Strogatz, S.H. (2007). Nonlinear dynamics and chaos: With applications to Physics, Biology, *Chemistry and Engineering*. Series: Studies in Nonlinearity. Levant Books, Kolkata, India.

Süel, M.G., Kulkarni, R.P., Dworkin, J., Garcia-Ojalvo, J. and Elowitz, M.B. (2007). Tunability and noise dependence in differentiation dynamics. *Science,* 315: 1716–1719.

Süel, M.G., Garcia-Ojalvo, J., Liberman, L.M. and Elowitz, M.B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. *Nature,* 440: 545–550.

Tyson, J.J. and Novak, B. (2013). Irreversible transitions, bistability and checkpoint controls in the eukaryotic cell cycle: A systems-level understanding. *In: Handbook of Systems Biology: Concepts and Insights. Eds.* Walhout, A.J.M., Vidal, M. and Dekker, J. (2013). Academic Press, New York. pp. 265–286.

vanKampen, N.G. (2007). Stochastic processes in physics and chemistry. Elsevier, Amsterdam, The Netherlands.

Vilar, M.G.J., Kueh, H.Y., Barkai, N. and Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *Proc. of National Acad. Sci.,* 99: 5988–5992.

Viswanathan, G.A. 2004. Transversal temperature patterns in packed-bed reactors. Ph.D. Dissertation, University of Houston, Houston, TX.

Wang, R.S., Saadatpour, A. and Albert, R. (2012). Boolean modelling in systems biology: An overview of methodology and applications. *Phys. Biol.,* 9: 055001.

Wilkinson, D.J. (2009). Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat. Rev. Gen.,* 10: 122–133.

Zhang, Y., Liu, R.Y., Heberton, G.A., Smolen, P., Baxter, D.A., Cleary, L.J. and Byrne, J.H. (2012). Computational design of enhanced learning protocols. *Nat. Neurosc.*, 15: 294–297.

Zhang, Y.L. and Dong, C. (2005). MAP kinases in immune responses. *Cellu. Mol. Immun.,* 2: 20–27.

10

Systems Biology - Modeling and Analysis of Biological Networks

Pramod R. Somvanshi[1], Anilkumar K. Patel[1], Sharad Bhartiya[1] and K.V. Venkatesh[1*]

ABSTRACT

Recent studies have shown that biological networks demonstrate control strategies which are also utilized in typical engineering systems. In systems biology, it is essential to understand the biological responses at a system level, by incorporating information from genetic, signaling and metabolic networks. Cells respond appropriately to diverse signals by coordinating complex network of protein, gene and metabolic interactions. Mathematical representation and quantification of the dynamic properties of biological networks is essential to decipher the inherent design principles evolved in biological systems. The networks can be modeled either by deterministic or stochastic formalisms. The essential parameters are estimated by data fitting and optimization techniques. Further, the models are subjected to steady state and dynamic analysis to study the dynamics, control and design principles of the system, which eventually helps the cell to achieve a desired phenotypic state. The modeling and analysis approaches used for describing the system level behavior in biological systems are described with examples in this chapter.

***Key words*:** Systems biology, Biological networks, Deterministic modeling, Stochastic modeling, Steady state and dynamic analysis

[1] Biosystems Engineering, Department of Chemical Engineering, Indian Institute of Technology, Bombay, Powai, Mumbai, India

**Corresponding author*: E-mail: venks@iitb.ac.in

INTRODUCTION

Molecular biology in the last four decades has deciphered components and their connections in biological systems through reductionist approach. Their studies have resulted in the description of the biological system through complex networks at the genetic, protein ant metabolic levels. It is increasingly becoming clear that in addition to the network description, quantification of the networks is important for linking the genotype to the phenotype.

Quantification of biological network requires construction of mathematical models to best describe the operation of the network (Schauer and Heinrich, 1983; Bailey, 1998; Koshland, 1998). Modeling and simulation of biological networks are becoming increasingly popular with models developed based on the quantitative experimental information of individual components such as time course data, dose response curves, protein concentrations and binding constants (Asthagiri and Lauffenburger, 2000; Bhalla, 2003; Eungdamrong and Iyengar, 2004; Sauro and Kholodenko, 2004; Aldridge *et al.*, 2006). Mathematical models provide a window to study the dynamics, control and design principles of a system, which eventually helps the system to achieve a desired state. Models developed can be effectively used to predict the network behavior, which can be subjected to experimental verification and further refinement. Models provide an ideal platform to test the effect of concentration and operating parameters, to study the effect of network perturbation (*in-silico* mutation), to analyze the roles and contributions of different interactions, to predict the emergent properties of a network and to identify missing information about the system. In general, a mathematical model developed should be able to generate valid hypothesis which can drive future experiments. Thus, models assist in system analysis, hypothesis generation and testing, experimental data validation and optimal product design.

MODELING IN SYSTEMS BIOLOGY

The comprehensive understanding based on quantitative experiments and computational modeling to gain insights into the physiology of a cellular process is termed as systems biology (Kitano, 2000; Ideker *et al.*, 2001; Kitano, 2002; Kitano, 2002; Ideker, 2004). One of the grand challenges in systems biology is to integrate the models of different nature and various spatiotemporal dynamics with bottom-up models and depict the exact outcome of the biological network. The systems approach to problem solving uses both computational and experimental data and is composed of several steps. These include modeling of the

system based on information about the interactions between its components, testing of the model for perturbations in structure and parameters of the system, and validation of the model by experiments. Mathematical models can then be used to identify recurring organizational principles. Interactions identified by *in vitro* experiments can lead to candidate models explaining the *in vivo* mechanisms. System-level properties depend on time- independent or steady-state stimulus response curves and temporal behavior of the system. Because cellular processes are noisy and uncertain, stochastic models may be helpful in describing them. Stochastic models can delineate the design of the system, which imparts properties for efficient performance. Here, we review some of the potential modeling frameworks that will help in understanding and quantifying biological networks. We draw attention to basic building blocks present in the pathways, various emergent properties of the network, modeling techniques and analysis methods employed in pathway modeling (Goel *et al.*, 2006; Machado *et al.*, 2011; Cardelli, 2005; Karlebach and Shamir, 2008; Wu, 2009; Noble, 2002).

Modeling Strategy

The objective of the mathematical modeling varies from predicting the dynamics of a system in response to a stimulus for understanding the emergent properties of the network. The choice of the quantification methods depends on the availability of the information that can be used to build a model. Based upon the quality and quantity of the data available and the nature of the system, modeling approaches can be broadly classified into deterministic and stochastic approaches. When the dynamics of the system is consistent with time and space with exact reproducibility and consistent with the experimental data then a deterministic approach is more suitable. Typically biological systems exhibit noise and are exposed to inherent fluctuations. Therefore, when there is inconsistency in the data and the system dynamics is noisy then stochastic modeling is more appropriate. Due to scarcity of quantitative data from experiments, mathematical models of biological systems have to depend on qualitative/semi-quantitative data. Input–output relationship of the network can be measured experimentally and the mathematical model can be constrained to predict system parameters to match experimental observation. Further models are subjected to sensitivity and stability analysis. The predictability of the models is refined so as to validate the observations from the experiments. The choice of parameter values depends on the experimental data obtained from wild type and mutants, while the rest has to be fixed based on parametric sensitivity analysis. Considering the complexity of biological system, it will be better to build models for smaller modules of

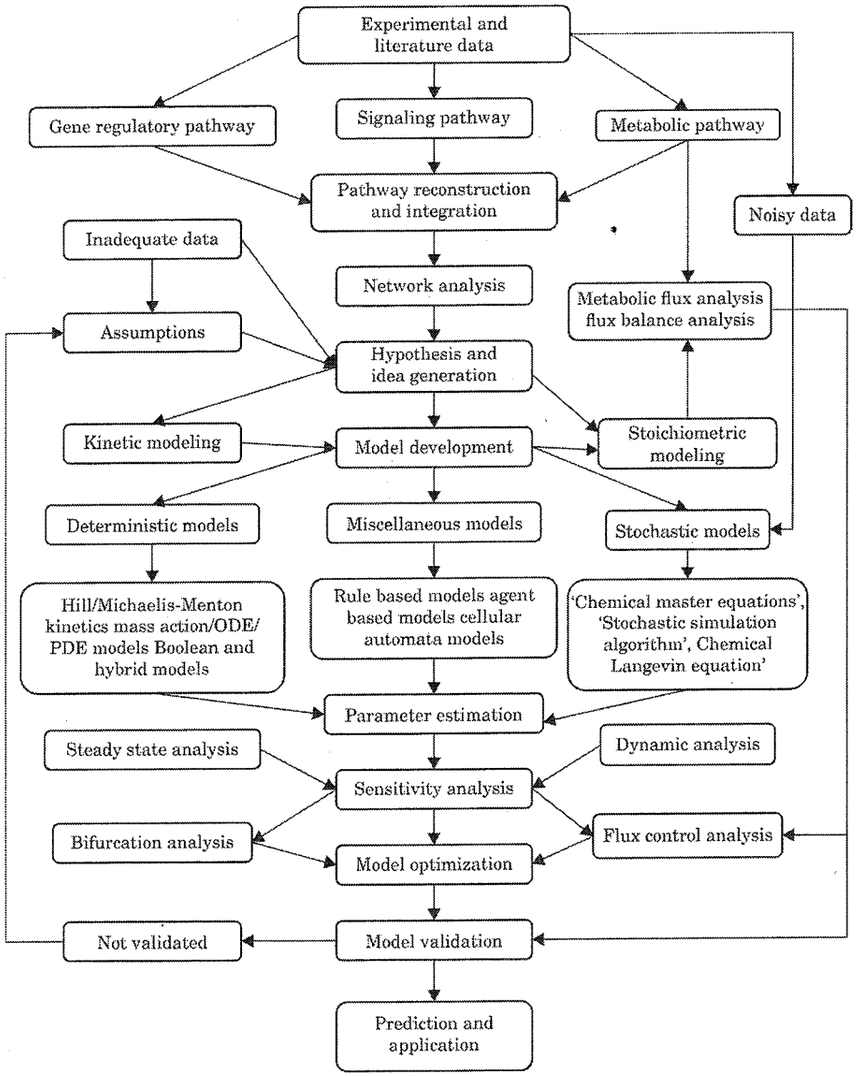


Fig. 1: The overall framework for modeling and analysis of Biological network. The process of modeling starts with precise definition of the problem and acquisition of relevant data from various sources *e.g.* omic databases. With the help of biological databases the pathway are reconstructed and integrated in the form of network. The hypothesis is generated based on the available knowledge. To test the hypothesis, the networks are further modeled using several mathematical techniques such as deterministic or stochastic modeling approaches. The parameters are estimated using optimization algorithms. Models are subjected to steady state and dynamic analysis to obtain the properties of a system. Further the models are validated and refined for their predictability.

the network before integrating the modules. This helps in studying the role and contribution of individual modules towards a network response. The overall schematic of the modeling and analysis approach is depicted in Fig. 1.

MODELING OF BIOLOGICAL NETWORKS

Graphical Representation

The easiest and the most straightforward way to model a biological network is to view it as a directed graph (as shown in Fig. 2). The communities of systems biologists and graphical modelers have formulated the standards for graphical representations of the biological systems, termed as Systems Biology Graphical Notation (SBGN). Various kinds of interactions and the process flow in the biochemical networks can be represented by the prescribed standards of the SBGN. The SBGN representations are classified as, Process Description, Entity Relationship and Activity Flow maps. In the Process Description maps, the events in the process are shown in the temporal order of the biochemical interactions. Each unit operation (reaction, transport, activation, inhibition, catalysis etc) and the species (protein, RNA, genes, metabolites etc.) in the network are represented by a unique symbol. In the Entity Relationship map, the relationship of each entity in the network with other entities is represented irrespective of the temporal information. The relationships can be the influences of the each component on other entities present in the network. In case of Activity Flow representation, the flows of the key information in terms of interactions between the entities of the network are represented, irrespective of their state transitions. The information such as activation, deactivation or up-regulation and down-regulation of the components

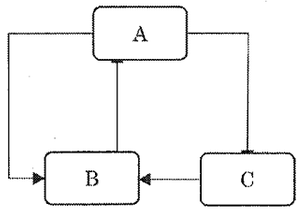


Fig. 2: A schematic of graphical representation of biological network. A, B, C are three components of the network with various interactions. The edges represent the interactions between these components, and the components form the nodes of the network. The arrow head represents the activation influence and the blunted arrow represents the inhibition influence among the components of the network. Therefore, the graph can be read as, A activates B, B inhibits A, A inhibits C, C activates B.

of the network is mainly focused in the Activity Flow representations. However, the symbols for these representations remain same for all the three kinds of graphical representations. The detailed list of symbols can be obtained at Cell Designer software (Funahashi *et al.*, 2008). An example of the Activity flow graph is given in Fig. 2.

Deterministic Modeling

In deterministic approach of modeling, the system is considered to be macroscopic, well mixed and the reaction is continuous. This is a simplification of the chemical reaction, which actually involve discrete, random collision between molecules. Moreover, the biochemical reaction occurs inside a cell, where the volume is small. Chemical reaction becomes deterministic in nature, if the reaction occurred at numerous times per generation, which average out the randomness. Deterministic models are broadly described based upon their discrete and continuous nature with respect to time and space, such as difference equations, ordinary differential equations, partial differential equations and maps (Breitling *et al.*, 2008, Cao *et al.*, 2010; Iba, 2008).

Boolean Modeling

Boolean networks are based on qualitative discrete framework and are most simple to address the dynamic systems. The model works on the principle of Boolean logic functions, where variables are quantified as binary output, *i.e.* ON or OFF and the value of the one variable is functionally related via a logical rule to the values of other variable (Fauré *et al.*, 2006). In case of loop systems the output is based on AND or OR input functions. Dynamics of the system is generated by updating the Boolean function, which causes system transition in accordance with logical rules. In a logical regulatory map, each node represents a protein or gene and arc (directed edges) with signs (positive/negative) representing the interaction. Each interaction is characterized by a source and target which is labeled by an integer (a threshold). This indicates the specific condition under which the interaction takes place. If the value of the source variable is equal to threshold, the interaction is said to be functional and their actions are described by logical parameters, which defines the activation of target (Kulasiri, 2008). The dynamics of the system is represented by the state transition graph, where nodes represent the states of the system and arc represents the transition between the states. This kind of approach reproduces the qualitative dynamic behavior of the network system. These models can equally predict the dynamic trend of the system even without using detailed kinetic parameters and differential equations.

The Boolean modeling approach is illustrated by an example as shown Fig. 3. Fig. 3 (I) shows a synthetic network that consists of three molecules A, B and C, where in, A has positive effect on C which activates B. Further, A inhibits B, which inhibits C. The Boolean Functions presenting the above mentioned interactions are listed in Fig. 3(II) and described further in the truth table (see Fig. 3(III)). Note that 1 and 0 indicate on and off states, respectively. The Boolean functions can be interpreted as: (1) $Bf_A(C)=C$, represents regulation of A by C; A = 1, if in the previous state, C=1; otherwise A =0. (2) $Bf_B(A,C)= C\ AND\ NOT\ A$, represents regulation of B by both A and C; B= 1 only if C =1 and A=0 in the previous state; B=0 otherwise. (3) $Bf_C(A,B)=A\ OR\ NOT\ B$, represents regulation of C by both A and B; C= 1 if A=1 or B =0 in the previous state; C=0 otherwise.

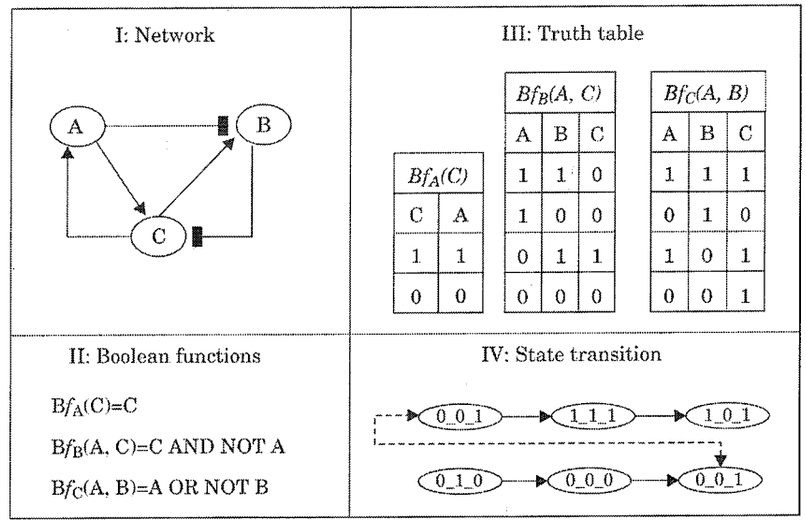


Fig. 3: An example of Boolean modeling. (I) A synthetic network consists of interactions between three proteins A, B and C. The edges with arrow indicate positive influence, while the edges with blunt cap indicate negative influence. (II) Regulatory Boolean functions that determine the effects of interactions (edges). (III) Truth Table, listing the states of each protein (1 or 0) based on the previous states of the all proteins involved in the interactions. (IV) Sequence of state transitions obtained based on two initial conditions, wherein Ist, IInd and IIIrd element within a state represents state of A, B and C, respectively.

CONTINUOUS MODELING

Law of Mass Action

The law of mass action states that the rate of the reaction is proportional to the probability of the collision of the reactants. This probability is

also proportional to the concentration of reactants to the power of their molecularity, a number of reactants participating in the specific reaction. The kinetics selected should be appropriate for the type of biochemical reaction, in most cases the mass- action kinetics are appropriate which assume the kinetics of elementary chemical reaction. The equation for a reversible reaction is given as:

$$A + B <=> C \tag{1}$$

$$V_a = k_b \times C - k_f \times A \times B \tag{2}$$

where A and B are the interacting species such as a ligand and a receptor or any two proteins, respectively. While C is the product of the interaction, V_a is the rate of reaction and K_f and K_b are rate constants of forward and backward reactions.

Michaelis Menten kinetics

In an enzyme catalyzed reaction, enzymes form a complex with the substrate to yield a product (See Eqn. 3). In such reaction if the substrate concentration is much higher than enzyme concentration where the enzyme is saturated with substrate, Michaelis Menten kinetics can be considered under the assumption of quasi steady state approximation (where the rate of change in enzyme- substrate complex is assumed zero). For irreversible enzyme kinetics the equation is given by

$$A + E <=> AE \longrightarrow E + P \tag{3}$$

$$V = V_{max}\left(\frac{S}{K_m + S}\right) \tag{4}$$

where, in Eq.(3) A , E, AE an P are substrate, enzyme, substrate-enzyme complex and protein, respectively. In Eq.(4) V_{max} represent the maximum reaction rate, 'S' represents the substrate concentration, 'K_m' represents Michaelis–Menten constant. In the case of a reversible reactions the equation takes the form

$$V = \frac{\frac{V_f S}{K_s}\left(1-\frac{\Gamma}{K}\right)}{1+\frac{S}{K_s}+\frac{P}{K_p}} = \frac{V_f \alpha\left(1-\frac{\Gamma}{K}\right)}{1+\alpha+\beta} \tag{5}$$

In Eq.(5) V_f and V_r are the limiting rates in forward and backward directions, Γ is the mass action ratio P/S under any condition, K is the

equilibrium constant, K_s and K_p are the Michaelis–Menten constants for the substrate and product, α and β substrate and product concentrations scaled by their Michaelis–Menten constants.

Hill Function

Another rate law which is mostly used in modeling the biological interaction is the Hills Equation. This equation is used to capture the cooperativity in bio-molecular reactions. Cooperativity exists when there are multiple binding sites on the enzymes, receptor and promoters regions. The bound subunit has the cooperative effect on the binding of the next subunit by increasing its affinity towards the binding region (Kulasiri, 2008). The general form of Hill equation is given by Eq. 6.

$$v = V_{max} Y = vmax * \frac{[L]^n}{[k_d]^n + [L]^n} \tag{6}$$

Whereas the inhibitory effects are modeled as

$$v = vmax * \frac{[k_d]^n}{[k_d]^n + [L]^n} \tag{7}$$

In case of reversible interactions the Hill equation is given by

$$V = \frac{V_f \alpha\left(1-\frac{\Gamma}{K}\right)(\alpha+\beta)^{n-1}}{1+(\alpha+\beta)^n} \tag{8}$$

where in Eq. (6 and 7) V_{max} represent the maximum reaction rate, L is the ligand/TF concentration, K_d is the equilibrium dissociation constant and n is Hill's coefficient. For Eq. (8) the parameters are as described for Eq. (5). The input-output relations as depicted in Eq. (6) (v=output and L=input) are also used to model other bio-molecular interactions, wherein, corresponding 'kd' represents the concentration of input at which the half of the maximal output response it obtained (also termed as half saturation constant).

ODE Modeling (Ordinary Differential Equations)

The rate expressions described above can be used to describe the rate of change of various components in a network. Ordinary differential equations are commonly used to describe the dynamics of a network system with various components as variables. This method involves representing a set of elementary reactions of allosteric interaction, covalent modification, feedback interactions and transport of proteins

as a system of chemical reactions. Chemical kinetic rate equations are used to describe elementary reactions. Each rate equation represents ordinary differential equation (ODE), (Polynikis *et al.*, 2009) which specify variation of component concentration with time. In a well mixed or homogeneous system, the solution of ODE with concentration as continuous variable forms the basis of deterministic modeling (Aldridge *et al.*, 2006) .

$$\frac{dC}{dt} = V_G - V_C \tag{9}$$

In. Eq. (9), the concentration of component C depends on the generation (V_G) and consumption (V_C) rate of the component, which in turn depends on the stoichiometry and kinetics of the reaction. Reaction can be zero order (synthesis), first order (degradation) or non linear, typically second order reactions or Michaelis Menten type kinetics. To illustrate the application of ODE for quantification of a biological network, here we present an example of a simple network (See Fig. 4). The figure shows a network, which is identical to the one presented in the description of Boolean Modeling. Here, we considered 6 reactions in a network, which follows typical mass action kinetics. (1) synthesis of A: it is dependent on concentration C with a reaction rate K_A*C, where K_A is rate constant, (2) degradation of A: it is first order with a rate constant of K_{dA}; (3) synthesis of B: It is dependent on concentration C with a reaction rate K_B*C , where K_B is rate constant; (4) degradation of B: It is triggered by A; second order with a rate constant K_{dB}; (5) synthesis of C: it is dependent on concentration of A with a reaction rate K_C*A, where KC is rate constant; (6) degradation of C: It is triggered by B; second order with a rate constant K_{dC}. The mass balance equation showing accumulation rates of A, B and C are shown in the Fig. 4 (II). The rate constants are listed in Fig. 4(III). The differential equations are integrated numerically and the time profiles of components are shown in Fig. 4(IV).

Compartmental Modeling

Dynamic analysis assumes well mixed system in a cell that yields temporal variations in the system. However, localization of components in different compartments in a cell influences the response characteristics. This necessitates the inclusion of variation in both space and time, due to different compartments. In compartmental models, a different compartment is treated as separate species and transport across the compartments are modeled as fluxes. Compartmental ODE modeling can capture the dynamics of spatially restricted reactions, however with

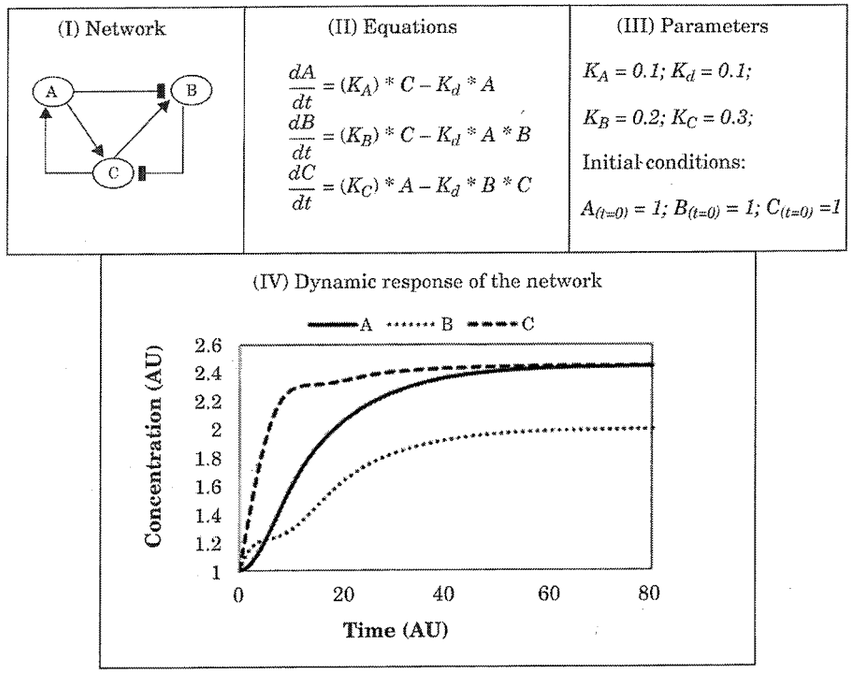
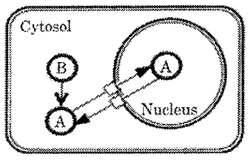


Fig. 4: (I) A synthetic network consists of interactions between three proteins A, B and C. The edges with arrow indicate positive influence, while the edges with blunt cap indicate negative influence. (II) ODE equations describing the effects of interactions (edges). (III) List of parameters and initial conditions (in arbitrary unit). (IV) Simulation of the model: The legends indicate lines assigned to the proteins.

assumptions that the transport rate across the compartment also affects the dynamics of the species in the compartment. Fig. 5 (I) shows two distinct compartments within a cell namely, cytosol and nucleus, both of them are assumed to be well mixed individually. In other words, concentration of any component is identical at each point within a compartment; however, it varies from compartment to compartment. Molecule A is able to translocate reversibly from cytosol to nucleus, whereas molecule B remains in the cytosol only. Therefore, the positive effect of B, as shown in the network, is limited to the cytosol compartment. The governing equations and parameters for the network are shown in Fig. 5 (II and III). Note that the balance of the concentration of A in the cytosol (A_{cyt}) and nucleus (A_{nuc}) involves for correction factor in the mass balance of A_{cyt} and A_{nuc}, since the volume of nucleus and cytosol are not equal. The resulting dynamics of the various species are shown in Fig. 5 (IV).

(I) Topology:



(II) Two compartment model:

$$\frac{d[A_{cyt}]}{dt} = K_A[B] - K_D[A_{cyt}] - K_{t1}[A_{cyt}] + K_{t2}\frac{[V_{nuc}]}{[V_{cyt}]}[A_{nuc}]$$

$$\frac{d[A_{nuc}]}{dt} = K_{t1}[A_{cyt}]\frac{[V_{cyt}]}{[V_{nuc}]} - K_{t2}[A_{nuc}]$$

$$\frac{dB}{dt} = k_0 - k_D[B]$$

(III) Parameters and initial conditions

Units are arbitrary:

$K_A = 1.00$; $K_D = 0.30$;
$k0 = 0.01$;
$V_{nuc} = 0.50$; $V_{cyt} = 1.00$;
$K_{t1} = 0.20$; $K_{t2} = 0.10$;
$A_{cyt,\,(t=0)} = 1.00$;
$A_{nuc,\,(t=0)} = 0.01$;
$B_{(t=0)} = 1.0$

(IV) Response of the compartmental model:

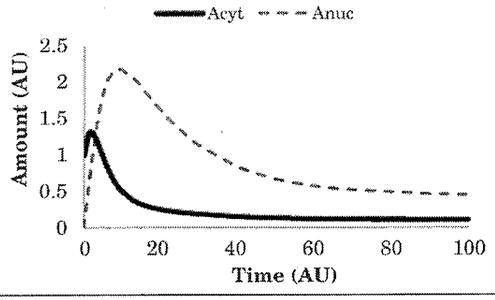


Fig. 5: (1) Network consists of two compartments, namely cytosol and nucleus. Only molecule A can translocate reversibly from cytosol to nucleus. B catalyzes the formation of A, and both A and B are subjected to natural degradation. However in nucleus, the degradation of A is neglected. (2) ODE equations representing the network (3) Parameters and Initial conditions used for the simulation (4) Response of the network.

PDE Modeling (Partial Differential Equations)

The assumptions made for compartmental ODE modeling does not hold good for non-homogenous systems, where there is explicit dependence of variable on spatially distributed processes such as a diffusion reaction. The dynamics of signaling pathway in relation to variation in space and time can be best described using a partial diffusion equations (PDE) (Kholodenko, 2006). The concentration of components in a compartment depends on the independent variables, such as diffusion, and biochemical reactions, which are described by diffusion-reaction equation.

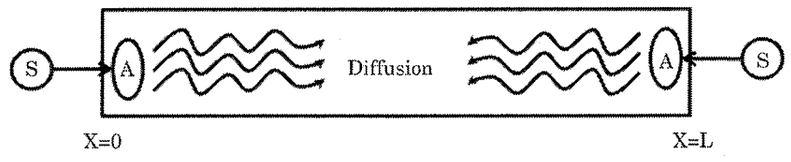
$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + v \qquad (10)$$

where, in Eq.(10), D represents the diffusion coefficient, C is the concentration of the component, t is the time, x is the spatial variable and v represents the resultant rate of generation and consumption of

the component. Nonlinear PDE requires more parameters as compared to ODE and also need to specify the boundary conditions in addition to the initial conditions (Tyson and Kagan, 1988). Moreover, solving PDE requires more computational time as compared to ODE.

To illustrate PDE modeling we have presented an example in which the diffusion of the species, A, is only in the horizontal direction. Fig. 6(a) shows the reaction diffusion scheme, wherein, (1) molecule A is activated at only at two edges of a compartment (at X=0 and X=L) (2) A is subjected to natural degradation and (3) A is subjected to diffusion due to concentration gradient only in X direction (*i.e.*, A is well mixed only in the Y direction). The concentration of A is then a function of both time (t) and space (X) as shown in Fig. 6. Note that the flux of A at the two edges (X=L and X=0) is directly proportional to the stimulus (*e.g.* 0.1 unit in the example). It should be noted that the incoming flux of A in the region, other than the two edges, is due to diffusion alone. Fig. 6(b) shows spatiotemporal distribution of A in the compartment. Initially, A is not present throughout the compartment. Once the stimulus is applied at the two edges, the concentration of A increases with time at each point. Initially the concentration is very low at the midpoint, however as time passes, the concentration of A at the midpoint increases due to diffusion.

(a)



(b) Equation and parameters

$$\frac{\partial[A]}{\partial t} = -K_d[A] + D\frac{\partial^2[A]}{\partial x^2}$$

$K_d = 0.1$ $(\text{sec})^{-1}$, $D = 10\ \frac{mm^2}{sec}$;

$J_L = 0.1\ \frac{mol}{(mm)^2 sec}$;

Boundary conditions:

Flux of A at X=L, and at X=0, J_0=0

(c) Plot

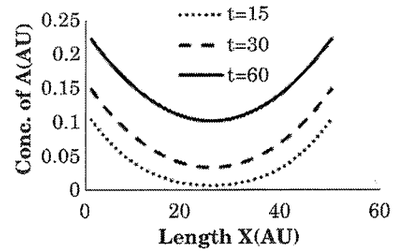


Fig. 6: (a) Cartoon representation of reaction diffusion scheme in which A is produced only at two edges of the compartment. Further, A is subject to degradation and diffusion. (b) Governing equations, parameters and boundary conditions used for the simulation. (c) Plot for Spatial profiles at three different time points.

Stoichiometric Modeling

Mathematical model can also be constructed only based on the stoichiometry of the biochemical reaction called as stoichiometric model. Such a model is useful when mechanistic details and kinetic parameters are not available. Topological structure of the reaction mechanism indicating which species are linked by reactions forms the basis of stoichiometric modeling. Such kind of modeling is widely used in analyzing metabolic networks, which essentially function along with the signaling network to convert the nutritional input into a cellular response. The stoichiometric modeling approach is also demonstrated to be useful for analyzing metabolic network from the perspective of an input–output relationship, crosstalk, measure of redundancy, contribution of individual reactions in signaling pathways and evaluation of co- related reaction rates (Papin and Palsson, 2004). Such properties essentially depend upon the network structures and the stoichiometry of the biochemical reactions.

This approach effectively involves formulating stoichiometric matrix of the network based on the reactions of the network involving allosteric binding, dimerization, phosphorylation and metabolic reactions. In stoichiometric modeling approach, quasi steady state approximation is employed ($0 = N.V$). Stoichiometry matrix N captures the structural relationship between the network components. This results in system of linear equations, which can be solved using linear optimization techniques (Vinod and Venkatesh, 2008). However, this often results in infinite number of solutions. To obtain appropriate solution it is necessary to constrain the optimization based on an objective functions. Network are subjected to mass balance and thermodynamic constraints to generate a set of systemic pathways that can fully characterize a network (Papin and Palsson, 2004). However, major drawback of stoichiometric models is the limited predictive power due to lack of regulatory information, which can only be included in the formulation of a kinetic model (Jamshidi and Palsson, 2010; Kauffman *et al.*, 2003; Ramakrishna *et al.*, 2001). To illustrate the modeling of metabolic flux, a general methodology and an example of flux balance analysis for a synthetic network are described as below.

Following are the steps which are generally followed for development of a model of metabolic flux and determining unknown variables like reaction flux or formation rate from the measurement of few other fluxes by imposing assumptions like pseudo-steady-state condition. The methodology uses following symbols.

Symbols

S_{ji}: Stoichiometric number corresponding to ith species in jth reaction

M: Total reactions; N: Total species; R_j: jth reaction; X_i: ith species

V_j: steady state (S.S.) flux of jth reaction;

r_i: Steady state formation (accumulation) rate of ith species

Step 1: Identify all the reactions (R_1, R_2 ... R_M) and their components (X_1, X_2, ... X_N).

$$R_1 : \sum_{i=1}^{i=N} S_{1i}X_i = 0,\ R_2: \sum_{i=1}^{i=N} S_{2i}X_i = 0\ ....,R_M : \sum_{i=1}^{i=N} S_{Mi}X_i = 0$$

Step 2: Obtain the stoichiometric matrix: $T = \begin{bmatrix} S_{11} & \dots & S_{1N} \\ \dots & \dots & \dots \\ S_{M1} & \dots & S_{MN} \end{bmatrix}$

Stoichiometric matrix T is formed based on stoichiometric coefficients of substrate and product consumed in a particular elementary reaction. Consider a reaction R_2: $4X_1 \rightarrow 3X_2$ in which the stoichiometric coefficient of X_1 is –4 while of X_2 is 3, therefore S_{21}=–4 and S_{22}=3.

Step 3: Obtain equations of flux balance:

$$[r_i] = T^T[V_j] \tag{11}$$

Step 4: Reorienting flux balance equations

Let $r_i = \begin{pmatrix} r_u \\ r_m \\ 0 \end{pmatrix}$, $T^T = \begin{pmatrix} T_u \\ T_{m,0} \end{pmatrix}$, by substituting for r_i in Eq.(11) we obtain

$$\begin{pmatrix} r_u \\ r_m \\ 0 \end{pmatrix} = \begin{pmatrix} T_u \\ T_{m,0} \end{pmatrix} V_j \tag{12}$$

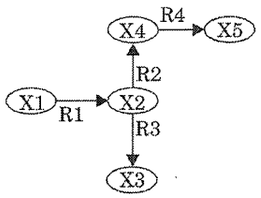
where, r_u is vector of unknown formation rates, r_m is vector of measured formation rates 0 is vector of zero for pseudo steady state metabolites, T_u and T_m,0 are transposes of stoichiometric matrices for the metabolites having unknown and known formation rates, respectively.

Step 5: Determination of unknown variables:

From $\begin{pmatrix} r_m \\ 0 \end{pmatrix} = T_2 V_j$, fluxes can be determined as $V_j = (T_2)^{-1} \begin{pmatrix} r_m \\ 0 \end{pmatrix}$ & from $r_u = T_1 V_j$, unknown formation rates can be determined

In Fig. 7, we provided one example of the above mentioned methodology to analyze metabolic flux balance, where in, a synthetic metabolic network consists of 4 reactions and 5 species. Here the number

(I) Metabolic network:



(II) List of reactions:

R1 : X2 – X1 = 0

R2 : X4 – X2 = 0

R3 : X3 – X2 = 0

R4 : X5 – X4 = 0

(III) Stoichiometric matrix:

Rj/Xi	X1	X2	X3	X4	X5
R1	–1	1	0	0	0
R2	–1	0	0	1	0
R3	0	–1	1	1	0
R4	0	0	0	–1	1

(IV) Flux balance:

$$\begin{pmatrix} r_{X1} \\ r_{X2} \\ r_{X3} \\ r_{X4} \\ r_{X5} \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} V_{R1} \\ V_{R2} \\ V_{R3} \\ V_{R4} \end{pmatrix}$$

(V) Reorientation of matrix:

Note: There are nine variables (i.e., rxi and VRj, i=1 to 5 and j=1 to 4)

(Formation rates of X1 and X3 are measured whereas X2 and X4 are assumed to be zero)

$$(r_{X5}) = (0 \quad 0 \quad 0 \quad 1) * \begin{pmatrix} V_{R1} \\ V_{R2} \\ V_{R3} \\ V_{R4} \end{pmatrix} \text{.....(1)}; \quad \begin{pmatrix} r_{X1} \\ r_{X2} \\ r_{X3} \\ r_{X4} \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} * \begin{pmatrix} V_{R1} \\ V_{R2} \\ V_{R3} \\ V_{R4} \end{pmatrix} \text{.....(2)}$$

From (2): $\begin{pmatrix} V_{R1} \\ V_{R2} \\ V_{R3} \\ V_{R4} \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} r_{X1} \\ r_{X2} \\ r_{X3} \\ r_{X4} \end{pmatrix}$; $\begin{pmatrix} V_{R1} \\ V_{R2} \\ V_{R3} \\ V_{R4} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} r_{X1} \\ r_{X2} \\ r_{X3} \\ r_{X4} \end{pmatrix}$;

hence, $V_{R1} = r_{x2} + r_{x3}$; $V_{R2} = -r_{x1} - r_{x2} - r_{x3}$; $V_{R3} = r_{x3}$; $V_{R4} = -r_{x1} - r_{x2} - r_{x3} - r_{x4}$

here, $r_{x2} = r_{x4} = 0$; *and let* $r_{x1} = -2$; $r_{x3} = 1$;

hence, $V_{R1} = 0 + 1 = 1$; $V_{R2} = -(-2) - 0 - 1 = 1$;

$V_{R3} = 1$; $V_{R4} = -(-2) - 0 - 1 - 0 = 1$, from (1); $(r_{x5}) = V_{R4} = 1$

Fig. 7: Example of Flux balance analysis for a synthetic network. (I) and (II) shows a synthetic metabolic pathway and its reactions (III) Stoichiometric matrix (IV) flux balance equations, (V) Reorientation of flux balance equation so that unknown variables can be explicitly represented as a function of known variables.

of unknowns are 9 (4 reaction fluxes and 5 formation/consumption rates of species) and number of equations are 5 (flux balance equations), therefore to determine the unknown variables, the value of 4 variables need to be fixed. This is achieved in this example by measuring the rate of formation of X_1 and X_3 and assuming steady state condition for X_2 and X_4 in Fig. 7.

Stochastic Modeling

Contrary to deterministic processes, many biological reactions occur infrequently leading to fluctuations, which can be attributed to inherent randomness in the molecular interactions which lead to stochastic nature of the reaction (El-Samad and Khammash, 2010). Lesser numbers of molecules and limited diffusion due to the structural organization of the cell contribute toward the fluctuation of biochemical reaction (Savant, 2007). Stochastic modelling approach involves predicting the probability of collision between molecules resulting in reactions at discrete time intervals (Gillespie, 2008). Stochastic modelling can be performed using stochastic differential equations , stochastic simulation algorithms and molecular dynamics simulation (Kulasiri, 2008). Stochastic models can be broadly characterized based upon their nature of random variables and probability distribution functions such as different kinds of Markovian processes and Monte-Carlo methods.

Master Equation Models

The discrete probability distribution of reaction as function of time is described by chemical master equation (CME) (Gillespie, 1977; Ridwan *et al.*, 2004). The master equations are derived from the Markov properties of chemical reactions by writing differential form of Chaman-Kolmogorov equation (Kampen, 1992). They assume the well mix system in which only one reaction occurs at infinitesimal time interval. Based on physical considerations master equation transform rates of chemical reactions into rates of transition probabilities which depicts the way in which the state of system changes with changes in probability distribution with time (Kulasiri, 2008). The general form of Chemical master equation can be given by

$$\frac{\partial P(X;t)}{\partial t} = \sum_{j=1}^{M} \left[a_j \left(X - v_j \right) P\left(X - v_j ; t \right) - a_j (X) P(X;t) \right] \quad (13)$$

where, $P(x,t)$ is the probability of the system state $X(t)=x$ at time t, a_j is propensity function and v_j is state change vector or stoichiometric vector.

Although these equations are difficult to solve analytically there are methods to solve them numerically by using stochastic simulation algorithms (SSA). One key simulation technique is the stochastic simulation approach to chemical reactions developed by Gillespie (1977) which implements Monte-Carlo method for getting the exact numerical solution (Gillespie, 1977). The other way is to approximate the CME to Fokker-Planck equation and Linear noise approximation and solve it numerically (El-Samad and Khammash, 2010). However, stochastic representations are complicated in nature and hence the system can be modeled as continuous reactions before attempting using stochastic simulations.

Stochastic Simulation Algorithm (SSA)/Gillespie Algorithm

Gillespie algorithm is used for obtaining numerical solutions for stochastic equations such as CME (Chemical Master Equations) (Gillespie, 2008). It is used to simulate the discrete probability distribution of the reactions in the network with respect to time (Gillespie, 1977). It follows Monte-Carlo method in which initially the defined number of reacting species is set to zero simulation time. Based on current abundance of the reacting species the propensity of each probable reaction is calculated. These propensities are used to simulate the time required for next reaction to occur and the simulation time is upgraded accordingly. Then random selection of reaction events is carried out with the probabilities in proportionate with reaction propensities and the number of reacting species is updated accordingly followed by recording time and state of the system. This process is repeated till the simulation ends with the exhaustion of reacting species. This algorithm precisely monitors the dynamics of each and every species which requires large amount of simulation time and makes the algorithm slow. Efforts are being made to increase the speed of these algorithms by certain approximations and leaping the time values in the simulation (Wilkinson, 2009; El-Samad and Khammash, 2010; Kulasiri, 2008)

Stochastic Differential Equations (SDE)/ Chemical Langevin Equation (CLE)

It is computationally quite expensive to stochastically simulate large and complex reaction network containing fast reactions. Chemical Langevin equations replaces the large differential equations with small stochastic equations which are easier to solve compared to CME. These models account for the external fluctuations in the system by introducing a noise term in the rate equation (Khanin and Higham, 2008). It has been demonstrated that even the CMEs can give rise to CLE under

specific conditions wherein intrinsic noise in the system is approximated. The general form of Chemical Langevin equation is given by

$$\frac{DX(t)}{dt} = \sum_{j=1}^{M} v_j a_j\left(X(t)\right) + \sum_{j=1}^{M} v_j \sqrt{a_j\left(X(t)\right)}\, \Gamma_j(t) \quad (14)$$

where, a_j is propensity function , v_j is state change vector or stoichiometric vector and Γ_j ' white noise' term. This equation is also called as 'white noise ' form of CLE where the second term in the equation defines the randomness of noise in the system (Kulasiri, 2008). To illustrate the methodology of CLE modeling we provide below an example of an enzyme-substrate reaction. Fig. 8(I) shows reaction scheme; enzyme E associates with substrate S to form a complex ES, which gets either converted to product P or dissociated back to E and S. The CLE equation of the scheme is shown in the Fig. 8(III). The parameters and random number description is also shown in the Fig. 8(V). The simulation results shows comparison of product concentration profiles obtained by the deterministic and stochastic modeling. It can be seen that in the product profile for the case of stochastic simulation fluctuates around mean profile (deterministic profile). This is due to the addition of the noise terms in each differential equation (last term in Eq. 15).

MISCELLANEOUS MODELING

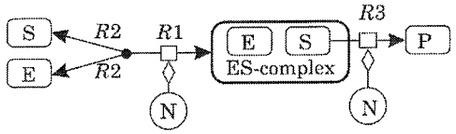
Molecular Dynamics Simulation

Biological interactions can be modeled using molecular dynamic simulation. It can account for the spatio-temporal evolution of both reacting and non reacting interaction in the network. This approach is extensively used in modeling folding and stability of proteins, conformational changes in proteins, enzymatic reactions, transport processes in biological systems and provides static and dynamic properties of the molecules based on the interactions (Martin and McCammon, 2002). This approach can be applied to the modeling of biological networks where only the dynamics of the reacting species are considered by assuming uniform spatial distribution (El-Samad and Khammash, 2010).

Rule-Based Formalisms

Knowledge-based or rule-based simulation formalisms allows higher variety of knowledge about the system to be expressed in a single

(I) Reaction network:



(II) Stoichiometric matrix

	E	S	ES	P
R1	−1	−1	1	0
R2	1	1	−1	0
R3	1	0	−1	1

(III) Stochastic differential equation model

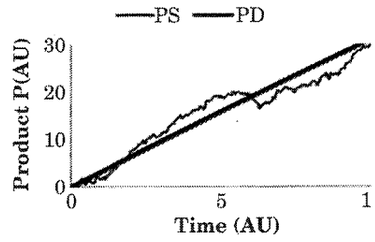
$$dY_i(t) = \sum_{j=R1}^{R3} s_j r_j(Y(t))dt + \sum_{j=1}^{3} s_j \sqrt{r_j(Y(t))}dW_j(t) \quad (15)$$

where Yi is concentration of species i(i=E,S,ES,P) s_j is an array of stochiometric coefficients for reaction j(j=R1,R2,R3), r_j is deterministic rate of reaction j for the given concentration at time t, which are as following:

r1=k1*S*E; r2=k2*ES; r3=k3*ES; dW indicates an infinitesimal for Brownian path W (white noise term), which is obtained by following equation

$dW = \sqrt{dt} * RV(0,1)$, where RV is random variable from a distribution between 0 and 1; dt is a small time step.

(IV) Comparison of deterministic and stochastic results



(V) Parameters and initial conditions:

k1=; k2=0.1; k=3=0.01; (units are arbitrary)

E = 100; S = 100; P = 0; ES = 0;

Here we considered poison distribution in RV(0, 1)

The simulation was done up to 10 units and dt was taken as 1/30 unit.

Fig. 8: CLE modeling approach applied to an enzyme-substrate reaction scheme. (I) Reaction network wherein, substrate S associates with enzyme E to form a complex ES. ES can undergo either dissociation or another reaction in which product P is formed along with enzyme E. (II) Stoichiometric matrix where rows corresponding to reaction and columns correspond to species, each element is stoichiometric coefficient of a species involved in a specific reaction (III) Model description (IV) Simulation results showing comparison for profiles P obtained in CLE modeling and deterministic modeling.(V) Parameters used for the simulation.

formalism. For example, the logical and physical structure of a gene such as the relative position of the regulatory sites at which transcription is initiated, prevented, and aborted, can be conveniently represented in the rule-based approach (Jong, 2002; Gutierrez-Rios *et al.*, 2003). Basically, rule-based formalisms consist of two components, a set of facts and a set of rules that are stored out in a knowledge base. Facts express knowledge about the objects of a regulatory system. The rules in the

knowledge base consist of two parts, a condition part and an action part. The condition part expresses conditions in terms of properties of objects; while the action part operates upon the objects, *e.g.*, by changing a property of an existing object.

Cellular Automata Models

Cellular automata (CA) models are lattice based models which perform complex computations using local information about the mechanistic details of the system. CA uses discrete lattice cells, each cell representing one possible discrete state (Gilbert *et al.*, 2006). The lattice cells are defined based on optimal dimensions of system at particular time, based on which the fate of neighboring lattice space is decided. At discrete unit time, cell updates its current state by transition-rules. CA can be applied for biophysical simulations (David *et al.*, 2005; Gordon *et al.*, 2005) base upon simple rules defined by the modeler. It essentially doesn't require development of differential equations and are much more easier computationally as compared to their mathematical counterparts.

Spatial and Agent Based Models

Spatio-temporal models such as Agent based modeling (Zambonelli and Omicini, 2004) and Spatial modeling are some of the other techniques to which modelers are resorting to understand the complex dynamics of the biological networks (Gilbert *et al.*, 2006). Agent-based models are the software agents which represent the discrete elements of the models. Using various artificial intelligence (Fei-Yue, 2005) paradigms such as artificial neural networks, these entities incorporate the spatio-temporal attributes of the elements. Spatial modeling is applied for time-space driven events such as travelling waves in biological process (wound healing and immunological systems), mechano-chemical theory of morphogenesis, pattern formations and spatial distributions using plot samples (Murray, 2003).

Model Parameter Estimation

There are several variables in the system which govern the dynamics of the system but do not represent the system state. The variables that define the state of the system are called as the state variables while those define the system dynamics are called as system parameters. The process of parameter estimation is also called as 'Model calibration' (Markschies, 2008). Rate constants, diffusivities, equilibrium constants and physical properties of the components are said to be system

parameters. It is essential to know this information for the simulation and analysis of model. Searching through literature and databases along with manual evaluation is primary way to get the parameters. However most of these are unknown and are difficult to measure experimentally. In such situation sometimes constructing the black box model is helpful wherein the system inputs are tried to correlate with the system output with various hypotheses (Sriram, 2010). If this approach doesn't work one can resort to reverse engineering techniques where observable information can be retracted to evaluate the system parameters. One can even try iterative process with assuming some initial guess and after many iterations and experimental validation the system parameters can be deduced, however this approach is limited by the smaller sample space.

Model Optimization and Validation

In cases of large number of parameter estimation, sophisticated computational techniques such as optimization methods and algorithms are widely used. The goal of parameter estimation using optimization technique is to locate the optimum for the ODE models and find the possible parameters that minimize the difference between experimental and simulation results which can be given as

$$min\phi(\theta), \qquad J=\frac{\partial \phi}{\partial \theta}, \qquad H=\frac{\partial^2 \phi}{\partial \theta \partial \theta^T} \tag{16}$$

where function $\Phi(\theta)$ represents the goodness of fit between experiment and simulation, it is a scalar function of the parameter vector θ. J is partial derivative of objective function (partial derivative matrix, a Jacobian), H is a Hessian matrix containing second derivative of objective function with respect to pairs of parameters.

This kind of parameters estimated by simulation to match experimental results is called as 'in-sample fit'. Sometimes, to evaluate the simultaneous effect of multiple parameters on the system, the higher order partial derivatives of the objective function with respect to each variable are taken which is represented by the Hessian matrix (Sriram, 2010). This kind of approach is implemented using Gradient search optimization algorithms. Various sophisticated algorithms using probabilistic approach and Monte-Carlo methods have been developed for optimizing biological data which are also prone to noise.

Although the parameters obtained from this technique might fit the experimental data, sometimes they fail to evaluate the models

predictability. To tackle this problem we need to make newer predictions and compare with new experimental data. After many iterations certain set of parameters are obtained that can well be fit to models predictive ability under various conditions. Such a process is called as 'out- sample fit' or model validation where the models can precisely predict the experimental output under varied conditions. Thus obtained model can be more reliable to simulate the outcome of complex systems (Markschies, 2008).

MODEL ANALYSIS

Steady State Analysis

Kinetic modeling of biochemical reactions can be simplified considerably if the overall reaction is studied with the aid of the quasi-steady-state or equilibrium approximations. The system is said to be at steady state if the concentration of system components do not change with time. Steady state behavior of the system can be obtained by setting the derivatives of all concentrations zero and solving a set of non linear algebraic equations simultaneously. The general form of equation can be given by

$$\frac{d[X^{ss}]}{dt} = V_r - k_t[X^{ss}] = 0 \tag{17}$$

where, X^{ss} is the steady state concentration of the component, V_r is the rate of production and K_t is the degradation rate. Steady state modeling approach provides insights into the emergent properties of the network and also helps to identify the role and contribution of individual regulatory structures. To illustrate utility of the steady state modeling, we have discussed different kinds of analysis with examples.

Parameter Estimation

Often steady state modeling is used to obtain parameter which can further be used for the dynamic modeling. We have provided one example of parameter estimation by steady state modeling, wherein, a hill equation is assumed to represent relationship between two molecules x and y. The parameters are estimated by fitting the steady-state predictions of the model to some known data points. The table in Fig. 9 shows data points for concentration of y and x. We assumed that concentration of y is related to the concentration of x with a Hill function and therefore estimated the Hill coefficient (n) and the half saturation constant Km.

(I) Data table:

x: Input, y: Output

x	y	x	y
0	0	0.6	0.525
0.1	0.001	0.7	0.635
0.2	0.012	0.8	0.759
0.3	0.060	0.9	0.845
0.4	0.159	1	0.885
0.5	0.325	2	1.000

(II) Linearization plot

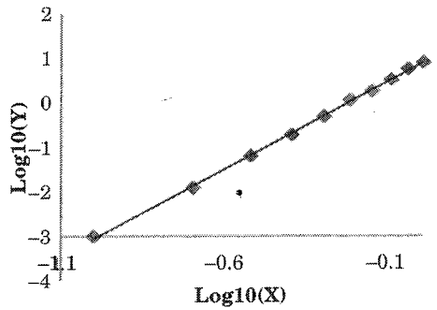


Fig. 9: (1) Data table listing concetration of x and y, where in x is changed and y is measured at steady state 2) The linear Plot of log (X) vs log (Y), where X and Y are as mentioned in the manuscript.

The hill function of x relating to y, which can be given as:

$$y = ymax * \left(\frac{x^n}{K_m^n + x^n}\right); \tag{18}$$

$$\text{let } Y = \frac{y}{y_{max} - y}, \tag{19}$$

then Hill equation relationship between x and y can be written in terms of X and Y as following

$$Y = \left(\frac{X}{K_m}\right)^n \tag{20}$$

Taking log10 on both sides gives,

$$\log(Y) = n * log(X) - n * log(K_m)$$

The constants n and K_m now can be obtained by plotting Log(Y) vs Log(X) for the given data points and then estimating the slope and the intercept of the fitted straight line. Fig. 9 (II) shows the linear relationship between log(Y) and Log(X) for the given data points, for which Slope=3.830 and intercept (*i.e.*, $-n*\log(K_m)$,) = 0.877; hence Hill coefficient n=3.83 and half saturation constant K_m=0.590 (See Fig. 9).

ANALYSIS OF AMPLIFICATION AND ULTRA-SENSITIVITY

For the input-output relationships between two variables, one can analyse signal amplification and ultra-sensitivity of the relationship by

fitting Hill function as described in the previous section. The Hill function with lesser *Km* as compared to the normal (half saturation constant) suggests that signal amplification is higher, while the higher value (more than 1) of the Hill coefficient '*n*' suggests that the input-output relationship is ultrasensitive. Fig. 10 compares steady-state input-output curves (*i.e.* dose response curves) for various values of hill function parameters *Km* and *n*. It can be seen from the Fig. 10(I), that; 1) for the given K_m (*i.e.* K_m=0.2) response curve is more like a on/off switch for n=4, while it is graded for n=1, therefore the input-output relationship given by the Hill function with n=4 can be termed as ultrasensitive. In Fig. 10 (II), for the given n (say n=4) the response is switched on earlier for the case with lower K_m value (*i.e.* K_m=0.2), therefore signal amplification in the relationship given by a Hill function with low half saturation constant is higher compare to the one with higher half saturation constant.

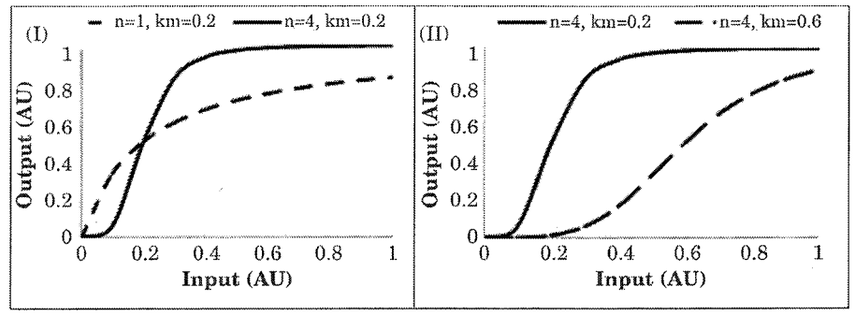


Fig. 10: Effect of Hill coefficients and half saturation constants on the input-output relationship. Dose response curves representing I) Ultra-sensitivity - the solid curve represents the sensitive response with the Hill coefficient n=4 as compared to the dashed curve that represents graded response with Hill coefficient n=1. Note that the K_m is identical in both the cases. II) Amplification- solid curve represents the amplified response with half saturation constant k_m=0.2, as compared to the dashed curve with K_m= 0.6. Note that the n identical in both the cases.

Analysis for the Robustness

Robustness of the network can be defined based on its capacity to restrict the variation in the steady state response with respect to the variation in the input. Various network structures (*e.g.* negative feedback loop, incoherent feed forward loop (IFFL)), play a role in the robustness of the system in maintaining the variable around a fixed value. Fig. 11 (I) shows one such IFFL, wherein, activator A and Inhibitor B are up-regulated by the stimulus S. Both A and B act antagonistically on molecule R, wherein, A enhances level of R whereas B reduces it. The

governing equations of the network are shown in the Fig. 11(II). The robustness of the network in maintaining R at a fixed value can be tested by obtained the steady state equations (see Fig. 11(III)) of R in terms of other variables and parameters. By substituting for activator's and inhibitor's concentration in terms of stimulus S, one can obtain concentration of R in terms of parameter and stimulus. For the given case, as both activator and inhibitor depends on S with a similar order of reaction (*i.e.* first order), the substitutions give equation of R in which stimulus effect is eliminated. Hence this system can be considered as very robust with respect to external disturbance. Fig. 11(IV) shows dose response curves for A, B and R. It can be seen that while A and B increase with stimulus S, the level of R is maintained at fixed value.

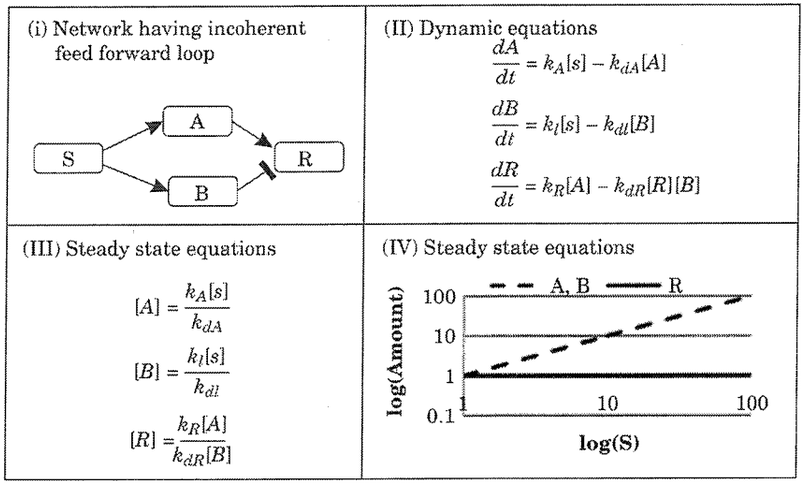


Fig. 11: (I) Incoherent feed-forward loop showing activation and deactivation of R by A and B, respectively. Both A and B are activated by a common stimulus S. (II) Dynamic equations which are used to obtain steady state equation (III) Reoriented steady state equations (IV) Dose response curves for A, B and R showing that while A and B varies with stimulus, R remains constant.

Analysis for Multiple Steady States

Biological systems are governed by the highly non-linear functions of the components involved in the system. Most of the times, the non-linear equations provides multiples solution, therefore, steady state model can be analyzed to obtain all real positive roots of the system. These roots denote the steady states of the model components. Typically linear stability analysis is used to characterize the nature of the steady states (stable/ unstable) which is described in the later section (Kapuy *et al.*, 2009).

To illustrate the utility of steady state modeling to obtain all possible phases of the system, we have shown here one simple example. The Fig. 12(I) shows a network motif, wherein, A is subject to activation by action of S. Further, the activated A restricts the backward reaction (*i.e.*, deactivation reaction). The governing equations are shown in the Fig. 12(II) which is obtained by equating the accumulation rate to zero. This equation is a non-linear algebraic equation which can give multiple solution of concentration of A. Among those roots, only non-zero and real roots are feasible. Fig. 12(III) shows plot of A for various value of stimulus. It can be seen that for the stimulus range S=0.05:0.25, there are three realistic roots (denoted by open triangles, circles and diamonds). However outside this range there is only one feasible root.

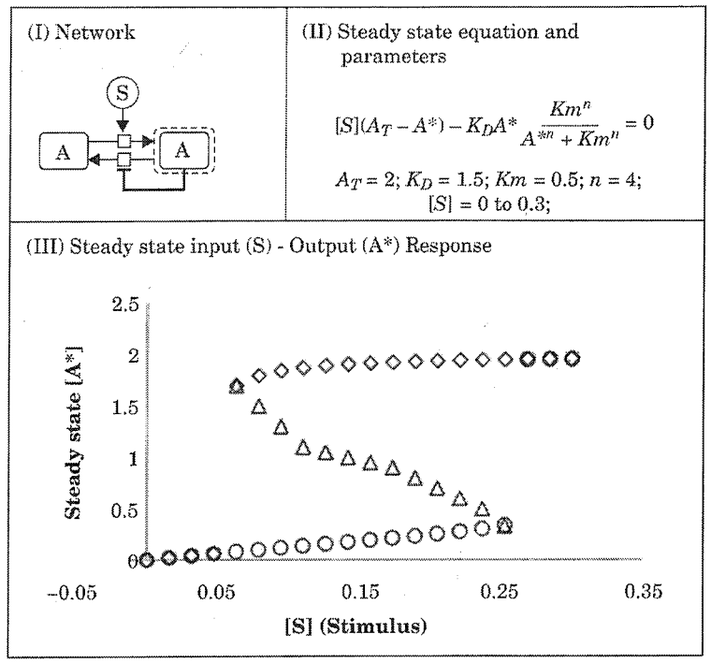


Fig. 12: (I) Network in which A is activated by stimulus S and the activated A inhibits its own deactivation, thereby making a positive feedback loop. (II) Governing equation of the reaction scheme, where A* is activated A. (III) The input output response showing multiple feasible solutions for A for various values of stimulus.

Dynamic Analysis

Dynamic behavior of regulatory pathways can be studied by solving a set of coupled non-linear ordinary differential equations (ODEs)

representing the individual components of the system, which is given by

$$\frac{dC}{dt} = N \cdot V \tag{21}$$

where, component concentrations of the network is represented as $C = [c1c2c3 \ldots cn]^T$ and reaction rates of the network (flux distribution) is represented as $V = [v1v2v3 \ldots vm]^T$.

Individual reaction rates depend on the kinetic description of the reaction involving regulation of enzymes and kinetic parameters. Further, N represents the stoichiometric matrix, where rows of the matrix correspond to network components and columns of the matrix correspond to reactions with each element of the matrix as the stoichiometric coefficient of a component in the associated reaction. N is invariant against time, kinetics and concentrations. Such formulation of the network is termed as kinetic modeling of regulatory pathways (Vinod and Venkatesh, 2008). These dynamic models can be analyzed through parametric variations to study hidden tendency of the system in showing a particular type of the response (Lee *et al.*, 2008).

Sensitivity Analysis

The models formulated based on the above stated methodologies can be analyzed by subjecting the model to perturbation in order to understand the influence of component concentrations and model parameter values on the overall response of the network. Such an approach is termed as parametric sensitivity analysis. Systemic behavior is evaluated with respect to variation in single parameter or multiple parameters, to evaluate the key parameters in the network and to study robustness of the network. Sensitivity analysis is carried by two methods: Local sensitivity analysis and global sensitivity analysis. In local sensitivity analysis the changes in the model output with respect to small variation in the parameters are studied. These variations are measured by the sensitivity coefficients which are the first order derivatives of the model with respect to model parameters (Cho *et al.*, 2003, Haseltine and Arnold, 2007; Sumner *et al.*, 2012). The equation for local sensitivity analysis can be given as

$$s = \frac{\partial G(N)}{\partial \theta} \tag{22}$$

where N is vector of species concentration, $G(N)$ is the system output and θ is the vector of parameters.

In biological systems often the parameters including rate constants and initial concentrations are varied in a large range depending upon the specific cell types and cellular environments. Therefore the global sensitivity analysis algorithms are required to handle the possible non-linear effects of parameters by simultaneous variations of all the parameters with Monte-Carlo simulations or multi parametric sensitivity analysis (Keasling, 2008).

Variation of multiple parameters results in multidimensional analysis with sensitivity of network behavior varying in space. Such an analysis provides a global view of the network behavior with operating zones based on the minimization of objective function. The solution space is represented by hills and valleys, with valleys representing the less sensitivity region in response to variation in parameters. With the help of optimization techniques global minima of the objective function can be determined (Moles, 2003). Such optimization techniques can also be used to estimate the parameters from experimental results. Moreover, the reference parameter set of the network can be defined and test parameter sets can be generated by random variation of parameters. The sensitivity of the global response can be plotted against the distance between the reference and test parameter set to get a scatter plot, which gives the measure of robustness of the network.

Flux Control Analysis

This analysis is more generally termed as Metabolic Control Analysis (MCA) which is used to quantify the relative control of each system variable on the network properties. This approach is generally used for quantifying metabolic networks but can be extended to study the dynamic properties of all biological networks. The sensitivity of network is analyzed by evaluating control coefficients for flux and concentration, elasticity coefficients and the response coefficients for the individual system components in the network. Control coefficient determines the control exerted by the activity of single enzyme on the pathway flux and the internal metabolites. Elasticity coefficient determines the sensitivity of bio-molecular reactions in the pathway with respect to local environment such as substrates, products and effectors. Response coefficient determines the action of the external parameters on the system variables. The equations for these coefficients are given by

$$C_i^J = \frac{E_i}{J}\left(\frac{\partial J}{\partial E_i}\right) = \frac{\partial ln|J|}{\partial E_i},\ i=1,2,3..n \text{ Flux control coefficient} \tag{23}$$

$$E_k^i = \frac{S_k}{v_i}\left(\frac{\partial v_i}{\partial S_k}\right) = \frac{\partial ln|v_i|}{\partial S_k}, i=1,2,3..n,\ k=\ 1,2,3,\ldots n\ \textit{Elasticity coefficient} \quad (24)$$

$$R_P^J = \frac{P}{J}\left(\frac{\partial J}{\partial P}\right) = \frac{\partial ln|J|}{\partial lnP},\ \textit{Response coefficient} \quad (25)$$

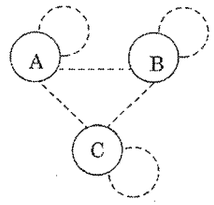
where J is the system flux, E is the enzyme, S is the metabolite, v_i is reaction rate and P is external parameter (Schuster,and Heinrich, 1992). The kinetic equations of the bio-molecular reactions are used to parameterize the control coefficients resulting into response coefficients. The system response to the perturbation is quantified by summation of the response coefficients which affect the desired output. This way the global dynamic properties of the network can be related with the individual reaction (Wildermuth, 2000; Steuer and Junker, 2008).

Linear Stability Analysis

In previous sections, we have shown that the steady state analysis can be used to provide information about number of realistic roots (steady solutions) for the system. However, to determine the nature of the solution, *i.e.* stable or unstable, linear stability analysis is used. Fig. 13 (shows various steps of linear stability analysis as following 1) establishing the network (Fig. 13 (I)); 2) obtaining differential equations of the network (Fig. 13 (II)); 3) obtain Jacobian of the system, which is matrix corresponding to partial derivative of steady equation with respect to variable (Fig. 13 (III)); 4) obtain roots of the steady state equation of the system; 5) obtain Eigenvalues for the possible roots of the steady state equations; Eigenvalues can be obtained by solving characteristic equation of the system as shown in the Fig. 13 (IV).

The sign of Eigenvalues and whether it is real or complex number, determines the nature of solution of the system. Various cases are shown in the Fig. 13 (V), where in, the solution is stable if all the Eigenvalues are real negative numbers and unstable if they are real positive numbers. For the case of complex Eigenvalues; the solution is asymptotically stable if real part of the Eigenvalue is negative and it is unstable otherwise. In case of stable node, as in case of real negative Eigenvalues, the minor fluctuations in the level of any variable (fluctuation in concentration of A, B or C) will not result in different steady state or will not make the system unstable in which concentration keeps on creasing with respect to time. However in case of un-stable node, as in case of real positive Eigenvalues, minor fluctuation (whether positive or negative) results in unstable system.

(I) Network



(II) Model equations

$$\frac{dA}{dt} = f_1(A, B, C)$$

$$\frac{dB}{dt} = f_2(A, B, C)$$

$$\frac{dC}{dt} = f_3(A, B, C)$$

(III) Jacobian matrixA

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial (A_0)} & \frac{\partial f_1}{\partial (B_0)} & \frac{\partial f_1}{\partial (C_0)} \\ \frac{\partial f_2}{\partial (A_0)} & \frac{\partial f_2}{\partial (B_0)} & \frac{\partial f_2}{\partial (C_0)} \\ \frac{\partial f_3}{\partial (A_0)} & \frac{\partial f_3}{\partial (B_0)} & \frac{\partial f_3}{\partial (C_0)} \end{bmatrix}$$

(IV) Characteristic equations:

$|J - \lambda I| = 0$; where J is matrix of partial derivative as obtained in III,

I is unit matrix and λ is eigen value. The characteristic equation in our case is a cubic equation as $\lambda^3 + c_1\lambda^2 + c_2\lambda + c_0 = 0$

(V) Eigen values and stability of the system

For real λ_i	*Nature of solution*	*For image $(\lambda_i) \neq 0$*	*Nature of solution*
$\lambda_i > 0$	Unstable node/ saddle point	Real $(\lambda_i) > 0$	Unstable focus (Oscillations with increasing amplitudes)
$\lambda_i < 0$	Stable node	Real $(\lambda_i) < 0$	Stable focus (Oscillations with decreasing amplitudes)
$\lambda_i = 0$	Multiple solution, with stable or unstable lines	Real $(\lambda_i) = 0$	Limit cycle (Oscillations with constant amplitudes)

Fig. 13: (I) A synthetic network consists of interactions between three proteins A, B and C. The dotted edges indicate possible interactions (positive/negative) (II) ODE equations that determine the effects of interactions. (III) Jacobian matrix, obtained by partial differentiation of each function (*i.e.* f1, f2 and f3 in (II)) with respect to the amount of each protein (*i.e.* A, B and C). (IV) List of possible cases classified according to the sign of Eigenvalues and also based on whether the Eigenvalue is real or values. The sign of real and imaginary part of value determine the nature of solution.

Bifurcation Analysis

Dynamic analysis methods such as stability and bifurcation analysis are often used to identify the qualitative changes occurring in a non-linear dynamical system with respect to parameter variation (Doedel, 1991; Chickarmane *et al.*, 2007). The dynamics of the system changes qualitatively either with system returning to original steady state called stable steady state or become unstable or shift to new steady state under perturbation. Such qualitative change in the location and stability of steady states is determined by the parameter values of the system with

different possibilities emerging in different parameter ranges (Goldbeter, 1996; Tyson, 1996).

Bifurcation analysis traces down the qualitative changes which occurs at points in the parameter space called bifurcation points (Zumsande and Gross, 2010; Sriram, 2010). The set of non-linear differential equation is solved at steady state and the stability of steady state is determined based on the Eigenvalues of Jacobian matrix. The most common bifurcations in biochemical networks are saddle node bifurcation and Hopf bifurcation, which usually leads to bistability and limit cycle oscillations, respectively (Vinod and Venkatesh, 2008). We have further illustrated two bifurcations commonly observed in biological systems.

Saddle node/ Limit point bifurcation

Limit point bifurcation analysis is used for investigating the transition points where the system behavior changes from stable to unstable. This essentially determines the thresholds where system behaves like switch and looses mono-stability. It happens when one real Eigenvalue of the Jacobian matrix crosses the imaginary plane (Gardner *et al.*, 2000).

Figure 14(I) shows network which consist of interaction between A and B. A exerts positive effect on the B, which in turn reduces deactivation of A. Thus there is a positive feedback loop in the system. Such positive feedback loop, is known to show bistability, wherein, two stable nodes are possible for a given value of a parameter. We presented here procedures to obtain the critical value of the parameter which separates two natures of the system *i.e.*, monostability and bistability. Fig. 14(II) shows governing equations of the network which can be used to obtain roots of the system for various values of parameter K_{DA} *i.e.*, the degradation rate constant of A. Note that we have fixed other parameters and have shown effect of parameter K_{DA} only. Further, the Eigenvalues can be obtained for each value of K_{DA} by using the Jacobian as shown in the Fig. 14(III). Note that the functions f_1 and f_2 are steady state equations of A and B, respectively. Since f_1 is function of K_{DA}, the partial derivatives df_1/dA and df_2/dB are also dependent on the parameter K_{DA}. We can obtain the characteristic equation as a function of A, B and K_{DA}. Thus there are three equations (two steady state equations and one characteristic equation) and four unknowns (A, B, K_{DA} and Eigenvalue). Fixing K_{DA} therefore provides Eigenvalue which can be analyzed as described in the previous section. Fig. 14 (IV and V) shows the plot of Eigenvalue and A as a function of parameter K_{DA}, respectively. It can be seen that for K_{DA} <0.55, there is only one

Eigenvalue and its sign is negative. This suggests that there is only one feasible root and it is a stable node. However for 1> K_{DA} [0.55,1] there are three Eigenvalues; two have negative signs suggesting two stable nodes and one has positive sign suggesting a saddle point (unstable node). Thus this suggests that system can achieve two distinct stable solutions over a range of K_{DA}. This is termed as bistability and the critical point after which there are multiple roots is termed as bifurcation point. This bifurcation is specified as saddle-node bifurcation because after bifurcation point two additional roots are created namely saddle and stable node.

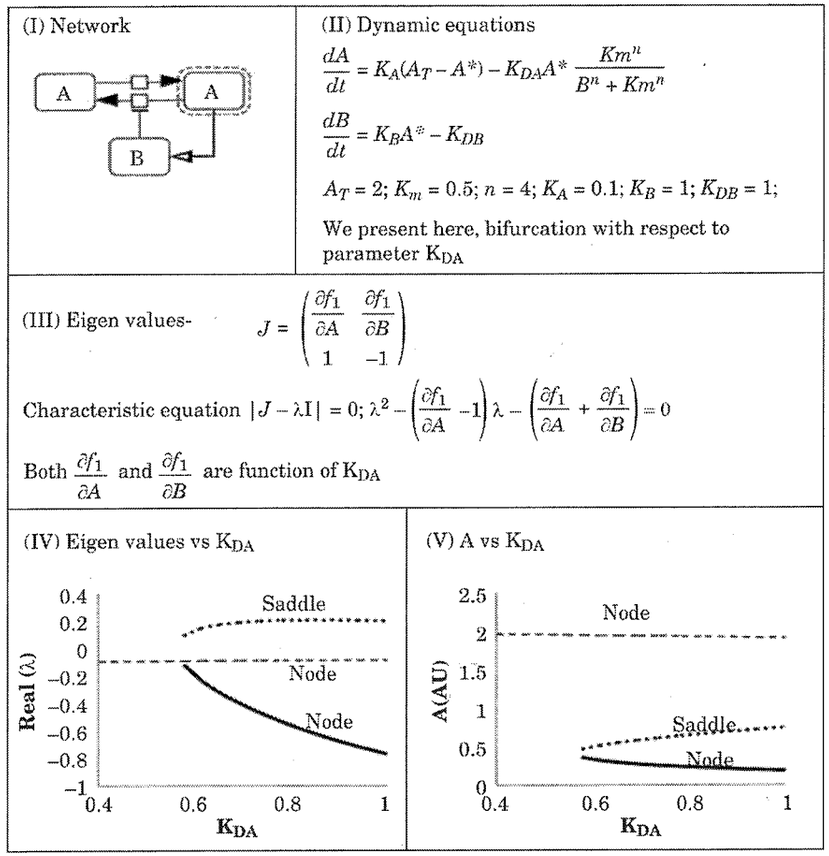


Fig. 14: (I) A network motif involving two molecules A and B. A has positive effect on B, which inhibits deactivation of A. Thus there is a positive feedback loop in the system. (II) Dynamic equations representing the network (III) Procedure for obtaining Eigenvalues (IV) Real part of Eigenvalue as a function of parameter K_{DA}. (V) Steady state A as a function of K_{DA}.

Supercritical Hopf bifurcation

Hopf bifurcation analysis determines the parameter values that are responsible for the transition of the system from stable steady state to oscillatory behavior. It determines the thresholds of the point where the system dynamics tend to oscillate leading to periodic solutions (Garcia-Ojalvo *et al.*, 2004; Hasty, 2002). This happens when the pair of imaginary Eigenvalues crosses the imaginary plane in the phase plane diagram. At this point the real part of the leading Eigenvalues is zero. As described in the previous as section, selection of parameter value can give distinct number of feasible solution in the system. There is another class of bifurcations in which the number of feasible solutions may not vary with parameters, however, the nature of solution could vary. Supercritical hofp bifurcation is one such type, wherein, the nature of solution changes to stable focus from stable limit cycle or vice a versa. A limit cycle is a dynamic nature of solution in which the solution oscillates between two values. A stable limit cycle is the one in which the stable oscillations are maintained and the amplitude do not change with time. Note that the amplitude may increase or decrease with parameter depending on the parameter values.

Figure 15(I) shows a network that consist of S0, S1 and R, wherein, R activates S0, which in turn, activates S1. S0, increases its own activation by an autocatalytic action. Further, there is a negative feedback on S0, mediated by S1. Such system with coupled auto-positive and negative feedback is known to demonstrate oscillations (Ferrell *et al.*, 2011). We will analyze the system for the conditions in which the oscillation dies out or are maintained. Such analysis can be done using Eigenvalue analysis as describe earlier. Fig. 15(II) shows dynamic equation which are used to obtain Eigenvalues as show in the Fig. 15 (III). It should be noted that the Eigenvalue of this system is a complex number and the figure shows only the real part of the complex Eigenvalue. It can be seen from the figure that, at k_{d1}=0.48 the Eigenvalue changes its sign (positive to negative). This point is termed as Hofp bifurcation point as solution transform from stable focus to limit cycle. It should be noted that there are many other systems in which the stable focus may be lost to an unstable focus, in which the amplitude of oscillation increase with time, however there always exists a point at which limit cycle exists. Unlike the limit cycle for the current system, such limit cycle is termed as unstable limit cycle. Fig. 15 (IV, V & VI) show temporal profiles of S0 for three distinct values of kd1, wherein, for kd1 values higher than Hopf point, the system shows damped oscillations; for kd1 lower or equal to Hopf point, system shows sustained oscillations, with increase in amplitude at lower kd1.

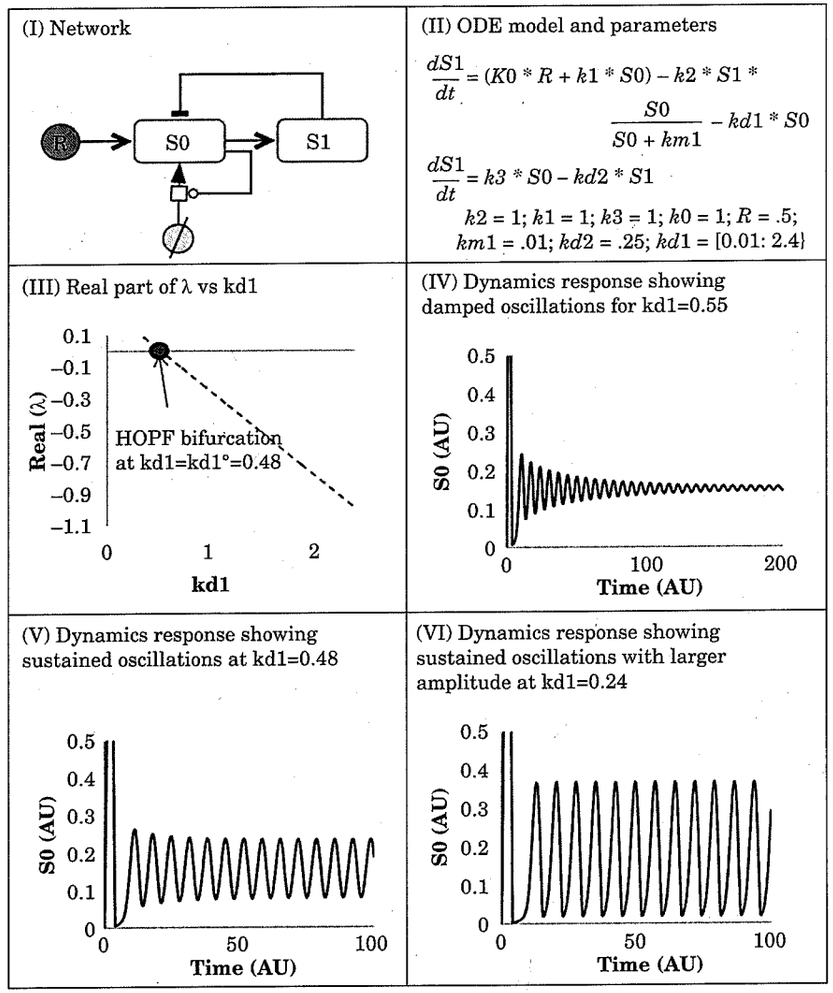


Fig. 15: Network showing interaction between two molecules S0 and S1, wherein, S0 undergoes activation due to stimulus R; S0 has auto-positive feedback and a negative feedback due S1; and S1 is activated by S0. (II) ODE modeling capturing above mentioned interactions (III) Plot of Eigenvalue vs kd1, wherein, kd1 is degradation rate constant of S0. (IV), (V) and (VI) are simulations results for kd1=0.55, 0.48 and 0.24, respectively.

System Identification

Often, known engineering principles are used to identify the key properties of the biological systems. The identification of the system is performed using, experimental and analytical techniques which involve application of the specific nature of inputs such as step change, wave,

impulse and ramp input. By application of dynamic analysis, the response obtained to these inputs are further analyzed to and obtain the response parameters such as, time constant, rise time, peak time, frequency response and amplitude response. Such kind of analysis enables to obtain the insights on the control structures prevailing in the biological networks.

Hierarchical regulation analysis

Apart from the above kinds of conventional analysis, biological systems need in depth analysis of the contributing networks. In hierarchical regulation analysis the contributions of various regulatory process are quantified based upon their molecular interaction networks and the output of the individual system (Westerhoff, 2006). It provides a platform for integrating the experimental data obtained at the various levels from metabolism, signaling and genetic analysis by various high-throughput methods. It becomes very essential to correlate the system outcome with various network levels to understand the mechanistic behavior that leads to certain phenotypes (Kuile and Westerhoffm, 2001; Even *et al.*, 2003; Rossell *et al.*, 2005, 2006). Most of the omics data can be analyzed using this approach and can be correlated to each other. This kind of analysis helps in dynamic analysis of the system with larger variations in the system parameters at various regulatory levels. By using hierarchical analysis we can understand how the information is processed at various network levels and precisely quantify the impact of each kind of network.

Model simulation

Simulation of biological system depends on the powerful numerical analysis methods to retrieve the solution of set of non-linear mathematical equations (ODEs, PDEs, stochastic, algebraic equations). Biological analysis softwares are publicly available for deterministic and stochastic simulations and for model analysis such as parameter estimation, parameter sensitivity analysis and bifurcation analysis. The details of the software are available in the www.sbml.org. Majority of them are graphical user interface (GUI) based modeling environment and provides an opportunity for user with limited computational and mathematical background to simulate the biological systems. Biological models can also be analyzed using general mathematical programming environment such as MATAB and MATHEMATICA (Ullah *et al.*, 2006; Schmidt *et al.*, 2007). The modeling environments also provide a provision to translate the developed mathematical models into Systems

Biology Markup Language (SBML), which facilitates the exchange of models among the modeling community.

Modeling tools

Here, we enlist useful software tools and platforms helpful for analyzing biological networks.

1. *Cell designer, (Funahashi, 2007, 2003)*: It is a platform used to build and visualize the biochemical networks. It is a tool that can be used for modeling and simulation aided with graphical user interface. Several kind of analysis can be performed with synchronization to SBML.
2. *COGRIM, (Clustering of Genes into Regulons using Integrated Modeling), (Chen et al., 2007)*: It allows clustering of genes into regulons using integrated modeling. It is based on Bayesian hierarchical model and Gibbs sampling implementation that integrates gene expression, ChIP binding, and transcription factor motif data.
3. *GeneACT, (Cheung et al., 2006)*: It provides detection of evolutionarily conserved transcription factor binding sites or microRNA target sites that are either unique or over-represented in differentially expressed genes from the DNA microarray data.
4. *BNArray,(Chen et al., 2006)*: It facilitates the construction of gene regulatory networks from DNA microarray data by using Bayesian networks.
5. *Jcell, (Spieth et al., 2006)*: It is Java-based framework for inferring regulatory networks from time series data.
6. *SynTReN, (den Bulcke and Moor, 2006)*: It is a generator of synthetic gene expression data for design and analysis of structure learning algorithms. It models different types of biological interactions and produces biologically plausible synthetic gene expression data.
7. *EXAMINE (EXpression Array MINing Engine), (Deng et al., 2005)*: It infers gene regulatory networks from time-series gene expression data sets.
8. *DBRF-MEGN method (Difference-Based Regulation Finding - Minimum Equivalent Gene Network), (Kyoda et al., 2004)*: It is implementation of an algorithm for deducing minimum equivalent gene networks from large-scale gene expression profiles of gene deletion mutants.

9. *CADLIVE (Computer-Aided Design of LIVing systEms), (Kurata et al., 2005)*: It allows construction of a large-scale biochemical network based on simulation-directed notation.
10. *BioTapestry Editor, (Longabaugh et al., 2005)*: It is Java-based interactive tool for building, visualization and simulation of genetic regulatory networks.
11. *Cytoscape,(Shannon et al., 2003)*: It is open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data. Various plugins in Cytoscape have also been developed *e.g.* BioNetBuilder (Avila-Campillo *et al.*, 2007), NetMatch (Ferro *et al.*, 2007), CABIN (Singhal and Domico, 2007).
12. *Gaggle, (Shannon et al., 2006)*: Is is an open-source software system for integrating bioinformatics software and data sources, for example, KEGG, BioCyc, String and software such as Cytoscape, Data Matrix Viewer, R statistical environment, and TIGR Microarray Expression Viewer.
13. *The Firegoose, (Bare, 2007)*: It provides two-way integration of diverse data from different bioinformatics web resources with desktop applications. The Firegoose is an extension to the Mozilla Firefox web browser and it enables data transfer between web sites and desktop tools.
14. *SiC (The Silicon Cells), (Snoep, 2009)*: It provides capability of computer simulations of biochemical networks in specific cells based on experimentally determined rate laws and parameter values.
15. *GEPASI, (Mendes, 1997)*: It is a software package for modeling biochemical systems. It simulates the kinetics of systems of biochemical reactions and provides a number of tools to fit models to data, optimize any function of the model, perform metabolic control analysis and linear stability analysis.
16. *Systems Biology Toolbox of MATLAB, (Schmidt and Jirstrand, 2005)*: The Systems Biology Toolbox for MATLAB offers an open and user extensible environment, wherein prototype, new algorithms, new applications for the analysis and simulation of biological systems can be developed and shared.
17. *SBML, (Evans, 2000)*: The Systems Biology Markup Language (SBML) is a computer-readable format for representing models of biochemical reaction networks. SBML is applicable to metabolic networks, cell-signaling pathways, regulatory networks. It simplifies the model sharing and integration across wide range of applications. Further, many different software have extended their environment to include SBML support.

SUMMARY

Understanding the design principles of biological network means a paradigm shift from evaluating interaction map of a network to quantifying the network interactions for gaining operational insights into cellular regulation. The process of modeling starts with precise definition of the problem that a researcher is going to address. The required data are gathered from various sources such as experiments, literature and omics databases. The pathways and interactions obtained from databases are reconstructed and integrated in the form of a network. The final network is developed that constitutes components from genetic, signaling and metabolic levels which are involved in the specific network. Based upon the acquired information suitable hypothesis are generated. To test this hypothesis, appropriate model development is essential. In absence of sufficient kinetic information, stoichiometric modeling is used to characterize the reaction fluxes which are based on steady state assumption and principle of mass balance. These methods are extensively applied for metabolic networks under the headings of Metabolic Flux analysis and Flux balance analysis. To study the dynamics of the system (time course evolution) kinetic modeling is preferred. Based upon the quality and quantity of the date available, kinetic modeling approaches such as deterministic modeling and stochastic modeling can be used. Parameter values are essential in obtaining quantitative match between simulation results and experimental data. These parameters are obtained from literature, experiments and estimates obtained through optimization techniques and algorithms. Models need to be analyzed for their stability and sensitivity. Models are subject to perturbation analysis by perturbing various ranges of the parameters. There are several model analysis techniques such as steady state analysis, dynamic analysis, sensitivity analysis, flux control analysis and bifurcation analysis. The models are simulated and the model predictions are validated with experimentally observed phenotypes and further refined to fine tune the predictability by the iterative processes.

Models provide an ideal platform to test the effect of concentration and operating parameters, to study network perturbation (*in-silico* mutation), to analyze the roles and contributions of different interactions, to predict the emergent properties of the network and to identify the missing information. Thus, models assist in system analysis, hypothesis generation and testing, experimental data validation and optimal product design.

REFERENCES

Albert Goldbeter, (1996). Biochemical oscillators and cellular rhythms.

Aldridge, B.B., Burke, J.M., Lauffenburger, D.A. and Sorger, P.K. (2006). Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.*, 8(11): 1195–1203.

Asthagiri, A.R. and Lauffenburger, D.A. (2000). Bioengineering models of cell signaling. *Annual Review of Biomedical Engineering*, 2(1): 31–53.

Bailey, J.E. (1998). Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress*, 14(1): 8–20.

Balla, U.S. (2003). Understanding complex signaling networks through models and metaphors. *Progress in Biophysics and Molecular Biology*, 81(1): 45–65.

Bare, J.C., Shannon, P.T., Schmid, A.K. and Baliga, N.S. (2007). The Firegoose: Two-way integration of diverse data from different bioinformatics web resources with desktop applications. *BMC Bioinformatics*, 8: 456.

Breitling, R., Gilbert, D., Heiner, M. and Orton, R. (2008). A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. *Brief Bioinform*, 9(5): 404–421

C.G. Moles, P.M., and Banga, J.R. (2003). Parameter estimation in biochemical pathways: A comparison of global optimization methods. *Genome Research*, 13(11): 2467–2474.

Cao, H., Romero-Campero, F., Heeb, S., Cámara, M. and Krasnogor, N. (2010). Evolving cell models for systems and synthetic biology. *Syst. Synth. Biol.*, 4(1): 55–84.

Cardelli, L. (2005). Brane calculi. Computational methods in systems biology. Springer, Berlin Heidelberg, pp. 257–278.

Chaves, M., Sontag, E.D. and Albert, R. (2006). Methods of robustness analysis for Boolean models of gene control networks. *IEE Proceedings - Systems Biology*, 153(4): 154–167.

Chen, G., Jensen, S. and Stoeckert, C. (2007). Clustering of genes into regulons using integrated modeling-COGRIM. *Genome Biology*, 8(1): R4.

Chen, X., Chen, M. and Ning, K. (2006). BNArray: An R package for constructing gene regulatory networks from microarray data by using Bayesian network. *Bioinformatics*, 22(23): 2952–2954.

Cheung, T., Kwan, Y Hamady., M. and Liu, X. (2006). Unraveling transcriptional control and cis-regulatory codes using the software suite GeneACT. *Genome Biology*, 7(10): R97.

Chickarmane, V., Kholodenko, B.N. and Sauro, H.M. (2007). Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. *Journal of Theoretical Biology*, 244(1): 68–76.

Cho, K.H., Shin, S.Y., Kolch, W. and Wolkenhauer, O. (2003). Experimental design in systems biology, based on parameter sensitivity analysis using a monte carlo method: A case study for the TNFα-mediated NF-κ β signal transduction pathway. *Simulation*, 79(12): 726–739.

Colvin, J., Monine, M., Gutenkunst, R., Hlavacek, W., Von Hoff, D. and Posner, R. (2010). RuleMonkey: Software for stochastic simulation of rule-based models. *BMC Bioinf.*, 11(1): 404.

Daniel Machado, Rafael S Costa, Miguel Rocha, Eugénio C Ferreira, Bruce Tidor and Isabel Rocha, (2011). Modeling formalisms in systems biology. *AMB Express*, 1: 45.

David, S.W., Robert, Y., David, A., Peter, T. and Joseph, C. (2005). Dynamic cellular automata: An alternative approach to cellular simulation. *In Silico Biology*, 5(2): 139–161.

de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1): 67–103.

den Bulcke, T.V., Leemput, K.V., Naudts, B., Remortel, P.V., Ma, H., Verschoren, A., and B.D.M. Moor, K. (2006)). SynTReN: A generator of synthetic gene expression data for design and analysis of structure learning algorithms. *BMC Bioinformatics*, 7(43).

Deng, X., Geng, H. and Ali, H. (2005). EXAMINE: A computational approach to reconstructing gene regulatory networks. *Biosystems*, 81(2): 125–136.

Doedel, E.J., H.B.K., and Kernevez, J.P. (1991). Numerical analysis and control of bifurcation problems (I) bifurcation in finite dimensions. *International Journal of Bifurcation and Chaos*, 1(3): 493–520.

Don Kulasiri, L.K.N., Samarasinghe, Sandhya and Xie, Zhi (2008). A review of systems biology perspective on genetic regulatory networks with examples. *Current Bioinformatics*, 3: 197–225.

Dr. Shrikant, V. and Savant, T.M. Inc. (2007). Stochastic simulation techniques in systems. Biology Proceedings of the 2007 American Control Conference, New York City, USA, .

El-Samad, H. and Khammash, M. (2010). Modelling and analysis of gene regulatory network using feedback control theory. *International Journal of Systems Science*, 41(1): 17–33.

Eungdamrong, N.J. and Iyengar, R. (2004). Computational approaches for modeling regulatory cellular networks. *Trends in Cell Biology*, 14(12): 661–669.

Evans, G.A. (2000). Designer science and the omic revolution. *Nat. Biotech.*, 18(2): 127–127.

Even, S., Lindley, N.D. and Cocaign-Bousquet, M. (2003). Transcriptional, translational and metabolic regulation of glycolysis in *Lactococcus lactis* subsp. *cremoris* MG 1363 grown in continuous acidic cultures. *Microbiology*, 149(7): 1935–1944.

Fauré, A., Naldi, A., Chaouiya, C. and Thieffry, D. (2006). Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, 22(14): e124–e131.

Fei-Yue, W. (2005). Agent-based control strategies for smart and safe vehicles. Vehicular Electronics and Safety, 2005. IEEE International Conference on.

Funahashi, A., Jouraku, A., Matsuoka, Y. and Kitano, H. (2007). Integration of Cell Designer and SABIO-RK. *In Silico Biology*, 7: S81–S90.

Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N. and Kitano, H. (2005). *Proceedings of the IEEE*, 96(8): 1254–1265

Funahashi, A., Tanimura, N., Morohashi, M. and Kitano, H. (2003). Cell designer: A process diagram editor for gene-regulatory and biochemical networks. *Biosilico*, 1: 159–162.

Garcia-Ojalvo, J., Elowitz, M.B. and Strogatz, S.H. (2004). Modeling a synthetic multicellular clock: Repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30): 10955–10960.

Gardner, T.S., Cantor, C.R. and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767): 339–342.

Gautam Goel, I-Chun Chou, and Eberhard O. Voit (2006). Biological systems modeling and analysis: A biomolecular technique of the twenty-first century. *Journal of Biomolecular Techniques*, 17: 252–269.

Gilbert, D., Fuß, H., Gu, X., Orton, R., Robinson, S., Vyshemirsky, V., Kurth, M.J., Downes, C.S. and Dubitzky, W. (2006). Computational methodologies for modelling, analysis and simulation of signalling networks. *Briefings in Bioinformatics*, 7(4): 339–353.

Gillespie, D. (2008). Simulation methods in systems biology. Formal methods for computational systems biology. Bernardo, M., Degano, P. and Zavattaro, G. Springer Berlin / Heidelberg. 5016: 125–167.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25): 2340–2361.

Glass, L. and Kauffman, S.A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *Journal of Theoretical Biology*, 39(1): 103–129.

Gordon, B., Melania, R.A., Eugene, C. and Michael, J.E. (2005). A life-like virtual cell membrane using discrete automata. *In Silico Biology*, 5(2): 163–178.

Haseltine, E.L. and Arnold, F.H. (2007). Synthetic gene circuits: Design with directed evolution. *Annual Review of Biophysics and Biomolecular Structure*, 36(1): 1–19.

Iba, H. (2008). Inference of differential equation models by genetic programming. *Inf. Sci.*, 178(23): 4453–4468.

Ideker, T. (2004). Systems biology 101[mdash] what you need to know. *Nat. Biotech.*, 22(4): 473–475.

Ideker, T., Galitski, T. and Hood, L. (2001). A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics*, 2(1): 343–372.

Imoto, S., Goto, T. and Miyano, S. (2002). Estimation of genetic networks and functional structures between genes by using Bayesian networks and nonparametric regression. Pacific Symposium on Biocomputing, pp. 175–186.

J. Hasty, M.D., Rottschafer, V. and Collins, J.J. (2002). Synthetic gene network for entraining and amplifying cellular oscillations. *Physical Review Letters*, 88(14): Article ID 148101, 148104.

James E. Ferrell, Tony Yu-Chen Tsai and Qiong Yang (2011). Modeling the cell cycle: Why do certain circuits oscillate? *Cell*, (2011): 144.

Jamshidi, N. and Palsson, B. (2010). Mass action stoichiometric simulation models: Incorporating kinetics and regulation into stoichiometric models. *Biophys. J.*, 98: 175–185

Jansson, A. and Jirstrand, M. (2010). Biochemical modeling with systems biology graphical notation. *Drug Discovery Today*, 15(9–10): 365–370.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1): 27–30.

Kapuy, O., Barik, D., Domingo Sananes, M.R., Tyson, J.J. and Novak, B. (2009). Bistability by multiple phosphorylation of regulatory proteins. *Prog. Biophys. Mol. Biol.*, 100: 47–56.

Karlebach, G. and Shamir, R. (2008). Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.*, 9(10): 770–780.

Kauffman, K., Prakash, P. and Edwards, J. (2003). Advances in flux balance analysis. *Curr. Opin. Biotechnol.*, 14(5): 491–496.

Keasling, J.D. (2008). Synthetic biology for synthetic chemistry. *ACS Chemical Biology*, 3(1): 64–76.

Khanin, R. and Higham, D.J. (2008). Chemical master equation and Langevin regimes for a gene transcription model. *Theoretical Computer Science*, 408(1): 31–40.

Kholodenko, B.N. (2006). Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.*, 7(3): 165–176.

Kitano, H. (2000). Perspectives on systems biology. *New Generation Computing*, 18(3): 199–216.

Kitano, H. (2002). Computational systems biology. *Nature*, 420(6912): 206–210.

Kitano, H. (2002). Systems biology: A brief overview. *Science*, 295(5560): 1662–1664.

Koshland, D.E.Jr. (1998). Biochemistry: Enhanced: The era of pathway quantification. *Science*, 280(5365): 852–853.

Kurata, H., Masaki, K., Sumida, Y. and Iwasaki, R. (2005). CADLIVE dynamic simulator: Direct link of biochemical networks to dynamic models. *Genome Research*, 15(4): 590–600.

Kyoda, K., Baba, K., Onami, S. and Kitano, H. (2004). DBRF–MEGN method: An algorithm for deducing minimum equivalent gene networks from large-scale gene expression profiles of gene deletion mutants. *Bioinformatics*, 20(16): 2662–2675.

Lee, J., Gianchandani, E., Eddy, J. and Papin, J. (2008). Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput. Biol.*, 4(5): e1000086

Longabaugh, W.J.R., Davidson, E.H. and Bolouri, H. (2005). Computational representation of developmental genetic regulatory networks. *Developmental Biology*, 283(1): 1–16.

Markschies, C. (2008). Mathematical modeling and kinetic analysis of cellular signaling pathways.

McAdams, H.H. and Shapiro, L. (1995). Circuit simulation of genetic networks. *Science*, 269(5224): 650–656.

Mendes, P. (1997). Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. *Trends in Biochemical Sciences*, 22(9): 361–363.

Noble, D. (2002). The rise of computational biology. *Nat. Rev. Mol. Cell Biol.*, 3(6):459–463.

Novere, N.L., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E. and Wegner, K.M. (2009). The systems biology graphical notation. *Nat. Biotech.*, 27(8): 735–741.

Papin, J.A. and Palsson, B.O. (2004). Topological analysis of mass-balanced signaling networks: A framework to obtain network properties including crosstalk. *Journal of Theoretical Biology*, 227(2): 283–297.

Polynikis, A., Hogan, S.J. and di Bernardo, M. (2009). Comparing different ODE modelling approaches for gene regulatory networks. *Journal of Theoretical Biology*, 261(4): 511–530.

Ramakrishna, R., Edwards, J.S., McCulloch, A. and Palsson, B.O. (2001). Flux-balance analysis of mitochondrial energy metabolism: Consequences of systemic stoichiometric constraints. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 280: R695–704.

Ridwan, A.M., Krishnan, A. and Dhar, P. (2004). A parallel Implementation of Gillespie's direct method. Computational science - ICCS 2004. Bubak, M., v. Albada, G.D., Sloot, P.M.A. and Dongarra, J.J. Springer Berlin / Heidelberg. 3037: 284–291.

Rossell, S., van der Weijden, C.C., Kruckeberg, A.L., Bakker, B.M. and Westerhoff, H.V. (2005). Hierarchical and metabolic regulation of glucose influx in starved *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 5(6–7): 611–619.

Rossell, S., van der Weijden, C.C., Lindenbergh, A., van Tuijl, A., Francke, C., Bakker, B.M. and Westerhoff, H.V. (2006). Unraveling the complexity of flux regulation: A new method demonstrated for nutrient starvation in *Saccharomyces*

cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7): 2166–2171.

Sauro, H.M. and Kholodenko, B.N. (2004). Quantitative analysis of signaling networks. *Progress in Biophysics and Molecular Biology*, 86(1): 5–43.

Schauer, M. and Heinrich, R. (1983). Quasi-steady-state approximation in the mathematical modeling of biochemical reaction networks. *Mathematical Biosciences*, 65(2): 155–170.

Schmidt, H. and Jirstrand, M. (2005). Systems biology toolbox for MATLAB: A computational platform for research in systems biology. *Bioinformatics*, 22(4): 514–515.

Schmidt, H., Drews, G., Vera, J. and Wolkenhauer, O. (2007). SBML export interface for the systems biology toolbox for MATLAB. *Bioinformatics,* 23(10): 1297–1298.

Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.,* 13(11): 2498–2504.

Shannon, P., Reiss, D., Bonneau, R. and Baliga, N. (2006). The Gaggle: An open-source software system for integrating bioinformatics software and data sources. *BMC Bioinformatics*, 7(1): 176.

Snoep, J.L.W. (2009). From isolation to integration, a systems biology approach for building the silicon cell. *In* Systems Biology - Definitions and Perspectives, pp. 13–30. Springer-Verlag Berlin.

Spieth, C., Supper, J., Streichert, F., Speer, N. and Zell, A. (2006). J. Cell—a Java-based framework for inferring regulatory networks from time series data. *Bioinformatics,* 22(16): 2051–2052.

Sriram, Y.Z. a. G. (2010). Mathematical modeling: Bridging the gap between concept and realization in synthetic biology. *Journal of Biomedicine and Biotechnology*, (Article ID 541609): 16.

Steuer, R. and Junker, B. (2008) Computational models of metabolism: Stability and regulation in metabolic networks. John Wiley & Sons, Inc142

Sumner, T., Shephard, E. and Bogle, I.D.L. (2012). A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modeling. *J. R. Soc. Interface*, 9(74): 2156–2166.

ter Kuile, B.H. and Westerhoff, H.V. (2001). Transcriptome meets metabolome: Hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters*, 500(3): 169–171.

Tyson, J.J. (1996). About time. A review of A. Goldbeter's biochemical oscillations and cellular rhythms: The Molecular Bases of Periodic and Chaotic Behaviour. *Nature,* 380: 213–214.

Tyson, J.J. and Kagan, M.L. (1988). Spatiotemporal organization in biological and chemical systems. *In*: From Chemical to Biological Organization (Markus, M., Mudi.ller S.C. and Nicolis, G. *Eds.*) pp. 14–21 (Springer-Verlag, Berlin).

Ullah, M., Schmidt, H., Cho, K.H. and Wolkenhauer, O. (2006). Deterministic modelling and stochastic simulation of biochemical pathways using MATLAB. *IEE Proceedings - Systems Biology*, 153(2): 53–60.

Vinod, P.K. and Venkatesh, K.V. (2008). Quantification of signaling networks. *Journal of The Indian Institute of Science*, 88(1): 1–26.

Westerhoff, F.J.B.H.V. (2006). Approaches to biosimulation of cellular processes. *J. Biol. Phys.,* 32: 273–288.

Wilkinson, D.J. (2009). Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat. Rev. Genet.,* 10(2): 122–133.

Wu, J. and Voit, E. (2009). Integrative biological systems modeling: Challenges and opportunities. *Front Comput. Sci. Chin.,* 3(1): 92–100.

Zambonelli, F. and Omicini, A. (2004). Challenges and research directions in agent-oriented software engineering. *Autonomous Agents and Multi-Agent Systems*, 9(3): 253–283.

Zumsande, M. and Gross, T. (2010). Bifurcations and chaos in the MAPK signaling cascade. *Journal of Theoretical Biology*, 265(3): 481–491.