16th European Symposium on Computer Aided Process Engineering
and 9th International Symposium on Process Systems Engineering
W. Marquardt, C. Pantelides (Editors)
© 2006 Published by Elsevier B.V.

THE SYSTEMS ENGINEERING OF CELLULAR PROCESSES

Vassily Hatzimanikatis* and Liqing Wang
Northwestern University
Evanston, IL 60208

Abstract

Mathematical description of metabolic systems allows the calculation of the expected responses of metabolism to genetic modifications and the identification of the most promising targets for the engineering of cellular processes. Metabolic control analysis (MCA) provides such a description in the form of sensitivity coefficients, called control coefficients. These indices are determined by perturbation experiments, or through (log)linear analysis of nonlinear mathematical models, around a reference steady state, and, therefore, the predictive power of MCA is limited to small changes in the metabolic parameters. We present here the Nonlinear Metabolic Control Analysis (NMCA), a framework that allows accurate description of the metabolic responses over wide range of changes in the metabolic parameters. The performance and capabilities of NMCA are illustrated using a model of the yeast glycolysis.

Keywords

Metabolic engineering, (log)linear model, metabolic control analysis, nonlinear metabolic control analysis, glycolysis.

Introduction

Complex systems are composed of interacting elements that give rise to properties which cannot be predicted by knowing the properties of their constituting elements alone (Ottino 2003). Metabolic pathways are such complex systems, where many different enzymes and proteins interact through common reactants, substrates, cofactors, and regulators to yield properties which cannot be described simply by the knowledge of the individual enzymes in the pathways. However, understanding the emerging properties of metabolic pathways is central to problems and questions in several life sciences disciplines, such as medicine, evolutionary biology, systems biology, and metabolic engineering. On the other hand, there is an exponential growth in the development of new technologies that provide a very good qualitative and quantitative description of metabolic systems. Genome sequencing and metabolic reconstruction allow the identification of the stoichiometry and pathway topology, and metabolic flux analysis (MFA) methodologies, developed within the field of metabolic engineering, provide with estimates of the fluxes (reaction rates) in metabolic networks (Forster et al. 2003; Varma and Palsson 1993a; Varma and Palsson 1993b). In this article we employ a mathematical and computational framework that allows the integration of the available information from these fields and the prediction of the responses of metabolic networks to changes in metabolic parameters.

From a systems perspective, metabolite concentrations and reaction rates are the *metabolic variables*, stoichiometry, enzyme kinetic parameters, and environmental conditions are the *metabolic parameters*. In studying the systemic origins of disease, or in identifying targets for genetic and metabolic engineering, we are interested in understanding how changes in metabolic parameters impact the variables of a metabolic system. Such an understanding will require a quantitative characterization of these effects, and the identification of the most important parameters, e.g., the parameters that have the highest impact on certain metabolic variables. The latter is what is in many cases called the *rate-limiting steps* in a metabolic pathway.

Quantitative characterization of this kind can be achieved either through sophisticated experiments that involve perturbation of the metabolic parameters and precise measurements of possibly every metabolic variable, or through simulations of detailed kinetic models of the metabolic pathways. However, both of these approaches have certain drawbacks. Namely, the experimental approach requires not only a large number of experiments, but also many advanced analytical techniques. Yet still, they are subject to a great degree of uncertainty. And for the computational approach, construction of such models entails a large amount of information, including but not limited to kinetics of the enzymes of

---

[*] To whom all correspondence should be addressed

the pathway and metabolite concentrations. Therefore a quantification framework that requires a minimum amount of information about enzyme kinetics and takes into account of the inherent uncertainty will significantly improve our understanding of the functions of metabolic pathways.

Among the developed mathematical frameworks, Metabolic Control Analysis (MCA) quantifies the link between genetic modifications or environmental changes and cellular process responses. MCA introduces the *control coefficients*, similar to the concept of transfer function in the system control theory, to quantify the fractional change of cellular output (metabolite concentrations and metabolic fluxes) in response to a fractional change of system parameters (e.g. enzymatic activities and growth conditions) (Fell and Sauro 1985; Hatzimanikatis and Bailey 1996; Hatzimanikatis and Bailey 1997; Heinrich and Rapoport 1974; Kacser and Burns 1973; Kholodenko and Westerhoff 1993; Reder 1988). An immediate application of MCA on rational metabolic engineering design is the ranking of potential targets based on the values of the control coefficients of the flux leading to the desired cellular product (Bowden 1999; Cascante et al. 2002; Schuster 1999; Westerhoff and Kell 1996). However, MCA, being a “local sensitivity” analysis framework, does not guarantee an accurate estimation of the system responses to large changes in enzyme activities and growth environment.

In order to address this issue we have previously developed a method called Nonlinear Metabolic Control Analysis (NMCA) that allows prediction of metabolic responses to large changes in metabolic parameters based on information about the properties of the system around a reference steady state (Hatzimanikatis 1999). We apply here this method on yeast glycolysis and illustrate how this method is applicable to the cases where partial information about the kinetics of the system is available.

Generalized (Log)linear MCA Formalism

For cells growing in a batch culture, the mass balances of intracellular metabolites can be described as

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{x}, \boldsymbol{p}_e, \boldsymbol{p}_s), \tag{1}$$

where $\boldsymbol{x}$ is the metabolite concentration vector, $\boldsymbol{N}$ is the stoichiometric matrix, $\boldsymbol{v}$ is the metabolic flux vector, $\boldsymbol{p}_e$ is the enzyme activity parameter vector, which includes both kinetic parameters and enzyme concentrations, and $\boldsymbol{p}_s$ is the vector of other system parameters such as temperature and pH.

Due to the presence of conserved moieties in the cellular metabolism, i.e. groups of compounds such as ATP, ADP, and AMP, whose total amount is assumed to remain invariant over the characteristic response time of the metabolic network, we can divide the original set of metabolite concentrations $\boldsymbol{x}$ into two categories: an independent metabolite concentration vector, $\boldsymbol{x}_i$, and a dependent metabolite concentration vector, $\boldsymbol{x}_d$ (Reder 1988). A third parameter set, $\boldsymbol{p}_m$, is also introduced into the system to represent the total concentration of the metabolites in each moiety group (Wang et al. 2004). The reduced form of mass balances with respect to independent metabolites can be represented as

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{N}_R\boldsymbol{v}(\boldsymbol{x}_i, \boldsymbol{x}_d(\boldsymbol{x}_i, \boldsymbol{p}_m), \boldsymbol{p}_e, \boldsymbol{p}_s), \tag{2}$$

where $\boldsymbol{N}_R$ consists of the rows in $\boldsymbol{N}$ corresponding to the independent metabolites.

In the previous work, we have demonstrated the calculation of control coefficients in a intracellular metabolic system based on the (log)linear model formalism (Hatzimanikatis and Bailey 1996; Hatzimanikatis and Bailey 1997; Hatzimanikatis et al. 1996; Wang et al. 2004). Assuming a stable steady state for the system in Equation (2), and after linearization and scaling we obtain the following equations for the control coefficients:

$$\mathbf{C}_{\mathbf{p}}^{\mathbf{x_i}} = -(\mathbf{N_R V E_i} + \mathbf{N_R V E_d Q_i})^{-1} \left[\mathbf{N_R V \Pi_m} \vdots \mathbf{N_R V \Pi_e} \vdots \mathbf{N_R V \Pi_s}\right], \tag{3}$$

$$\boldsymbol{C}_p^v = (\boldsymbol{E}_i + \boldsymbol{E}_d\boldsymbol{Q}_i)\boldsymbol{C}_p^{x_i} + \left[\boldsymbol{\Pi}_m \vdots \boldsymbol{\Pi}_e \vdots \boldsymbol{\Pi}_s\right]. \tag{4}$$

Concentration control coefficients, $\boldsymbol{C}_p^x$, and flux control coefficients, $\boldsymbol{C}_p^v$, are defined as the fractional change of metabolite concentrations and metabolic fluxes, respectively, in response to fractional changes of system parameters. In this formalism, $\boldsymbol{V}$ is the diagonal matrix whose elements are the steady state fluxes; $\boldsymbol{E}_i$ and $\boldsymbol{E}_d$ are the matrices of the elasticities with respect to metabolites, defined as the local sensitivities of metabolic fluxes to independent and dependent metabolite concentrations, respectively; $\boldsymbol{\Pi}_m$, $\boldsymbol{\Pi}_e$, and $\boldsymbol{\Pi}_s$ are the matrices of the elasticities with respect to parameters, defined as the local sensitivities of metabolic fluxes to system parameters, $\boldsymbol{p}_m$, $\boldsymbol{p}_e$, and $\boldsymbol{p}_s$, correspondingly; and $\boldsymbol{Q}_i$ is a weight matrix that represents the relative abundance of dependent metabolites with respect to the abundance of the independent ones. A second weight matrix, $\boldsymbol{Q_m}$, is also defined, for the relative abundance of dependent metabolites with respect to the levels of their corresponding total moieties, which leads to the following expression for the matrices of elasticities with respect to parameters, $\boldsymbol{\Pi_m}$:

$$\boldsymbol{\Pi_m} = \boldsymbol{E}_d \boldsymbol{Q_m} \tag{5}$$

It is worth mentioning that the estimation of control coefficients using this framework does not require explicit knowledge of the values of the concentration of metabolites.

Nonlinear Metabolic Control Analysis (NMCA)

In the previous work (Hatzimanikatis 1999), we have demonstrated how we can use the "local" information from MCA (Equations (3)-(5) above) to calculate responses to large changes in metabolic parameters. The method is based on the observation that the dependency of metabolite concentrations and metabolic fluxes can be formulated as the following ordinary differential equation problem:

$$\frac{dz_i}{dq_k} = C_{p_k}^{x_i} = f\left(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{E}\left(\boldsymbol{x}\right)\right) \tag{6}$$

$$\frac{dw_j}{dq_k} = C_{p_k}^{v_j} = g\left(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{E}\left(\boldsymbol{x}\right)\right) \tag{7}$$

where:

$$z_i = \ln \frac{x_i}{x_{i,o}} \quad \Rightarrow \quad x_i = x_{i,o} \cdot e^{z_i} \tag{8}$$

and

$$w_j = \ln \frac{v_j}{v_{j,o}} \quad \Rightarrow \quad v_j = v_{j,o} \cdot e^{w_j} \tag{9}$$

with $x_{i,o}$ and $v_{j,o}$ the initial, reference steady-state values of metabolite x_i and metabolic flux v_j, respectively.

Equations (6)-(9) can be now solved simultaneously using any ODE solver and could provide the dependency of steady-state values of metabolites and metabolic fluxes on large changes in the metabolic parameter p_k, with $p_k = p_{k,o} \cdot e^{q_k}$. For every integration point, we also check the local stability characteristic of the system in order to guarantee that the system transitions to new steady states through a sequence of *stable* steady states. In case the system crosses into an unstable steady state, we terminate the integration. Details of the method will be provided elsewhere (Wang and Hatzimanikatis, in preparation).

Analysis of the Rate Limiting Steps in Yeast Batch Fermentation

We applied the NMCA framework on the central carbon metabolism of yeast *S. cerevisiae* growing in a batch reactor based on the model by Teusink et al. (Teusink et al. 2000) (Figure 1). Teusink et al. provided information about the kinetic parameters of most of the enzymes in the system and the concentration of most of the metabolites. A total of 14 parameter and metabolite values were missing. In order to overcome this limitation we consider the following three cases:

I. The missing parameters and metabolites were chosen such that the corresponding enzymes were at high saturation at the reference, initial steady state.
II. The missing parameters and metabolites were chosen such that the corresponding enzymes were at low saturation at the reference, initial steady state.
III. The missing parameters and metabolites were chosen with all possible 2^{14} combinations of the corresponding enzymes between high and low saturation at the reference, initial steady state.

In all cases we studied the responses of the metabolites and metabolic fluxes to large changes in the activity of the glucose transport enzyme, HXT.

In case I we have been able to calculate the responses of metabolites and metabolic fluxes for a 100% change in HXT activity (Figures 2 and 3). In case II though, beyond 40% changes in the activity of HXT the system became unstable going through a saddle node bifurcation (Figures 4 and 5). This transition is also demonstrated through the sharp increase in the relative concentrations of some of the metabolites, such as PEP and PYR (Figure 4). In both cases, we compared the results from NMCA analysis with the changes in the values of metabolites and metabolic fluxes if we had used the control coefficients from the reference steady states using the following equations:

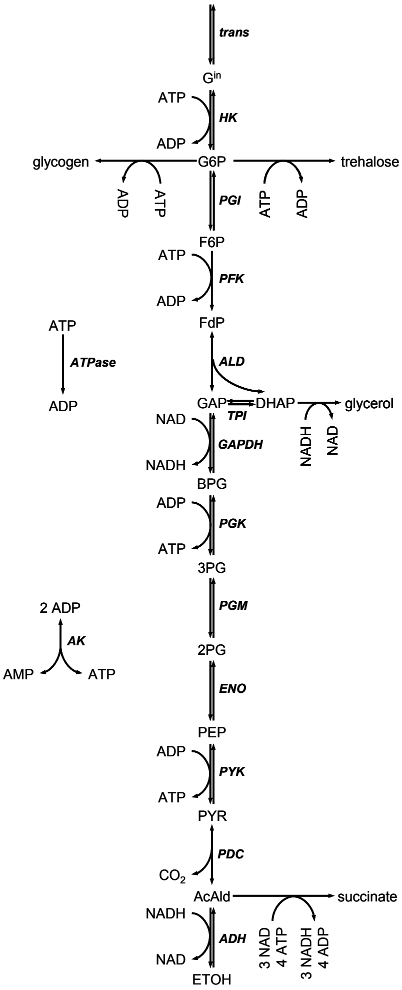

Figure 1. ***Anaerobic glycolytic pathway model of nongrowing yeast, Saccharomyces cerevisiae, with glucose as the sole carbon source.*** *Chemical species:* G^{in}*, intracellular glucose; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FdP, fructose 1,6-diphosphate; GAP, glyceraldehydes-3-phosphate; DHAP, dihydroxy acetone phosphate; BPG, bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; PYR, pyruvate; AcAld, acetaldehyde; ETOH, ethanol; ATP, adenosine triphosphate; ADP, adenosine diphosphate; AMP, adenosine monophosphate; NADH, nicotinamide adenine dinucleotide. Pathway steps and enzymes (in bold):* trans, *glucose cross-membrane transport;* HK, *hexokinase;* PGI, *phosphoglucose isomerase;* PFK, *phosphofructokinase;* ALD, *fructose 1,6-diphosphate aldolase;* TPI, *triose phosphate isomerase;* GAPDH, *glyceraldehydes-3-phosphate dehydrogenase;* PGK, *phosphoglycerate kinase;* PGM, *phosphoglycerate mutase;* ENO, *enolase;* PYK, *pyruvate kinase;* PDC, *pyruvate decarboxylase; ADH, alcohol dehydrogenase; ATPase, net ATP consumption; AK, adenylate kinase.*

$$X_* = X_{*,o} \cdot \left(P_{HXT}/P_{HXT,o}\right)^{C^X_{HXT}} \tag{10}$$

and

$$V_* = V_{*,o} \cdot \left(P_{HXT}/P_{HXT,o}\right)^{C^V_{HXT}} \tag{11}$$

In both cases the MCA prediction is quite accurate for up to 25% increase in the HXT activity. The MCA fails as the change in the activity increases, and as expected, it cannot capture the metabolite accumulation associated with the bifurcation observed in case II. It also appears that MCA can predict the fluxes much better in case II for up to almost 40%, while it fails to do so for the same range in case I. This observation suggests that the accuracy of MCA depends on the values of the metabolites at the reference, initial steady state. This effect is also illustrated in the studies of case III (Figure 6). In this case we have studied all the possible combinations of the 14 unknown metabolite concentrations that could support stable steady states up to 100% changes in the activity of HXT. In general, it appears that the MCA overestimates the responses of the fluxes relative to NMCA. Probably the most interesting result is the fact that the propagation of the uncertainty due to the uncertainty in the reference, initial parameters is not significant and it appears that this uncertainty propagation is smaller in the NMCA predictions. Future studies could provide important insights on the conditions that reduce, or increase, the uncertainty propagation.

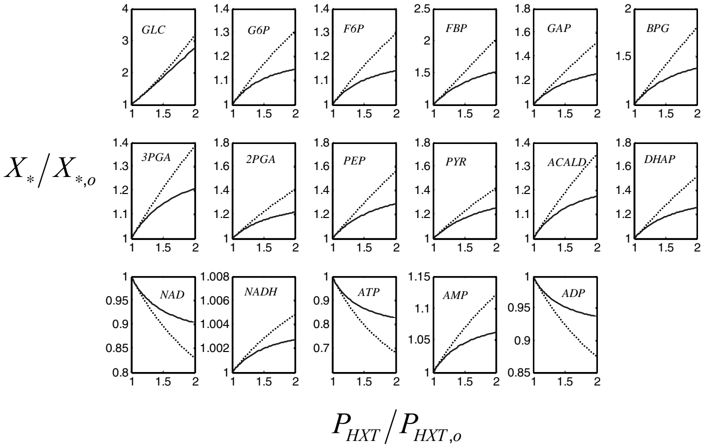

Figure 2. ***Relative changes in metabolite concentrations for 100% increase in the activity of HXT, case I.*** *Chemical species notation same as in Figure 1. Solid (blue) line: NMCA predictions; dashed (black) line: MCA predictions.*

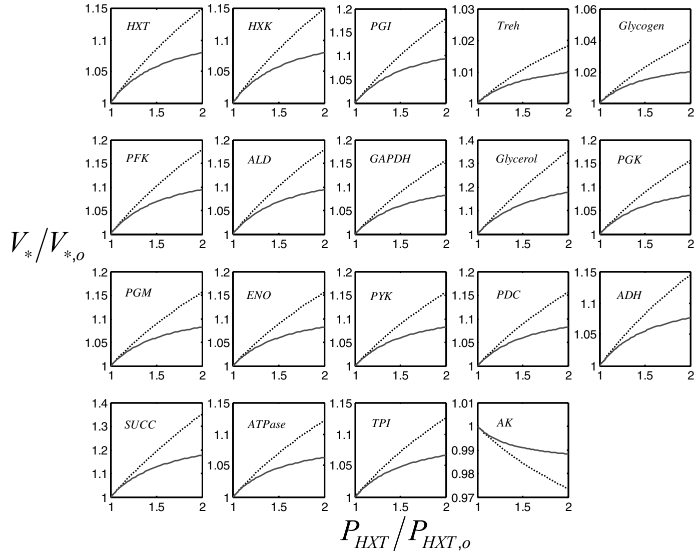

Figure 3. ***Relative changes in metabolite fluxes for 100% increase in the activity of HXT, case I.*** *Metabolic reaction notation same as in Figure 1. Solid (red) line: NMCA predictions; dashed (black) line: MCA predictions.*

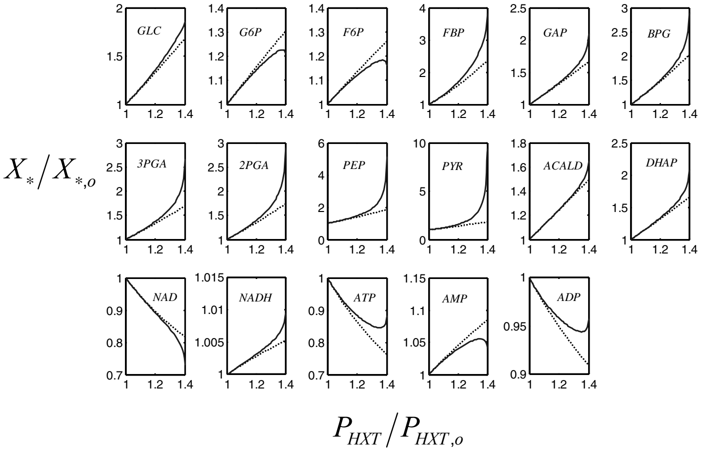

Figure 4. ***Relative changes in metabolite concentrations for 40% in the activity of HXT, case II.*** *Chemical species notation same as in Figure 1. Solid (blue) line: NMCA predictions; dashed (black) line: MCA predictions.*

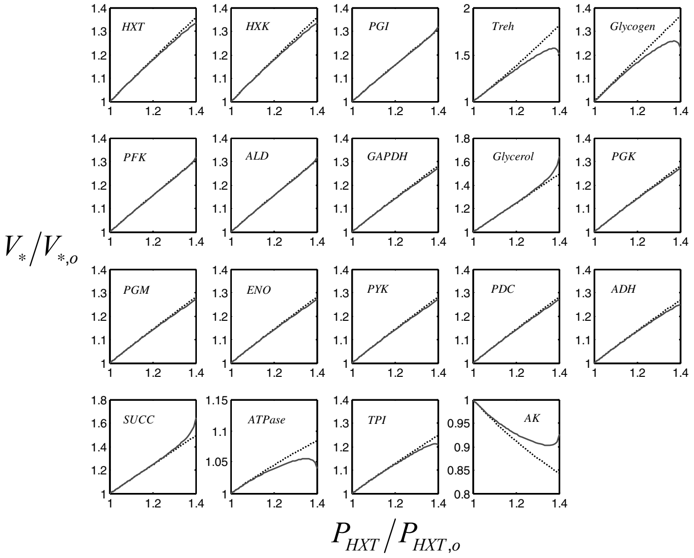

Figure 5. ***Relative changes in metabolite fluxes for 40% increase in the activity of HXT, caseII.*** *Metabolic reaction notation same as in Figure 1. Solid (red) line: NMCA predictions; dashed (black) line: MCA predictions.*

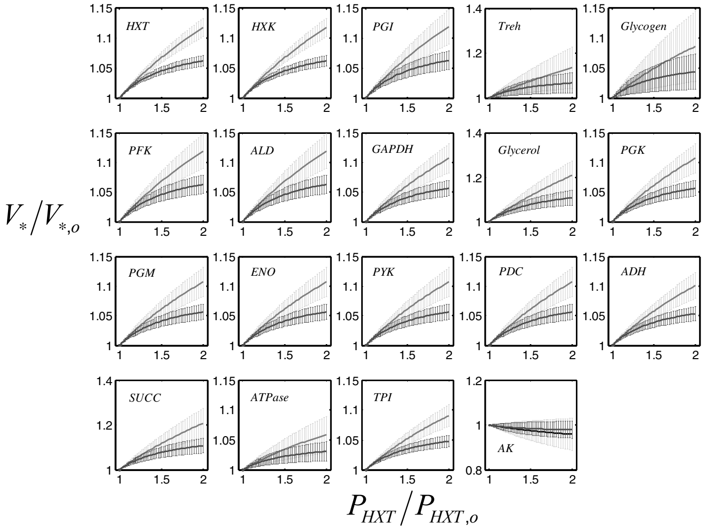

Figure 6. ***Relative changes in metabolite fluxes for 100% increase in the activity of HXT, caseIII.*** *Metabolic reaction notation same as in Figure 1. Red line: NMCA predictions; Grey line: MCA predictions. Error bars quantify the standard deviation over all the samples.*

It also expected that increase in the activity of an enzyme will lead to redistribution in the value of the control coefficients. In cases I and II, the control coefficient of ethanol production, i.e., flux through enzyme ADH, with respect to HXT is very small in the initial state and therefore there is no significant redistribution in the values of the control coefficients (Figure 7). This also explains the relative small changes in metabolic fluxes to changes in the HXT activity (Figures 1, 2, and 6).

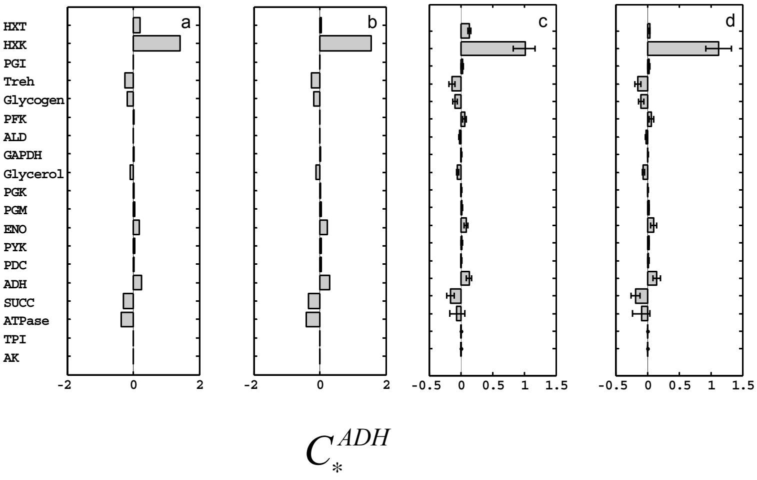

Figure 7. ***Control coefficient of ethanol production with respect to the enzymes in the network.*** *The subscript "*" correspond to the enzymes labeled on the Y-axis with notation same as in Figure 1. Panels (a) control coefficients at the reference, initial state, case I; (b) control coefficients after 100% increase in HXT activity, case I;. (c) control coefficients at the reference, initial state, case III; (d) control coefficients after 100% increase in HXT activity, case III.*

However, in case II, the initial control coefficient of ethanol production with respect to HXT activity is relative large (Figure 8a) and therefore the system experiences significant changes even with 40% increase in HXT activity (Figures 4 and 5). However, after 40% increase in HXT activity, the control coefficient of ethanol production with respect to HXT activity has vanished, and PDC became the rate limiting step as indicated by the significant increase in control coefficient of ethanol production with respect to PDC activity (Figure 8b). It also appears that in the reference state ATPase, a process associated with energy balancing in the cell, is the second most rate limiting process. However, its control also vanishes after 40% increase in HXT activity (Figure 8). Therefore NMCA analysis suggests that in a strategy to further improve the ethanol production using the overexpression of two enzymes, HXT and PDC are the most important targets. However, MCA analysis would have led us to choose HXT and ATPase, a choice that would have led to lower improvements.

Conclusions

We presented an analysis of yeast glycolysis using a metabolic control analysis framework, Nonlinear Metabolic Control Analysis (NMCA) that takes into account large changes in enzyme activities. This analysis demonstrated the power of the framework in identifying targets for genetic and metabolic engineering towards improvement of cellular processes. We also demonstrated the ability of the method to perform metabolic control analysis even in the absence of complete information about the system. Future efforts will combine NMCA with our previously developed method for metabolic control analysis under uncertainty (Wang et al. 2004a). Finally, the method appears to be ideally suitable identifying multiple targets for genetic and metabolic engineering.

Acknowledgments

The authors are grateful for the financial supported provided by the Department of Energy (DE-AC36-99GO103), the National Aeronautics and Space Administration (NAG 2-1527), and DuPont through a DuPont Young Professor Award to VH. LW received partial support by the Chinese Government through the State Excellence Scholarship program for students studying overseas.

References

Bowden AC. 1999. Metabolic control analysis in biotechnology and medicine. Nat Biotechnol 17(7):641-3.

Cascante M, Boros LG, Comin-Anduix B, de Atauri P, Centelles JJ, Lee PWN. 2002. Metabolic control analysis in drug discovery and disease. Nature Biotechnology 20(3):243-249.

Fell DA, Sauro HM. 1985. Metabolic control and its analysis. Additional relationships between elasticities and control coefficients. Eur J Biochem 148(3):555-61.

Forster J, Famili I, Fu P, Palsson BO, Nielsen J. 2003. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res 13(2):244-53.

Hatzimanikatis V. 1999. Nonlinear metabolic control analysis. Metab Eng 1(1):75-87.

Hatzimanikatis V, Bailey JE. 1996. MCA has more to say. J Theor Biol 182(3):233-42.

Hatzimanikatis V, Bailey JE. 1997. Effects of spatiotemporal variations on metabolic control: Approximate analysis using (log)linear kinetic models. Biotechnology and Bioengineering 54(2):91-104.

Hatzimanikatis V, Floudas CA, Bailey JE. 1996. Analysis and design of metabolic reaction networks via mixed-integer linear optimization. Aiche Journal 42(5):1277-1292.

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

Kacser H, Burns JA. 1973. The control of flux. Symp Soc Exp Biol 27:65-104.

Kholodenko BN, Westerhoff HV. 1993. Metabolic channelling and control of the flux. FEBS Lett 320(1):71-4.

Ottino JM. 2003. Complex systems. Aiche Journal 49(2):292-299.

Reder C. 1988. Metabolic control theory: a structural approach. J Theor Biol 135(2):175-201.

Schuster S. 1999. Use and limitations of modular metabolic control analysis in medicine and biotechnology. Metab Eng 1(3):232-42.

Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV and others. 2000. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem 267(17):5313-29.

Varma A, Palsson BO. 1993a. Metabolic Capabilities of Escherichia-Coli .1. Synthesis of Biosynthetic Precursors and Cofactors. Journal of Theoretical Biology 165(4):477-502.

Varma A, Palsson BO. 1993b. Metabolic Capabilities of Escherichia-Coli .2. Optimal-Growth Patterns. Journal of Theoretical Biology 165(4):503-522.

Wang L, Birol I, Hatzimanikatis V. 2004a. Metabolic Control Analysis under Uncertainty: Framework Development and Case Studies. Biophys. J. 87(6):3750-3763.

Westerhoff HV, Kell DB. 1996. What bio technologists knew all along...? J Theor Biol 182(3):411-20.