Published in final edited form as:

Prog Biophys Mol Biol. 2011 August ; 106(2): 400–409. doi:10.1016/j.pbiomolbio.2011.03.003.

Mathematical Descriptions of Biochemical Networks: Stability, Stochasticity, Evolution

Simon Rosenfeld

National Cancer Institute, 6130 Executive Blvd., EPN, Rm 3108, Rockville, Maryland 20852

Simon Rosenfeld: sr212a@nih.gov

Abstract

In this paper, we review some fundamental aspects, as well as some new developments, in the emerging field of network biology. The focus of attention is placed on mathematical approaches to conceptual modeling of biomolecular networks with special emphasis on dynamic stability, stochasticity and evolution.

Introduction

In the widely cited paper “Can a Biologist Fix a Radio?”, Lazebnik, juxtaposes engineering and biological approaches to the analysis of complex systems[1]. Briefly, a superiority of engineering thinking, as viewed in this paper, is in its ability to describe a large variety of complex systems in an unambiguous formalized language thus avoiding fuzziness and uncertainties so characteristic for biology. Lazebnik concludes that “formal approaches would make biological research more meaningful, more productive and might indeed lead to miracle drugs.” Developments in systems biology during the last decade, especially with the advent of high throughput measurements (microarrays, protein mass spectrometry, etc.), show that the vision outlined in [1] is now quickly coming to fruition. For example, broad application of the concept of feedback loops [2, 3], wide use of computational modeling assisted by biology-oriented algorithmic languages [4–6], and other traditionally engineering approaches have emerged as routine tools in systems biology. The state of the art in formalization of biochemical systems has been recently elucidated in [7]. A review and extensive bibliography can be also found in the paper by this author (hereafter referred to as SR) [8].

More and more frequently, the schematic diagrams of biological regulatory systems are mimicking the wiring diagrams in solid state microchip electronics; a perfect example of such schematic charts is presented by now famous depiction of the endomesoderm gene network in the sea urchin embryo produced by the E. Davidson’s lab (<http://sugp.caltech.edu/endomes/#EctodermNetwork>). These schematic diagrams, apart from their scientific content and aesthetic appeal, produce an impression of solidity, determinacy and unlimited reusability that are so typical for high-precision electronics imprinted in silicon. And herein lies a dangerous flaw in what at first sight appears to be a nearly perfect fit between the biological and engineering ways of thinking. For this impression of solidity and determinacy is misleading, if not to say false, when applied to biological systems. Fundamentally, cellular functionality is a vast collection of biochemical

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

reactions between hundreds of thousands of molecular species compartmentalized into a large number of comparatively independent modules and acting simultaneously at many hierarchical levels and time scales. There is nothing solid and predetermined at all in the intra-cellular organization; randomness and fluctuations with a wide spectrum of temporal characteristics and a large variety of probabilistic structures is a hallmark of behavior in such systems.

Randomness in biochemical systems has several fundamental origins. On quantum level, formation of heavy-weight molecules such as amino-acids, bases, or enzymes from hydrogen, oxygen, carbon, nitrogen and other atoms, usually includes the phenomenon of *degeneracy* of energy levels, that is, existence of a dense manifold of different quantum states corresponding to the same energy level. Thermodynamically, each of these states has the same probability to occur in a large ensemble of such molecules. Although equivalent energetically, they may be different structurally, thus introducing internal variability into any process in which they participate. A biologically relevant example is protein folding. It often happens that a polypeptide chain's free energy conformational space does not have a well defined minimum; this may result in markedly different structures and chemical behaviors of proteins after folding [9]. Moving upward to the level of chemical interactions between individual biomolecules, one may find that what is commonly referred to as *chemical reaction* is in fact a large set of complex processes dependent on many internal variables. For example, a chemical interaction between two protein molecules, with or without enzymatic assistance, may include many unsuccessful attempts to find a high-affinity configuration of individual energy landscapes, such configurations that would be capable of withstanding thermal fluctuation and aggressive attacks of rival molecules. Determinacy of outcomes in such attempts may only be guaranteed in asymptotically large systems in which deviations from ensemble average may be safely ignored [10]. In living cells, the number of biomolecules of each sort is not necessarily large and sometimes numbers only in the dozens [11]. As a result, chemical interactions in such systems become essentially discrete processes with large relative fluctuations [12].

Continuing to move up in complexity to the level of gene expression, one finds bewildering number of sources of stochasticity. Thus, Elf and Ehrenberg [13] observe that “the copy numbers of the individual messenger RNAs can often be very small, and this frequently leads to highly significant relative fluctuations in messenger RNA copy numbers and also to large fluctuations in protein concentrations.” There are inevitable statistical variations in the random partitioning of small numbers of regulatory molecules between daughter cells when cells divide [14]. McAdams and Arkin [15] indicate that “time delays required for protein concentration growth depend on environmental factors and availability of a number of other proteins, enzymes and supporting molecules. As a result, the switching delays for genetically coupled links may widely vary across isogenic cells in the population.” Multiple closely spaced ribosomes may process the same strand of mRNA simultaneously. Since the spacings between ribosomes are random, the number of proteins translated from the same transcript may also fluctuate randomly [14]. Recent experiments [16] demonstrated that even in an individual cell, the production of a protein and supporting enzymes is a stochastic process following a complex pattern of *bursting* with random distribution of intensities and durations. Similarly, it was found in [17] that quantitative relations between transcription factor concentrations and the rate of protein production fluctuate dramatically in individual living cells, thereby limiting the accuracy with which genetic transcription circuits can transfer signals. The phenomenon of burstiness is wide spread in genetic regulation. Thus, the authors of [18] report that “transcription occurs in pulses in muscle fibers.” Likewise, it was found in [19] that “transcription of individual genes in eukaryotic cells occurs randomly and infrequently.” Similar observations have been made in [12, 20–22]. Despite many peculiar circumstances surrounding each individual instance of burstiness, a common cause

may be envisioned behind many of them. This cause has been proposed by SR and termed *stochastic cooperativity paradigm* (SCP)[23]. In qualitative terms, the SCP reflects a simple observation that in order for a transcription event to occur, a large number of regulatory proteins should simultaneously appear in the gene's promoter region, become transcription factors, and assemble the reading machinery with RNA polymerase being a key element. Since each of these transcription factors-to-be are the proteins that originate from other genes, confluence of all the prerequisites necessary for successful transcription cannot be anything but a highly sporadic event. Mathematically, the SCP may be formulated in terms of nonlinear dynamical instability and heavy-tailed stochastic processes [23, 24]. Moving further up to higher hierarchical level, one may find that things do not become more orderly on the next level of biological organization, that is, on the level of extracellular matrix and tissue [25, 26]. This brief overview illustrates glaring differences between the biological structures and rigid orderly formats that are often used for their depiction, let alone those which are used for electronic devices intended for unlimited flawless repetition of their functions. It seems that literally nothing in an individual cell is predetermined and guaranteed; everything fluctuates, constantly changing in form, shape and timing. Nevertheless, cellular organization as a whole is self-reproducible to fine detail, capable of withstanding constantly changing environmental conditions and perfectly attuned to the intricate task of maintaining life. Therefore, a natural question arises what are those ultimate, self-reproducible laws that manifest themselves through all the diversity and fluidity of biological forms? Very briefly, the answer is this: these are the physico-chemical properties of biomolecules and their interactions that maintain cellular structural integrity, self-similarity and functional stability. Chemical interactions between biomolecules are undoubtedly a dominating process producing the phenomenon of life. In a sense, any cell is a vast system of intertwined biochemical reactions in which complex compartmentalization, separation of time scales, spatial heterogeneity and hierarchical structure are the epiphenomena of comparatively simple and universal laws of chemical interactions. Somewhat simplifying the matters, it may be rightfully said that two sets of parameters, namely the *kinetic rates* and *stoichiometric coefficients*, implicitly predetermine major features of these systems. A large number of theoretical disciplines have been engaged in deciphering the relations between the properties of elementary interactions and macroscopic structurization, with mathematics, statistical physics, chemical kinetics and nonlinear dynamics being the key players.

In this paper, we review the most fundamental approaches to mathematical description of large biochemical systems. During its long history, this field of research has produced a massive body of literature with hundreds of excellent books and countless number of outstanding reviews. Our limited goal in this paper will be in focusing on comparatively new developments pertaining to dynamical stability and stochasticity in their relation to biological robustness and evolution.

2. Physical Environment of Intracellular Biochemistry

Prior to delving into the intricacies of mathematical theory of intracellular biochemical reactions, it is useful to provide a brief sketch of the environment within which these reactions take place. In [27, 28], Goodsell represents a vivid picture of the interior of a living cell compiled from electron microscopy and X-ray crystallography. He likens this picture to a crowded airport terminal where each person slowly moves to his destination pushing and shoving other people, and where each individual trajectory is hundreds of times longer than it would be in the space free of other passengers. Given high density and big size of macromolecules within the cell, it can be easily imagined that, in fact, the molecules spent a substantial amount of time being clustered in a variety of temporary intermediate associations and exploring their very complex internal energy landscapes. The complexity of

this picture is a far cry from a naive *collision* scheme adopted in many theoretical models [29, 30].

Things become even more complex, and in a sense mysterious, when one looks into the biochemistry of gene expression. A key premise of molecular biology known as *Central Dogma* assumes that there is a unidirectional flow of genetic information from the DNA to protein structures with mRNAs being the intermediaries. Although it is generally accepted that the DNA contains instructions for assembling the proteins and not vice versa, the role of proteins in making the Central Dogma workable is crucial: they serve as transcription factors in the very process of decoding the genes. In order for this circulation of matter and information to be possible, the proteins produced by ribosomes in the cytoplasmic area should be able to reach their pre-specified places in the gene's regulatory regions in a timely manner. Since typically from 30 to 100 regulatory proteins per gene are used as transcription factors [31], a corresponding number of genes should go through their individual cycles of expression in a perfectly synchronized manner; otherwise, a mere shortage of a few transcription factors may lead to drop-out from the regulatory process and a halting of big sections of transcription machinery [23, 24]. Obviously, simultaneous random walk of thousands of molecular species through all kinds of impediments in densely packed intracellular environments is not conducive to such synchronization. Among numerous questions pertaining to this issue, there is one of outstanding importance: how exactly do the regulatory proteins reach their gene-specific places in the corresponding genes' regulatory sites? It is now generally accepted [32–34] that in prokaryotic cells, protein translocation is a combination of 1D non-specific random *slide* along the DNA strand and 3D *jumping* from one section of the strand to another; dense DNA coiling is conducive to such a process. It has been shown theoretically that *in principle* 1D + 3D diffusion is capable of producing effect of the *faster-than-diffusion* translocation. Possibility of such *accelerated diffusion* has been confirmed experimentally [30, 35] *in vitro*. However, it should be noted that both theoretical considerations and experimental setups have been mainly dealing with highly idealized conditions with only two key players in the field, that is, the *uncoiled naked* DNA strand and a single protein. In reality, similar to cytoplasmic environment, the DNA strand is densely populated with numerous transcription factors, holoenzymes, stoppers, isolators and other elements [31], and therefore, motion of individual proteins in such a medium has to be much slower than that observed in the *in vitro* experiments. *In vivo*, a search process can hardly be called a *slide*; the words *running hurdles* seem to be more appropriate. It is this author's opinion that diffusion, whether facilitated or not, is too slow a process to be capable of supporting the seamless functioning of gene expression machinery. It should also be noted that stochastic, diffusion-like search cannot be too fast due to the fundamental limitation known as *speed-stability paradox* [36]: rapid search in conformation space requires a smooth energy landscape, but in such a landscape, the designated junction of protein and DNA cannot be stable due to low affinity. Relatively high stability is a prerequisite for high affinity and specificity underlying the very notion of regulatory site.

Needless to say, theoretical models and laboratory experiments available at this time may be mostly applied to prokaryotic cells. So far the theory is insufficient and incapable of explaining the bewildering intricacies of transcription machinery in eukaryotic cells. Looking at the densely packed eukaryotic DNA chromatin structures with tight coiling around nucleosomes, it is hard to imagine at all how a protein molecule may travel within this structure in search for the appropriate regulatory site. The chromatin remodeling and nucleosome repositioning which are suspected to participate in transcription [37] are far more complex processes than *simple diffusion*. However, the question "If not diffusion then what?" is formidable. Further inquiry into this issue may lead to deep reevaluation of basic principles of biological thinking. This is because current biological views, the Central Dogma included, heavily rely on the existence of *fast physical processes which are not*

known to physics. It is quite possible that a deeper level of reality, with quantum nonlocality being an all-pervading guiding principle, plays a much more substantial role than it is usually thought in biology [38]. Stated in [29]: "...surprises are likely to be in store. The protectorate of physics now grants admission tickets to inquiring cell biologists... We are all welcome to the show."

Since inception, the Central Dogma has been challenged along innumerable lines of attack including the very name of it [39–41]. It is not our goal here to provide a review of this complex issue of very high biological and epistemological significance.

Nevertheless, it seems to be in order to give a very brief summary of the most salient biological facts that limit generality of the Central Dogma. Thus, existence of the *reverse transcription* shows that there is some flow of biological information from RNA to DNA [42]. The phenomenon of alternative splicing shows that the mRNA information content is modulated by a number of *epigenetic* signals. Therefore, it is an oversimplification to think that the role of mRNA is to be just a carrier of an intermediate copy of the genetic code [43]. The mRNAs are transcribed not only from those parts of DNA that contain genes. The non-coding parts of DNA (used to be called *junk DNA*) also produce RNAs. Hence, the role of RNA is wider than to be simply a messenger for delivering genetic code to the protein-synthesizing ribosomes; they have a number of other (largely unknown) regulatory functions. Notably, these functions are inheritable but are not encoded in genes [44]. Post-translational modifications show that building a protein molecule does not end up with creation of the polypeptide chain and initial folding. It is continued under the impact of factors independent on genetic information [45]. The most striking challenge to the Central Dogma has been posed by recent discovery of the *prions*, the proteins that inherit conformational changes completely bypassing the gene expression machinery [46]. This list may be continued [40, 47]. These empirical findings demonstrate how much more complex are the biological realities than a fairly straightforward paradigm of the Central Dogma.

3. Biochemical Networks: What They Are and What They Are Not

In systems biology, the term *network* may refer to a large variety of logical, mathematical and computational constructs, often having very little in common. Many of them are not actually representations of any *system*. Rather, they may be likened to the *constellations* in night skies which consist of the stars grouped only by human imagination but otherwise not connected through any physical process. For example, in a widely used construct of protein-protein interaction network (PPI), the word *interactions* often means nothing more than evidence of being involved in a common process derived from the literature by mere text mining [48, 49]. Whether or not the proteins implicated in such interactions are actually linked biochemically often remains unknown and unquestioned. Although biochemically spurious, such networks may nevertheless be useful in organizing knowledge and presenting it in a graphical form. Global analysis of such networks may help to identify a number of important properties, such as existence of the *small world* [50], *network motifs* [51], and *vulnerability to damage* [52], to name just a few. Moreover, changes in the network's *topological structure* may be indicative of development of a disease and serve as a systemic marker [53]. Our point, however, is that one should not confuse the *heuristic information* represented in the *form of network* with the *real networks*, that is, the ones in which interactions mean actual biochemical and/or physical processes governed by natural laws. In accordance with general theory of systems (see [8] and references therein), big biochemical systems possess a number of emergent properties, the ones that cannot be derived directly from knowledge of elementary chemical interactions. These are emergence, robustness and modularity. Comprehensive understanding of such emergent properties requires the system-

level conceptualization and cannot be derived from the reductionist perspective focused on the system's components [54].

These considerations are fully applicable to the concept of *genetic regulatory network*. Presumably, such networks are intended to depict various aspects of *gene-to-gene interactions*, and there are countless publications which use this term within their titles or key words. However, strictly speaking, since the genes are just different sections of the DNA molecule, they do not *interact biochemically*. They are the parts of a bigger game and interact through at least two intermediaries, that is, through the mRNAs and proteins. This clarification is not just a terminological hair-splitting. Taken separately, *gene-to-gene interaction networks* do not represent any dynamical systems, and therefore there cannot be any physical or biochemical laws that govern their behavior. They may only serve as a way of organizing empirical information and as a vehicle for *heuristic* considerations regarding *possible* processes on a wider biochemical stage. Only when taken together, do the genome, transcriptome and proteome form a minimal system describable by biochemical laws. Contrary to this view, a frequent motif in microarray data analysis is that the genes that are statistically *co-expressed* are (probably!) *co-regulated*. This kind of circumstantial statistical evidence often becomes the basis for declaring the existence of a gene-to-gene interaction. Needless to say, such an evidence is not very reliable, as has been repeatedly demonstrated experimentally [55].

The focus of this paper is on the networks built from biomolecules and their complexes, and the word *interaction* denotes actual biochemical processes. Biochemical networks are not just graphical representation of knowledge, they are dynamical systems governed by the laws of chemical kinetics. As such, they possess a vast array of properties subject to study by the methods of mathematics, statistical physics and nonlinear dynamics. These methods are largely universal across a wide spectrum of natural phenomena; universality of mathematical formulations often helps to reveal fundamental commonalities in the behaviors of such dissimilar systems as biochemical networks, social networks, Ising's spin lattices, systems of oscillators, predator-prey populations, stock market, internet and others [56].

4. Mathematical Descriptions of Biochemical Networks

Mathematical description of biochemical reactions is deeply rooted in chemical kinetics. At its very core, a system of chemical reactions may be written in the form of the *Law of Mass Action*

$$\frac{dx_i}{dt} = F_i(x_1, \dots, x_n) = \alpha_i \prod_{m=1}^N x_m^{P_{im}} - \beta_i \prod_{m=1}^N x_m^{Q_{im}}, \quad i=1, \dots, N \quad (1.1)$$

where α_i, β_i are the kinetic rates, and P_{im}, Q_{im} are the *matrices of stoichiometric coefficients* in the direct and inverse reactions, respectively [57, 58]. Depending on the nature and complexity of the system under investigation, the quantities $\{x_i\}$ may represent concentration of various biochemical constituents participating in the process, including individual molecules or their aggregates. There is no unique way of representing the biochemical machinery in mathematical form: depending on the level of *structural granularity* and temporal resolution, the same process may be seen either as an individual chemical reaction or as a complicated system of reactions. Formally, equations (1.1) are appropriate for a system in which each constituent is generated by one direct and one reverse reaction. The reality of large biochemical systems is, of course, far more complex. In particular, there may be several competitive reactions producing and degrading the same

constituents but following different intermediate pathways. For these cases, a more appropriate form of the equations would be

$$\frac{dx_i}{dt} = F_i(x_1, \dots, x_n) = \sum_{n=1}^{L_i} \alpha_{ni} \prod_{m=1}^{N_{ni}} x_m^{P_{nim}} - \sum_{n=1}^{M_i} \beta_{ni} \prod_{m=1}^{N_{ni}} x_m^{Q_{nim}}, \quad i=1, \dots, N \quad (1.2)$$

known as the law of *Generalized Mass Action*. Here L_i, M_i are the numbers of concurrent reactions of production and degradation, α_{ni}, β_{ni} are the *matrices of rates*, and P_{nim}, Q_{nim} are the *tensors of stoichiometric coefficients*. In principle, however, this more complex system is reducible to the form (1.1) by appropriate redefinition of chemical constituents [59].

The *Power Law Formalism* expressed by the equation (1.1) has been introduced by Savageau in 1969 as a method for studying complex biochemical phenomena [57, 60, 61]. In ensuing years, this method has been applied to numerous problems in biology [62–69]. Later on, the Power Law Formalism has been applied in a more general context of *organizationally complex systems*, beyond purely biological meaning of this word [70, 71]. The term *S-Systems* appeared later for denoting the *synergy-saturation systems*. In addition to its subject matter meaning, this term was useful for distinguishing the S-Systems methodology from other methodologies containing the key words *power law* in their names. Extensive practical work with S-Systems revealed many useful analytical properties which put them into a position of a much more potent instrument than it was originally anticipated. In particular, it turned out that similarly to *neural networks* the S-Systems may serve as *universal approximators*, which means that *in principle* any system of multidimensional nonlinear differential equations may be recast into the form of S-System [72]. Furthermore, it was discovered that S-Systems can be transformed into the *Lotka-Volterra system* known in *population dynamics*; these two nonlinear dynamical models have been developed within two completely independent contexts and were originally thought to be unrelated to each other [73]. In addition, the S-Systems has shown their usefulness in dealing with purely computational problems, such as those in probability theory [74], in numerical integration of differential equations [75], in finding the roots of algebraic equations [76]. A comprehensive summary of these developments, as well as some newer results, are given in [58].

In the S-System approach, the state of a system is characterized in terms of time-dependent *constituents*, $\{x_i(t)\}$. These constituents, as well as kinetic rates and stoichiometric matrices, should be appropriately defined at each level of biological organization. To make this premise clearer, let us imagine that on a certain level of abstraction it could be quite sufficient to describe gene expression by the following pair of biochemical reactions

where TF and RNAP stand for transcription factor and RNA Polymerase, respectively. The structure of stoichiometric matrices is quite obvious from this notation. As to the kinetic rates, they should be defined and measured on this particular level of abstraction. A more close look at the processes symbolically depicted as a chemical reaction (1.3) would reveal a much more complicated picture involving hundreds of biomolecules and thousands of elemental steps, each representing a separate chemical reaction [31, 77]. If one would like to create a model accounting for this level of details then the definitions of *constituents* may drastically change, and the model would become much more complicated and rich in

properties. However, in the S-Systems methodology, the *functional form of the model* (1.1) will remain the same. That's why the S-System is not just another model, one of many possible. It is a *unified framework of thinking* about big and complex hierarchical systems. The originators of the S-System touched upon this fundamental property in many of their works. They refer to the process of convolution of a detailed model of lower level of abstraction into a more concise but more coarse-grained model at a higher level of abstraction as *aggregation*. They write in [78]: "There are alternative hierarchical levels at which biochemical systems can be described by this formalism. The lowest level might correspond to the elemental chemical kinetic description for each of the steps in each of the enzyme-catalyzed reactions. At this level, the power-law description is exactly the same as that of conventional chemical kinetics... At a higher level, one can aggregate system components, and the net rates of increase and decrease of these aggregate variables are again mathematical expressions that can be represented by power-law functions. At a still higher level, one can describe the growth of organisms and their interactions at a biochemical level by means of this same power-law formalism."

Obviously, any hierarchical model of reality, however complex it is, is always open at the top towards further generalization, and at the bottom towards further elaboration and enrichment. In the S-System, however, the *functional form of dynamics* remains the same while the focus of attention moves from the bottom to the top in this hierarchy. Voit and Savageau called this property of S-Systems a *telescopic property* [70, 78].

A major limitation of ordinary differential equations (ODE), such as (1.1) and (1.2), in application to intracellular biochemistry is that they are valid only for the *well stirred* systems. In such systems, any chemical constituent produced anywhere in the system becomes immediately available for all other chemical transformations. As discussed in Section 2, in the tight nuclear environment this assumption may be questionable because of slow delivery of transcription factors to regulatory sites. A way of overcoming this drawback is in accounting for spatial distributions of biochemical constituents and compartmentalization of the model. These measures automatically lead to replacement of the system of ODEs by a system of partial differential equations (PDE) with diffusion terms. A detailed review of these approaches may be found in [79].

As with any discipline seeking support in mathematics, a fundamental limitation comes from the fact that any mathematical model leaves out of its scope a vast universe of *unmodeled realities* thus introducing uncontrollable distortions into any theoretical description. During centuries-long history of this question, a number of powerful techniques have been developed to address the problem. In mathematical biology, existence and influence of the unmodeled realities is usually taken into account through introducing the concept of *intrinsic noise*. Formally, this measure leads to the replacement of the ODE of chemical kinetics by the system of stochastic differential equations (SDE) thus enabling reformulation of dynamics in probabilistic terms [13, 20, 80]. It should be noted, however, that the very usage of the word *noise* has a connotation that *in principle* the system allows for deterministic description and it is only the nuisance factors beyond our control that make such a description impossible. Such a premise is not always justifiable even in low-dimensional nonlinear systems (with Lorenz attractor being a celebrated example [81]) in which the dynamics may assume the form of *deterministic chaos*. This means that the motion in the system, being in principle deterministic and fully predictable, is nevertheless so complex and tangled that may formally satisfy the criteria of stochasticity; hence, for all practical purposes they may be regarded as random [82]. This situation is especially likely in nonlinear systems of very high dimension with strong interactions between components. Equations of chemical kinetics represent a perfect example of such systems. Many aspects

of stochastic behavior in strongly nonlinear high dimensional systems has been discussed earlier in the works [23, 24, 83, 84] by SR.

Serious limitations of the basic model of chemical kinetics in applications to intracellular biomolecular dynamics forced researchers to look for alternative approaches. One such approach has emerged comparatively recently and represents a powerful combination of two, largely independent, mathematical disciplines: graph theory and nonlinear dynamics. It has been termed *Chemical Reaction Networks* (CRN) theory, and in recent decades has been slowly but steadily percolating into intra-cellular biochemistry [85, 86]. A major motivation for introducing the CRN and other graph-theoretical techniques is a regrettable, yet inescapable, fact that *quantitative parameterization* of any large-scale biochemical model may be prohibitively time-, labor- and money-consuming, if feasible at all. Hence, in the vast majority of models such a parameterization may be simply unavailable. In addition, there is no guarantee that *in vitro* measurements of kinetic rates can be truly representative of *in vivo* dynamics. At the same time, *qualitative* relations between the major biochemical players expressed in the form of graphical networks, such as the charts of metabolic pathways [87], are often known much better. Therefore, prior to labor-consuming effort of dynamical modeling, it seems natural to derive as much as possible from the careful examination of the network's topological structure. CRN theory provides a theoretical framework for such an approach and often allows for acquiring surprisingly rich information. In particular, topological analysis allows for making judgments regarding stability, vulnerability to damage, mechanisms of self-repair, scenarios of self-organization and other properties. Moreover, some versions of the network dynamics are capable of elucidating the mechanisms of prebiotic evolution through spontaneous birth of small *autocatalytic molecular sets* followed by rapid self-organization into complex biomolecular structures [88]. In general, topological analysis is capable of elucidating *systemic properties*, that is, the ones that characterize the network's *global behavior*. These global structural properties are hardly possible to derive directly from the equations of chemical kinetics (see {Rosenfeld, 2008 721/id} for more detail.) The power of graph-theoretical methods is seen from the simple fact that neither kinetic rates nor actual concentrations of biochemical constituents are used in topological analyses. This means that the conclusions derived from the topological analysis are applicable to a wide class of systems possessing similar topologies, but otherwise entirely different [87].

5. Nonlinear Dynamics and Stability of Biochemical Networks

Biochemical networks are a special case of nonlinear dynamical systems. As such, their stability may be studied by general principles of nonlinear dynamics. As discussed in [23, 24], standard criteria of stability known in nonlinear dynamics (Routh-Hurwitz, Lyapunov) would impose a set of highly stringent constraints of high algebraic order on kinetic rates and stoichiometric coefficients. There are no known first principles or fundamental laws in statistical mechanics, thermodynamics and nonlinear dynamics that would draw a large biochemical system into the state in which these constraints would emerge naturally, unless the system is *externally controlled* or/and is *artificially designed* to be stable (see more in [24] by SR.)

The *fundamental Deficiency Zero Theorem* (DZT) [89, 90] provides an additional avenue for studying stability. The DZT states that *weakly reversible* chemical networks, i.e., the ones in which each direct chemical reaction is balanced by the *chain* of inverse reactions, are *globally asymptotically stable* under the condition that the network has the *deficiency zero*. The *deficiency* of a network is an integer quantity, $\delta = m - l - s$, where m is the number of complexes, l is the number of linkage classes, and s is the rank of stoichiometric space (see [90] for the definitions and details.) It is also known that if a network is not weakly

reversible, and/or if its deficiency is not exactly zero, then such a network is unstable. The DZT is a powerful statement, and it is hard to escape a temptation to declare the DZT to be a *design principle* of nature and to hypothesize that *evolutionary pressure* takes care about survival of only the systems which satisfy the DZT. However, such a hypothesis, although highly attractive, would be a far reaching extrapolation of the facts at hand to completely unknown territory. It would also task the theory with a fundamental problem of uncovering the *natural* mechanisms that draw a system into the state in which a large number of parameters come to such a perfect match that the *deficiency* comes exactly to zero. We conjecture that in the vast majority of intracellular biochemical networks, the DZT *cannot be valid*. This is because it is difficult to envision how in a system where the number of chemical species is in hundreds of thousands and the number of chemical reactions is perhaps in millions such a precise balance can be maintained. This conclusion is in line with that made in the seminal paper by R. May [91] who asked a fundamental question “Will a large complex system be stable?” A detailed discussion of stability may be found also in [24, 84] by SR. It has been shown in these works that in large biochemical networks, the Jacobian eigen-spectrum in the vicinity of equilibrium has a comparable number of eigenvalues with negative and positive real parts. This is an indication of strong instability.

We conclude this brief overview by the conjecture that *generally* large biochemical networks cannot be dynamically stable. This inherent dynamical instability has numerous implications. First, instability means that a large biochemical network with the number of units in tens of thousands and link density in hundreds cannot behave in a smooth assembly-line manner. Spontaneous failures like traffic jams, bottlenecks, backlogs, delays, loss of synchronization, etc., are unavoidable circumstances surrounding their functioning. In the absence of independent *external* forces, that is, those not involved in the processes *to-be-regulated* and capable of *supervising* and quickly *repairing* these failures, each such event initiates a domino effect of secondary failures thus moving the system *unidirectionally* towards destabilization. In such an unstable environment, any molecular event requiring a team-work of many participants (e.g., transcription), and where each member has to reach its designated site through many unpredictable hurdles, cannot be other than sporadic [23].

Engineering concepts of *feedback loops* have been extensively used in interpretations of complex biological data [2, 92]. However, the prerequisites for applicability of such concepts to *biochemical networks* are rarely formulated explicitly, or even mentioned, in biological literature. Since biochemical networks, whether high- or low-dimensional, are strongly nonlinear dynamical systems, such prerequisites are very far from trivial. First of all, the system has to possess at least one asymptotically stable regime; as seen from the above analysis, a number of intricate criteria should be satisfied to make such asymptotic stability possible. If a stable asymptotic regime does exist, then small deviations from this regime may be analyzed using the concepts of linear theory. This step is equivalent to the Jacobian analysis of stability in which the eigenvalues with negative and positive real parts would correspond to negative and positive feedback loops, respectively. Only the negative feedback loops provide stability, and the conditions for their existence, as explained above, are very complex. In biological studies, all these important prerequisites are often either *assumed* to be valid or simply *postulated* to be the natural outcomes of evolution. Not only is such a hypothesis baseless, but to a large extent incorrect. Careful examination of the issue shows that stability cannot be a preferred choice of evolution. Quite the contrary, internal instability is a prerequisite for evolution; evolution is impossible in stable systems.

The question of dynamical stability of large biochemical networks has many implications in computational biology. It is often the case that the software packages specifically designed for computational modeling of intracellular biochemistry leave the question of stability largely unaddressed. These packages -- usually equipped with easy-to-use graphical

interfaces and convenient scripting languages -- allow one to quickly design any imaginable system in chemical kinetics (see [87], Table 1.3, and [93]). What is easy to overlook in such designs is the question of dynamical stability. If no precautions are made then, according to the central conjecture by R. May [91], the probability that a system turns out to satisfy the conditions of stability due to a miraculous coincidence is miniscule. Therefore, almost for sure such a model will be dynamically unstable. As known from mathematics, if a system does not possess the property of stability then, in a time-course dynamics, its computational convergence to a certain limit may have nothing to do with the properties of the system-to-be-modeled; it may be a purely computational artifact.

Internal instability imposes certain limits on reproducibility of the measurement made on the system. Lack of reproducibility is a common complaint regarding high-throughput measurements with microarrays and protein mass-spectrometry [94], and often is seen as a major obstacle for using them in clinical applications [95]. This comment is not assumed to draw a *direct link* between inherent dynamical instabilities of large biochemical networks and lack of high throughput reproducibility. Rather, it is just a cautionary note that when dealing with inherently unstable systems, the very concept of reproducibility should be carefully crafted; the system is not going to manifest reproducibility in whatever an unwitting researcher wants it to. Such measures as temporal and ensemble averaging may mitigate the problem, but in the same time they may blur fine details and population contrasts. In this context, let us take a closer look at the DNA microarray measurements. The belief that the amount of mRNA harvested from the cell reflects the corresponding *gene's activity* is a cornerstone of using DNA microarrays for measuring expression levels. More specifically, the assumption is that the number of mRNA copies transcribed from a gene per unit of time is proportional to the number of the corresponding mRNA copies harvested from the cell. In formal terms, this conjecture means that the transcription levels are thought to be proportional to the transcription rates, and that this statement is applicable to each gene individually, regardless of the states of other genes. Generally, this assumption is incorrect, and the specific conditions under which it may be correct are rarely spelled out in the literature. Mathematically, the issue is reduced to the question of equilibrium and stability. In the simplest scenario, a stable mRNA concentration depends not only on the rate of its production but also on the half-life of degradation. The rate of mRNA production is gene-specific, whereas the rate of degradation is not. If degradation is fast, then the transcription level may be low despite the fact that the corresponding genes are very active. That means that high degradation rates have a tendency to wipe out the traces of high gene's activity. Conversely, highly stable transcripts will be accumulated with time with little or no degradation thus masking a possibly low transcription rate of the corresponding genes. This simple scheme is grossly complicated by the fact that none of the genes is acting on its own because the transcription factors of a certain gene result from expressions of other genes. Essentially, this means that certain relations between transcription levels and transcription rates may only exist when the entire system is in equilibrium, and even then the relations will be much more complex than a mere proportionality of the individual transcription levels to the *corresponding* transcription rates. This, however, is not the end of story. It is not sufficient that such equilibrium does exist; as repeatedly emphasized above, *it should also be dynamically stable*. A more detailed discussion of these issues may be found in the [83, 96] by SR where it has been shown that the relations between transcription levels and transcription rates become more and more fuzzy when the network's complexity increases. In a qualitative agreement with this finding, it has been demonstrated experimentally that in budding yeast the expression levels and expression rates of about half of genes had either very low or zero correlations [97].

6. Dynamical Instability, Biological Robustness and Evolution

The word *instability* has a negative connotation for a biologist. Observed resilience of biological entities in their strife for survival and their almost unlimited repertoire of responses to any adverse development seems to be a convincing argument against dynamical instability. Another common motif is that evolution had four billion years for the cell's perfection and elimination of unwanted characteristics by natural selection; presumably, the biochemical instability has been one of them. Such a view basically equates the evolution to a goal-setting omnipotent and omnipresent force licensed within the domain of molecular biology and capable of overriding the laws of chemical kinetics and nonlinear dynamics. In addition to obvious philosophical weakness, this view includes, without serious examination, the property of biochemical instability into the list of properties threatening the integrity of life. Interestingly enough, the concept of *intrinsic noise* enjoys a much more favorable status in biology [98]. The opponents of instabilities ignore the fact that intrinsic noise observed at a certain level of organization is nothing else as manifestation of instabilities existing on a deeper level of organization in the same system [99].

Observed biological robustness of living organisms is not a counterargument to dynamical biochemical instability. *Robustness* differs from *stability* in that it deals with maintaining *the system's functions* as opposed to *the system's states* [100]. Seeming contradiction between functional stability of a vast organizational structure consisting of a large number of biochemical networks and possible dynamical instability in each of them is fictitious; it attempts to oppose different levels of biological organization. A logically satisfactory way of looking into these issues is through the paradigm called *dual causality* formulated by Pálsson [101]. He writes: "Unlike physiochemical sciences, biology is subject to dual causality or dual causation. Biology is governed not only by the natural laws but also by genetic programs. Thus, while biological functions obey the natural laws, their functions are not predictable by the natural laws alone. Biological systems function and evolve under the confines of the natural laws according to basic biological principles, such as generation of diversity and natural selection. The natural laws can be described based on physico-chemical principles and used to define the constraints under which organisms must operate. How organisms operate under these constraints is a function of their evolutionary history and survival." Within this paradigm of *dual causality*, inherent dynamical instability represents the *natural laws* and *physico-chemical principles* whereas biological robustness is the result of evolutionary history in which this dynamical instability has been effectively used for gaining evolutionary advantages and survival.

The role of fluctuations in evolution has been extensively discussed in the literature [102]. According to current views, short-term dynamical instability is essential for the very process of *search for a long-term stability*. Without fluctuations created by instabilities, a system would be incapable of exploring the topology of the fitness landscape around its current state and making steps towards long-term stability [103, 104]. Biochemical instability is the material basis for functional plasticity, adaptivity and evolvability.

7. Biochemical Instabilities and Evolutionary Views of Cancer

"Is cancer a disease?" This question may seem utterly inappropriate in this issue of the journal. However, recently Dr. R. Austin of Princeton University made a presentation bearing this title at the National Cancer Institute (available as videocast [105]). Very briefly, the point made in this presentation is that it is conceivable that cancer is an evolutionary response of human race to new threats posed by rapid deterioration of the environment and numerous changes in lifestyle. Therefore, it does not make much sense to wage a *war against cancer* since it would be a war against the evolution itself.

Evolution has become a frequent topic in cancer research, and there are many contexts in which the link between cancer and evolution is being addressed [106–108]. Quite in a spirit of the aforementioned presentation, the authors of [109] consider cancer as a *robust intrinsic state of endogenous molecular-cellular network shaped by evolution*. They write: “The molecular and cellular agents, such as oncogenes and suppressor genes, and related growth factors, hormones, cytokines, etc., form a nonlinear, stochastic, and collective dynamical network, the endogenous molecular–cellular network. This endogenous network may be specified by the expression or activity levels of a minimum set of endogenous agents, resulting in a high dimensional stochastic dynamical system. The nonlinear dynamical interactions among the endogenous agents can generate many locally stable states with obvious or non-obvious biological functions.” According to the [109], stochasticity may accidentally cause a transition from one stable state to another.

Notably, the word *stochasticity* has been mentioned three times in the brief passage above, and it seems that the authors take the very existence of stochasticity for granted. Where might this stochasticity come from? As discussed above in this paper, stochasticity is an unavoidable consequence of high nonlinearity coupled with high dimensionality. A very general view of stochasticity in high-dimensional nonlinear systems has been offered in a classical, highly cited work by Zwanzig [110]. It was demonstrated in this work that in any nonlinear system with a wide spectrum of characteristic times, *fast* variables may be considered as *chaotic* and treated probabilistically, thus introducing a stochastic element into an otherwise deterministic system. Stochasticity of biochemical networks is just a partial case of this general premise. It should be also mentioned in this context that, as shown in [111], in the vicinity of equilibrium, whether stable or unstable, any nonlinear system with conflicting terms (e.g., production and degradation) may be described by the equations analytically similar to the equations of chemical kinetics. In a sense, this means that any large-scale dynamics may be considered, at least locally, as a *kind of chemistry* per appropriate definitions of chemical constituents and kinetic rates (see Appendix A in [24]). This means that the term *biochemical instability* adopted in this paper may be extended far beyond the biochemistry per se.

Any attempt to place large scale dynamics into the evolutionary perspective invariably leads to the concept of *fitness landscape*. The word *landscape* invokes the image of a river flowing with many twists and turns through the terrain with gradually decreasing elevations towards the ocean. An essential difference between the fitness landscapes of evolutionary dynamics and the terrain landscapes is that the latter are two-dimensional whereas the former are those in the space of very high dimension. In mathematical terms, an overall picture of evolution may be seen as *stochastic percolation through the multidimensional fitness landscape towards the states with the highest fitness*.

8. Conclusion

Henri Poincare wrote in *Science and Hypothesis*: ‘Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house.’

To date, a wealth of *facts* has been accumulated regarding individual cellular components and their functions. Bioinformatics, in alliance with computer science, has made great strides in extracting this information from millions of sources and presenting it in mentally digestible forms. Yet, continuing the metaphor by Henri Poincare, the databases are not yet houses as well; they are just easy-to-access warehouses of stones. What makes a house is a well designed structure in which every stone plays a certain role and, in interaction with other stones, supports the house as a whole. In science, the way of whole-structure building

is provided by mathematical modeling. Mathematics represents a systematic and orderly way of describing and organizing knowledge. In the majority of scientific disciplines, mathematical reasoning has proven to be an unparalleled and indispensable tool for understanding complex dynamics.

Eugene Wigner, the Nobel laureate in physics, in his famous paper “The Unreasonable Effectiveness of Mathematics in the Natural Sciences” says that “...enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious”[112]. Molecular systems biology is now making decisive steps towards engaging the full power of mathematical reasoning.

Acknowledgments

It is with great pleasure that the author expresses his gratitude to Dr. P. Prorok and Dr. N. Vydelingum of the National Cancer Institute, Division of Cancer Prevention, for valuable help in preparation of this manuscript.

Reference List

1. Lazebnik Y. Can a biologist fix a radio?--Or, what I learned while studying apoptosis. *Cancer Cell*. 2002; 2:179. [PubMed: 12242150]
2. Becskei A, Serrano L. Engineering stability in gene networks by autoregulation. *Nature*. 2000; 405:590. [PubMed: 10850721]
3. Thomas, R.; D'Ari, R. *Biological Feedback*. CRC Press; 1990.
4. Hoops S, et al. COPASI--a COmplex PAthway SIMulator. *Bioinformatics*. 2006; 22:3067. [PubMed: 17032683]
5. Hucka M, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*. 2003; 19:524. [PubMed: 12611808]
6. Weidemann A, et al. SYCAMORE--a systems biology computational analysis and modeling research environment. *Bioinformatics*. 2008; 24:1463. [PubMed: 18463116]
7. Gauges, R.; Kummer, U.; Pahle, J.; Willy, P. Computation of Biochemical Pathways and Genetic Networks, BioSim Event. In: Gauges, R.; Kummer, U.; Pahle, J.; Willy, P., editors. *Fifth Workshop on Computation of Biochemical Pathways and Genetic Networks*, BioQuant; University of Heidelberg; 2008.
8. Rosenfeld S, Kapetanovic I. Systems Biology and Cancer Prevention: All Options on the Table. *Gene Regulation and Systems Biology*. 2008; 2:307. [PubMed: 19787092]
9. Grosberg, A. Statistical Mechanics of Protein Folding: Some Outstanding Problems. In: Attig, N.; Binder, K.; Grubmuller, H.; Kremer, K., editors. *Computational Soft Matter: From Synthetic Polymers to Proteins*, Lecture Notes. John von Neumann Institute for Computing; 2004. p. 375-400.
10. Gardiner, CW. *Handbook of Stochastic Methods: For Physics, Chemistry, and the Natural Sciences*. Springer-Verlag; 1983.
11. Guptasarma P. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of *Escherichia coli*? *Bioessays*. 1995; 17:987. [PubMed: 8526893]
12. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. *Cell*. 2005; 123:1025. [PubMed: 16360033]
13. Elf J, Ehrenberg M. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res*. 2003; 13:2475. [PubMed: 14597656]
14. McAdams HH, Arkin A. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet*. 1999; 15:65. [PubMed: 10098409]
15. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A*. 1997; 94:814. [PubMed: 9023339]
16. Cai L, Friedman N, Xie XS. Stochastic protein expression in individual cells at the single molecule level. *Nature*. 2006; 440:358. [PubMed: 16541077]
17. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Gene regulation at the single-cell level. *Science*. 2005; 307:1962. [PubMed: 15790856]

18. Newlands S, et al. Transcription occurs in pulses in muscle fibers. *Genes Dev.* 1998; 12:2748. [PubMed: 9732272]
19. Ross IL, Browne CM, Hume DA. Transcription of individual genes in eukaryotic cells occurs randomly and infrequently. *Immunol Cell Biol.* 1994; 72:177. [PubMed: 8200693]
20. Blake WJ, Kaern M, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. *Nature.* 2003; 422:633. [PubMed: 12687005]
21. Golding I, Cox EC. RNA dynamics in live *Escherichia coli* cells. *Proc Natl Acad Sci U S A.* 2004; 101:11310. [PubMed: 15277674]
22. Yu J, Xiao J, Ren X, Lao K, Xie XS. Probing gene expression in live cells, one protein molecule at a time. *Science.* 2006; 311:1600. [PubMed: 16543458]
23. Rosenfeld S. Stochastic cooperativity in non-linear dynamics of genetic regulatory networks. *Math Biosci.* 2007; 210:121. [PubMed: 17617426]
24. Rosenfeld, S. *EURASIP Journal of Bioinformatics and Systems Biology.* 2009. Origins of Stochasticity and Burstiness in High-Dimensional Biochemical Networks.
25. Sonnenschein C, Soto AM. Somatic mutation theory of carcinogenesis: why it should be dropped and replaced. *Mol Carcinog.* 2000; 29:205. [PubMed: 11170258]
26. Sonnenschein, C.; Soto, AM. *Cancer Biol.* 2008. Theories of carcinogenesis: An emerging perspective, *Semin.*
27. Goodsell DS. Inside a living cell. *Trends Biochem Sci.* 1991; 16:203. [PubMed: 1891800]
28. Goodsell DS, Olson AJ. Soluble proteins: size, shape and function. *Trends Biochem Sci.* 1993; 18:65. [PubMed: 8480362]
29. Pederson T. Diffusional protein transport within the nucleus: a message in the medium. *Nat Cell Biol.* 2000; 2:E73. [PubMed: 10806486]
30. Wang YM, Austin RH, Cox EC. Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys Rev Lett.* 2006; 97:048302. [PubMed: 16907618]
31. Kadonaga JT. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell.* 2004; 116:247. [PubMed: 14744435]
32. Coppey M, Benichou O, Voituriez R, Moreau M. Kinetics of target site localization of a protein on DNA: a stochastic approach. *Biophys J.* 2004; 87:1640. [PubMed: 15345543]
33. Halford SE, Szczelkun MD. How to get from A to B: strategies for analysing protein motion on DNA. *Eur Biophys J.* 2002; 31:257. [PubMed: 12122472]
34. Slutsky M, Mirny LA. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys J.* 2004; 87:4021. [PubMed: 15465864]
35. Graneli A, Yeykal CC, Robertson RB, Greene EC. Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proc Natl Acad Sci U S A.* 2006; 103:1221. [PubMed: 16432240]
36. Slutsky, M. PhD thesis. MIT, Physics Dpt; 2005. Protein-DNA Interaction, Random Walk and Polymer Statistics.
37. Shivaswamy S, et al. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* 2008; 6:e65. [PubMed: 18351804]
38. Ogryzko VV, Schroedinger Erwin. Francis Crick and epigenetic stability. *Biol Direct.* 2008; 3:15. [PubMed: 18419815]
39. Morange M. What history tells us XIII. Fifty years of the Central Dogma. *J Biosci.* 2008; 33:171. [PubMed: 18535351]
40. Shapiro JA. Revisiting the central dogma in the 21st century. *Ann N Y Acad Sci.* 2009; 1178:6. [PubMed: 19845625]
41. Thieffry D, Sarkar S. Forty years under the central dogma. *Trends Biochem Sci.* 1998; 23:312. [PubMed: 9757833]
42. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature.* 1970; 226:1211. [PubMed: 4316301]
43. House AE, Lynch KW. Regulation of alternative splicing: More than just the ABCs. *J Biol Chem.* 2008; 283:1217. [PubMed: 18024429]
44. Shapiro JA, von SR. Why repetitive DNA is essential to genome function. *Biol Rev Camb Philos Soc.* 2005; 80:227. [PubMed: 15921050]

45. Eisenhaber B, Eisenhaber F. Prediction of posttranslational modification of proteins from their amino acid sequence. *Methods Mol Biol.* 2010; 609:365. [PubMed: 20221930]
46. Lindquist SL. Prion proteins: one surprise after another. *Harvey Lect.* 2002; 98:173. [PubMed: 16033161]
47. Henikoff S. Beyond the central dogma. *Bioinformatics.* 2002; 18:223. [PubMed: 11847069]
48. Ideker, TE. *Network Genomics, Systems Biology.* Springer; Berlin, Heidelberg: 2007. p. 89-115.
49. Schwikowski B, Uetz P, Fields S. A network of protein-protein interactions in yeast. *Nat Biotechnol.* 2000; 18:1257. [PubMed: 11101803]
50. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004; 5:101. [PubMed: 14735121]
51. Milo R, et al. Network motifs: simple building blocks of complex networks. *Science.* 2002; 298:824. [PubMed: 12399590]
52. Albert R, Jeong H, Barabasi AL. Error and attack tolerance of complex networks. *Nature.* 2000; 406:378. [PubMed: 10935628]
53. Tuck DP, Kluger HM, Kluger Y. Characterizing disease states from topological properties of transcriptional regulatory networks. *BMC Bioinformatics.* 2006; 7:236. [PubMed: 16670008]
54. Aderem A. Systems biology: its practice and challenges. *Cell.* 2005; 121:511. [PubMed: 15907465]
55. Farina L, De SA, Morelli G, Ruberti I. Dynamic measure of gene co-regulation. *IET Syst Biol.* 2007; 1:10. [PubMed: 17370425]
56. Strogatz SH. Exploring complex networks. *Nature.* 2001; 410:268. [PubMed: 11258382]
57. Savageau MA. Biochemical systems analysis. 3. Dynamic solutions using a power-law approximation. *J Theor Biol.* 1970; 26:215. [PubMed: 5434343]
58. Voit, EO. *S-System Approach to Understanding Complexity.* Van Norstand Reinhold; NY: 1991. Canonical Nonlinear Modeling.
59. Sorribas A, Savageau MA. Strategies for representing metabolic pathways within biochemical systems theory: reversible pathways. *Math Biosci.* 1989; 94:239. [PubMed: 2520170]
60. Savageau MA. Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J Theor Biol.* 1969; 25:365. [PubMed: 5387046]
61. Savageau MA. Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation. *J Theor Biol.* 1969; 25:370. [PubMed: 5387047]
62. Savageau MA, Steward JP. Repression of the threonine synthetase system in *Escherichia coli*. *Arch Biochem Biophys.* 1970; 137:181. [PubMed: 4907804]
63. Savageau MA. Concepts relating the behavior of biochemical systems to their underlying molecular properties. *Arch Biochem Biophys.* 1971; 145:612. [PubMed: 4332048]
64. Savageau MA. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature.* 1971; 229:542. [PubMed: 4925348]
65. Savageau MA, Kotre AM, Sakamoto N. A possible role in the regulation of primary animation for a complex of glutamine: -ketoglutarate amidotransferase and glutamate dehydrogenase in *Escherichia coli*. *Biochem Biophys Res Commun.* 1972; 48:41. [PubMed: 4557512]
66. Savageau MA. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature.* 1974; 252:546. [PubMed: 4431516]
67. Savageau MA. Optimal design of feedback control by inhibition. *J Mol Evol.* 1974; 4:139. [PubMed: 4469274]
68. Savageau MA. Genetic regulatory mechanisms and the ecological niche of *Escherichia coli*. *Proc Natl Acad Sci U S A.* 1974; 71:2453. [PubMed: 4601590]
69. Savageau MA. Significance of autogenously regulated and constitutive synthesis of regulatory proteins in repressible biosynthetic systems. *Nature.* 1975; 258:208. [PubMed: 1105191]
70. Savageau MA. Growth of complex systems can be related to the properties of their underlying determinants. *Proc Natl Acad Sci U S A.* 1979; 76:5413. [PubMed: 16592715]
71. Savageau MA. Mathematics of organizationally complex systems. *Biomed Biochim Acta.* 1985; 44:839. [PubMed: 4038284]

72. Savageau MA, Voit EO. Recasting nonlinear differential equations as S-Systems: A canonical nonlinear form. *Math Biosci.* 1987; 87:83.
73. Voit EO, Savageau MA. Equivalence between S-Systems and Volterra Systems. *Math Biosci.* 1986; 78:41.
74. Savageau MA. A Suprasystem of Probability Distributions. *Biomedical Journal.* 1982; 24:323.
75. Irvine D, Savageau MA. Efficient Solution of Nonlinear Ordinary Differential Equations Expressed in S-System Canonical Form. *SIAM Journal on Numerical Analysis.* 1990; 27:704.
76. Savageau MA. Finding Multiple Roots of Nonlinear Algebraic Equations Using S-System Methodology. *Applied Mathematics and Computations.* 1993; 55:187.
77. Lemon B, Tjian R. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* 2000; 14:2551. [PubMed: 11040209]
78. Voit EO, Savageau MA. Accuracy of alternative representations for integrated biochemical systems. *Biochemistry.* 1987; 26:6869. [PubMed: 3427048]
79. deJong H. Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol.* 2002; 9:67. [PubMed: 11911796]
80. Kepler T, Eiston T. Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. *Biophysical Journal.* 2001; 81:3116. [PubMed: 11720979]
81. Lorenz E. Deterministic non-periodic flow. *J Atmos Sci.* 2006; 20:130.
82. Strogatz, S. *Nonlinear Dynamics and Chaos: With applications to Physics, Biology, Chemistry and Engineering.* Springer-Verlag; 1994.
83. Rosenfeld S. Stochastic Oscillations in Genetic Regulatory Networks. *EURASIP Journal of Bioinformatics and Systems Biology.* 2006:1.
84. Rosenfeld, S. Why do high-dimensional networks seem to be stable?-A reflection on stochasticity of dynamically unstable nonlinear systems. In: Gauges, R.; Kummer, U.; Pahle, J.; Willy, P., editors. *Fifth Workshop on Computation of Biochemical Pathways and Genetic Networks;* Heidelberg: University of Heidelberg; 2008. p. 101-112.
85. Bailey JE. Complex biology with no parameters. *Nat Biotechnol.* 2001; 19:503. [PubMed: 11385433]
86. Barkai N, Leibler S. Robustness in simple biochemical networks. *Nature.* 1997; 387:913. [PubMed: 9202124]
87. Steuer R, Junker B. Computational Models of Metabolism: Stability and Regulation in Metabolic Networks. *Advances of Chemical Physics.* 2009 in press.
88. Jain, S.; Krishna, S. Graph theory and the evolution of autocatalytic networks. In: Bornholdt, S.; Shuster, HG., editors. *Handbook of Graphs and Networks: From the Genome to the Internet.* Wiley-VCH; 2003. p. 356-395.
89. Feinberg M. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch Rational Mech Anal.* 1995; 132:311.
90. Siegel D, Chen Y. Global stability of deficiency zero chemical networks. *Canad Appl Math Quart.* 1994; 2:413.
91. May RM. Will a large complex system be stable? *Nature.* 1972; 238:413. [PubMed: 4559589]
92. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet.* 2002; 31:64. [PubMed: 11967538]
93. Alves R, Antunes F, Salvador A. Tools for kinetic modeling of biochemical networks. *Nat Biotechnol.* 2006; 24:667. [PubMed: 16763599]
94. Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet.* 2005; 365:454. [PubMed: 15705441]
95. Ioannidis JP. Is molecular profiling ready for use in clinical decision making? *Oncologist.* 2007; 12:301. [PubMed: 17405894]
96. Rosenfeld, S. Stochastic Oscillations in Genetic Regulatory Networks. Applications to Microarray Experiment. In: Rizzi, A.; Vichi, M., editors. *COMPSTAT-2006.* Physica-Verlag; 2006. p. 1609-1618.

97. Garcia-Martinez J, Aranda A, Perez-Ortin JE. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell*. 2004; 15:303. [PubMed: 15260981]
98. Scott M, Ingalls B, Kaern M. Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks. *Chaos*. 2006; 16:026107. [PubMed: 16822039]
99. Gaspard P. What is the role of chaotic scattering in irreversible processes? *Chaos*. 1993; 3:427. [PubMed: 12780050]
100. Kitano H. Biological robustness. *Nat Rev Genet*. 2004; 5:826. [PubMed: 15520792]
101. Palsson, B. *Systems Biology Properties of Reconstructed Networks*. Cambridge University Press; 2006.
102. Wagner, A. *Robustness and evolvability in living systems*. Princeton University Press; 2005.
103. Krylov N. Relaxation processes in statistical systems. *Nature*. 1944; 153:709.
104. Nicolis, G.; Prigogine, I. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*. John Wiley & Sons; 1977.
105. Austin, RH. Is Cancer a Disease? Videocast of Presentation, NCI Scientific Retreat "Conversations About the Future of Cancer Research". Jan 27. 2009 <http://videocast.nih.gov/PastEvents.asp?c=0&s=131>, .NCI OD, 2009
106. Breivik J. Don't stop for repairs in a war zone: Darwinian evolution unites genes and environment in cancer development. *Proc Natl Acad Sci U S A*. 2001; 98:5379. [PubMed: 11344279]
107. Kitano, H. The theory of biological robustness and its implication in cancer. Ernst. Schering. Res. Found. Workshop; 2007. p. 69
108. Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. *Nat Rev Cancer*. 2006; 6:924. [PubMed: 17109012]
109. Ao P, Galas D, Hood L, Zhu X. Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution. *Med Hypotheses*. 2008; 70:678. [PubMed: 17766049]
110. Zwanzig R. Ensemble method in the theory of reversibility. *The Journal of Chemical Physics*. 1960; 33:1338.
111. Tournier, L. Approximation of dynamical systems using S-Systems theory: Application to biological systems. *International Symposium on Symbolic and Algebraic Computations*; 2005. p. 317-324.
112. Wigner E. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Communications in Pure and Applied Mathematics*. 1960; 13