Systems Biology Markup Language: Case Study of T-Cell Signal Transduction Network

Ana Tušek, Želimir Kurtanjek
University of Zagreb, Faculty of Food Technology and Biotechnology
Pierottijeva 6, 10000 Zagreb, CROATIA
atusek@net.hr zkurt@pbf.hr

Abstract. *Application of system theory in biology has led to development of computer languages for mathematical modeling which are in open source and apply common standards for model trans-platform exchangeability and integration with extensive model banks, databases, numerical and statistical methods and graphical tools. Standardization has led to Systems Biology Markup Language (SBML) based on XML format, while open source policy enables integration of computer resources and provides synergistic integration of scientific results and methodologies from various scientific fields.*

Different software projects have been initiated resulting in software such as: SBToolbox, JDesigner, SBWorkbench (SBW), and CellDesigner. The main use of the systems biology software is to investigate biochemical networks, gene regulatory networks, signal transduction, cell communication, dynamical complex behavior, and most importantly to lead to understanding of life regulation.

In this work is presented use of Systems Biology Markup Language and modeling tools applied for analysis of signal transduction in T-cell as a part of a human immune system.

Keywords. SBML, T-cell receptor, signal transduction.

1. Introduction

By computational biology is possible to build models and simulate biochemical networks. There are many computational tools for model building, but many of them are incompatible. Because, systems biology community needed information standards, in year 2000 scientist from Japan Science and Technology Corporation formed a team of researches to work on new software. They started to developing SBML language based on XML format.

SBML is a computer-readable format for representing models of biochemical networks. SBML can represent metabolic networks, cell-signaling pathways, regulatory networks and many other.

The main purposes of SBML are enabling models to be shared and published in form that other researches can use even in a different software environment, ensuring the survival of the model, providing synergy for databases and other software tools.

Models in SBML are composed of: compartments, species, reactions, parameters, and other similar units, definitions, rules, events, constraints.

```
<?xml version="1.0" encoding="UTF-8" ?>
- <sbml xmlns="http://www.sbml.org/sbml/level2"
  xmlns:celldesigner="http://www.sbml.org/2001/ns/celldesigner"
  level="2" version="1">
- <model id="untitled">
  - <annotation>

      <celldesigner:modelVersion>2.5</celldesigner:modelVersion>
    <celldesigner:modelDisplay sizeX="2500" sizeY="2000" />
   - <celldesigner:listOfIncludedSpecies>
    - <celldesigner:species id="s6" name="TCRp">
     - <celldesigner:notes>
      - <html
          xmlns="http://www.w3.org/1999/xhtml">
          <!-- Notes by CellDesigner  -->
          <body />
        </html>
       </celldesigner:notes>
     - <celldesigner:annotation>
```

Figure 1. Part of TCR signal transduction model in SBML format.

SBML continuously undergoes changes because it is still in development. Major releases of SBML are called *Levels*, and represent great changes in structure of language. Current level is Level 2. Minor releases are called *Versions* and are made to correct, adjust and improve characteristics of language [1].

In this work SBML complaint software CellDesigner was used for constructing the model of T-cell receptor (TCR) signal transduction. Model was simulated using SBML ODE Solver from CellDesigner. To compare simula-

tion results, the simulation was also conducted in SBToolbox for MatLab. SBML format enables model to be directly exported from CellDesigner to SBToolbox and vice versa. The SBToolbox allows access to all data and data structures, resulting in full control over the tasks to be perform [2]. For analysis of dynamic properties of the signal transduction representation of SBML model based on ordinary differential equations (ODEs) is used.

2. CellDesigner

CellDesigner is SBML compliant software for drawing gene regulatory and biochemical networks. Biochemical networks are drawn based on the process diagram and using graphical notation proposed by Kitano [3].

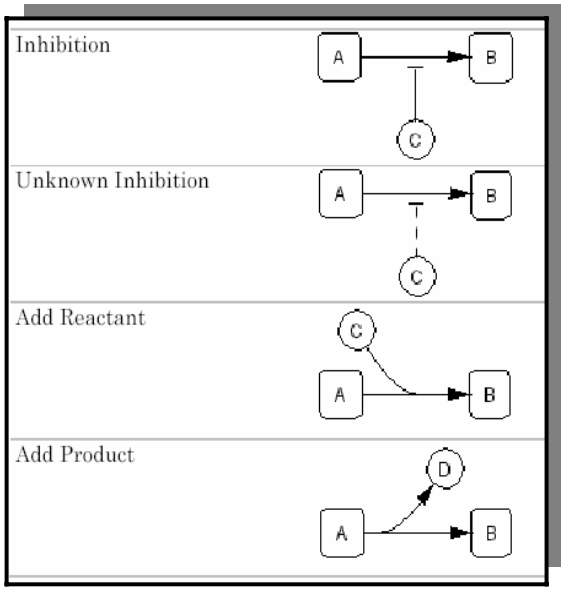


Figure 3. Examples of reaction symbols [4].

Major features of CellDesigner are its graphical notation, layout function editing and direct simulation with the Control Panel. Other features are intuitive user interface, extensive note description; output lists to CV format, and output of images into JPG, PNG and SVG formats.

Simulations in CellDesigner can be conducted in two ways. The first is by using SBML ODE Solver from CellDesigner. The second is by using SBW menu to call SBML complaint simulators. ControlPanel in CellDesigner allows specifying the details of parameters, changing amount of species, conducting parameter search and interactive simulation with intuitive manner.

3. TCR signal transduction

In this work is applied model of T-cell receptor signal transduction developed by Zheng [5]. Model is composed of 24 differential equations and 69 parameters, 16 of them represent the total level of protein expression in a cell and other 53 are related to the reaction kinetics.

T-cell activation starts with recognition of peptide-MHC complexes through extracellular domains of TCR, α i β chains. Intracellular signaling is initiated through cytoplasmic domains of transmembrane proteins associated with T-cell receptor. Phosphorylation of cytoplasmic domains of TCR complex is done by protein tyrosine kinase (SFK). This phosphorylation leads to TCR activation. In resting T-cells, SFKs are inactive, they become active through post-translation modification, phosphorylation. SFKs are also regulated by dephosphorylation.

In resting T-cells, constitutively active Csk (the cytoplasmic kinase) is bind to membrane with transmembrane protein Cbp, in this way Csk suppresses SFK activation. Cbp phosphorylation in resting cell is critical for the inhibitory function of Csk on SFK. T-cell stimulation with antigen induces rapid dephosphorylation of Cbp, release of Csk and SFK activation. Activated SFK phosphorylates TCR, Zap and many other proteins. Activation of Zap through phosphorylation plays a key role in early T-cell signaling.

Opposite to kinases, phosphatases catalyze the removal of phosphate groups. CD45 and SHP-1 are protein phosphatases (PTP) essential for regulating early T-cell signaling. CD45 is transmembrane PTP that dephosphorylates SFK and makes it active. SHP-1 is a cytosolic PTP. Activated SHP-1 dephosphorylates Zap70 and many other proteins, so the role of SHP-1 is to suppress TCR signaling.

Zap 70 and SFK activate downstream signaling pathways in a T-cell. Activated Zap70 kinase, phosphorylates LAT (transmembrane adaptor protein). Grb2 associates with SOS. Grb2-SOS complex is then binding to phosphorylated LAT. SOS catalyses the exchange of RasGDP to RasGTP. The activated RasGTP induces the cascade of reaction that lead to activation of Erk.

CellDesigner model of T-cell receptor signal transduction network is given in Fig2.

Mathematical structure of the model is a network of protein interactions expressed as bimolecular reactions. The bimolecular model of interactions is applied as an approximation of Micha-

elis-Menten kinetics supported by the assumption of low molecular number per cell. Set of number balances of various membranes bound and cytoplasm free proteins are given in the forms as:

$$\frac{dn_i}{dt} = k_{i+} \cdot n_k \cdot n_l - k_{i-} \cdot n_i \qquad (1)$$

where n_i, n_k, and n_l are numbers of interacting protein molecules, k_+ and k_- are specific reaction rates (coefficients) in forward and backward directions. The complexity arises from interwoven positive and negative feedbacks as effectors of the specific rates. Applied are general linear models

$$k_i = k_{i0} \cdot \left(\sum_{j=1}^{S} \alpha_j \cdot n_j - \sum_{l=1}^{P} \beta_l \cdot n_l \right) \qquad (2)$$

In the model (2) the parameters α_j are associated with the effect of positive feedback on the corresponding reaction rate, while β_l are related to the negative. The complexity of the reaction networks and numerous effectors interactions are presented in the graphical form by CellDesigner reaction map (Fig. 3). The model includes 69 parameters which are estimated from very limited number of experimental data. They are interpreted as random variables with assumed corresponding uniform probability density functions in a finite range. One of the important issues is to determine influence of the parameter variations on the overall property of the signal transduction. In this work is applied one-to-one local sensitivity analysis of the parameters on the output signal.

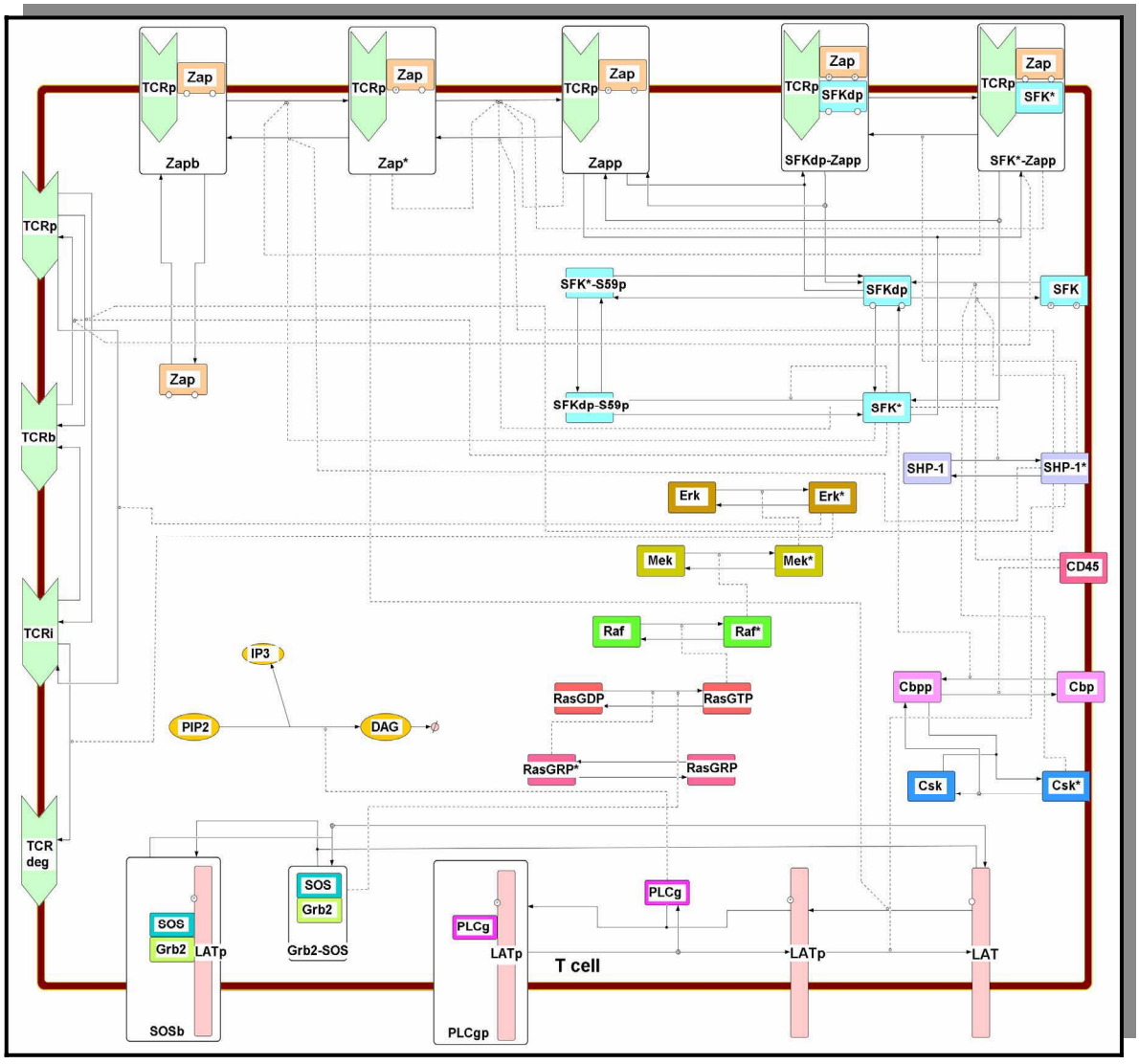


Figure 3. CELL DESIGNER model of T-cell receptor signal transduction network.

As the output signal is considered the steady state number of Erk* protein molecules present in cytoplasm responsible for promotion of the nucleus gene transcription leading to T-cell differentiation, proliferation and kinase excretion. The relative parametric sensitivity coefficients of the steady state number of activated Erk* proteins are calculated by the relations:

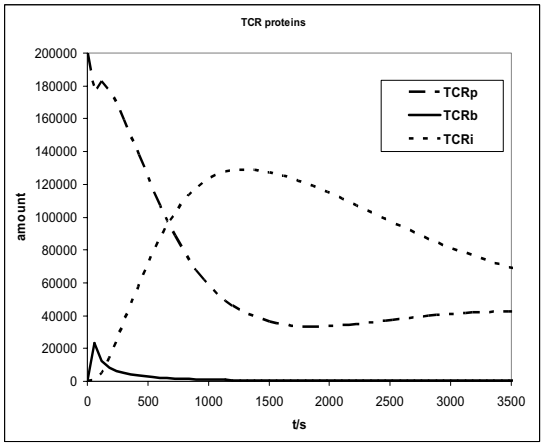
$$S_{Erk*/j}\% = \frac{k_j}{n_{Erk*}} \cdot \frac{\partial n_{Erk*}}{\partial k_j} \cdot 100 \qquad (3)$$

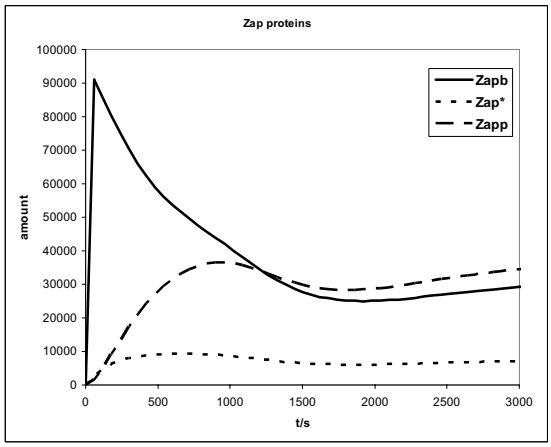
The coefficients $S_{\mathrm{Erk*/j}}$ are evaluated by the model simulation until a steady state is reached, and use of the finite difference approximation with 10% parameter increase. Positive relative sensitivity coefficients of 100 % corresponds to linearly proportional effect of a parameter on the output signal, while -100 % corresponds to the inverse proportionality of a parameter on the output signal Erk*. Small values of the sensitivity coefficient indicate insensitivity of the output signal on the parameter, while the values above ±100 % correspond to the exponential dependencies.

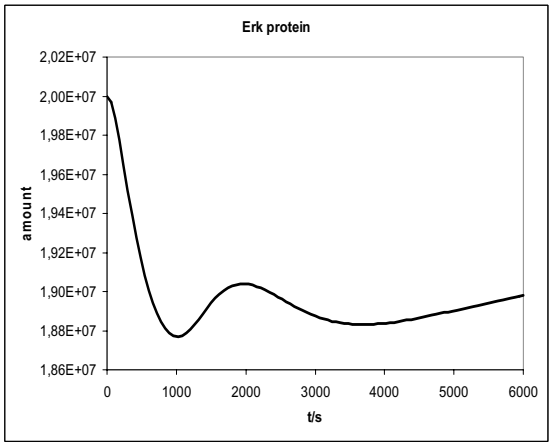
4. Results and Discussion

Simulation results, depicted in Fig. 4., present transient propagation of the TCR, Zap, Erk and Erk* proteins.

Sensitivities of the steady state Erk* signal on the model parameters are graphically presented by arrows in Fig. 5.







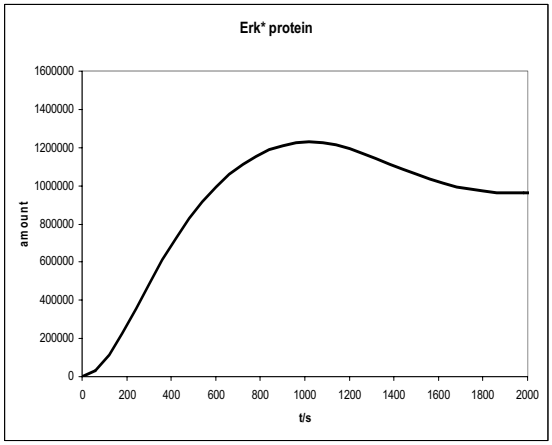


Figure 4. Transient responses of TCR, Zap, Erk and Erk* proteins.

Direction of the arrows from left to right indicate the positive feedback of the parameter, i.e. increases propagation of the signal toward the output, while the opposite direction stands for reversible reactions.

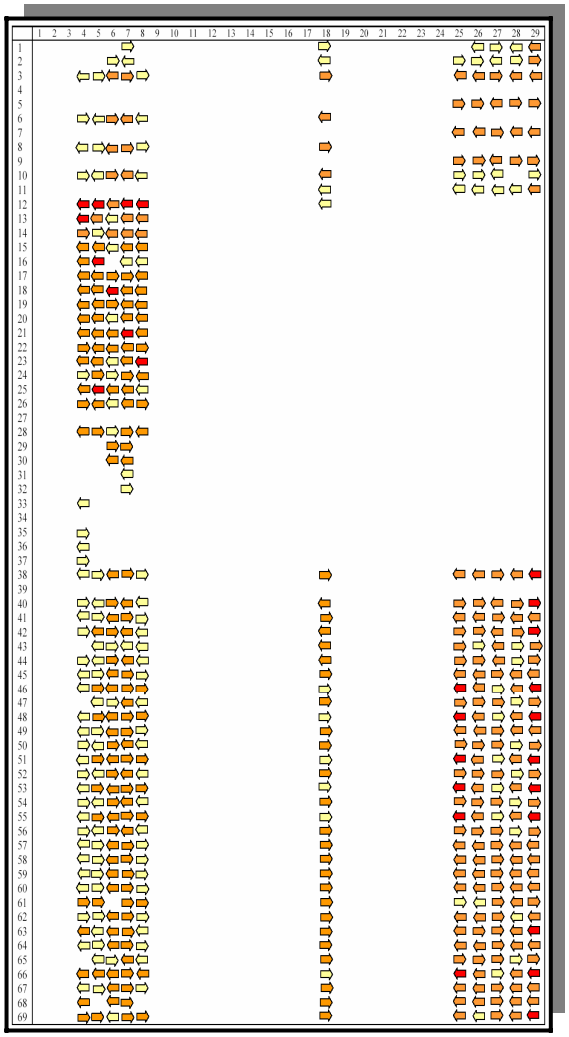


Figure 5. Relative sensitivities of Erk* steady state number on the model parameters. Columns in the table represent reactions and rows represent parameters of the model.

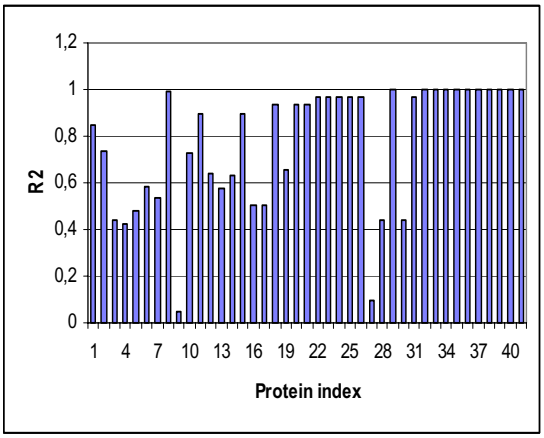


Figure 6. Correlations between signals along the network starting from the initial T-cell receptor activation and the output signal Erk*. Protein index 8 corresponds to double phosphorylated Zapp.

5. Conclusions

Application of CellDesigner for modeling of complex network of signal transductions in a single T-cell proved to be effective tool for integration of graphical interface and numerical methods for simulation of high dimensional, nonlinear and stiff system of differential equations (ODE).

Simulation results indicate that the initial engagement of T-cell receptor is the fastest process, leading to a sudden and a very sharp increase in number of phosphorylated TCRp.

SFK proteins exhibit the most influential effect on the initial stage of the signal transduction, and consequently their dominance is reflected on dynamics of the output signal.

Sensitivity analysis reveal numerous and simultaneous action of positive and negative regulation effects of the interaction parameters. The most influential parameters with positive feedback are from the interaction of LATp and PLCγ (second order of magnitude), while the most prominent negative effects resulting in the signal suppression are from the dephosphorylation of SFK.

The correlations between proteins and the output signal activated Erk*, show that most of the signals are highly correlated. The strongest and the longest distance correlation are determined with double phosphorylated Zapp.

Modeling and simulation of the signal transduction in activation of T-cell results in better understanding of the system properties and may lead to design of new experiments which can reveal critical protein interactions responsible for aberrations of immune response.

6. References

[1] Finney A, Hucka M, Le Novere N. Systems Biology Markup Language (SBML) Level2: Structures and Facilities for Model Definition 2006.

[2] Schmidt H, Jirstrand M. Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. Bioinformatics 2006; 22(4): 514-515.

[3] Kitano H. A Chow Y-C, Anger F graphical notation for biochemical networks. Biosilico 2003; 1(5): 169-176.
[4] Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for graphical representation of biochemical networks. Nature biotechnology 2005; 23(8): 961-966.
[5] Zheng Y. A systems biology study to delineate the T-cell receptor-activated ERK-MAPK signaling pathway. Ph.D. Thesis 2005; Weldon School of Biochemical Engineering, Purdue University, West Lafayette, Indiana.