

# Diverse metabolic model parameters generate similar methionine cycle dynamics

Matthew Piazza<sup>a</sup>, Xiao-Jiang Feng<sup>a</sup>, Joshua D. Rabinowitz<sup>a,b,\*</sup>, Herschel Rabitz<sup>a</sup>

<sup>a</sup>Department of Chemistry, Princeton University, Princeton, NJ 08544, USA

<sup>b</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

Received 20 July 2007; received in revised form 12 December 2007; accepted 17 December 2007

Available online 23 December 2007

## Abstract

Parameter estimation constitutes a major challenge in dynamic modeling of metabolic networks. Here we examine, via computational simulations, the influence of system nonlinearity and the nature of available data on the distribution and predictive capability of identified model parameters. Simulated methionine cycle metabolite concentration data (both with and without corresponding flux data) was inverted to identify model parameters consistent with it. Thousands of diverse parameter families were found to be consistent with the data to within moderate error, with most of the parameter values spanning over 1000-fold ranges irrespective of whether flux data was included. Due to strong correlations within the extracted parameter families, model predictions were generally reliable despite the broad ranges found for individual parameters. Inclusion of flux data, by strengthening these correlations, resulted in substantially more reliable flux predictions. These findings suggest that, despite the difficulty of extracting biochemically accurate model parameters from system level data, such data may nevertheless prove adequate for driving the development of predictive dynamic metabolic models.

© 2007 Elsevier Ltd. All rights reserved.

**Keywords:** Parameter identification; Data inversion; Biochemical modeling; Metabolism; Cyclic pathway

## 1. Introduction

Recent technological advances are enabling improved quantification of intracellular metabolite concentrations and fluxes (Brauer et al., 2006; Hollywood et al., 2006; Lu et al., 2006; Sauer, 2006). Constraints-based approaches to modeling metabolism, such as flux-balance analysis, have substantial power for predicting metabolic fluxes in certain organisms (Edwards et al., 2001; Schilling and Palsson, 1998; Steffen et al., 2002; Yuan et al., 2006). Despite this progress, there remains the need to also develop dynamic metabolic models which incorporate understanding of metabolite regulation to predict both metabolite concentrations and fluxes (Di Ventura et al., 2006; Gombert and Nielsen, 2000). The development of such models has been limited in part by the challenge of model parameter identification from laboratory data.

The nonlinear nature of metabolic dynamics calls for nonlinear parameter estimation (also known as data inversion) methods. Gradient-based approaches, such as the Gauss–Newton method, are fast for solving convex optimization problems. However, when the inverse problem is non-convex, these approaches may fail to provide globally optimal solutions (Glowinski and Stocki, 1981; Linga et al., 2006). Global optimization methods such as simulated annealing (Eftaxias et al., 2002; Kirkpatrick et al., 1983), scatter search (Rodriguez-Fernandez et al., 2006b), evolutionary computation (Fogel, 2000; Goldberg, 1989; Moles et al., 2003; Tsuchiya and Ross, 2001), and variants (Rodriguez-Fernandez et al., 2006a), are generally more reliable in these situations. The application of global methods has been reported in a few studies modeling metabolic networks (Himmelblau et al., 1967; Moles et al., 2003; Polisetty et al., 2006).

Another challenge of bionetwork model identification arises from the existence of laboratory noise. For most identification methods, the mean values of the data points are utilized to yield one solution for each unknown model parameter, and variations in the extracted parameters are subsequently

\*Corresponding author at: Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. Tel.: +1 609 258 8985.

E-mail addresses: [joshr@princeton.edu](mailto:joshr@princeton.edu),  
[joshr@genomics.princeton.edu](mailto:joshr@genomics.princeton.edu) (J.D. Rabinowitz).

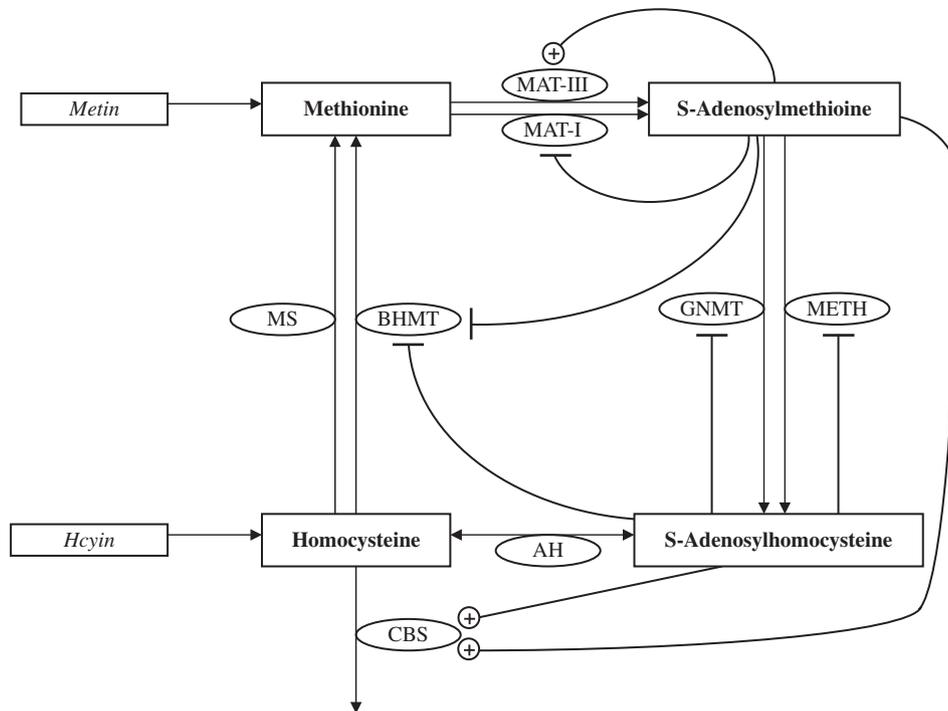


Fig. 1. Methionine cycle as modeled by Reed et al. (2004). Modeled metabolites are in bold within rectangles; enzymes are in ovals; metabolite influxes are in italics. Unmodeled metabolites are not shown. The conversion of methionine (Met) to *S*-adenosylmethioine (AdoMet) is catalyzed by either methionine adenosyl transferase I (MATI) or methionine adenosyl transferase III (MATIII), which are isozymes derived from the same polypeptide chain. The former is feedback inhibited by AdoMet, while the latter is upregulated by it in the concentration ranges used in this study. AdoMet is demethylated by either glycine *N*-methyltransferase (GNMT) or a general AdoMet dependent methyltransferases (METH) to give *S*-adenosylhomocysteine (AdoHcy). Both enzymes are feedback inhibited by AdoHcy. AdoHcy is converted to homocysteine (Hcy) by AdoHcy hydrolase. Hcy can be methylated by Met synthase (MS) or by betaine:homocysteine methyltransferase (BHMT, inhibited by AdoHcy and AdoMet) to re-form Met. Alternatively, it can combine with serine via cystathionine  $\beta$ -synthase (CBS, up-regulated by AdoHcy and AdoMet) to exit the cycle.

estimated via linear propagation of the data noise. As a result, distributions of the identified parameters tend to look regular due to the linear error propagation (Kutalik et al., 2007), giving the impression that the mean parameter values are reliable estimates of the true values. However, the combination of data noise and system nonlinearity (hence nonlinear error propagation) can easily result in wide and highly irregular parameter distributions, where the true parameter values can be located anywhere in the distributions.

In order to address the above issues, Feng and Rabitz developed a general nonlinear method that yields not one, but a representative family of model parameters that can reproduce the bionetwork behavior to within the laboratory error ranges (Feng and Rabitz, 2004). This family then automatically forms the parameter distribution without having to perform explicit error propagations. Initial application of this method (to simulated *in vitro* reactions related to proofreading of tRNA loading) provided evidence that parameter distributions are better captured by this method than by traditional techniques.

In this work, we apply this method to parameter estimation for a metabolic network model that simulates the methionine cycle (Martinov et al., 2000; Reed et al., 2004). The focus is to understand the properties of the extracted parameter families (including the reliability of the predictions that they make) as a function of the input data provided. Biologically, the

methionine cycle was selected because it is a representative of cyclic metabolic pathways (an important metabolic network motif), contains multiple positive and negative feedback mechanisms, and describes a physiologically and medically important system (Fig. 1). Abnormalities of methionine cycle metabolism are associated with liver disease (Avila et al., 2002), cancer (Lu and Mato, 2005), neural tube defects (Refsum, 2001), Alzheimer's disease (Clarke et al., 1998), and cardiovascular disease (Mangoni and Jackson, 2002).

We find that exceptionally diverse and inter-correlated parameter sets exist that generate time-dependent methionine cycle behavior indistinguishable (within modest error estimates) from the published model. These parameter sets not only recapitulate the data used in their identification, they also result in similar model predictions under different perturbations. Sequential addition of flux and biochemical constraints modestly reduced the parameter ranges. Despite the remarkable diversity of the identified parameters, those extracted from concentration data generally yielded reliable concentration predictions. Addition of flux constraints into the parameter search identified parameter sets that were also flux predictive. These findings suggest that, while identification of the true (biochemically accurate) metabolic model parameter values (Teusink et al., 2000) from metabolomic data will generally be difficult owing to system nonlinearity and data noise,

development of dynamic models with substantial predictive power may nevertheless prove feasible. This overall message is consistent with recent analysis of parameter sensitivities in other systems biology models (Gutenkunst et al., 2007).

## 2. Materials and methods

### 2.1. Methionine cycle model

A previously published mathematical model of the methionine cycle developed initially by Martinov et al. (2000) and expanded by Reed et al. (2004) was used to generate the simulated data. The model consists of four nonlinear ordinary differential equations describing the rate of change of methionine (Met), *S*-adenosylmethionine (AdoMet), *S*-adenosylhomocysteine (AdoHyc), and homocysteine (Hyc) (see Fig. 1). These four metabolites interconvert via eight different enzyme catalyzed reactions:

$$\begin{aligned}\frac{d\text{Met}}{dt} &= \text{Metin} + V_{MS} + V_{BHMT} - V_{MATI} - V_{MATIII} \\ \frac{d\text{AdoMet}}{dt} &= V_{MATI} + V_{MATIII} - V_{METH} - V_{GNMT} \\ \frac{d\text{AdoHyc}}{dt} &= V_{METH} + V_{GNMT} - V_{AH} \\ \frac{d\text{Hyc}}{dt} &= \text{Hcyin} + V_{AH} - V_{CBS} - V_{MS} - V_{BHMT}.\end{aligned}$$

Table 1  
Functional forms of rate expressions and associated parameters

Reaction rate	Rate expression	Relevant parameters
$V_{MATI}$	$\frac{V_{\max}^{MATI}}{1 + \left(\frac{K_m^{MATI}}{[\text{Met}]}\right) \left(1 + \frac{[\text{AdoMet}]}{K_i^{MATI}}\right)}$	$V_{\max}^{MATI}, K_m^{MATI}, K_i^{MATI}$
$V_{MATIII}$	$\frac{V_{\max}^{MATIII}}{1 + \left(\frac{K_m^{MATIII}}{[\text{Met}]^2 + K_{m2}^{MATIII}[\text{Met}]}\right) \left(\frac{20000}{1 + 5.7 \left(\frac{[\text{AdoMet}]}{[\text{AdoMet}] + 600}\right)^2}\right)}$	$V_{\max}^{MATIII}, K_{m2}^{MATIII}$
$V_{GNMT}$	$\frac{V_{\max}^{GNMT}}{\left(1 + \left(\frac{K_m^{GNMT}}{[\text{AdoMet}]}\right)^{2.3}\right) \left(1 + ([\text{AdoMet}]/K_i^{GNMT})\right)}$	$V_{\max}^{GNMT}, K_m^{GNMT}, K_i^{GNMT}$
$V_{METH}$	$\frac{V_{\max}^{METH}}{1 + \left(\frac{1}{[\text{AdoMet}]}\right) \left(1 + \frac{[\text{AdoHyc}]}{4}\right)}$	$V_{\max}^{METH}$
$V_{AH}$	$\alpha_1 \cdot ([\text{AdoHyc}] - \alpha_2 \cdot [\text{Hcy}])$	$\alpha_1, \alpha_2$
$V_{BHMT}$	$(0.7 - (0.025)([\text{AdoMet}] + [\text{AdoHyc}] - 150)) \cdot \frac{V_{\max}^{BHMT} [\text{Hcy}]}{K_m^{BHMT} + [\text{Hcy}]}$	$V_{\max}^{BHMT}, K_m^{BHMT}$
$V_{MS}$	$\frac{V_{\max}^{MS}}{1 + \frac{K_m^{MS}}{[\text{Hcy}]}}$	$V_{\max}^{MS}, K_m^{MS}$
$V_{CBS}$	$(\beta_1([\text{AdoMet}] + [\text{AdoHyc}]) - \beta_2) \cdot [\text{Hcy}]$	$\beta_1, \beta_2$

Table 1 lists the functional forms of the flux expressions for each of the eight reactions. The formulation of the model here is identical to that of Reed et al. (2004), except for elimination of any over-parameterization. Specifically, in the expression for  $V_{METH}$ , the parameter  $K_{m2}^{METH}/[A]$  from the original publication was absorbed into  $V_{\max}^{METH}$ , where  $[A]$  represents an arbitrary methylation substrate. In the case of  $V_{MS}$ , the reaction was initially formulated as an ordered bi–bi reaction of the substrates. However, since  $[5\text{mTHF}]$  is a constant in the model,  $V_{MS}$  was re-written as a simple, one substrate Michaelis-Menten expression. Table 2 lists the parameter values used in the revised, reference model. This model was used to generate all the data supplied to the inversion algorithm in the simulations.

### 2.2. Simulated perturbations

Perturb-and-observe experiments were simulated with influxes to the system limited to methionine (*Metin*) and/or homocysteine (*Hcyin*), as described in Section 3. The magnitudes of the perturbations were chosen to approach but not exceed the largest values that could be simulated by the model while retaining consistency with the thermodynamic constraints (i.e., larger perturbations resulted, due to limitations of the published models, in negative fluxes through reactions that are chemically essentially

Table 2  
Model parameter values and search ranges

Parameter	Published model value	Literature values	Biochemical parameter range constraints
$V_{max}^{MATI}$	561 $\mu\text{M}/\text{h}$	–	–
$K_m^{MATI}$	41 $\mu\text{M}$	41 $\mu\text{M}$ (Sullivan and Hoffman, 1983)	33–49 $\mu\text{M}$
$K_i^{MATI}$	50 $\mu\text{M}$	230 $\mu\text{M}$ (Sullivan and Hoffman, 1983)	172–287 $\mu\text{M}$
$V_{max}^{MATHI}$	22870 $\mu\text{M}/\text{h}$	–	–
$K_{m2}^{MATHI}$	21.1 $\mu\text{M}$	–	–
$V_{max}^{GNMT}$	10600 $\mu\text{M}/\text{h}$	–	–
$K_m^{GNMT}$	4500 $\mu\text{M}$	100 $\mu\text{M}$ (Heady and Kerr, 1973; Kerr, 1972)	75–125 $\mu\text{M}$
$K_i^{GNMT}$	20 $\mu\text{M}$	35 $\mu\text{M}$ (Kerr, 1972)	26.2–43.7 $\mu\text{M}$
$V_{max}^{METH}$	421 $\mu\text{M}/\text{h}$	–	–
$\alpha_1$	100 $\text{h}^{-1}$	–	–
$\alpha_2$	10	–	–
$\beta_1$	1.7 $\mu\text{M}^{-1} \text{h}^{-1}$	–	–
$\beta_2$	30 $\text{h}^{-1}$	–	–
$V_{max}^{MS}$	86.1 $\mu\text{M}/\text{h}$	–	–
$K_m^{MS}$	0.0205 $\mu\text{M}$	–	–
$V_{max}^{BHMT}$	2500 $\mu\text{M}/\text{h}$	–	–
$K_m^{BHMT}$	12 $\mu\text{M}$	32 $\mu\text{M}$ (Garrow, 1996), 12 $\mu\text{M}$ (Finkelstein et al., 1972)	9–40 $\mu\text{M}$

irreversible). Concentrations of ancillary compounds (not described by the ordinary differential equations of the model; e.g., 5mTHF) were assumed to be constant. The greatest change in the dynamics of the system occurs within 1 h after each perturbation. Simulated data points were collected every 2.4 min over this 60 min period to supply the inversion algorithm with dynamically rich data for parameter identification. A total of 200 metabolite time points were used for data inversion (4 metabolites  $\times$  25 time points/perturbation  $\times$  2 perturbations). When fluxes (amounts of metabolites per unit volume enzymatically transformed per unit time; i.e., metabolite flows) were also used in the inversion, flux data for each of the five interconversions in the pathway were gathered at each time point as well. A relative error of 15% was assigned to each concentration and flux data point to simulate experimental noise. The typical reported error in metabolomic and fluxomic experimental studies ranges from  $\sim$ 15% to  $>$ 50% (Ikeda et al., 1996; Ruijter and Visser, 1997) and sampling 25 time points per perturbation somewhat exceeds the time-resolution of the most intensive dynamic metabolomic studies to date (Brauer et al., 2006).

### 2.3. The inversion algorithm

A routine based on genetic algorithms (GAs) (Goldberg, 1989) was used for the model parameter identification. An initial population of the unknown parameter vectors was generated, with each parameter value randomly selected within the range of  $10^{-2}$  to  $10^5$ . The “fitness” score of each parameter vector in the population was determined by an objective function which measured how well the parameter vector generated the simulated data. The half of the population with the better fitness scores was retained and

the rest were replaced by new parameter vectors generated through crossover and mutation of existing parameter vectors, with the total (parameter vector) population size held constant at 5000. The mating process operated as follows: two parent parameter vectors were chosen statistically with a probability derived from their respective fitness scores. Each element of each parent parameter vector had a probability  $N_r = 0.7$  of exchanging values with the same element of the corresponding mate parameter vector (genetic recombination) and a probability  $N_m = 0.1$  of being replaced with a random value determined by a Gaussian distribution (genetic mutation). Once the new population of parameter vectors was generated, the process of selection and mating continued until the objective function was minimized.

The objective function can take a variety of forms depending on the type of input data and the nature of the system. The objective here was to minimize the difference between data generated by the model (with the parameter values selected via the inversion algorithm) and the simulated laboratory data. The form of the objective function  $F$  used is given by Eq. (1):

$$F = \frac{1}{N_t} \sum_{i=1}^{N_t} \left[ \frac{w_c}{N_{eq}} \sum_{j=1}^{N_{eq}} \left\{ \begin{array}{l} 1 : |X_{ij}^{cal} - X_{ij}^{lab}| \leq \varepsilon_{ij}^c \\ \frac{|X_{ij}^{cal} - X_{ij}^{lab}|}{\varepsilon_{ij}^c} : |X_{ij}^{cal} - X_{ij}^{lab}| > \varepsilon_{ij}^c \end{array} \right\} + \frac{w_{flx}}{N_{flx}} \sum_j \left\{ \begin{array}{l} 1 : |V_{ij}^{cal} - V_{ij}^{lab}| < \varepsilon_{ij}^f \\ \frac{|V_{ij}^{cal} - V_{ij}^{lab}|}{\varepsilon_{ij}^f} : |V_{ij}^{cal} - V_{ij}^{lab}| \geq \varepsilon_{ij}^f \end{array} \right\} \right], \quad (1)$$

where  $N_t$  is the number of time points at which data was gathered,  $N_{eq}$  is the number of species whose concentrations were measured,  $X_{ij}^{lab}$  is the simulated laboratory

concentration measurement for the  $j$ th species at the  $i$ th time point,  $\varepsilon_{ij}^c$  is the laboratory error associated with the  $j$ th species at the  $i$ th time point (set here to 15%),  $X_{ij}^{cal}$  is the concentration calculated by the model (using the GA-generated parameter vectors) for the  $j$ th species at the  $i$ th time point,  $N_{flux}$  is the number of system fluxes measured,  $V_{ij}^{lab}$  is the  $j$ th simulated laboratory flux measured at the  $i$ th time point,  $\varepsilon_{ij}^f$  is the laboratory error associated with the  $j$ th flux at the  $i$ th time point (set here to 15%);  $V_{ij}^{cal}$  is the  $j$ th system flux calculated by the model (using the candidate parameter vector) at the  $i$ th time point;  $w_c$  is a weight factor for the concentration data, and  $w_{flux}$  is a weight factor for the flux data. The objective function was calculated by evaluating the following conditional for each data point: if the absolute difference between  $X_{ij}^{lab}$  and  $X_{ij}^{cal}$  falls within the error  $\varepsilon_{ij}^c$ , then  $X_{ij}^{cal}$  receives a score of 1; otherwise  $X_{ij}^{cal}$  was scored using the continuous function  $|X_{ij}^{cal} - X_{ij}^{lab}|/\varepsilon_{ij}^c$  that increases linearly from 1 as the deviation of  $X_{ij}^{cal}$  from  $X_{ij}^{lab}$  increases. The scores for each data point were summed and averaged over all time points and measured species/fluxes. For weighting factors related by  $w_c + w_{flux} = 1$ ,  $F = 1$  is the minimum/optimal value. Candidate parameter sets that receive a fitness score  $F = 1$  were considered acceptable solutions. For the inversion of concentration data alone,  $w_{flux} = 0$  and  $w_c = 1$ . For the inversion of both concentration and flux data,  $w_{flux} = 0.5$  and  $w_c = 0.5$ .

#### 2.4. Constraints describing properties of isolated enzymes

Biochemically constrained search ranges for select parameters were based on previously published data from *in vitro* kinetic studies performed for the various enzymes of the methionine cycle. The third column of Table 2 lists the constraints and relevant references from which they were derived. Here we provide a brief justification for them. The constant  $K_m^{MATI}$  for MATI was determined by Sullivan and Hoffman (1983) to be  $41 \pm 4 \mu\text{M}$ . We selected a constraint of  $41 \pm 8 \mu\text{M}$ , with the additional breadth of the range selected to correct for potential *in vitro* versus *in vivo* differences in the  $K_m$ .  $K_i^{MATI}$  was calculated based on a product inhibition study (Sullivan and Hoffman, 1983) which found that at  $[\text{Met}] = 25 \mu\text{M}$  and  $[\text{AdoMet}] \sim 375 \mu\text{M}$  resulting in  $\sim 50\%$  inhibition of MATI activity. Based on the reaction rate formulation for  $V_{MATI}$  in Table 1, this implies that the  $K_i$  for MATI is  $\sim 230 \mu\text{M}$ . Experimental values but not error estimates for GNMT,  $K_m^{GNMT}$  and  $K_i^{GNMT}$  have been reported (Heady and Kerr, 1973; Kerr, 1972). Absent experimental error estimates for  $K_i^{MATI}$ ,  $K_m^{GNMT}$ , and  $K_i^{GNMT}$ , we assigned a constrained search range of  $\pm 25\%$  the nominal value. The Michaelis-Menten constant  $K_m^{BHMT}$  for the BHMT reaction rate was independently determined by Finkelstein et al. (1972) and Garrow (1996) as 12 and 32  $\mu\text{M}$ , respectively. We assigned a biologically constrained search range for this parameter of 9–40  $\mu\text{M}$  (25% below the lower experimental estimate to 25% above the higher one).

#### 2.5. Computational implementation

The ODE model of the methionine cycle was integrated using a stiff Gear integrator (Petzold and Hindmarsh, 1997). The GA was from GALib (Wall, 1995). The parameter distributions were generated as follows: identified parameter sets from three independent inversion runs were combined into a single data set with redundant copies of any particular identified parameter vector removed. For each parameter, the logarithmic parameter search range  $[-2, 5]$  was subdivided into 50 equally sized bins of length 0.14. The frequency with which identified parameter values fell in each of the bins was determined, producing a distribution of outcomes along the range  $[-2, 5]$ . Several thousands of parameter sets were identified in each run, which was sufficient to provide a converging distribution for each parameter.

### 3. Results

#### 3.1. Diverse parameter sets reproduce simulated dynamic concentration data

The simulated concentration data shown in Fig. 2 was supplied to the inversion algorithm to identify families of parameter vectors that reproduce the data. The algorithm identified over 30,000 parameter vectors consistent with the simulated concentration data within error estimates of  $\pm 15\%$ . The resulting parameter values are plotted in the histograms of Fig. 3 (blue lines; the green lines show analogous results with inclusion also of flux data and the red lines with further addition of biochemical constraints; see below). In these histograms, the Y-axis represents the fractional occurrence of particular parameter value ranges (shown on the X-axis) among the parameter vectors consistent with the simulated data. A sharp distribution in Fig. 3 indicates that only a narrow range of the particular value could reproduce the simulated data of Fig. 2. A spiky distribution indicates that a diversity of values are consistent with the data, but the likelihood is higher for certain parameter values. A broad, smooth distribution indicates that all (or very many) values of that parameter are acceptable.

The most notable feature of the blue lines in Fig. 3 is the exceptional breadth of most of the parameter ranges, which cover the entire dynamic range of  $10^7$  for eight parameters ( $V_{max}^{MATI}$ ,  $K_m^{MATI}$ ,  $K_i^{MATI}$ ,  $V_{max}^{GNMT}$ ,  $K_m^{GNMT}$ ,  $K_i^{GNMT}$ ,  $V_{max}^{MS}$ , and  $K_m^{MS}$ ) and exceeded a range of  $10^3$  for seven parameters ( $V_{max}^{MATIII}$ ,  $K_m^{MATIII}$ ,  $V_{max}^{METH}$ ,  $\alpha_1$ ,  $\beta_2$ ,  $V_{max}^{BHMT}$ , and  $K_m^{BHMT}$ ). While some of the parameters in this latter category, such as  $V_{max}^{METH}$ , favor a smaller range of values, scattered cases demonstrate that other divergent values are acceptable. Only two parameters,  $\alpha_2$  and  $\beta_1$ , were identified to within a  $\sim 10$ -fold range.

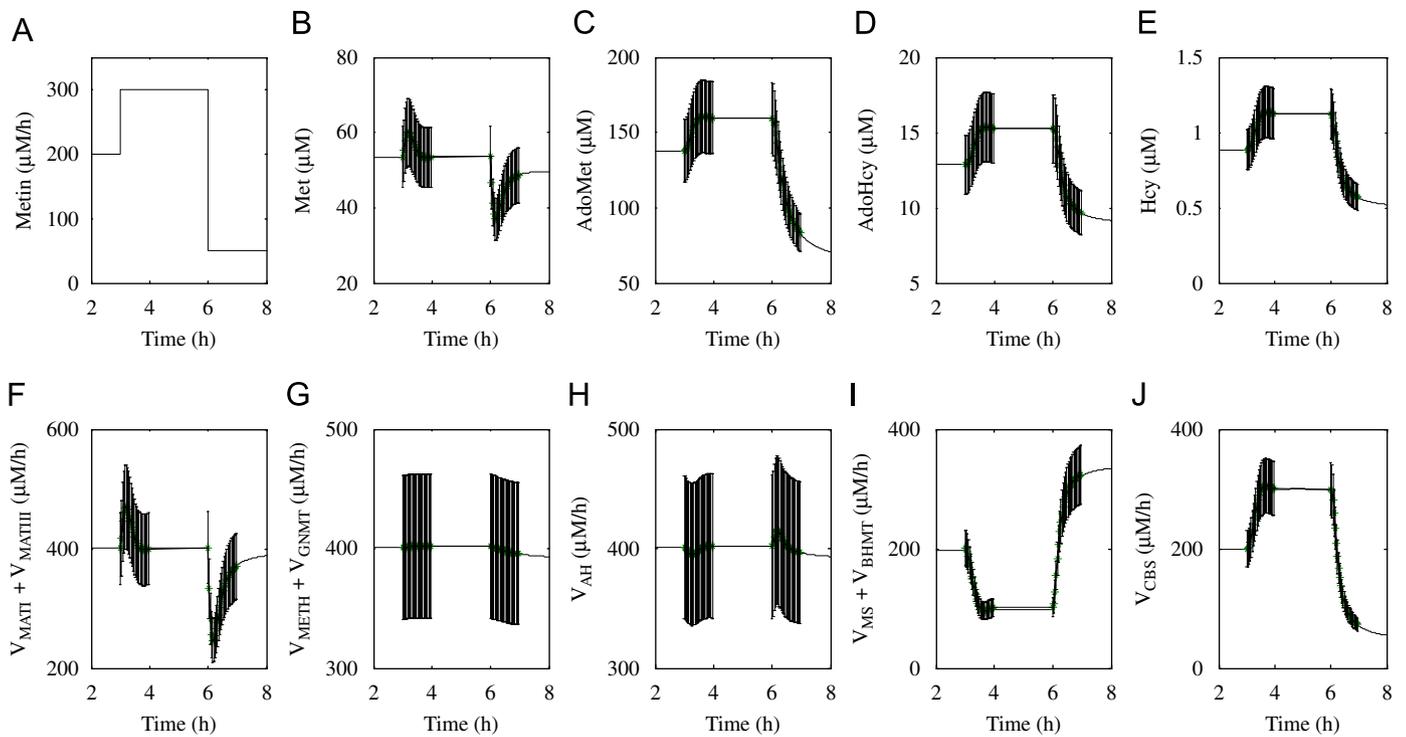


Fig. 2. Simulated methionine flux perturbation and dynamic metabolic responses. Data points used for parameter identification are surrounded by error bars of  $\pm 15\%$  to simulate experimental noise. (A) Methionine influx (*Metin*). (B) Methionine concentration (*Met*). (C) *S*-Adenosylmethionine concentration (*AdoMet*). (D) *S*-Adenosylhomocysteine concentration (*AdoHcy*). (E) Homocysteine concentration (*Hcy*). (F) Flux from *Met* to *AdoMet* ( $V_{MATI} + V_{MATIII}$ ). (G) Flux from *AdoMet* into *AdoHcy* ( $V_{METH} + V_{GNMT}$ ). (H) Flux between *AdoHcy* and *Hcy* ( $V_{AH}$ ). (I) Flux from *Hcy* to *Met* ( $V_{MS} + V_{BHMT}$ ). (J) Flux from *Hcy* to cystathionine ( $V_{CBS}$ ).

### 3.2. Correlations within identified parameter vectors explain the parameter value diversity

While indicating that divergent parameter vectors can produce similar kinetic behavior, the above results do not imply that random parameters selected from the identified individual parameter ranges can do so. Notably, we found that only 1 in  $2 \times 10^6$  parameter vectors, when randomly sampled from the identified ranges, replicated the concentration data within the 15% error range. Thus, while any particular parameter can generally take on a huge diversity of values, the network behavior is conserved only for parameter vectors with highly correlated members.

To better understand the model features resulting in the remarkable diversity and correlation of feasible parameter values, we examined the roles of particular parameters in the ODE model. The parameters that exhibit the most diversity in identified values are associated with the enzymes *MATI*, *GNMT*, and *MS*. Each of these enzymes is involved in one of the three conversions of the methionine cycle that is catalyzed by multiple enzymes (Fig. 1). Most of the identified parameter sets partition the fluxes for these conversions unequally such that *MATI*, *GNMT*, and *MS* have relatively minor flux contributions (with the major fluxes coming from *MATIII*, *METH*, and *BHMT*, respectively). Because the *MATI*, *GNMT*, and *MS*

fluxes are small, changes in the values of the parameters associated with these reactions generally have little effect on the overall network behavior. Only parameter values which result in large fluxes through these enzymes are prohibited. Acceptably low fluxes are ensured either by keeping  $V_{max}$  low or by offsetting increasing  $V_{max}$  values with corresponding increases in  $K_m$ . Fig. 4A–C show that the  $V_{max}$  and  $K_m$  combinations identified for *MATI*, *MS*, and *GNMT* indeed result in low fluxes through these enzymes: the black lines in each of these plots represents the parameter combinations which result in these enzymes producing 10% of the estimated total flux of the pathway step (and the alternative enzymes of *MATIII*, *METH*, or *BHMT* 90% of the flux). The vast majority of the parameter combinations for *MATI*, *MS*, and *GNMT* fall below this line and thus contribute  $< 10\%$  to the total flux. Thus, one reason for the diversity of acceptable parameter values is that, for pathway steps catalyzed by two different enzymes, if one enzyme contributes very little to the total flux, changes in the parameter values for that enzyme (as long as they do not increase its flux greatly) will have very little impact on the overall network behavior.

The parameter ranges for the flux-dominant enzymes, such as *MATIII* and *BHMT*, while better defined than for the minor flux enzymes, are nevertheless broad

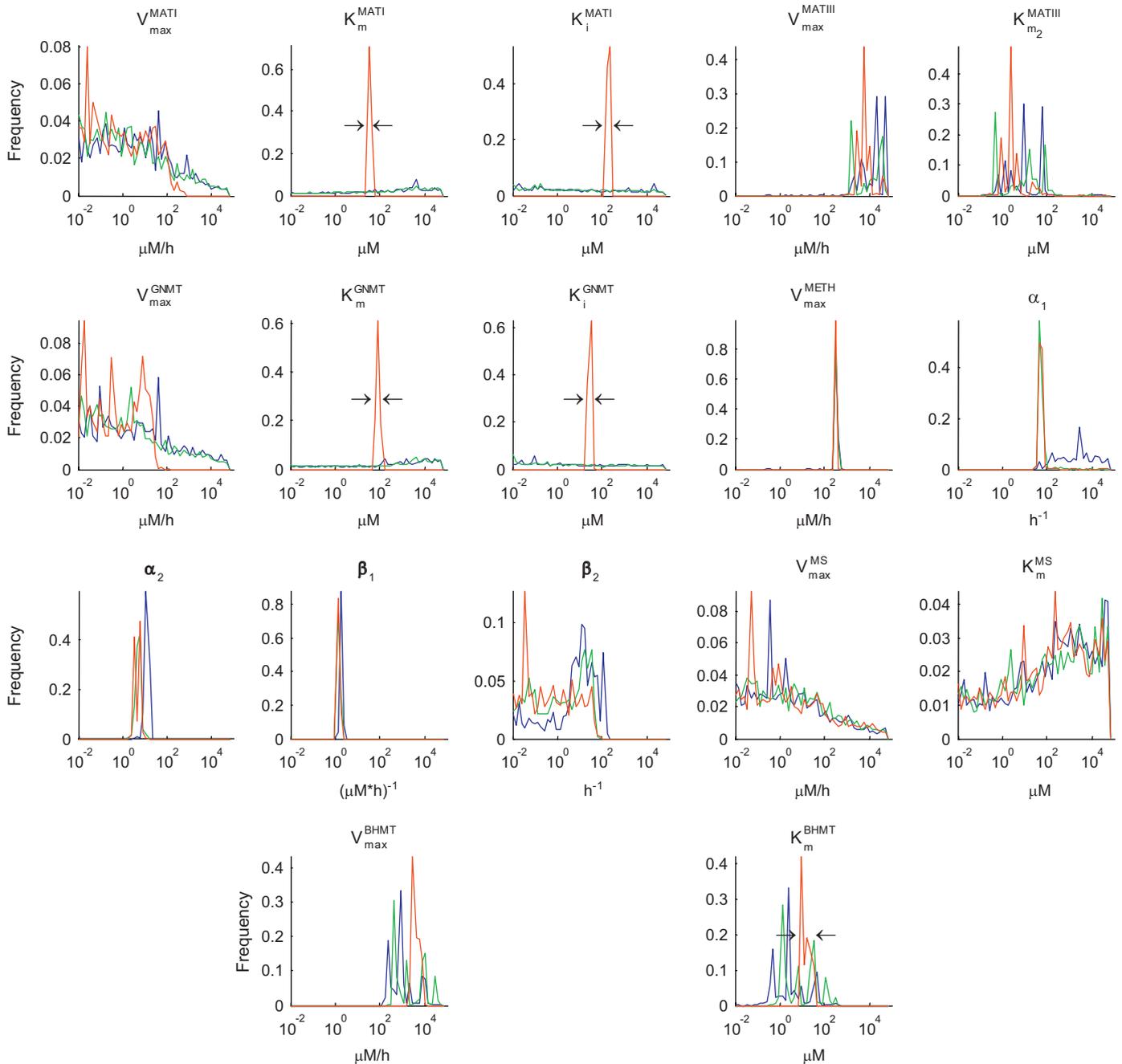


Fig. 3. Histograms of parameter values that reproduce the dynamic data shown in Fig. 2 within the error bars. The results are from inversion of the simulated concentration data only (blue histograms), of both concentration and flux data (green histograms), and of both concentration and flux data with additional biochemical constraints on the search ranges for selected parameters (red histograms; constraints are shown by black arrows surrounding the red peaks for  $K_m^{MATI}$ ,  $K_i^{MATI}$ ,  $K_m^{GNMT}$ ,  $K_i^{GNMT}$ , and  $K_m^{BHMT}$ ; see Table 2 for exact ranges). For all parameters, the X-axis spans from  $10^{-2}$  to  $10^5$ . Global search was conducted over this  $10^7$ -fold range, except for the indicated constraints mentioned above for the red histograms. The X-axis has been binned into 50 equally spaced segments of the parameter search range (on the  $\log_{10}$  scale). The Y-axis values represent the frequency with which the identified parameter values fell within each bin.

(most identified values fall within a 1000–10,000-fold range). One possible reason for such breadth could be correlated changes in two or more parameter values which offset each other to conserve the flux. For unsaturated enzymes described by Michaelis-Menten kinetics, paired increases in  $V_{max}$  and  $K_m$  offset, and the flux is unchanged. Both MATIII and BHMT are far from saturation in both

the original model and most of the new parameter sets. As shown in Fig. 4D and E, for each of these enzymes,  $V_{max}$  and  $K_m$  are strongly correlated (Pearson's  $R^2$  0.95 and 0.97 for MATIII and BHMT, respectively). Thus, another reason for the diversity of acceptable parameter values is that, for unsaturated enzymes, correlated increases in  $V_{max}$  and  $K_m$  do not alter the enzymatic flux.

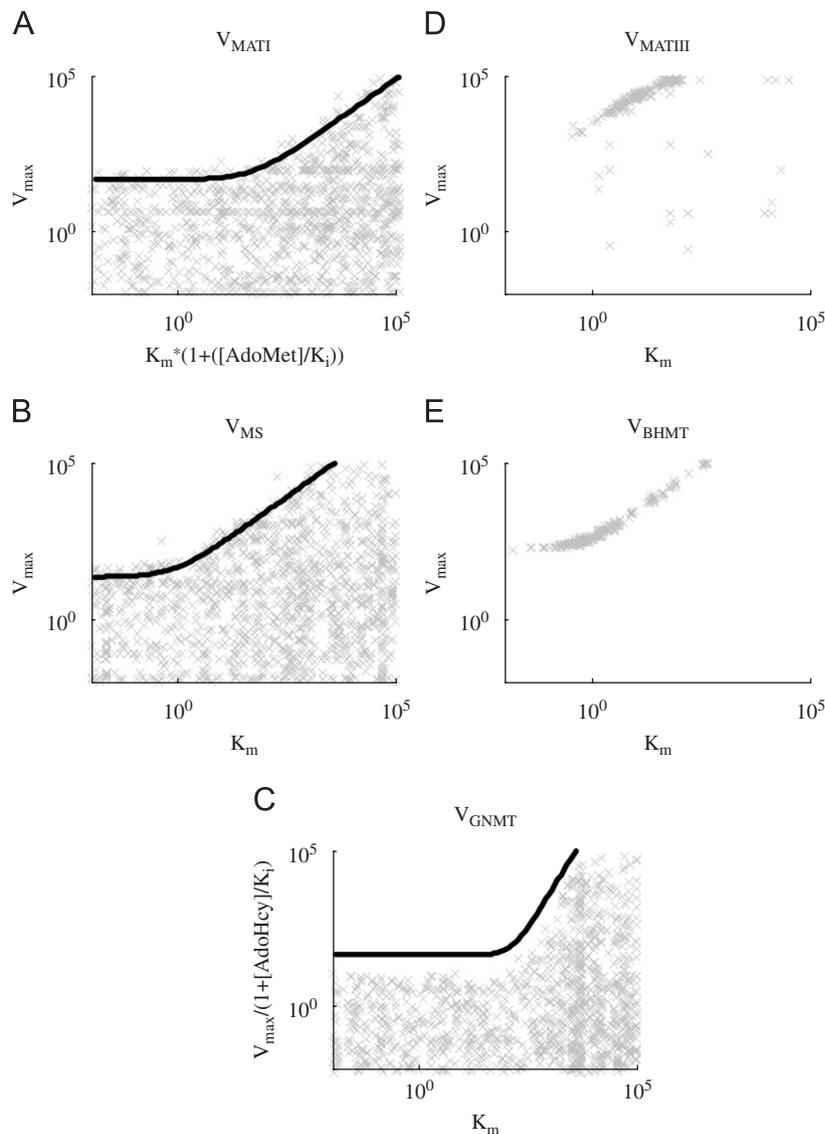


Fig. 4. Correlations between parameter values. Symbols indicate parameter combinations that reproduce the concentration data shown in Fig. 2. MATI (A) and MATIII (D) catalyze the same reaction, as do MS (B) and BHMT (E). The conversion catalyzed by GNMT (C) is also catalyzed by METH, but no correlation plot shown for METH, as it is characterized by only a single parameter value. (A, B and C) Lower flux enzymes. The solid dark lines correspond to parameter combinations producing 10% of the overall reaction flux of the reference model. For these enzymes, the observed parameter values are restricted to combinations of  $V_{max}$  and  $K_m$  that ensure low flux ( $<10\%$  of total). (D and E) Higher flux enzymes, for which  $V_{max}$  and  $K_m$  are roughly linearly correlated ( $R^2 \geq 0.95$ ).

### 3.3. Identified parameter vectors generate similar concentration data predictions

We examined the reliability of the new models (using as their parameters the values identified by the inversion of the simulated concentration data) in making concentration and flux predictions for perturbations (of both methionine and homocysteine influxes, Fig. 5A) different from those used for the parameter identification. The new models (gray lines), despite their high degree of parameter diversity, generally resulted in concentration predictions which closely mimicked the reference system (Fig. 5B). However, certain fluxes, most notably the rate of homocysteine methylation ( $V_{MS} + V_{BHMT}$ ), were not accurately

predicted (Fig. 5C). Eq. (1) provides a quantitative metric of the predictive power of the new models with respect to concentration data ( $w_c = 1$ ,  $w_{flux} = 0$ ) and flux data ( $w_c = 0$ ,  $w_{flux} = 1$ ). Predictions that are accurate with  $\pm 15\%$  error receive the optimal score of 1; worse predictions yield higher scores. The average score of the new models for concentration predictions was almost perfect (mean, 1.0043; S.D., 0.0002). Conversely, the average score for flux predictions was far from perfect ( $23 \pm 0.15$ ). For comparison, random parameter vectors (not generated by inversion of the data shown in Fig. 2) yielded scores of  $280 \pm 150$  and  $97 \pm 3.3$  for concentration and flux predictions, respectively. Hence, while the identified parameter sets had some predictive power for both concentrations

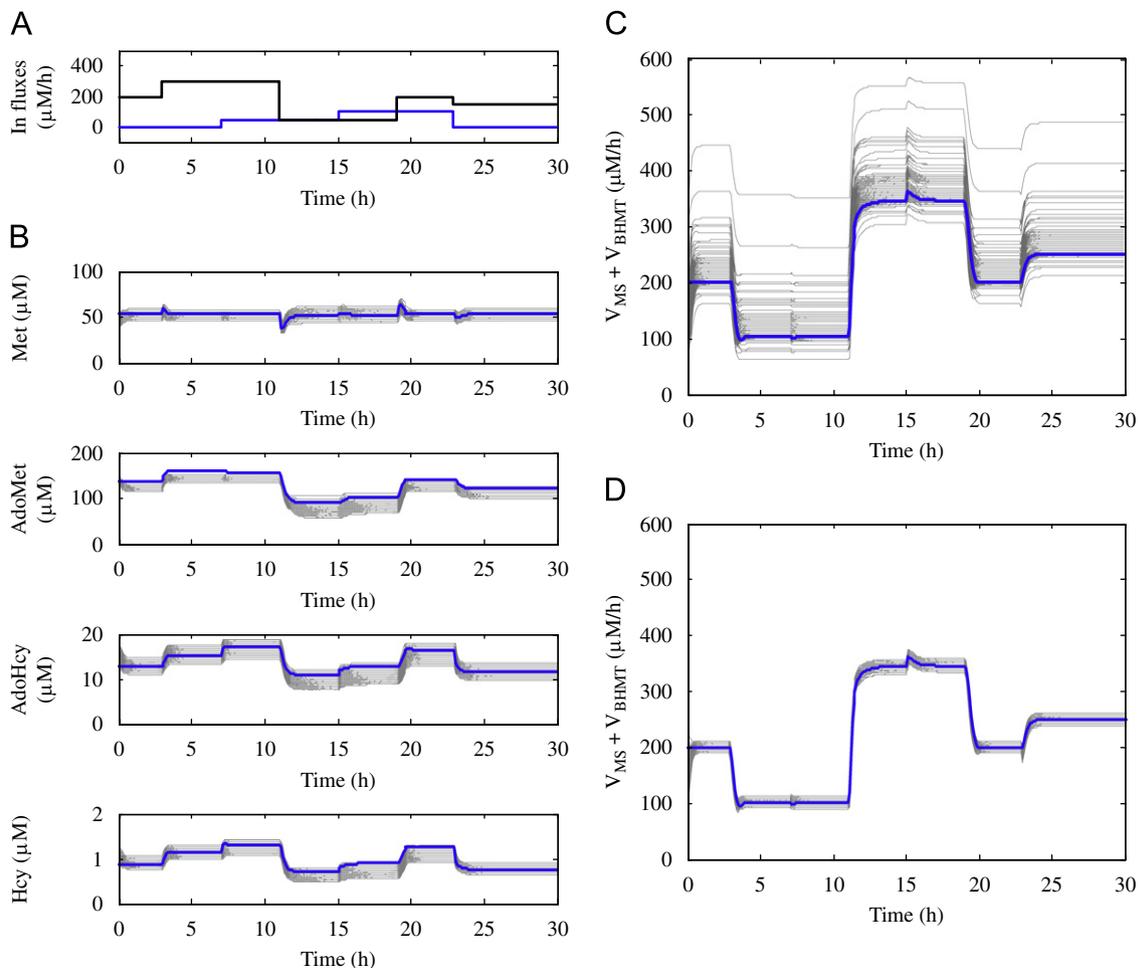


Fig. 5. Predictions of models that contain parameter vectors identified by inversion of the simulated data shown in Fig. 2. Blue lines show the output of the reference model; gray lines ( $N = 200$ ) show the output of the models containing the newly identified parameter vectors. (A) Simulated methionine (*Metin*, black line) and homocysteine (*Hcyin*, blue line) influx perturbations used to generate the output shown in the other panels. (B) Metabolite concentration responses. (C and D) Homocysteine methylation reaction flux ( $V_{MS} + V_{BHMT}$ ) response. The parameter vectors used in (B) and (C) were identified by inversion of concentration data only, and in (D) by inversion of both concentration and flux data.

and fluxes, they were much more accurate for concentrations.

### 3.4. Inversion of flux data yields parameter vectors producing reliable flux predictions

Flux data from the perturbation simulation in Fig. 2 (in addition to concentration data) was incorporated into the inversion procedure to further constrain the parameter search. The inversion algorithm identified  $>30,000$  parameter vectors consistent with the simulated concentration and flux data to within 15% error (Fig. 3, green). The predictive capability of the associated new models was examined using the methionine–homocysteine perturbations of Fig. 5A. The concentration predictions were comparable to those of models trained on concentration data alone (see Fig. 5B); however, the flux predictions were greatly improved (compare Fig. 5C and D). Quantitatively, the new models received scores for concentrations and fluxes of  $1.0041 \pm 0.0003$  and  $1.2 \pm 0.01$ , respectively.

The parameters identified using both flux and concentration data show a discernable reduction in the ranges of acceptable parameter values compared to parameters obtained from inversion of only concentration data (Fig. 3). For  $V_{max}^{MATIII}$  and  $V_{max}^{METH}$ , the breadth of identified distributions dropped  $\sim 10^4$  fold. This reduction in diversity results in parameter ranges that span a region of the parameter space richer in parameter sets that match the simulation data:  $\sim 1$  in  $10^4$  parameter vectors randomly sampled within these ranges regenerated the simulated concentration data from Fig. 2 within the 15% error (compared to 1 in  $2 \times 10^6$  parameter vectors from the broader ranges found by inversion of only concentration data).

As in the case of the parameters identified from concentration data alone, the diversity of flux and concentration-constrained parameters for MATI, GNMT, and MS, can be explained by the lax requirements on these enzymes to maintain  $<10\%$  flux contributions. Similarly, the identified parameters for MATIII and BHMT remain correlated in a manner that holds their associated enzyme

fluxes constant, with the strength of the correlation enhanced by inclusion of the flux data (Pearson's  $R^2$  increase from 0.95 to 0.99 and from 0.97 to 0.998 for MATIII and BHMT, respectively).

### 3.5. Inclusion of biochemical constraints

Even after including flux constraints, considerable diversity of identified parameters remained (Fig. 3). A major source of the observed diversity arises from strong correlations between the individual parameters. This suggests that narrowing the ranges of selected parameter values (using data from literature on their *in vitro* biochemical properties) should also reduce the ranges of other parameters. We accordingly restricted the search ranges for certain parameters describing MATI, GNMT, and BHMT (see Table 2). As before, the inversion algorithm was supplied with concentration and flux data as in Fig. 2. More than  $3 \times 10^4$  parameter vectors that fit the simulated data well were obtained. The distributions of the resulting parameter values are shown in the red lines in Fig. 3. The black arrows surrounding the plots for  $K_m^{MATI}$ ,  $K_i^{MATI}$ ,  $K_m^{GNMT}$ ,  $K_i^{GNMT}$ , and  $K_m^{BHMT}$  indicate the biochemically restricted boundaries of the search ranges. The biochemical restrictions on the six parameters somewhat limited the feasible ranges for the other parameters. Most notably, the distributions of both  $V_{max}^{MATI}$  and  $V_{max}^{GNMT}$  obtained clear upper boundaries absent in the previous inversions (compare the red to the blue and green lines on Fig. 3). Thus, incorporation of suitable biochemical data to restrict the parameter search may significantly reduce the diversity of the consistent parameter vectors. Interestingly, random parameter vectors selected within the parameter ranges obtained with the biochemical constraints were substantially more likely to yield models consistent with the simulation data with  $\sim 1$  in  $10^3$  random parameter vectors yielding results consistent with the simulated concentration data. Thus, the biochemically constrained parameter vectors lie within a region of parameter space that is especially rich in vectors yielding the desired model behavior.

## 4. Discussion

A series of recent modeling studies have made major strides towards identifying topological features of bionetworks that give rise to emergent properties such as the ability to sense and respond to environmental signals, maintain metabolic homeostasis, and form complex developmental patterns, all in a manner that is robust to environmental and genetic noise (Barkai and Leibler, 1997; Bhartiya et al., 2006; Brandman et al., 2005; Kollmann et al., 2005; Meir et al., 2002; Stelling et al., 2002; von Dassow et al., 2000; Wagner, 2005). A complement to understanding the topological features of bionetworks that lead to their qualitative properties is finding the precise parameter values which lead to specific quantitative

phenotypes. This type of detailed quantitative understanding is of practical importance in that parameter values ultimately determine quantitative biochemical phenotypes such as bioreactor yields, insulin sensitivity, etc.

With a focus on the parameter identification problem and its importance in metabolic modeling, here we examined the ranges of specific model parameter values capable of producing particular quantitative behavior of a previously published methionine cycle model (Reed et al., 2004). Using a nonlinear global parameter identification algorithm, we found that remarkably broad ranges of parameter values were consistent with simulated metabolite concentration and flux data within modest error estimates. Such wide and irregular distributions are unlikely to be obtained from linear error propagation as occurs in traditional identification methods. Within these broad parameter value ranges, parameter sets consistent with published biochemical kinetics data were identified. This contrasts to previous methionine cycle models, which resorted to using kinetic constants inconsistent with biochemical literature for certain enzymes (notably for  $K_i^{MATI}$  and  $K_m^{GNMT}$ , see Table 1). The ability to rectify this deficiency points to the value of employing a global search algorithm for selecting more realistic metabolic model parameters.

The primary findings of the present work are that broad ranges of parameter values produce similar model behavior, and that models with highly divergent parameter values nevertheless can make similar quantitative predictions. Similar inferences were made recently for a diversity of systems biology models based on local sensitivity analysis of system response to parameter variations (Gutenkunst et al., 2007). Rather than using sensitivity analysis, our conclusions result directly from the wide and correlated parameter distributions identified by global inversion. Thus, they are a useful complement to those of Gutenkunst et al.

The results of both the present work and Gutenkunst et al. (2007) have clear implications for robustness: at some level, the methionine cycle model used here (Reed et al., 2004), as well as the models studied by Gutenkunst et al., are robust to parameter value variation. However, robustness generally implies similar behavior with parameter value variation in any direction, whereas Gutenkunst et al. showed that parameter value variation in only certain directions is tolerated without markedly altering the quantitative output of the model. Similarly, we find that highly divergent parameter values can produce similar quantitative model behavior, but only if certainly parameter value correlations are maintained. Thus, the breadth of the observed parameter distributions found here does not necessarily imply robustness in the traditional sense.

A basic question regarding our identification of such broad parameter value ranges via data inversion is whether similar results would be obtained with real, rather than simulated, data. The parameter identification here was aided by the fact that the model subjected to inversion had

exactly the same form of the one generating the simulated data. This may not be the case in laboratory settings where there will generally be (1) unknown and/or unmodeled variables in the real network that impact the observed dynamics or (2) failure of the precise mathematical form of the flux equations to completely describe the actual chemical events occurring in the real biological system (due to complex enzymatic mechanisms, diffusion, etc.). In addition, a constraint in the present work was the need to limit the extent of perturbations to avoid non-physiological behavior of the model (e.g., thermodynamically infeasible fluxes). In experimental work, responses to more extreme perturbations might prove useful for better defining feasible parameter ranges. In addition, more extreme perturbations might highlight limitations to model predictive power.

The ability to generalize from the findings here also depends on the degree to which the methionine cycle is representative of a typical metabolic sub-network. Particular features of the methionine cycle found in some but not all other metabolic pathways include (1) the network's cyclic topology, (2) the presence of a large number of regulatory interactions, and (3) the existence of multiple enzymes catalyzing almost all of the pathway steps.

With the above caveats in mind, it is worthwhile to point out some key lessons for future modeling efforts that are suggested by the present results. One is that inclusion of additional types of data (in this case, the flux data), even when leading to only modestly better definition of the feasible parameter ranges, can nevertheless yield important gains in predictive power. Such gains may be hard to realize by simply increasing the amount of data of a given type, and support the application of multiple analytical methods in parallel experimentally.

Other lessons are suggested by the causes of the diversity of the parameter values found here. One cause was the ability of low-flux producing enzymes to adopt a broad range of parameter values while making minimal flux contributions. This suggests an approach to investigating metabolic networks in which one first experimentally identifies the flux-dominant enzymes (e.g., via gene knock-out or siRNA experiments) and then focuses initial modeling efforts on them. Subsequent experiments in cells in which the flux-dominant enzymes have been knocked out could then be used to determine biochemically accurate parameters for the minor-flux enzymes.

Another cause of the parameter diversity was correlations between the flux parameters (typically  $V_{max}$  and  $K_m$ ) for enzymes with dominant fluxes. Such correlations suggest that several of the mathematical flux formulations were over-parameterized, given the relatively narrow ranges of metabolite concentrations observed in the perturbation simulations, which precluded the enzymes from taking on the full range of Michaelis-Menten behaviors. When substrate ranges are limited (e.g., to a <4-fold range), simple linear flux formulations (e.g., flux =  $k$  [substrate]) may be preferable to representations

based on Michaelis-Menten kinetics or biochemical systems theory (Holmberg, 1982; Savageau, 1969; Savageau, 1991). Alternatively, as shown in Section 3, when *in vitro* kinetic data are available, they can be used to tightly constrain  $K_m$  values in Michaelis-Menten formulations, eliminating the problem of dramatic swings in  $V_{max}$  to accommodate widely changing values of  $K_m$ .

A complementary approach to incorporation of diverse types of data is specifically tailoring of experimental perturbations and data sampling for sensitive parameter identification. This can be done through computational simulation of the extent of refinement of the feasible parameter ranges likely to be obtained with new data (Feng et al., 2006). One can imagine especially favorable results by coupling such computationally guided experimental design to laboratory research in a closed-loop fashion.

## Acknowledgments

The authors acknowledge the support of a National Science Foundation (NSF) Dynamic Data-Driven Application System-Small Multidisciplinary Research Projects (DDDAS-SMRP) Grant (0540181), NSF CAREER Grant (0643859), and Environmental Protection Agency STAR grant, as well as the National Institutes of Health Center for Quantitative Biology at Princeton University (P50 GM071508), the Beckman Foundation, and the American Heart Association.

## References

- Avila, M.A., Garcia-Trevijano, E.R., Martinez-Chantar, M.L., Latasa, M.U., Perez-Mato, I., Martinez-Cruz, L.A., del Pino, M.M.S., Corrales, F.J., Mato, J.M., 2002. S-Adenosylmethionine revisited: its essential role in the regulation of liver function. *Alcohol* 27, 163–167.
- Barkai, N., Leibler, S., 1997. Robustness in simple biochemical networks. *Nature* 387, 913–917.
- Bhartiya, S., Chaudhary, N., Venkatesh, K.V., Doyle, F.J., 2006. Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series. *J. R. Soc. Interface* 3, 383–391.
- Brandman, O., Ferrett, J.E., Li, R., Meyer, T., 2005. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science* 310, 496–498.
- Brauer, M.J., Yuan, J., Bennett, B.D., Lu, W.Y., Kimball, E., Botstein, D., Rabinowitz, J.D., 2006. Conservation of the metabolomic response to starvation across two divergent microbes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19302–19307.
- Clarke, R., Smith, A.D., Jobst, K.A., Refsum, H., Sutton, L., Ueland, P.M., 1998. Folate, vitamin B-12, and serum total homocysteine levels in confirmed Alzheimer disease. *Arch. Neurol.* 55, 1449–1455.
- Di Ventura, B., Lemerle, C., Michalodimitrakis, K., Serrano, L., 2006. From *in vivo* to *in silico* biology and back. *Nature* 443, 527–533.
- Edwards, J.S., Ibarra, R.U., Palsson, B.O., 2001. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* 19, 125–130.
- Eftaxias, A., Font, J., Fortuny, A., Fabregat, A., Stuber, F., 2002. Nonlinear kinetic parameter estimation using simulated annealing. *Comput. Chem. Eng.* 26, 1725–1733.
- Feng, X.J., Rabitz, H., 2004. Optimal identification of biochemical reaction networks. *Biophys. J.* 86, 1270–1281.

- Feng, X.J., Rabitz, H., Turinici, G., Le Bris, C., 2006. A closed-loop identification protocol for nonlinear dynamical systems. *J. Phys. Chem. A* 110, 7755–7762.
- Finkelstein, J.D., Harris, B.J., Kyle, W.E., 1972. Methionine metabolism in mammals-kinetic study of betaine-homocysteine methyltransferase. *Arch. Biochem. Biophys.* 153, 320–324.
- Fogel, D.B., 2000. What is evolutionary computation? *IEEE Spectr.* 37, 26–32.
- Garrow, T.A., 1996. Purification, kinetic properties, and cDNA cloning of mammalian betaine-homocysteine methyltransferase. *J. Biol. Chem.* 271, 22831–22838.
- Glowinski, J., Stocki, J., 1981. Estimation of kinetic-parameters-initial guess generation method. *AIChE J.* 27, 1041–1043.
- Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Pub. Co., Reading, MA.
- Gombert, A.K., Nielsen, J., 2000. Mathematical modelling of metabolism. *Curr. Opin. Biotechnol.* 11, 180–186.
- Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., Sethna, J.P., 2007. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3, e189.
- Heady, J.E., Kerr, S.J., 1973. Purification and characterization of glycine N-methyltransferase. *J. Biol. Chem.* 248, 69–72.
- Himmelblau, D.M., Jones, C.R., Bischoff, K.B., 1967. Determination of rate constants for complex kinetics models. *Ind. Eng. Chem. Fundam.* 6, 539–543.
- Hollywood, K., Brison, D.R., Goodacre, R., 2006. Metabolomics: current technologies and future trends. *Proteomics* 6, 4716–4723.
- Holmberg, A., 1982. On the practical identifiability of microbial-growth models incorporating Michaelis-Menten type nonlinearities. *Math. Biosci.* 62, 23–43.
- Ikeda, T.P., Shauger, A.E., Kustu, S., 1996. *Salmonella typhimurium* apparently perceives external nitrogen limitation as internal glutamine limitation. *J. Mol. Biol.* 259, 589–607.
- Kerr, S.J., 1972. Competing methyltransferase systems. *J. Biol. Chem.* 247, 4248–4252.
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* 220, 671–680.
- Kollmann, M., Lovdok, L., Bartholome, K., Timmer, J., Sourjik, V., 2005. Design principles of a bacterial signalling network. *Nature* 438, 504–507.
- Kutalik, Z., Tucker, W., Moulton, V., 2007. S-system parameter estimation for noisy metabolic profiles using Newton-flow analysis. *IET Syst. Biol.* 1, 174–180.
- Linga, P., Al-Saifi, N., Englezos, P., 2006. Comparison of the Luus-Jaakola optimization and Gauss-Newton methods for parameter estimation in ordinary differential equation models. *Ind. Eng. Chem. Res.* 45, 4716–4725.
- Lu, S.C., Mato, J.M., 2005. Role of methionine adenosyltransferase and S-adenosylmethionine in alcohol-associated liver cancer. *Alcohol* 35, 227–234.
- Lu, W.Y., Kimball, E., Rabinowitz, J.D., 2006. A high-performance liquid chromatography-tandem mass spectrometry method for quantitation of nitrogen-containing intracellular metabolites. *J. Am. Soc. Mass Spectrom.* 17, 37–50.
- Mangoni, A.A., Jackson, S.H.D., 2002. Homocysteine and cardiovascular disease: current evidence and future prospects. *Am. J. Med.* 112, 556–565.
- Martinov, M.V., Vitvitsky, V.M., Mosharov, E.V., Banerjee, R., Ataulakhanov, F.I., 2000. A substrate switch: a new mode of regulation in the methionine metabolic pathway. *J. Theor. Biol.* 204, 521–532.
- Meir, E., von Dassow, G., Munro, E., Odell, G.M., 2002. Robustness, flexibility, and the role of lateral inhibition in the neurogenic network. *Curr. Biol.* 12, 778–786.
- Moles, C.G., Mendes, P., Banga, J.R., 2003. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res.* 13, 2467–2474.
- Petzold, L.R., Hindmarsh, A.C., 1997. *Livermore Solver for Ordinary Differential Equations*. Lawrence Livermore National Laboratory.
- Polisetty, P., Voit, E., Gatzke, E., 2006. Identification of metabolic system parameters using global optimization methods. *Theor. Biol. Med. Model.* 3, 4.
- Reed, M.C., Nijhout, H.F., Sparks, R., Ulrich, C.M., 2004. A mathematical model of the methionine cycle. *J. Theor. Biol.* 226, 33–43.
- Refsum, H., 2001. Folate, vitamin B12 and homocysteine in relation to birth defects and pregnancy outcome. *Br. J. Nutr.* 85, S109–S113.
- Rodriguez-Fernandez, M., Mendes, P., Banga, J.R., 2006a. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems* 83, 248–265.
- Rodriguez-Fernandez, M., Egea, J.A., Banga, J.R., 2006b. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics* 7.
- Ruijter, G.J.G., Visser, J., 1997. Determination of intermediary metabolites in *Aspergillus niger* (vol. 25, p. 295, 1996). *J. Microbiol. Methods* 30, 171.
- Sauer, U., 2006. Metabolic networks in motion: C-13-based flux analysis. *Mol. Syst. Biol.*
- Savageau, M.A., 1969. Biochemical systems analysis. I. Some mathematical properties of rate law for component enzymatic reactions. *J. Theor. Biol.* 25, 365–369.
- Savageau, M.A., 1991. Biochemical systems-theory-operational differences among variant representations and their significance. *J. Theor. Biol.* 151, 509–530.
- Schilling, C.H., Palsson, B.O., 1998. The underlying pathway structure of biochemical reaction networks. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4193–4198.
- Steffen, M., Petti, A., Aach, J., D’Haeseleer, P., Church, G., 2002. Automated modelling of signal transduction networks. *BMC Bioinformatics* 3.
- 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.
- Sullivan, D.M., Hoffman, J.L., 1983. Fractionation and kinetic-properties of rat-liver and kidney methionine adenosyltransferase isozymes. *Biochemistry* 22, 1636–1641.
- Teusink, B., Passarge, J., Reijenga, C.A., Esgalardo, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B.M., van Dam, K., Westerhoff, H.V., Snoep, J.L., 2000. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* 267, 5313–5329.
- Tsuchiya, M., Ross, J., 2001. Application of genetic algorithm to chemical kinetics: systematic determination of reaction mechanism and rate coefficients for a complex reaction network. *J. Phys. Chem. A* 105, 4052–4058.
- von Dassow, G., Meir, E., Munro, E.M., Odell, G.M., 2000. The segment polarity network is a robust development module. *Nature* 406, 188–192.
- Wagner, A., 2005. Circuit topology and the evolution of robustness in two-gene circadian oscillators. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11775–11780.
- Wall, M., 1995. *The GALib Genetic Algorithm Package* (copyright 1995–1996 Massachusetts Institute of Technology; copyright 1996–1999 Matthew Wall). Available at: <http://lancet.mit.edu/ga>.
- Yuan, J., Fowler, W.U., Kimball, E., Lu, W.Y., Rabinowitz, J.D., 2006. Kinetic flux profiling of nitrogen assimilation in *Escherichia coli*. *Nat. Chem. Biol.* 2, 529–530.