and inhibiting muscles and neurons, and show how controlled spatial and temporal stimuli will advance understanding even in some well-described neural circuits. For example, it is well known that anterior touch causes reversals in *C. elegans,* whereas posterior touch causes forward acceleration, consistent with a simple escape response. Until now, it has been difficult to quantitatively assess the competition between the two, but by varying the intensity between anterior and posterior illumination, Stirman *et al.*[3] provide some insight into how the two stimuli are integrated into a behavioral output.

Given that *C. elegans* locomotion is thought to have an important proprioceptive component, experiments on restrained worms are unlikely to provide a complete picture of how the motor circuit generates (and modulates) the undulatory waves necessary for crawling on surfaces or swimming through fluids. Experiments using traditional electrophysiology in freely moving worms are extremely challenging, but optogenetics presents two alternatives. Using calcium imaging to monitor muscle and motor neuron activity in freely behaving worms could provide critical data on how the circuit operates during normal locomotion. Even in the absence of knowledge of the underlying neural activity, targeted interventions in freely moving worms can also be revealing, as Leifer *et al.*[2] demonstrate by silencing only a portion of the ventral nerve cord motor neurons or body-wall muscles, which does not change the anterior body wave but prevents its posterior propagation past the silenced region. More extensive data and their direct comparison with competing theories of the motor circuit will be an important advance.

These two papers[2,3] highlight how genetics-based tools are revolutionizing the study of behavioral neuroscience in simple animal models. Engineered proteins targeted to specific neurons have been used to map neural circuits[8], monitor neural activity during behavior[9,10] and now to precisely manipulate neuronal function at the single-cell level. With the development of automated tools to effectively exploit these reagents, the possibility of understanding the complete neural circuitry of *C. elegans* has become a realistic goal.

The lessons learned from these studies, as well as the technologies used to obtain them, should provide important stepping stones toward a similar mechanistic understanding of more complicated brains, including those of flies, fish and ultimately humans.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Miesenbock, G. *Science* **326**, 395–399 (2009).
2. Leifer, A.M., Fang-Yen, C., Gershow, M., Alkema, M.J. & Samuel A.D.T. *Nat. Methods* **8**, 147–152 (2011).
3. Stirman, J.N. *et al. Nat. Methods* **8**, 153–158 (2011).
4. White, J.G., Southgate, E., Thomson, J.N. & Brenner, S. *Phil. Trans. Royal Soc. B* **314**, 1–340 (1986).
5. Guo, Z.V., Hart, A.C. & Ramanathan, S. *Nat. Methods* **6**, 891–896 (2009).
6. Waggoner, L.E., Zhou, G.T., Schafer, R.W. & Schafer, W.R. *Neuron* **21**, 203–214 (1998).
7. Pierce-Shimomura Morse, T.M. & Lockery, S.R. *J. Neurosci.* **19**, 9557–9569 (1999).
8. Feinberg, E.H. *et al. Neuron* **57**, 353–363 (2008).
9. Kerr, R. *et al. Neuron* **26**, 583–594 (2000).
10. Chronis, N., Zimmer, M. & Bargmann, C.I. *Nat. Methods* **4**, 727–731 (2007).

New approaches to modeling complex biochemistry

John A Bachman & Peter Sorger

Combining rule-based descriptions of biochemical reactions with agent-based computer simulation opens new avenues for exploring complex cellular processes.

A major challenge in modeling complex genetic and biochemical circuits is turning imprecise 'word models' (text and drawings) into precise mathematical statements suitable for computational analysis. Many biochemical processes common to eukaryotic signaling networks are remarkably difficult to describe and simulate with rigor and precision. Assembly of multiprotein complexes or multisite post-translational modification can create a very large number of distinct molecular species from interactions involving only a handful of gene products. Such 'combinatorial complexity' exceeds the ability of conventional modeling methods to enumerate species and track their dynamics, leading to *ad hoc* assumptions about which types of interactions matter and which do not.

The network-free stochastic simulator (NFSim), described in this issue of *Nature Methods*[1], is a new entry in an evolving class of general-purpose computational tools that addresses the challenge of combinatorial complexity by combining a rule-based representation with agent-based simulation. In a rule-based representation, patterns of interaction among proteins and other biomolecules are specified using a specialized computer language. By analogy to organic chemistry, only information important for a specific reaction is included in a rule; all nonparticipating structures are left unspecified. A rule-based approach differs from conventional reaction-centric models in which equations (commonly ordinary differential equations (ODEs)), one for every possible molecular species or complex, are enumerated to describe the time dynamics of the system. For combinatorially complex systems, equation-based models are hard to error-check, extend and reuse, in contrast to rule-based models, which are concise, comprehensible and easily extended. Research to date suggests that rule-based approaches enable simulation and analysis of classes of complex reactions that would otherwise be intractable[2].

Available rule-based modeling tools are differentiated by the way they perform simulations (**Fig. 1**). In BioNetGen[3], the technology on which NFSim is built, a rule set is typically used to generate the full network of all possible chemical reactions, which are then simulated using ODEs or the stochastic simulation algorithm of Doob-Gillespie[4]. However, for systems with a sufficiently high degree of combinatorial complexity, enumeration of all chemical reactions becomes prohibitively costly (as the authors[1] of NFSim demonstrate, this can occur in simulations of multisite phosphorylation with as few as six to eight sites).

To overcome this limitation, NFSim represents the system as a finite pool of interacting molecules, or 'agents', and the simulation unfolds stochastically by the repeated action of the rules on the pool of molecules. Because the rules themselves are used to direct the simulation, generation of the full reaction network is unnecessary. In this respect, NFSim adopts

John A. Bachman and Peter Sorger are in the Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, USA.
e-mail: peter_sorger@hms.harvard.edu

an approach pioneered by Kappa, an alternative but closely related rule-based modeling tool that also uses a network-free simulation algorithm[5]. The NFSim authors[1] demonstrate increased simulation efficiency over Kappa and other existing rule-based tools for a small set of example models, an improvement attributed to extensive software optimization.

Although rules are more intuitive and concise than systems of differential equations, existing rules have limited expressive power. For example, if the activity of a multisubunit enzyme varies with conformation, a set of similar rules is needed for each distinct configuration of the enzyme. NFSim introduces 'local functions', a feature in which the rate of a reaction rule can depend on the properties of a protein complex in arbitrarily definable ways reflective of empirical data. The authors' application of NFSim to the study of bacterial chemotaxis is a useful example of the benefits of this approach[1] as their model could not be described using either a strictly ODE- or rule-based approach.

Quantitative information on even the best-understood biochemical pathways is poor, and practical models of real biology usually include both mechanistic and nonmechanistic aspects. The functional rate laws in NFSim allow users to encode the kinetic relationships between model species in terms of arbitrary mathematical functions, including Boolean functions and conditional expressions (for example, 'if inhibitor is off, turn synthesis of your favorite gene on') that often correspond to current levels of experimental understanding.

The ability of NFSim to blend a mechanistic representation of molecular interactions with mechanistically agnostic expressions enables 'modeler-driven coarse-graining'. For example, a precise model of cyclin degradation through the cell cycle could be embedded in a rule-based model without including detailed information on the biochemistry of protein ubiquitination; the actual degradation machinery would be simply represented as an abstract degradation equation. This approach differs from what could alternatively be called 'model-driven coarse-graining', also known as model reduction, in which formal, analytical or numerical techniques are applied to a detailed model to derive a set of simplifying assumptions that allow the dynamics of the original model to be retained.

The efficiency and multiresolution modeling features offered by NFSim, in combination with its accompanying Matlab tools for parameter estimation and analysis of model output, make it a valuable general-purpose addition to the toolbox of the biological modeler. However, it is not without limitations: whereas the agent-based simulation approach adopted by NFSim tames the problem of combinatorial complexity by obviating the need to generate the entire reaction network, this approach, as all agent-based approaches, is computationally very costly when the number of molecules, or more precisely the number of molecular interactions, is high (greater than ~10^4–10^6 molecules of each species, using current technology). Also, the use of NFSim's features for multiresolution modeling may preclude the application of powerful model checking[6] and model reduction algorithms that have recently been described for ODE-based[7] and rule-based[8,9] models. Additional research is also required to determine the extent to which the flexible, hybrid models that can be encoded in NFSim remain analytically tractable.

Rapid progress in the development of NFSim, Kappa, little b[10] and related metalanguages for representing biochemical reactions are the harbingers of an entirely new way of representing and studying cellular networks. We predict that within a decade these methods will be mainstream components of modern quantitative biology.

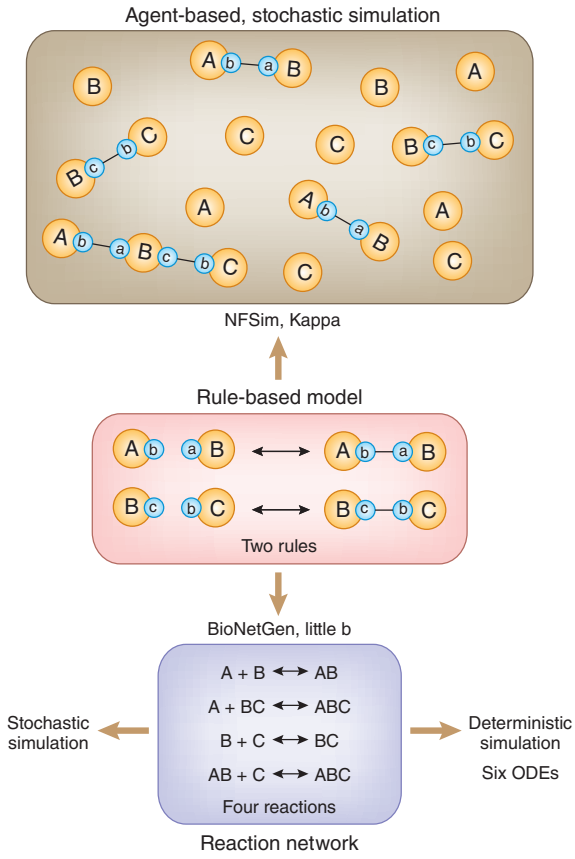

Figure 1 | Modeling the assembly of a three-protein complex, ABC, using a rule-based approach. A and C bind B at independent sites (sites a and c, respectively), allowing the assembly process to be represented using only two rules. When the chemical reaction network is generated, these rules expand to four chemical reactions and six ODEs. In systems with sufficiently great combinatorial complexity, generating the full reaction network becomes impossible. NFSim allows the assembly process to be simulated using a pool of interacting molecules, which are represented computationally as 'agents'. Changes in the state and connectivity of the agents are determined by the rules, and quantities of relevant molecular species are tracked during simulation.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Sneddon, M.W., Faeder, J.R. & Emonet, T. *Nat. Methods* **8**, 177–183 (2011).
2. Hlavacek, W.S. *et al. Sci. STKE* **344**, re6 (2006).
3. Faeder, J.R., Blinov, M.L. & Hlavacek, W.S. *Methods Mol. Biol.* **500**, 113–167 (2009).
4. Gillespie, D.T. *J. Comput. Phys.* **22**, 403–434 (1976).
5. Danos, V., Feret, J., Fontana, W. & Krivine, J. *Lect. Notes Comput. Sci.* **4807**, 139–157 (2007).
6. Danos, V., Feret, J., Fontana, W. & Krivine, J. *Lect. Notes Comput. Sci.* **4905**, 83–97 (2008).
7. Conzelmann, H., Fey, D. & Gilles, E.D. *BMC Syst. Biol.* **2**, 78 (2008).
8. Borisov, N.M., Chistopolsky, A.S., Faeder, J.R. & Kholodenko, B.N. *IET Syst. Biol.* **2**, 342–351 (2008).
9. Feret, J., Danos, V., Krivine, J., Harmer, R. & Fontana, W. *Proc. Natl. Acad. Sci. USA* **106**, 6453–6458 (2009).
10. Mallavarapu, A., Thomson, M., Ullian, B. & Gunawardena, J. *J. R. Soc. Interface* **6**, 257–270 (2009).