MATHEMATICAL MODELING OF SERINE AND GLYCINE BIOSYNTHESIS REGULATION IN ESCHERICHIA COLI

Turnaev I.I.*, Ibragimova S.S.*, Usuda Y.2, Matsui K.2
1 Institute of Cytology and Genetics, SB RAS, Novosibirsk, 630090, Russia; 2 Institute of Life Sciences, Ajinomoto, Co., Inc., Kawasaki, Japan

* Corresponding author: e-mail: turn@bionet.nsc.ru

Key words: mathematical modeling, gene network, regulation, serine and glycine biosynthesis, Escherichia coli

SUMMARY

Motivation: Development of an electronic cell as a resource for use on computers for modeling and analysis of intracellular processes is important to further advancements in systems biology and bioinformatics. To contribute to the progress, it is of importance that mathematical models of cellular metabolic pathway regulation, in particular, regulation of serine and glycine biosynthesis be developed.

Results: The gene network for regulation of serine and glycine biosynthesis in E. coli has been reconstructed using the GeneNet technology. Using Hill’s generalized functions, mathematical models have been developed for the enzyme reactions. Based on these models, a mathematical model for serine and glycine has been developed.

Availability: The model is available on request, the gene network is accessible at http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/viewer/index.shtml

INTRODUCTION

SER is the most important cell metabolite. In cells grown on minimum medium with glucose as a source of carbon, about 15 % cell carbon follows through the PGDH reaction and onwards to the synthesis of SER and its derivatives (Pizer, Potochny, 1964). SER being a growth factor, cell division starts at the SER concentration of 0.05 mM (Prub, Matzumura, 1996). E. coli depends on the SER biosynthetic pathway, because mutations, which result in the loss of PGDH, lead to a growth requirement for SER (Pizer, Potochny, 1964). High levels of SER in growth medium lead to growth inhibition. The intracellular SER pool is regulated by feedback inhibition of PGDH, the product of the serA gene, by SER through conformational change in the enzyme. SER serves as a precursor for GLY, TRY and CYS. SER is also incorporated into phospholipids. SER is converted to GLY by SHMT, the glyA gene product. This reaction is the only source of GLY and the major source of one-carbon units to the cell (Stauffer, 1996).

We report herein the reconstruction of the gene network for serine and glycine biosynthesis and a mathematical model of this process.

6 The abbreviations used are: SER, L-serine, PGDH, 3-phosphoglycerate dehydrogenase, GLY, glycine, TRY, tryptophan, CYS, cysteine, SHMT, serine hydroxymethyltransferase.
METHODS AND ALGORITHMS

The gene network reconstruction was performed using the GeneNet software program (Ananko et al., 2005). The efficiency of the functioning of enzyme systems were modeled using generalized Hill functions (Likhoshvai, Ratushny, 2006).

RESULTS AND DISCUSSION

The gene network for regulation of serine and glycine biosynthesis in *E. coli* cells has been reconstructed (Fig. 1; http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/viewer/index.shtml). A listing of the components of the gene network for serine and glycine biosynthesis and degradation is presented in Table 1.

![Gene network for regulation of serine and glycine biosynthesis and degradation](image)

*Figure 1.* The gene network for regulation of serine and glycine biosynthesis and degradation reconstructed using the GeneNet software program.

<table>
<thead>
<tr>
<th>Item</th>
<th>Protein</th>
<th>RNA</th>
<th>Operon</th>
<th>Reaction</th>
<th>Inorganic substance</th>
<th>Literature source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>32</td>
<td>10</td>
<td>10</td>
<td>108</td>
<td>36</td>
<td>226</td>
</tr>
</tbody>
</table>

*Table 1.* A listing of the components of the gene network for serine and glycine biosynthesis and degradation.
A listing of enzyme reactions in the reconstructed gene network, the names of enzymes catalyzing the corresponding reactions and the names of genes encoding the corresponding enzymes are presented in Table 2.

Table 2. A listing of enzyme reactions in the gene network for serine and glycine biosynthesis and degradation

<table>
<thead>
<tr>
<th>No.</th>
<th>Enzyme name</th>
<th>Gene</th>
<th>Reaction</th>
<th>Low-molecular regulators</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3-Phosphoglycerate dehydrogenase (PGDH)</td>
<td>serA</td>
<td>$3\text{PG} + \text{NAD} \rightarrow \text{NADH} + \text{PHP}$</td>
<td>SER, H$^+$, NADH, PHP</td>
</tr>
<tr>
<td>2</td>
<td>Phosphoserine transaminase (PSAT)</td>
<td>serC</td>
<td>$\text{PHP} + \text{GLU} \rightarrow \text{AKG} + 3\text{PSER}$</td>
<td>AKG, 3PSER</td>
</tr>
<tr>
<td>3</td>
<td>Phosphoserine phosphatase (PSP)</td>
<td>serB</td>
<td>$3\text{PSER} \rightarrow \text{PI} + \text{SER}$</td>
<td>PI, SER</td>
</tr>
<tr>
<td>4</td>
<td>Glycine hydroxymethyltransferase (SHMT)</td>
<td>glyA</td>
<td>$\text{THF} + \text{SER} \rightarrow \text{GLY} + \text{MTHF}$</td>
<td>GLY, MTHF</td>
</tr>
</tbody>
</table>

A database containing experimental data on the dynamics of the gene network components has been developed (Khlebodarova et al., 2006). Using Hill’s generalized functions, mathematical models have been developed for the enzyme reactions listed in Table 2, and the models’ parameters, evaluated.

Based on the models of separate enzyme reactions, a dynamic model for serine and glycine biosynthesis in *E. coli* has been developed, a general schematic of which is presented in Fig. 1. The mathematical model comprises 9 differential equations.

**Enzyme reaction rates:**

$$V_1(\text{PGDH activity}) = \frac{k_1 [\text{PGDH}][3\text{PG}][\text{NAD}]}{k_{1a} + k_{1b} [\text{PGDH}][3\text{PG}][\text{NAD}]} \cdot \left(1 + \frac{k_{1c}}{1 + [\text{SER}]/k_{1d}}\right) \cdot \left(1 + \frac{k_{1e}}{1 + [\text{PI}]/k_{1f}}\right) \cdot \left(1 + \frac{k_{1g}}{1 + [\text{H}]/k_{1h}}\right)$$

$$V_2(\text{PSAT activity}) = \frac{k_{2a} [\text{PSAT}][\text{GLU}][\text{PHP}]}{k_{1a} \cdot [\text{PSAT}][\text{GLU}][\text{PHP}]} \cdot \left(1 + \frac{k_{2b}}{1 + [\text{AKG}]/k_{2c} + [\text{PHP}]/k_{2d} + [\text{3PSER}]/k_{2e}}\right)$$

$$V_3(\text{PSP activity}) = \frac{k_{1a} [\text{PSP}][1 + [\text{PI}]/k_{1b}][3\text{PSER}]}{k_{1a} \cdot [\text{PSP}][1 + [\text{PI}]/k_{1b}]} \cdot \left(1 + \frac{k_{2}}{1 + [\text{SER}]/k_{1d}}\right)$$

$$V_4(\text{SHMT activity}) = \frac{k_{3a} [\text{SHMT}][\text{THF}][\text{SER}]}{k_{3a} \cdot [\text{SHMT}][\text{THF}][\text{SER}]} \cdot \left(1 + \frac{k_{3b}}{1 + [\text{MTHF}]/k_{3c} + \text{THF} \cdot (1 + [\text{MTHF}]/k_{3d})} - (1 + [\text{GLY}]/k_{3e} + [\text{SER}]/k_{3f})\right)$$

**Differential equations:**

$$\frac{d[\text{PGDH}]}{dt} = k_1 - k_{1a} [\text{PGDH}]$$  \hspace{1cm} $$\frac{d[\text{PSAT}]}{dt} = k_2 - k_{2a} [\text{PSAT}]$$  \hspace{1cm} $$\frac{d[\text{PSP}]}{dt} = k_3 - k_{3a} [\text{PSP}]$$  \hspace{1cm} $$\frac{d[\text{SHMT}]}{dt} = k_4 - k_{4a} [\text{SHMT}]$$

$$\frac{d[\text{3PG}]}{dt} = k_1 - V_1 - k_{1a} [\text{3PG}]$$  \hspace{1cm} $$\frac{d[\text{PHP}]}{dt} = V_1 - V_2$$  \hspace{1cm} $$\frac{d[\text{AKG}]}{dt} = V_2 - V_3$$  \hspace{1cm} $$\frac{d[\text{3PSER}]}{dt} = V_3 - V_4 - k_{1d} [\text{3PSER}]$$  \hspace{1cm} $$\frac{d[\text{GLY}]}{dt} = V_4 - k_{3f} [\text{GLY}]$$

**Constant values:** $k_1 = 1.15 \times 10^{-3}$ mM/s, $k_2 = 5.3 \times 10^{-4}$ mM/s, $k_3 = 2.5 \times 10^{-5}$ mM/s, $k_4 = 0.4$ mM/s, $k_5 = 1.15 \times 10^{-4}$ mM/s, $k_{10} = 10^{-3}$ s$^{-1}$, $k_{11} = 0.4$ s$^{-1}$, $k_{12} = 0.0103$ s$^{-1}$, $k_{13} = 8.03 \times 10^{-3}$ s$^{-1}$, $k_{14} = 0.55$ s$^{-1}$, $k_{15} = 1.1$ mM, $k_{17} = 7.8 \times 10^{-3}$ mM, $k_{18} = 0.78$ mM.

---

7 The abbreviations used are: 3PG, 3-phosphoglycerate, PHP, 3-phosphohydroxypropionate, PSAT, 3-phosphoserine aminotransferase, 3PSER, 3-phosphoserine, PSP, 3-phosphoserine phosphatase, LEU, Leucine, Lrp, Leucine-responsive regulatory protein, MTHF, 5,10-methyltetrahydrofolate, MET, methionine, MetR, Homocysteine transcriptional activator, HCYS, homocysteine, PurR, Hypoxanthine transcriptional repressor, THF, tetrahydrofolate, PI, phosphate, Crp, cAMP receptor protein.
Modelling of molecular genetic systems in bacterial cell

$k_{19} = 110 \text{ mM, } k_{20} = 0.047 \text{ mM, } k_{21} = 0.074 \text{ mM, } k_{22} = 3.5 \times 10^{-3} \text{ mM, } k_{23} = 2.2, k_{26} = 1.8 \text{ mM, } k_{27} = 3 \times 10^{-4} \text{ mM, } k_{28} = 40, k_{29} = 0.6, k_{30} = 1.75 \text{ s}^{-1}, k_{31} = 0.015 \text{ mM, } k_{32} = 20 \text{ mM, } k_{33} = 100 \text{ mM, } k_{34} = 1.5 \text{ mM, } k_{35} = 1.43 \text{ s}^{-1}, k_{36} = 1.5 \times 10^{-2} \text{ mM, } k_{37} = 0.15 \text{ mM, } k_{38} = 500 \text{ mM, } k_{39} = 1 \text{ mM, } k_{40} = 36, k_{41} = 1.5, k_{42} = 1.83 \text{ s}^{-1}, k_{43} = 0.8 \text{ mM, } k_{44} = 0.04 \text{ mM, } k_{45}, k_{46} = 1 \text{ mM, } k_{47} = 0.083 \text{ mM.}

In addition to the enzyme processes (expressions V1, V2, V3, V4), the model describes the constitutive synthesis of the active forms of pathway enzymes (parameters $k_1, k_2, k_3$), constitutively describes predecessor 3PG uptaking (parameter $k_4$), linearly describes enzyme degradation (constants $k_6, k_7, k_8, k_9$), the utilization of the substrate 3PG (constant $k_{10}$) and the utilization of the end products SER and GLY in other pathways in *E. coli* (constants $k_{13}, k_{14}$). Enzyme activity is described (expressions V1, V2, V3, V4) with due regard to regulation exerted on the enzymes by low-molecular compounds (Table 2). For example, the effects of H⁺ concentrations on PGDH activity are taken into account as proposed by Michaelis (Cornish-Bowden, 1979). The concentrations of NAD, NADH, GLU, AKG, THF, MTHF, PI and pH, involved in the regulation of enzyme activity and being a source of groups and energy, are included in the equations as parameters.

The values of the synthesis, degradation and utilization parameters and the physiological values of substance concentrations included in the model as parameters were selected in such a way that they best fit in experimental data and that stationary state of the mathematical model fit in experimental data on substance concentrations and enzyme activity in the gene network for serine and glycine biosynthesis.

Equilibrium concentration values for serine and glycine are shown in Fig. 3A (dotted lines). Curves for the time it takes the system to reach equilibrium after assigning 0 mM to SER or GLY concentrations were calculated using the model. As the calculations demonstrate, SER (solid line) reaches equilibrium (dotted line) in 100 seconds, but does not go beyond, plateauing at about 0.9 mM for 4 to 10 minutes and then declines to equilibrium during 100 seconds. It takes GLY 7–8 minutes to reach equilibrium. The time course of SER can be explained from looking at SER and PHP interactions (Fig. 3C). As a predecessor, PHP activates SER synthesis. SER inhibits PHP synthesis, which suggests that SER down regulates its own synthesis (Fig. 2). Four segments can be recognized in the curve presented in Fig. 3C: 1) The straight line ending with the first turn (coordinates 4, PHP × 0.25, SER (mM). PHP is growing more rapidly than SER, because SER in deficient amounts fails to inhibit PHP synthesis. 2) The segment ending with the turn (coordinates 3.8, PHP × 0.86, SER). SER is growing more rapidly than PHP, because down regulation of PHP synthesis by SER is enhanced. Shortly before reaching the turn, SER passes by the equilibrium point, because excess PHP (Fig. 3B) enhances SER synthesis. SER goes beyond the equilibrium point because SER self-controlled synthesis is temporarily unaffected by PHP synthesis. SER stays beyond until excess PHP has exhausted. 3) The segment ending with the turn (coordinates 0.5, PHP × 0.9, SER). Because excess PHP has exhausted, SER is plateauing beyond the equilibrium point, and PHP is declining. 4) From 0.5, PHP × 0.9, SER on, after excess PHP has entirely exhausted, PHP concentrations again become dependent on SER-controlled SER synthesis and both PHP and SER return to equilibrium. The calculated time it takes SER to return to the physiological level is about 100 seconds, because from that point on SER do not overstep the physiologic limits.

In the future, we are planning to expand the model by including the regulatory mechanisms of gene expression in the serine and glycine biosynthetic pathway.
Figure 2. A general schematic of serine and glycine biosynthesis including synthesis and degradation.

Figure 3. Dynamics calculations: a – serine, glycine and b – PHP. The X axis: time (in seconds); the Y axis: concentrations of serine, glycine and PHP (in mM). c – recovery to equilibrium on the interaction diagram for serine and PHP.

ACKNOWLEDGEMENTS

The authors are grateful to Vitaly Likhoshvai and Aleksandr Ratushny for valuable discussions and Irina Lokhova for bibliographical support. This work was supported in part by the Russian Government (Contract No. 02.467.11.1005), by Siberian Branch of the Russian Academy of Sciences (the project “Evolution of molecular-genetical systems:
computational analysis and simulation” and integration projects Nos 24 and 115), and by the Federal Agency of Science and Innovation (innovation project No. IT-CP.5/001).

REFERENCES


