom ME (2002) Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C_3 and C_4 metabolism. *Biotechnology and Bioengineering*, **78**, 296–312.

Waterman MS (1995) *Introduction to Computational Biology*, Chapman & Hall: New York.

Westerhoff HV and van Dam K (1987) *Thermodynamics and Control of Biological Free-Energy Transduction*, Elsevier: Amsterdam.

Abstract

Constraint-based analysis of large-scale metabolic systems is the computational exploration of metabolic fluxes and concentrations constrained by the physical chemical laws of mass conservation and thermodynamics. This chapter reviews the mathematical formulation of the constraints on reaction fluxes and reactant concentrations that arise as a consequence of the stoichiometry of a specified network of biochemical reactions. Linear algebraic constraints arising from steady state mass balance form the basis of the related computational technologies metabolic flux analysis (MFA) and flux-balance analysis (FBA). Thermodynamic laws, while introducing inherent nonlinearities into the mathematical description of the feasible space, facilitate the introduction of reactant concentrations to the constraint-based framework. Together, the mass-balance and thermodynamic constraints form the basis of an approach to modeling and analysis of biochemical systems that is alternate and complementary to detailed chemical kinetics.

Keywords

metabolic networks, **constraint-based modeling**, **network thermodynamics**, **computational modeling**

Author queries

[AQ1] We have noted that there are no cross-references included in the text of your article. Could you confirm that this is correct? If you would like to add some in now, please indicate their position on the proof, and use the look-up list on www.wiley.com/ggpb (author tools) to cite the article Unique IDs.

[AQ2] This reference has not been listed in the reference list. Please provide the reference details.

[AQ3] Please provide the publication details for this reference.