Simulation of systems marked by combinatorial complexity

Michael Monine[1,2,*], Jin Yang[1], Ambarish Nag[1], Byron Goldstein[1], Richard G. Posner[3], Joshua Colvin[3], William S. Hlavacek[1,2], James R. Faeder[4]
[1]Theoretical Biology and Biophysics Group, Theoretical Division, [2]Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545
[3]Translational Genomics Research Institute, Phoenix, AZ 85004
[4]Department of Computational Biology, University of Pittsburgh School of Medecine, Pittsburgh, PA 15261
*email: mmonine@lanl.gov

Background

Signal transduction generally involves multivalent protein-protein interactions that can produce myriad protein complexes and post-translational modifications. By using reaction rules to represent such interactions and their consequences one can generate the whole reaction network implied by those interactions, either before a simulation of the reaction kinetics or during the course of a simulation on the fly [1,2]. However, reaction networks implied by sets of rules are typically large and challenge conventional simulation procedures. Even network generation on the fly in the course of a simulation can fail. Here, to address this problem, we propose a network-free simulation technique based on the well-known Gillespie method [3,4].

Stochastic Simulation Approach

The main feature of our modification of the Gillespie method is a separation of reactants into interaction classes that are selected by a reaction rule. Each class does not necessarily contain only one type of molecule; rather, it can be a group of different molecules that can participate in reactions with *the same kinetics*. The proposed method allows for a speed up of stochastic simulations of multicomponent systems by optimizing calculation of interaction rates and sampling of reactions.

Reactants are modeled as software objects with multiple binding sites. Reactants can share their binding sites to create bonds. We build a graph of rules, where each rule is associated with a particular reaction type i and gives a map of transformations of reactant classes, $\mathcal{M}_i$. If we define reactant class $\mathbf{A}$ as a vector of addresses of reactant sites of type A, then each reaction with A will cause a shift of reactant sites (picked at random) from $\mathbf{A}$ to another class or group of classes. Thus, $\mathcal{M}_i$ represents an acyclic graph of nodes ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, etc.) connected by directed edges, where edges denote directions, in which reactant site addresses should be moved. Figure 1 shows an example of class update rules for a system of trivalent ligands and bivalent receptors. Once we keep track of all classes, the fastest way to calculate reaction rates (a_i) is to use the following relation:

$$a_i = \sum_{j=1}^{J_i} a_{ij} = \begin{cases} k_i N_A & \text{for unimolecular reaction } A \to C, \\ k_i N_A N_B & \text{for bimolecular reaction } A + B \to C, \\ k_i N_A (N_A - 1)/2 & \text{for bimolecular reaction } A + A \to A_2, \end{cases} \quad (1)$$

where a_{ij} is a microscopic rate, J_i is the number of ways to react, k_i is ith reaction rate constant, and N_A, N_B denote the numbers of reactants in each class. After we have the full set of reaction rates, sampling of reactions becomes straightforward.

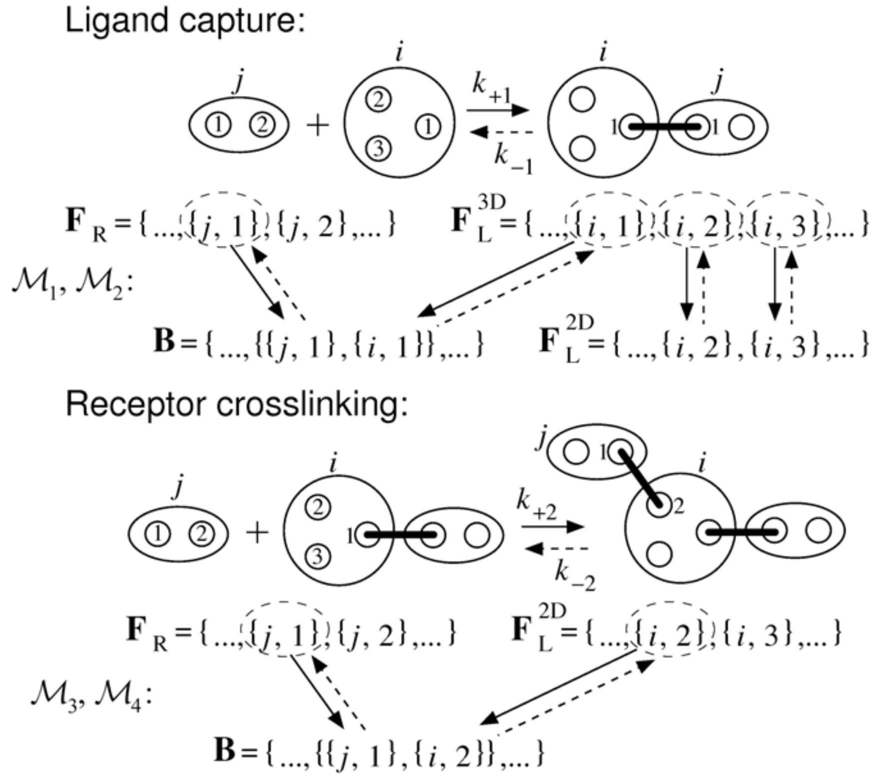

Figure 1. Four reaction rules for the trivalent ligand – bivalent receptor system. Four reaction types are distinguished: ligand binding from the solution (with rate constant k_{+1}), binding of a surface-tethered ligand (k_{+2}), and dissociation ($k_{-2} = k_{-1} = k_{\mathrm{d}}$). The occupancy of binding sites is recorded in four reactant classes: lists $\mathbf{F}_{\mathrm{L}}^{3\mathrm{D}}$, $\mathbf{F}_{\mathrm{L}}^{2\mathrm{D}}$, $\mathbf{F}_{\mathrm{R}}$ contain addresses of unoccupied binding sites on free ligands, surface-tethered ligands and receptors, respectively; list $\mathbf{B}$ contains pairs of addresses of ligand-receptor bound sites. After two reactive sites on ligand and receptor are selected, the corresponding addresses are moved from the lists of reactants to the list of bonds. Updates associated with the dissociation reaction are performed similarly.

Application

We apply our model to simulate binding of a haptenated fluorescent multivalent ligand, dinitrophenol-coupled phycoerythrin (DNP-PE), to bivalent high affinity receptors for IgE, anti-DNP IgE-FcεRI complexes [5]. In the flow cytometry experiments, only cumulative data on ligand binding and receptor aggregation can be obtained [6-8]. Therefore, the simulations, in which dynamics of aggregates can be observed, are of great benefit to understanding the kinetics of aggregation. We also simulate interactions of intracellular proteins, such as LAT, SOS1 and GRB2 [9]. Crosslinking of LAT produces clusters and plays a functional role in TCR-mediated signaling. The simulation results support the experimental finding that to form clusters, the GRB2-SOS1 complex should have the composition 2:1 and LAT should have three docking sites for GRB2.

References

[1] Hlavacek W.S. et al., (2006) *Sci. STKE*, **2006**, re6.
[2] Blinov M.L. et al., (2004) *Bioinformatics*, **20**, 3289.
[3] Gillespie D.T., (1977) *J. Phys. Chem.*, **81**, 2340; Gillespie D.T., (2007) *Ann. Rev. Phys. Chem.*, **58**, 35.
[4] Cao Y., Petzold L., (2004) *J. Chem. Phys*., **121**, 4059.
[5] Holowka D., et al., (2007) *Immunol. Rev*., **217**, 269.
[6] Posner R.G. et al., (2004) *Biochemistry*, **43**, 11352.
[7] Hlavacek W.S. et al., (1999) *Biophys. J*., **76**, 2421.
[8] Tessema M. et al., (2006) *Cytometry*, **69**, 326.
[9] Houtman J.C.D. et al., (2006) *Nat. Struct. Mol. Biol*., **13**, 798.