Computational Modeling Approaches for Studying of Synthetic Biological Networks

Elizabeth Pham[1], Isaac Li[1] and Kevin Truong[*1,2,]

[1]*Institute of Biomaterials and Biomedical Engineering, University of Toronto, 164 College Street, Toronto, Ontario, M5S 3G9, Canada;* [2]*Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Circle, Toronto, Ontario, M5S 3G4, Canada*

Abstract: Synthetic biology is an emerging field that strives to build increasingly complex biological networks through the integration of molecular biology and engineering. The growth of the field has been supported by progress in the design and construction of synthetic genetic and protein networks. This has led to the possibility of assembling modular components to attain novel biological functions and tools. In addition, these synthetic networks give rise to insights that facilitate the investigation of interactions and phenomena in naturally-occurring networks. Integration of well-characterized biological components into higher order networks requires computational modeling approaches to rationally construct systems that are directed towards a desired outcome. A computational approach would improve the predictability of the underlying mechanisms that would otherwise be difficult to deduce through experimentation alone. The analysis and interpretation of both qualitative and quantitative models also becomes increasingly important towards taking a systems-level perspective on synthetic genetic and protein networks. This review will first discuss the analogy of synthetic networks to circuit engineering. It will then look at computational modeling approaches that can be applied to biological systems and how synthetic biology will help to develop more accurate *in silico* representations of these systems.

Keywords: Synthetic biology, genetic circuits, protein circuits, computational modeling, networks, biological complexity.

INTRODUCTION

Synthetic biology encompasses an engineering-based approach to designing biological networks. It shares the holistic perspective of systems biology as its ultimate goal is to construct *de novo* networks of high complexity and interconnectivity. Systems biology aims at understanding systems as a whole by studying the interactions between the components of biological networks. One goal of systems biology is to provide an in-depth comprehensive body of knowledge of the interactions and kinetics governing biological systems at the molecular level. Progress in synthetic biology will address fundamental principles of these biological interactions, as well as lead to practical applications in drug discovery and biotechnology. In order to move towards a higher-order, systems-level perspective, it is necessary to examine the composition, structure, and kinetics of cellular networks, rather than the characteristics of the isolated parts alone.

An important post-genomic research area is the analysis and elucidation of the dynamic interactions of genes and proteins in naturally-occurring systems. The massive amount of data generated from genomic sequencing has led to research in the '-omics', including functional genomics and proteomics. Linkages between the molecular and system levels were recently made possible by advances in these areas. The current drive is to analyze systems in terms of their responses to perturbations and to uncover network features such as robustness and degeneracy.

*Address correspondence to this author at the Institute of Biomaterials and Biomedical Engineering, University of Toronto, 164 College Street, Toronto, Ontario, M5S 3G9, Canada; Tel: (416) 978-7772; Fax: (416) 978-4317; Email: kevin.truong@utoronto.ca

This is where synthetic biology can be incorporated to achieve systems-level analysis. Molecular characteristics of biological interactions have been identified and categorized into specific functional modules. A systematic means of piecing together different modules to progressively build more complex networks will not only lead to a systems-level understanding, but also reveal the underlying kinetics governing how the individual modules interact and respond to each other. This results in a more continuous stream of knowledge, from molecular to modular to systems descriptions. Not only is there a gap in our understanding and knowledge of all biological phenomena, even for biological systems in which all the components are known, it is still unclear precisely how these components interact to make cellular processes work. It is important to note, however, that connecting modules into more complex networks will still not lead to a bottom-up reconstruction of a complete organism. The use of synthetic modules will assist in understanding how subnetworks may interact and crosstalk, but emergent properties of an organism arise due to inherent nonlinearities of complexity and not only by the accumulation of modular properties.

Natural biological systems include gene regulatory networks, protein signaling cascades, and metabolic pathways. These complex systems are both structurally and functionally diverse, with multifunctional sets of elements that interact selectively and nonlinearly, yet able to produce highly specific behaviours. The vast amount of biological data from molecular biology has revealed many sequences and properties of genes and proteins, but is not sufficient for interpreting system behaviour.

Recently, computational modeling approaches have been employed to study natural biological systems and would be

1574-8936/08 $55.00+.00

applicable, in fact highly recommended, for synthetic networks. These approaches integrate advances in algorithms and statistics to analyze biological data. Through a combination of both experimental and computational approaches, we can gain deeper understanding of the function of biological processes. Therefore, it is worthwhile to look at how computational approaches could be used to complement construction and experimentation of synthetic networks. (Fig. **1**)

CIRCUIT ENGINEERING ANALOGY

Biological networks can often be treated as electrical circuits. It is thus useful to adopt a vocabulary from circuit engineering theory to describe concepts such as input, output, control, logic, feedback, amplification, adaptation, and robustness. Both circuits and biological systems transform information from one form into another based on a set of defined rules. Stimuli function as inputs, while signals modulating the behaviour of the system are processed and generated as outputs by some established protocol inherently defined within the network itself.

Mapping out the components of a biological system and making the connections between interacting components can be analogous to drawing a circuit diagram. In order to deduce the mechanisms controlling these biological circuits, a parts list needs to be generated and the transformation between input and output needs to be established. While the former has been successful through genomic sequencing and protein studies, the latter requires more rigorous analysis. Building a circuit to perform a particular function is much easier than deducing the function of an existing black-box circuit solely through correlating its outputs with its inputs.

Circuit control theory has been used to develop a theoretical understanding of an adaptation mechanism through negative feedback [1, 2]. However, this approach has limitations as control theory assumes that inputs are provided to the system, but in biology, such inputs or stimuli are often created and refined continuously within the system itself. In another study, an integrative modeling approach was used to run a circuit simulation of the lysis-lysogeny decision circuit of bacteriophage lambda, making use of the parallels between genetic and electrical circuits [3]. Similarly, other frameworks integrating control theory and biological control processes have been proposed to describe genetic regulatory networks and adaptation in bacterial chemotaxis [2, 4].

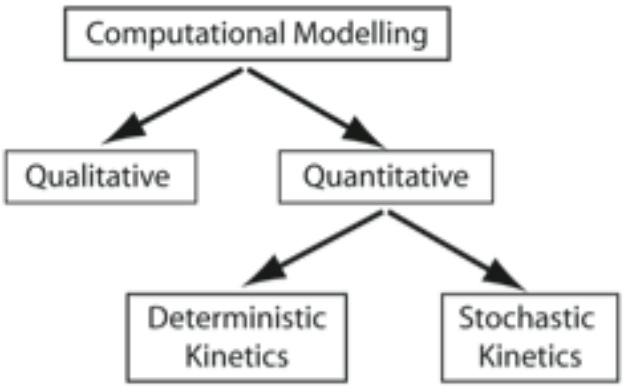


Fig. (1). Global picture of computational modeling approaches that can be used for studying synthetic networks.

These function well as system descriptions, drawing parallels between biological processes with more established control theory. While such an analogy allows for a framework in which to systematically identify and analyze synthetic biological networks, it does not address the need to computationally study these networks for a more quantitative perspective.

Beyond some of the borrowed terminology and circuit theory, this is where the circuit engineering analogy starts breaking down. While in some scenarios biological networks can be treated as if they were circuits, there are no counterparts for many electrical circuit features and established rules such as Kirchoff's laws. Even for a large complex electrical circuit, prior knowledge of the properties and characteristics of the individual components is used to build the larger circuit. This prior knowledge allows every circuit to be accompanied by a set of equations reliably describing its function and behaviour. Such an unambiguous understanding does not exist for all biological systems. Even for well-studied systems, no set of defined equations or approximations correspond exactly to how that system behaves. In many cases, even the smaller components making up a biological system are still under rigorous study.

SYNTHETIC NETWORKS CONSTRUCTED

The rational construction and analysis of synthetic networks provides a framework for computational modeling studies. Using the analogy of logic flow from circuit engineering, both genetic-based and protein-based synthetic networks have been designed and tested [5]. A library of networks with novel connectivities between transcriptional regulators and the corresponding promoters was previously developed for combinatorial synthesis of biological networks of varying levels of complexity [5]. Examples of synthetically engineered gene circuits include autoregulatory systems displaying stability through negative feedback [6], toggle switches [7], logic gates, and repressilators [8]. (Fig. **2**) Similarly, engineered protein circuits have been constructed to function as Boolean logic gates of AND, OR, and NOT [9, 10]. (Fig. **3**)

It has been suggested that the study of biological systems is moving towards modular biology [11]. Biology is moving away from the molecular perspective of correlating functions with individual proteins and subsequently those with individual genes, to a more modular perspective of analyzing how those proteins and genes interact to produce a higher function. Cellular behaviour is carried out and regulated by 'modules' that are themselves made up of many species of interacting molecules, ranging from nucleotides to the many proteins coded by genetic sequences. Modules have evolved to perform specific functions, much like electrical circuits have been engineered for specific purposes. These modules can be classified by function, such as genetic switches, flip-flops, logic gates, amplifiers and oscillators. Another means of classification is to define 'network motifs' to represent interconnections that are more commonly found, such as feed-forward loops, single-input modules (SIM), and dense overlapping regulons (DOR) [12]. General principles and mechanisms that govern the behaviour and structure of modules can be elucidated through studies of synthetic networks, with the help of molecular engineering and computer sci-

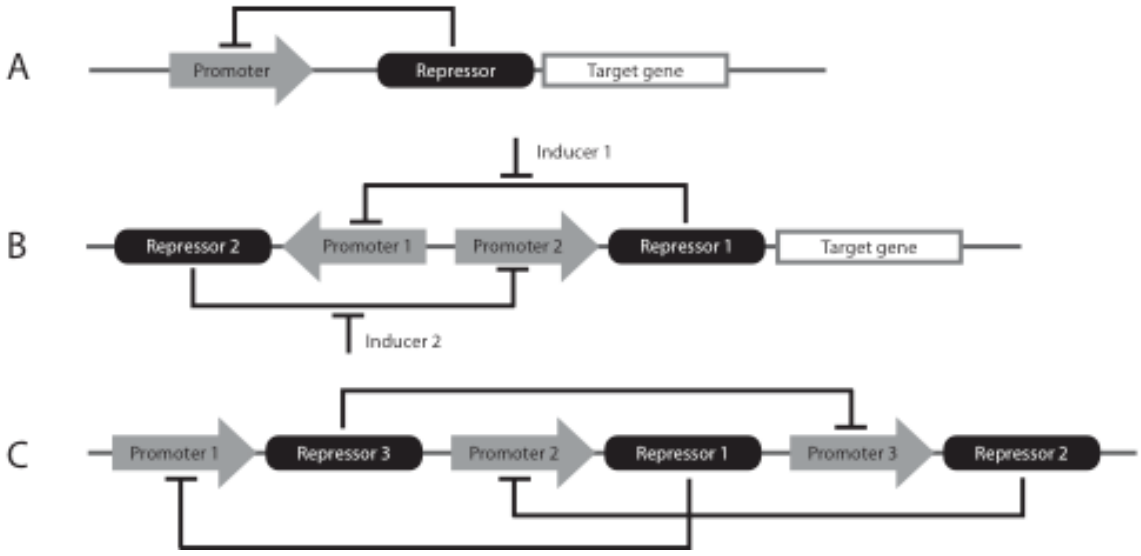


Fig. (2). Examples of engineered synthetic genetic circuits. A) A negative-feedback circuit consists of a promoter that drives the expression of both a target gene and a repressor gene that inhibits expression of its own promoter. B) a toggle switch circuit regulates the expression of a target gene by inducer molecules. Promoter 1 drives the expression of Repressor 2; Promoter 2, both Repressor 1 and the target gene. Repressor 1 inhibits expression of Promoter 1; Repressor 2, Promoter 2. When an inducer molecule blocks the expression of Repressor 2, the inhibition of Promoter 2 is released and the target gene is expressed; conversely, when an inducer molecule blocks Repressor 1, the target gene is not expressed. C) an oscillator circuit: three repressor proteins forma cyclic negative feedback loop, each inhibiting the expression of another repressor protein; when Promoter 1 drives the expression of Repressor 3, Promoter 3 is blocked, causing an accumulation of Repressor 1; this accumulation inhibits the expression of Repressor 3, which then causes an accumulation of Repressor 2 and inhibition of Repressor 1; when the expression of Repressor 1 is blocked, this causes an accumulation of Repressor 3 and the cycle repeats, resulting in oscillating concentrations of the three repressor proteins.

ence. However, looking only at modular components of biological networks is still insufficient for understanding the system itself.

To head towards a systems-level analysis, computational modeling approaches become even more important. Reductionism has been a dominant approach to studying biology, reducing a system into the components and attempting to reconnect those components through assumptions and approximations. However, a larger issue that cannot be addressed by reductionism is the lack in understanding of the dynamic and nonlinear behaviour of the systems, which can only be obtained by taking a holistic approach. Based on *in silico* prediction and optimization from computational models, more complex circuits can be rationally assembled from subnetworks. These larger circuits can then be used for further experimental study and implementation.

THE NEED FOR COMPUTATIONAL APPROACHES

A computational model is needed to assemble the wealth of data together in order to predict network function and be-

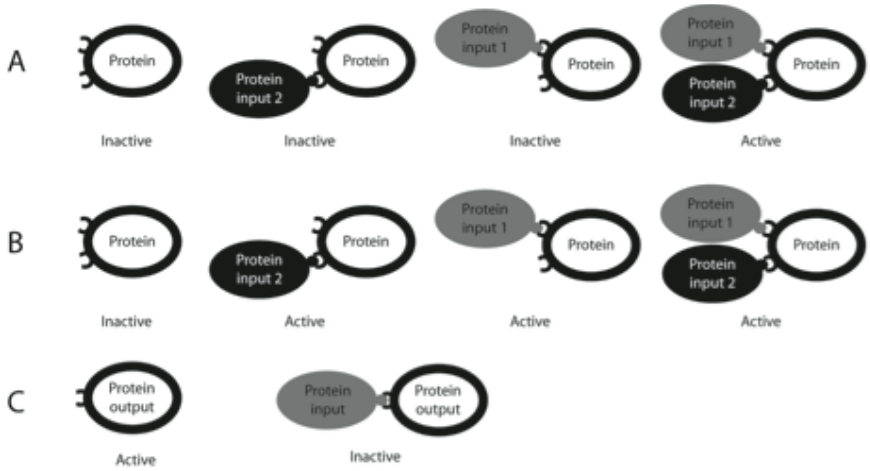


Fig. (3).Examples of engineered synthetic protein circuits. A) An AND gate: the protein (white) is active only in the presence of both input proteins, B) an OR gate: the protein is active when one or both input proteins are present, and C) a NOT gate: the protein is no longer active in the presence of an input protein.

haviour, predictions that can subsequently be tested experimentally. Simulations of genetic networks have revealed behaviours that were not apparent from studies of the isolated interactions alone. A systematic analysis of four simple signaling pathways determined that the integrated network of cross-talking pathways showed properties that were not apparent in any of the individual pathways [12, 13]. These results show promise in using computational means to elucidate system properties and behaviours.

Computational models can provide effective descriptions of biological networks at different levels of resolution. In many biological networks, molecular species exist in very low copy numbers despite huge gene numbers, making it difficult to detect these small populations of proteins and mRNA even with current high-throughput techniques such as microarrays. Thus, well defined models and simulations will significantly assist in providing insights for constructing and analyzing synthetic networks.

QUALITATIVE COMPUTATIONAL MODELING APPROACHES

Many attempts thus far have focused on mapping and causally modeling the different components of biological networks. Hypotheses are then proposed to describe the system behaviour. Using the logical binary approach, also known as the Boolean network model, on and off states have been used to describe the state of genes and proteins in a circuit [14, 15]. In an attempt to screen active networks that show significant changes in gene expression under specific conditions, one group tested an algorithm that combined a statistical scoring measure with mRNA expression data to uncover potential regulatory pathways in a systematic and integrative fashion [16]. As a proof of concept, high scoring networks were compared to known natural regulatory networks in literature. Similarly, another group causally represented regulatory connections between different genes as weights so that the net effect of regulation on a particular gene expression was calculated by summing the weights of its inputs [17].

Early characterization of any network is establishing the elements involved, the connectivity between the elements, and any associated parameters controlling the interactions. While a qualitative mapping of the elements in a system provides a visual representation and suggests directionality, in order to allow predictive power to the design of synthetic networks, these hypotheses must be tested against quantitative models. Because a biological system is not simply the sum of all its parts, even the most accurate diagram of its interconnections will not fully reveal the kinetics of that system without quantitative considerations. (Fig. **4-A**)

QUANTITATIVE COMPUTATIONAL MODELING APPROACHES

Conventional methods of creating network models involves performing a series of experiments, identifying specific interactions, conducting extensive literature research for confirmation, and repeating. Several methods are available to reveal regulatory relationships based solely on mRNA expression data from microarray studies. Microarray analysis is superior because it is automated and a large amount of data can be found in parallel. However, the many mechanisms occurring in a single system including post-transcriptional and post-translational modifications cannot be incorporated all at once on a microarray without losing precision and accuracy. An alternative means of incorporating many different mechanisms that occur simultaneously in a system is through algorithms for *in silico* experiments and computational modeling.

Quantitative computational approaches consist of two distinct aspects, namely data mining and simulation. Data mining attempts to extract hidden patterns from huge quantities of experimental data in order to form hypotheses. Such data mining has been used extensively in bioinformatics to predict exon-intron regions and protein structures, as well as gene regulatory networks from expression profiles [16]. Here, statistical correlations between expression levels are inferred to determine likely interactions. Alternatively, simulation-based analysis tests hypotheses with *in silico* experiments, these predictions can then be tested by *in vitro* and *in vivo* studies. Simulations predict the kinetics of systems, incorporating assumptions and approximations to complete the models. They are generally based on heuristics and statistical considerations, the validity of which can be tested experimentally. Simulations often require the integration of multiple hierarchies of models that span several orders of magnitude in terms of scale, abundance, binding affinities, and rate constants [18]. Iterative comparisons between experimental observations and computational models will reveal inconsistencies that need to be dealt with. Assumptions and approximations can then be refined to better correlate generated models with experimental results.

Current advances in high-throughput experimentation produce the large amount of quantitative data needed to support simulation-based studies. Advances in software and computational power have allowed for more realistic, complex biological models including those for bifurcation of the cell cycle, metabolic analysis, and oscillatory circuits [19-22] (Table **1**).

Due to the complexity of biological systems, computer simulations and heuristics are often used as part of experimental research methods to determine relationships between inputs to outputs of biological networks, both natural and synthetic. The choice of what is to be modeled depends on the availability of biological knowledge that can be incorporated into the model without adding more unknowns and variability. Computational modeling approaches will help to predict the underlying mechanisms of these networks, predictions which can then be supported or eliminated through experimentation. To attain a systems-level understanding, the question of *what to look for* is difficult to resolve. In many cases, it is challenging to intuitively predict how a system will behave and how it will respond to perturbations.

Abstract models of biological systems can be developed based on general hypotheses. While this may be useful for obtaining new insights into biological processes and allows the classification of different systems under broad headings, abstract and intuitive reasoning alone is not sufficient to handle the complexity of biological networks. Intuition is not enough for constructing functional synthetic networks and so the design of synthetic networks needs to be accompanied and guided by computational models. Even then, this is a daunting task due to the high degree of complexity in bio-

Table 1. Software Tools Developed for the Modeling and Simulation of Biological Interactions

Tools Available	Description	Source
BioJake	Visualization tool for the manipulation of metabolic pathways	[23]
BioSPICE	Software system for access to current computational tools	[24, 25]
CellDesigner	Software for diagrammatic editing of biological networks	[26]
CellWare	Integrative multi-algorithmic simulation tool for deterministic and stochastic cellular events	[27, 28]
COPASI	Platform-independent tool for the simulation of biochemical events	[29]
Dizzy	Software tool for modeling integrated large-scale networks deterministically and stochastically	[30]
Dynetica	Simulation tool for studying kinetic models of dynamic networks	[31]
E-CELL	Software environment for simulation of integrative models of cellular behaviour	[32]
Gepasi	Software system for modeling chemical and biochemical reaction networks	[33-35]
Pathway Tools	Software environment for creating model-organism databases	[36]
StochSim	Stochastic simulation tool for chemical reactions	[37]
STOCKS	Stochastic kinetic simulation tool for biochemical processes	[38]
Systems Biology Work-bench (SBW)	Software framework for communication between software applications	[39]
Virtual Cell	Computational framework for modeling and testing biological networks	[40, 41]

logical systems. Not only are experiments difficult to carry out on complex systems, but it is also[1] challenging to obtain an accurate representation through simulation. This point again emphasizes the need to closely study simpler components and subsystems before making that leap to studying entire complex systems. Once the components and subsystems are characterized and models verified with experimental data, a systematic means can be used to progressively build these simpler synthetic networks into one of higher complexity with the aid of computational modeling approaches.

Deterministic Chemical Kinetics Approach

Chemical kinetic models represent a biological process as a system of chemical reactions. (Fig. **4-B)** As such, concentrations of each molecular species involved in that process defines the state of the system at any given time. The interactions between molecular species are simplified into chemical reactions, where reactions occurring would alter the population of molecular species. Each potential chemical reaction is represented by a differential equation involving reaction rates, and other relevant reaction parameters such as binding affinities. (Fig. **4-C)** A time course of the concentration changes of the molecular species is simulated in order to produce a system transition path. This deterministic approach assumes a predictable process governed by a set of differential equations, namely reaction-rate equations [42]. Networks representing bacterial chemotaxis and bacteriophage infections have been modeled using this chemical kinetics approach and have been verified experimentally through quantitative kinetics studies [3, 12, 43-45].

Biological modeling tools do not provide the same precision as tools established for circuit engineering. This lack of predictive power means that many approximations are incorporated into quantitative models, depending on how much prior knowledge is available. One of the complications presented by biological systems is varying kinetic rate scales and binding affinities. Starting with mapped connectivities from qualitative modeling of small subnetworks, it is possible to form and test algorithms for inferring network relevant parameters by quantifying molecular responses to a given perturbation. However, emerging technologies to facilitate these studies face the challenge that the larger the quantity of data generated, often the lower the quality and resolution. As well, parameters are often underdetermined, that is, the number of unknown parameters outnumbers those already established from previous experimental work. Often, many parameters need to be determined by fitting specific experimental observations using trial and error. Employing this brute force approach, one group followed the time courses of major cyclin-dependent kinase activities in budding yeast cell cycles [20]. Quantitative simulations were then required to predict relevant parameters. However, these predictions may give rise to multiple solutions that must be further discerned by more detailed studies to eliminate improbable solutions and distinguish between different network topologies.

Despite the approximations required to develop quantitative models, recent attempts to advance the mathematical modeling of genetic regulation have been reported for relatively simple networks. These models have aimed to elucidate the underlying kinetics of transcriptional regulation [46]. Steady state analysis can be conducted without prior knowledge of any exact rate constants by solely looking at the structure of the network. Results are then confirmed by intensive experimentation. For those rate constants that are known from previous studies, changes to stimuli and reaction constants can then be correlated with changes in system behaviour. Models incorporating assumed kinetic rates are useful for approximating system behaviour and function even though exact kinetic rates are not readily available. In other cases, relative kinetic rates have been used to approxi-

[1] This is where Table **1** will be placed as a new page. See pdf version.

mate system behaviour. Since exact kinetic rates of the components were unknown, randomly chosen kinetic rates were assigned to different components and a large number of computational models were generated [47]. Experimentation was then used to fit the predicted data.

Stochastic Kinetics Approach

While some biological processes can be modeled using simple chemical kinetics and deterministic approaches, many are more accurately represented by random events, in which case stochastic considerations are used instead of specifying differential equations. To capture probabilistic fluctuations in gene expression and genetic regulatory networks such as that in the lysis-lysogeny decision circuit of bacteriophage lambda, stochastic approaches provided more accurate representations [48-50]. Stochastic models can also be used to deduce the effects of noise within a synthetic network, potentially leading to the manipulation of the network itself in order to improve the signal-to-noise ratio within these networks [51, 52].

A stochastic approach regards changes over time as random-walk processes, with no set of differential equations defined, and takes into account inherent fluctuations that are not considered in the deterministic kinetic approach. Stochastic effects may be particularly significant in some biological systems with small molecular populations involved. Although this stochastic basis is more accurate in modeling, it is more difficult to solve mathematically. (Fig. **4-D)** However, numerical simulations are possible using Monte Carlo principles [42, 53, 54]. Instead of considering reaction parameters as reaction rates, they are treated as reaction probabilities. A software called STOCKS (STOChastic Kinetic Simulations) was developed to run Monte Carlo simulations of biochemical processes such as the binding of transcriptional regulators using a stochastic simulation algorithm [38].

Simulations using stochastic considerations have been reported for biological systems involving genetic and enzymatic reactions between molecular populations that were relatively small, including synthetic oscillatory networks [8, 55], transcriptional regulation [55], and circadian rhythms [56, 57]. For large populations of molecular species, the predictions obtained from stochastic approaches match with deterministic ones. However, at smaller population sizes, stochastic effects become more dominant, in which case, deterministic approaches become insufficient [48, 58, 59]. Unfortunately, for many biological networks, stochastic simulations are still computationally expensive due to the huge differences in timescales of biological interactions and population sizes. Various improvements, approximations, and hybrid approaches have been presented [60-66]. In one such study, stochastic simulations were done on multi-scaled systems to study reactions occurring in three different regimes (slow, medium, and fast) as well as coupled reactions. The presented approach showed substantial improvement over using the basic stochastic simulation approach when applied to the study of expression and activity of Lac proteins in *E. coli* [67]. In another, a simple genetic circuit was modeled and simulated using a modified Gillespie algorithm with a quasi-steady-state assumption. This assumption was shown to greatly simplify the stochastic model and to significantly reduce the computational complexity required, speeding up the algorithm [60].

Despite providing a more complete representation of biological networks, stochastic approaches still face the challenge of dealing with several orders of magnitude in terms of scale and properties including binding affinities, specificities, and kinetic rates. Therefore, even statistics-based theories have limitations. Although they provide insights into macroscopic properties of a network, they may have inaccurate predictions about specific interactions. These limitations can be addressed with new developments in integrative modeling.

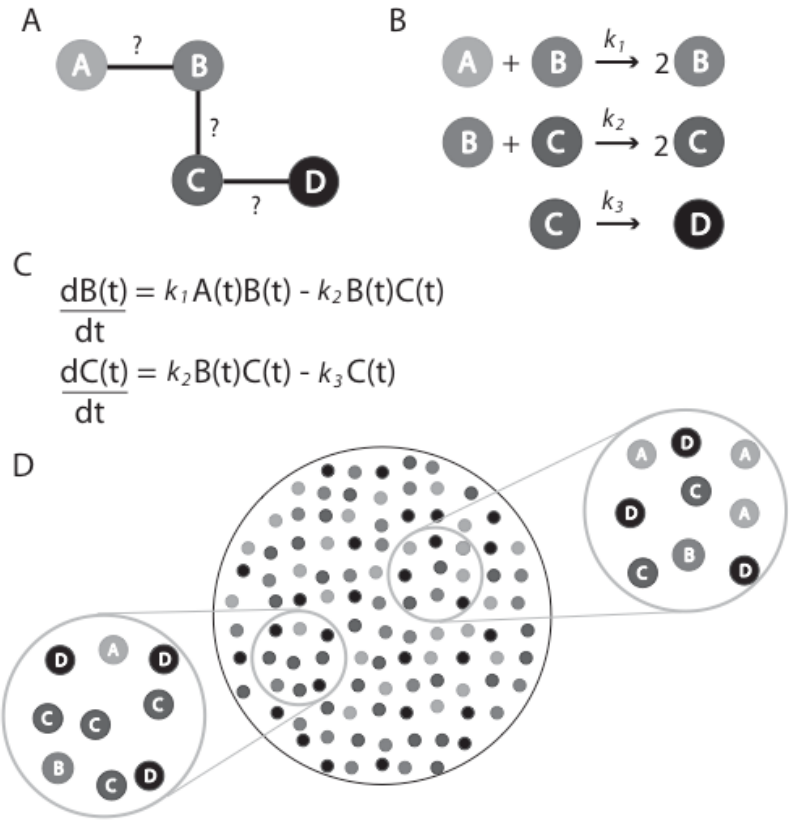
INTEGRATION OF EXPERIMENTATION AND COMPUTATIONAL MODELING

It is important to integrate computational modeling approaches with experimentation. Even when *in silico* models are generated, proper interpretation and a valid means to verify the predictions is required. Detailed simulations of processes occurring within synthetic networks allow researchers the ability to quantitatively predict the behaviour of the network. From this hypothesis, experimental protocols can be designed to verify model predictions. Also, the next piece of information useful for refining the model can be uncovered, initiating the next round of experimentation.

The modeling of networks relies on the accurate characterization of the subsystems. Working with a mathematical model makes it easier to test assumptions and detect contradictions that arise experimentally. For these reasons, integration of simulation and experimental information will dictate the design of explanatory and predictive models for biological systems [68]. Since building a complex biological system entails bringing together many aspects of biological processes and incorporating approximations to fill in any gaps in knowledge, hypotheses formed need to be iteratively verified through experimentation and continually-refined model simulations.

Choosing an Appropriate Approach

Bioinformatics provides an interface to manage, characterize, and interpret data found from studies of both natural and synthetic networks. Bringing together knowledge of subsystems through modeling into a larger model will reveal new properties of the integrated model [69]. However, precision is often difficult to attain due to the lack of precise data and prior knowledge necessary to develop an accurate simulation model. Features of natural systems like spatial distribution, localization, and cross-talk between different pathways are rarely accounted for. In addition, genes may be regulated at the transcriptional or translational level. Transcribed mRNA may be alternatively spliced. Translated proteins may be one of many different isoforms. Proteins often regulate their own production by means of transcriptional control and are involved in many protein-protein interactions. Metabolites are also key components of biological networks. Natural biological loops involving both genes and proteins are not new concepts but there is a drive to attain a quantitative perspective on this regulation and complexity. The level of detail and precision necessary in forming a complete model depends on the application and what is being studied. The level of detail will dictate the number of



$$\frac{dB(t)}{dt} = k_1 A(t)B(t) - k_2 B(t)C(t)$$

$$\frac{dC(t)}{dt} = k_2 B(t)C(t) - k_3 C(t)$$

Fig. (4). Example of different approaches to describing a system. A) Molecules A, B, C, and D describe the protein elements in a protein interaction system, where A is converted into B upon interacting with another B molecule, B is converted into C upon interacting with another C molecule, and C is decayed into molecule D. A qualitative mapping of the connections between each protein does not describe factors such as stoichiometry and kinetic rates. B) A representation of this system as a series of chemical reactions reveals the stoichiometry of each element involved at each intermediate. C) Deterministically, these chemical reaction relations can be described as a set of differential equations where A(t), B(t), C(t) and D(t) are concentrations over time. k1, k2 and k3 are rate constants at each intermediate step. A deterministic approach assumes a spatially homogeneous distribution of molecular species. D) As reactions occur at different rates depending on the local concentrations of reactants, local concentration fluctuations will arise that are different from the rest of the population. These local fluctuations can be studied stochastically for a more spatially accurate representation of the system.

model parameters involved, many of which may need to be approximations and assumptions due to the lack of prior experimental data. Models become unreliable when there is a lack of data or too many variables at once requiring approximations. As a result, the computational modeling approach adopted would be limited by data availability.

The computational modeling approach chosen also largely depends on the application intended and the resolution necessary to describe a certain biological process. Qualitative approaches will be required to establish the connectivity between components, while quantitative approaches are required to observe concentration perturbations. Considerations that need to be taken include whether the model is to be deterministic or stochastic, qualitative or quantitative. Previous computational studies based on quantitative and stochastic models have mainly been on small and simple networks. The reasons for this restriction are a lack of complete quantitative data to input as parameters, only partial characterization of the networks by experimental studies, and expensive computational complexity required to simulate network behaviours. However, the future for quantitative approaches appears promising as better experimental procedures, including high-throughput, large-scale techniques such as microarrays and mass spectrometry are developed [70, 71]. Databases are also being developed to collect shared, published experimental parameter data [72] (Table **2**).

In particular, the construction of useful and predictive synthetic networks allows the direct prediction and measurement of model parameters. As both the complexity and design of the networks are under control of the designer, there are fewer ambiguities and uncertainties. As well, there is a firmer foundation upon which more complex networks can be built. With these developments, quantitative, stochastic approaches will become increasingly practical.

APPLICATIONS OF SYNTHETIC BIOLOGY

Elucidating Network Kinetics of Natural Biological Networks

The design of synthetic networks allows chosen subnetworks of natural biological systems to be isolated. Modeling and experimental studies can be focused first on understanding the isolated subsystem before progressively increasing complexity. Accurate models of synthetic networks provide

Table 2. Databases Developed to Store, Categorize, and Share Data from Biological Studies and Modeling

Database	Description	Website	Source
Alliance for Cellular Signaling (AfCS)	Collection of databases and tools to study signaling processes	http://www.afcs.org/	[73]
BioModels	Database of published, peer-reviewed, quantitative models of biochemical and cellular networks	http://www.ebi.ac.uk/biomodels/	[74]
BioSilico	Integrated web-based database system for metabolic pathways	http://biosilico.kaist.ac.kr	[75]
BRENDA	Information system on enzyme properties and functions	http://www.brenda.uni-koeln.de/	[76]
EcoCyc	Pathway database describing biological networks of *E. coli*	http://ecocyc.org	[77]
ENZYME	Repository for enzyme nomenclature	http://expasy.org/enzyme/	[78]
Kyoto Encyclopedia of Genes and Genomes (KEGG)	Knowledge database system for analysis of gene functions and pathways; includes the databases: GENES, PATHWAY, and LIGAND	http://www.genome.ad.jp/kegg/	[79]
MetaCyc	Database describing metabolic pathways in model organisms	http://MetaCyc.org	[80]
ERGO (WIT)	Database system for comparative analysis of sequenced genomes and metabolic reconstructions	http://wit.integratedgenomics.com/GOLD/	[81]

fundamental insights and act as a foundation with which to describe natural biological networks, including genetic regulatory networks and protein signaling pathways. The ultimate goal of synthetic biology is to construct increasingly complex networks, concomitantly assembling increasingly more complete models of natural systems. The advantage of this approach is that at each stage, subsystems have been characterized by modeling and experimentation, thus keeping the number of unknowns at a minimum. Practically, this approach reduces the degree of trial-and-error experimentation required for the understanding of complex biological networks. Once the structures of synthetic networks are mapped out and their functional dynamic properties are understood, an ever-growing library of circuits will facilitate the classification and comparison of subsequent circuits to provide yet more insights into the complexity of natural biological systems. Synthetic biology allows the study of natural regulatory networks and cellular behaviours using *de novo* networks, potentially leading to future applications in biotechnology and medicine.

Undoubtedly, signaling networks are complex and highly interconnected, interacting at several levels to[2] regulate biological functions within cells [82]. Synthetic genetic regulatory systems mimicking those of mammalian cells have led to the potential of designing mammalian cells with desired properties for tissue engineering, gene therapy, and biopharmaceutics [83]. Furthermore, many diseases result from malfunctioning of natural biological networks including both signaling pathways and transcriptional regulation. In diseases like cancer, single abnormalities in signaling pathways do not lead to complications, but the combined effect of multiple abnormalities to several key pathways result in substantial consequences. Understanding how individual components function within the context of a larger, complex signaling network provides a molecular view of which interactions are involved in causing the diseased state.

[2] This is where Table **2** will be placed as a new page. See pdf version.

Rewiring of Transcriptional Regulatory Networks Controlling Cell Fate

Stem cells are of great interest for their therapeutic potential in regenerative medicine. They are characterized by two properties: self-renewal, the ability to remain in an undifferentiated state, and potency, the capacity to differentiate into more specialized cell types. Embryonic stem cells (ESCs) are said to be pluripotent and are able to differentiate into any cell type of the three germ layers. Until recently, differentiation marked an endpoint of unidirectional development. However, there is accumulating research demonstrating that the state of a differentiated cell, such as a fibroblast cell, can be reprogrammed to induce a pluripotent, ESC-like state. Reprogramming here describes the collective epigenetic changes that cause corresponding changes in the gene expression profile of a cell. Recent studies in both mouse and human cells have identified a small set of ectopically expressed transcription factors capable of forcing fibroblast cells into a pluripotent state. These cells have appropriately been dubbed iPS cells (induced pluripotent stem cells) [84-87].

iPS cells demonstrate the feasibility of rewiring the transcriptional circuitry of specialized cells at the molecular level to manipulate cell fate *in vitro*. By introducing a set of four transcription factors several groups have generated pluripotent cells that are similar to ESCs in their morphology, expression of specific ESC marker genes, proliferation, surface antigens, and epigenetic status [85-88]. It is interesting to note that while some groups used the transcription factors Oct4, Sox2, Klf4, and c-Myc [84-86, 89], others used Oct4, Sox2, Nanog, and Lin28 [87]. Biochemical and genetic studies have established that Oct4, Sox2 and Nanog are responsible for the maintenance of ESC pluripotency and self-renewal [90, 91]. It was thus unexpected that Takahashi *et al* found Nanog to be dispensable as a core reprogramming factor. One explanation is that the introduction of Oct4 and Sox2 can promote the endogenous activation of Nanog. Results have shown that Oct4, Sox2 and Nanog work together to form regulatory network motifs of autoregulatory and feed-forward loops [25].

A better understanding of the differentiation process, the networks regulating it, and the mechanisms controlling cellular reprogramming will further advance iPS studies. Computational modeling of synthetic networks will provide insights that help to identify other factors, both genes and proteins that can be tested for induction of iPS cells. Beyond the identification of factors that can induce pluripotency, considerations to the kinetics and efficiency of this induction are just as important. As can be seen from the use of different transcription factors to generate iPS cells, several of these factors such as c-Myc and Klf4 are not core reprogramming factors. While they can be replaced by other transcription factors, these substitutions may result in delayed and less efficient induction [84]. A computational analysis of how these factors regulate the epigenetic state of a cell will help to define a more specific role for each of the factors used to generate iPS cells.

Furthermore, results from one group have suggested that exogenous expression of Oct4, Sox2, Klf4, and c-Myc may only be necessary during the initial activation of the reprogramming process to trigger changes in the endogenous transcriptional expression program that would lead to pluripotency [88]. Computational models can be used to study how long an exogenous signal must be maintained before sufficient changes in a cell are made for pluripotency. This will also help to address what epigenetic changes such as chromatin reorganization are occurring during the reprogramming of a cell. Computational studies are necessary to further characterize the induction process. The identified transcription factors are sufficient to generate iPS cells; however, it may be a subset of the downstream protein products of these factors that are the necessary inputs. In fact, it may be advantageous to model the modular networks of the input factors instead of the individual factors themselves. For instance, Nanog is part of much larger, more complex protein-interaction network regulating the pluripotency of ESCs [92]. Combinatorial control of genes and proteins within a network is far greater in complexity than experiments alone can elucidate. The observed low efficiency of reprogrammed cells [85] indicates that not all cells are equally responsive to the introduced transcription factors and suggests that other stochastic events may need to be taken into consideration. Considerations of dose-dependencies of the input factors are also relevant for iPS studies.

Computational studies will provide a more comprehensive map of transcriptional regulatory circuitry involved in maintaining different cell states. It would be interesting to see what combination of a few factors and events will stimulate iPS cells, which were originally fibroblast cells, to differentiate into different cell types such as neural, cardiac, and blood cells [86] (Fig. **5**). One recent study has used iPS cells to treat sickle cell anemia in a mouse model [93]. These results demonstrate the therapeutic potential of using reprogrammed cells for the treatment of disease. The generation of patient- and disease-specific pluripotent cells allows researchers to avoid controversial ethical difficulties of using human embryos to extract ESCs and prevents potentially fatal issues of tissue rejection.

Novel Functional and Biochemical Applications

Besides assembling synthetic networks to help elucidate the underlying kinetics of natural biological systems, these networks can be used to monitor and manipulate the kinetics of cellular control, including control at the DNA level using artificial genetic circuits. Any constructed synthetic network adds to a repertoire of modular elements that can be assembled together for specific, novel purposes in medical and biotechnological applications. The introduction of functional synthetic circuits of genes and proteins to control cellular kinetics *in vivo* may be possible in the future with new advancements in the field of synthetic biology. The motivation for using synthetic networks is to compensate for any deficiencies and malfunctioning components of naturally existing networks. Computational models and simulations will reveal any side-effects and ensure that circuits function as intended.

Synthetic networks may serve as an interface between sensory inputs and biological response outputs. Future biotechnology applications include using developed synthetic networks as biosensors of parameter fluctuations in a natural biological system. It may be possible to supplement or replace an existing natural biological function in diseased cells, including the re-engineering of viral regulatory networks in the development of oncolytic viral vectors to target cancer cells [94, 95]. The kinetics of an assembled network can be tweaked for particular purposes based on computational modeling predictions, which is easier with synthetic networks than natural ones.

Detailed models of synthetic systems will soon provide significant insights into drug discovery, such as revealing the effects of feedback mechanisms that may offset the effective dose of drugs [96]. It is conceivable that simulation-based screening of pharmaceutical drugs to confirm both the expected function of the drug and to reveal any unexpected side effects may be used to evaluate future drug discoveries.

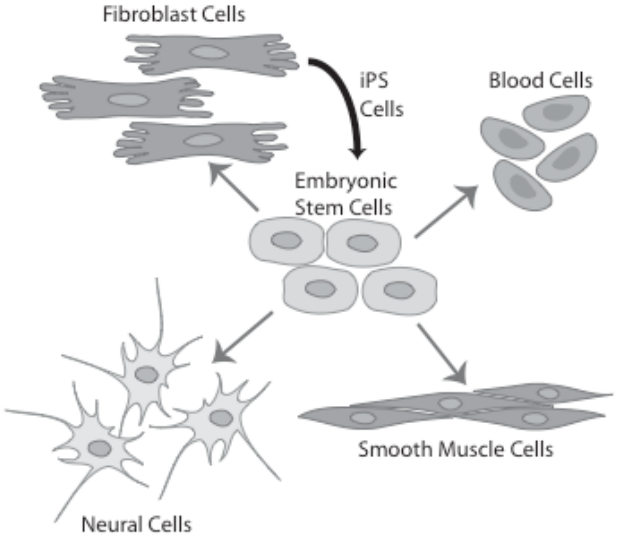


Fig. (5). Cell differentiation and reprogramming. Embryonic stem cells are pluripotent and can be differentiated to many different cell types including fibroblast cells, blood cells, neural cells, and smooth muscle cells. iPS cells have been generated by reprogramming fibroblast cells into pluripotent cells. These patient-derived iPS cells can then be differentiated into other cell types for therapeutic purposes.

Feedback circuits can also be developed to monitor drug release and pharmacokinetics.

Synthetic biology research also ties nicely into molecular-electronics research, where biological molecules can act as analogues of silicon-based integrated circuits. A move towards computing with biological molecules has begun to surface in literature, including molecular-based logic circuitry [97]. Modular synthetic networks can function as logic gates, and the combination of these logic gates into higher complexity systems will feed into the study of computational devices relying on proteins and their interactions [10, 98]. From the construction and characterization of simpler elements such as switches and logic gates, it is possible to build more elaborate devices to perform higher-level functions such as memory devices [13].

CONCLUSIONS

Genomic sequencing has equipped researchers with the ability to generate a huge amount of data. Such sequences provide the 'parts list' for deducing the mechanisms controlling biological circuits. Studies of protein interactions with other proteins, DNA, RNA, and small molecules provide a connectivity map, one which is constantly being refined and revised. Beyond this detailed diagram of connectivity, a study of the kinetics of these interactions is necessary for a systems-level perspective and a more holistic view of biological complexity.

A subset of the terminology and theory can be adopted from circuit engineering and control theory to describe synthetic networks. However, the inherent complexity of biological networks requires the aid of computational modeling approaches to complement such descriptions and experimentation to provide a more complete perspective. Thus, accompanying the design of synthetic networks is the need for extensive modeling capabilities.

Qualitative modeling can be used to map out the many interactions and interconnectivities between network components. For more detailed analysis, quantitative modeling approaches can be used to study parameters that govern the kinetics of biological networks. Deterministic approaches model networks as coupled differential equations. However, more relevant for biological systems are stochastic approaches that provide a more complete representation of the fluctuations in these networks.

There is a large gap between mathematically definable complexity and the inherent complexity of biological systems. Inference is often drawn from limited observations. Thus, to benefit from the full potential of synthetic networks, significant advancement in computational modeling approaches are required to attain a better representation of biological systems. The modeling approach chosen depends on the level of resolution necessary to understand designed synthetic networks. Nonetheless, these modeling approaches are meant to be combined with experimental analysis of any synthetic network constructed.

In addition to the current importance for elucidating the behaviour of natural biological systems, constructed synthetic networks may have their own future applications in drug discovery, biocomputing, and novel cellular functions.

PERSPECTIVE AND OUTLOOK

The ultimate goal of computational efforts is to study the behaviour of higher level biological functions *in silico*. The current trend of synthetic biology is to construct and study select gene regulatory networks and signaling pathways. However, the integration of gene regulation with metabolism, signaling pathways, among other biological processes is more interesting and biologically relevant. Though it may seem a super model – a single model covering all aspects of cellular function – may have some appeal as it would suggest a completion of our quest to fully understand biological function, this super model is of very little practical use. Instead, an ensemble of different models at different levels of resolution, abstraction, time scales, and modular behaviour will allow the understanding of each aspect in finer detail. This knowledge thus enables novel synthetic networks to be assembled in the development of new functions and behaviours for applications in medicine and biotechnology.

REFERENCES

[1] Mello BA, Tu Y. Perfect and near-perfect adaptation in a model of bacterial chemotaxis. *Biophys J* **2003**; 84: 2943-56

[2] Yi TM, Huang Y, Simon MI, Doyle J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Natl Acad Sci U S A* **2000**; 97: 4649-53.

[3] McAdams HH, Shapiro L. Circuit simulation of genetic networks. *Science* **1995**; 269: 650-6.

[4] Tanaka RJ, Okano H, Kimura H. Mathematical description of gene regulatory units. *Biophys J* **2006**; 91: 1235-47.

[5] Guet CC, Elowitz MB, Hsing W, Leibler S. Combinatorial synthesis of genetic networks. *Science* **2002**; 296: 1466-70.

[6] Becskei A, Serrano L. Engineering stability in gene networks by autoregulation. *Nature* **2000**; 405: 590-3.

[7] Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in Escherichia coli. *Nature* **2000**; 403: 339-42.

[8] Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature* **2000**; 403: 335-8.

[9] Bray D. Protein molecules as computational elements in living cells. *Nature* **1995**; 376: 307-12.

[10] Unger R, Moult J. Towards computing with proteins. *Proteins* **2006**; 63: 53-64.

[11] Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. *Nature* **1999**; 402: C47-52.

[12] Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **2002**; 31: 64-8.

[13] Bhalla US, Iyengar R. Emergent properties of networks of biological signaling pathways. *Science* **1999**; 283: 381-7.

[14] Akutsu T, Miyano S, Kuhara S. Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. *Pac Symp Biocomput* **1999**; 17-28.

[15] Chaves M, Sontag ED, Albert R. Methods of robustness analysis for Boolean models of gene control networks. *Syst Biol (Stevenage)* **2006**; 153: 154-67.

[16] Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* **2002**; 18 Suppl 1: S233-40.

[17] Weaver DC, Workman CT, Stormo GD. Modeling regulatory networks with weight matrices. *Pac Symp Biocomput* **1999**; 112-23.

[18] Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* **2001**; 2: 343-72.

[19] Borisuk MT, Tyson JJ. Bifurcation analysis of a model of mitotic control in frog eggs. *J Theor Biol* **1998**; 195: 69-85.

[20] Chen KC, Csikasz-Nagy A, Gyorffy B, Val J, Novak B, Tyson JJ. Kinetic analysis of a molecular model of the budding yeast cell cycle. *Mol Biol Cell* **2000**; 11: 369-91.

[21] Edwards JS, Ibarra RU, Palsson BO. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol* **2001**; 19: 125-30.

[22] Morohashi M, Winn AE, Borisuk MT, Bolouri H, Doyle J, Kitano H. Robustness as a measure of plausibility in models of biochemical networks. *J Theor Biol* **2002**; 216: 19-30.

[23] Salamonsen W, Mok KY, Kolatkar P, Subbiah S. BioJAKE: a tool for the creation, visualization and manipulation of metabolic pathways. *Pac Symp Biocomput* **1999**; 392-400.

[24] Garvey TD, Lincoln P, Pedersen CJ, Martin D, Johnson M. BioSPICE: access to the most current computational tools for biologists. *Omics* **2003**; 7: 411-20.

[25] Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG and others. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **2005**; 122: 947-56.

[26] Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol* **2005**; 23: 961-6.

[27] Dhar P, Meng TC, Somani S, *et al*, Cellware--a multi-algorithmic software for computational systems biology. *Bioinformatics* **2004**; 20: 1319-21.

[28] Dhar PK, Meng TC, Somani S, *et al*, Grid cellware: the first grid-enabled tool for modelling and simulating cellular processes. *Bioinformatics* **2005**; 21: 1284-7.

[29] Hoops S, Sahle S, Gauges R, *et al*, COPASI--a COmplex PAthway SImulator. *Bioinformatics* **2006**; 22: 3067-74.

[30] Ramsey S, Orrell D, Bolouri H. Dizzy: stochastic simulation of large-scale genetic regulatory networks (supplementary material). *J Bioinform Comput Biol* **2005**; 3: 437-54.

[31] You L, Hoonlor A, Yin J. Modeling biological systems using Dynetica--a simulator of dynamic networks. *Bioinformatics* **2003**; 19: 435-6.

[32] Tomita M, Hashimoto K, Takahashi K, *et al*, and others. E-CELL: software environment for whole-cell simulation. *Bioinformatics* **1999**; 15: 72-84.

[33] Baigent S. Software review. Gepasi 3.0. *Brief Bioinform* **2001**; 2: 300-2.

[34] Mendes P. GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems. *Comput Appl Biosci* **1993**; 9: 563-71.

[35] Mendes P. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem Sci* **1997**; 22: 361-3.

[36] Karp PD, Paley S, Romero P. The Pathway Tools software. *Bioinformatics* **2002**; 18 Suppl 1: S225-32.

[37] Le Novere N, Shimizu TS. STOCHSIM: modelling of stochastic biomolecular processes. *Bioinformatics* **2001**; 17: 575-6.

[38] Kierzek AM. STOCKS: STOChastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics* **2002**; 18: 470-81.

[39] Sauro HM, Hucka M, Finney A, *et al*, Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *Omics* **2003**; 7: 355-72.

[40] Schaff J, Fink CC, Slepchenko B, Carson JH, Loew LM. A general computational framework for modeling cellular structure and function. *Biophys J* **1997**; 73: 1135-46.

[41] Schaff JC, Slepchenko BM, Loew LM. Physiological modeling with virtual cell framework. *Methods Enzymol* **2000**; 321: 1-23.

[42] Gillespie DT. Exact Stochastic Simulation of Coupled Chemical Reactions. *J Phys Chem* **1977**; 81: 2340-2361.

[43] Barkai N, Leibler S. Robustness in simple biochemical networks. *Nature* **1997**; 387: 913-7.

[44] Kobiler O, Rokney A, Friedman N, Court DL, Stavans J, Oppenheim AB. Quantitative kinetic analysis of the bacteriophage lambda genetic network. *Proc Natl Acad Sci U S A* **2005**; 102: 4470-5.

[45] Bray D, Bourret RB, Simon MI. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol Biol Cell* **1993**; 4: 469-82.

[46] Smolen P, Baxter DA, Byrne JH. Modeling transcriptional control in gene networks--methods, recent results, and future directions. *Bull Math Biol* **2000**; 62: 247-92.

[47] McDaniel R, Weiss R. Advances in synthetic biology: on the path from prototypes to applications. *Curr Opin Biotechnol* **2005**; 16: 476-83.

[48] Arkin A, Ross J, McAdams HH. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics* **1998**; 149: 1633-48.

[49] Mao L, Resat H. Probabilistic representation of gene regulatory networks. *Bioinformatics* **2004**; 20: 2258-69.

[50] Meng TC, Somani S, Dhar P. Modeling and simulation of biological systems with stochasticity. *In Silico Biol* **2004**; 4: 293-309.

[51] Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science* **2002**; 297: 1183-6.

[52] Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* **2002**; 99: 12795-800.

[53] Gillespie DT. Stochastic Simulation of Chemical Kinetics. *Annu Rev Phys Chem* **2006**; .

[54] Gillespie DT. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *J Comput Phys* **1976**; 22: 403-434.

[55] Kepler TB, Elston TC. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys J* **2001**; 81: 3116-36.

[56] Gonze D, Goldbeter A. Circadian rhythms and molecular noise. *Chaos* **2006**; 16: 026110.

[57] Gonze D, Halloy J, Goldbeter A. Robustness of circadian rhythms with respect to molecular noise. *Proc Natl Acad Sci U S A* **2002**; 99: 673-8.

[58] Hubble J. Monte Carlo simulation of biospecific interactions. *Biotechnol Lett* **2000**; 22: 1483-1486.

[59] McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A* **1997**; 94: 814-9.

[60] Rao CaA, A. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J Chem Phys* **2003**; 118: 4999-5010.

[61] Chatterjee A, Mayawala K, Edwards JS, Vlachos DG. Time accelerated Monte Carlo simulations of biological networks using the binomial tau-leap method. *Bioinformatics* **2005**; 21: 2136-7.

[62] Salis H, Kaznessis Y. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *J Chem Phys* **2005**; 122: 54103.

[63] Bentele MaE, R. 2005. General Stochastic Hybrid Method for the Simulation of Chemical Reaction Processes in Cells. Computational Methods in Systems Biology. Heidelberg: Springer Berlin/Heidelberg. p 248-251..

[64] Haseltine EL, Rawlings JB. On the origins of approximations for stochastic chemical kinetics. *J Chem Phys* **2005**; 123: 164115.

[65] Haseltine ELaR, J.B. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J Chem Phys* **2002**; 117: 6959-6969.

[66] Puchalka J, Kierzek AM. Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophys J* **2004**; 86: 1357-72.

[67] Burrage K, Tian T, Burrage P. A multi-scaled approach for simulating chemical reaction systems. *Prog Biophys Mol Biol* **2004**; 85: 217-34.

[68] Dhar PK, Zhu H, Mishra SK. Computational approach to systems biology: from fraction to integration and beyond. *IEEE Trans Nanobioscience* **2004**; 3: 144-52.

[69] Palcy S, Chevet E. Integrating forward and reverse proteomics to unravel protein function. *Proteomics* **2006**; 6: 5467-80.

[70] Smith MG, Jona, G, Ptacek, J, Devgan, G, *et al*, Global analysis of protein function using protein microarrays. *Mechanisms of Ageing and Development* **2005**; 126: 171-175.

[71] Tucker CL, Gera JF, Uetz P. Towards an understanding of complex protein networks. *Trends Cell Biol* **2001**; 11: 102-6.

[72] Thieffry D, Huerta AM, Perez-Rueda E, Collado-Vides J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. *Bioessays* **1998**; 20: 433-40.

[73] Zavzavadjian JR, Couture S, Park WS, *et al*, The alliance for cellular signaling plasmid collection: a flexible resource for protein localization studies and signaling pathway analysis. *Mol Cell Proteomics* **2006**; .

[74] Le Novere N, Bornstein B, Broicher A, *et al*, BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res* **2006**; 34: D689-91.

[75] Hou BK, Kim JS, Jun JH, Lee DY, *et al*, In YH, Lee SY. BioSilico: an integrated metabolic database system. *Bioinformatics* **2004**; 20: 3270-2.

[76] Pharkya P, Nikolaev EV, Maranas CD. Review of the BRENDA Database. *Metab Eng* **2003**; 5: 71-3.

[77] Keseler IM, Collado-Vides J, Gama-Castro S, *et al*, EcoCyc: a comprehensive database resource for Escherichia coli. *Nucleic Acids Res* **2005**; 33: D334-7.
[78] Bairoch A. The ENZYME database in 2000. *Nucleic Acids Res* **2000**; 28: 304-5.
[79] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **2000**; 28: 27-30.
[80] Caspi R, Foerster H, Fulcher CA, *et al*, MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res* **2006**; 34: D511-6.
[81] Overbeek R, Larsen N, Pusch GD, *et al*, : integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Res* **2000**; 28: 123-5.
[82] Weng G, Bhalla US, Iyengar R. Complexity in biological signaling systems. *Science* **1999**; 284: 92-6.
[83] Weber W, Fussenegger M. Artificial mammalian gene regulation networks-novel approaches for gene therapy and bioengineering. *J Biotechnol* **2002**; 98: 161-87.
[84] Park IH, Zhao R, West JA, *et al*, Daley GQ. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **2008**; 451: 141-6.
[85] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**; 126: 663-76.
[86] Takahashi K, Tanabe K, Ohnuki M, *et al*, Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**; 131: 861-72.
[87] Yu J, Vodyanik MA, Smuga-Otto K, *et al*, Induced pluripotent stem cell lines derived from human somatic cells. *Science* **2007**; 318: 1917-20.
[88] Maherali N SR, Xie W, Utikal J, *et al*, Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **2007**; 1: 55-70.
[89] Lowry WE, Richter L, Yachechko R, *et al*, Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci U S A* **2008**; 105: 2883-8.
[90] Masui S, Nakatake Y, Toyooka Y, *et al*, Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **2007**; 9: 625-35.
[91] Boyer LA, Mathur D, Jaenisch R. Molecular control of pluripotency. *Curr Opin Genet Dev* **2006**; 16: 455-62.
[92] Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. A protein interaction network for pluripotency of embryonic stem cells. *Nature* **2006**; 444: 364-8.
[93] Hanna J, Wernig M, Markoulaki S, *et al*, Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **2007**; 318: 1920-3.
[94] Ramachandra M, Rahman A, Zou A, Vaillancourt M, *et al*, Re-engineering adenovirus regulatory pathways to enhance oncolytic specificity and efficacy. *Nat Biotechnol* **2001**; 19: 1035-41.
[95] Zwiebel JA. Cancer gene and oncolytic virus therapy. *Semin Oncol* **2001**; 28: 336-43.
[96] Cho CR, Labow M, Reinhardt M, van Oostrum J, Peitsch MC. The application of systems biology to drug discovery. *Curr Opin Chem Biol* **2006**; 10: 294-302.
[97] Heath JR, Stoddart JF, Williams RS. More on molecular electronics. *Science* **2004**; 303: 1136-7; author reply 1136-7.
[98] Sundaram K, Srinivasan TS. Computer simulated modeling of biomolecular systems. *Comput Programs Biomed* **1979**; 10: 29-33

Received: February 01, 2008 **Accepted: March 29, 2008**