OPTIMAL METABOLIC REGULATION USING A CONSTRAINT-BASED MODEL

WILLIAM J. RIEHL[1]
briehl@bu.edu

DANIEL SEGRÈ[1,2]
dsegre@bu.edu

[1]*Graduate Program in Bioinformatics, Boston University, 44 Cummington St., Boston, Massachusetts, 02215, USA*
[2]*Departments of Biology and Biomedical Engineering, Boston University, 24 Cummington St., Boston, Massachusetts, 02215, USA*

Regulation of metabolic enzymes plays a crucial role in the maintenance of metabolic homeostasis, and in the capacity of living systems to undergo physiological adaptation under multiple environmental conditions. Metabolic regulation is achieved through a complex interplay of transcriptional and post-transcriptional mechanisms, some of which have been experimentally characterized for specific pathways and organisms. Many of the details, however, including the values of most kinetic parameters, have proven difficult to elucidate. Hence, understanding the principles that underlie metabolic regulation strategies constitutes an ongoing challenge. In the context of genome-scale steady state models of metabolic networks, it has been shown that evolution may drive metabolic networks towards reaching computationally predictable optimal states, such as maximal growth capacity. Here we develop a new computational approach based on the hypothesis that the regulatory systems operating on metabolic networks have evolved towards an optimal architecture as well. Specifically, we hypothesize that the topology of metabolic regulation networks has been selected for optimally maintaining the system balanced around one or more steady states. Based on these hypotheses, we use methods related to flux balance analysis to construct a model of metabolic regulation based primarily on a metabolic network's topology, bypassing the requirement for the details of all kinetic parameters. This model predicts an optimal regulatory network of metabolic interactions that can resolve perturbations to a given steady state in a metabolic system. We explore the ability of the model to predict optimal regulatory responses in both a simple toy network and in a fragment of the well-described glycolysis pathway.

Keywords: metabolic regulation; flux balance analysis; enzyme kinetics; metabolism; optimality; logistic map; chaos

1. Introduction

Genome-scale stoichiometric models of cellular metabolism, such as flux balance analysis (FBA), can provide predictions of mass flow through metabolic networks in a population of cells under steady state conditions [3, 12]. While some recent stoichiometric models and FBA predictions include Boolean regulatory expressions [2, 14], understanding how to best formulate a joint modeling framework for metabolism and its regulatory control constitutes an important ongoing challenge. Regulatory networks allow cells to undergo physiological changes in response to dynamically changing environmental conditions, or possibly to cope with externally imposed genetic modifications (e.g. gene knockouts). In addition, even at unperturbed steady states, regulatory networks may have an important role in ensuring homeostatic stability against stochastic noise [8]. In particular, one might expect that if a metabolic network is

perturbed slightly out of steady state, the regulatory system should help quickly restore homeostasis [10, 11]. Although much is known about transcriptional and post-transcriptional regulation in several central metabolic pathways, a lot of regulatory interactions, and most parameters, still remain to be discovered. One may therefore ask whether it is possible to develop a mathematical framework to study metabolic regulation even in the absence of detailed parameters and of specific experimental knowledge of all interactions. Along these lines, previous work has analyzed the evolution of metabolic networks and their regulation, especially as it pertains to optimality of construction and use of resources [4]. If metabolism has evolved over time to be optimal for several different functions – e.g. generation of biomass, reliable transduction of energy, production of signals and antibiotics – then it is likely that the regulatory mechanisms that control these processes have also evolved to be optimal for the maintenance of different metabolic states. In this work we approach the question of how metabolism is regulated by proposing a model of optimal metabolic regulation that predicts the regulatory capacity of metabolites in a network by using the reactions involved in the system.

2. A Constraint-based Model of Regulation

In genome-scale stoichiometric models of metabolism, the metabolic network is represented as a stoichiometric matrix, S, where each element S_{ij} represents the number of moles (or molecules) of metabolite i produced (positive sign) or consumed (negative sign) in reaction j. If the metabolic network is at steady state, then the system can be described by the following set of equations:

$$\frac{dX_i}{dt} = \sum_{j=1}^{N} S_{ij} v_j = 0 \qquad i = 1,2,...,M \tag{1}$$

where X_i is the concentration of metabolite i, v_j is the flux through reaction j, N is the total number of reactions, and M is the total number of metabolites. The flux v_j represents the rate at which reaction j proceeds at steady state. The approach of flux balance analysis (FBA) takes advantage of the fact that the steady state approximation transforms these nonlinear differential equations in the concentrations into linear algebraic equations in the fluxes. The space of feasible flux states identified by these linear constraints is further restricted by linear inequalities that define nutrient availability, followed by a Linear Programming (LP) search for a set of fluxes that is optimal for a given linear objective function (for a more detailed introduction to FBA, see [5]).

Deviating from the traditional formulation of a flux balance model, we ask now how to describe the regulatory response to a metabolic perturbation that takes the system out of its original steady state. In particular, we ask how the regulatory system optimally restores homeostasis. Equation (1) describes the steady state fluxes in an FBA model. If the fluxes are not in steady state, the equations need to take into account the time dependent accumulation (or depletion) of metabolite pools:

$$\frac{dX_i}{dt} = \sum_j S_{ij} v_j = E_i(t) \tag{2}$$

where $E_i(t)$ represents the excess production (or consumption) of metabolite i at time t. If a control mechanism exists to return the system to a steady state, it would act to change the fluxes so that after some time Δt, the excess production will become zero. In terms of our variables, this would mean that a flux correction Δv, which is a function of the excess production E, should modify the flux distribution v, yielding a regulated set of fluxes v_r:

$$v_r = v + \Delta v \tag{3}$$

Our goal will be to find how regulation should cause such a change Δv to occur.

We base our model of regulation on the hypothesis that the system's control mechanism can detect and respond to this overproduction through allosteric and kinetic effects. If the perturbations to the fluxes are small, it is possible that the rapid control conferred by metabolic regulation (allosteric effects, feedback inhibition, cofactor activation, etc.) will restore homeostasis in an optimal manner [7]. In the simplest formulation of our model, we assume that the overproduction of certain metabolites can directly regulate the fluxes that produce them. Specifically, we implement a feedback mechanism limited to non-competitive inhibition without transcriptional effects. To explain how this may work, we start with an analysis of the simple uni-directional Michaelis-Menten equation:

$$v = \frac{V_{\max} X}{K_m + X} \tag{4}$$

where v is the reaction rate (analogous to the fluxes from FBA), $V_{\max}$ is the maximum reaction rate, X is the concentration of substrate, and K_m is the Michaelis constant. If an inhibitor molecule with concentration I, and constant of inhibition K_I is included in the system, affecting this enzymatic reaction in a non-competitive way, the Michaelis-Menten Equation (4) is rewritten as

$$v = \frac{V_{\max} X}{K_m + X} \left(\frac{1}{1 + I/K_I} \right) \tag{5}$$

Thus, the change in flux Δv caused by the presence of the inhibitor I can be written as:

$$\Delta v = v_r - v = \frac{V_{\max} X}{K_m + X} \left(\frac{-I/K_I}{1 + I/K_I} \right) \tag{6}$$

If the concentration of inhibitor present is considerably less than its activity, then we can further reduce (6) to:

$$\Delta v = \frac{V_{\max} X}{K_M + X}\left(-\frac{I}{K_I}\right) = v\left(-\frac{I}{K_I}\right) \tag{7}$$

In our regulation model we are going to assume that any metabolite *j* could potentially act as a regulator for any flux *i*, based on a kinetic law similar to the inhibitor effect described in Equation (7). In other words, we will assume that the flux correction is proportional to the amount of regulating metabolite. Note that because we assume that the metabolites can act as general regulators, they can have both an activator effect (positive value) as well as an inhibitory effect (negative value). Hence, we extend Equation (7) to define the regulatory change as the cumulative effect of all metabolites acting on all fluxes based on a matrix Λ of weights, to be determined.

$$\Delta v_i = v_i \sum_{j=1}^{M} \Lambda_{ij} X_j \qquad i = 1,2,...,N \tag{8}$$

Λ_{ij} is the element (*i*, *j*) of the *N* x *M* matrix Λ, representing the regulatory effect that metabolite *j* has on flux *i*. This matrix element is related to the classical definition of the constant of inhibition by the following equation:

$$\Lambda_{ij} = -\frac{1}{K_{I,ij}} \tag{9}$$

The question we are going to focus on is whether, given specific perturbed flux states, we can infer a regulatory matrix Λ leading to a Δv that brings the network back to a steady state (Equation 3). One problem we need to face is that we do not necessarily know the concentrations of each metabolite X_j. However, we do assume the knowledge of the perturbed non-steady state fluxes (i.e. the excess production *E*). From Equation (2) we know that $\Delta X_i = E_i \Delta t$. If we simplify our model by taking a fixed Δt set to unity, and by assuming (see also Discussion) that the regulatory response depends on the concentration change (ΔX_i) rather than on the total concentration (X_i), we can rewrite Equation (9) as:

$$\Delta v_i = v_i \sum_{j=1}^{M} \Lambda_{ij} E_j \qquad i = 1,2,...,N \tag{10}$$

or, in matrix form, as:

$$\Delta v = \Lambda E v \tag{11}$$

Inserting Equation (11) into Equation (3) we obtain an explicit expression for the newly regulated flux state v_r as a function of the perturbation and the regulatory matrix:

$$v_r = v + \Lambda E v \tag{12}$$

or, in a component-wise form,

$$v_{r,i} = v_i + v_i \sum_{j=1}^{M} \Lambda_{ij} \left(\sum_{k=1}^{N} S_{jk} v_k \right) \qquad i = 1,2,...,N \tag{13}$$

The matrix of putative regulatory interactions, Λ, therefore describes the regulatory effect that all metabolites in a network could have on all reactions. However, it is unlikely that all metabolites can regulate all reactions. Rather, we may expect that evolutionary adaptation may have shaped the regulatory network to use only a small set of optimally useful regulatory interactions. We use LP to predict these interactions, in the form of the matrix Λ, subject to different objectives and constraints.

We define two different possible linear constraints for the mode of optimal regulation: the perturbed system could either be regulated to reach any closely available steady state, or to more specifically go back and restore the original unperturbed steady state. Each of these strategies could have a biological relevance, depending on whether general stability or a specific set of fluxes is functionally advantageous.

We also define three possible linear programming objectives for identifying the optimal Λ matrix. First, one can minimize the number of regulatory interactions (i.e.: nonzero elements of Λ). This would concentrate the regulatory power of metabolites in the network to just a few key players. However, it may be best to minimize the overall regulatory effort, even if this implies using a non-minimal number of regulatory arrows. We search these types of optima in two different ways: either we minimize the sum of the absolute values of the elements of Λ, or the sum of squares of the values of Λ. Both of these may reduce the regulatory ability of any single metabolite, distributing the interactions among different elements in the system.

3. Model Results

3.1 *Toy linear metabolic pathway*

We initially applied this algorithm to very simple metabolic networks to be able to explore exhaustively all possible modes of regulation. The simple model we used was one of a linear metabolic pathway of two metabolites and three uni-directional reactions (Figure 1). Any steady state in this pathway will be such that all fluxes must have the same value: in this case the initial steady state each flux has a value of 1 mmol/gDWh (i.e., $v_i = 1$, for all i).

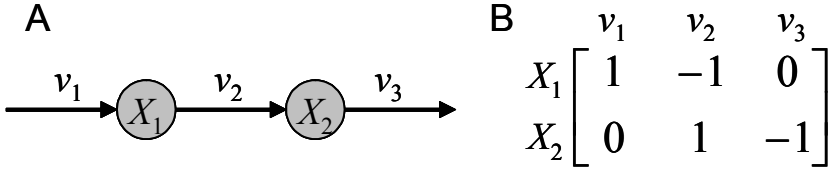

Figure 1: Linear metabolic pathway for testing. (a) Map of model. (b) Stoichiometric matrix of model.

We model perturbations as changes to one or more values of the fluxes, moving the system out of steady state. An optimal regulatory response would entail the minimal amount of feedback necessary to restore the fluxes through the system to steady state. We performed multiple perturbations to the linear system and gathered the resulting regulatory changes that occurred.

3.2. *Single flux perturbations*

The set of perturbations explored involved a perturbation to a single flux in the system. Each flux was individually increased and decreased by 10%, and Λ was calculated using all combinations of the above methods. We found that the resulting regulatory structures had two overall approaches to control (Figure 2). If we calculate a Λ to resolve both perturbations at once with the goal of reaching any steady state, the control exerted acts on all fluxes in the system except for the one perturbed, adjusting flow through the system to match the perturbation. (Figure 2A, top row) However, when we calculated Λ to return to the original steady state, there were two different effects. First, we found that only one perturbation at a time could be resolved – because of the linear dependence of the two perturbed vectors, there is no single Λ that can resolve both an increase and a decrease to the same initial steady state. Second, we found that regardless of the size of the perturbation, the same regulatory structure is encountered. In either case, the metabolite that is being overproduced (or overconsumed) is predicted to exert control over the fluxes that produce (consume) it.

With regard to the different optimization objectives used, we find that when either the first or last flux (v_1 or v_3) was perturbed, the method used to optimize the regulatory network was irrelevant: in all cases, the same regulatory network was predicted. However, when the second flux was perturbed, the different objectives yielded different, although related, regulatory networks (Figure 2B). This suggests that there could be multiple optimal regulatory schemes that would have the same regulatory effect on the network through different mechanisms. These could be a feedback system (as described by the minimization of the number of regulatory interactions), a feedforward system (in the minimization of the sum of absolute values of Λ) or a combination of both (in the minimization of the sum of squares of Λ). Any of these mechanisms has the potential to restore homeostasis and, depending on the metabolic pathway, may be most appropriate under different experimental conditions.

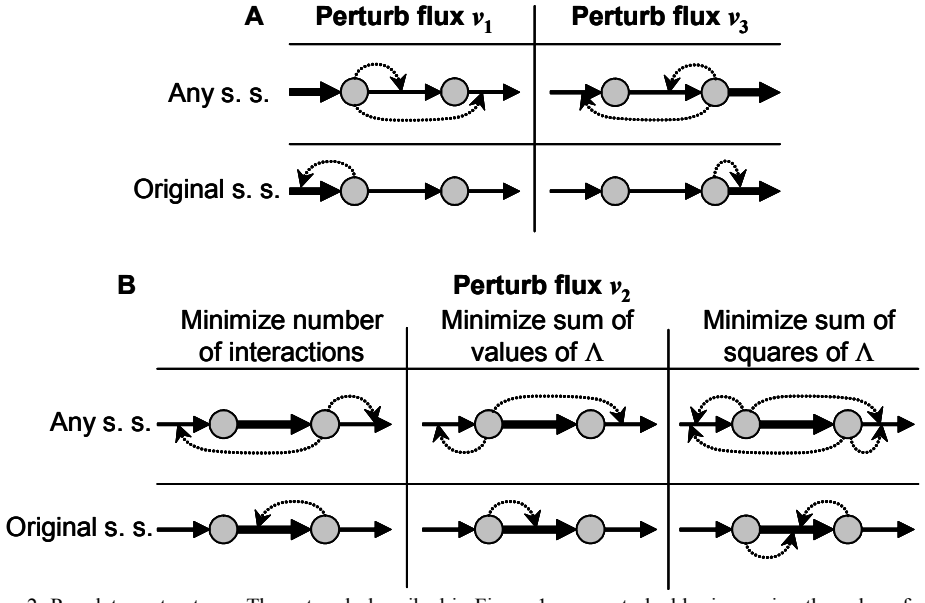

Figure 2: Regulatory structures. The network described in Figure 1 was perturbed by increasing the value of a single flux. Bold lines indicate perturbed fluxes. Separate Λs were calculated using each of the optimization schemes and objectives described in the text. (A) When fluxes v_1 or v_3 are perturbed, each optimization scheme yields the same regulatory structures (dashed edges). These structures differ based on the regulatory objective, whether they seek any steady state (top) or to return to the original steady state (bottom). (B) When flux v_2 is perturbed, each combination of objective and optimization scheme yields a different regulatory structure.

3.3. *Single perturbation robustness*

Next, we expand on the method of restoring a perturbation to a given steady state by studying the robustness of such a mechanism. Given a Λ constructed around restoring a single perturbed flux to a given steady state, how well does the same mechanism perform against different perturbations to the same flux? We approach this question by perturbing only v_3 (from the model used previously, Figure 1). A single Λ was calculated, then applied to several different perturbations of v_3, with perturbed values ranging from -0.75 to 3.25 (with a value of 1 being the target steady state). We find that when applying one regulatory system to multiple perturbations, the relationship between the perturbed flux and the resulting regulated flux varies quadratically (Figure 3). Robustness is also observed for any Λ: we found a range of perturbed fluxes such that when regulation with a single Λ is applied to them, they approach the steady state solution. However, if a perturbation is out of that range, the application of regulation will move the system even further from steady state. This range appears to be calculable based on the target steady state of the system and the perturbation used to calculate Λ.

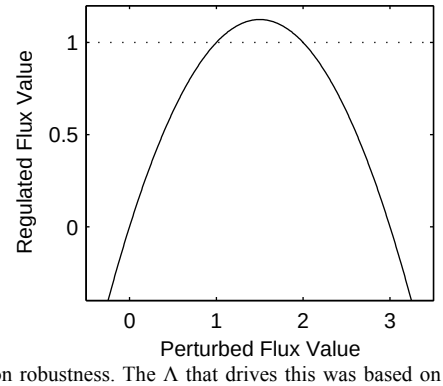

Figure 3: Single perturbation robustness. The Λ that drives this was based on a perturbation of flux v_3 of the linear pathway described in Figure 1A. The steady state value of regulation is 1 (dotted line) and the perturbation on which Λ was calculated is 2. This flux was then perturbed at values between -0.75 and 3.25 (x-axis) and values after regulation were calculated (y-axis). Note that both "perturbations" of 1 and 2 are regulated to the same steady state value. Perturbation values between 0 and 3 all approach the steady state value of 1 after regulation. Perturbation values outside of this range move further from the steady state solution.

3.4. *Single flux perturbation trajectories*

As noted above, after applying a regulatory scheme to a perturbation different than the perturbation for which Λ was optimized, Λ can still regulate the perturbation by either bringing it closer or moving it further from steady state. This leads to the hypothesis that after several iterations of applying the same Λ to the adjusted fluxes, the perturbation may either be fully dampened, or diverge. For certain ranges of perturbation (as described above), this is indeed the case. In Figure 4, we plot the trajectory of several iterations of this regulation, which can be thought of as representing a dynamical process of metabolic regulation. This can be described by the recurrence relation:

$$v_n = v_{n-1} + \Lambda E_{n-1} v_{n-1} \tag{14}$$

where n indicates the time iteration step. For example, v_0 is the original perturbed flux, while v_n denotes the n^{th} application of the regulation described in Equation (13).

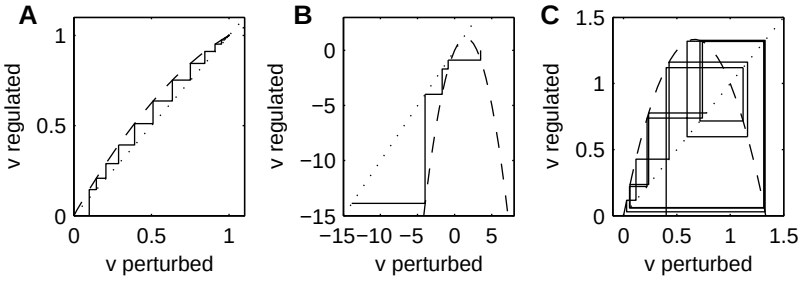

Figure 4: Single perturbation trajectory. Λ matrices were calculated to restore steady state to perturbations to the flux v_3 in the network in Figure 1. Plots A and B use a Λ calculated to return a perturbation of 2 to 1 while plot C uses a Λ that returns a perturbation of 1/3 to 1. In each plot, the dotted line is the diagonal, the dashed line is the parabola described in Figure 3, and the solid line is the trajectory after several iterations of Equation (14). (A) Convergent regulation. A perturbed flux value of 0.1 will return to steady state after several regulatory steps. (B) Divergent regulation. A perturbed flux value of 3.5 will approach $-\infty$. (C) Chaotic regulation. For some values of Λ and initial perturbations (here, the initial perturbation is 0.4), any regulation performed may behave chaotically, never converging on a steady state or diverging toward infinity.

This dynamical regulation process behaves similar to a logistic map [13], displaying regimes of convergence, divergence or apparent chaotic trajectories, depending on the values of the parameters Λ and v. With regard to metabolic regulation this finding potentially implies that chaotic or divergent behavior might be easily encountered by regulatory networks, unless specific ranges of parameters are avoided. This may pose constraints on possible regulatory networks optimized through evolutionary adaptation.

4. Glycolysis

An obvious question is whether our method can be used to predict the topology and dynamics of regulation in real-world networks. As a simple example, we chose a simplified (condensed) version of the glycolytic pathway, previously used for similar testing of computational approaches (Figure 5) [15]. Similarly to what done for the simple linear pathways (Figure 2), we approach this network by perturbing each flux individually and predicting the optimal network to restore homeostasis.

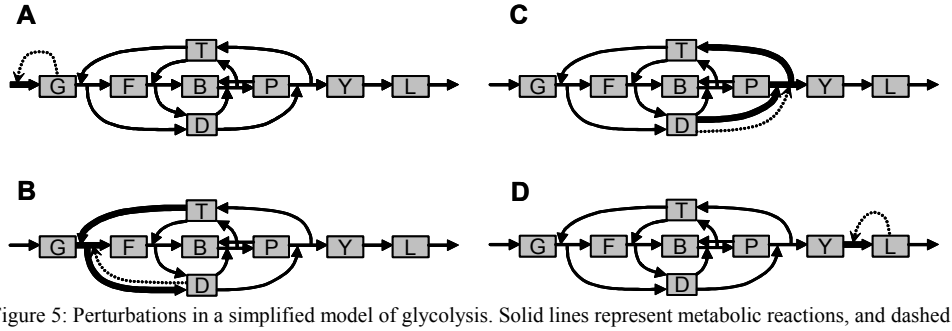

Figure 5: Perturbations in a simplified model of glycolysis. Solid lines represent metabolic reactions, and dashed lines represent predicted optimal metabolic regulation. Reactions represented as bold lines are the ones being perturbed. G = glucose, F = fructose-6-phosphate, B = fructose-1,6-bisphosphate, P = phophoenolpyruvate, Y = pyruvate, L = lactate, T = adenosine triphosphate, D = adenosine diphosphate.

In all cases, only one regulatory metabolite was necessary for optimal regulation that restores a given steady state. For each of the reactions involving an energy-carrier, ADP was predicted to act as the main regulatory molecule (Figures 5B, 5C, and data not shown). Lactate also acted as a negative feedback regulator on its own production (Figure 5D), and glucose acted as a negative regulator on the influx of glucose (Figure 5A).

5. Discussion

In this work we developed new algorithms and methods for predicting optimal metabolic regulation based on the topology and stoichiometry of a metabolic network. Thus far, we have applied these algorithms to small pathways that are linear in nature in order to understand how accurate and robust the predictions are. Initially we found that while a single regulatory scheme can be robust for some perturbed values (Figures 3 and 4), it quickly becomes clear that a single regulatory approach predicted by this method is incapable of effectively regulating all perturbations. For example, a regulatory scheme focused on regulating perturbations to a single flux will have little or no effect on other fluxes. We also observed that multiple applications of a single regulatory system can produce unexpected, apparently chaotic results (Figure 4C). While some of these results may be unrealistic consequences of the mathematical approximations used, they may also capture some fundamental properties of biological regulation systems evolved to respond to multiple perturbations. Recent work has shown, for example, that some metabolic states are more stable than others, and that perturbations occurring on top of unstable states can lead to cell death [9].

It is worth emphasizing that each of these predicted optimized regulatory mechanisms represents just that: the optimal amount of regulation necessary to respond to a given perturbation. In all cases explored (perturbations to a single flux in the network), the optimal controlling metabolite turns out to be either a reactant or product in the perturbed reaction. However, it remains a point of interest that for many perturbations in glycolysis, the controlling metabolite predicted most often was ADP. This is interesting because both ADP and ATP are known to be strong regulators (either activators or inhibitors) of glycolysis. This may point to the utility of this method as both a quantitative (degree of regulation necessary) and a qualitative (type of metabolite functioning as a regulator) prediction generator.

The current model involves simplifying hypotheses and approximations, some of which may be unjustified from the biochemical point of view. These include the assumption that the regulatory response is based on concentration changes, rather than absolute concentration values; the fact that we do not include flux relaxation induced by plain kinetic effects; the use of arbitrary values for flux perturbations; the implementation of a dynamical process based on discrete time points; and the limitation to non-competitive inhibition as the only form of feedback. In ongoing work, we are addressing each of these assumptions to determine their impact on our results, and possible

strategies for more realistic implementations. We plan to expand on this work and use it to explore more complex systems. At first, we will use this method to understand how it predicts regulation of different and multiple perturbations to a system. We expect that when two or more fluxes are perturbed, the regulatory network will quickly become complex and intricate. Next, we plan to explore the regulation of networks with complex topologies that include branching and cyclical pathways. Eventually we intend to apply this predictive method to whole-genome models of flux balance, such as the *Escherichia coli* model produced by Feist *et al*. [6] or the *Saccharomyces cerevisiae* model produced by Blank, *et al*. [1].

Acknowledgements

The authors wish to thank Hsuan-Chao Chiu, Niels Klitgord, and Evan Snitkin for meaningful discussion and critical reading of the manuscript. Linear Programming calculations were performed using the software Xpress, kindly provided by Dash Optimization under free academic license. This work was partially supported by the NASA Astrobiology Institute, the US Department of Energy and the US National Institutes of Health (NIGMS).

References

[1] Blank, L.M., Kuepfer, L. and Sauer, U., Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast, *Genome Biol*, 6(6):R49, 2005.

[2] Covert, M.W., Schilling, C.H. and Palsson, B., Regulation of gene expression in flux balance models of metabolism, *J Theor Biol*, 213(1):73-88, 2001.

[3] Covert, M.W. and Palsson, B.O., Constraints-based models: regulation of gene expression reduces the steady-state solution space, *J Theor Biol*, 221(3):309-25, 2003.

[4] Ebenhöh, O. and Heinrich, R., Stoichiometric design of metabolic networks: multifunctionality, clusters, optimization, weak and strong robustness, *Bull Math Biol*, 65(2):323-57, 2003.

[5] Edwards, J.S. and Palsson, B.O., Metabolic flux balance analysis and the in silico analysis of Escherichia coli K-12 gene deletions, *BMC Bioinformatics*, 1(1, 2000.

[6] Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V. and Palsson, B.O., A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information, *Mol Syst Biol*, 3(121, 2007.

[7] Fell, D., Understanding the Control of Metabolism, Portland Press Ltd., 1997.

[8] Goyal, S. and Wingreen, N.S., Growth-induced instability in metabolic networks, *Phys Rev Lett*, 98(13):138105, 2007.

[9] Grimbs, S., Selbig, J., Bulik, S., Holzhutter, H.G. and Steuer, R., The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks, *Mol Syst Biol*, 3(146, 2007.

[10] Hatzimanikatis, V., Floudas, C.A. and Bailey, J.E., Optimization of regulatory architectures in metabolic reaction networks, *Biotechnology and Bioengineering*, 52(4):485-500, 1996.

[11] Heinrich, R. and Rapoport, T.A., A linear steady-state treatment of enzymatic chains. General properties, control and effector strength, *Eur J Biochem*, 42(1):89-95, 1974.

[12] Kauffman, K.J., Prakash, P. and Edwards, J.S., Advances in flux balance analysis, *Curr Opin Biotechnol*, 14(5):491-6, 2003.

[13] May, R.M., Simple mathematical models with very complicated dynamics, *Nature*, 261(5560):459-67, 1976.

[14] Shlomi, T., Eisenberg, Y., Sharan, R. and Ruppin, E., A genome-scale computational study of the interplay between transcriptional regulation and metabolism, *Mol Syst Biol*, 3:101, 2007.

[15] Vance, W., Arkin, A. and Ross, J., Determination of causal connectivities of species in reaction networks, *Proc Natl Acad Sci U S A*, 99(9):5816-21, 2002.