In silico identification of the key components and steps in IFN-γ induced JAK-STAT signaling pathway

Zhike Zi[a], Kwang-Hyun Cho[b,c,*], Myong-Hee Sung[d], Xuefeng Xia[a], Jiashun Zheng[a], Zhirong Sun[a,*]

[a] *Institute of Bioinformatics, MOE Key Laboratory of Bioinformatics, Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China*
[b] *College of Medicine, Seoul National University, Chongno-gu, Seoul, 110-799, Korea*
[c] *Korea Bio-MAX Institute, Seoul National University, Gwanak-gu, Seoul, 151-818, Korea*
[d] *Biometric Research Branch, National Cancer Institute, National Institutes of Health, 6130 Executive Blvd., Bethesda, MD 20892-7434, USA*

Received 13 October 2004; revised 7 December 2004; accepted 5 January 2005

Available online 19 January 2005

Edited by Robert B. Russell

Abstract Systems biology efforts are increasingly adopting quantitative, mechanistic modeling to study cellular signal transduction pathways and other networks. However, it is uncertain whether the particular set of kinetic parameter values of the model closely approximates the corresponding biological system. We propose that the parameters be assigned statistical distributions that reflect the degree of uncertainty for a comprehensive simulation analysis. From this analysis, we globally identify the key components and steps in signal transduction networks at a systems level. We investigated a recent mathematical model of interferon gamma induced Janus kinase-signal transducers and activators of transcription (JAK-STAT) signaling pathway by applying multi-parametric sensitivity analysis that is based on simultaneous variation of the parameter values. We find that suppressor of cytokine signaling-1, nuclear phosphatases, cytoplasmic STAT1, and the corresponding reaction steps are sensitive perturbation points of this pathway.
© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Janus kinase-signal transducers and activators of transcription; Interferon gamma; Robustness analysis; Multi-parametric sensitivity analysis; Systems biology

1. Introduction

Considerable efforts have been made so far in the realm of systems biology for dynamical modeling and systems analysis of cellular signal transduction pathways and other networks. Quantitative mechanism-based models could allow researchers to predict the comprehensive behavior of the specified system over time and to track its dynamics for each set of fixed system parameters [1–8]. However, all of the parameters including rate constants and initial components concentrations in the mathematical models must be experimentally measured or inferred to specify the model. Even for those models with experimentally estimated parameters, it is still uncertain whether the particular set of parameters closely approximates the corresponding biological system because some of the kinetic parameters are usually taken or estimated from measurements reported by different laboratories using different in vitro models and conditions. Given the inherent uncertainties in the structure and parameter values of the models, parameters can be assigned statistical distributions that reflect the degree of uncertainty and then simulation analysis can be performed by sampling from the distributions. It is therefore of vital importance not only to study the dynamical properties governed by the particular kinetic parameters but also to further investigate the effects of their perturbations on the overall system. The purpose of this work is trying to answer the question: which signaling components and rate constants are more critical to the output behavior of the system? Investigation of such a question has been one of the major problems raised in systems biology [9].

In this study, we chose the interferon-γ (IFN-γ) induced Janus kinase-signal transducers and activators of transcription (JAK-STAT) pathway for analysis. IFN-γ, or type II IFN, was first identified in PHA-activated lymphocyte supernatants with distinctive antiviral activity [10] and is a pleiotropic cytokine widely involved in the regulation of both innate and adaptive immune responses. The IFN-γ induced JAK-STAT pathway is a stress-responsive mechanism that transduces signals from the cell surface to the nucleus. The binding of the cytokine to its cell-surface receptor results in receptor dimerization and the subsequent activation of JAK tyrosine kinases, which are constitutively associated with the receptors. The receptors are then phosphorylated by activated JAKs and serve as docking sites for the STAT1. STAT1 is phosphorylated by JAK, dimerizes, and subsequently leaves the receptor and translocates to the nucleus, where it activates gene transcription. The STAT1 dimers in the nucleus can be dephosphorylated to be STAT1 monomers and transported back to the cytosol by nuclear export [11,12]. Dysregulation of JAK-STAT

*Corresponding authors. Fax: +82 2 887 2692 (K.-H. Cho), Fax: +86 1062772237 (Z. Sun).
E-mail addresses: ckh-sb@snu.ac.kr (K.-H. Cho), sunzhr@mail.tsinghua.edu.cn (Z. Sun).

Abbreviations: IFN-γ, interferon gamma; JAK, Janus kinase; STAT, signal transducers and activators of transcription; IFNR, interferon-γ receptor; RJ, IFNR–JAK complex; STAT1, signal transducer and activator of transcription 1; SHP-2, SH2 domain-containing tyrosine phosphatase 2; SOCS1, suppressor of cytokine signaling-1; PPN, nuclear phosphatase; PPX, unidentified phosphatase in the cytoplasm; MAPK, mitogen-activated protein kinase; MPSA, multi-parametric sensitivity analysis

0014-5793/$30.00 © 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2005.01.009

signaling is associated with various immune disorders and cancers. The signaling strength, kinetics, and specificity of the JAK-STAT pathway are modulated at many levels by distinct regulatory proteins including the suppressor of cytokine signaling (SOCS) proteins, SH2 domain-containing tyrosine phosphatase 2 (SHP-2), and various cytoplasmic and nuclear protein tyrosine phosphatases (PTPs) [13–16]. In this study, we use STAT to represent STAT1 in particular, in the absence of kinetic data distinguishing rate constants for the different STAT isoforms. The basic steps and regulatory scheme of JAK-STAT pathway are shown in Fig. 1.

Here, we propose a global approach for systematic analysis of the JAK-STAT signaling pathway against variations in kinetic parameters and initial concentrations of signaling proteins. The multi-parametric sensitivity analysis (MPSA) method used in this study is based on a Monte-Carlo method over a broad range of simultaneous variation of parameters in uniform distribution followed by a statistical assessment. With this method, we globally identify the key components and steps that are critical to the dynamical behaviors of this signaling pathway.

2. Materials and methods

2.1. The mathematical model

We employ the mathematical model developed by Satoshi Yamada et al. in 2003 [17] for the IFN-γ induced JAK-STAT signaling pathway in liver cells. Since it does not include synthesis of new transcription factors, the direct transcriptional activation by this signaling pathway is to be referred to as the primary IFN-γ response [16]. Fig. 1 and Supplementary Figure 1 show the dynamic scheme of this pathway and all the biochemical reactions included in the model. The model is constructed by ordinary differential equations composed of 32 state variables and 51 parameters. Detailed chemical reactions as well as their parameters are described in Supplementary Table 1.

Experimental studies have shown that phosphorylated STAT1 dimers in the nucleus (STAT1n∗–STAT1n∗) mediates and is necessary, although not sufficient, for the induction of IFN-γ-inducible genes [11,18,19]. Therefore, STAT1n∗–STAT1n∗ can be regarded as an indirect indicator for target gene activation and we considered STAT1n∗–STAT1n∗ as the output of the signal transduction system in our analysis. The simulated time course of STAT1n∗–STAT1n∗ using the reference set of parameters shows that it is detected within 15 min and reaches its maximum between 30 min and one hour, and then it decreases by SOCS1 action (see Fig. 2). Longer simulations showed that STAT1n∗–STAT1n∗ arrives at a steady state after 8 h.

2.2. Multi-parametric sensitivity analysis

The MPSA method was proposed by Hornberger and Chang [20,21] and further developed by Choi et al. [22] in the field of hydrology. MPSA is a tool that can be used to define the relative importance of the factors related to the model [23]. The idea of MPSA is to inject uncertainty of the parameters into the model by randomly selecting parameter values from probability distributions rather than using fixed values. This is achieved using a Monte-Carlo method in which the model is run repeatedly using sets of parameters drawn randomly from the distributions. Because the natural distributions of parameter values for real biological systems are unknown, we used uniform probability distribution [24]. The range of the parameter distributions are usually determined from the available literature or guided by the experiences of the researchers. For the MPSA with respect to the rate constants, due to the large number (51) of parameters to vary simultaneously, it was necessary to sample a representative set from all possible combinations of parameter values. Latin hypercube sampling method was used to generate random sets of parameter values for simulations in this case (see below). A criterion is coded into the algorithm to classify the output of each model simulation as either acceptable or unacceptable. The final step of MPSA is statistical evaluation of the occurrences of the acceptable and unacceptable cases, summarized for each parameter. The larger the difference between the cumulative frequencies of the two cases, the more significant is the given parameter. The detailed procedure of MPSA is described in the following:

Step 1. Select the parameters to be tested.

Step 2. Set the range of each selected parameter large enough to cover all feasible variations.

Step 3. For each parameter, generate a series of independent, random numbers from a uniform distribution within the defined range and obtain parameter combinations (see below for sampling methods).

Step 4. Simulate the model for each chosen set of parameter values and calculate the corresponding objective function. The objective function is defined as the sum of squared errors between the observed and perturbed system output values. That is

$$f_{\mathrm{obj}}(k) = \sum_{i}^{n} (x_{\mathrm{obs}}(i) - x_{\mathrm{cal}}(i,k))^2 \tag{1}$$

where f_{obj} is the objective function that describes how much the system output deviates from the observed data by varying the parameters, $x_{\mathrm{obs}}(i)$ denotes an observed system output value at the ith sampling time (this is to be substituted by the simulation result from the reference parameter values), $x_{\mathrm{cal}}(i,k)$ denotes the perturbed system output value at the ith sampling time for the parameter variation set k, and n is the number of sampling time points. We set 50 sampling time points in our analysis.

Step 5. Determine whether the chosen set of parameter values is 'acceptable' or 'unacceptable' by comparing the objective function value to a given threshold. If the objective function value is greater than the threshold, the set of parameter values is classified as 'unacceptable'. If the value is less than the threshold, it is classified as 'acceptable'. A previous work [22] indicated that MPSA results are not affected by the choice of a subjective threshold and here we used the average of the objective function over all parameter variations as the threshold value.

Step 6. Statistically evaluate the parameter sensitivity. To this end, we quantitatively compare the distributions of the individual parameter values associated with the acceptable and the unacceptable cases. For each selected parameter, the cumulative frequency is computed for both acceptable and unacceptable cases. We evaluate the sensitivity by a direct measure of the separation of the two cumulative frequency distributions. We use the following Kolmogorov–Smirnov (K–S) statistic:

$$\text{K–S} = \sup_{x} |S_{\mathrm{a}}(x) - S_{\mathrm{u}}(x)| \tag{2}$$

where S_{a} and S_{u} are the cumulative frequency functions corresponding to acceptable cases and unacceptable cases, respectively, and x is the given parameter. The statistic K–S is determined as the maximum vertical distance between the cumulative frequency distribution curves for n acceptable and m unacceptable cases. A larger value of K–S indicates that the system is sensitive to variation in the given parameter.

In Step 3 for the 51 rate constants, selecting just two values for each parameter would generate 2^{51} simulations to run, which is not practical. Instead, we used the Latin hypercube sampling method to sample 2000 random parameter vectors while evenly covering individual parameter ranges (some background information about Latin hypercube sampling method is available in the Supplementary material). This way we could computationally manage the large number (51) of rate constants being varied simultaneously, while ensuring maximal sampling through each parameter dimension [25]. Briefly, for the jth parameter, we divide the range of the parameter into N (= 2000) subintervals of equal size. Then randomly sample N values (p_{ij}, $i = 1,\ldots,N$), one from each subinterval, for the jth parameter. To combine these values of individual parameters to generate sets of parameter values, we randomly permute the N values for each parameter to get the parameter vectors, i.e., we individually permute the elements of each column of the matrix p_{ij} and use the N rows as the parameter

Download English Version:

https://daneshyari.com/en/article/10873634

Download Persian Version:

https://daneshyari.com/article/10873634

Daneshyari.com