Irun R. Cohen

Modeling immune behavior for experimentalists

Authors' address
Irun R. Cohen
Department of Immunology, Weizmann Institute of Science, Rehovot, Israel

Correspondence to:
Irun R. Cohen
Department of Immunology
Weizmann Institute of Science
Rehovot 76100, Israel
Tel.: +972 8 934 2911
Fax: +972 8 943 4103
E-mail: irun.cohen@weizmann.ac.il

Immunological Reviews 2007
Vol. 216: 232–236

Immunological Reviews
0105-2896

Summary: This article outlines the requirements for useful models of biologic systems. Such models should fulfill three conditions: (i) suit the bottom-up data of the living system, not merely adhere to top-down logic; (ii) abet experimentation by stimulating new ideas for novel experiments; and (iii) engage the mind of the experimentalist with understandable, visual representations. Seven characteristics of a useful model are discussed.

Keywords: *theoretical immunology, immune system modeling, computer modeling*

The modeling problem

Experimental immunology, from its birth as a field of science, has been influenced if not driven by theories and often contending theories: humoral versus cellular immunity, *horror autotoxicus*, instruction (out) versus selection (in), forbidden clones, suppressor T cells (out) versus regulatory T cells (in), idiotypic networks (in then out), non-self versus self recognition, deletion versus regulation (both in), danger, natural autoimmunity, the immunologic homunculus, among others. Many of these notions and theories have lacked precise definitions; yet, they have motivated experimenters to test, affirm, support, deny, or falsify them. Theories that drive experimenters to experiment are successful scientifically, even if they turn out to be wrong.

In contrast to the great influence on experimentation of much imprecise thinking, precise mathematical formulation has had essentially no impact on either experimentation or concept in immunology. This anomaly has frustrated the modelers only; the experimenters, until now, did not fret much about precise modeling. Now, however, more and more experimenters want good models.

The reductionist problem

Physics, the paragon of sciences, has taught us that the seemingly erratic behavior of the material world can be reduced

to immutable laws of nature. We attain true understanding only through probing beneath appearances to discover these underlying laws. This reductionist program, however, seems to have failed biology. After decades of brilliant experimentation, biologists, with immunologists among them, have brought to light most, if not all, of the genes, molecules, and processes that generate living systems: the genome project has sequenced all the DNA, the proteome project is rapidly advancing, signal transduction is becoming clear, and the metabolome and the reactome are being tamed. Yet, no underlying, immutable laws of life have emerged. The seminal idea of evolution through natural selection, which preceded modern biology, is still biology's only law of nature. Indeed, biology seems to be suffering from too much information. The more data we have access to, the more confused we have become. The living cell and all the creatures it has generated are simply too complex to be understood just by inspecting the data. We have reduced the organism to its component parts, but we cannot reduce the complexity of the organism to understanding. There are simply too many parts, and each part does too many different things in the various states of development, differentiation, disease, repair, healing, and aging. The organism and all its systems are formed by molecules that are too pleiotropic, too redundant, too wasteful, and too degenerate for neat logic (1, 2).

The problem of emergent properties

Most frustrating is the fact that the characteristic behaviors of living organisms are emergent properties; emergent properties are the properties of the system as a whole that are not featured by any of the component parts of the system (2–4). The exquisite specificity of the immune response, for example is greater than the specificity of any of the variably degenerate antigen receptors that participate individually in the response (5). Life is an emergent property of cells made entirely of dead molecules. Thoughts are the emergent properties of brains made entirely of thoughtless neurons, and so forth. Reducing living systems to component parts and processes simply does not grant understanding. Biologists need models to try to put humpty dumpty together again, as all the king's horses and all the king's men have failed.

The problem of scale

Living systems and their emergent properties embrace many scales (or levels) simultaneously. The formative scales of living systems involve both continuous interactions and discrete objects. Continuous molecular interactions, for example at a 'lower' scale, create a discrete cell at a 'higher' scale. We may proceed up or down the scales. Going down the scale, we can see that continuous atomic interactions create a discrete molecule. Going up the scale, we can see that continuous cellular interactions create a discrete organism. Continuous interactions between organisms create a higher scale society or species.

Understanding biologic causality involves knowing, for example how the modification of molecular interactions crosses a scale to influence cellular behavior (and, ultimately, the behavior of the organism, the society, and the species). Emergent properties, in other words, become visible by the way causality crosses from one scale to the next (3).

The problem of complexity

Living objects are packages of information: sequences of DNA, all kinds of RNA molecules, protein amino acid sequence, protein structure, lipid and carbohydrate chemistry and structure, membranes, and all kinds of other molecules, large and small. Information is also dynamic: gene activation, the proteome, signaling pathways, enzymatic pathways, replication, death, and so forth. Biologic systems never rest; everything is on the move.

These vast arrays of information, by themselves, generate ever more information. Information is autocatalytic (6): genomes duplicate and diverge, and existing molecules continuously get exploited for new functions as we go up the evolutionary tree. Consider introns, alternative splicing, post-translational modifications, etc. It takes about 20 000 genes to make a round worm, but you can get a human from about 30 000 genes. The same sets of genes are exploited in increasingly complex ways in the human compared with the worm. Life, in other words, evolves increasing complexity, automatically.

The autocatalysis of information and, hence, complexity makes it difficult to grasp the logic of the organism. There is a design problem. The more complex is not necessarily the more fit (2, 6); indeed, the more complex may be the more fragile. The box of round worms (*Caenorhabditis elegans*) survived the Columbia re-entry disaster, but the human astronauts did not. The complex dinosaurs did not survive the consequences of the meteoric impact, but simpler creatures survived and took over to regenerate and evolve a higher degree of complexity, including us. Who knows how long we will last? Simpler *Escherichia coli* is more likely to survive everything.

Evolution produces workable not optimal solutions, while she goes on generating complexity. Human reason and

intuition, therefore, cannot be trusted for understanding a biologic system before the data are in. Models have to fit data first and logic is secondary (1).

The problem of dynamics

Living systems exist by virtue of their dynamics (2, 3); a system at rest is a dead system. Experiments, by their nature, are usually static and performed at fixed times and fixed concentrations of reactants, while the real living system proceeds to evolve with time and varying concentrations. Living systems are reactive systems that stop only upon death (2, 3). Unfortunately, dynamic models have to be constructed from data that are essentially static. Experiments are and must be contrived and limited to a two-dimensional or three-dimensional snapshot of a moving, four-dimensional picture. Models that fit biologic systems have to bridge the gap between our piecemeal and static data and the continuously interacting and changing organism and its environment that is the living system. Moreover, once we construct a dynamic model, it should allow us to see and experiment with the dynamics of the system. A useful model should be interactive; you should be able to play with it and see what happens.

Child's game

It is natural that humans would suppose that the world was constructed in the way a human would construct it – according to an intended plan. How do we discover the basic plan of life? The tools of biology are ideas, observations, experiments, and logical interpretations. Biologists are experts in using the tools of enlightened human reason. The problem is that the emergence of life is not a product of enlightened human reason; life, childlike, seems to have evolved through the trial-and-error tinkering of nature with itself (7). Biology then is like a game played by experts that aims to discover the principles of a game played by a curiously inventive child. The organism, like a spaceship from another world, has appeared without blueprints, so we do not know the principles that were used to build the machine. The user's manual is not in the glove compartment, so we do not know how the machine flies. There are no service contract or repair recommendations, so we do not know how to care for the machine. Biology aims to deduce the blueprints, and medicine aims to write the user's manual. In this sense, biology and medicine are always engaged in constructing models. At this stage of the game, we hope that we can get help from computers to build more accurate and more useful models.

The aims of modeling

Models for living systems for biologists should fulfill three conditions: (i) suit the reality of living systems, not merely adhere to top-down logic; (ii) abet experimentation by stimulating new ideas for novel experiments; and (iii) engage the minds of the experimentalist with intuitively understandable representations; numbers alone will not do.

Model requirements

To be useful for biologists, models should stimulate experimenters to experiment, while they take account of the data in hand. To reach these aims, an ideal model should have the following seven characteristics. (i) Bottom-up: The useful model must begin with the data, bottom-up. The data are the essence. (ii) Object-oriented: Biologists experiment with objects (molecules, cells, organisms, etc) and their interactions. Biologists think about objects, their changes of state, and how they interact. So a model useful to biologists should deal with objects and not only with concepts or numbers. (iii) Dynamic: The model must allow us to deal with the system's dynamics faithfully. (iv) Multi-scalar: To see emergent properties and to study them, the model must be able to cross scales. We need to be able to zoom into cells to see how molecular interactions affect the cells, and we need to zoom out from the cells to see how cellular interactions affect the organism, etc. (v) Modular: Because we are continuously gathering new data from new experiments, a useful model should be modular, so that we can add new data to the existing model without having to redo the whole thing. (vi) Interactive: A useful model has to allow us to do experiments *in silico*, as an introduction to doing real experiments *in vitro* and *in vivo*. We want to be able to play with the model and see the consequences of modifying this or that component, concentration, reaction time, etc. (vii) Realistic: Experimentalists tend to be experimentalists because they are more at home with the concrete than they are with abstract or numerical representations of the concrete. A useful model should show us things that trigger a 'Wow' response. A good model has to talk to us in a language we understand. Much human understanding is visual. The human mind processes large amounts of information visually (8), and we feel, we understand a thing better when we see it. 'Now I see', usually means, 'now I understand'. Indeed, biologists traditionally like to draw pictures of cells (round objects) and arrows (interactions) when they want to understand and explain cellular interactions. Thus, a model that features visual output is more likely to be perceived as realistic and understandable.

A comment on data-based modeling

How can we build a good model bottom-up from the data when the data are necessarily incomplete? We do not know all the facts; indeed, living systems are so complex that even our grandchildren will never know all the facts (9). Because the data are incomplete, our models will be incomplete. Can an incomplete model at all work? Can an incomplete model simulate reality?

Fortunately, living systems tend to have evolved to be robust; they still perform when parts are missing or changed. Look at the phenotypes of knock-out mice: despite removing key genes, we often get viable mice (10). Biologic systems can work amazingly well without various component parts. Or consider polymorphism: individuals with different versions of key molecules can still do more or less the same things. This robustness and resilience emerge from the complexity of the organism; redundancy, pleiotropism, and degeneracy allow the living system to make do. So too can we get working simulations of the real thing, despite the fact that the simulation might be missing various parts of the real system (11). Still, a model, even a bottom-up model, will always be tentative. But that is OK with science; scientific knowledge is tentative by nature. We just have to keep reality-testing our models, by way of prediction, continuous experimentation, and observation (12). We need the model to help us see what experiments are important to do and the model to explain the results.

An important advantage of data-based modeling is that it minimizes the top-down teleology that taints so much of biologic thinking and immunologic thinking in particular. Rather than assume that the immune system functions to distinguish between self and not-self (13) or to recognize danger (14), it is more in the spirit of science to prefer hard data to soft ideas for guidance in formulating and interpreting experiments (1).

Test models

Strategies for building immune system models that fulfill the requirements listed above are beginning to appear. Simmune, devised by Meier-Schellersheim and his colleagues (15), is one such model. Another model is Reactive Animation formulated by Efroni, Harel and this author (11, 16–19). The details of these models are beyond the scope of this article; interested persons can find the details in the publications. It is important for now to note that these models are based on defined objects and their interactions and changes of state, allow one to zoom in and out across scales, are modular so that new data can be added (or data in error removed), and are dynamic and are interactive, inviting experimentation and prediction. An innovative feature of Reactive Animation is that it allows the data themselves to create a realistic moving picture of the system in action (11).

Life science–computer science alliance

The bottom line is that biology, and immunology with it, will progress into the new area of systems biology, the synthesis of the component parts into a dynamic system, with the aid of the tools and thinking of computer science. Useful modeling will emerge from the alliance between life science and computer science (20). Obviously, the way we have to model biologic systems will affect the way we shall have to educate our students. The new research modeling will require a new model of research education.

References

1. Efroni S, Cohen IR. Simplicity belies a complex system: a response to the minimal model of immunity of Langman and Cohen. Cell Immunol 2002;**216**:22–30.
2. Cohen IR. Tending Adam's Garden: Evolving the Cognitive Immune Self. London: Academic Press, 2000.
3. Cohen IR, Harel D. Explaining a complex living system: dynamics, multi-scaling and emergence. J R Soc Interface (in press).
4. Deguet J, Demazeau Y, Magnin L. Elements about the emergence issue: a survey of emergence definitions. Complexus 2006;**3**:24–31.
5. Maverakis E, van den Elzen P, Sercarz EE. Self-reactive T cells and degeneracy of T cell recognition: evolving concepts-from sequence homology to shape mimicry and TCR flexibility. J Autoimmun 2001;**16**:201–209.
6. Cohen IR. Informational landscapes in art, science and evolution. Bull Math Biol 2006;**68**:1213–1229.
7. Carroll SB. Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom. New York: W.W. Norton, 2005.
8. Molnar-Szakacs I, Kaplan J, Greenfield PM, Iacoboni M. Observing complex action sequences: the role of the fronto-parietal mirror neuron system. Neuroimage 2006;**33**:923–935.
9. Fleck L. Genesis and Development of a Scientific Fact. Chicago: University of Chicago Press, 1979.
10. Ponomarev I, et al. Transcriptional signatures of cellular plasticity in mice lacking the alpha1 subunit of GABAA receptors. J Neurosci 2006;**26**:5673–5683.
11. Efroni S, Harel D, Cohen IR. Emergent dynamics of thymocyte development and lineage determination. PLoS Comput Biol (accepted for publication).
12. Mayr E. Cause and effect in biology. Science 1961;**134**:1501–1506.
13. Burnet FM. Self and Not-Self. Cambridge: Cambridge University Press, 1969.
14. Davies EL, et al. Heat shock proteins form part of a danger signal cascade in response to lipopolysaccharide and GroEL. Clin Exp Immunol 2006;**145**:183–189.
15. Meier-Schellersheim M, et al. Key role of local regulation in chemosensing revealed by a new molecular interaction based modeling method. PLoS Comput Biol 2006;**2**:e82.

16. Efroni E, Harel D, Cohen IR. Towards rigorous comprehension of biological complexity: modeling, execution and visualization of thymic T cell maturation. Genome Res 2003;**13**:2485–2497.
17. Efroni S, Harel D, Cohen IR. Reactive animation: realistic modeling of complex dynamic systems. Computer 2005;**38**:38–47.
18. Efoni S, Harel D, Cohen IR. A theory for complex systems: reactive animation. In: Paton R, McNamara L, eds. Multidisciplinary Approaches to Theory in Medicine. Studies in Multidisciplinarity, Vol. 3. Amsterdam: Elsevier, 2005:309–324.
19. Harel D, Efroni S, Cohen IR. Reactive Animation. Lecture Notes in Computer Science 2003;**2852**:136–153.
20. Harel D. A grand challenge for computing: full reactive modeling of a multi-cellular animal. Bull EATCS 2003;**81**:226–235.