James R. Faeder
Department of Computational and Systems Biology, University of Pittsburgh
e-mail: faeder@pitt.edu

Rule-Based Modeling of Molecular and Cellular Processes

Cells possess complex sensory mechanisms that are governed by the biochemical interactions of proteins. A typical signaling protein possesses multiple interaction sites, whose activity can be modified both by direct chemical modification (termed "post-translational modification") and by the effects of modification or interaction at other sites (termed "allostery"). This complexity at the protein level leads to combinatorial complexity at the level of signaling networks - an individual protein has many potential states of modification and interaction, which gives rise to an ever-multiplying set of possible complexes and poses a major barrier to traditional methods of modeling and simulation [1]. Here, I will review major developments in modeling, both from my work and that of others, that have helped to tame these difficulties.

The need to simplify the development of signal transduction models and to expand their scope has motivated the development of rule-based modeling languages, such as BioNetGen [2] and Kappa [3], which provide a rich and yet concise description of signaling proteins and their interactions. Their success is demonstrated by the growing community of users and the substantial number of models that have been developed and published. While greatly facilitating the translation of knowledge about signaling biochemistry into models, however, rule-based languages do not directly address the combinatorial challenges involved in the simulation of such models, which arise from the size of the reaction network implied by the rules. For these, new agent-based stochastic simulation methods have been developed for rule-based models with computational requirements that are independent of the number of possible species (i.e., complexes) and proportional to the number of molecules (e.g., proteins) being simulated. In addition, general and efficient implementations are now available that enable the rapid simulation of rule-based models of virtually any complexity. NFsim is one such simulator that stands out because of its efficiency and the ability to course-grain complex interactions through the incorporation of high-level functions into the rate laws that govern rule application [4]. The use of stochastic simulations, however, exacerbates the already difficult problems common to all complex models of relating model parameters to model behavior and of estimating parameter values based on experimental observations and data. For these, new statistical model checking algorithms and tools have been developed that allow model properties to be determined from a minimal number of simulation runs [5]. Taken together, rule-based modeling languages and their associated tools address the issue of combinatorial complexity in cell regulatory networks, allowing the development, simulation, and analysis of models with unprecedented scope and detail and, we hope, predictive capability.

References

[1] W. S. Hlavacek and J. R. Faeder (2009) *Sci. Signaling* **2** pe46.
[2] J. R. Faeder, M. L. Blinov, and W. S. Hlavacek (2009) *Meth. Mol. Biol.* **500**, 113–167.
[3] V. Danos, J. Feret, W. Fontana, and J. Krivine (2007) *Lect. Notes. Comput. Sci* **4807**,139-157.

[4] M. W. Sneddon, J. R. Faeder, and T. Emonet (2011) *Nature Methods* **8**, 177–183.
[5] E. M. Clarke, et al. (2008) *Lect. Notes. Comput. Sci.* **5307**, 231-250.