# **Statistical Global Metabolic Control Analysis**

ZELIMIR KURTANJEK Faculty of Food Technology and Biotechnology University of Zagreb Pierottijeva 6, 10 000 Zagreb CROATIA zkurt@pbf.hr http://www.pbf.hr

*Abstract*: This work introduces the concept of global sensitivity based on simultaneous variation of a complete set of enzymes, metabolite concentrations, and cofactors in finite ranges of concentrations. Perturbations are defined by finite ranges of concentrations and corresponding probability density distributions. Effects of perturbations of the homeostatic state variables including input and output products of a considered flux, and cofactors on the flux analysis are evaluated. The flux sensitivities are determined as first and second order relative multidimensional variances. The first order effects are reflection of random perturbation of each individual enzyme, metabolite and cofactor. The dispersions of fluxes due to each perturbation are evaluated by numerical simulation and the Fourier Amplitude Sensitivity Test algorithm. The implications of the proposed theory are demonstrated by computer simulation of theoretical problem on regulation of branched metabolism pathway at steady state, and also unsteady control flux analysis based on experimental data of *E. coli* central metabolism upon perturbation by glucose impulse.

Key-Words: MCA metabolic control analysis, flux control coefficients, homeostasis, global sensitivity

# **1** Introduction

Metabolic Control Analysis (MCA) is a mathematical theory stemming from electrical engineering network analysis applied to biological systems. Availability of annotated genome, metabolomics and proteomics of numerous industrial important microorganism leads fundamental research of industrial microbiology "in silico" [1]. From engineering point of view, open are possibilities for computer design of synthetic genome and proteinsfor development of new technologies, most importantly for bioenergetics based on synthetic microorganism with integrated photosynthesis and fermentation metabolisms (bioethanol) or biodiesel production by algae [2]. The main obstacles toward this far reaching goal are not in chemical synthesis of genome, but rather in biological and computer analysis of intricate metabolism control on a molecular level. At present, most of MCA analysis is based on steady state (homeostatic constraint) analysis and study of "one factor at a time" infinitesimal effects of perturbations of each individual enzyme and metabolite concentration on metabolic fluxes and individual reaction rates.

Control of metabolic networks is result of evolution encoded in a genome and epigenetic processes resulting to adaptation to external conditions. The result is highly structured and hierarchical systems of interrelated pathways responding to multi-objective life support goal functions. Although annotated genomes of numerous industrially important microorganisms are known, and related pathways can be discerned from the corresponding stoichiometric matrix of metabolic reactions, control of fluxes through the pathways on the cellular level is very complex and open to experimental and theoretical (mathematical) research. For example, analysis of the flux control to lipid production in autotrophic and heterotrophic algae is of interest for optimization of microbial biodiesel production

Most of approaches are based on assumptions that pathways can be isolated (and commonly of a linear structure) from the whole, a top to down decomposition (from catabolism to anabolism) is applicable, and steady state (cell homeostasis) conditions is presumed..

However, these assumptions are usually "ad hoc" and systems analysis can assist experimental research to prove or disapprove these assumptions. The aim of this work is to apply systemic or global approach to metabolic control analysis for a model of a steady state branched pathway and the unsteady flux analysis of *E. coli* central metabolism under glucose impulse, and to infer conclusions from comparison of the local versus the global approach.

## **2** Problem Formulation

The problem of metabolic control analysis is usually formulated to a single pathway of importance. For a flux  $J_i$  the relative effect (local sensitivity) of involved enzyme  $e_j$  the flux control coefficient is defined as [2-5]

$$C_{ij}^{J} = \frac{e_{j}}{J_{i}} \cdot \frac{dJ_{i}}{de_{j}}$$
(1)

It is assumed that the flux is at a steady state, and that perturbation of the i-th enzyme does not reflect on homeostatic state (homeostatic constraint), as depicted in Fig. 1. Concentration control coefficients are defined by

$$C_{ij}^{x} = \frac{e_{j}}{x_{i}} \cdot \frac{dx_{i}}{de_{j}}$$
(2)

where the relative sensitivity of metabolite concentration  $x_i$  on perturbation of enzyme concentration  $e_j$  is considered. Similarly are defined the concentration control coefficients with the respect to a change in to the pathway external concentrations. From the kinetic modeling view point, the local property termed as flux elasticity is of the most importance

$$\varepsilon_{ij}^{x} = \frac{x_{j}}{J_{i}} \cdot \frac{\partial v_{i}}{\partial x_{j}}$$
(3)

The various metabolic control coefficients are interrelated by the summation and connectivity theorems [2-5].

In view of distribution of enzyme distribution over a population of cell generations, and especially due to inherently large errors in modeling and estimation of kinetic parameters in metabolic networks, here is applied a stochastic approach in evaluation of sensitivities. Proposed is the alternative approached based on the concept of systemic global sensitivities. Enzyme concentrations  $e_i$  are considered as independent random variables in a finite range of concentrations  $[e_{i,min}, e_{i,max}]$  with associate probability distribution functions  $\rho_i(e_i)$ . The joint probability distribution function is a product of individual distributions, as statistical independence of enzyme concentrations is assumed. Effect of total (ensemble) enzyme distribution on a metabolic flux is determined on the basis of the first and the second moments. The first moment is

$$E(J_{i}) = \int_{e_{1,\min}}^{e_{1,\max}} \cdots \int_{x_{N,\min}}^{e_{N,\max}} J_{i}(e_{1}, \cdots e_{N}) \cdot \rho_{1}(e_{1}) \cdots \rho_{N}(e_{N}) \cdot de_{1} \cdots de_{N}$$

$$(4)$$

The second moment is the variance of the flux

$$\sigma^2(J_i) = E(J_i^2) - E(J_i)^2$$
<sup>(5)</sup>

The average enzyme ensemble of the square term is given by

$$E(J_i^2) = \int_{e_{1,\min}}^{e_{1,\max}} \cdots \int_{x_{N,\min}}^{e_{N,\max}} J_i^2(e_1, \cdots e_N) \cdot p_1(e_1) \cdots p_2(e_N) \cdot de_1 \cdots de_N$$
(6)

Effect  $S_{ij}^{J}$  of each individual enzyme  $e_j$  on the flux  $J_i$  is evaluated by the ratio of the conditional variance of a given enzyme and the total dispersion of the ensemble of enzymes

$$S_{ij}^{J} = \frac{\sigma^2 \left( E\left(J_i \middle| e_j\right) \right)}{\sigma^2 \left(J_i\right)} \tag{7}$$

The multidimensional integrals can be numerically evaluated by random number generation and Monte Carlo method. This work is focused on numerically more efficient method of pattern search called the Fourier Amplitude Sensitivity Test (FAST) [6,7].

Assumed is a uniform probability distribution of enzyme concentration in a preselected concentration range. The concentration range is scaled to the standard range [-1,1]. The uniform probability distribution is generated by piece wise linear functions arcsin(sin(s)) of a scan variable *s* and two randomly selected parameters, frequency  $\omega_j$ , and phase  $\varphi_j$ , corresponding to each enzyme  $e_j$ 

$$e_{j} = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin\left(\sin\left(\pi \cdot \omega_{j} \cdot \mathbf{s} + \varphi_{j}\right)\right) \quad (8)$$

Randomly are selected integer frequencies in a preselected range, and the scan variable is incrementally covering the complete range. Random selection of the parameters ensures that the functions (8) are mutually independent, i.e. uncorrelated. For each scan variable the flux is evaluated and the resulting data are interpolated to provide a continuous function function  $J_i(s)$  is expanded into Fourier series

$$A_{\omega} = \frac{1}{2 \cdot \pi} \cdot \int_{-\pi}^{\pi} J_i(s) \cdot \cos(\omega \cdot s) \cdot ds \quad (9)$$

$$B_{\omega} = \frac{1}{2 \cdot \pi} \cdot \int_{-\pi}^{\pi} J_i(s) \cdot \sin(\omega \cdot s) \cdot ds \qquad (10)$$

The total dispersion is calculated from the spectra

$$D_T = 2 \cdot \sum_{\omega=1}^{\infty} \left( A_{\omega}^2 + B_{\omega}^2 \right) \tag{11}$$

Contribution of each enzyme  $e_j$  in the total dispersion is calculated from the harmonics of the corresponding fundamental frequency  $\omega_i$ 

$$D_{j} = 2 \cdot \sum_{\omega=k \cdot \omega_{j}}^{\infty} \left( A_{\omega}^{2} + B_{\omega}^{2} \right)$$
(12)

The global flux control coefficient (termed flux global sensitivity) is determined as the ratio of the dispersion corresponding to each enzyme and the total dispersion

$$S_{ij}^{J} = \frac{D_j}{D_T} \tag{13}$$

Numerical simulation and Fourier analysis is efficiently and numerically accurate and stably evaluated by Wolfram Research "*Mathematica*" software [11].



Fig.1. Schematic presentation of a pathway and homeostatic variables in a cell. *S* is the metabolite consumed by the pathway (substrate), and  $P_1$  and  $P_2$  are the metabolites produced by the pathway. *A* and *AH* are the cofactors globally regulated on a cell level.

### **3** Problem Solutions

#### 3.1. Steady state branched network

The first case study is a problem of control analysis of a branched pathway. The model is used as a standard case due to its kinetic complexity and potential application for a flux redirection at a branch point [4,8]. The pathway, Fig. 2, consumes an external substrate (metabolite) S which is originally considered at a constant concentration level as a part of the homeostatic (steady state) assumption. The pathway branches and produces two products (metabolites)  $P_1$  and  $P_2$  which are also assumed to be constant concentration levels.



Fig. 2. Structure of a branched pathway. Metabolites engaged in the pathway are denoted from  $M_1$  to  $M_6$ . For each reaction are designated corresponding enzymes from  $E_1$  to  $E_8$ . The fluxes through the branch point are denoted as  $J_1$  to  $J_3$  [4,8].

The network includes 8 enzymes with complex kinetics, negative and positive feedback interactions, allosteric regulation, and moiety due to cofactor generation and consumption. As a part of the homeostatic constraint is assumed that the cofactors are conserved, i.e. its total amount is constant. The enzyme kinetics include 1 reversible Hill rate (including 7 parameters reflecting allosteric control), 3 reactions with ordered bi-bi kinetics (10 parameters), and 3 reactions with uni-uni kinetics (4 parameters). The model is highly nonlinear with pos-





Fig. 3. Local flux control coefficients evaluated at steady state S = 1 mmol/L,  $P_1 = 0.1 \text{ mmol/L}$ , and  $P_2 = 0.2 \text{ mmol/L}$ .



Fig. 4. Local concentration control coefficients evaluated at the given steady state.

The metabolic flux control coefficients and local concentration coefficients are evaluated by simulation of the perturbed dynamic balances and application of eq. (1-2). The results are presented in Fig. 3-4. The main result is that the total flux  $J_1$  is dominantly regulated by the enzyme  $E_2$  at the branch point, while  $E_4$  is the key enzyme for the first branch  $J_2$ , while a set of enzymes  $E_2$ ,  $E_4$  and  $E_7$ regulate the second branch J<sub>3</sub>. Global flux control coefficients are presented in Fig. 5-6. The first set of results, Fig. 5., is obtained by perturbation of the enzyme concentrations in the range 1:4 at the given steady state (homeostatic condition). The second set of results, Fig. 6., is obtained when simultaneous perturbation of the enzymes and the flux external variables (substrate S, products  $P_1$  and  $P_2$ , and total cofactors concentrations A and AH) are perturbed in the same range. The results reveal a very different control. The first enzyme  $E_1$  controls the total flux  $J_1$  for the steady and the perturbed homeostasis. In contrast to the local flux control results the enzyme  $E_2$  at the branch point does not



Fig. 5. Global flux control coefficients S<sup>J</sup> evaluated at homeostatic conditions for perturbations of the enzyme concentrations (activities) in the range from 1 to 4.



flux "external" substrate, products, and cofac-

tor variables.

exert the control over the branches. The dominant enzymes for  $J_2$  and  $J_3$  are enzymes  $E_3$  and  $E_6$  which are bellow the branch point in the respective branches. This conclusion is also irrespective to steady or perturbed homeostasis. The strong effect of hemostasis perturbation is revealed on the role of the enzyme  $E_8$ , at the end of the second branch responsible for synthesis of  $P_2$ . It becomes a key enzyme which controls not only the second branch flow but also the total flux.

### 3.2. Unsteady response of *E. coli* central metabolism

The second problem is focused on metabolic flux control analysis through *E. coli* central metabolism under unsteady conditions. The key metabolites are on-line measured in intervals of 0.25 seconds during the interval of first 15 seconds after a glucose impulse. The modeling details are given in [9-10]. The model is given as a set of 24 enzymes, 10 mass balances, and 132 kinetic parameters. Structure of the model is

$$\frac{d}{dt}\mathbf{x} = \mathbf{N} \cdot \mathbf{v} \tag{14}$$

where  $\mathbf{x}$  are intracellular components (metabolites), N is a constant stochiometric matrix, and  $\mathbf{v}$  is the vector of rates (intra and extra cellular fluxes). The rates are complex functions of enzyme  $\mathbf{e}$ , metabolite  $\mathbf{x}$ , cofactor and allosteric effector concentrations  $\mathbf{c}$ :

$$\mathbf{v} = \mathbf{v}(\mathbf{e}, \mathbf{x}, \mathbf{c}) \tag{15}$$

The global flux control coefficients are determined for the input and output fluxes of the central metabolism. From the model is revealed that there is a secondary input flux besides PTS which is a surge flow from OAA (oxaloacetate) to PEP. The global sensitivity coefficients are calculated at the level of perturbations  $\pm 50\%$ . Sensitivities are determined for the parameters and also for the enzymes associated to a corresponding set of parameters by

$$S_{ENZYME} = \sum_{i=k}^{i=l} S_i \quad \text{for } i = k, k+1, k+2\cdots, l$$

$$[k, l] \in \text{Model}(\text{Enzyme}) \tag{16}$$

In Fig. 7a-b. are presented the sensitivities of PTS flux evaluated at 1 second and 5 seconds after the impulse. The result shows a sudden shift of the flux control from PTS enzyme(s) to PFK. The PTS



Fig. 7a. Global PTS flux control coefficients on the enzyme (kinetic parameters) evaluated 1 second after the glucose impulse.



Fig. 7b. Global PTS flux control coefficients of on the enzyme (kinetic parameters) evaluated 5 seconds after the glucose impulse.



Fig. 8. Global sensitivities  $S_i$  of the flux from pyruvate to biomass evaluated during 5 s after the glucose impulse. The maximum sensitivities are: (1) pyruvate to biomass  $S_{130}(vfpyrbm) = 0.165$ ;

 $S_{139}(Km44) = 0.152;$ 

- (2) aldolase  $S_{34}$ (vmaxaldo) =0.1527;
- (3) phosphofructokinase  $S_{19}(L) = 0.00961$ ;
- (4) PTS  $S_4(K_{i1}) = 0.083;$
- (5) pyruvate-dehydrogenase  $S_{86}(K_{m28}) = 0.094$ .

complex is regulating the very initial intake rush of glucose in the time span bellow 1 second. Soon after, the control of glucose intake is regulated by PFK, in particular due to the interaction (inhibition) of PFK by PEP. Results show that the PFK control remains as throughout the experiment. The second result depicted in Fig. 8 shows global control sensitivities of the pathway from pyruvate to biomass. It can be inferred that this flux has a very distributed control over 5 to 6 enzymes involved. The given "snap shot" for the first 5 seconds remains approximately as a steady profile throughout the interval of the experiment.

# **4** Conclusions

In view of stochastic effects involved in metabolic flux modeling and control analysis proposed is a statistical concept of the global flux control sensitivity. It is based on assumption of probability density functions for parameters involved. The parameters include levels of enzyme concentrations (activities expressed as the maximum rates), substrate consumed and products released by a pathway. Besides the parameters associated with a specific pathway, included are also sensitivities to parameters associated with a concept of homeostasis. Specifically, is considered perturbation of cofactors affecting a pathway but regulated on a systemic or cellular level.

The studied examples show the main implications of the proposed concept. For the case of a steady flux through a branched pathway shows the essential difference between the local one by one parameter sensitivity and the global sensitivity flux control coefficients. The flux control is changed from the enzyme at the branch point to the enzymes on the branches. The effect of perturbation of the pathway products on the flux control became dominant.

The results of unsteady flux control analysis shows a drastic shift of flux control in time. For example, it is shown that PTS enzyme(s) are dominant only for a fraction of second upon glucose impulse, and the control is shifted to PFK regulated by interaction with PEP. For a pathway from pyruvate to biomass flux control is distributed over several enzymes and the same control profile remains throughout the experiment.

In conclusion, this approach is an extension of the local one-to-one sensitivity analysis which can pro-

vide a broader horizon for theoretical analysis and experimental verification aimed to rational application of genetic engineering.

#### References:

- G. Stephanopoulos, A. Aristidou, J. Nielsen, Metabolic Engineering – Principles and Methodologies, Academic Press, 1998.
- [2] G. Stephanopoulos, Metabolic engineering: Enabling technology for biofuels production, *Metabolic Engineering*, Vol. 10., 2008, pp. 293-294.
- [3] D. Visser, J.J. Heijnen, The Mathematics of Metabolic Control Analysis, *Metabolic Engineering*, Vol. 4, 2002, pp. 115-123.
- [4] D. Visser, J.J. Heijnen, Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics, *Metabolic Engineering*, Vol. 5, 2003, pp. 164-176.
- [5] J.J. Heijnen, Approximative Kinetic Formats Used in Metabolic Network Modeling, *Biotechnology and Bioengineering*, Vol. 91, No. 5, 2005, pp. 534-545,
- [6] R.I. Cukier, C.M. Fortuin, K.E. Shuler, A.G. Petschek, J.H. Schaibly. Study of the sensitivity of the coupled reaction systems to uncertainties in rate coefficients. I theory, *The Journal of Chemical Physics*, Vol. 53, 1970, pp. 3873-3878.
- [7] A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, Sensitivity analysis of chemical models, *Chemical Reviews*, Vol. 105, No. 7, 2005, pp. 2811-2827.
- [8] P. Mendes, D.B. Kell, Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation, *Bioinformatics*, Vol. 14, No. 10, 1998, pp. 869-883.
- [9] S. Čerić, Ž. Kurtanjek, Model identification, parameter estimation and dynamic flux analysis of *E. coli* central metabolism, *Chemical and Biochemical Engineering Quarterly*, Vol. 20, No. 3. 2006, pp. 243-253.
- [10] A. Tušek, Ž. Kurtanjek, Model and Global Sensitivity Analysis of *E. Coli* Central Metabolism, *Proceedings MathMod 09*, Vienna, Austria, 2009, pp. 1778-1785.
- [11] Wolfram Research "*Mathematica*", v. 6.0, 2008.