Rule-Based Modeling of Biochemical Systems with BioNetGen

James R. Faeder[1], Michael L. Blinov[2], and William S. Hlavacek[3]

***Short Abstract* — Rule-based modeling involves the representation of molecules as structured objects and molecular interactions as rules for transforming the attributes of these objects. The approach allows systematic incorporation of site-specific details about protein-protein interactions into a model for the dynamics of a signal-transduction system, as well as other applications. The consequences of protein-protein interactions are difficult to specify and track with a conventional modeling approach because of the large number of protein phosphoforms and protein complexes that these interactions potentially generate. Here, we focus on how a rule-based model is specified in the BioNetGen language (BNGL) and how a model specification is analyzed using the BioNetGen software tool. We also discuss new developments in rule-based modeling that should enable the construction and analysis of comprehensive models for signal transduction pathways and similarly large-scale models for other biochemical systems.**

***Keywords* — mathematical modeling; combinatorial complexity; protein-protein interactions; signal transduction.**

THE components of the BioNetGen software package are shown in Fig. 1 (1, 2). There are two basic ways to construct a BioNetGen model: either write an input BNGL file using a text editor or the on-line Virtual Cell interface (see http://vcell.org/bionetgen) or use the graphical model editor called RuleBuilder, which can output a BNGL file. BNGL files are read by the BNG2 engine, written in Perl, which provides a range of translation and simulation capabilities. A model consists of five basic elements: parameters, molecule definitions, rules, observables, and actions. Molecules are the basic building blocks of a BioNetGen model, and are used to represent proteins and other structured biological molecules, such as metabolites, genes, or lipids. Molecules may contain components, which represent the functional elements of molecules, and may bind other components, either in the same molecule or another molecule. Components may be associated with state variables, which take on a finite set of possible values that may represent conformational or chemical states of a component, e.g., tyrosine phosphorylation. Rules define how molecules in the system can interact, and thus either explicitly or implicitly define the biochemical reaction network. Observables define model outputs, and actions specify the operations that are to be carried out either to generate or simulate a network.

Work on BioNetGen has been supported by NIH grants GM035556, RR18754, and GM076570 and DOE contract DE-AC52-06NA25396.

[1]Department of Computational Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA. E-mail: faeder@pitt.edu

[2]Center for Cell Analysis and Modeling, University of Connecticut Health Center, Farmington, CT. E-mail: blinov@uchc.edu

[3]Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM and Department of Biology, University of New Mexico, Albuquerque, NM . E-mail: wish@lanl.gov

Once a model specification is processed by BNG2, models can be exported in one of a variety of formats, including MATLAB M-files and Systems Biology Markup language (SBML) (3). BNG2 can also generate a network of species and reactions for simulation using either an ODE solver or Gillespie's stochastic simulation algorithm (4). The dashed line in Fig. 1 indicates that the network generation and stochastic simulation engines are coupled so that for large reaction networks species and reactions can be generated on-the-fly as needed (5, 6). Recently, particle-based simulation methods (7, 8) have been developed that avoid the computational costs associated with explicit generation of the reachable network of species and interactions, which can become a major bottleneck for the simulation of even moderate-sized rule-based models (9). The programs DYNSTOC and NFSim, being developed in collaboration with Joshua Colvin and Richard G. Posner at Translational Genomics Research Institute and Michael Sneddon and Thierry Emonet at Yale University, will be able to read an XML encoding of any BioNetGen model and perform a particle-based simulation (10). This capability will allow simulation of rule-based models that comprehensively describe multiple coupled signaling pathways within a cell.

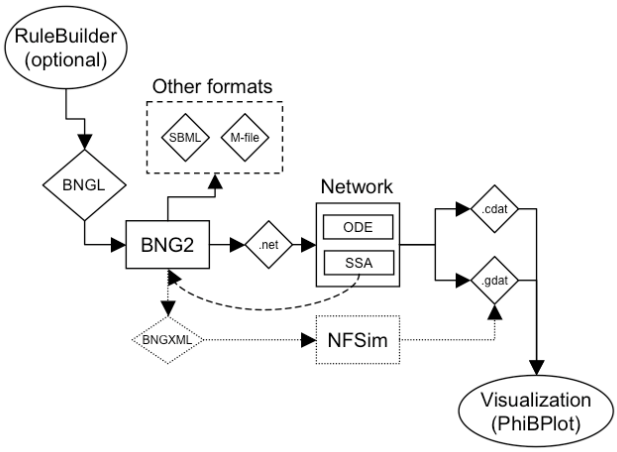

Fig. 1. BioNetGen architecture (see **http://bionetgen.org**).

REFERENCES

[1] Blinov, M. L., *et al.* (2004) *Bioinformatics* **20,** 3289-91.
[2] Faeder, J. R., Blinov, M. L., and Hlavacek, W. S. (2008) *in Methods in Molecular Biology: Systems Biology* (Maly, I. V., Ed.), Humana Press, Totowa, NJ.
[3] Hucka, M., *et al.* (2003) *Bioinformatics* **19,** 524-31.
[4] Gillespie, D. T. (1976) *J. Comp. Phys.* **22,** 403-34.
[5] Lok, L., and Brent, R. (2005) *Nat. Biotechnol.* **23,** 131-36.
[6] Faeder, J. R., *et al.* (2005) *Complexity* **10,** 22-41.
[7] Danos, V., *et al.* (2007) *Lect. Notes Comput. Sci.* **4807,** 139-57.
[8] Yang, J., *et al.* (2007) *arXiv:0712.3773*.
[9] Hlavacek, W. S., *et al.* (2006) *Sci. STKE* **2006,** re6.
[10] Sneddon, M. W., *et al.* (2008) *Second Annual q-bio Conference on Cellular Information Processing*, Santa Fe, NM.