Gene 518 (2013) 84–90

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Parameter optimization and sensitivity analysis for large kinetic models using a real-coded genetic algorithm

Yukako Tohsato *, Kunihiko Ikuta, Akitaka Shionoya, Yusaku Mazaki, Masahiro Ito

Department of Bioinformatics, College of Life Sciences, Ritsumeikan University, Kusatsu, Shig 525-8577, Japan

ARTICLE INFO

Available online 27 December 2012

Keywords:
Parameter sensitivity
Global parameter sensitivity
Parameter correlation
Kinetic model
Real-coded genetic algorithm

ABSTRACT

Dynamic modeling is a powerful tool for predicting changes in metabolic regulation. However, a large number of input parameters, including kinetic constants and initial metabolite concentrations, are required to construct a kinetic model. Therefore, it is important not only to optimize the kinetic parameters, but also to investigate the effects of their perturbations on the overall system. We investigated the efficiency of the use of a real-coded genetic algorithm (RCGA) for parameter optimization and sensitivity analysis in the case of a large kinetic model involving glycolysis and the pentose phosphate pathway in *Escherichia coli* K-12. Sensitivity analysis of the kinetic model using an RCGA demonstrated that the input parameter values had different effects on model outputs. The results showed highly influential parameters in the model and their allowable ranges for maintaining metabolite-level stability. Furthermore, it was revealed that changes in these influential parameters may complement one another. This study presents an efficient approach based on the use of an RCGA for optimizing and analyzing parameters in large kinetic models.

1. Introduction

Mathematical modeling is a powerful approach for understanding and predicting dynamic behavior in the regulatory mechanisms of metabolic pathways in response to genetic modifications and environmental changes. Several kinetic models of cell metabolism, based on nonlinear ordinary differential equations, have been developed to detect time-dependent changes in metabolic concentrations. For example, models of glycolysis and the pentose phosphate (PP) pathway in *Escherichia coli* K-12 (Chassagnole et al., 2002) and the tricarboxylic acid (TCA) cycle in *Dictyosteliumdiscoideum* (Wright et al., 1992) have been constructed. Furthermore, large-scale model integration has been performed. For example, the glycolysis and PP pathway model developed by Chassagnole et al. (2002) have been integrated with models of the TCA cycle (Kadir et al., 2010; Usuda et al., 2010) and amino acid biosynthesis (Lee et al., 2010).

However, kinetic models require a large number of parameters, including kinetic constants and initial metabolite concentrations. Information about these parameters has been stored in databases, such as BRENDA (Schomburg et al., 2004), SABIO-RK (Wittig et al., 2012), and BioModels (Li et al., 2010). In many cases, however, the parameters stored in databases are insufficient for the construction of an accurate metabolic model, since the kinetic parameters are usually obtained or estimated from measurements reported by different laboratories using different in vitro models and conditions. Parameter estimation and optimization, achieved by comparing the simulation results of a kinetic model and experimental data, are integral parts of kinetic modeling.

Kinetic parameters have distinct solutions (*multimodality*), distinct scales for parameters (*ill-scaling*), and interdependency among subsets of the parameters (*parameter dependency*) under the same objective function. For the optimization of kinetic parameters with these properties, various approaches based on evolutionary algorithms, such as simulated annealing (Chassagnole et al., 2002), evolutionary programming (Costa et al., 2010), and genetic algorithms (Fang et al., 2009; Ishii et al., 2007; Lee et al., 2010; Matsubara et al., 2006), have been adopted.

A genetic algorithm (GA) is a biologically inspired method of function optimization based on evolutionary theory. It is well suited to problems involving numerous parameters, because of its multi-start random-search approach. However, the performance of a GA depends on how it encodes solutions as chromosomes and on its parameters. Real-coded GAs (RCGAs) (Ono and Kobayashi, 1997), which use strings of real numbers for parameters to be optimized as chromosomes, are known to be more efficient for functional optimization than binary-coded GAs (Goldberg, 1989), which use strings of binary bits 0 and 1, with respect to the ability to avoid converging to local minima of the solution space (Ono and Kobayashi, 1997). Previous adaptations of RCGAs seem to be insufficient for evaluating parameters and indicating methodology (Fang et al., 2009; Ishii et al., 2007; Lee et al., 2010; Matsubara et al., 2006).

Furthermore, it is important to understand and quantify the effect of the optimized input parameters on the model output. Sensitivity analysis

* Corresponding author at: Department of Bioinformatics, College of Life Sciences, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan. Tel./fax: +81 77 561 5290.
E-mail address: yukako@sk.ritsumei.ac.jp (Y. Tohsato).

0378-1119/$ – see front matter
http://dx.doi.org/10.1016/j.gene.2012.11.080

is one approach to determine which input parameters have the largest impact on model output (Maggio et al., 2010). There are two types of sensitivity analysis: local and global. Local sensitivity analysis may be used to determine which parameters are relatively important at a single point in the parameter space, whereas global sensitivity analysis seeks a measure of relative importance over the entire parameter space. Global methods have a higher computational cost than local methods, but they provide more realistic results, since parameter interactions can be identified (Maggio et al., 2010). There are many techniques for global sensitivity analysis, including Monte Carlo simulations. Local and global sensitivity analyses have been applied to kinetic models based on Chassagenole's model (Costa et al., 2010; Degenring et al., 2004; Maggio et al., 2010).

In this study, we evaluated RCGA parameters with respect to efficiency of parameter estimation and optimization in a large kinetic model. Parameter sensitivity analyses coupled with the RCGA were used to determine parameters having a large effect on the variability of the system output, lower and upper limits of parameters for the maintenance of metabolic concentrations, and their correlation control.

2. Materials and methods

2.1. Kinetic model structure

The dynamic model of *E. coli* formulated by Chassagnole et al. (2002) was used as a benchmark. This model, which describes the dynamic metabolic behavior of glycolysis and the PP pathway after a glucose pulse, includes 30 enzymatic reactions and 25 metabolites consisting of 18 balanced metabolites and 7 unbalanced cofactors (e.g., atp, adp, and nad). The corresponding metabolic network is shown in Fig. 1. The rate of change of the concentration of a metabolite in this metabolic network is given by the following equation:

$$\frac{dC_i}{dt} = \sum_{j=1} N_{ij} v_j - \mu C_i \tag{1}$$

where C_i is the concentration of the metabolite i, v_j is the rate of the reaction j, N_{ij} is the stoichiometric coefficient of the metabolite i in the reaction j, and μ is the specific growth rate. Thus, the term μC_i represents the dilution effect due to growth. All formulas in the dynamic model of *E. coli* can be found in the original paper (Chassagnole et al., 2002). For example, the mass balance for g6p and f6p metabolites is given by Eqs. (2) and (3), and PTS and PGI biosynthesis are described by Eqs. (4) and (5).

$$\frac{dC_{g6p}}{dt} = v_{PTS} - v_{PGI} - v_{G6PDH} - v_{PGM} - \mu C_{g6p} \tag{2}$$

$$\frac{dC_{f6p}}{dt} = v_{PGI} - v_{PFK} + v_{TKb} + v_{TA} - 2v_{MurSynth} - \mu C_{f6p} \tag{3}$$

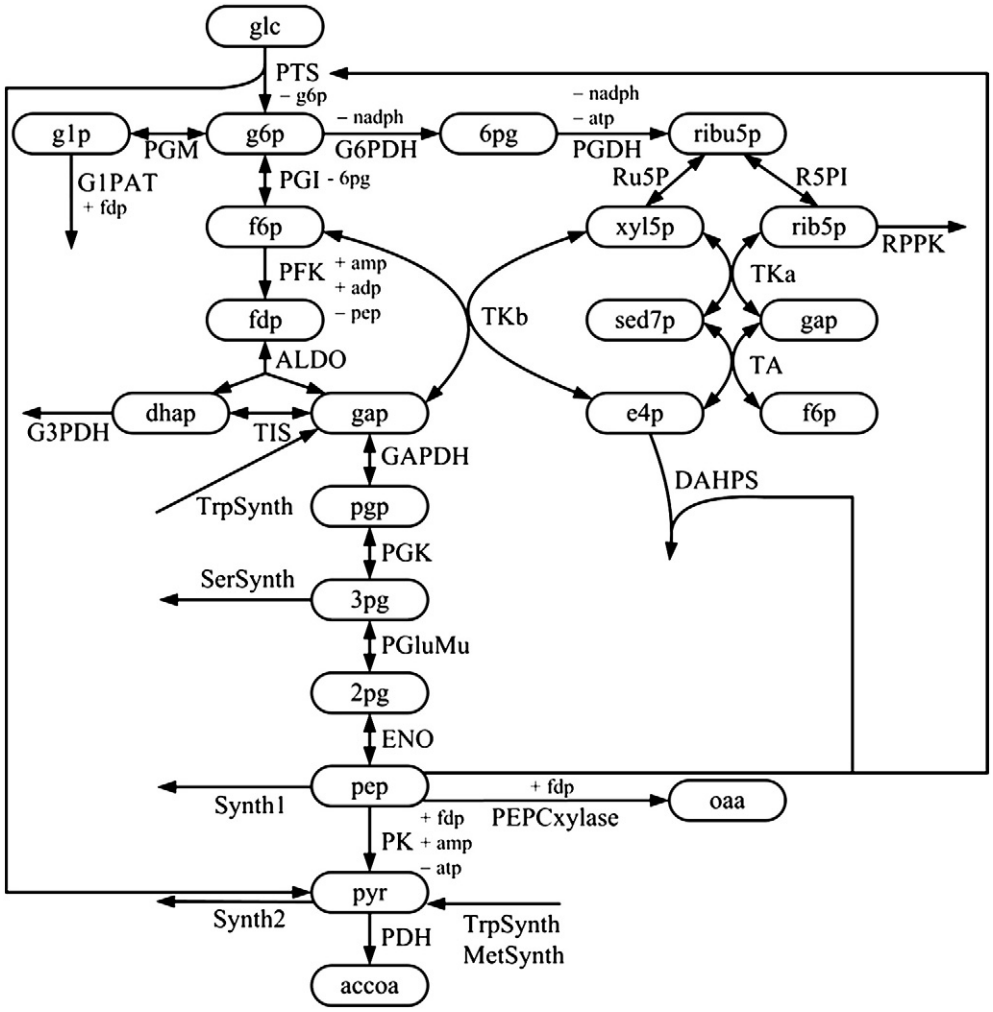

Fig. 1. Glycolysis and pentose phosphate (PP) pathway in *E. coli*. Enzyme names are written in upper case. Arrows indicate the directions of reactions. Metabolite names are written in lower case in ellipses. Names of cofactors, which are written in lower case, are shown beside enzymatic reactions. Positive signs and negative signs indicate activators and inhibitors, respectively. The abbreviations correspond to the formal names given in Supplementary Table 1.