Experimental Design in Systems Biology, Based on Parameter Sensitivity Analysis Using a Monte Carlo Method: A Case Study for the TNFα-Mediated NF-κB Signal Transduction Pathway

Kwang-Hyun Cho
Sung-Young Shin
School of Electrical Engineering
University of Ulsan
Ulsan, 680-749, Korea
ckh@mail.ulsan.ac.kr

Walter Kolch
Beatson Institute for Cancer Research
Cancer Research UK
Beatson Laboratories
Switchback Road, Glasgow, G61 1BD, United Kingdom
Institute of Biomedical and Life Sciences
University of Glasgow
University Avenue, Glasgow, G12 8QQ, United Kingdom

Olaf Wolkenhauer
Department of Biomolecular Science and
Department of Electrical Engineering & Electronics
UMIST, Manchester, M60 1QD, United Kingdom
olaf.wolkenhauer@umist.ac.uk

Mathematical modeling and dynamic simulation of signal transduction pathways is a central theme in systems biology and is increasingly attracting attention in the postgenomic era. The estimation of model parameters from experimental data remains a bottleneck for a major breakthrough in this area. This study's aim is to introduce a new strategy for experimental design based on parameter sensitivity analysis. The approach identifies key parameters/variables in a signal transduction pathway model and can thereby provide experimental biologists with guidance on which proteins to consider for measurement. The article focuses on applying this approach to the TNFα-mediated NF-κB pathway, which plays an important role in immunity and inflammation and in the control of cell proliferation, differentiation, and apoptosis. A mathematical model of this pathway is proposed, and the sensitivity analysis of model parameters is illustrated for this model by employing the Monte Carlo method over a broad range of parameter values.

Keywords: Systems biology, experimental design, signal transduction pathway, mathematical modeling, parametric sensitivity, Monte Carlo method, TNFα, NF-κB

1. Introduction

Systems biology is attracting increasing numbers of researchers from various disciplines. The current interest is not due to particular biological breakthroughs facilitated by mathematical models but is largely based on the promise that signal- and systems-oriented approaches have. One area in which mathematical modeling and simulation has been considered for some time is the inter- and intradynamics of cell signaling. Once a reasonable mathematical model for even a small part of the pathway is established, then the potential benefits are considerable—supporting

SIMULATION, Vol. 79, Issue 11–12, November–December 2003
xxx-xxx

DOI: xxx

experimental design, generating hypotheses, and suggesting experiments to test them. Building such a mathematical pathway model comprises mainly two aspects: (1) deciding on the model structure and (2) estimating the involved parameter values. In general, it is difficult to formalize the derivation of a suitable model structure since signal transduction pathways almost certainly have nonlinear dynamics. However, one possible systematic way to decide on a model structure was proposed [1–3] by considering the fact that a protein involved in the pathway has, in most cases, a catalytic role. Given a model structure, parameter estimation remains the limiting step in the modeling and simulation of signal transduction pathways [4]. Most problems are due to a lack of quantitative measurements and data sets covering all variables involved. Another, maybe even greater, complication is posed by the fact that measurements of protein activities are done with purified proteins in vitro. The parameters measured may not reflect the "true" activity in the cell. However, for technical reasons, it is very difficult to measure accurate protein activities in cells. For all of those reasons, the experiments will continue to challenge the methodologies that are being developed for dynamic pathway modeling and simulation. With parameter estimation being the most pressing problem in modeling, the principal aim of this study is introducing a new strategy for experimental design based on parameter sensitivity analysis and providing guidance on which proteins to measure and why.

Signal transduction pathways enable the cell to receive, process, and respond to biochemical stimuli (information). The components of a pathway interact not only with each other but also with components of other pathways, resulting in complex signaling networks. Cellular signal transduction pathways can be constructed with experimentally determined coefficients and analyzed by computational methods to understand their role in complex biological processes [5, 6]. Mathematical modeling of signal transduction pathways can support quantitative hypothesis testing. Simulation can be further used to detect unexpected state components and to predict concentration profiles for different experimental designs. Ultimately, the development of computational models and the integration of these models with experiments can provide valuable insight into the complex systems-level behavior of pathways [7–9]. Bhalla and Iyengar [5] proposed a computational model of the cAMP signal transduction pathway by using the general neural simulation system, GENESIS. Schoeberl et al. [10] developed a mathematical model describing the dynamics of the epidermal growth factor (EGF) signal transduction pathway by using ordinary differential equations (ODEs), implemented using Matlab libraries. Schoeberl, Gilles, and Scheurich [11] also investigated a mathematical model of the TNFα receptor interaction by using the simulation tool DIVA, and the computational simulation illustrated that the apoptotic cross-talk of TNFR1 and TNFR2 depends on TRAF2 depletion and that RIP plays a key role in the cross-talk.

A study of the literature suggests that to complete the mathematical modeling of a signal transduction pathway, parameter estimation is essential. There have been in-depth studies regarding parameter estimation, which has led to several developments in general methods of the parameter estimation of nonlinear dynamic systems [12–18] and in numerical estimation methods, such as the principal axis method and the simplex method [19, 20]. However, there still remain difficulties due to the measurement noise and according to the increasing number of parameters to be estimated. Using conventional statistical approaches, we would require enormous time-course measurements for all of the proteins involved in the signal transduction pathway. Long time-course experiments with large numbers of time points and replicate measurements, as well as those covering all variables in a pathway, are very difficult to obtain if not impossible. For example, the mathematical model of the TNFα signal transduction pathway, presented in this article, and the EGF signal transduction pathway in Schoeberl et al. [10] consist of 31 and 94 state variables, respectively. This means that it would be necessary to measure concentration changes via time-course measurements for at least about 24 and 70 proteins, respectively, even after eliminating redundancies of the model equations from structural properties (mainly introduced by feedback loops) based on identifiability concepts [12, 13, 17, 19, 20]. To overcome this hurdle, we investigate a new approach to experimental design that first identifies relatively sensitive and hence important parameter values by applying multiparametric sensitivity analysis (MPSA) to the mathematical model. This allows us to design an experiment that measures changes of protein concentration only for key variables in the system.

Figure 1 illustrates the steps taken in the development of a mathematical pathway model based on Phair and Misteli's scheme [7] for the role of mathematical models in quantitative hypothesis testing. The dashed lines and arrows in Figure 1 illustrate the part of the proposed parametric sensitivity analysis in relation to mathematical modeling, parameter estimation, and experimental design. We can summarize the proposed experimental design, modeling, and simulation procedure as follows: (1) assemble published relevant experimental data and qualitative information to decide on a model structure, (2) identify relatively sensitive (i.e., important) parameters through sensitivity analysis of model parameters and define a hypothesis, (3) design an experiment to measure the corresponding changes in protein concentrations based on the sensitivity analysis, (4) conduct the experiment and thereby gather time-course-measured data, (5) estimate kinetic parameter values from the measured data, (6) simulate the mathematical model with estimated parameter values and study the dynamics of the signal transduction pathway, (7) compare the simulation results with the experimental data, and (8) repeat the validation of the model until the simulation results are in accord with the experimental data.

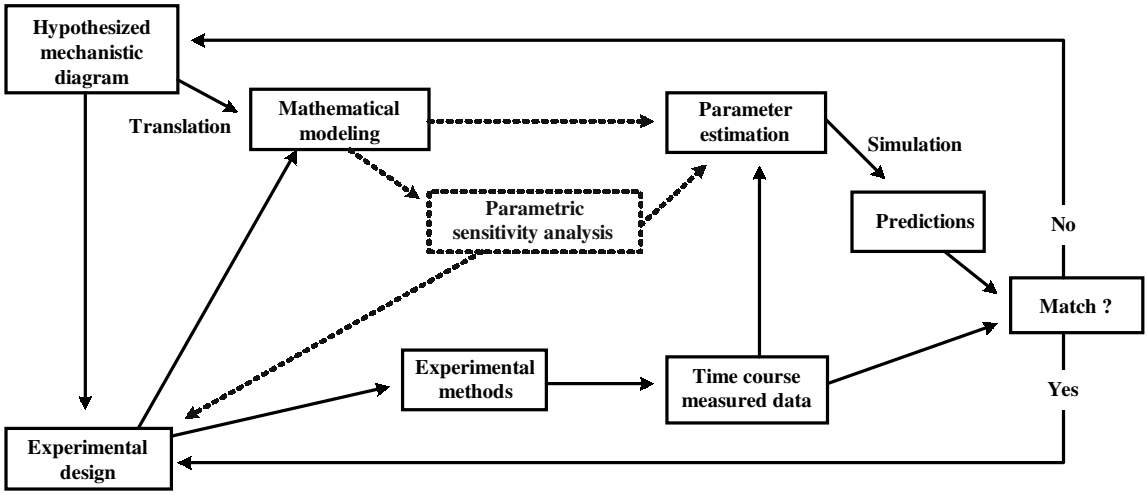

Figure 1. Modified Phair's scheme for mathematical modeling, parameter sensitivity analysis, and experimental design. The dashed lines and arrows indicate the role of the proposed parametric sensitivity analysis in relation to mathematical modeling, parameter estimation, and experimental design.

Sensitivity analysis is a general-purpose technique that is often used to analyze how sensitive a system is with respect to (w.r.t.) the change of parameter values. This parameter sensitivity analysis can be used to validate the model's response and to design experiments that support the estimation of parameters. Shi [21] presented a statistical mechanic model for cell receptors' enhancement of clustering and oligomerization. To explore the extent of clustering and to calculate the signaling, Monte Carlo simulations using the Metropolis algorithm were studied. Keasling, Kuo, and Vahanian [22] proposed a mathematical model for the *Escherichia coli* cell cycle and employed a Monte Carlo simulation method to study the initiation at each increment in cell mass growth, which can account for the initiation of chromosome replication. Sensitivity analysis was also used to identify relatively sensitive parameters in the multiparametric system of the natural attenuation of mining constraints and volatile organic compounds [23, 24] and was referred to as MPSA based on Monte Carlo simulation.

This article investigates the TNFα-mediated NF-κB signal transduction pathway using MPSA. TNFα is a potent proinflammatory cytokine that plays an important role in immunity and inflammation, as well as in the control of cell proliferation, differentiation, and apoptosis [25–28]. After we identify the most sensitive parameters, we analyze how sensitive the TNFα-mediated NF-κB signal transduction pathway is w.r.t. the changes of these parameter values.

The article is organized as follows. Section 2 briefly introduces the multiparametric sensitivity analysis technique, together with an illustrative example. Section 3 describes the mathematical modeling of the TNFα signal transduction pathway. Section 4 shows the application of MPSA to the TNFα signal transduction pathway. Section 5 summarizes the results from sensitivity analysis and the corresponding simulations. Finally, conclusions and suggestions for further studies are made in section 6. The detailed mathematical model of the TNFα signal transduction pathway, the nominal values and the range of parameters for the simulation, and the initial value of each protein for simulation are summarized in the appendix.

2. Parametric Sensitivity Analysis

Figure 2 illustrates a basic graphical "template model" of the signal transduction pathway [1–3, 29]. This template component accounts for one step in the signal transduction of a signaling cascade. In Figure 2, an enzyme E combines with substrate S to form an ES complex with an association coefficient k_1. The complex ES has two possible fates. It can be dissociated into E and S with a dissociation coefficient k_2, or it can further proceed to form a product P with a production rate coefficient k_3. This basic template model is to be used to illustrate how we employ the MPSA for the parametric sensitivity analysis of a signal transduction pathway. The corresponding mathematical model using differential equations is as follows:

$$\begin{aligned}
\dot{m}_1(t) &= -k_1 \cdot m_1(t) \cdot m_2(t) + k_2 \cdot m_3(t),\\
\dot{m}_2(t) &= -k_1 \cdot m_1(t) \cdot m_2(t) + k_2 \cdot m_3(t) + k_3 \cdot m_3(t),\\
\dot{m}_3(t) &= k_1 \cdot m_1(t) \cdot m_2(t) - k_2 \cdot m_3(t) - k_3 \cdot m_3(t),\\
\dot{m}_4(t) &= k_3 \cdot m_3(t),
\end{aligned} \tag{1}$$

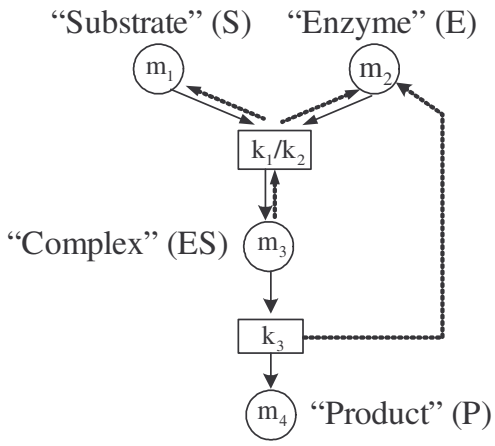

Figure 2. A graphical basic template model of a single step in the signal transduction pathway. A circle represents a state variable (protein concentration), a rectangle represents the relevant kinetic parameter(s), solid arrows indicate the forward reactions, and dotted arrows indicate the inverse reactions.

where m_i with $i = 1, 2, 3, 4$ are the state variables (m_1 for S, m_2 for E, m_3 for ES, and m_4 for P, respectively); k_j with $j = 1, 2, 3$ are the system parameters; and $\dot{m}_i(t) = \frac{dm_i(t)}{dt}$ denotes derivatives w.r.t. time. Figure 3 shows the computational simulation results of the basic signal transduction pathway. At an early stage, the substrate and the enzyme decrease, but the ES complex and the product increase along with time. However, the ES complex and the enzyme become turned over **[TURNED OVER?]** around 1.5 minutes passed **[PLS. CLARIFY]**. This is because the enzyme first binds to the substrate to catalyze it, and after catalysis, the enzyme returns to the initial state.

In general, the system parameters k_1, k_2, and k_3 of the template signal transduction pathway in Figure 2 can be regarded as random variables since the parameter values depend on unknown variations in the cell environment. Furthermore, the dynamics of the system can be changed in turn according to the variation of the parameter values. However, we note that the dynamics must be differentially influenced by the variation of each parameter. Starting from this point of view, we consider the following problems: (1) How can we build a "core" mathematical model of the given signal transduction pathway given the limited experimental data available? (2) How can we provide guidance for an experimental design that generates experimental data for the "core" mathematical model? Here, the "core" model means a simplified mathematical model that still captures the main (i.e., the most important) characteristics of the pathway. To tackle these problems, we propose first a sensitivity analysis of the model w.r.t. parameters, providing a ranked list that is used to select those variables for which to measure changes in protein concentrations. These data are then used for the estimation of parameter values of the model.

The sensitivity analysis is a general technique that is widely used to analyze how sensitive a system is w.r.t. the changes in parameter values and is usually defined as the ratio of the change in the system transfer function to the change of a parameter value [30]. This concept also can be extended to nonlinear systems, such as a cellular signal transduction pathway, by introducing a sensitivity function and sensitivity equation [31]. Hence, without loss of generality, the sensitivity gain can be written as

$$S_P^M = \frac{\partial M/M}{\partial P/P} = \frac{\text{Percentage change in M}}{\text{Percentage change in P}}, \qquad (2)$$

where M represents the system (output) function, and P denotes one of the system parameters. MPSA is employed to identify the parameters for which the system is most sensitive. The approach is based on the Monte Carlo method [23, 32–34] over a broad range of each parameter value. The procedure of MPSA includes the following steps:

Step 1: Select the parameters to be tested.

Step 2: Set the range of each selected parameter value large enough to cover all feasible variations, guided by experience, measurements, and the literature. We suggest here to set the range between one-fifth of a nominal value and five times the nominal value.

Step 3: Generate a series of independent random numbers with a uniform distribution within the range.

Step 4: Simulate the model for each set of parameter values and calculate the corresponding objective function.

Step 5: Determine whether the chosen set of parameter values is acceptable or unacceptable by comparing the objective function value to a given criterion. If the objective function for a specific set of parameters is greater than the criterion, the set of parameter values is classified as unacceptable; otherwise, it is classified as acceptable. Here the criterion is defined as the average of the objective functions for the entire range of parameters.

Step 6: Evaluate the parametric sensitivity by comparing two distributions of the parameter values associated with the acceptable and the unacceptable results. Here we consider simply "cumulative frequencies," respectively, for each parameter via corresponding correlation coefficients. If the two distributions are not similar (w.r.t. a relatively large correlation coefficient), the parameter is classified as insensitive; otherwise, the parameter is classified as

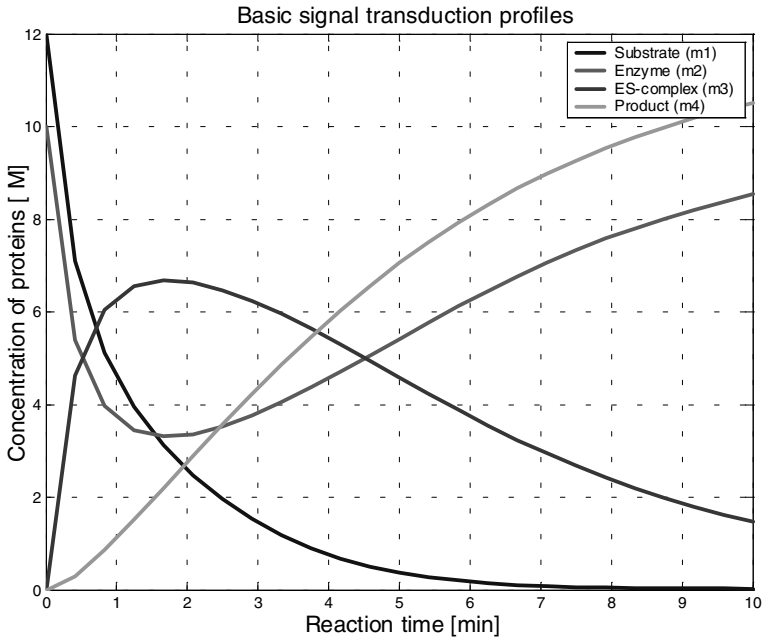

Figure 3. Simulation results of the template single-step signal transduction pathway

sensitive. Relative importance can be further evaluated statistically if needed.

The objective function for the sensitivity analysis is defined as a sum of squared errors between observed and perturbed system output values:

$$f_{obj}(k) = \sum_{i=1}^{q} \left(f_{obs}(i) - f_{per}(i,k) \right)^2, \qquad (3)$$

where $f_{obj}(k)$ is the objective function that describes how much the system (model) output deviates from the observed data w.r.t. the variation of parameters; $f_{obs}(i)$ denotes an observed output value at the ith sampling time, which is to be substituted by a simulation result from the nominal parameter values; $f_{per}(i,k)$ denotes the perturbed output value at the ith sampling time with a variation of k, which is a simulation result according to the variation of the parameter k; and q is the number of sampling time points. If the objective function is less than a specified criterion, then the result is classified as acceptable or unacceptable when the criterion is defined as the average of all objective function values. The proposed definition of the criterion implies the average of the perturbed output values (i.e., differences compared with that of nominal parameters) over the whole variation of parameters such that it forms a basis for the measurement of relative perturbation for each set of parameter values. This is because the main purpose of comparison is to discern the degree of relative change of the objective function for each parameter value. This is only one example of formulating the criterion; another way of formulation is also possible. However, note that the relative sensitivity analysis result will still remain the same even if we choose 33, 50, or 66 divisions of the sorted objective functions as criteria (as suggested in the original MPSA method). This qualitatively justifies the proposed definition of the criterion in that the criterion supports the proposed idea of identifying "relatively" big change.

This idea of parameter sensitivity analysis is to be illustrated for the basic template model of the signal transduction pathway in Figure 2. In this example, the concentration of the "product" is regarded as an output of the system. We first select k_1 and k_3 as the parameters to be analyzed by sensitivity analysis and assume that the range for parameter variations is $k_1 \in [0.092, 0.368][s^{-1}M^{-1}]$ and $k_3 \in [0.14, 0.318][s^{-1}]$, respectively, whereas k_2 is fixed at $0.016[s^{-1}]$. For the sensitivity analysis of parameters k_1 and k_3, we generate 15 random numbers with a uniform distribution within the specified ranges of parameters and perform a simulation study to calculate the objective function values over the set of selected parameter values. Figure 4 shows the resulting objective function distribution w.r.t. the two parameters k_1 and k_3. The upper two graphs in Figure 5 show the frequency distributions of two classes—unacceptable and acceptable, respectively—and the lower graphs in Figure 5 illustrate the corresponding cumulative frequency distributions for k_1 and k_3. The frequency distribution in the upper left graph of Figure 5 represents the

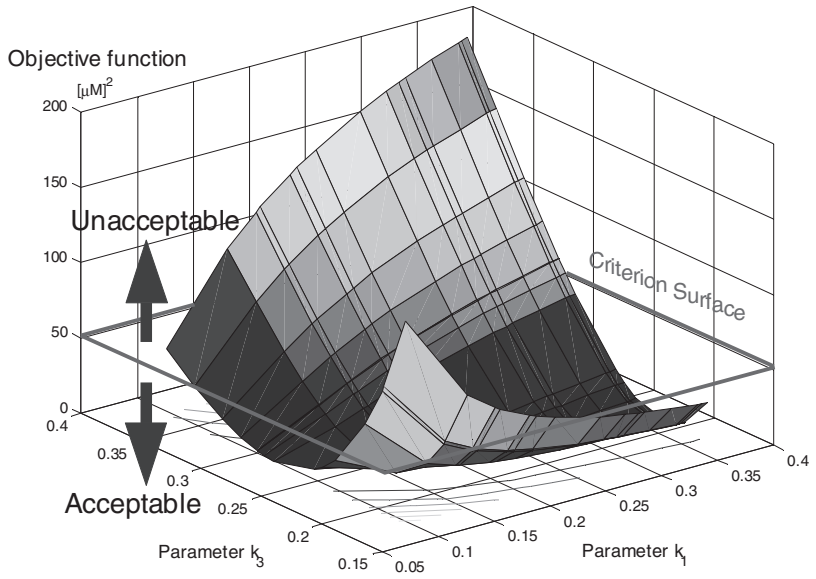

Figure 4. Objective function distribution of the parameter sensitivity analysis with respect to the parameters k_1 and k_3 in the basic template model of the signal transduction system

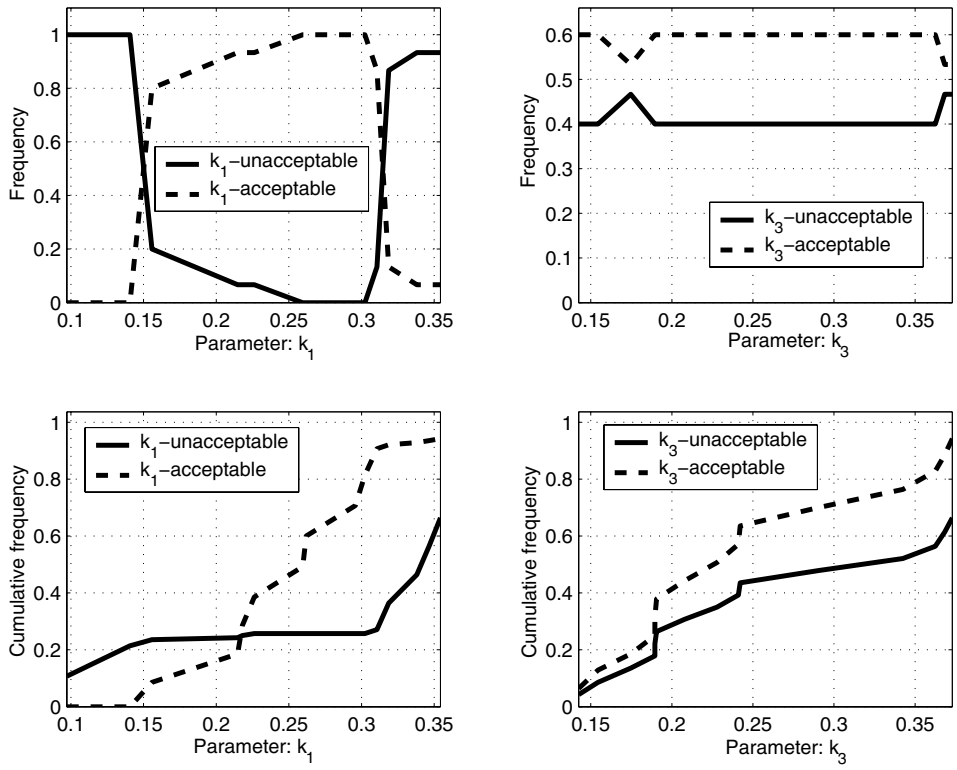

Figure 5. (Cumulative) frequency distributions of the parameter sensitivity analysis results for the example. The solid line denotes the acceptable case, and the dashed line indicates an unacceptable case.

frequency of the objective function value to be classified as unacceptable and acceptable, according to the variation of k_3 at each value of k_1 and, similarly, to the variation of k_1 at each value of k_3 (see upper right graph, Fig. 5). The cumulative frequency distributions in the lower plots of Figure 5 represent the accumulation of the frequencies for k_1 and k_3, respectively. For the purpose of a quantitative comparison of the differences of cumulative frequency distributions, the corresponding correlation coefficients are computed. The correlation coefficient of the cumulative frequency distributions for k_1 is 0.7107 and that for k_3 is 0.9995, respectively. We can therefore conclude that the resulting distribution for the parameter k_1 differs to a greater extent than that of k_3; in addition, the parameter k_1 can be classified as more sensitive, and the parameter k_3 can be classified as insensitive. As the result of this analysis, we find that the association coefficient of the substrate and enzyme k_1 is relatively more sensitive (hence more important) than the production rate coefficient k_3, which means that the output of the system is more dependent on the association coefficient k_1 for this system. This resembles and confirms the empirical observation that in classical Michaelis-Menten enzyme kinetics, changes in the Km (i.e., the concentration of reaction partners) have a relatively greater impact than changes in the V_{max}, the maximal velocity of the reaction.

3. A Case Study for the TNFα-Mediated NF-κB Signal Transduction Pathway: The Structure of the Mathematical Model

TNFα is a potent proinflammatory cytokine that plays an important role in immunity and inflammation, as well as in the control of cell proliferation, differentiation, and apoptosis. Figure 6 illustrates the so-called "circuit diagram" describing the TNFα-mediated NF-κB signal transduction pathway and is based on the graphical representation introduced in Figure 2. In addition, Figure 6 implies the underlying mathematical model of the signal transduction pathway, which is summarized as the set of nonlinear ODEs (4) in the appendix. The graphical model in Figure 6 is structured into an apoptosis module, a cell proliferation module, and a regulation module. In general, TNFα exerts its effects through two distinct receptors, TNFR1 (represented by m_2 in the model) and TNFR2. Binding of the inherently trimeric TNFα (m_1) to TNFR1 and TNFR2 induces receptor trimerization and recruitment of several signaling proteins to the cytoplasmic domains of the receptor. The first protein recruited to TNFR1 is TRADD (m_4), which serves as a platform to recruit at least three additional mediators: FADD (m_{21}), RIP1 (m_{10}), and TRAF2 (m_6). Even though TRAF2 and RIP1 can be independently recruited to the TNFR1 complex (m_5), TNFα-mediated IKK **[PLS. SPELL OUT]** activation depends on the interaction between RIP1 and IKKγ regulatory subunit, whereas TRAF2 is required for the recruitment of the IKK complex to the activated TNFR1 (m_9). This complex also contains the IKKα- and IKKβ-catalytic subunits. FADD binds to TRADD and induces the caspase cascade via the binding and proteolytic activation of caspase-8 (m_{22}). Thus, caspase-8 is the initiating caspase, whereas caspase-3 and caspase-7 represent major executional capases (so-called an effector caspase (m_{29})) that cleave protein targets during apoptosis. NF-κB (m_{17}) is regulated primarily by phosphorylation of the inhibitory protein IκB (m_{14}). In response to TNFα and other agonists, IκB is phosphorylated by the IKK complex (m_{13}), resulting in ubiquitination, degradation of IκB, and nuclear translocation of the released NF-κB. The antiapoptotic activity of NF-κB depends on gene induction. In fact, NF-κB induces the expression of a number of genes whose products can inhibit apoptosis: c-IAP, FLICE, A1, TRAF1, and TRAF2. The most intensively studied protein is c-IAP (m_{31}), which directly binds and inhibits the effector caspases such as caspase-3 and caspase-7. Several key components in the NF-κB activation pathway seem to be targeted by the caspases, resulting in the termination of its antiapoptotic activity. At least two of the signaling proteins are involved in the TNFα-induced NF-κB activation—RIP and TRAF2, which are caspase substrates. In particular, caspase-8 cleaves RIP. This proteolysis, which produces an NH2-terminal truncated fragment (RIPc (m_{20})) and a COOH-terminal truncated fragment (RIPn (m_{20})), eliminates the ability of RIP to activate IKKs and hence NFκB.

We can summarize the signal transduction mechanism of the pathway as follows: the apoptosis module comprises signal cascades of the programmed cell death signal to the effector protein that cleaves the DNA in the nucleus. The cell proliferation module is involved in cell proliferation. The regulation module controls cell death and is thus an important mechanism to sustain the homeostasis of the cell.

4. A Case Study for the TNFα-Mediated NF-κB Signal Transduction Pathway: Sensitivity Analysis w.r.t. Parameters and Simulation

The nominal value of each parameter, the defined range of variation, and the initial value of each signaling protein concentration for simulation are summarized in Tables 1 and 2. The proposed parametric sensitivity analysis is performed for selected system parameters such as the association coefficient of TNFα and TNFR1 (k_1), the association coefficient of TRADD and the TNFα/TNFR1 complex (k_3), the association coefficient of TRAF2 and the TNFα/TNFR1/TRADD complex (k_7), the association coefficient of FADD and the TNFα/TNFR1/TRADD complex (k_{20}), the association coefficient of RIP1 and caspase-8 (k_{17}), and the production rate coefficient of NF-κB and c-IAP (k_{19}). This is because these parameters are known to play a critical role in the signal transduction pathway. The nominal value of each parameter and the initial value of each signaling protein concentration are excerpted from the accumulated knowledge on this pathway in the literature [35, 36] and prior experiences.

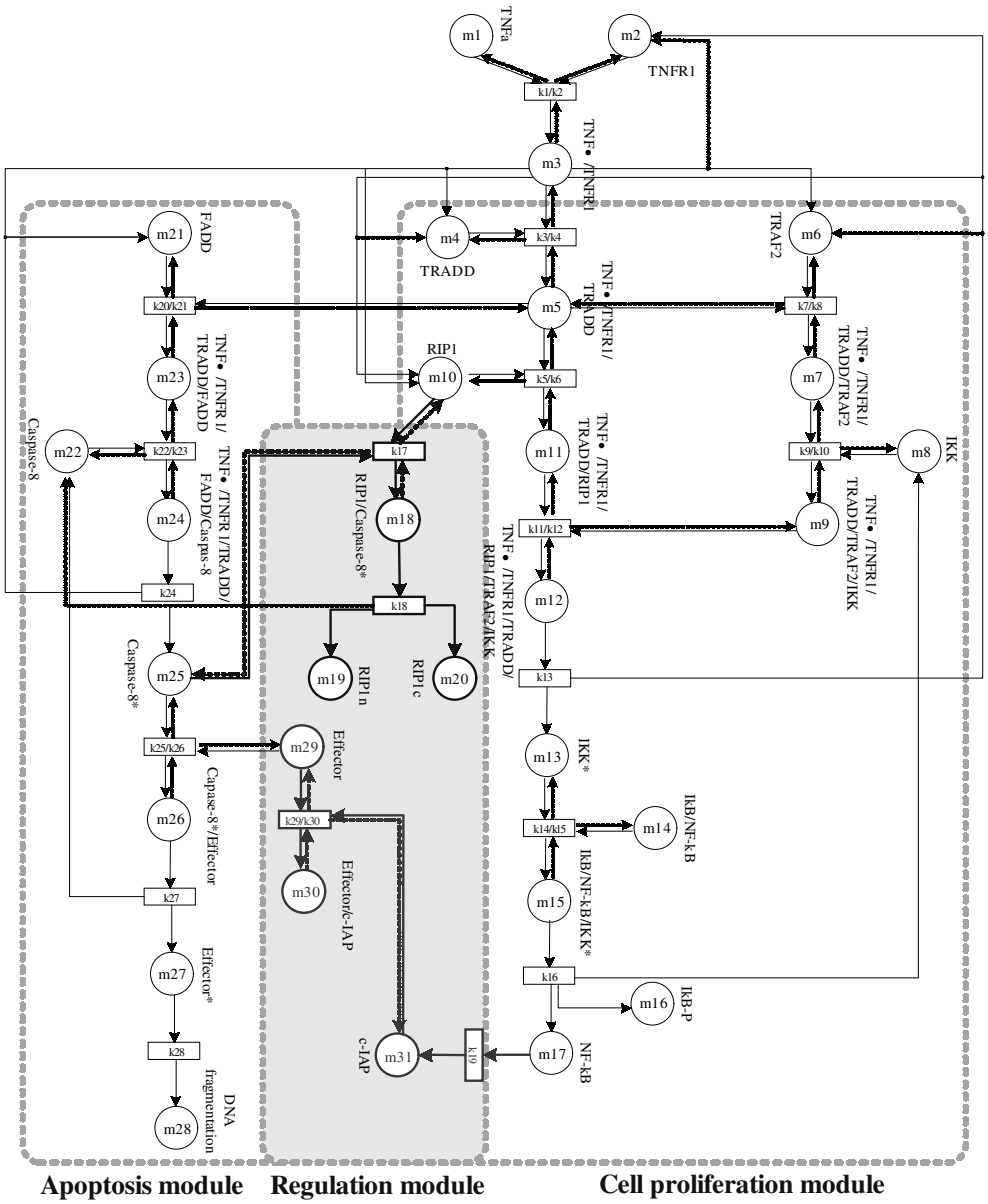

Figure 6. A graphical model (circuit diagram) of the TNFα-mediated NF-κB signal transduction pathway composed of an apoptosis module, a regulation module, and a cell proliferation module

The resulting cumulative frequency distributions for the acceptable and unacceptable cases of the selected parameters (30 samples are chosen from each parameter interval in this case) are shown in Figure 7. The relative sensitivity is determined based on the statistical difference of these distributions. To do this in a quantitative manner, the corresponding correlation coefficient of the cumulative frequency distributions for each selected parameter is computed as follows: the correlation coefficient is 0.061 for k_1, 0.9996 for k_3, 1.0 for k_7, 0.9994 for k_{17}, 0.9995 for k_{19}, and 0.9995 for k_{20}. We can therefore conclude that the parameter k_1 is most different from the others and accordingly classify the parameter k_1 as the most sensitive (important) parameter of the mathematical model in Figure 6. This implies that the output of the signal transduction system—the concentration of NF-κB (m_{17}), in this case—is mostly dependent on the association coefficient of TNFα and TNFR1 (k_1). This suggests an experimental design that measures the concentration change of the corresponding proteins for proper estimation of the key parameter. This, in turn, leads

Table 1. Summary of the nominal parameter values and the corresponding range of variation

Corresponding Component and Range of Variation	Symbol	Value	Units
TNFα/TNFR1 association	k_1	0.185	μ M^{-1} s^{-1}
Range of variation		0.00925∼2.775	
TNFα/TNFR1 dissociation	k_2	0.00125	s^{-1}
TNFα/TNFR1/TRADD association	k_3	0.185	μ $M^{-1}s^{-1}$
Range of variation		0.00925∼2.775	
TNFα/TNFR1/TRADD dissociation	k_4	0.00125	s^{-1}
TNFα/TNFR1/TRADD/RIP1 association	k_5	0.185	$\mu M^{-1}s^{-1}$
TNFα/TNFR1/TRADD/RIP1 dissociation	k_6	0.00125	s^{-1}
TNFα/TNFR1/TRADD/TRAF2 association	k_7	0.185	μ $M^{-1}s^{-1}$
Range of variation		0.00925∼2.775	
TNFα/TNFR1/TRADD/TRAF2 dissociation	k_8	0.00125	s^{-1}
TNFα/TNFR1/TRADD/TRAF2/IKK association	k_9	0.185	μ $M^{-1}s^{-1}$
TNFα/TNFR1/TRADD/TRAF2/IKK dissociation	k_{10}	0.00125	s^{-1}
TNFα/TNFR1/TRADD/RIP1/TRAF2/IKK association	k_{11}	0.185	μ $M^{-1}s^{-1}$
TNFα/TNFR1/TRADD/RIP1/TRAF2/IKK dissociation	k_{12}	0.00125	s^{-1}
TNFα/TNFR1/TRADD/FADD association	k_{20}	0.185	μ $M^{-1}s^{-1}$
Range of variation		0.00925∼2.775	
TNFα/TNFR1/TRADD/FADD dissociation	k_{21}	0.00125	s^{-1}
TNFα/TNFR1/TRADD/FADD/caspase-8 association	k_{22}	0.185	μ $M^{-1}s^{-1}$
TNFα/TNFR1/TRADD/FADD/caspase-8 dissociation	k_{23}	0.00125	s^{-1}
IKK*/IκB/NF-κB association	k_{14}	0.185	μ $M^{-1}s^{-1}$
IKK*/IκB/NF-κB dissociation	k_{15}	0.00125	s^{-1}
Caspase-8/effector association	k_{25}	0.185	μ $M^{-1}s^{-1}$
Caspase-8/effector dissociation	k_{26}	0.00125	s^{-1}
Effector/c-IAP association	k_{29}	0.0925	μ $M^{-1}s^{-1}$
Effector/c-IAP dissociation	k_{30}	0.00125	s^{-1}
RIP1/caspase-8 association	k_{17}	0.0925	μ $M^{-1}s^{-1}$
Range of variation		0.0462∼0.1387	
RIP1/caspase-8 dissociation	k_{31}	0.000625	s^{-1}
Complex Degradation and Product Generation Coefficients			
TNFα/TNFR1/TRADD/RIP1/TRAF2/IKK → IKK* + TNFα + TNFR1 + TRADD + RIP1 + TRAF2	k_{13}	0.37	s^{-1}
IKK*/IκB/NF-κB → IκB-P + IKK + NF-κB	k_{16}	0.37	s^{-1}
NF-κB → c-IAP	k_{19}	0.092	s^{-1}
Range of variation		0.0462∼0.1387	
RIP1/caspase-8 → RIP1c + RIP1n	k_{18}	0.37	s^{-1}
TNFα/TNFR1/TRADD/FADD/caspase-8 → TNFα + TNFR1 + TRADD + FADD + caspase-8	k_{24}	0.37	s^{-1}
Caspase-8*/effector → caspase-8 + effector*	k_{27}	0.37	s^{-1}
Effector* → DNA fragmentation	k_{28}	0.37	s^{-1}

to a "core" mathematical model capturing the essential dynamics of the signal transduction pathway given that only a limited set of experimental data can be generated.

Figures 8 and 9 illustrate the simulation studies, investigating how sensitive the TNFα-mediated NF-κB signal transduction pathway is w.r.t. a change in parameter k_1. From these simulation results, we confirm strong nonlinear effects of the signal transduction pathway, particularly according to the variation of the key parameter k_1. Note that TNFα binds to TNFR1 relatively slowly for k_1 below 0.02, and the receptor is bound by TNFα relatively quickly for k_1 over 0.02 (the peak reaction occurs at k_1 equal to 0.0462). As TNFα binds to TNFR1, the concentration of the TNFα/TNFR1 complex increases accordingly. The simulation results also reveal that the steady-state value of the RIP1 complex varies along with the change of the parameter k_1. This variance of the steady-state value is significant since it determines, in turn, the next-step reaction

Table 2. Summary of the initial values for the simulation of the case study

Symbol	Corresponding Component	Initial Value [μM]	Symbol	Corresponding Component	Initial Value [μM]
m_1	TNFα	30	m_{17}	NF-κB	0
m_2	TNFR1	15	m_{18}	RIP1/caspase-8	0
m_3	TNFα/TNFR1	0	m_{19}	RIP1n	0
m_4	TRADD	15	m_{20}	RIP1c	0
m_5	TNFα/TNFR1/TRADD	0	m_{21}	FADD	10
m_6	TRAF2	10	m_{22}	Caspase-8	10
m_7	TNFα/TNFR1/TRADD/TRAF2	0	m_{23}	TNFα/TNFR1/TRADD/ FADD	0
m_8	IKK	10	m_{24}	TNFα/TNFR1/TRADD/ FADD/caspase-8	0
m_9	TNFα/TNFR1/TRADD/ TRAF2/IKK	0	m_{25}	Caspase-8*	0
m_{10}	RIP1	10	m_{26}	Caspase-8*/effector	0
m_{11}	TNFα/TNFR1/TRADD/ TRAF2/RIP1	0	m_{27}	Effector*	0
m_{12}	TNFα/TNFR1/TRADD/ TRAF2/RIP1/IKK	0	m_{28}	DNA_fragmentation	0
m_{13}	IKK*	0	m_{29}	Effector	10
m_{14}	IκB/NF-κB	10	m_{30}	Effector/c-IAP	0
m_{15}	IκB/NF-κB/IKK*	0	m_{31}	c-IAP	0
m_{16}	IκB-P	0			

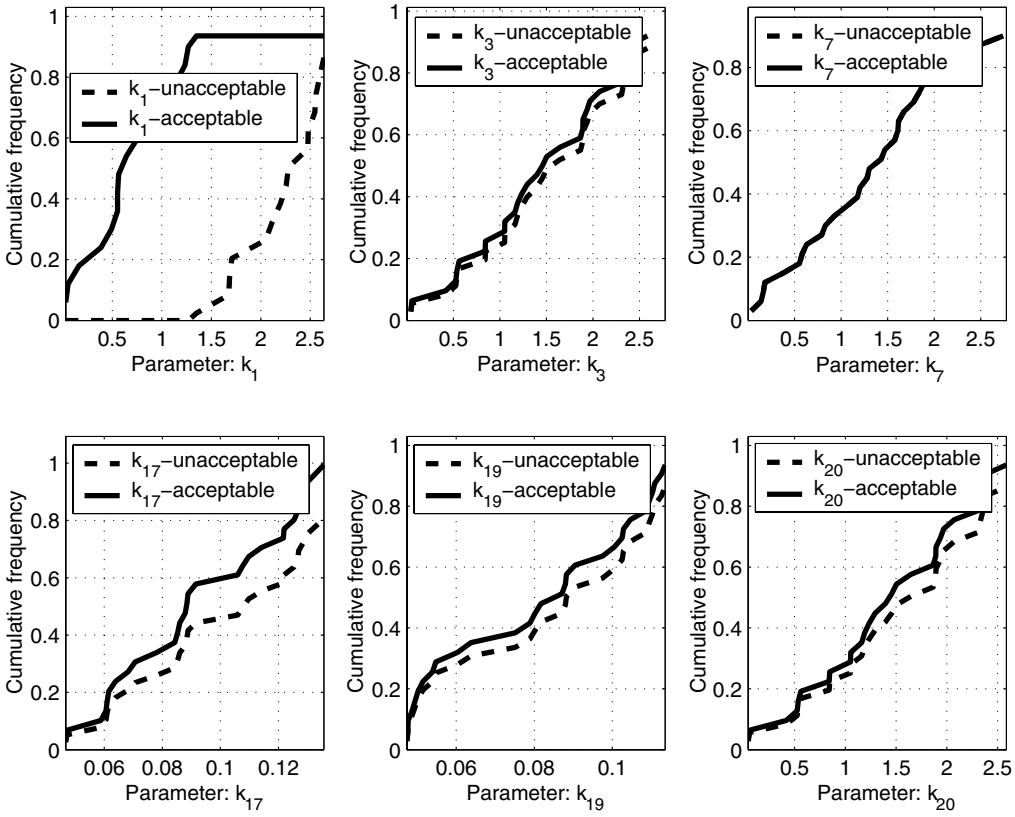

Figure 7. Parametric sensitivity analysis results of the TNFα-mediated NF-κB signal transduction pathway. A solid line denotes an acceptable case, and a dashed line indicates an unacceptable case.

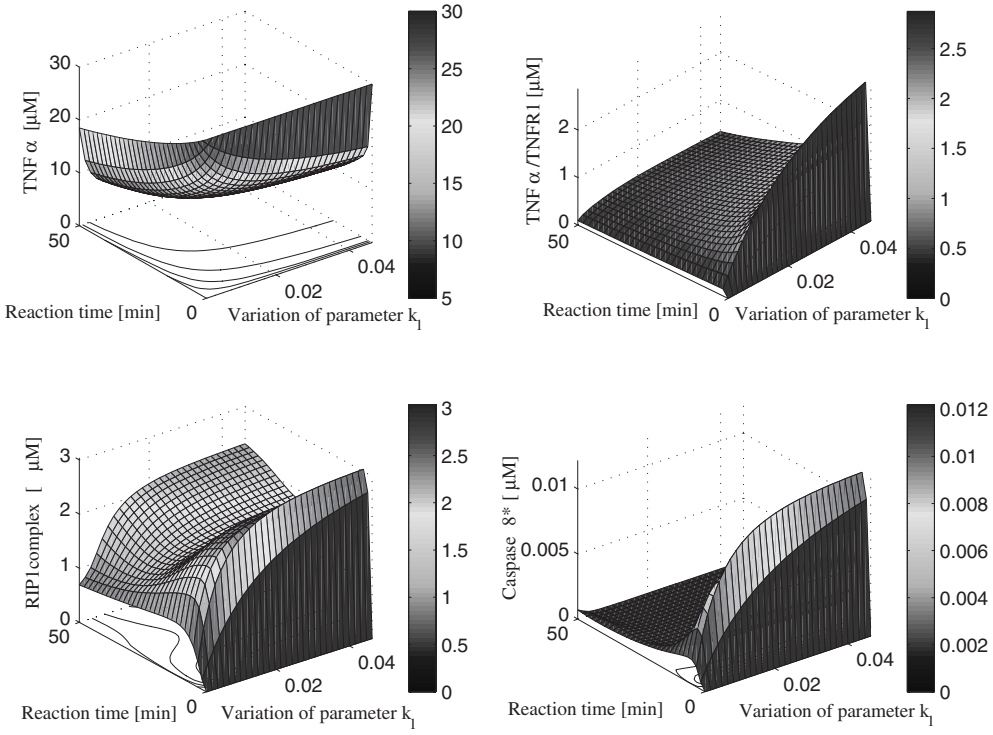

Figure 8. Simulation studies for the signal transduction sensitivity analysis with respect to the change of k_1

rate coefficient in the signal transduction pathway. Besides, it is remarkable that in Figure 8, the activity of caspase-8* is relatively slow for the lower value of k_1; however, the dynamics of caspase-8* becomes very active and fast for higher values of k_1. The dynamics of the IKK complex in Figure 9 is similar to that of caspase-8*, except for some difference of absolute magnitudes. The rate of concentration change for DNA fragmentation and the activity of IKK* and NF-κB become slow for the lower value of k_1 as well. Furthermore, we know that the steady-state values of these proteins are also influenced by the change of the parameter k_1. In particular, it turns out that the dynamics of the signaling proteins is suppressed at a lower value of the parameter and is supposed to be influenced by the initial concentration of the TNFα rather than the higher parameter value. One of the biological implications of this simulation study is that the dynamics of the TNFα-mediated NF-κB signal transduction pathway is heavily influenced by a cell environment since the parameter values depend heavily on the cell environment.

5. Conclusions and Further Studies

This study has investigated the sensitivity analysis of signal transduction pathway models with respect to parameter changes. The results of this analysis have led to an experimental design indicating which measurements should be taken for a "core" mathematical model that captures the essential characteristics of the system. A case study of the TNFα-mediated NF-κB signal transduction pathway has been provided to illustrate the proposed parameter sensitivity analysis and the corresponding experimental design. In particular, the case study has shown that there exist relatively sensitive parameters, and the association coefficient of the TNFα and TNFR1 (k_1) has been identified as the most sensitive (hence important) parameter amid them. This suggested subsequently an experimental design measuring the relevant changes in protein concentrations for the estimation of the identified key parameters. In addition, the simulation study has shown that the nonlinear dynamics of the pathway is heavily dependent on the (key) parameter values, and the steady-state value of each signaling protein is particularly suppressed at lower parameter values during signal transduction.

A sensitivity analysis w.r.t. sampling time points is also crucial for an experimental design and is to be explored in further studies. Once data have been collected, a suitable approach for parameter estimation is important. The approach presented in Horbelt, Voss, and Timmer [12] and applied to the JAK-STAT signaling pathway [13] is one

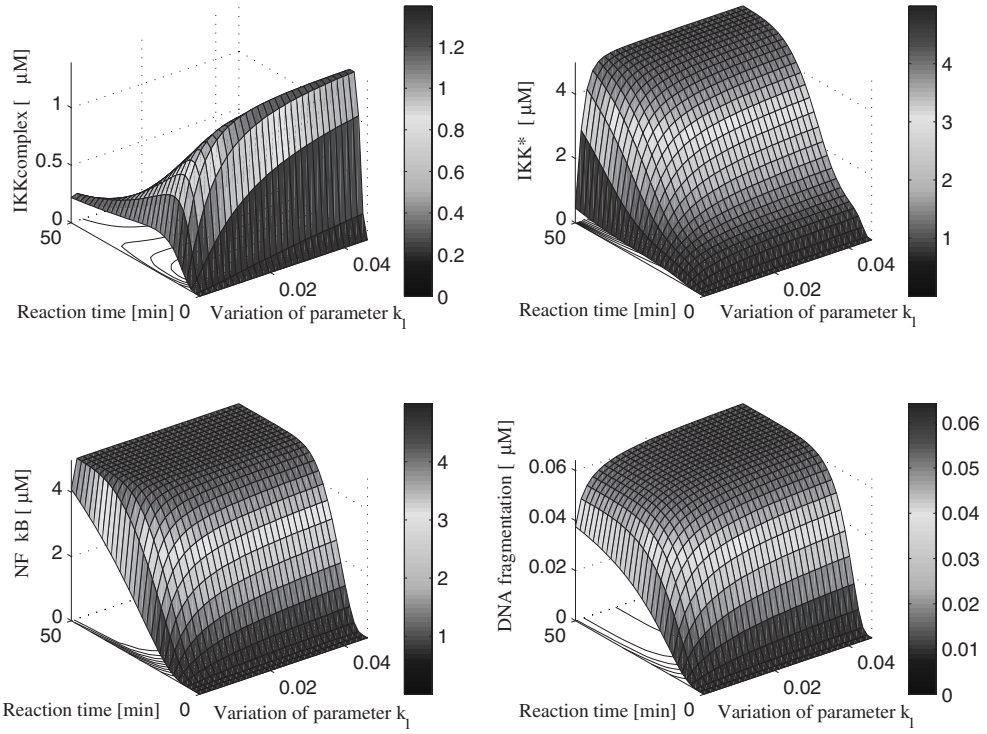

Figure 9. Simulation studies for the signal transduction sensitivity analysis with respect to the change of k_1 (continued)

possible direction. It would be also interesting to investigate the differences in the identifiability [19, 20] of the original model and the model reconstructed by the key parameters/variables identified from the proposed approach. Here we also note that while the identifiability theory [12, 13, 17, 19, 20] provides us with information regarding how many variables (proteins) have to be measured to estimate all or a specific subset of parameters, the proposed sensitivity analysis evaluates how much a parameter can affect a system output from the signal transduction point of view, such that parameters can be ordered according to the sensitivity analysis results, and thereby gives us the information about the core part of the system. Moreover, once we build a model from the data obtained by the designed experiments, the analysis regarding dependency of models with regard to initial conditions would be useful if we perform Monte Carlo simulation on the final mathematical model.

In this article, we suggested setting the parameter range between one-fifth of a nominal value and five times the nominal value based on experiences. The most suitable parameter ranges for a global sensitivity analysis depend on the hidden dynamics of the pathway, and a systematic way for its determination would be an interesting further research topic. In addition, the correlation coefficient has been employed in this study for the quantitative comparison of the cumulative distribution functions because it is the simplest way for the quantitative comparison of correlation. A proper statistical study is out of the scope of this study. Other tests, such as the Kolomogorov Smirnov test, could further improve the proposed approach.

Regarding computational complexity, note that we suggested first selecting a set of parameters to be tested by the proposed approach. However, in general, if we have no a priori knowledge of the relative importance of parameters/variables at all, then we should simulate for all possible combinations of the randomly generated parameter values in pairs, which, however, might lead to a large amount of simulation time, depending on the complexity of the signal transduction pathways, and thereby might limit the application of the proposed method in some cases. However, we could approach this problem by considering several submodels of the whole pathway in a vertical or horizontal way based on a mathematical justification. In addition, there are now remarkable advances in handling this kind of simulation complexity. Most of all, despite this potentially enormous simulation time, the proposed approach is certainly much more advantageous than a direct experiment for all the required measurements. They **[THESE TWO APPROACHES? PLS. CLARIFY]**would be incomparable.

6. Appendix

The following set of nonlinear ODEs represents the developed mathematical model of the TNFα-mediated NF-κB signal transduction pathway shown in Figure 6.

$$\frac{dm_1}{dt} = -k_1 \cdot m_1 \cdot m_2 + k_2 \cdot m_3 \quad (4)$$

$$\frac{dm_2}{dt} = -k_1 \cdot m_1 \cdot m_2 + k_2 \cdot m_3 + k_{24} \cdot m_{24} + k_{13} \cdot m_{12}$$

$$\frac{dm_3}{dt} = k_1 \cdot m_1 \cdot m_2 - k_2 \cdot m_3 - k_3 \cdot m_3 \cdot m_4 + k_4 \cdot m_5$$

$$\frac{dm_4}{dt} = -k_3 \cdot m_3 \cdot m_4 + k_4 \cdot m_5 + k_{13} \cdot m_{12} + k_{24} \cdot m_{24}$$

$$\frac{dm_5}{dt} = k_3 \cdot m_3 \cdot m_4 - k_4 \cdot m_5 - k_5 \cdot m_5 \cdot m_{10} + k_6 \cdot m_{11} - k_7 \cdot m_5 \cdot m_6 + k_8 \cdot m_7 - k_{20} \cdot m_5 \cdot m_{20} + k_{21} \cdot m_{23}$$

$$\frac{dm_6}{dt} = -k_7 \cdot m_5 \cdot m_6 + k_8 \cdot m_7 + k_{13} \cdot m_{12} + k_{24} \cdot m_{24}$$

$$\frac{dm_7}{dt} = k_7 \cdot m_5 \cdot m_6 - k_8 \cdot m_7 - k_9 \cdot m_7 \cdot m_8 + k_{10} \cdot m_9$$

$$\frac{dm_8}{dt} = -k_9 \cdot m_7 \cdot m_8 + k_{10} \cdot m_9 + k_{16} \cdot m_{15}$$

$$\frac{dm_9}{dt} = k_9 \cdot m_7 \cdot m_8 - k_{10} \cdot m_9 - k_{11} \cdot m_{11} \cdot m_9 + k_{12} \cdot m_{12}$$

$$\frac{dm_{10}}{dt} = -k_5 \cdot m_5 \cdot m_{10} + k_6 \cdot m_{11} - k_{17} \cdot m_{10} \cdot m_{25} + k_{31} \cdot m_{18} + k_{13} \cdot m_{12} + k_{24} \cdot m_{24}$$

$$\frac{dm_{11}}{dt} = k_5 \cdot m_5 \cdot m_{10} - k_6 \cdot m_{11} - k_{11} \cdot m_9 \cdot m_{11} + k_{12} \cdot m_{12}$$

$$\frac{dm_{12}}{dt} = k_{11} \cdot m_9 \cdot m_{11} - k_{12} \cdot m_{12} - k_{13} \cdot m_{12}$$

$$\frac{dm_{13}}{dt} = k_{13} \cdot m_{12} - k_{14} \cdot m_{13} \cdot m_{14} + k_{15} \cdot m_{15}$$

$$\frac{dm_{14}}{dt} = -k_{14} \cdot m_{13} \cdot m_{14} + k_{15} \cdot m_{15}$$

$$\frac{dm_{15}}{dt} = k_{14} \cdot m_{13} \cdot m_{14} - k_{15} \cdot m_{15} - k_{16} \cdot m_{15}$$

$$\frac{dm_{16}}{dt} = k_{16} \cdot m_{15}$$

$$\frac{dm_{17}}{dt} = k_{16} \cdot m_{15} - k_{19} \cdot m_{17}$$

$$\frac{dm_{18}}{dt} = k_{17} \cdot m_{10} \cdot m_{25} - k_{31} \cdot m_{18} - k_{18} \cdot m_{18}$$

$$\frac{dm_{19}}{dt} = k_{18} \cdot m_{18}$$

$$\frac{dm_{20}}{dt} = k_{18} \cdot m_{18}$$

$$\frac{dm_{21}}{dt} = -k_{20} \cdot m_5 \cdot m_{21} + k_{21} \cdot m_{23} + k_{24} \cdot m_{24}$$

$$\frac{dm_{22}}{dt} = -k_{22} \cdot m_{22} \cdot m_{23} + k_{23} \cdot m_{24} + k_{18} \cdot m_{18} + k_{27} \cdot m_{26}$$

$$\frac{dm_{23}}{dt} = -k_{22} \cdot m_{22} \cdot m_{23} + k_{23} \cdot m_{24} + k_{20} \cdot m_5 \cdot m_{21} - k_{21} \cdot m_{23}$$

$$\frac{dm_{24}}{dt} = k_{22} \cdot m_{22} \cdot m_{23} - k_{23} \cdot m_{24} - k_{24} \cdot m_{24}$$

$$\frac{dm_{25}}{dt} = k_{24} \cdot m_{24} - k_{17} \cdot m_{10} \cdot m_{25} + k_{31} \cdot m_{18} - k_{25} \cdot m_{25} \cdot m_{29} + k_{26} \cdot m_{26}$$

$$\frac{dm_{26}}{dt} = k_{25} \cdot m_{25} \cdot m_{29} - k_{26} \cdot m_{26} - k_{27} \cdot m_{26}$$

$$\frac{dm_{27}}{dt} = k_{27} \cdot m_{26} - k_{28} \cdot m_{27}$$

$$\frac{dm_{28}}{dt} = k_{28} \cdot m_{27}$$

$$\frac{dm_{29}}{dt} = -k_{25} \cdot m_{25} \cdot m_{29} + k_{26} \cdot m_{26} - k_{29} \cdot m_{29} \cdot m_{31} + k_{30} \cdot m_{30}$$

$$\frac{dm_{30}}{dt} = k_{29} \cdot m_{29} \cdot m_{31} - k_{30} \cdot m_{30}$$

$$\frac{dm_{31}}{dt} = -k_{29} \cdot m_{29} \cdot m_{30} + k_{30} \cdot m_{30} + k_{19} \cdot m_{17}$$

7. Acknowledgment

This work was supported by a grant from the Korean Ministry of Science and Technology (Korean Systems Biology Research Grant, M10309000006-03B5000-00211).

8. References

[1] Cho, K.-H., S.-Y. Shin, H.-Y. Lee, and O. Wolkenhauer. Forthcoming. Investigations into the analysis and modeling of the TNFα mediated NF-κB signaling pathway. *Genome Research*.

[2] Cho, K.-H., S.-Y. Shin, H.-Y. Kim, O. Wolkenhauer, B. McFerran, and W. Kolch. 2003. Mathematical modeling of the influence of RKIP on the ERK signaling pathway. In *Computational methods in systems biology*, edited by C. Priami, 127-41. Berlin: Springer Verlag.

[3] Wolkenhauer, O., W. Kolch, and K.-H. Cho. 2003. Mathematical systems biology: Genomic cybernetics. In *Computations in cells and tissues: Perspectives and tools of thought*, edited by R. Paton. London: Springer Verlag.
[4] Voit, E. O. 2000. *Computational analysis of biochemical systems*. Cambridge, UK: Cambridge University Press.
[5] Bhalla, U. S., and R. Iyengar. 1999. Emergent properties of networks of biological signaling pathways. *Science* 283:381-7.
[6] Weng, G., U. S. Bhalla, and R. Iyengar. 1999. Complexity in biological signaling systems. *Science* 284:93-6.
[7] Phair, R. D., and T. Misteli. 2001. Kinetic modeling approaches to in vivo imaging. *Nature Reviews: Molecular Cell Biology* 2:898-907.
[8] Phair, R. D. 1997. Development of kinetic models in the nonlinear world of molecular cell biology. *Metabolism* 46:1489-95.
[9] Neves, S. R., and R. Iyengar. 2002. Modeling of signaling networks. *BioEssays* 24:1110-7.
[10] Schoeberl, B., C. Eichler-Jonsson, E. D. Gilles, and G. Muller. 2002. Computational modeling of the dynamics of MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology* 20:370-5.
[11] Schoebel **[SCHOEBERL?]**, B., E. D. Gilles, and P. Scheurich. 2001. A mathematical vision of TNF receptor interaction. In *Proceedings of 2nd International Conference on Systems Biology (ICSB2001)*, **CITY?**, pp. 158-67.
[12] Horbelt, W., H. U. Voss, and J. Timmer. 2002. Parameter estimation in nonlinear delayed feedback systems from noisy data. *Physical Letters A* 299:513-21.
[13] Swameye, I., T. G. Müller, J. Timmer, O. Sandra, and U. Klingmüller. 2003. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proceedings of the National Academy of Science* 100:1028-33.
[14] Schittkowski, K. 1994. Parameter estimation in systems of nonlinear equations. *Numerical Mathematics* 68:129-42.
[15] Hegger, R., H. Kantz, F. Schmüser, M. Diestelhorst, R.-P. Kapsch, and H. Beige. 1998. Dynamical properties of a ferroelectric capacitor observed through nonlinear time series analysis. *Chaos* 8:727-36.
[16] Müller, T. G., N. Noykova, M. Gyllenberg, and J. Timmer. 2002. Parameter identification in a dynamical model of anaerobic waste water treatment processes. *Mathematics in Bioscience* 177-178:147-60.
[17] Timmer, J., T. Müller, I. Swameye, O. Sandra, and U. Klingmüller. Forthcoming. Modeling the non-linear dynamics of cellular signal transduction. *International Journal of Bifurcation and Chaos* 14.
[18] Fomina, T., K. Holmström, and V. B. Melas. **YEAR?**. Nonlinear parameter estimation for inorganic chemical equilibrium analysis. Research Report 2000-3, Department of Mathematics and Physics, Mälardalen University [Online]. Available: www.ima.mdh.se/forskning/research-reports.ima.mdh.e.shtml
[19] Ljung, L. 1999. *System identification*. Englewood Cliffs, NJ: Prentice Hall.
[20] Audoly, S., L. D'Angio, M. P. Saccomani, and C. Cobelli. 1998. Global identifiability of linear compartmental models: A computer algebra algorithm. *IEEE Transactions on Biomedical Engineering* 45:36-47.
[21] Shi, Y. 2002. Clustering and signaling of cell receptors. *Physica A* 311:199-212.
[22] Keasling, J. D., H. Kuo, and G. Vahanian. 1995. A Monte Carlo simulation of the *Escherichia coli* cell cycle. *Journal of Theoretical Biology* 176:411-30.
[23] Choi, J., J. W. Harvey, and M. H. Conklin. 1999. Use of multi- parametric sensitivity analysis to determine relative importance of factors influencing natural attenuation of mining contaminants. U.S. Geological Survey Water-Resources Investigations Report 99-4018A [Online]. Available: http://toxics.usgs.gov/pubs/wri99-4018/Volume1/sectionC/ 1405_Choi/pdf/1405_Choi.pdf
[24] Paker, W. J. 1997. A multi-parametric sensitivity analysis of a model describing the fate of volatile organic compounds in tricking filters. *Journal of Air & Water Management Association* 47:871-80.
[25] Baud, V., and M. Karin. 2001. Signal transduction by tumor necrosis factor and its relatives. *Trends in Cell Biology* 11:372-7.
[26] Karin, M., and A. Lin. 2002. NF-κB at the crossroads of life and death. *Nature Immunology* 3:221-7.
[27] Ting, A. T., and D. Endy. 2002. Decoding NF-κB signaling. *Science* 298:1189-90.
[28] Kyriakis, J. M. 2001. Life-or-death decisions. *Nature* 414:265-6.
[29] Stryer, L. 1988. *Biochemistry*. New York: W. H. Freeman.
[30] Kuo, B. C. 1995. *Automatic control system*. Englewood Cliffs, NJ: Prentice Hall.
[31] Khalil, K. K. 2002. *Nonlinear systems*. Englewood Cliffs, NJ: Prentice Hall.
[32] Law, A. M., and W. D. Kelton. 2000. *Simulation, modeling and analysis*. New York: McGraw-Hill.
[33] Woolfson, M. M., and G. J. Pert. 1999. *An introduction to computer simulation*. New York: Oxford University Press.
[34] Newman, M. E. J., and G. T. Barkema. 1999. *Monte Carlo methods in statistical physics*. New York: Oxford University Press.
[35] Hoffmann, A., A. Levchenko, M. L. Scott, and D. Baltimore. 2002. The IκB-NF-κB signaling module: Temporal control and selective gene activation. *Science* 298:1241-5.
[36] Fussenegger, M., J. E. Bailey, and J. Varner. 2000. A mathematical model of caspase function in apoptosis. *Nature Biotechnology* 18:768-74.

Kwang-Hyun Cho *is* **(POSITION?)** *in the School of Electrical Engineering, University of Ulsan, Ulsan, Korea.*

Sung-Young Shin *is* **(POSITION?)** *in the School of Electrical Engineering, University of Ulsan, Ulsan, Korea.*

Walter Kolch *is* **(POSITION?)** *at the Beatson Institute for Cancer Research and* **(POSITION?)** *at the Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, United Kingdom.*

Olaf Wolkenhauer *is* **(POSITION?)** *in the Department of Biomolecular Science and Department of Electrical Engineering & Electronics, UMIST, Manchester, United Kingdom.*