

Good practices for building dynamical models in systems biology

Evren U. Azeloglu and Ravi Iyengar*

Dynamic models can offer deep understanding of information processing mechanisms in physiology, cell signaling, and biological regulation when they are appropriately detailed. Here, we describe some of the key aspects of the model-building process, including proper parameterization and error analysis, as well as common mistakes, such as model-tweaking and oversimplification, which can decrease the value of the models.

Spherical cows and the dangers of model-tweaking

Dynamic modeling of signaling networks provides mechanistic insights of cellular regulation by identifying systems-level emergent properties—that is, properties that arise from network components and their directional interactions—and it helps to discover drug targets and to understand drug action. Most cellular activities can be represented by chemical reactions; ordinary differential equation (ODE)-based models can capture their dynamics and produce testable predictions. Building ODE models for chemical reactions is straightforward if the initial concentrations of the reactants and the forward and reverse reaction rates are known. The complexity of regulatory biology, from molecules to cells to organisms, is daunting, so building simplified models offers a practical solution. However, simplifications often lead to “spherical cows,” a humorous description used by physicists to describe simplified models that deviate substantially from reality. In regulatory biology, spherical cows can arise from two kinds of operations. In the first, a simplified model does not consider the details, producing a biologically unrealistic result. Typically, such a model will have far fewer equations than would a realistically detailed model. In the second, the unknown or difficult-to-measure model parameters are arbitrarily selected to obtain or “fit” a desired output. Such operations produce models that are unlikely to provide mechanistic insights into biological processes or to predict the behavior of the

system in a pathological context. Here, we describe a set of good practices for building models with incomplete knowledge of system topology or kinetic parameters.

Einstein’s assertion in 1934 that “it can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience” (1) is an excellent guiding principle for building pathways and networks that form the basis for systems of coupled ODEs that make a model. Consider the pathway from β -adrenergic receptors through protein kinase A to the transcription factor CREB. How much detail is required to describe the dynamics of this system? Minimally, four coupled equations could suffice; however, in a study of this pathway in regulating the function of kidney podocytes, we found that we needed to build a progressively more complicated network with additional components so that model simulations matched experimentally observed time courses (2), thus adhering to Einstein’s dictum. We tested our model using standard signal transduction assays and comprehensive error analysis; the “right-sized” model had the least error when considering system dynamics. Thus, error analysis provides a method to test the depth of detail a model needs to contain. To obtain the “right-sized” model that provides mechanistic insight into how the input-output (I/O) relations arise, it is insufficient to obtain just a match of the I/O relationship between theory and experiments. Such matches need to be obtained for relationships between intermediary components as well. When such concordance is achieved, the model provides deep insight into how information is processed within signaling networks to achieve the observed I/O relationships.

Building ODE models with incomplete parameter sets

A major problem in building dynamical models is the lack of experimentally measured parameter values. In most cases, the concentration of a protein of interest within a cell is not known; however, it can be estimated either from protein purification tables (3) or by quantitative immunoblotting by using specific antibodies (4). To simultaneously estimate the abundance of large numbers of proteins, mass spectrometry is useful. Although usually absolute protein concentrations cannot be obtained with mass spectrometry, changes in protein amounts can be easily measured (5).

Kinetic parameters for biochemical reactions, such as forward and reverse rate constants, have been rarely measured explicitly, and often these are determined with purified proteins under conditions that are dissimilar from those within the cell. Estimating rate constants from time course data is an effective approach. Time course experiments performed with intact cells, combined with activity assayed separately under cell-free conditions, provide some of the best data for estimating rate constants to obtain kinetic parameters. Such parameter estimations can be readily done in Matlab by using curve-fitting algorithms. Typically in our laboratory, the parameters are obtained from a number of experimental systems and assays. One common criticism is that parameters obtained by combining data from multiple systems are “canonical” and may not be relevant to a specific cell type. However, our studies on the adenosine 3',5'-monophosphate (cAMP) pathway show that the concentration of cAMP predicted from simulations, by use of canonical parameters obtained from multiple cells and tissues (6), are in agreement with values measured in different types of cells (7).

Although obtaining initial concentrations and kinetic parameters from different approaches or cell types is acceptable, parameterization of the model must be completed before simulations are started (Fig. 1). Information about how and where the parameters were obtained and a detailed description of the initial conditions, including starting assumptions and simplifications, are necessary (8, 9). Large models can be constrained in a modular fashion (8). One should not change the conditions to match experimental observations once the simulations are started; that is, one should not alter starting kinetic parameters in order to

Department of Pharmacology and Systems Therapeutics, Systems Biology Center New York, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.

*Corresponding author. E-mail: ravi.iyengar@mssm.edu

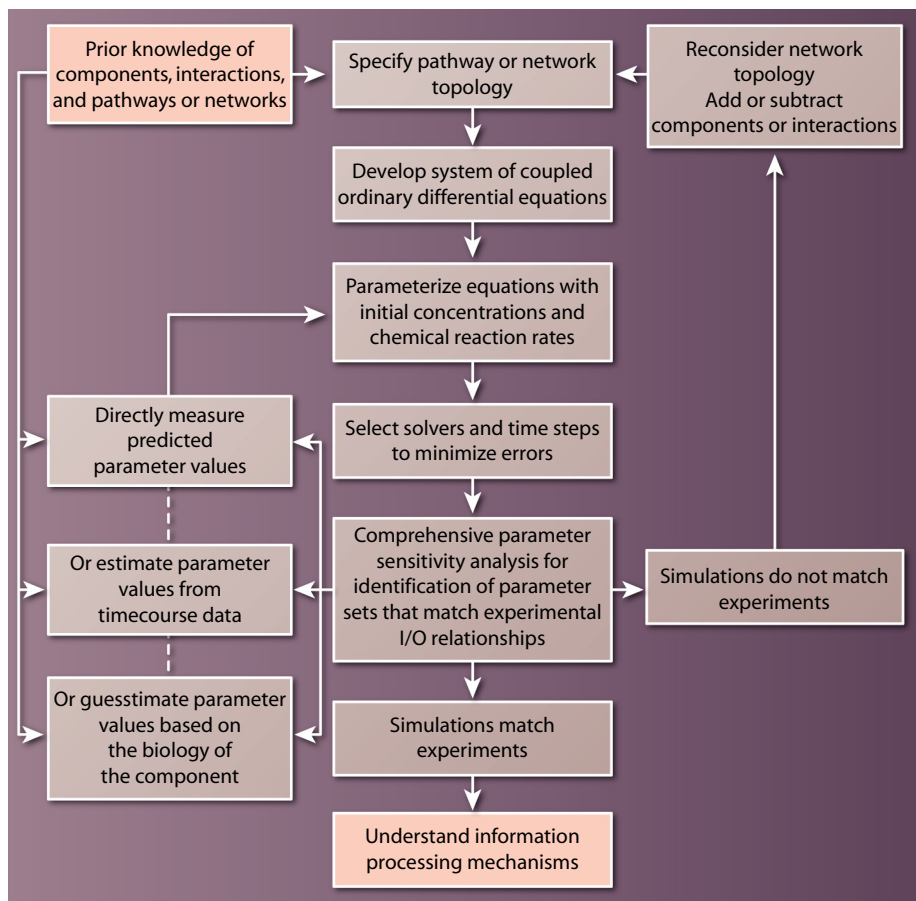


Fig. 1. Key steps in the development of detailed dynamical models to understand information-processing mechanisms within regulatory networks that control I/O relationships. In addition to the steps shown here, it is essential that all models that are based on chemical kinetics conform to microscopic reversibility principles. Once topology and parameters for initial conditions are set without bias (that is, prior expectations for systems behavior), simulations should be run without altering these conditions. Parameter variation should be systematic, unbiased, and as comprehensive as possible. Topology should be altered on the basis of biological knowledge only when simulation outputs do not match experimental observations for I/O relationships; the entire process is iterative between model development, simulations, and experiments.

obtain a desired simulation output. This is model “tweaking,” an impermissible operation that can lead to spherical cow models and provides no mechanistic insight into systems behavior.

Good models have minimal computational error. Smaller time steps will reduce errors at the expense of computational time. If there are parameters in very different temporal domains, issues of stiffness (that is, numerical instabilities) can arise, leading to errors in calculation. Commercial software suites have solvers that deal with stiffness.

An appropriate way to determine the best parameter sets for the simulations is to conduct an unbiased parameter variation exercise. Massive parameter sweeps that until recently were not feasible are now quite inexpensive. Multiple parameter sets may yield dynamics that match the experimental observations. Such findings often provide mechanistic insight regarding system redundancies and robustness that can produce phenotypic convergence (10). Sometimes no parameter sets will produce simulations that match experiments, or

only biologically implausible parameters yield simulations that match experiments. In such cases, one should reexamine the topology of the model and add additional details (2). If the steps outlined in Fig. 1 are followed and the models are integrated with experiments, the ODE-based models can provide deep mechanistic insights into biological regulation.

References and Notes

1. A. Einstein, On the method of theoretical physics. *Philos. Sci.* **1**, 163–169 (1934).
2. E. U. Azeloglu, S. V. Hardy, N. J. Eungdamrong, Y. Chen, G. Jayaraman, P. Y. Chuang, W. Fang, H. Xiong, S. R. Neves, M. R. Jain, H. Li, A. Ma'ayan, R. E. Gordon, J. C. He, R. Iyengar, Interconnected network motifs control podocyte morphology and kidney function. *Sci. Signal.* **7**, ra12 (2014).
3. S. R. Neves, Obtaining and estimating kinetic parameters from the literature. *Sci. Signal.* **4**, tr8 (2011).
4. D. J. Carty, R. T. Premont, R. Iyengar, Quantitative immunoblotting of G-protein subunits. *Methods Enzymol.* **195**, 302–315 (1991).
5. A. E. Merrill, J. J. Coon, Quantifying proteomes and their post-translational modifications by stable isotope label-based mass spectrometry. *Curr. Opin. Chem. Biol.* **17**, 779–786 (2013).
6. S. R. Neves, P. Tsokas, A. Sarkar, E. A. Grace, P. Rangamani, S. M. Taubenfeld, C. M. Alberini, J. C. Schaff, R. D. Blitzer, I. I. Moraru, II, R. Iyengar, Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks. *Cell* **133**, 666–680 (2008).
7. L. R. Castro, N. Gervasi, E. Guiot, L. Cavellini, V. O. Nikolaev, D. Paupardin-Tritsch, P. Vincent, Type 4 phosphodiesterase plays different integrating roles in different cellular domains in pyramidal cortical neurons. *J. Neurosci.* **30**, 6143–6151 (2010).
8. U. S. Bhalla, R. Iyengar, Emergent properties of networks of biological signaling pathways. *Science* **283**, 381–387 (1999).
9. P. Rangamani, M. A. Fardin, Y. Xiong, A. Lipshtat, O. Rossier, M. P. Sheetz, R. Iyengar, Signaling network triggers and membrane physical properties control the actin cytoskeleton-driven isotropic phase of cell spreading. *Biophys. J.* **100**, 845–857 (2011).
10. A. A. Prinz, D. Bucher, E. Marder, Similar network activity from disparate circuit parameters. *Nat. Neurosci.* **7**, 1345–1352 (2004).

Acknowledgments: This work was supported by NIH grants R01-GM54508 and Systems Biology Center grant P50-GM071558.

10.1126/scisignal.aab0880

Citation: E. U. Azeloglu, R. Iyengar, Good practices for building dynamical models in systems biology. *Sci. Signal.* **8**, fs8 (2015).

Good practices for building dynamical models in systems biology

Evren U. Azeloglu and Ravi Iyengar

Sci. Signal. **8** (371), fs8.
DOI: 10.1126/scisignal.aab0880

ARTICLE TOOLS	http://stke.sciencemag.org/content/8/371/fs8
RELATED CONTENT	http://stke.sciencemag.org/content/sigtrans/2/64/eg3.full http://stke.sciencemag.org/content/sigtrans/8/371/eg4.full http://stke.sciencemag.org/content/sigtrans/8/371/rs2.full http://stke.sciencemag.org/content/sigtrans/8/374/ra40.full http://stm.sciencemag.org/content/scitransmed/7/294/294ed7.full http://stke.sciencemag.org/content/sigtrans/9/443/mr1.full
REFERENCES	This article cites 10 articles, 4 of which you can access for free http://stke.sciencemag.org/content/8/371/fs8#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the [Terms of Service](#)

Science Signaling (ISSN 1937-9145) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Signaling* is a registered trademark of AAAS.