

# Chemical Reaction Kinetics Is Back: Attempts to Deal with Complexity in Biology

*Developing a Quantitative Molecular View to Understanding Life*

**PETER SCHUSTER**

*Peter Schuster is Professor of Theoretical Chemistry at Vienna University. His publications, about 300 papers and nine books, deal with the analysis and prediction of molecular structures from small molecules to biopolymers and the dynamics of evolutionary processes. The books include "The Hydrogen Bond", three collective volumes together with Georg Zundel and Camille Sandorfy; "The Hypercycle—A Principle of Natural Self-Organization" together with Manfred Eigen; and "Evolutionary Dynamics—Exploring the Interplay of Selection, Accident, Neutrality, and Function", a collective volume together with James P. Crutchfield. His recent research focuses on optimization of RNA structures as a model for molecular evolution and the role of selectively neutral variants. Peter Schuster is at the Institut für Theoretische Chemie und Molekulare Strukturbiologie der Universität Wien, Waeringerstraße 17, A-1090 Wien, Austria.*

**N**ew strategies to handle complexity in biology have been and are developed under the catch phrase "systems biology." What stands in the core of this recent field of research is the concept to understand and model cells and organisms as high-dimensional dynamical systems and to determine the necessary input parameters by experiment. Regulation of gene activities and metabolic functions are encapsulated in differential equations that have their origin in chemical

**Over many decades, molecular biology has been extraordinarily successful in applying a qualitative molecular view to understanding life.**

reaction kinetics. Needless to say, this approach has to envisage enormous complexity. On the other hand, solution of large numbers of kinetic equations, up to one thousand and more commonly rather stiff equations, is routine in combustion chemistry and flame modeling. What's new, however, is the fact that cellular reaction networks have a number of unique properties unknown in physics and chemistry. They are not only self-regulated but they are also capable of reproduction, they are robust and don't change their state under often not so small changes in the environment, and they can tolerate loss of one or the other constituent without losing function. Both the experimental [1] and the computational approach to systems biology [2, 3] have made substantial progress within the last few years. Somehow, the mathematical analysis of the basic properties of genetic and metabolic networks is lagging behind [4, 5]. Despite undoubted success [6], many fundamental questions are still unanswered.

Over many decades, molecular biology has been extraordinarily successful in applying a qualitative molecular view to understanding life. This qualitative image of nature is based on yes-or-no answers rather than the conventional quantitative

Correspondence to: Peter Schuster, E-mail: pks@tbi.univie.ac.at

description applied in physics and chemistry. It uses rough pictures replacing the commonly very fine details of molecular structures. Function is illustrated by means of cartoons rather than equations. This was not always so. Until the 1970s, biochemical kinetics was central to the research in molecular biosciences and explored cooperative processes, such as allosterically induced conformational changes of biopolymers, and the mechanisms of enzyme reactions. Within the last three decades of the twentieth century, however, mathematics and quantitative thinking were largely banned from molecular biology. I've heard hard-nosed professors of molecular biology at European universities saying, "Molecular biology as I understand it, is qualitative in nature!" Let me take, for a moment, the position of devil's advocate: Molecular genetics, one might well say, is even in a pre-Linnaean state because there is no sign for the beginning of the development of a systematic and generally accepted nomenclature of genes and gene products. Instead, molecular biologists continue to name the genes they've discovered, after the entries in their laboratory notebooks, or, they use more or less arbitrary names taken from various sources. I am not denying that there are serious attempts toward a more systematic nomenclature, but they have not (yet) made it into daily laboratory work. Starting in the mid-1990s, biologists began to feel the lack of comprehensive theory and quantitative thinking. As Sir Peter Medawar had already said a decade earlier, "No new principle has declared itself from below a heap of facts." Even the pioneer of molecular biologists and Nobel laureate, Sidney Brenner, urged the development of a novel quantitative and comprehensive theoretical biology in an interview that he gave for the German magazine *Laborjournal*. [7].

Why are the biologists now calling for the return of quantitative aspects? Have qualitative thinking and the con-

struction of raw images come to an end [8]? Has the period of gathering facts reached the limits because the volume and the diversity of data escape the imaginative power of the human brain? I don't think so, but the current mass production of experimental data indeed provides new and hitherto unknown challenges: How are databases created that allow for fast and unambiguous retrieval and provide convenient tools for comparison of information from very different sources? The conventional techniques of data processing work fine with precisely defined objects, for example, sequences and structures at atomic resolution. Gene expression data from microarrays are different in this respect because they have an indispensable quantitative component that is poorly reflected by a yes-no-maybe

**Gene expression data from microarrays are different in this respect because they have an indispensable quantitative component that is poorly reflected by a yes-no-maybe classification.**

classification. Theory is required to bring order into the more complex "heaps of facts" before they can be processed efficiently by computer techniques. The really fundamental problem, however, arises from the numbers of genes, which lie between a few thousand in bacteria and some 30,000 in humans, and the nature of their interactions: Enormously complex regulation networks, rather than simple cascades, are formed by gene interaction and a new kind of network theory that allows for asking the appropriate questions, is required. For a moment, let us imagine the complexity of a mammalian cell: 30,000 genes have the potential to produce the same number of gene products, but, in every cell the majority of gene activities have to be down regulated to leave us with a few thousand

**Enormously complex regulation networks, rather than simple cascades, are formed by gene interaction and a new kind of network theory that allows for asking the appropriate questions, is required.**

active structural or housekeeping genes and about the same number of specific regulatory genes that define the state of the cell and its role within the organ to which it belongs. All this is executed by means of a complex network, interweaving gene activities in subtle manner. The problem is to cut, or better yet, to release this Gordian knot.

Chemical reaction kinetics such as combustion or polymerization have plenty of experience with high-dimensional ordinary or partial differential equations. On the other hand, nonlinear chemical systems have been investigated in great detail [9]. Being autocatalytic processes, these multistep reactions represent excellent examples of multiple steady states, oscillations, deterministic chaos, and spatial pattern formation. The way from the relatively simple nonlinear chemical model systems to the characterization of the states of cells by means of attractors is elaborate and hard to go into detail, but it is straightforward. Complex chemical reactions can also be a suitable study model for the development of novel reverse engineering tools [10] for the study of biological complexity. Perhaps engineering theory is a good method for providing insight into the interplay between resilience, robustness, modularity, and hierarchical control in biological systems. Because most of the kinetic rate parameters of cellular processes are unknown and their determination through measurements is difficult, expensive, and often almost impossible, the solution of the inverse mathematical problem of reaction kinetics consisting of the determination of parameters from measured data is a great challenge

(The forward problem is the computation of solution curves from rate parameters and initial conditions).

It is already commonplace to say that understanding complexity in biology and in other disciplines will not be possible without a joint effort integrating experience from many branches of science and engineering, including mathematics,

computer science, physics, electrical engineering, and chemistry, into modern biology. In reality, to achieve in such a great synthesis toward the life sciences of the future is a different story, but interdisciplinary research is no longer placed at the side-table of funding agencies. The reorientation of molecular life sciences has clear-cut consequences for an up-to-

date education of biology students. University curricula have to be adapted to these new developments. Fortunately, this fact has already been appreciated and has reached current awareness in the United States [11, 12] and in other countries. With very few exceptions, however, the universities in continental Europe are still lagging behind.

## REFERENCES

1. Tewari, M. et al. Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF- $\beta$  signaling network. *Mol Cell* 2004, 13, 469–482.
2. Hasty, J.; McMillen, D.; Isaacs, F.; Collins, J.J. Computational studies of gene regulatory networks: In numero molecular biology. *Nat Rev Genet* 2001, 2, 268–279.
3. Kitano, H. Computational systems biology. *Nature* 2002, 420, 206–210.
4. Strogatz, S.H. Exploring complex networks. *Nature* 2001, 410, 268–276.
5. Friedman, N. Inferring cellular networks using probabilistic graphical models. *Science* 2004, 303, 799–805.
6. Almaas, E.; Kovács, B.; Vicsek, T.; Oltvai, Z.N.; Barabási, A.-L. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 2004, 427, 839–843.
7. Sydney Brenner im Gespräch, Eine einsame Stimme aus der Prägenomik Ära (An interview with Sydney Brenner: A lonely voice from the pre-genomic era). *Laborjournal* 2002, Heft 4, 28–33.
8. Fagerström, T.; Jagers, P.; Schuster, P.; Szathmari, E. Biologists put on mathematical glasses. *Science* 1996, 274, 2039–2040.
9. Sagués, F.; Epstein, I.R. Nonlinear chemical dynamics. *J Chem Soc Dalton Trans.* 2003, 1201–1207.
10. Csete, M.E.; Doyle, J. C. Reverse engineering of biological complexity. *Science* 2002, 295, 1664–1669.
11. Bell, E. The future of education in the molecular life sciences. *Nat Rev Mol Cell Biol* 2001, 2, 221–225.
12. Ideker, T. Systems biology 101—What you need to know. *Nat Biotech* 2004, 22, 473–475.