Journal of Bioinformatics and Computational Biology
Vol. 5, No. 2(b) (2007) 521–531
© Imperial College Press

Imperial College Press
www.icpress.co.uk

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

GENERALIZED HILL FUNCTION METHOD FOR MODELING MOLECULAR PROCESSES

VITALI LIKHOSHVAI* and ALEXANDER RATUSHNY†
Institute of Cytology and Genetics SB RAS
Novosibirsk, 63090, Russia
**likho@bionet.nsc.ru*
†*aratushny@systemsbiology.org*

Received 17 October 2006
Revised 16 January 2007
Accepted 2 February 2007

Development of an *in silico* cell is an urgent task of systems biology. The core of this cell should consist of mathematical models of intracellular events, including enzymatic reactions and control of gene expression. For example, the minimal model of the *E. coli* cell should include description of about one thousand enzymatic reactions and regulation of expression of approximately the same number of genes. In many cases detailed mechanisms of molecular processes are not known. In this study, we propose a generalized Hill function method for modeling molecular events. The proposed approach is a method of kinetic data approximation in view of additional data on structure functional features of molecular genetic systems and actually does not demand knowledge of their detailed mechanisms. Generalized Hill function models of an enzymatic reaction catalyzed by the tryptophan-sensitive 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase and the *cydAB* operon expression regulation are presented.

Keywords: Mathematical modeling; molecular genetic systems; regulation; gene expression; enzyme reaction; Hill function.

1. Introduction

Development of an *in silico* cell is often called a challenge of the postgenomic epoch in systems biology. Such large-scale computer resources demand huge volumes of cross-disciplinary studies involving biology, mathematics, and computer science. The cores of such systems must include knowledge bases storing comprehensive experimental and theoretical information on the cell and its processes, as well as mathematical and computer models describing cell operation in space and time and allowing experiments *in silico*. Construction and analysis of these models will provide new opportunities for understanding processes occurring in living systems and allow statement and solution of fundamental and applied problems. To construct models suitable for making predictions, it is necessary to describe cellular processes in sufficient detail.

†Present address: Institute for Systems Biology, Seattle, WA 98103-8904, USA.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

For example, analysis of metabolic activity of cells under various conditions demands not only description of metabolic reactions but also consideration of gene control. Thus, concerning the *E. coli* cell, the minimal model should include description of about one thousand enzymatic reactions and regulation of expression of approximately the same number of genes.[1,2] The main information for constructing models concerns the structure of elements of the system and variation of their operation under various experimental conditions. Use of experimental data is hampered by the fact that they are often contradictory. Detection of contradictions is a separate problem, which often requires painstaking construction of numerous model versions and their analysis. Another difficulty is that different elements of the system are differently understood. Some enzymes have been studied in detail with regard to their structure and kinetics, and corresponding mathematical models have been proposed. The same is true for some genetic systems. However, the levels of our knowledge on various elements are broadly different. Thus, we cannot describe all processes equally comprehensively, and our description should be determined by the volume of knowledge of the system. The modeling method should allow choosing the most appropriate level of detailedness in the model. It is pertinent to consider in brief the hierarchy of description levels in the model, recognized as relatively separate.

The most comprehensive models can be constructed in terms of biochemical reactions. With the assumption of perfect circulation, they yield sets of ordinary differential equations, which operate on substance concentrations or probabilities of certain states. Such models are nonsteady. Obviously, they require large volumes of data, because they involve large numbers of intermediate states and constants demanding verification.

Another detailedness level is occupied by models that are also constructed on the grounds of biochemical systems but with the assumption of a certain steadiness level. This approach is commonly applied to enzymatic kinetics, where the King&Altman method is helpful.[3,4] It can be applied to genetic processes with no substantial modifications. However, this approach is still sensitive to the level of knowledge of the system.

The last level recognized by us and considered in this report involves construction of approximate models. We place there all approaches dealing with construction of approximate models based on description of actual data with as little speculation as possible. One of such methods, involving approximation with a class of generalized Hill functions (GHF), is presented here. We use these functions because they provide natural generalization of steady-state biochemical models. With certain assumptions, such models can be expressed by rational polynomials of integer degrees. The name of this function class emphasizes that for the first time such functions were considered by Hill, who described kinetics of oxygen attachment to hemoglobin. Later, they were broadly used for modeling regulation of genetic processes. GHF modeling method mediates between flux methods which allows to calculate a steady-state distribution of fluxes in regulatory and metabolic networks[5] and portrait modeling of molecular systems.[6,7]

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

The approach proposed in this paper supplements the generalized chemical kinetic modeling method developed earlier.[6,7] GHF models of an enzymatic reaction catalyzed by the tryptophan-sensitive 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase and the *cydAB* operon expression regulation are presented.

2. General Structure of Elementary Subsystems Models

Construction of elementary subsystems (ES) models is the mandatory stage of the dynamic simulation of molecular genetic systems (MGS). Such elementary subsystems may be: (i) enzymatic reactions; (ii) formation of complexes with biologically significant functions (active enzyme species, transcription factors, RNA polymerases, ribosomes, etc.); (iii) molecular genetic events, including initiation and termination of replication, transcription, and translation, with regard to their regulation; (iv) macromolecule degradation processes; and so on.

ES models describing the rates of changes of substance concentrations are written in the differential form:

$$\frac{dX}{dt} = F(X, Y, K), \tag{1}$$

where X is the list of dynamic variables; K, list of parameters; Y, list of functional activities; $F(X, Y, K)$, monitor rule.

For calculating functional activities from Eq. (1), we use the algebraic function:

$$Y = G(X, Z, K), \tag{2}$$

where Z is the list of output functional activities, which also calculated by using the algebraic function of view (2).

Target object models (TOMs) are constructed on the grounds of a set of elementary models. The algorithm includes two stages. At the first stage, a combined set of differential equations (SDE) is constructed. The dynamic variables of the SDE are variables entering the union of all lists X for all selected elementary models (1). The right side for each variable is determined as the sum of the right sides F of all elementary models containing this variable in their lists X. The second stage involves repetitive substitution of equations G from elementary models of view Eq. (2) for functional activities of list Y in the right sides of the SDE. The resulting TOM is used for solving corresponding tasks.

Let us dwell on methods for defining the types of functions F and G in GHF terms.

2.1. *Definition of generalized Hill functions*

Let X designate a set of variables. Denote the set of indices for elements of a subset of the set of all subsets of set X as $A = \{\alpha\} : \aleph = \{X_\alpha : X_\alpha \subseteq X, \alpha \in A\}$. Null set $\emptyset$ may enter into $\aleph$. Let K, N, and Δ designate the sets of parameters called efficiency coefficients (EC), Hill coefficients (HC), and activity coefficients (AC),

respectively: $\{k_\alpha : \alpha \in \mathrm{A}\} = K$, $\{n_{\alpha,x} : x \in X_\alpha, \alpha \in A\} = N$, $\{\delta_\alpha : \alpha \in A\} = \Delta$. By implication of $k_\alpha, n_{\alpha,x} > 0, \delta_\alpha \geq 0$.

(a)

$$h(x|x \in X) = R(X)/Q(X) = \sum_\alpha \delta_\alpha \prod_{x \in X_\alpha} (x/k_\alpha)^{n_{\alpha,x}} \Big/ \sum_\alpha \prod_{x \in X_\alpha} (x/k_\alpha)^{n_{\alpha,x}}$$

— This rational polynomial is a GHF (By convention, if X_α $\emptyset$, then $\prod_{x \in X_\alpha} (x/k_\alpha)^{n_{\alpha,x}} \equiv 1$),

(b) If $h(X, K, N, \Delta)$ is a GHF, then $\tilde{h}(X, K(X, K, N, \Delta), N(X, K, N, \Delta), \Delta)$, obtained from $h(X, K, N, \Delta)$ by substitution of K and N parameters to the GHF, is also a GHF.

2.2. *Description of the GHF model constructing procedure*

Experimental data are reported in terms of the (discrete) function $E : S \subseteq DR'' \to R$, where n is the dimension of vector X. Point S_i of S is associated with the corresponding value of the function $E(S)$ and determines the rate of an enzymatic reaction or gene expression efficacy. The problem is to construct the continuous (on all arguments) GHF having the least deviation from function E at corresponding points (on S). Note that this problem is soluble by virtue of the validity of theorems about generalized Hill function completeness.[8]

The process of construction begins with choice of the vector X variable to be considered at the first stage of the process. As a rule, the first variable has the most informative scripts (experimental data). The script is understood as a discrete function constructed from experimental data for variation of the chosen substance under identical conditions with regard to all other substances. It is obvious that the more points there are in one script, the more information such function contains. Without loss of generality, we may select the first variable of vector X. Then we construct the common structured GHF of only one selected variable which allows reproducing each script, but it is possible with different sets of kinetic and Hill parameters. We designate the constructed function as $\mathbf{H1}(x_1)$. The result of the first stage is the generalized Hill function $\mathbf{H1}(x_1)$ of a universal form and tables of all its parameters obtained by varying variables $X_1 = X \backslash \{x_1\}$. Then we consider the coefficients of the function $\mathbf{H1}(x_1)$ as generalized Hill functions of n-1 variables and the tables of parameters as an analogue of experimental data (it should be noted that not all of $\mathbf{H1}(x_1)$ parameters vary at different conditions. Some of them can be constant, and for such parameters we finish the process).

The next stage is to construct the GHF for each parameter. The procedure of its definition is the same as described earlier. For each parameter, we select the most informative variable, prepare the scripts and construct the universal function of one variable and the corresponding table of parameters. After definition of the GHF for each parameter we pass to the following stage. We repeat the above described procedure until all variables are exhausted.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

As a result we get set of generalized Hill functions of view (a) with certain coefficients and the order of their composition to uniform generalized Hill function of view (b).

3. Results and Discussion

Here, we present examples of modeling an enzymatic reaction catalyzed by the tryptophan-sensitive 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase and regulation of the expression of the *cydAB* operon by using generalized Hill function method.

3.1. *GHF model of the tryptophan-sensitive 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase*

We consider regulation of the reaction catalyzed by the enzyme tryptophan-sensitive 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase (DAHPS(Trp)), coded by the *aroH* gene, as an example of simulation of molecular processes in the gene network of aromatic amino acid biosynthesis in *E. coli.* This enzyme is one of three differentially regulated isozymes that catalyze the first step of aromatic biosynthesis, the condensation of phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) to form 3-deoxy-D-arabino-heptulosonate-7-phosphate (3DDAH7P). Here is the reaction of synthesis of 3DDAH7P from E4P and PEP, catalyzed by DAHPS:

$$\mathrm{E4P} + \mathrm{PEP} + \mathrm{H2O} \rightarrow \mathrm{3DDAH7P}$$

The enzyme is homodimeric and has two independent inhibitor binding sites. DAHPS(Trp) displays sigmoid kinetics with respect to both substrates, E4P and PEP. Both catalytic activity and substrate affinity of the DAHPS(Trp) are dependent on the species of activating metal ion. L Tryptophan (Trp) binding decreases k_{cat}, decreases positive homotropic cooperativity for both substrates and activates the enzyme at low concentrations of E4P.[9]

The enzymatic reaction has a very intricate mechanism of DAHPS(Trp) regulation by Trp. With regard to the effect of Trp on various parameters of the molecular system under consideration, the rate of the enzymatic reaction can be expressed as:

$$V = \frac{k_{\mathrm{cat}} \cdot e_0 \cdot (S_1/(K_{m,S_1} \cdot f_2))^{h_{S_1} \cdot f_3} \cdot (S_2/K_{m,S_2})^{h_{S_2} \cdot f_4}}{P/K_{i,P} + [1 + (S_1/(K_{m,S_1} \cdot f_2))^{h_{S_1} \cdot f_3}] \cdot [1 + (S_2/K_{m,S_2})^{h_{S_2} \cdot f_4}]} \cdot f_1,$$

$$f_1 = \frac{1}{1 + kl_{R,V_{\max}} \cdot \frac{R}{k_{R,V_{\max}} + R}}, \quad f_2 = \frac{1}{1 + kl_{R,K_{m,S_1}} \cdot \frac{R^{h_{R,K_{m,S_1}}}}{k_{R,K_{m,S_1}}^{h_{R,K_{m,S_1}}} + R^{h_{R,K_{m,S_1}}}}}, \tag{3}$$

$$f_3 = \frac{1}{1 + kl_{R,h_{S_1}} \cdot \frac{R^{h_{R,h_{S_1}}}}{k_{R,h_{S_1}}^{h_{R,h_{S_1}}} + R^{h_{R,h_{S_1}}}}}, \quad f_4 = \frac{1}{1 + kl_{R,h_{S_2}} \cdot \frac{R^{h_{R,h_{S_2}}}}{k_{R,h_{S_2}}^{h_{R,h_{S_2}}} + R^{h_{R,h_{S_2}}}}}.$$

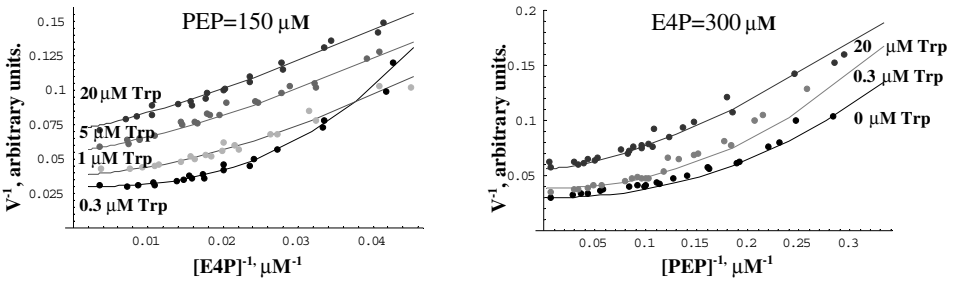
J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

where V is the rate of the reaction; e_0 is the concentration of the enzyme DAHPS(Trp); S_1, S_2, P, and R are the concentrations of E4P, PEP, 3DDAH7P, and Trp, respectively; k_f is the catalytic constant; $K_{m,S1}$ and $K_{m,S2}$ are the Michaelis-Menten constants for the substrates E4P and PEP, respectively; $K_{i,P}$ is the constant of inhibition by the 3DDAH7P product; h_{S1} and h_{S2} are constants determining the nonlinearity of the effect of the substrates E4P and PEP on the reaction rate, respectively; $kl_{R,V\max}$ is the constant determining the maximum degree of reaction rate inhibition by Trp; $k_{R,V\max}$ is the constant of efficiency of the effect of Trp on the maximum rate of the reaction; $kl_{R,k}$ is the constant determining the maximum effect of Trp on the constant designated by k, where k assumes a character value from the set $\{K_{m,S1}, h_{S1}, h_{S2}\}$; $k_{R,k}$ is the constant determining the efficiency of the effect of Trp on constant k; and $h_{R,k}$ is the constant determining the nonlinearity of the effect of Trp on constant k.

Although simple, the model provides good agreement between experimental data and simulation results. Figures 1(a) and (b) present the results of calculations according to model (3) compared with experimental data on the effect of various Trp concentrations on the rate of the reaction catalyzed by the enzyme DAHPS(Trp). Apparently, the proposed model of the regulation of the metabolic reaction under study is fairly precise.

3.2. *GHF model of the cydAB operon expression regulation*

The transcription initiation region of the *cydAB* operon has a complex structure: 5 transcription starts and 10 binding sites within transcription sites ArcA (5 sites),



(a) (b)

Fig. 1. Effect of various Trp concentrations on the rate of the enzymatic reaction catalyzed by the DAHPS(Trp). Experimental conditions: (a) PEP = 150 μM, (b) E4P = 300 μM. On the X axis the reverse concentrations of substrates are plotted; on the Y axis, the reverse rates on the enzymatic reactions. Dots indicate experimental data according to Akowski&Bauerle[9]; curves are the results of simulation according to model (3); parameter values: $k_f = 20.6\,\mathrm{s}^{-1}$; $K_{m,S1} = 35\,\mu$M; $K_{m,S2} = 5.3\,\mu$M; $h_{S1} = 2.6$; $h_{S2} = 2.2$; $K_{i,P1} = 1\,$mM; $kl_{R,V\max} = 1.7$; $k_{R,V\max} = 5\,\mu$M; $kl_{R,KmS1} = 0.85$; $k_{R,KmS1} = 25\,\mu$M; $h_{R,KmS1} = 0.6$; $kl_{R,hS1} = 1.1$; $k_{R,hS1} = 1\,\mu$M; $h_{R,hS1} = 1$; $kl_{R,hS2} = 0.47$; $k_{R,hS2} = 1\,\mu$M; $h_{R,hS2} = 2$.

H-NS (3 sites), and Fnr (2 sites) (Fig. 2). Therefore, the portrait of this subsystem involves 510 mono- and bimolecular reactions with 197 dynamic variables (194 promoter region states and 3 regulators) (Fig. 3). Verification of its parameters demands detailed dynamic information on the operation of the subsystem depending on transcription factor concentrations. At present, such data are absent from available sources. However, this subsystem has been studied in the context of oxygen concentration variation.[10] Therefore, we model it with generalized Hill functions.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

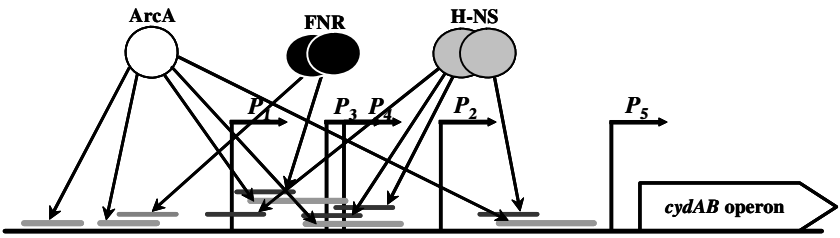


Fig. 2. Structure of the promoter of the *cydAB* operon, coding for cytochrome bd oxidase. Retrieved from Ecocyc database (http://ecocyc.org). Factor ArcA is a transcription activator, H-NS is a repressor, and Fnr can be both an activator (left binding site of the second promoter) and a repressor.

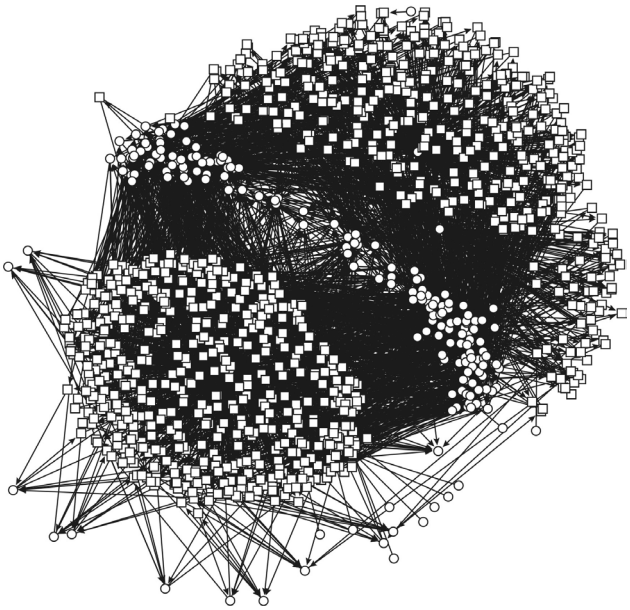


Fig. 3. Bipartite graph of the model constructed in terms of mono- and bimolecular reactions to describe the regulation of transcription initiation in the promoter region of the *cydAB* operon. The graph contains 510 elementary processes (square nodes) and 194 promoter region states and 3 regulators (circular nodes).

Thus, we obtain the following model equation describing the level of *cydAB* expression (f_{Δ_a,Δ_f}) in the wild type ($\Delta_a = 1, \Delta_f = 1$) and knockout strains $\Delta arcA$ ($\Delta_a = 0, \Delta_f = 1$) and Δfnr ($\Delta_a = 1, \Delta_f = 0$) depending on oxygen concentration:

$$f_{\Delta_a,\Delta_f} = \frac{k_{\Delta_a,\Delta_f} + \Delta_a[(O_2/k_{a11})^{h_{a1}} + (O_2/k_{a21})^{h_{a2}}] + \Delta_f(O_2/k_{f1})^{h_{f1}} + \Delta_a\Delta_f(O_2/k_{af1})^{h_{af1}}}{1 + \Delta_a[(O_2/k_{a12})^{h_{a1}} + (O_2/k_{a22})^{h_{a2}}] + \Delta_f(O_2/k_{f2})^{h_{f2}} + \Delta_a\Delta_f(O_2/k_{af2})^{h_{af2}}} \tag{4}$$

where O_2 is oxygen concentration in the medium; k_{Δ_a,Δ_f}, basal expression of the *cydAB* operon; k_{ai}, k_{fi}, k_{afi} are constants of the effect of ArcA, Fnr, and their combined effect on *cydAB* operon expression, respectively; h_{ai}, h_{fi}, h_{afi}, constants describing the nonlinearity of the effect of ArcA, Fnr, and their combined effect on *cydAB* expression, respectively; expressed in terms of oxygen concentration.

Figure 4 shows the results of comparison between calculations according to model (4) and experimental data describing the dependence of the rate of *cydAB* transcription on oxygen concentration for wild-type *E. coli* cells ($f_{1,1}$) and mutant strains $\Delta arcA(f_{0,1})$ and $\Delta fnr(f_{1,0})$.[10]

By now, the accumulated knowledge[11] has allowed construction of about 300 elementary enzymatic reaction models and about 20 elementary models describing the regulation of expression of various genes according to the method presented here.[12–22]

4. Conclusion

Reconstruction of mechanisms of gene expression regulation and functioning of enzyme reactions is one of key problems of molecular genetic systems modeling. If the mathematical model is represented in terms of differential equations the

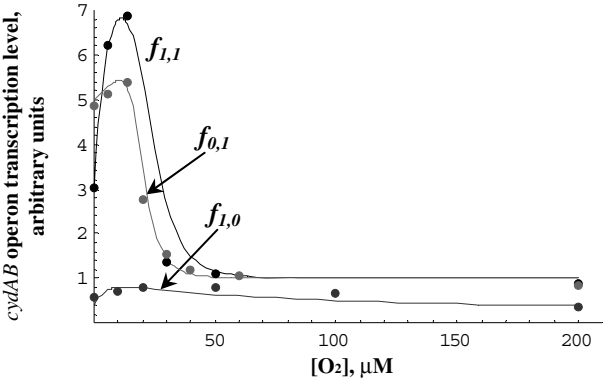


Fig. 4. Dependence of the *cydAB* operon transcription on oxygen concentration in wild-type *E. coli* cells ($f_{1,1}$) and mutant strains $\Delta arcA$ ($f_{0,1}$) and Δfnr ($f_{1,0}$). Dots indicate experimental data.[10] Curves indicate calculations from Eq. (4). The curves were calculated with the following parameters: $k_{1,1} = 3$; $k_{0,1} = 5$; $k_{a11} = 20\,\mu$M; $k_{a12} > 500\,\mu$M; $h_{a1} = 1$; $k_{a21} = 20\,\mu$M; $h_{a2} = 6$; $ka_{22} = 20\,\mu$M; $k_{1,0} = 0.4$; $k_{f1} = 9\,\mu$M; $h_{f1} = 1$; $k_{f2} = 11\,\mu$M; $h_{f2} = 1.4$; $k_{af1} = 1.2\,\mu$M; $h_{af1} = 1$; $k_{af2} = 100\,\mu$M; $h_{af2} = 6$.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

given problem can be formulated as a problem of reconstruction of right parts of these equations. The way of the problem solving depends on the amount of studied structural and dynamic characteristics of molecular subsystem under investigation. Revealing of the MGS biochemical scheme is one of the basic purposes at studying patterns of its functioning. The way of modeling from the biochemical scheme to mathematical model is attractive but it is difficult to use the one in many cases. Often biochemical schemes are very complex and intricate. There is a huge amount of different variants of schemes and it is impossible to check them up for consistency. It is possible to construct non-steady models only for well studied molecular subsystems. In some cases biochemical schemes can be proposed but their non-steady models cannot be verified because of the lack of necessary experimental data. One of the decisions of this problem can be the construction of the simplified model. The assumption of presence of fast and slow processes in the system is often applied to simplify the model.[23] For example, in case of enzymatic reaction modeling the King&Altman algorithm[1,2] is commonly used for description of steady-state rates of reactions by rational polynoms. GHF are generalization of rational polynoms and one can use these functions for modeling biochemical processes and gene expression regulation. Modeling in terms of GHF generally does not demand a reconstruction of the biochemical scheme. The GHF approach is a method of kinetic data approximation in view of additional data (if they are available) on structure functional features of MGS. Thus application of the GHF method allows to bypass problems of reconstruction of detailed mechanisms of molecular subsystems and to build the model for the whole MGS at the same time. In process of experimental data accumulation GHF models can be replaced with models describing mechanisms of molecular subsystems in more details.

In the future we plan to improve the proposed method and to develop algorithms for the GHF reconstruction based on available kinetic data. We plan also to develop algorithms for verification of MGS biochemical schemes based on GHF models.

Acknowledgments

The authors are grateful to Ilya Akberdin, Kirill Bezmaternykh, Tamara Khlebodarova, Sergei Lashin, Eric Mjolsness, Viatcheslav Mordvinov, Eugenia Nedosekina and Nikolay Kolchanov for valuable discussion. The work was supported in part by the Russian Foundation for Basic Research (Grants No. 05-07-98011, 05-07-98012, 06-04-49556), by NSF:FIBR (Grant No. EF-0330786), by the Federal Agency of Science and Innovation (Contract No. 02.467.11.1005), by the Russian Academy of Sciences (Contracts No. 10.4 and 10104-34/P-10/147-269/260606-038) and by the Siberian Branch of RAS (Budget project No. 23.1).

References

1. Keseler IM, Collado-Vides J, Gama-Castro S, Ingraham J, Paley S, Paulsen IT, Peralta-Gil M, Karp PD, EcoCyc: A comprehensive database resource for *Escherichia coli*, *Nucleic Acids Res* **33**(Database issue):D334–D337, 2005.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

2. Salgado H, Gama-Castro S, Peralta-Gil M, Diaz-Peredo E, Sanchez-Solano F, Santos-Zavaleta A, Martinez-Flores I, Jimenez-Jacinto V, Bonavides-Martinez C, Segura-Salazar J, Martinez-Antonio A, Collado-Vides J, RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions, *Nucleic Acids Res* **34**(Database issue):D394–D397, 2006.
3. King EL, Altman C, A Schematic method of deriving the rate laws for enzyme-catalyzed reactions, *J Phys Chem* **60**:1375–1378, 1956.
4. Cornish-Bowden A, An automatic method for deriving steady-state rate equations, *Biochem J* **165**:55–59, 1977.
5. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO, Integrating high-throughput and computational data elucidates bacterial networks, *Nature* 6, **429**(6987):92–96, 2004.
6. Likhoshvai VA, Matushkin YuG, Vatolin YuN, Bazhan SI, A generalized chemical kinetic method for simulating complex biological systems. A computer model of λ phage ontogenesis, *Computational Technologies* **5**(2):87–99, 2000.
7. Likhoshvai VA, Matushkin IuG, Ratushnyi AV, Ananko EA, Ignatéva EV, Podkolodnaia OA, A generalized chemical-kinetic method for modeling gene networks, *Mol Biol* (*Mosk*) **35**(6):1072–1079, 2001.
8. Bazaikin YaV, Golubyatnikov VP, Storozhuk KV, private report, 2004.
9. Akowski JP, Bauerle R, Steady-state kinetics and inhibitor binding of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (tryptophan sensitive) from *Escherichia coli*, *Biochemistry* **36**:15817–15822, 1997.
10. Tseng CP, Albrecht J, Gunsalus RP, Effect of microaerophilic cell growth conditions on expression of the aerobic (cyoABCDE and cydAB) and anaerobic (narGHJI, frdABCD, and dmsABC) respiratory pathway genes in *Escherichia coli*, *J Bacteriol* **178**:1094–1098, 1996.
11. Khlebodarova TM, Ananko EA, Nedosekina EA, Podkolodnaya OA, Poplavsky AS, Smirnova OG, Ibragimova SS, Mishchenko EL, Stepanenko IL, Ignatieva EV, Proskura AL, Nishio Y, Imaizumi A, Usuda Y, Matsui K, Podkolodny NL, Kolchanov NA, In silico cell II. Information sources for modeling *Escherichia coli* metabolic pathways, *Proc* 5^{th} *Int. Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 13–18, 2006.
12. Ratushny AV, Usuda Y, Matsui K, Podkolodnaya OA, Mathematical modeling of elementary processes of the gene network controlling histidine biosynthesis in *Escherichia coli*, *Proc* 5^{th} *Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 25–29, 2006.
13. Nedosekina EA, Usuda Y, Matsui K, Modeling of the effects of threonine, valine, isoleucine and pyridoxal 5′-monophosphate on biosynthetic threonine dehydratase reaction, *Proc* 5^{th} *Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 30–34, 2006.
14. Ratushny AV, Nedosekina EA, Regulation of pyrimidine biosynthesis in *Escherichia coli*: Gene network reconstruction and mathematical modeling, *Proc* 5^{th} *Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 35–39, 2006.
15. Ratushny AV, Smirnova OG, Usuda Y, Matsui K, Regulation of the pentose phosphate pathway in *Escherichia coli*: Gene network reconstruction and mathematical modeling of metabolic reactions, *Proc* 5^{th} *Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 40–44, 2006.
16. Ratushny AV, Khlebodarova TM, Mathematical modeling of regulation of cyoABCDE operon expression in Escherichia coli, *Proc* 5^{th} *Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 49–54, 2006.

17. Khlebodarova TM, Lashin SA, Apasieva NV, Gene network reconstruction and mathematical modeling of *E. coli* respiration: Regulation of F0F1-ATP synthase by metal ions, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 55–59, 2006.
18. Nedosekina EA, Mathematical modeling of regulation of *Escherichia coli* purine biosynthesis pathway enzymatic reactions, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 60–67, 2006.
19. Mishchenko EL, Lashin SA, Usuda Y, Matsui K, Reconstruction and mathematical modeling of the gene network controlling cysteine biosynthesis in *Escherichia coli*: Regulation of serine acetyltransferase activity, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 68–72, 2006.
20. Smirnova OG, Ratushny AV, Gene network reconstruction and mathematical modeling of salvage pathways: Regulation of adenine phosphoribosyltransferase activity by structurally similar substrates, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 73–77, 2006.
21. Turnaev II, Ibragimova SS, Usuda Y, Matsui K, Mathematical modeling of serine and glycine biosynthesis regulation in *Escherichia coli*, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 78–83, 2006.
22. Turnaev II, Smirnova OG, Regulation of pyruvate biosynthesis in *Escherichia coli*: Gene network reconstruction and mathematical modeling of enzymatic reactions of the pathway, *Proc 5th Int Conf Bioinform Genome Regulation and Structure*, Novosibirsk, Russia, Vol. 2, pp. 84–88, 2006.
23. Tikhonov AN, Systems of differential equations containing small parameters in the derivatives, *Mat Sbornik N. S.* (*Russian*) **31**(3):575–586, 1952.

J. Bioinform. Comput. Biol. 2007.05:521-531. Downloaded from www.worldscientific.com
by VIRGINIA POLYTECHNIC INSTITUTE on 04/24/15. For personal use only.

Vitali A. Likhoshvai received his Ph.D. degree in Molecular Biology from the Institute of Molecular Biology, Koltsovo, Russia, in 1985. Currently, he is a senior researcher at the Sector of Molecular Evolution of the IC&G.

Alexander V. Ratushny received his Ph.D. in Bioinformatics from the Institute of Cytology and Genetics, Novosibirsk, Russia, in 2006. Currently, he is a Postdoctoral Fellow at the Institute for Systems Biology, Seattle, WA.