Comment on "Regularizing capacity of metabolic networks"

Petter Holme[1] and Mikael Huss[2]
[1]*Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131, USA*
[2]*School of Computer Science and Communication, Royal Institute of Technology, 10044 Stockholm, Sweden*
(Received 26 April 2007; published 4 February 2008)

In a recent paper, Marr, Müller-Linow, and Hütt [Phys. Rev. E **75**, 041917 (2007)] investigate an artificial dynamic system on metabolic networks. They find a less complex time evolution of this dynamic system in real networks, compared to networks of null models. The authors argue that this suggests that metabolic network structure is a major factor behind the stability of biochemical steady states. We reanalyze the same kind of data using a dynamic system modeling actual reaction kinetics. The conclusions about stability, from our analysis, are inconsistent with those of Marr *et al.* We argue that this issue calls for a more detailed type of modeling.

DOI: 10.1103/PhysRevE.77.023901 PACS number(s): 82.39.−k, 05.45.−a, 89.75.Kd

Within living organisms, matter is constantly converted between different molecular species. It is often assumed that concentrations of metabolites tend to settle into steady states (rather than showing periodic or chaotic behavior) [1]. Furthermore, experimental studies of metabolic pathways are typically performed under steady-state conditions [2]. Although the steady state is, strictly speaking, a mathematical abstraction, it is nevertheless a useful reference state [3]. The steady-state assumption is also fundamental to traditional metabolic control analysis [2]. In a recent paper [4] Marr, Müller-Linow, and Hütt hypothesize that the network structure of metabolism is an important factor for promoting a steady-state, rather than a complex, dynamics. The authors run an artificial dynamic system on the networks. In short, a vertex i can have two states, 0 or 1. If the sum of the state variables in the neighborhood of i exceeds a fixed threshold, then i changes to the other state. This scheme is very different from real biochemical dynamics and traditional models of reaction kinetics. First, the state variables in metabolic models are usually continuous (concentrations [2]) or sometimes discrete variables (molecule counts [5]), but never binary. Second, the sum of these variables is conserved (if in- and outflow is neglected); whereas, in the dynamics of Marr *et al.*, this is not the case. This dynamics in fact does not reach a steady state; instead, the authors analyze the complexity of the output time series with entropylike measures. They conclude that real metabolic networks give a less complex output than different ensembles of model networks, and argue that this implies that the structure of real metabolic networks may promote steady-state dynamics. However, no theory for how the values of the "entropies" of the binary dynamics relate to the stability of metabolic flux is given. Many authors have used binary dynamics to model other processes of cellular biology, such as signal transduction (e.g., Ref. [6]) or genetic regulation (e.g., Ref. [7]). However, these models explicitly try to describe systems in which the components, at least under some circumstances, show switchlike, binary behavior [6]. Moreover, the lack of detailed understanding of, e.g., gene expression makes it more natural to use simplified dynamics to uncover underlying principles. Metabolic reactions follow known principles of chemistry and are described by well-established differential equation models. Therefore, in contrast to the mentioned information processes, there is no need for, or natural interpretation of, a binary description of the metabolism. Therefore the conclusions of Marr *et al.* can be validated by comparing their results to results obtained using such standard differential equation modeling. We follow this approach to find no significant difference between real and null-model networks. Our simulations in this Comment are rather limited, but sufficient to make the point that the study of Marr *et al.* cannot rule out the possibility that stable steady states can be explained more simply, as direct consequences of the reaction kinetics, rather than as a result of the interaction between the network topology and the dynamic system.

The examples of Marr *et al.* of real metabolic networks are taken from a work of Ma and Zeng (MZ) [8]. They are *substrate graphs* [9] where a substrate is linked to the products of a reaction [Fig. 1(a)]. MZ also preprocessed the data by omitting ubiquitous "currency metabolites" [10]. We also use the MZ networks as the starting point for our simulations. In constructing the networks, information is lost; so if one wants to simulate the reaction system with a substrate graph as a starting point, one needs (explicitly or implicitly) to recreate the reactions via a model. From two assumptions about reaction systems, we propose a simple scheme to create a plausible set of reactions that can be reduced to a given substrate graph. We first assume that all reactions are of a *2-2 form* $A+B \leftrightarrow C+D$, or *2-1 form* $A+B \leftrightarrow C$. These are the most common forms of biochemical reactions. We do not include more complex reactions, mostly because it would significantly increase the computational complexity. Our second assumption is that the number of reactions creating an edge in the substrate graph is rather small (we confirm *a posteriori* that the average number of reactions per vertex is similar to that of the real substrate graph). The following algorithm is a simple way of fulfilling these assumptions.

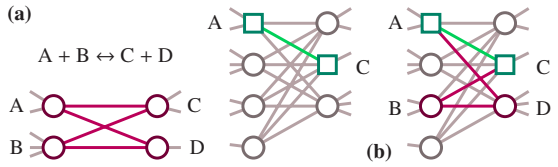

FIG. 1. (Color online) Illustration of the construction of substrate graphs from chemical reactions (a), and our method for recreating a reaction system from a substrate graph (b).

(1) Start with all edges unmarked.

(2) Pick one unmarked edge (A,C).

(3) Find the maximal, full bipartite subgraph $K_{(A,C)}$ (a subgraph consisting of two vertex sets, and edges between every pair of vertices in the different sets, but no edge between vertices in the same set) that contains (A,C). [See Fig. 1(b).]

(4) Pick one four-cycle of $K_{(A,C)}$ including (A,C) [say (A,C,B,D,A)]. [See Fig. 1(b).]

(5) Add $A+B\rightarrow C+D$ and $C+D\rightarrow A+B$ to the set of reactions and mark the edges (A,C), (C,B), (B,D), (D,A).

(6) If there are unmarked edges go to step 2.

$K_{(A,C)}$ gives all reactions of the 2-2 form that induce the edge (A,C) in a substrate graph. If $K_{(A,C)}$ is empty at step 3, then (A,C) must have been generated by a reaction of a 2-1 form. Therefore, instead of $K_{(A,C)}$, consider the set $L_{(A,C)}$, of (A,C) and its adjacent edges, and deduce reactions of the 2-1 form at step 5.

Once we have a set of reactions, derived from a substrate graph, we simulate the biochemical dynamics by standard Michaelis-Menten (MM) kinetics [11] supplemented by noise in the enzyme and substance concentrations. MM kinetics are used to model enzyme-catalyzed reactions, which are common in metabolism. The MM description builds on the principle of mass action [11]—the forward rate of a reaction is proportional to the product of the concentrations of the substrates—and uses additional assumptions to simplify the resulting rate laws. For the reaction i, $A+B\rightarrow C+D$ (which is assumed to be enzyme catalyzed), we use a version of the two-substrate MM rate law to calculate the flux

$$r_i = \frac{E_i + \eta_E}{k_0^{-1} + 1/k_1[A] + 1/k_2[B] + 1/k_{12}[A][B]}, \tag{1}$$

where η_E is a random variable modeling fluctuations in the enzyme concentration. The time evolution of the concentration of a substance A is then determined by

$$\frac{d[A]}{dt} = \eta_{\text{conc}} + \sum_i \pm r_i, \tag{2}$$

where the sum is over all A's reactions, the sign in front of r_i is positive (negative) if A is a product (substrate) of i, and η_{conc} is a random noise term modeling fluctuations due to in- and outflow of A. We seek to use the same information as in Ref. [4]; therefore we do not use empirical parameter values (which are, anyway, hard to obtain). Instead, we assign parameters in arbitrary units, but we choose them to be reasonably relative to one another. η_E and η_{conc} are normally distributed $N(0,0.002)$ (the first argument is the mean; the second is the standard deviation). New values for η_E and η_{conc} are drawn at every time step of the integration. The initial values of substance concentrations, enzyme concentrations, and reaction coefficients (k_0, k_1, k_2, and k_{12}) are drawn from $N(1,1)$, $N(0.2,0.2)$, and $N(0.1,0.1)$, respectively. We choose $\mu_k=1\times10^{-3}$, $\sigma_k=5\times10^{-4}$, $\mu=1$, and $\sigma=0.5$. The equations are integrated with a second-order Runge-Kutta-Helfand-Greenside scheme with time step 0.1, total running time 2500 time units, and 20 averages over different sets of initial configurations. These runs are, for each network, averaged over 20 realizations of the reaction-system construction. For comparison, we also run the dynamics on 100 samples of one of the null-model networks of Marr *et al.*—random graphs with the same degree sequence as the original network. These are obtained by random rewiring—we go through all edges and for each edge (i,j) randomly pick an edge (i',j') and replace these edges with (i,j') and (i',j) [unless this would introduce a multiple edge or a self-edge, in which case a new random edge (i',j') is selected].

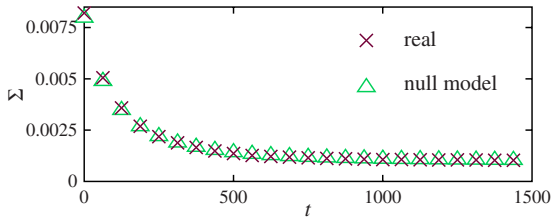

FIG. 2. (Color online) Time evolution of the average standard deviation Σ of the flux. Standard errors are smaller than the symbol size.

In the simulations, a vast majority of the substance concentrations converge to steady states. A few subnetworks of oscillating or chaotic concentrations may exist [12], but in this Comment we focus on bulk properties. To study the approach to equilibrium, and fluctuations, we measure the average standard deviation of the flux through the substances,

$$\Sigma = \sqrt{\frac{1}{n}\sum_i \Phi_i^2 - \left(\frac{1}{n}\sum_i \Phi_i\right)^2}, \tag{3a}$$

where

$$\Phi_i = \frac{1}{2}\sum_j |r_j|. \tag{3b}$$

The sums in Eq. (3a) are over all substances; the sum in Eq. (3b) is over i's reactions; the factor 1/2 comes from the double count of mass flow in and out of substance i, and n is the number of substances. In a plot of $\Sigma(t)$, the stability of the steady state can be monitored by two quantities. First, when the system approaches an equilibrium, $\Sigma(t)$ will decrease—a faster decrease implies a more stable system. Second, a lower equilibrium level means that the system has fewer cyclic or chaotic components, and responds faster to perturbations from the noise, and therefore has more stable steady states. In Fig. 2 we plot $\Sigma(t)$ for MZ's human metabolic network and the null-model networks. The two curves almost overlap. For different null-model realizations the curves may deviate slightly from those of the real network, but there the null model cannot be rejected with any high level of significance. The equilibrium level is of the order of the noise, which means that periodic and chaotic behavior is almost fully suppressed. To be more systematic, we note that $\Sigma(t)$ displays two aspects of stability—how fast equilibrium is reached and the height of the equilibrium level. Since the curves do not fit any simple functional form, we measure the

half-time [the time to reach midway between $\Sigma(0)$ and $\Sigma(\infty)$] using spline interpolation. The p value (fraction of null-model observations lower than the real value) of the half-time is 67%; the corresponding value of the equilibrium level is 65%. The effect of the difference in network structure between the real and null-model networks is thus, in this case, negligible. We test a few other sets of parameter values, organisms (*Escherichia coli* and *Mus musculus*), and a more straightforward mass action kinetics; and, in all cases, arrive at the above conclusion. A full scan of the parameter space would be interesting, however, in this Comment we just make the point that the conclusion of Marr *et al.* can be inconsistent with more realistic simulations, and leave the insignificance of network structure for the stability of metabolic steady states as a conjecture.

To conclude; simple, stylized dynamic systems—"dynamic probes"—are valuable tools for studying complex biological systems. We believe that these should be designed to model the real dynamics as closely as possible. Marr *et al.* [4] propose a dynamic probe to study metabolic steady states that violates many of the known features of reaction kinetics; it does not even reach a steady state—the objective of the study as given in their Abstract. By the collective effort of researchers, our understanding of metabolism continuously advances. This does not, however, include the approach of Marr *et al.* as their dynamics does not make use of biochemical information. Relative to biological information processes, metabolism is a rather simple system, and is believed to be well described by simple differential equation models. For the simulations we carry out, the null-model networks are not less efficient than the real networks in suppressing complex dynamics. If this holds in general, then steady-state stability is a fundamental property of chemical reaction systems. The study of Marr *et al.* does not rule out such a simple explanation of steady-state behavior—an explanation that is a common opinion in biochemical literature (cf. Hofmeyr and Cornish-Bowden's dictum "mass action is the intrinsic driving force for self-organization of reaction networks" [3]). If the reaction kinetics is the sole cause of metabolic steady states, the steady-state dynamics is a constraint to, rather than an outcome of, natural selection. This situation is reminiscent of the power-law degree distribution of metabolic networks. Such distributions can also be seen in astrochemical networks that are not subject to natural selection [13]. We believe the question of dynamic stability in metabolism should be studied at a more detailed level than networks, avoiding the reduction of database information to substrate graphs. This is not to say that graph theory is useless in the study of metabolism. On a large scale, metabolic networks are different from the null-model networks. This is a valid conclusion from Ref. [4]—the authors manage to separate the real metabolic networks from both the null-model network we use, and various other types of model networks derived from the original graph. We believe much information about the organization of metabolism lies in the answer to how this separation occurs. On smaller scales, network theory can be used to find, e.g., functional modules [10] and functions of individual metabolites [14].

P.H. acknowledges financial support from the Wenner-Gren Foundations and National Science Foundation (Grant No. CCR-0331580). The authors are grateful for comments from Jing Zhao and Martin Rosvall.

[1] C. Giersch, Eur. J. Biochem. **174**, 509 (1988).
[2] D. A. Fell, *Understanding the Control of Metabolism* (Portland Press, Miami, 1997).
[3] J.-H. S. Hofmeyr and A. Cornish-Bowden, FEBS Lett. **476**, 47 (2000).
[4] C. Marr, M. Müller-Linow, and M.-T. Hütt, Phys. Rev. E **75**, 041917 (2007).
[5] Y. Togashi and K. Kaneko, J. Phys.: Condens. Matter **19**, 065150 (2007).
[6] O. Brandman, J. E. Ferrell, R. Li, and T. Meyer, Science **310**, 496 (2005).
[7] S. A. Kauffman and K. Glass, J. Theor. Biol. **39**, 103 (1973).
[8] H. Ma and A.-P. Zeng, Bioinformatics **19**, 270 (2003).
[9] J. Zhao, H. Yu, J. Luo, Z. W. Cao, and Y.-X. Li, Chin. Sci. Bull. **51**, 1529 (2006).
[10] M. Huss and P. Holme, IET Syst. Biol. **1**, 280 (2007).
[11] H. Gutfreund, *Kinetics for the Life Sciences* (Cambridge University Press, Cambridge, U.K., 1995).
[12] J. Gómez-Gardeñes, Y. Moreno, and L. M. Floría, Physica A **352**, 265 (2005).
[13] A. Wagner, *Robustness and Evolvability in Living Systems* (Princeton University Press, Princeton, NJ, 2005).
[14] R. Guimerà and L. A. Nunes Amaral, Nature (London) **433**, 895 (2005).