Chapter 2
Metabolic Network Dynamics: Properties and Principles

Neema Jamshidi and Bernhard Ø. Palsson

2.1 Introduction

Dynamic descriptions of biological processes, especially metabolism, have been of interest for many years (Segel 1975). The size and complexity of these models, however, have stagnated for the last 20 years or so, in spite of dramatic improvements in computational capabilities. The development of large-scale kinetic models (hundreds to thousands of dynamic variables) has been deemed infeasible now for a number of years. Traditional approaches for parameterization of kinetic models require time and labor intensive biochemical assays on individual enzymes. This presents a challenge and a practical limitation to the number of enzymes that can be described using kinetic rate expressions and hence limits the size of networks that can be described dynamically. Furthermore, the confines of a microtiter plate or test tube are often significantly different than those of the intracellular environment. Hence, even once these measurements are carried out, they may not be relevant since the conditions were so different than the in vivo environment. Thus the development of genome-scale kinetic models with this approach has been recognized as infeasible. However, biology is a technology-driven science and new technologies have driven the understanding of biology through the ability to make deeper and broader measurements (e.g., fluxomics and metabolomics). Thus, these new data should analogously motivate and drive the development of new computational approaches.

Future development of network dynamics in biology, particularly with metabolism, will involve two branches, the construction of dynamic networks and the subsequent analysis and simulations of the resulting networks. This chapter will focus on the aspects of the latter; however, the first half will concern basic properties and features of dynamic networks, which will be relevant for both construction and analysis of networks. There are a number of reasons for interest in kinetic models: (1) the ability to make predictions about fluxes as well as concentrations, (2)

N. Jamshidi (✉)
Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0142, USA
e-mail: neema@ucsd.edu

W. Dubitzky et al. (eds.), *Understanding the Dynamics of Biological Systems: Lessons Learned from Integrative Systems Biology*, DOI 10.1007/978-1-4419-7964-3_2,
© Springer Science+Business Media, LLC 2011

a more direct tie-in with experimental measurements, and (3) the ability to make a more direct connection with environmental as well as genetic perturbations by modifying the initial conditions and catalytic and binding constants. The analysis of kinetic models is a necessarily mathematical and computational topic, and there is often a tendency to lose the forest for the trees. The first few sections of this chapter will aim to place the dynamics in a broader context, so that it can be seen how it relates to steady-state flux-based models. While the principles and equations described herein will be applicable to most biological networks, our focus will be metabolism. Furthermore, we will focus on the dynamic hierarchy of metabolic networks with the aim of try to dissect and understanding the interactions that occur between components on different time scales.

2.2 Dynamic Mass Balances and Fundamental Subspaces

Representing biological interactions in terms of mathematical expressions enables one to be precise and unambiguous about what is being discussed. More importantly, however, this enables the ability to benefit from underlying mathematical properties reflected in the equations and the ability to apply physical constraints. For example, through mathematical representation of a metabolic network, one can enforce mass and energy conservation and then explore the implications of these constraints.

The dynamic mass balance equations that describe the dynamic states of biochemical reaction networks are (Heinrich et al. 1977; Reich and Selkov 1981):

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{S} \cdot \mathbf{v}(\mathbf{x}, \mathrm{k}) \tag{2.1}$$

in which $\mathbf{x}$ is an m-dimensional vector of concentrations of the metabolites in the network in R^m, $\mathbf{v}$ is an n-dimensional vector of reaction fluxes in R^n, k represents a set of rate parameters, and $\mathbf{S}$ is the $m \times n$ stoichiometric matrix, containing the stoichiometric coefficients of reactants and products for each reaction in the network (Palsson 2006). The stoichiometric matrix, $\mathbf{S}$, is a mathematical representation of a metabolic pathway or network. Each column in the matrix corresponds to an enzymatic (or nonenzymatic) biochemical conversion which may be reversible or irreversible. The reaction flux vector contains rate expressions for all of the biochemical conversions described by $\mathbf{S}$.

If we consider a nonzero vector, there can be two types of results when the vector is multiplied by a matrix: it can result in a null vector (vector with 0 in all of the entries) or it can be nonzero. Vectors that yield null vectors when multiplied by a matrix lie in the null space. Pre-multiplication (multiplication from the left-hand side) of any column vector is actually a mapping from the column to the row space. Figure 2.1 pictorially illustrates what is described by (2.1). The four subspaces in Fig. 2.1 can be viewed as a 2×2 table with each quadrant defined by fluxes or concentrations across the top and dynamics or conservation quantities on the side.

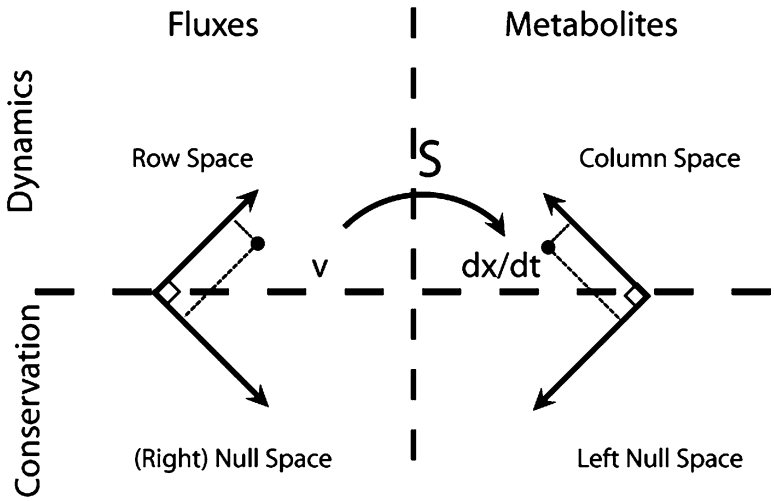


Fig. 2.1 *The fundamental linear sub-spaces and their metabolic network interpretations.* The stoichiometric matrix, **S**, maps **v** which resides in the row and null spaces to d**x**/dt, which resides in the column and left null spaces. The null spaces describe conserved quantities (Palsson 2006); in the right null space this corresponds to conservation of flux and in the left null space this refers to conserved moieties within the network. The row and column spaces describe dynamic states. The row space/right null space and column space/left null space are orthogonal complement pairs, respectively (Strang 1988)

A complete study of the system properties of (2.1) would result in the characterization of all four subspaces of **S** (Strang 1988). The right null and left null spaces of **S** have been studied extensively over the past decades (Palsson 2006; Heinrich and Schuster 1996; Famili and Palsson 2003). The left lower box in Fig. 2.1, for example, is the set of flux balances reflecting mass conservation (total mass accumulated = total mass entering the system – total mass exiting the system). The bounds of the right null space confine the complete set of allowable steady-state flux distributions by enforcing the principle of mass conservation. This subspace has proven to be extremely insightful from a biological standpoint and has been studied extensively during the past decades (Palsson 2006; Heinrich and Schuster 1996). The left null space contains the time-invariant pools, which reflect conserved moieties or functional groups of metabolites in a particular system. Although there has been relatively less investigation into the properties of this subspace in metabolic networks, its significance and meaning has been well described (Palsson 2006; Famili and Palsson 2003).

The row and column spaces are the orthogonal complements to the null spaces, and while dynamic simulations of metabolism have been carried out since the very earliest days of the field of biochemistry (Segel 1975), dynamics are rarely discussed in terms of the subspaces and in relation to their orthogonal complements. Truly appreciating the general principles and underlying factors of dynamics requires recognition of their role and relationship to different subspaces. For example,

thinking about dynamics in terms of the row and column spaces (see Fig. 2.1) leads one to immediately recognize that the null spaces describe "conserved" quantities (i.e., mass) and the row and column spaces of **S** describe the driving forces and the direction of motion for the network variables, respectively.

The topics and challenges of this chapter will focus on the top row of Fig. 2.1; however, occasional mention to the bottom row will be made, because these subspaces are not independent of one another. For example, if the total NAD moiety in a network is assumed to be constant (which would be identifiable through analysis of the left null space), there would be obvious implications in the analysis of the column space for NAD and NADH (knowing the dynamics of one would immediately inform the dynamics of the other).

2.2.1 Key Considerations in Networks

In the spirit of appreciating the components needed to construct dynamic networks, it is important to be cognizant of the nature of molecular interactions as well as some basic assumptions that are regularly made when building kinetic models.

The Michaelis–Menten rate equation is perhaps the most famous and commonly used rate expression for describing reaction kinetics. The rate expressions have proven to be extremely useful when the underlying assumptions have not been violated; however, in vivo conditions such as completely saturated enzymes are not always met. Hence, these rate expressions cannot be claimed to be valid in general. More to the point, the interactions that occur in biological networks, including macromolecular interactions and enzymatic catalysis, are all fundamentally bilinear interactions (Fig. 2.2). That is, most reactions involve two molecules combining to form a third. The general rate law for any of these steps is given by, $v = kx_1x_2$, in which k is a bilinear rate constant, and x_1 and x_2 are the concentrations of the

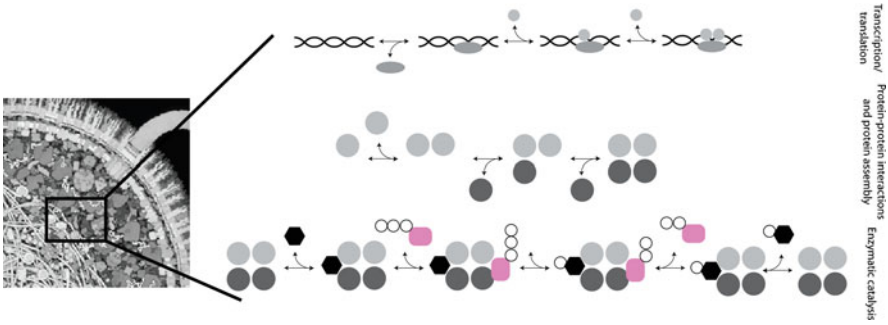


Fig. 2.2 Molecular interactions are almost always combinations of bilinear association or dissociations

interacting components. Fortunately, due to increased processing speed and memory in computers, it is possible to begin describing large networks with complex regulatory schemes in terms of their bilinear interactions.

There is a long history of investigations into the dynamics and kinetics of metabolism. During this course, various mathematically driven operations and procedures have been developed. However, since biology is a technology-driven field, theories and formalisms are only as useful as their ability to integrate available data and to make testable predictions. The approach and formalism here, focusing on the **S** and **G** matrices, are predicated on capturing the key biological features of systems while also enabling the integration of available data types. Along these lines, the gradient matrix has not been defined in any way; it arises naturally from the linearization of the flux vector comprised of net elementary reaction rates. The gradient matrix describes the responses of the reactions to be changed in the concentrations. As will be described later, it is through **G** that the dual nature of the relationship between fluxes and concentrations can be developed.

The metabolic dynamics described by (2.1) also assumes that concentrations can be meaningfully defined (i.e., the number of compounds within the specified volume) and in the absence of any spatial gradients. There are clearly examples when these assumptions fail to be satisfied, for example, in transcription of genes when stochastic effects take place or with excitation–contraction coupling between muscle contraction and energy metabolism, when temporo-spatial gradients have significant effects. We will not address these issues here, but only caution that the modeler should take heed of the physicochemical environment of the phenomena that being modeled and to be cognizant of when particular assumptions may or may not be appropriate.

2.2.2 Properties of Dynamic Systems

Unfortunately, network dynamics are often discussed and viewed with a sense of "magic", and an implication that somehow nonlinearity can make something appear out of nothing. However, if one understands the parts of a model and how they fit together, the results and predictions will be much more palatable and lead to an improved understanding of a network model and its behavior rather than increased confusion. We discuss some key properties of dynamic systems and how they contribute to the properties of networks. There are three matrices that will be of interest in this chapter: the stoichiometric, gradient, and Jacobian matrices. The stoichiometric matrix is a mathematical representation of the "links and nodes" of a network. The columns correspond to the links (or reactions) and the rows correspond to nodes (or chemical species/metabolites). The gradient matrix represents the dependence of the links on the nodes (in the linear regime). The Jacobian matrix is used to describe the overall dynamic relationships in the network; as will be seen however, this matrix can be composed from the stoichiometric and gradient matrices. The key point here is that the fundamental matrices of interest are the stoichiometric and gradient matrices, and these are in fact biological data matrices.

2.2.2.1 Underlying Structure of the Jacobian

Linearization of (2.1) as described in Sect. 2.6.1 results in the ability to define the Jacobian matrix as a product of the stoichiometric and gradient matrix,

$$\mathbf{J} = \mathbf{S} \cdot \mathbf{G} \tag{2.2}$$

The gradient matrix can then be factored, such that

$$\mathbf{J} = \mathbf{S} \cdot \mathbf{K} \cdot \mathbf{M} \tag{2.3}$$

in which $\mathbf{K}$ is an $n \times n$ diagonal matrix whose entries are the lengths of the rows of $\mathbf{G}$ with units 1/time.

Hence, these entries are pseudo time constants or characteristic times corresponding to the reactions in the network. Consequently, the rows in $\mathbf{G}$ indicate the direction that each reaction lies. So the $\mathbf{K}$ and $\mathbf{M}$ matrices describe kinetic and thermodynamic driving factors in the network.[1] This decomposition into the kinetic and thermodynamic influences was carried out without any involved mathematical procedures and has been determined by matrices with biologically meaningful interpretations.

2.2.2.2 Structural Similarity

Reaction rates are commonly expressed as the net sum of elementary reactions. When this is done, it follows that $\mathbf{S}$ and $\mathbf{G}^T$ are structurally similar (Jamshidi and Palsson 2008a) with the corresponding row and column entries have zero or nonzero values. This similarity underlies the stoichiometric influence in network dynamics. In spite of these similarities, there are also key differences between these matrices, which will be touched up on in Sect. 2.4.

2.2.2.3 Flux-Concentration Duality

The first property leads to the ability to define a pair of dual Jacobian matrices. One for the concentrations,

$$\mathbf{J}^x = \mathbf{S} \cdot \mathbf{G} \tag{2.4}$$

and one for the fluxes,

$$\mathbf{J}^v = \mathbf{G} \cdot \mathbf{S} \tag{2.5}$$

[1] Note that $\mathbf{M}$ does *not* refer to the modal matrix in this chapter.

The systems described by each of these equations is the same; however, the independent variables is different, in one the variables are the concentrations and in the other the variables are the fluxes. Note that to convert from the concentration Jacobian to the flux Jacobian, not advanced mathematics or decompositions were used, simply reversing the order of multiplication of two matrices.

2.2.2.4 Hierarchical Dynamics

A key feature of biological networks is the presence of many interactions that occur on a wide range of different time scales. Analysis of these properties has been active for many decades, and there is a rich history in time scale separation and modal decomposition of metabolic networks (Heinrich et al. 1977; Palsson and Lightfoot 1984; Okino and Mavrovouniotis 1998). This will not be detailed here, suffice it to say that one approach that has been successfully carried out for biochemical network analysis has been modal decomposition, which involves the diagonalization of the Jacobian matrix, and the redefinition of concentration variables into "modal" variables which move on dynamically independent time scales.

One overall goal of dynamic analyses of networks is to simplify network structure and to determine which interactions are relevant at particular time scales of interest. This enables one to filter out interactions that are either too fast or too slow to be of interest and to also characterize the progressive pooling of metabolites across slower and slower time scales.

2.3 Dual Jacobian Matrices

One consequence of recognizing the gradient matrix is that it leads to the definition of dual Jacobian matrices and highlights the nature of the relationship between concentration and flux dynamics. The duality between fluxes and concentrations results from the ability to define flux and concentration Jacobian matrices, as mentioned above. The nature of this relationship while mathematically interesting is also of importance from the biological perspective, and hence we will spend some additional time discussing the nature of the relationship. Measurements and perturbations are carried out in terms of concentration variables; however, analysis of the fluxes is what enables interrogation of the systemic properties of networks.

Each network only has a single stoichiometric matrix, $\mathbf{S}$ and gradient matrix, $\mathbf{G}$. However, biological networks can be analyzed in terms of compound (node) variables or in terms of flux (link) variables. Thus, there are two Jacobian matrices describing the same network, $\mathbf{J}^x = \mathbf{S} \cdot \mathbf{G}$ and $\mathbf{J}^v = \mathbf{G} \cdot \mathbf{S}$, depending on which variables, concentrations or fluxes, are used as state variables. The former gives a reaction-centric view of the dynamics, while the latter gives a compound-centric view. These are complementary views of the same system. The relevance of network topology in dynamic systems is highlighted by the fact that the Jacobian matrices

are weighted adjacency matrices containing weighted inner products of the reaction rows and columns ($\mathbf{J}^v$) and compound rows and columns ($\mathbf{J}^x$). Thus, $\mathbf{J}^v$ and $\mathbf{J}^x$ are structurally similar to the reaction adjacency matrix and the metabolite adjacency matrix, respectively (Palsson 2006).

Modal decomposition of the Jacobian has been previously applied for the analysis of biological networks. We note that the two Jacobian matrices share the same eigenvalues. The eigenvectors/rows $\mathbf{J}^x$ relate to pool formation on various time scales (Heinrich et al. 1977; Palsson and Lightfoot 1984), while the eigenvectors/rows of $\mathbf{J}^v$ relate the formation of groups of fluxes that move these pools (Jamshidi and Palsson 2008a). The key point to appreciate is that both views describe the same set of network interactions, but in terms of different dynamic variables; dynamic concentration variables in one case and dynamic flux variables in the other.

2.4 Stoichiometry Versus Gradients

Having stepped through the construction and deconstruction of biological networks, it is hopefully sufficiently impressed upon the reader that network dynamics can be comprehensively characterized through the definition of two matrices: the stoichiometric matrix, **S**, and the gradient matrix, **G**. This is a bold statement, and thus it will be followed by a bold caveat. The gradient matrix is rarely known in general for any condition; hence, experimental and measurement limitations require that it be characterized under a limited set of condition(s). Thus, one generally will only approximate the elements in the matrix and often be restricted to a linearized region close to a particular steady state.

It is important to recognize that **S** and **G** are data matrices, and they are not just of theoretical relevance but have very practical significance and import for the construction and subsequent analysis of kinetic networks (Jamshidi and Palsson 2008a). As mentioned above, when a network is described in terms of bilinear net elementary reactions, which are in general the most appropriate expressions, **S** and $\mathbf{G}^T$ have similar structures. In spite of these structural similarities, there are many important differences between these matrices, which we mention briefly here.

The stoichiometric matrix describes the chemical transformations and interconversions that occur among compounds in a network, and it is through the **S** matrix that mass conservation can be enforced. The gradient matrix, on the other hand, accounts for the kinetic interactions that occur within a network and is constrained by thermodynamic bounds. With these differing physical constraints, there is subsequently different data types that are used to populate these matrices. Genomic and bibliomic data are needed to construct **S** matrices. Alternatively, metabolomic, fluxomic, thermodynamic (e.g., equilibrium constants), and if possible kinetic data are needed to define **G**.

The stoichiometric matrix contains integer entries; hence, it is a "knowable" matrix with the potential of no error associated with its elements. In contrast, the

elements of the gradient matrix are non-integer values and are subject to often significant experimental errors; hence, these entries may often only be known to an order of magnitude. The values within the gradient matrix may differ by more than 10 orders of magnitude; hence, **G** is ill-conditioned and this underlies the stiffness of biological models, which may lead to difficulties when integrating the set of differential equations. However, it is also this wide range of values that leads to the characteristic time-scale separation in biological networks.

Thus, given these similarities and differences between **S** and **G**, the resulting biological interpretation of the matrices also differs. Each **S** matrix is effectively a genomic representation of a species. Thus, different species will have different stoichiometric matrices, and changes in the **S** results from "distal causation" (Mayr 1961). Conversely, the gradient matrix reflects the genetic features of individuals, and so although a species has a single **S** matrix, a population will have a large set of differing **G** matrices. Thus, the gradient matrix represents individual differences within populations and results from changes in "proximal causation" (Mayr 1961).

2.5 Example: Folate Metabolism

The remainder of this chapter will investigate some of the dynamic properties in a dynamic folate metabolic network. As mentioned above, there are many different avenues of analysis to focus on in dynamic networks, and the focus here will be on the metabolite pooling structure within the network on progressively slower time scales and the effects of environmental perturbations on this pooling. Reed et al. (2006) carried out multiple studies with folate one-carbon metabolism in humans with interesting observations with respect to nutrition and genetic variation. Since this is a validated yet small and relatively simple kinetic model, we focus on this network to discuss and highlight some of the discussion points earlier in this chapter.

2.5.1 Constituent Matrices and Subspaces

This network is described by 10 dynamic concentration variables and 20 reaction fluxes. As reflected by the network map (Fig. 2.3), there are often multiple interconversions between the same metabolites. The stoichiometric matrix and the gradient matrix appear in Tables 2.1 and 2.2. Note that all of the entries in **S** are integers, whereas almost all of the entries in **G** are real numbers. Since the model was not constructed strictly from mass action kinetics and various assumptions were made (e.g., Michaelis–Menten kinetics), the transpose of **G** and **S** is not similar in this case (however, if the relationships were explicitly described using mass action kinetics, the similarity between the two matrices would be preserved). Also note that the methionine input flux (metin) is zero for all of the entries in the gradient matrix.

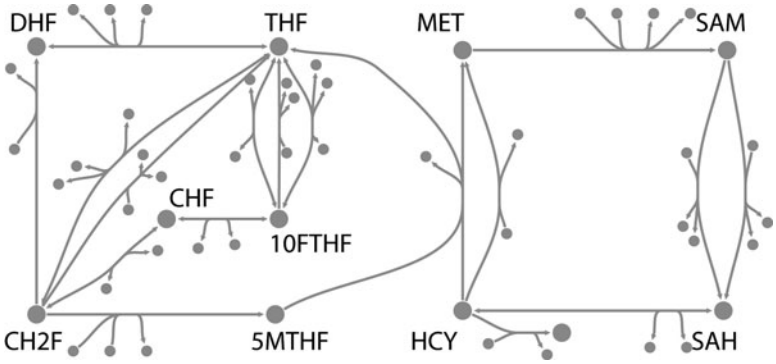


Fig. 2.3 *A map of the folate one-carbon metabolism network.* A map of the folate one-carbon metabolism network for the model described by Reed et al. (2006). Only the dynamic metabolite variables have been labeled. Abbreviations: *5MTHF* 5-methyltetrahydrofolate, *THF* tetrahydrofolate, *DHF* dihydrofolate, *CH2F* 5,10-methylenetrahydrofolate, *CHF* 5,10-methenyltetrahydrofolate, *10FTHF* 10-formyltetrahydrofolate, *MET* methionine, *SAM* S-adenosylmethionine, *SAH* S-adenosylhomocysteine, *HCY* homocysteine

Table 2.1 *The stoichiometric matrix for the folate and methionine cycles metabolism.* Only part of the values of the matrix is shown

	1	2	3	4	5	6	7	8	9	10	...	20
	Vbhmt	Vcbs	Vdnmt	Vgnmt	Vmati	Vmatiii	Vmthfr	Vne	Vaicart	Vdhfr	...	Metin
m5mthf	0	0	0	0	0	0	1	0	0	0	...	0
thf	0	0	0	0	0	0	0	−1	1	1	...	0
dhf	0	0	0	0	0	0	0	0	0	−1	...	0
ch2f	0	0	0	0	0	0	−1	1	0	0	...	0
chf	0	0	0	0	0	0	0	0	0	0	...	0
m10fthf	0	0	0	0	0	0	1	0	−1	0	...	0
met	1	0	0	0	−1	−1	0	0	0	0	...	1
sam	0	0	−1	−1	1	1	0	0	0	0	...	0
sah	0	0	1	1	0	0	0	0	0	0	...	0
hcy	−1	−1	0	0	0	0	0	0	0	0	...	0

This is because this flux was assumed to be constant in the network. As will be seen later, however (see Sect. 2.5.3), even though metin is a constant, varying this value can result in changes throughout the gradient matrix.

The rank, as well as the size of the row and column spaces, of the stoichiometric matrix is 9, and the rank of the gradient matrix is 10. Given the dimensions and rank of each of these matrices, the size of the null spaces can be calculated (Table 2.3). The single dimension in the left null space of **S** reflects conservation of folate within the network. Since folate is never directly synthesized or degraded in this network, it appears in the left null space.

As discussed in Sect. 2.4, the entries in the stoichiometric matrix are integers, whereas those in the gradient matrix are real valued. The condition number for the

Table 2.2 The gradient matrix for the network with a methionine input flux of 200 μM/h. Only part of the values of the matrix is shown. Note that since the methionine input flux is a constant, all of its entries in the gradient matrix are 0

		m5mthf	thf	dhf	ch2f	chf	m10fthf	met	sam	sah	hcy
1	Vbhmt	0	0	0	0	0	0	0	−0.090	−0.090	19.31
2	Vcbs	0	0	0	0	0	0	0	0.077	0.077	104.2
3	Vdnmt	0	0	0	0	0	0	0	0.228	−1.257	0
4	Vgnmt	−67.46	0	0	0	0	0	0	0.294	−3.772	0
5	Vmati	0	0	0	0	0	0	0.746	−0.230	0	0
6	Vmatiii	0	0	0	0	0	0	1.805	0.885	0	0
7	Vmthfr	0	0	35.49	0	0	0	−0.231	−0.231	0	0
8	Vne	0	150.0	0	−23.20	0	0	0	0	0	0
9	Vaicart	0	0	0	0	0	47.80	0	0	0	0
10	Vdhfr	0	0	8847	0	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…	…
20	Metin	0	0	0	0	0	0	0	0	0	0

Table 2.3 The dimensions of the stoichiometric and gradient matrices and the sizes of their right and left null subspaces

	Rows	Columns	Left Null Space	Right Null Space
S	10	20	1	11
G	20	10	10	0

system described by theses matrices is approximately 7.8×10^4. This is a relatively large number and reflects the fact that there is a wide range of concentrations for different metabolites in the network and that some of the biochemical interactions in the network occur much quicker (or slower) with respect to other reactions within the network.

2.5.2 Hierarchical Pooling of Metabolites

As discussed in Sect. 2.2.2, a characteristic feature of metabolism, particularly in higher order organisms, is aggregate pool formation of metabolites when one moves from very fast to very slow time scales (Palsson and Lightfoot 1984). This concept is illustrated in Fig. 2.4 for the glycolytic pathway. For this example pooling between chemical isomers occurs on the earliest time scales (these time scales are all faster than milliseconds). There are chemical as well as physiological relevance to pooling of metabolites, and this process occurs in an organized, hierarchical manner. One challenge in biology is to understand this process, because being able to pool metabolites of interest on a particular time scale enables modularization and simplification of an otherwise complex set of interactions. Furthermore, once the network is modularized, it may be possible to identify metabolites (or summed grouped

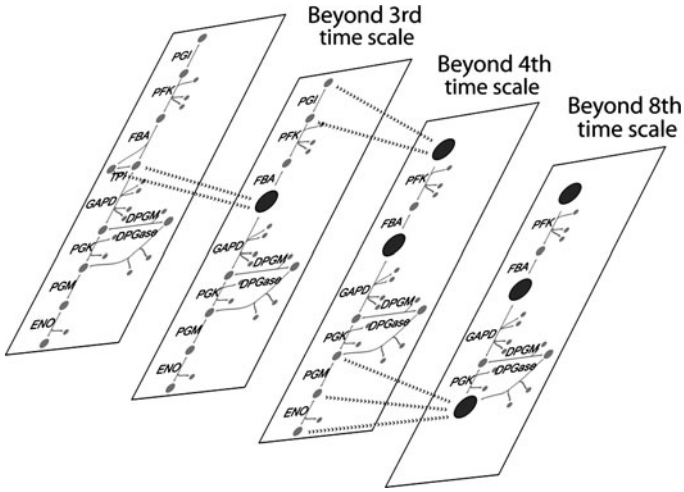


Fig. 2.4 *Beginning from the fastest time scale and moving forward.* Beginning from the fastest time scale and moving forward, components that move together on subsequent time scales are lumped into an aggregate pool variable. A hierarchical reduction of this network is shown in Fig. 2.6

of metabolites) that reflect different "functional states." This may help reduce the number of experimental measurements required to characterize the function of a particular network.

Aggregate pools of metabolites can be identified using different approaches. If there are a small number of possible perturbations of interest, then simulation-driven methods can be used (Kauffman et al. 2002). This approach is limited, however, if one wants to characterize all possible responses of the system. An alternative is to adopt an analytical approach through the analysis of the Jacobian around a particular steady state (Jamshidi and Palsson 2008b). The network can be dynamically decoupled, and then any correlations between metabolites (or fluxes) can be assessed on every single one of the independent time scales. Through calculation of correlations between all of the metabolites (or fluxes) on progressive time scales, removing the time scales one by one (beginning with the fastest), and recalculating correlations between the components, one can identify the pools that form. This procedure (see Jamshidi and Palsson (2008b)) was carried out for the folate network with a methionine input flux of 200 μM/min and are visually depicted in Fig. 2.5.

There were seven independent time scales for the network under these conditions, and there is clear separation of the folate carrier branch from the methionine cycle, although SAM immediately pools with the folate metabolites. Analysis of the kinetics in the context of the stoichiometric matrix identified that the pooling of SAM with the folate cycle was not stoichiometric determined (i.e., there are no reactions that directly involve metabolites from the folate cycle and SAM), but these were kinetically driven events.

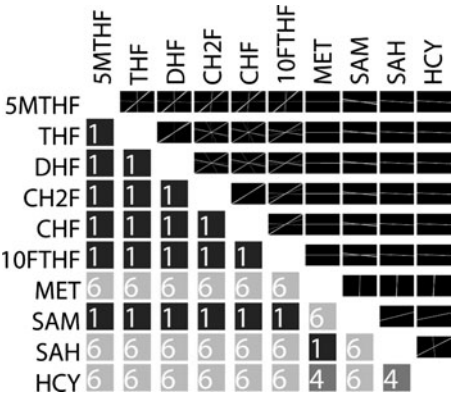


Fig. 2.5 *Pooling among metabolites on progressive time scales.* Methionine input flux at 200 μM/h. Time scale hierarchy of metabolic pool formation in the human red blood cell. The *lower left triangle* indicates the modes after which pooling occurs between the corresponding metabolites (one being the fastest time scale). The *upper right triangle* are plots of the slopes between the two metabolites for the remaining time scales after pool formation (the origin is always included in these approximations of the slopes), color coded according to the time scale at which pooling occurs. A correlation cutoff of 0.9 was used for the pooling criteria determination

2.5.3 Environmental Perturbations

A benefit of building a model in silico is the ability to carry out various perturbations and to observe the changes the occur. The methionine input flux is described by a zero-order rate expression, and as noted in Table 2.2, all of its entries in the gradient matrix are 0. However, changes in the methionine input will cause the system to shift from one steady state to another. This change may result in altered network dynamics. The methionine input flux was considered at halved as well as doubled rates. A cursory glance at the numerical entries shows that many of the values are significantly different under the different conditions. This implies that different homeostatic states have different dynamic properties and quantitatively different systemic – in response to perturbations.

One can immediately see differences in the entries of the gradient matrix, as well as the **K** and **M** matrices for these different conditions. The network-wide changes are more easily highlighted in the tiled pooling arrays of the networks for methionine input fluxes of 400 and 50 μM/h, as shown in Figs. 2.7 and 2.8. At these alternate methionine input flux states, the pooling among metabolites has completely changed. Most notably pooling within the folate cycle occurs much later at the lower methionine input flux rate. At the much higher methionine input flux rate, we see that there are effectively two time scales in which pooling occurs, the first time scale (∼0.5 ms) and the seventh time scale (∼45 s).

These results highlight not only the importance of environmental conditions in the analysis of dynamics in metabolic networks but also the potential for different dynamic properties at different steady states in networks.

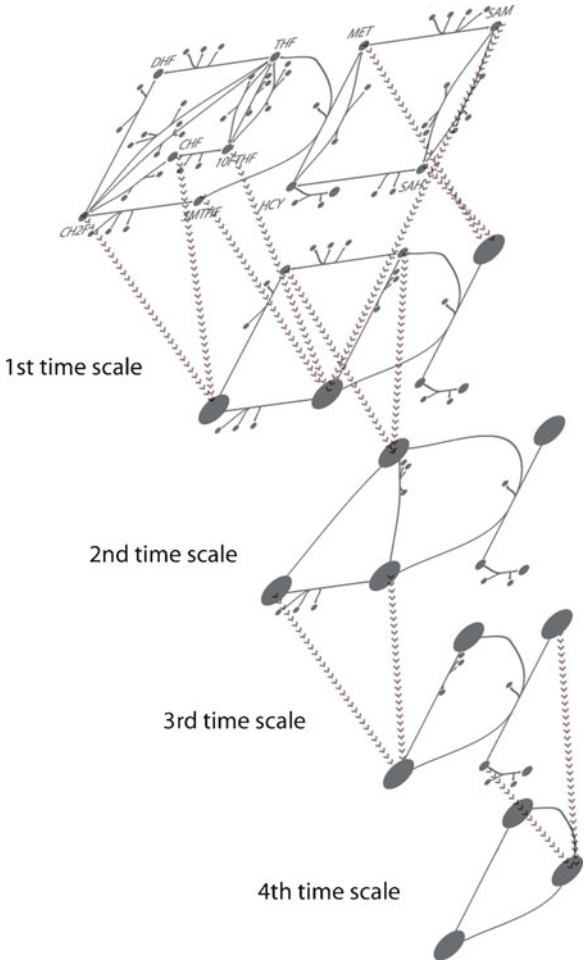
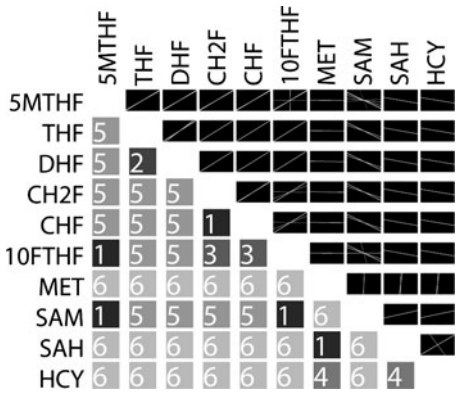


Fig. 2.6 *Hierarchical reduction of the network in Fig. 2.3*. Progressive pooling of metabolites in folate and methionine cycles was determined according to Fig. 2.5

Fig. 2.7 *Pooling among metabolites on progressive time scales.* Methionine input flux at 50 μM/h



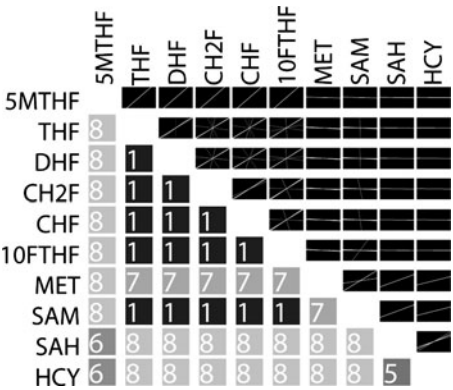


Fig. 2.8 *Pooling among metabolites on progressive time scales.* Methionine input flux at 400 μM/h

2.6 Conclusions

To date there has not been a successful, generalized strategy to build genome-scale kinetic models. This has been principally due to the large number kinetic parameters required to define the system which is further confounded by the fact that in vitro measurements of kinetic constants are often not representative of their numerical values in vivo. These challenges have lead to the infeasibility of achieving cell scale models using such approaches. Identification of the key structural and dynamic properties of networks and the inherent relationships between fluxes and concentrations will help to achieve dynamic descriptions of genome-scale models. Here, we showed how the dynamics of a biochemical reaction network can be described by dual Jacobian matrices, which is enabled by recognition of the fact that dynamic interactions are constrained by network topology. Fluxes and concentrations are dual variables in biochemical reaction networks, but they are related via changes in fluxes and concentrations. These relationships are described by the gradient matrix. The ability to convert from one set of variables into another is not just of mathematical interest, but highlights the underlying roots of the relationship between fluxes and concentrations. Ultimately the characterization of biological systems is to understand how the system responds to perturbations. To date, dynamic descriptions of networks have been confined to the column space; however, the relationships described here allow one to describe the network in terms of column space or row space variables. This is of particular interest in biological networks, as the perturbation variables are generally concentration variables in the column and left null spaces, whereas the response variables are the fluxes, in the row and right null spaces. Thus, concentration variables perturb a network, and the flux variables respond to the perturbation and tie the network together. These are complementary variables that are tied together in the network by the stoichiometric and gradient matrices.

A key motivation in being able to build larger models is to then analyze, understand, and hopefully simplify these networks. Herein, we focused on some of the approaches for simplification of network dynamics in the context of dynamic

hierarchies. A goal in these efforts is to window in on a time scale of interest and to determine the simplified, pooled structure of a network. This effectively filters out processes that occur to slowly or too quickly to be of interest and may highlight grouped metabolites that can be used as surrogates for network functional states, such as the redox state or energy charges of a cell.

2.6.1 Future Directions: Constructing Genome-Scale Models

Previously when models of biochemical reactions and networks have been constructed, it has been through the statement of assumptions such as quasi-equilibrium and quasi-steady state, followed by incorporation of data into the models and curve fitting parameters; thus the statements of assumptions are in effect "preprocessing" the model. The description and decomposition of models described here are carried out from a different perspective. The mechanistic, bilinear interactions are represented in the stoichiometric matrix, the various high-throughput data types (nucleic acid, protein, and small metabolite concentrations) are incorporated into the model, and then the decisions are made for assumptions. These assumptions can be varied and adjusted depending on the question of interest and the time scale(s) of interest. Thus, this is a mechanistic, data-driven approach, in which assumptions are a "postprocessing" step of model construction. This approach recognizes and appreciates the stoichiometric and gradient matrices as the key matrices in building large-scale networks. There has been some progress in this area for outlining the approaches to build kinetic models. We have recently developed an approach that is practical, feasible, and successful in test cases to date (Jamshidi and Palsson 2010). As progress continues to be made in the "-omics" field, particularly metabolomics, we anticipate the development of genome-scale kinetic models in the near future.

Appendix: Details About Matrices

Forming the Gradient Matrix

Dynamic analysis of complex systems is normally carried out with the linearization of the right-hand side of (2.1). Noting that $\mathbf{S}$ is a matrix with constant coefficients, linearization of (2.1) comes down to the Taylor series expansion of reaction rates $\mathbf{v}(\mathbf{x})$:

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(\mathbf{x}_0) + \left.\frac{d\mathbf{v}}{d\mathbf{x}}\right|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}\left.\frac{d^2\mathbf{v}}{d\mathbf{x}^2}\right|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0)^2 + \ldots \tag{2.6}$$

Neglecting all second order and higher terms yields,

$$\mathbf{v}(\mathbf{x}) \approx \mathbf{v}(\mathbf{x}_0) + \left.\frac{d\mathbf{v}}{d\mathbf{x}}\right|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0) \tag{2.7}$$

When the reference state, $\mathbf{x}_0$, is specified as a steady state for the system, then by definition,

$$\mathbf{S} \cdot \mathbf{v}(\mathbf{x}_0) = \mathbf{0} \tag{2.8}$$

so that the linearized form of (2.1) is,

$$\frac{\mathrm{d}(\mathbf{x} - \mathbf{x}_0)}{\mathrm{d}t} = \mathbf{S} \cdot \left.\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{x}}\right|_{\mathbf{x}_0} \cdot (\mathbf{x} - \mathbf{x}_0) \tag{2.9}$$

So quite naturally one can define the gradient matrix, $\mathbf{G}$,

$$\mathbf{G} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\mathbf{x}} \tag{2.10}$$

We note that this is not an arbitrary or a definition of mathematical convenience, but simply the result of linearization of fluxes around a specified reference point. We further note that the gradient matrix is equal to the nonlogarithmic form of the elasticity matrix in metabolic control analysis (Hatzimanikatis and Bailey 1996).

The stoichiometric matrix has been investigated in detail in the literature (Palsson 2006). Since the gradient matrix has only recently been recognized (Jamshidi and Palsson 2008a), time will be spent highlighting and contrasting its key features with the stoichiometric matrix.

The Jacobian Matrix for Concentrations

Specifying the reference point in (2.7), $\mathbf{x}_0$, to be a steady state for the system, (2.1) becomes:

$$\frac{\mathrm{d}\mathbf{x}'}{\mathrm{d}t} = \mathbf{S} \cdot \mathbf{G} \cdot \mathbf{x}' \tag{2.11}$$

in which $\mathbf{x}'$ is the deviation variable, $(\mathbf{x} - \mathbf{x}_0)$. $\mathbf{J}^x = \mathbf{S} \cdot \mathbf{G}$ is the Jacobian for the system of equations describing the concentration variables. Note that this factorization separates the chemistry that specifies network topology (through $\mathbf{S}$), and the kinetics and thermodynamics that give the driving forces and their time scale of action (residing in $\mathbf{G}$). These two effects can be effectively separated by scaling the rows of $\mathbf{G}$ to unity as (Jamshidi and Palsson 2008a):

$$\mathbf{G} = \mathbf{K}^v \cdot \mathbf{M}^v \tag{2.12}$$

where the rows in $\mathbf{M}^v$ represent the direction of the driving forces (the thermodynamics) in the row space. $\mathbf{M}^v$ is a row-normalized gradient matrix, and each row corresponds to a reaction. The matrix $\mathbf{K}^v$ is diagonal. Its elements represent the time scales on which the thermodynamic force of a reaction acts, of the kinetics. In this formulation, the rows of the gradient matrix are drivers and the columns of the stoichiometric matrix define the directions of motion.

The Jacobian Matrix for Fluxes

The concentrations and fluxes are two sets of variables that characterize the dynamic state of a network. Either can in principle be used as the set of independent variables and the other computed as set of dependent variables. Stoichiometric matrices for biochemical networks are, however, normally rectangular with $m < n$, and rank, $r < m$. $\mathbf{S}$ is thus not invertible and (2.1) cannot be directly converted into a system of dynamic equations in terms of fluxes.

The gradient matrix enables the change of the system of equations from the concentration variables to a system of equations in terms of flux variables. Defining the flux deviation variable, $\mathbf{v}' = \mathbf{G} \cdot \mathbf{x}'$, and premultiplying (2.9) by the gradient matrix yields:

$$\frac{\mathrm{d}\mathbf{v}'}{\mathrm{d}t} = \mathbf{G} \cdot \mathbf{S} \cdot \mathbf{v}' \tag{2.13}$$

Thus the Jacobian matrix is $\mathbf{J}^v = \mathbf{G} \cdot \mathbf{S}$, when treating the fluxes as the independent variables. In a similar way as above, we can scale every column in $\mathbf{J}^v$ and factor the gradient matrix as:

$$\mathbf{G} = \mathbf{M}^x \cdot \mathbf{K}^x \tag{2.14}$$

yielding $\mathbf{J}^v = \mathbf{M}^x \cdot \mathbf{K}^x \cdot \mathbf{S}$. Here, $\mathbf{M}^x$ has the columns of $\mathbf{G}$ normalized to unity, and the diagonal matrix $\mathbf{K}^x$ contains the length of these columns, which correspond to compounds. Note that the elements of $\mathbf{M}^x$ represent the kinetic potential of compounds.

$\mathbf{J}^v$ is thus reassembled compound by compound, whereas $\mathbf{J}^x$ was assembled reaction by reaction. In this formulation, drivers (the rows of $\mathbf{S}$) are the sums of the fluxes in and out of a node multiplied by the kinetic potential of the compound. The directions of motions are given by the columns of $\mathbf{M}^x$, and the elements in the diagonal matrix $\mathbf{K}^x$ determine the weights or influence of the motions. The direction of a column in $\mathbf{M}^x$ designates the kinetically balanced outflow of a compound from a node, if the concentration of the compound in that node is perturbed from steady state.

References

I. Famili and B. Ø. Palsson. The convex basis of the left null space of the stoichiometric matrix leads to the definition of metabolically meaningful pools. *Biophys J*, 85:16–26, 2003

V. Hatzimanikatis and J. Bailey. MCA has more to say. *J Theor Biol*, 182:233–242, 1996

H. Heinrich and S. Schuster. *The regulation of cellular systems*. Springer, Heidelberg, 1996

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

N. Jamshidi and B. Ø. Palsson. Formulating genome-scale kinetic models in the post-genome era. *Mol Syst Biol*, 4:171, 2008a

N. Jamshidi and B. Ø. Palsson. Top-down analysis of temporal hierarchy in biochemical reaction networks. *PLoS Comput Biol*, 4:e1000177, 2008b

N. Jamshidi and B. Ø. Palsson. Mass action stoichiometric simulation models: Incorporating kinetics and regulation into stoichiometric models. *Biophys J*, 98(2):175–185, 2010

K. J. Kauffman, J. D. Pajerowski, N. Jamshidi, B. Ø. Palsson, and J. S. Edwards. Description and analysis of metabolic connectivity and dynamics in the human red blood cell. *Biophys J*, 83:646–662, 2002

E. Mayr. Cause and effect in biology. *Science*, 124:1501–1506, 1961

M. S. Okino and M. L. Mavrovouniotis. Simplification of mathematical models of chemical reaction systems. *Chem Rev*, 98:391–408, 1998

B. Ø. Palsson. *Systems biology: Determining the capabilities of reconstructed networks*. Cambridge University Press, Cambridge, 2006

B. Ø. Palsson and E. N. Lightfoot. Mathematical modelling of dynamics and control in metabolic networks. I. On Michaelis-Menten kinetics. *J Theor Biol*, 111:273–302, 1984

M. C. Reed, H. F. Nijhout, M. L. Neuhouser, J. F. Gregory, B. Shane, S. J. James, A. Boynton, and C. M. Ulrich. A mathematical model gives insights into nutritional and genetic aspects of folate-mediated one-carbon metabolism. *J Nutr*, 136:2653–2661, 2006

J. Reich and E. Selkov. *Energy metabolism of the cell: A theoretical treatise*. Academic, London, 1981

I. Segel. *Enzyme kinetics*. Wiley, New York, 1975

G. Strang. *Linear algebra and its applications*. Harcourt Brace Jovanovich, San Diego, 1988

http://www.springer.com/978-1-4419-7963-6

Understanding the Dynamics of Biological Systems
Lessons Learned from Integrative Systems Biology
Dubitzky, W.; Southgate, J.; Fuß, H. (Eds.)
2011, XIV, 238 p., Hardcover
ISBN: 978-1-4419-7963-6