LA-UR- 08-7821

Approved for public release; distribution is unlimited.

Title:	Modeling Multivalent ligand-receptor interactions with steric constraints on configurations of cell-surface receptor aggregates
Author(s):	Michael Monine, Los Alamos National Laboratory Richard Posner, Translation Genomics Research Institute Paul Savage, Brigham Young University James Faeder, University of Pittsburgh William Hlavacek, Los Alamos National Laboratory
Intended for:	PLOS Computational Biology

Form 836 (7/06)

Modeling multivalent ligand-receptor interactions with steric constraints on configurations of cell-surface receptor aggregates

Michael I. Monine[1], Richard G. Posner[2,3],
Paul B. Savage[4], James R. Faeder[5], and William S. Hlavacek[1,6,*]

[1]Theoretical Biology and Biophysics Group, Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[2]Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 85004, USA

[3]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011, USA

[4]Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602, USA

[5]Department of Computational Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA

[6]Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA

*Corresponding author. Address: Mail Stop K710, Los Alamos, NM 87545, Tel: 505-665-1355, Fax: 505-665-3493, E-mail: wish@lanl.gov

Abstract

Signal transduction generally involves multivalent protein-protein interactions, which can produce various protein complexes and post-translational modifications. The reaction networks that characterize these interactions tend to be so large as to challenge conventional simulation procedures. To address this challenge, a kinetic Monte Carlo (KMC) method has been developed that can take advantage of a model specification in terms of reaction rules for molecular interactions. A set of rules implicitly defines the reactions that can occur as a result of the interactions represented by the rules. With the rule-based KMC method, explicit generation of the underlying chemical reaction network implied by rules is avoided. Here, we apply and extend this method to characterize the interactions of a trivalent ligand with a bivalent cell-surface receptor. This system is also studied experimentally. We consider the following kinetic models: an equivalent-site model, an extension of this model, which takes into account steric constraints on the configurations of receptor aggregates, and finally, a model that accounts for cyclic receptor aggregates. Simulation results for the equivalent-site model are consistent with an equilibrium continuum model. Using these models, we investigate the effects of steric constraints and the formation of cyclic aggregates on the kinetics and equilibria of small and large aggregate formation and the percolation phase transition that occurs in this system.

Introduction

Attempts to simulate cell signaling systems must address the problem of combinatorial complexity (1), the large numbers of chemical species and reactions that arise from molecular interactions in these systems. This complexity strains conventional model-specification and simulation approaches, which take a reaction network (i.e., a list of reactions) as input. An alternative approach to model specification is to use rules to represent interactions and their consequences (2). In this approach rules represent generalized reactions, or reaction classes, which arise from molecular interactions (3), and the rules serve as generators of chemical species and reactions. A rule defines a transformation and the molecular properties and context necessary for a molecule or molecular complex to undergo the transformation. When the contextual requirements are limited, as when the local environment of a site determines its reactivity, a set of rules can provide a compact high-level specification of a model for a system characterized by a large underlying chemical reaction network.

Several software tools have been developed to enable rule-based modeling, including StochSim (4), BioNetGen (5, 6), Moleculizer (7), Simmune (8), and DYNSTOC (9). Three approaches have emerged for the simulation of rule-based models. In the "generate-first" approach, rules are iteratively applied before a simulation to generate a complete list of the possible species and reactions, followed by simulation of the network using either ODEs or Gillespie's Stochastic Simulation Algorithm (SSA) (2). This approach, which is efficient for small-to-medium sized networks (10^4 species or less), becomes computationally intractable as network size grows. Simulation via this approach can be made more tractable by introducing exact or approximate reductions of a model implied by a given set of rules (10). The second approach is the "on-the-fly" method, which both speeds simulation and limits network generation from rules, which is expensive (6, 7). On-the-fly simulation takes advantage of a feature of methods for simulating discrete-event reaction kinetics, such as the SSA (11, 12), that allows lazy processing of reaction rules. Namely, to advance a simulation using such a method, we require only knowledge of the species currently populated and their potential reactions, not all potential species and reactions. However, in some cases, even the partial reaction network required for on-the-fly simulation can be exceedingly large (2, 13). The third simulation approach avoids network generation altogether, and we will refer to it as the "network-free" approach. A variation of this approach, which is implemented in the StochSim software tool (4), was introduced by Morton-Firth and Bray

(14). However, the method of Morton-Firth and Bray tends to be inefficient for systems marked by stiffness (i.e., systems with processes occurring on multiple time scales), and StochSim has limited rule-processing capabilities.

Recently, Danos, Krivine, and co-workers have presented a more efficient formulation of the network-free approach (15). Instead of individual reactions being fired in the simulation procedure, reaction rules are fired, followed by selection of reactants and rule-defined generation of products. Although conceptually straightforward, this extension introduces a requirement to track individual molecules and their reactive sites, so that the ability of molecules to participate in the different reaction classes defined by rules can be determined along with the cumulative propensities of the different reaction types. Overall efficiency of the method depends on careful implementation of the bookkeeping procedures needed to accomplish these tasks.

Here, we present an application and extension of the rule-based kinetic Monte Carlo (KMC) simulation method of Yang et al. (13), which extends the method of Danos et al. (15). The difference between this method and the method of Danos et al. is the introduction of a step in which the non-local context of potential reactants is examined to determine their reactivity. This step introduces null events (i.e., time steps in which no reaction or system updates occur). Introduction of null events has been shown to provide for efficient processing of rules for reaction types that depend on the non-local context of reactive sites (13), such as ring-closure reactions that produce cyclic molecular complexes. Here, we show how the method of Yang et al. (13) can be extended to account for geometric constraints on the formation of molecular complexes.

We develop a model of trivalent ligand and bivalent receptor interactions. In this system, which we also study experimentally, a synthetic molecule containing three symmetrically-arrayed hapten groups interacts with anti-hapten monoclonal IgE antibody bound to FcεRI, the high-affinity IgE receptor, on the surface of rat basophilic leukemia (RBL) cells. The system involves the following basic interactions: binding of a freely diffusing ligand from a solution phase to cell-surface receptors, crosslinking of receptors by surface-bound ligands, and dissociation of ligand-receptor bonds. These interactions are considered in a kinetic model studied by Yang et al. (13) that is based on the equivalent-site assumption (16) and corresponds to the equilibrium continuum model of Goldstein and Perelson (17). In this model, ligand-induced receptor aggregates are acyclic and have tree-like topology.

The Goldstein-Perelson model predicts a sol-gel percolation transition (PT) and formation, in the gel phase, of a superaggregate containing a significant fraction of the receptors present on a cell (17). Using the TLBR model we demonstrate how the dynamics of receptor aggregation depends on ligand concentration using values of kinetic parameters estimated in part from fitting equilibrium binding data.

We also develop an extension of the TLBR model to account for steric effects on receptor aggregates. Such effects can be expected to arise from the geometrical properties of interacting molecules. As a simplification, we assume that ligand and receptor properties are such that receptor aggregates form on a hexagonal lattice. In the extended TLBR model, the kinetics of binding is modified by the structures of interacting complexes. Overlaps in space, unfolding, and flips in 3D are prohibited in the model. Algorithmically, such constraints are taken into account by incorporating null-events in the simulation procedure. The overall system evolution rate is estimated based on the assumption that there is no dependence of reaction rates on aggregate structures, and then, for any particular binding reaction, accurate corrections to the rate are made to determine whether or not the event occurs. To account for steric constraints on molecular structures of ligand-receptor aggregates, we track the connectivities and rotation angles of molecules in aggregates. Note that, although the spatial configurations of molecules are tracked, the extended TLBR model is still based on the assumption of a well-mixed reaction compartment.

Finally, we demonstrate that the influence of steric effects on the overall kinetics of binding and aggregate formation can be effectively lumped into binding rate constants by introducing an empirical binding probability that depends only on the sizes of interacting aggregates. The parameters in a fitting function for this probability are estimated from simulations using the extended TLBR model.

Models

We assume a homogeneous distribution of bivalent receptors on a cell membrane. The medium surrounding a cell contains a homogeneous distribution of trivalent ligands. The system is well-mixed. Due to the multivalent nature of interacting ligands and receptors, the number of distinct ligand-induced receptor aggregates can be very large (2, 13, 17). Among the different possible structures of receptor aggregates, there are chains, trees and cyclic aggregates, as illustrated in Fig. 1. Here, we

consider three models for the system described above: the equilibrium-continuum model of Goldstein and Perelson (17), a kinetic version of this model, which has been called the TLBR model (13), and an extension of the TLBR model that incorporates steric constraints on the configurations of receptor aggregates. We also consider versions of the extended TLBR model with and without inclusion of cyclic aggregates.

Goldstein-Perelson model

The Goldstein-Perelson model (17) is based on the equivalent-site approximation (16, 27). Interactions between ligands and receptors include only two types of associative binding reactions: binding of a single site on a ligand in solution to a single cell-surface receptor site, and crosslinking of two receptors by a ligand that is already bound to a receptor site. These interactions are characterized by the following dimensionless parameters:

$$c = 3K_1C_\infty^L, \tag{1}$$

$$\beta = K_2C^R, \tag{2}$$

where the equilibrium association constant K_1 characterizes the binding of a ligand site in solution to a receptor site, the equilibrium crosslinking constant K_2 characterizes the binding of a tethered ligand site to a receptor site, C_∞^L is the concentration of free ligand in solution at equilibrium, and C^R is the total concentration of receptors. Note that a linear chain elongation reaction is assumed to be characterized by the same crosslinking constant K_2 as a reaction that forms a branch (i.e., a three-way junction). Examples of receptor aggregates included in the model are illustrated in Fig. 1(b,c).

The Goldstein-Perelson model accounts for all possible receptor-containing species except cyclic aggregates. A key element of this model is a partition function, q, which is a sum of the concentrations of all possible linear and branched aggregates. Derivation of q is presented elsewhere (17). This quantity is given by the following expression:

$$q = \frac{1}{\gamma q_0}\left(1 - \frac{1-(1-4\gamma q_0^2)^{3/2}}{6\gamma q_0^2}\right), \tag{3}$$

where q_0 is the partition function for all linear chains and $\gamma = 4\beta^2 c/(1+c)^3$. The quantity q_0 is given

by

$$q_0 = w/(1 - \delta w), \tag{4}$$

where $\delta = 4\beta c/(1+c)^2$ and w is the fraction of unaggregated receptors, which is given by

$$w = (1+c)^2 x, \tag{5}$$

where x is the fraction of free (unbound) receptors at equilibrium. Variables C_∞^L and x obey the following conservation equations:

$$x\frac{\partial q}{\partial x} - 1 = 0, \tag{6}$$

$$C_\infty^L \left(1 + C^R \frac{\partial q}{\partial C_\infty^L}\right) - C^L = 0, \tag{7}$$

where C^L is the total ligand concentration in the system. To find the values of C_∞^L and x, we substitute q given by Eq. 3 into Eqs. 6 and 7, and solve the resulting algebraic equations numerically using iterative methods, e.g., the Newton-Raphson method or secant method (28).

According to the Goldstein-Perelson model, the fraction of receptor aggregates containing n receptors and $m \geq 2$ branches, w_n, is given by

$$w_n = 2\delta^{n-1} w^n \sum_{m=0}^{m_{max}} \begin{pmatrix} n-1 \\ 2m \end{pmatrix} \begin{pmatrix} 2m \\ m \end{pmatrix} \frac{(\gamma\, \delta^{-2})^m}{(m+1)(m+2)}, \tag{8}$$

where $m_{max} = \text{floor}[(m-2)/2]$, the maximum number of branches in an aggregate of size n. The function floor[ ] takes the integer number of the argument. The model also predicts the fraction of receptors in the gel phase, f_g, which is given by

$$f_g = 1 - \frac{(1+c)}{\beta}[1 + (1 + 1/c)^{1/2}]. \tag{9}$$

In the sol phase, the system contains single cell-surface receptors and small aggregates. When the value of the dimensionless parameter β is below a critical value, $\beta_c = 4$, f_g is negative (undefined) for any value of c, and this condition corresponds to the sol phase. For $\beta > \beta_c$, f_g (Eq. 9) crosses the zero at two points, where $c = c_-$ and c_+, and the sol and gel phases coexist (results not shown). The

condition $c_- < c < c_+$ corresponds to the gel phase, which is characterized by the formation of an infinite-sized aggregate. Thus, Eq. 9 with $f_g = 0$ describes the boundary of a percolation transition (PT) in the space of the parameters β and c.

TLBR model

The TLBR model (13) is a kinetic version of the Goldstein-Perelson model. In the kinetic model, individual molecules are tracked explicitly. The data structures used for this purpose are presented below. Here, we introduce the assumptions underlying the model.

All sites of interacting molecules are assumed to be equivalent. As illustrated in panels (a) and (b) of Fig. 2, binding of a ligand from solution to cell surface-receptors takes place with single-site forward rate constant k_{+1}, and crosslinking of receptors occurs with single-site rate constant k_{+2}. All association reactions are reversible, and dissociation occurs with single-site rate constant k_{off}. In contrast with the Goldstein-Perelson model, which is a continuum model, the kinetic model accounts for a finite-sized system. The size of the simulated system is defined by the total numbers of ligands, N_L, and receptors, N_R. The kinetic rate constants k_{+1} and k_{+2} are related to K_1 and K_2 as follows: $K_1 = k_{+1}N_AV/k_{\text{off}}$ and $K_2 = k_{+2}N_AV/k_{\text{off}}$, where V and N_A denote the volume of the simulated system and Avogadro's number, respectively. The relationships between k_{+1}, k_{+2} and the dimensionless parameters c and β of the continuum theory (17) are given as follows:

$$c = 3k_{+1}N_{L,\infty}/k_{\text{off}}, \tag{10}$$

$$\beta = k_{+2}N_R/k_{\text{off}}, \tag{11}$$

where $N_{L,\infty}$ denotes the number of free ligands in solution at equilibrium.

Extended TLBR model

In the TLBR model, spatial configurations of receptor aggregates are assumed to be unimportant. This assumption is relaxed in the extended TLBR model, in which the rates of ligand-receptor interactions are assumed to be dependent on the configurations of receptor aggregates. In the extended TLBR model, ligand binding from solution and receptor crosslinking are treated in the same way as in the TLBR model, but final acceptance of reactions depends on a check of the spatial context of interacting

molecules. If a sampled reaction is not feasible for steric reasons, the reaction is rejected. To keep track of the configurations of ligand-receptor aggregates, we use data structures that record the orientation angles of molecules and their binding sites, as described below.

Assumptions made in our consideration of steric constraints include the following. The ligands and receptors are rigid. We assume that the ligand binding sites are symmetrically arrayed in a plane (18, 29), such that they lie at the vertices of an equilateral triangle. The receptor binding sites lie at the ends of a line segment. Thus, cell-surface reactions are treated as if they occur on a hexagonal lattice, as shown in Fig. 2(c) and Fig.3. Once a free ligand binds a surface receptor, the resulting complex only has freedom to rotate in two dimensions. Flips in three dimensions are not allowed. Two binding sites of different aggregates are not allowed to react if other parts of these aggregates overlap in space. In other words, steric clashes are prohibited.

In the extended TLBR model, we account for cyclic aggregates as follows. If ligand and receptor sites are unbound and positioned next to each other in the same aggregate, the rate constant for a binding reaction between these sites is defined as j_{+6}. Note that, because of our assumptions about the steric properties of ligands and receptors, the smallest cycle or ring contains six receptors, and only a single rate constant (j_{+6}) is required to characterize all ring-closure reactions. All dissociation reactions, including ring-opening reactions, are characterized by k_{off}.

Simulation outputs

To characterize receptor aggregation using the equivalent-site and extended TLBR models, we calculate the following quantities: the fraction of aggregates containing n receptors, w_n^s; the fraction of receptors in the gel phase, f_g^s; and the mean aggregate size, $\langle S \rangle$. The quantities w_n^s and f_g^s are analogous to w_n and f_g defined in the Goldstein-Perelson model (Eqs. 8 and 9). In simulations, the value of w_n^s is calculated as the number of aggregates of size n divided by N_R, and f_g^s is calculated as the number of receptors in the largest aggregate divided by N_R. The equilibrium values of w_n^s and f_g^s are evaluated as averages over time and/or multiple simulation runs. The mean aggregate size is given by

$$\langle S \rangle = \frac{\sum_{n=2}^{N_R} n\, w_n^s}{\sum_{n=2}^{N_R} w_n^s}. \tag{12}$$

Scaling of system size in simulations

The efficiency of simulating the equivalent-site and extended TLBR models depends on the system size, and therefore, to make simulations more tractable, we scale parameter values that are related to the system volume. We specify reference values for the volume (V^*) and the total numbers of receptors and ligands (N_R^* and N_L^*). These values are given in Table 1. We also specify reference values for association rate constants (k_{+1}^* and k_{+2}^*) that are consistent with the best-fit values of K_1 and K_2 and other reference values (Table 1). The TLBR model parameters are scaled such that $V = \chi V^*$, $N_R = \chi N_R^*$, $N_L = \chi N_L^*$, $k_{+1} = \chi^{-1} k_{+1}^*$ and $k_{+2} = \chi^{-1} k_{+2}^*$, where χ is a volumetric scaling factor. Note that the dissociation rate constant, k_{off}, does not depend on the system volume.

Materials and methods

Reagents

The ligand, a model antigen, was synthesized as described elsewhere (18). The ligand (compound 6a) is comprised of an Alexa-488 label and three symmetrically-arrayed 2,4-dinitrophenyl (DNP) groups. The effective receptor consists of FcεRI expressed on RBL-2H3 cells, tightly coupled to DNP-specific monoclonal IgE antibody. The average lifetime of an IgE-FcεRI complex is greater than 12 hours (31), which is much longer than the time scale of a binding experiment. The mouse monoclonal anti-DNP IgE coupled to FcεRI on RBL cells was isolated from hybridoma H1 26.82 by affinity purification (32). Isolation of IgE involved, in the final steps, ion exchange chromatography, to remove bound DNP-glycine, and gel filtration, to separate monomeric IgE from IgE aggregates.

Cells

RBL-2H3 cells (33) were grown adherent in 75 cm^2 flasks. Cell cultures, which were used typically five days after passage, were maintained at 37°C. Culture media consisted of MEM 1X with Earle's salts without glutamine (Gibco BRL), 20% fetal bovine serum (HyClone, Logan, UT), 1% glutamine, 1% v/v penicillin, and 1% v/v streptomycin (Gibco BRL). To harvest cells, we rinsed and then incubated the cells, for five minutes at 37°C, with trypsin-EDTA. Cells harvested for experiments were washed and resuspended in buffered salt solution (pH 7.7), which was freshly passed through a 0.22 micron

filter. Buffered salt solution (BSS) consisted of 135 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$ 1.8 mM $CaCl_2$, 5.6 mM glucose, 0.1% gelatin, and 20 mM Hepes. Cell suspensions in BSS were supplemented with 10 mM sodium azide and 10 mM 2-deoxy-glucose (Sigma) to inhibit receptor recycling and cellular degranulation during binding experiments. To sensitize cells to DNP, we incubated cells overnight, while cells were still in culture, with excess (10 mg) anti-DNP IgE. Cells, which express FcϵRI (at roughly 300,000 copies per cell) were exposed to IgE for at least 12 hours prior to harvesting.

Flow cytometric binding assays

Binding experiments were performed as described elsewhere (34). Briefly, we incubated a suspension of sensitized cells, with varying concentrations of Alexa-488-labeled ligand at room temperature. After incubating for at least 90 minutes, we used a Becton Dickinson FACScan flow cytometer, which was controlled with Cell Quest software, to collect histograms of fluorescence. Flow cytometric data were recorded as the mean fluorescence (520 nm) of the cell suspension. To correct for nonspecific binding of ligand to cells, we performed a control experiment using cells lacking surface IgE and subtracted the mean fluorescence measured in this experiment from that measured in the corresponding experiment with sensitized cells.

Fitting

To estimate the equilibrium constants K_1 and K_2, we fit the Goldstein-Perelson model to flow cytometric binding data. The experimental data characterize the relative amounts of surface-associated ligand at equilibrium for a series of n total ligand concentrations: $C_1^L, ..., C_n^L$. The quantity that describes the ratio of cell-bound ligands to receptor sites for a total ligand concentration C_i^L is given by

$$\lambda_i = \frac{1}{2C^R}(C_i^L - C_{\infty,i}^L), \tag{13}$$

where i is the index of an experimental data point, and C^R is the total concentration of receptors, which we take to correspond to 300,000 receptors per cell and a cell density of 10^6 cells/ml, and $C_{\infty,i}^L$ is the concentration of free ligand in experiment i. Note that $0 \leq \lambda_i \leq 1$. At a fixed value of C^R (or volume V), the parameters K_1 and K_2 are substituted into Eqs. 1 and 2, which are used to evaluate the partition function q (Eq. 3 combined with Eqs. 4 and 5). The conservation law equations, Eqs. 6

and 7, are then solved for $C^L_{\infty,i}$ and x_i (the fraction of free receptors in the ith experiment) at a fixed value of C^L_i. Then, the value of $C^L_{\infty,i}$ is substituted into Eq. 13 to calculate λ_i.

Before starting the fitting procedure, a series of flow cytometric measurements is scaled such that $0 \leq \mathrm{FL1}(C^L_i) \leq 1$, where $\mathrm{FL1}(C^L_i)$ denotes the mean fluorescence of the cell suspension at ligand concentration C^L_i. Each value of $\mathrm{FL1}(C^L_i)$ is taken to be related to λ_i by a constant factor α, which is an adjustable parameter in fitting. The quality of the fit between the model and experimental data is determined by the root mean square of a vector of deviates,

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} v_i^2}, \tag{14}$$

where $v_i = \lambda_i - \alpha\,\mathrm{FL1}(C^L_i)$.

The fitting procedure is performed as follows. Initial values of K_1 and K_2 are chosen randomly with log-uniform probability in the ranges 10^7 – 10^{10} M^{-1} and 10^9 – 10^{12} M^{-1}, respectively. Then the Levenberg-Marquardt non-linear least squares algorithm (28) is used to find the best-fit values of K_1 and K_2.

To determine confidence intervals for best-fit parameter estimates, we use a bootstrap procedure (28). Fitting as described above is performed for a set of synthetic data sets, which are obtained by sampling the experimental data set with replacement. For each combination of K_1 and K_2 values, the equilibrium model is evaluated for the full set of ligand concentrations $(C^L_1, ..., C^L_n)$, and then λ_i values are calculated (Eq. 13), which are used to obtain RMS values (Eq. 14). We collect 2000 successive sets of K_1, K_2 and α values for which RMS$<$ 0.02. For each parameter, the 2000 values are sorted in rank order. The 320th and 1680th values are used to define a confidence interval at the 68% confidence level.

Data structures of the kinetic models

To simulate the TLBR or extended TLBR model, we use the rule-based kinetic Monte Carlo (KMC) method (13), which requires that we track individual molecules and sites. Here, we describe the data structures that are used to accomplish this task.

Molecules are represented as structured objects with binding sites. The molecules and sites are considered distinguishable for purposes of simulation and the time evolution of each individual reac-

tant is explicitly tracked. Sites are partitioned into sets, each of which contains sites that are reactive in one particular type of reaction (13). Each site can be a member of only one set at a time. Indices of molecules and binding sites are stored in fixed-length and variable-length lists. We define fixed-length lists as indexed data structures; the length is determined by the number of molecules. Unassigned elements in such lists are allowed. Fixed-length lists are used to store the connectivity between particular molecules and their binding sites. Variable-length lists are defined as unsorted data structures, in which every new element is added at the end. If any element within such a list is removed, the last element replaces the one removed. Variable-length lists are used to store reactive sites.

Suppose we simulate a system including multivalent molecules of types P and Q. Each molecule of type P or Q and each reactive site in P or Q are associated with unique indices. The ith molecule of type P, which we take to have n reactive sites, is associated with a set of addresses $P_i = \{p^Q_{i1}, p^Q_{i2}, ..., p^Q_{in}\}$, each of which consists of a pair of pointers $p^Q_{ij} = \{m^Q_{ij}, s^Q_{ij}\}$, where m^Q_{ij} is the index of the molecule of type Q bound to the jth binding site on P_i through the site in this molecule of type Q with index s^Q_{ij}. If site j in the ith molecule of type P is not bound, p^Q_{ij} is set to $\{-1, -1\}$, which by convention is understood to mean that m^Q_{ij} and s^Q_{ij} are unassigned.

Thus, taking into account the numbers and types of sites on molecules, we define $L_i = \{p^R_{i1}, p^R_{i2}, p^R_{i3}\}$ and $R_j = \{p^L_{j1}, p^L_{j2}\}$ for every ith ligand and jth receptor. L_i and R_j are stored in fixed-length lists, $\boldsymbol{L}$ and $\boldsymbol{R}$, respectively, where $i = 1, ..., N_L$ and $j = 1, ..., N_R$. To model cyclic aggregate formation (i.e., ring closure), we assume that each molecule can have six different orientations, as shown in Fig. 3(a). Sites of all ligands and receptors are enumerated clockwise, and each site is characterized by a rotation angle: 0, $\pi/3$, $2\pi/3$, π, $4\pi/3$, $5\pi/3$ or 2π relative to an arbitrarily chosen reference axis. We define fixed-length lists of rotation angles for sites of ligands and receptors, $\boldsymbol{A^L}$ and $\boldsymbol{A^R}$. Elements of these lists are $A^L_i = \{\alpha^L_{i1}, \alpha^L_{i2}, \alpha^L_{i3}\}$ and $A^R_j = \{\alpha^R_{j1}, \alpha^R_{j2}\}$, where α^L_{i1} and α^R_{j1} are rotation angles of sites with index 1 in the ith ligand and jth receptor, respectively, and $i = 1, ..., N_L$ and $j = 1, ..., N_R$. Note that $\alpha^L_{i2} = \alpha^L_{i1} + 2\pi/3$, $\alpha^L_{i3} = \alpha^L_{i1} + 4\pi/3$ and $\alpha^R_{j2} = \alpha^R_{j1} + \pi$.

The reactants are also classified into variable-length lists of free reactive sites and bonds (i.e., pairs of sites that are connected to each other). In the TLBR model, as can be seen in Fig. 2(a,b), we define the following lists of free reactive sites available for association reactions: sites on ligand molecules in solution, F^{3D}_L; sites on surface-tethered ligands, F^{2D}_L; and sites on receptors, F_R. Each list of free reactive sites is defined as $\{p_1, p_2, ..., p_N\}$, where $p_i = \{m_i, s_i\}$ is an address of a ligand or receptor

with index m_i and a site with index s_i. The value of i denotes a position in a list, and N is the number of sites in the list. In the extended TLBR model, we also introduce a list of sites available for participation in ring-closure reactions: $H = \{..., \{p_i^L, p_i^R\}, ...\}$, where p_i^L and p_i^R are addresses of ligand and receptor sites that are positioned next to each other but do not have a bond between them. In both models, a list of bonds, B, comprises bound sites (pairs of addresses) capable of participating in dissociation reactions, $\{..., \{p_i^L, p_i^R\}, ...\}$, where p_i^L and p_i^R are site addresses of two bound molecules, ligand and receptor, respectively. The value of the index i denotes a position in a list, and p_i^L and p_i^R correspond to the non-empty addresses in the fixed-length lists of molecules.

Finally, the full set of sites that can participate in the four reaction classes of the extended TLBR model, X, is defined as follows:

$$X = \left\{ \begin{array}{c} X_{11} \cup X_{12} \\ X_{21} \cup X_{22} \\ X_{31} \\ X_{41} \end{array} \right\} \tag{15}$$

where X_{in} denotes the nth set of sites that are potentially reactive in reaction class i. Note that $X_{11} = F_L^{3D}$, $X_{12} = F_R$, $X_{21} = F_L^{2D}$, $X_{22} = F_R$, $X_{31} = B$ and $X_{41} = H$.

In the equivalent-site TLBR model, sets X_1, X_2 and X_3 (Eq. 15) are constructed based on the local context of interacting molecules in accordance with Yang et al. (13). In the extended TLBR model, we add set X_4, which is constructed by identifying free sites in position to form cycles on a hexagonal lattice. Reaction rate constants associated with reaction classes 1-4 are $k_1 = k_{+1}$, $k_2 = k_{+2}$, $k_3 = k_{\text{off}}$ and $k_4 = j_{+6}$, respectively.

Updates of the system state when reactions occur

The rules are defined in such a way that any reaction causes exchange of site addresses between reactant lists. To illustrate how the reactant lists are updated to account for a reaction, we consider binding of a free ligand l to a free receptor r. Initially, three site addresses of ligand l are stored in list F_L^{3D}, and two site addresses of receptor r are stored in list F_R. Suppose site 3 of ligand l is chosen to react with site 2 of receptor r. Then, firing the reaction is associated with a shift of ligand site address $p^L = \{l, 3\}$ from list F_L^{3D} and receptor site address $p^R = \{r, 2\}$ from list F_R to the list of bonds, B. This reaction modifies the reactivity of the other two sites of ligand l, because this molecule is now

tethered to the surface. Thus, an additional update for sites of ligand l is required: the addresses of the free sites, $\{l,1\}$ and $\{l,2\}$, are shifted from F_L^{3D} to F_L^{2D}. For a receptor crosslinking reaction, an update involves a shift of randomly selected p^L and p^R from lists F_L^{2D} and F_R, respectively, to B; no other changes are required. The ring-closure rule requires a pair of selected sites to be moved from H to B.

In the extended TLBR model, each aggregate containing a molecule affected by a reaction is examined to identify pairs of ligand and receptor sites that are adjacent but not bound to each other. If a new pair is found, H is updated. Every binding reaction in the extended model requires alignment of two interacting binding sites, i.e., rotation of one of the interacting molecules or aggregates relative to the other, as illustrated in Fig. 3(b). Rotation of multimolecular aggregates by angle ϕ is achieved by assigning $A_l^L = \{\alpha_{l1}^L + \phi, \alpha_{l2}^L + \phi, \alpha_{l3}^L + \phi\}$ and $A_r^R = \{\alpha_{r1}^R + \phi, \alpha_{r2}^R + \phi\}$ for each ligand l and receptor r, respectively. After the interacting sites are aligned, one of the two interacting molecules (e.g., receptor r) is positioned at spatial coordinate $(0,0)$ on a lattice with spacing determined by the relative radii of ligands and receptors (or arbitrarily), as shown in Fig. 3(c). Another interacting molecule (e.g., ligand l) is positioned next to the binding site of r, preserving the correct orientation. Spatial positions of the rest of the molecules in these aggregates are retrieved from the site connectivities and rotation angles stored in $\boldsymbol{L}$, $\boldsymbol{R}$, $\boldsymbol{A^L}$ and $\boldsymbol{A^R}$. Finally, the coordinates of all molecules are checked for spatial overlaps; if overlaps are found, then the binding reaction is rejected. If the reaction is accepted, updates for the fixed-length lists of rotation angles are made. Note that the rectangular lattice used in the above procedure may have arbitrary spacing and need not reproduce the exact geometry of receptor aggregates, which form on a hexagonal lattice. Only relative positions, which can be easily transformed into real coordinates for visualization in a post-processing step, are important.

General stochastic algorithm for distinguishable particles

For completeness, below, we present the rule-based KMC method reported by Yang et al. (13), but here we use the data structures introduced above for clarity. We consider four reaction classes with reactants X (Eq. 15). The number of reactant lists for the ith reaction class is given by n_i, where $n_i = 1$ for a unimolecular reaction and $n_i = 2$ for a bimolecular reaction. Thus, X_{in} is the list of reactant sites matching the criteria to be nth-type reactants of reaction class i. Note that the set X_{in} can be of any type: F_L^{3D}, F_L^{2D}, F_R, H or B. Connectivity between two binding sites is recorded in

the fixed-length lists of molecules $\boldsymbol{L}$, $\boldsymbol{R}$, $\boldsymbol{A^L}$ and $\boldsymbol{A^R}$.

The main steps of the algorithm are summarized below.

1. Initialize.

   a. Specify an initial set of molecules in $\boldsymbol{L}$ and $\boldsymbol{R}$. We consider the case where all sites are initially free.

   b. Construct the sets of free reactive sites F_L^{3D}, F_L^{3D} and F_R.

2. Perform the following steps in a loop until $t > t_{\text{end}}$.

   a. Compute the reaction rate for each reaction class i, $r_i = k_i \prod_{n=1}^{n_i} |X_{in}|$, where k_i is the rate constant for class i and $|X_{in}|$ denotes the number of reactive components in class X_{in}. The cumulative reaction rate is given by $r_{\text{tot}} = k_i \sum_{i=1}^{N_X} r_i$, where N_X is the number of reaction rules.

   b. Choose the time step according to $\tau = -\ln(\rho_1)/r_{\text{tot}}$, where ρ_1 is a uniform random number on (0,1).

   c. Choose a reaction class by finding the smallest J such that $\sum_{i=1}^{J} r_i \geq \rho_2 r_{\text{tot}}$.

   d. Choose a reactant site for each reactant type M of the n_J types of sites in reaction class J by picking a site address $p^M \in X_{JM}$ uniformly over the $|X_{JM}|$ possible sites. Note that $p^M = \{m, s\}$, where m is an index of a reactant molecule of type M and s is its reacting site.

   e. Check the final acceptance probability of a reaction (if it is modified by global properties). The rate constant for the modified reaction, k'_J, must obey $k'_J \leq k_J$, in which case a reaction is taken to occur only if $P > \rho_3$, where P is the final acceptance probability defined as $P = k'_J/k_J$, and ρ_3 is a uniform random number on (0,1). If no reaction occurs, proceed to (g).

   f. Update reactant lists and rates. For each reactant site p^M of n_J,

      i. Remove p^M from X_{JM}.

ii. Remove other site addresses of m from other reactant lists in which it appears if matching criteria are no longer satisfied after a reaction fires.

iii. Add p^M and other site addresses of m to other reactant lists if matching criteria are satisfied after a reaction fires.

iv. For each site in molecules connected to m, update all reactant lists with matching criteria that can be affected by transformation of m.

v. Update lists of molecules in which connectivity of sites is changed.

g. Compute simulation results of interest at specified time intervals, e.g., w_n^s, f_g^s and $\langle S \rangle$.

h. Update time by setting $t \leftarrow t + \tau$.

Note that Step 2.e is required if global properties (e.g., connectivities and spatial configurations of molecules in complexes) need to be taken into account. The structure of a complex is determined using the breadth-first traversal method (30). If the acceptance condition for a reaction is not satisfied (i.e., because of steric constraints), then we set $k'_J = 0$, which leads to rejection of the attempted reaction. In the extended TLBR model, formation of cycles is allowed, but only if $j_{+6} > 0$ and a cycle includes exactly six receptors, which follows from restriction of receptor aggregates to a hexagonal lattice.

Results

The Goldstein-Perelson model fits equilibrium binding data

We used flow cytometry to monitor the association of Alexa-488-labeled ligand with cell-surface receptors at equilibrium as a function of ligand dose (Fig. 4). We then fit the Goldstein-Perelson model to the binding data to determine best-fit parameter values. Figure 4 illustrates the agreement between the experimental data and the model fit. Similar agreement can be obtained for any combination of K_1 and K_2 values at points indicated in the parameter space of Fig. 5. The points in this plot correspond to fits for which RMS$<$ 0.02. The best-fit values for K_1 and K_2 (and their 68% confidence intervals) are given in Table 2.

Large aggregates are predicted

For best-fit parameter values, both the TLBR model and the Goldstein-Perelson model predict a sol-gel percolation transition and extensive receptor aggregation, or formation of a gel phase, as ligand concentration increases from 0.2 nM to 30 nM. The simulated fraction of receptors in the gel phase, f_g^s, agrees well with f_g, which is given by Eq. 9. Discrepancy between f_g^s and f_g, when both these quantities approach zero (results not shown), can be attributed to finite-size effects. These effects vanish as $\chi \to \infty$.

Correlation between receptor aggregation and secretory response

Based on fitting of the binding data of Fig. 4, extensive receptor aggregation is predicted by the equivalent-site TLBR model at ligand concentrations that yield a strong secretory response (compare Fig. 6 of this study and Fig. 3 of (18)). This finding contradicts studies indicating that extensive receptor aggregation inhibits secretory responses of RBL cells (19–21). We consider the following possible explanations for this discrepancy: (I) large receptor aggregates are not inhibitory for the ligand used here and in (18); (II) predictions of the equivalent-site TLBR model cannot be trusted because this model is oversimplified (e.g., it does not account for steric constraints on receptor aggregates); (III) large aggregates may be inhibitory, but these aggregates form late in the response to ligand and the early dynamics of receptor aggregation plays a dominant role in triggering secretion.

To investigate hypotheses II and III, we study the equivalent-site TLBR model and the extended TLBR model with and without rings. In particular, we determine the effects of geometrical constraints on the kinetic and equilibrium behavior of ligand-induced receptor aggregation.

The percolation transition

The phase diagram of Fig. 7 characterizes receptor aggregation as predicted by the equivalent-site and extended TLBR models over a broad parametric range. In calculating this phase diagram, for simplicity, we fix the values of N_L and N_R, and then vary k_{+1} and k_{+2}; the volume is also fixed. In Fig. 7(a), we show the Goldstein-Perelson predicted boundary of the PT in the (c, β)-parametric plane (solid line). The parameter β characterizes crosslinking. Higher values of β correspond to stronger binding of surface-tethered ligands to receptors. Comparison of the equivalent-site TLBR

model against the Goldstein-Perelson model (Eqs. 8 and 9) reveals that f_g^s is calculated accurately in both the sol and gel regions (results not shown, see also Yang et al. (13)). The Goldstein-Perelson boundary where $f_g = 0$ corresponds to simulated values of $f_g^s \approx 0.05$. Because of the finite-size effects in the simulation, the transition of f_g^s near the boundary is smeared out, i.e., the PT is not sharp, as shown in Figs. 6 and 8(a). At parameter values corresponding to the region above the PT boundary, f_g^s increases dramatically, and agreement between the continuum Goldstein-Perelson model and finite-size simulation improves. Near the PT boundary, both the Goldstein-Perelson model and simulations predict coexistence of the sol and gel phases. Moving away from the coexistence region to higher values of β, the fraction of receptors in the gel phase increases.

Typical simulated aggregate size distributions are shown in Fig. 7(b-d). Below the PT boundary (point b in Fig. 7(a)), the distribution is such that aggregates containing five or fewer receptors prevail (Fig. 7(b)). In contrast, above the PT boundary (point c in Fig. 7(a)), simulation predicts formation of a superaggregate (Fig. 7(c)). The peak of the distribution of Fig. 7(c) is observed at an aggregate size close to N_R, meaning that the superaggregate contains the majority of receptors of the system. The distribution shown in Fig. 7(d) is characterized by an interesting shape, which is determined by the valence of ligands. Within the sol-gel coexistence region, where the rate of free ligand binding is relatively low compared to the rate of crosslinking (point d of the PT diagram), maximum saturation of surface ligand sites is reached. The simulated distributions below and near the PT boundary agree with the Goldstein-Perelson predicted ones (not shown). Note that the continuum Goldstein-Perelson model cannot predict the size of a superaggregate in the gel phase.

As shown in Fig. 8(a), the equilibrium continuum model predicts a singularity of f_g at the PT boundary, whereas in the finite-sized TLBR model, the transition of f_g^s is not sharp. Recall that f_g^s is defined as the maximum aggregate size divided by N_R, which is finite in the simulations. In the region of parameters below and near the PT boundary, the system contains small aggregates, as shown by the distributions in Fig. 7(b,d). A characteristic of these aggregates is that their sizes are not affected by the system volume (results not shown). Therefore, an increase of the system volume (or the volumetric scaling factor χ, keeping all other parameters the same) improves the agreement between f_g^s and f_g.

Steric constraints on receptor aggregates

In contrast with the equivalent-site TLBR model, the model with steric constraints, but without cyclic aggregates ($j_{+6} = 0$), predicts a gradual change of f_g^s as the crosslinking parameter β is increased, even as the volumetric scaling factor χ goes to infinity, as illustrated in Fig. 8(a). Receptor aggregation is generally suppressed. When steric constraints are taken into account, the fraction of receptors in the gel phase, as well as the mean aggregate size, decreases significantly over the considered range of ligand concentrations (0.01 - 100 nM). However, the maximum of both f_g^s and $\langle S \rangle$ is observed at the same ligand concentration as in the equivalent-site TLBR model (results not shown).

The possibility for ring closure reactions ($j_{+6} > 0$) changes the percolation behavior. By including these reactions, the effects of steric constraints are reversed. As shown in Fig. 8(b), an increase in the value of rate constant j_{+6} makes the PT steeper (i.e., more sensitive to the value of β). As $j_{+6} \to \infty$, the PT occurs at a much lower value of β compared to that in the equivalent-site model (about 10 times lower), and the transition becomes steep. The approximate PT boundary simulated using the extended TLBR model with $j_{+6} = 100$ s^{-1} is indicated in Fig. 7(a) by a dashed line; the boundary corresponds to $f_g^s \approx 0.05$. Formation of stable rings increases the size of the gel region in the phase diagram of Fig. 7(a) (compare the dashed and solid lines) because each hexagonal cycle in a receptor aggregate has up to six ligand sites free for receptor crosslinking.

Influence of the ring-closure rate constant, j_{+6}, on receptor aggregate structure is illustrated in Fig. 9. The fraction of receptors in the gel phase, f_g^s, as a function of j_{+6} at different system sizes (i.e., at different values of χ) is shown in Fig. 9(a). A critical range of j_{+6} values, in which the PT occurs, spans 0.1 to 1 s^{-1}. A decrease or increase of j_{+6} out of this range does not cause significant changes. For $j_{+6} < 0.1$, formation of "open cycles" (Fig. 2(c)) is the rate-limiting step in aggregation kinetics; for $j_{+6} > 1$, ring closure stabilizes open cycles once such aggregates form. Increase of the system size makes the transition of f_g^s more pronounced, but does not influence the range of sensitivity to j_{+6}. The value of f_g^s can also serve as a characteristic of the density of aggregates. Figure 9(b) shows a fragment of a highly-branched aggregate predicted to form in one particular simulation of the extended TLBR model with $j_{+6} = 0$ (no ring closure). At another limit, where $j_{+6} = 10^4$ s^{-1} and ring closure is fast, cycles are stable and simulated aggregates have a very dense structure characterized by $f_g^s \approx 1$, as shown in Fig. 9(c).

Dynamics of aggregation

Here, we illustrate how the dynamics of receptor aggregation are influenced by steric constraints and ring formation. Also, we reiterate our earlier results, which characterize dynamical behavior of receptor aggregation according to the equivalent-site TLBR model (13).

In our earlier work (13), we have shown that in the gel phase, the existence of which is predicted for ligand concentrations corresponding to strong secretory responses (18) (Fig. 6), small receptor aggregates form transiently before the formation of a superaggregate. As receptor aggregates grow in size during the initial time course, small and intermediate-sized aggregates may possibly have a stimulating effect and induce a secretory response. At the same time, the number of aggregates that contain up to 10 receptors increases. As seen from the simulation results shown in Fig. 4 of Ref. (13), the transient predominance of small aggregates can last for a few minutes with reasonable parameter values.

Similar dynamical behavior is obtained using the extended model without the ring-closure reaction ($j_{+6} = 0$), as shown in Fig. 10 (a and b). Ligand binding and receptor-crosslinking rate constants are equal to those used by Yang et al. (13) and are within the confidence intervals of our best-fit parameter values (Table 2). When steric constraints are taken into account, the number of small aggregates that contain 3, 5, 7, and 9 receptors increases transiently at both low and high ligand doses. At high ligand dose, the transient behavior occurs faster. The number of small aggregates does not tend to zero, but remains at the level of $0.1S_{max}$, where S_{max} is the fraction of receptors in the largest aggregate.

Our previous study (13) also indicated that two ligand doses stimulating receptor aggregation to the same extent at equilibrium can generate qualitatively distinct time courses of receptor aggregation. We choose two ligand doses, 0.33 nM and 8.3 nM, that yield about the same fraction of receptors in the gel phase at equilibrium, $f_g \approx 0.5$ (Fig. 6). As shown in Fig. 11, the mean aggregate size, $\langle S \rangle$, as a function of time for the two ligand doses, merges at the same steady-state level of $\langle S \rangle \approx 17$ (curves a and b). The extended model without the ring-closure reaction ($j_{+6} = 0$) demonstrates qualitatively similar behavior (curves c and d in Fig. 11). A feature of both models at high ligand concentration, 8.3 nM, is an overshoot in the fraction of receptors in the gel phase (superaggregate). The reason for this overshoot has been explained by Yang et al. (13).

When the ring-closure reaction is fast ($j_{+6} \to \infty$), the transient behavior changes. The two time

courses of $\langle S \rangle$ for low and high ligand doses (curves e and f in Fig. 11) merge at $\langle S \rangle \approx 250$, and the number of small aggregates tends to zero at both doses (results not shown). Also, as shown in Fig. 11, the overshoot seen in the case of $j_{+6} = 0$ at high ligand dose disappears.

Steric effects can be modeled using effective rate constants

In this section we discuss how the analysis of interactions with steric constraints can be simplified. Explicitly keeping track of the spatial structure of aggregates can be avoided as follows. First, simulation of the extended TLBR model (with $j_{+6} = 0$) is modified so that at every attempt to bind ligand l and receptor r we evaluate the sizes of aggregates in which the interacting ligand and receptor are members, s_l and s_r (both values are the number of receptors in each aggregate). For each combination of s_l and s_r, an empirical binding probability, $P(s_l, s_r)$, is calculated as the number of successful binding events divided by the total number of attempts. Estimated $P(s_l, s_r)$ is plotted in Fig. 12(a). The functional form of $P(s_l, s_r)$ is given approximately as follows:

$$P(s_l, s_r) = \min\left\{1,\ a\left(\frac{1}{b + s_l^d} + \frac{1}{b + s_r^d}\right)\right\}, \tag{16}$$

where the coefficients are $a = 5.88$, $b = 5.75$ and $d = 1.19$ for the ligand and receptor geometries considered in this study. As can be expected, $P(s_l, s_r)$ values are close to 1 if both or at least one of the interacting aggregates contains only a few molecules. For example, $P(s_l, s_r) = 1$ for any combination of $s_l = 0, 1, 2$ and $s_r = 1, 2$, because such small aggregates cannot overlap in space. Note that ligand and receptor molecules that are members of the same aggregate and positioned next to each other, but do not have a bond between them, cannot bind single receptors and ligands from solution because their free sites are not accessible. Therefore, $P(0, s_r)$ and $P(s_l, 1)$ can be less than 1.

In simulation of the equivalent-site TLBR model, we use tabulated values of $P(s_l, s_r)$ in step 2.e of the simulation algorithm to reject on average the same fraction of binding attempts as when steric clashes are found explicitly in simulation of the extended TLBR model without rings. For every association reaction, the sizes of interacting aggregates, s_l and s_r, are determined and then used to evaluate $P(s_l, s_r)$. Figures 12(b,c) show a comparison of the simulation results obtained with explicit and implicit enforcement of steric constraints. The aggregate size distributions reveal good agreement between two the simulation approaches. Thus, we conclude that the influence of steric effects on the

overall kinetics of binding and aggregate formation can be effectively lumped into a modified binding rate constant, k'_J, as follows:

$$k'_J = P(s_l, s_r)k_J, \tag{17}$$

where the empirical binding probability $P(s_l, s_r)$ depends only on the sizes of interacting aggregates (Eqs. 16). The modified simualtion method is more efficient than the original with explicit tracking of aggregate configurations because it does not require a determination of the geometries of interacting aggregates for every binding reaction once the empirical binding probability function has been determined.

Discussion

In this work, we present equilibrium binding data characterizing the interaction of a trivalent ligand (18) with a bivalent receptor. To fit this data and estimate the model parameters describing ligand capture from solution and receptor crosslinking, we use the Goldstein-Perelson model (17) (Fig. 4). We found that at the best-fit parameter values, the model predicts the formation of a gel phase for an interval of ligand doses that corresponds to strong secretory responses (18). This model prediction contradicts the results of previous studies that indicate that large aggregates are inhibitory (19–21). Here, we consider the following explanations of this contradiction. First, the Goldstein-Perelson model, although it fits the binding data well, is oversimplified, i.e., it treats sites as equivalent and does not account for steric effects, which can be expected to limit the formation of large aggregates. Second, dynamics of receptor aggregation may play a dominant role in initiating cellular responses. Small receptor aggregates that are forming during the initial transient response to ligand may be sufficient to stimulate a secretory response.

Recently, we have developed a general-purpose KMC simulation method that can be used to study multivalent interactions (13). Based on this method, we have built a kinetic model of interactions between trivalent ligands and bivalent receptors, the equivalent-site TLBR model. We have now extended number to account for structural properties of interacting molecules. Because no information is available in the literature about how these steric effects are expected to affect ligand-receptor binding to the best of our knowledge, we carried out an extensive analysis of steric effects in this system, considering both equilibrium and kinetic effects.

As expected, we found that steric constraints tend to suppress receptor aggregation. To demonstrate this behavior, we simulated the extended TLBR model, which takes into account steric constraints on receptor aggregate configurations that arise from assumed ligand and receptor geometries. The extended TLBR model is built on the assumption that ligands and receptors are rigid and that receptor aggregates form on a hexagonal lattice. A sharp PT observed in the equivalent-site TLBR model as a function of the receptor crosslinking parameter β disappears in the extended TLBR model without ring formation. Investigation of finite-size effects in these simulations revealed that an increase of system size does not sharpen the transition.

Behavior of the extended TLBR model changes dramatically if cyclic aggregates are included. In this model extension, we assume that steric effects constrain cyclic aggregates on hexagonal cells. If the rings considered in our model are stable, they essentially turn individual bivalent receptors into 6-valent "receptors" for ligand. At high ring-closure rates, the region of the PT is extended (compared to the equivalent-site TLBR model predictions) and receptor density of aggregates increases.

Dynamics in the extended TLBR model without rings are qualitatively similar to the kinetics of the original equivalent-site TLBR model (Figs. 10 and 11 of this work and Fig. 4 of Yang et al. (13)). In both these models, at high rates of ligand capture from solution, the extent of aggregation passes through a maximum during the initial time course and then decreases to the equilibrium level (Fig. 11). Even at conditions that yield extensive aggregation, the fraction of small aggregates does not tend to zero. However, when rings are added and the ring-closure reaction constant is high, the dynamical behavior of aggregation changes. The ring-closure reaction enhances receptor crosslinking and eliminates overshoot in receptor aggregation even at high ligand doses (Fig. 11). Small aggregates disappear due to formation of stable hexagonal cells.

Another important finding in our analysis of steric effects is that these effects can be captured using an empirical function that characterizes the dependence of binding probability on the sizes of associating aggregates. This function is estimated from our fully-detailed simulations using the extended TLBR model without rings (Fig. 12). Thus, steric constraints can be effectively lumped into factors that multiply binding rate constants (Eq. 17), and complicated data structures and graph traversal methods otherwise used to track molecular geometry can be avoided. Using the implicit description of steric constraints may be advantageous for fitting procedures and other computational protocols.

The methodology developed here for studying steric constraints on receptor aggregates can be applied to an array of other problems where the steric properties of interacting molecules constrain or facilitate interactions. The modeling approach may even be extended to account for the structures of interacting molecules in three-dimensional space. One example of a potential application area is modeling of viral capsid assembly. This process has recently been modeled using mass-action kinetic models (22) and molecular dynamics simulations (23). The method presented here could potentially be applied to study the kinetics of capsid assembly more efficiently while maintaining a realistic representation of geometrical factors. Another potential area of application is simulation of actin filament dynamics (24–26).

Acknowledgments

This work was supported by NIH grants RR18754, GM076570, AI35997, and CA109552, DOE contract DE-AC52-06NA25396, and the Arizona Biomedical Research Comission. M.I.M. acknowledges support from the Center for Nonlinear Studies. J.R.F. acknowledges institutional support. We thank Byron Goldstein for helpful discussions.

References

1. Hlavacek, W. S., J. R. Faeder, M. L. Blinov, A. S. Perelson, and B. Goldstein, 2003. The complexity of complexes in signal transduction. *Biotechnol. Bioeng.* 84:783–794.

2. Hlavacek, W. S., J. R. Faeder, M. L. Blinov, R. G. Posner, M. Hucka, and W. Fontana, 2006. Rules for modeling signal-transduction systems. *Sci. STKE* 2006:re6.

3. Pawson, T., and P. Nash, 2003. Assembly of cell regulatory systems through protein interaction domains. *Science* 300:445–452.

4. Le Novére, N., and T. S. Shimizu, 2001. StochSim: modelling of stochastic biomolecular processes. *Bioinformatics* 17:575–576. NO! Cite Morton-Firth and Bray (1998) J. Theor. Biol. first

5. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek, 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* 20:3289–3291.

6. Faeder, J. R., M. L. Blinov, B. Goldstein, and W. S. Hlavacek, 2005. Rule-based modeling of biochemical networks. *Complexity* 10:22–41.

7. Lok, L., and R. Brent, 2005. Automatic generation of cellular reaction networks with Moleculizer 1.0. *Nat. Biotechnol.* 23:131–136.

8. Meier-Schellersheim, M., X. Xu, B. Angermann, E. J. Kunkel, T. Jin, and R. N. Germain, 2006. Key role of local regulation in chemosensing revealed by a new molecular interaction-based modeling method. *PLoS Comput Biol.* 2:e82.

9. Colvin, J., M. I. Monine, J. R. Faeder, W. S. Hlavacek, and R. G. Posner. Submitted. DYNSTOC: a tool for simulating large-scale rule-based models .

10. Borisov, N. M., A. S. Chistopolsky, J. R. Faeder, and B. N. Kholodenko, 2008. Domain-oriented reduction of rule-based network models. *IET Syst. Biol.* 2:342–351.

11. Gillespie, D. T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.* 22:403–434.

12. Cao, Y., H. Li, and L. Petzold, 2004. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.* 121:4059–4067.

13. Yang, J., M. I. Monine, J. R. Faeder, and W. S. Hlavacek, 2008. Kinetic Monte Carlo method for rule-based modeling of biochemical networks. *Phys. Rev. E.* 78:31910–31917.

14. Morton-Firth, C. J., and D. Bray, 1998. Predicting temporal fluctuations in an intracellular signalling pathway. *J. Theor. Biol.* 192:117–128.

15. Danos, V., J. Feret, W. Fontana, and J. Krivine, 2007. Scalable simulation of cellular signaling networks. *Lect. Notes Comp. Sci.* 4807:139–157.

16. Perelson, A. S., and C. DeLisi, 1980. Receptor clustering on a cell surface. I. Theory of receptor cross-linking by ligands bearing two chemically identical functional groups. *Math. Biosci.* 48:71–110.

17. Goldstein, B., and A. S. Perelson, 1984. Equilibrium theory for the clustering of bivalent cell surface receptors by trivalent ligands. Application to histamine release from basophils. *Biophys. J.* 45:1109–1123.

18. Posner, R. G., D. Geng, S. Haymore, J. Bogert, I. Pecht, A. Licht, and P. B. Savage, 2007. Trivalent antigens for degranulation of mast cells. *Org. Lett.* 9:3551–3554.

19. Becker, K. E., T. Ishizaka, H. Metzger, K. Ishizaka, and P. M. Grimley, 1973. Surface IgE on human basophils during histamine release. *J. Exp. Med.* 138:394–409.

20. Metzger, H., 1992. Transmembrane signaling: the joy of aggregation. *J. Immunol.* 149:1477–1487.

21. Metzger, H., 2002. Molecular versatility of antibodies. *Immunol. Rev.* 185:186–205.

22. Endres, D., and A. Zlotnick, 2002. Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys. J.* 83:1217–1230.

23. Olson, A. J., Y. H. E. Hu, and E. Keinan, 2007. Chemical mimicry of viral capsid self-assembly. *Proc. Natl. Acad. Sci. USA* 104:20731–20736.

24. Pollard, T. D., L. Blanchoin, and R. D. Mullins, 2000. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Bimol. Struct.* 29:545–576.

25. Ponti, A., M. Machacek, S. L. Gupton, C. M. Waterman-Storer, and G. Danuser, 2004. Two distinct actin networks drive the protrusion of migrating cells. *Science* 305:1782–1786.

26. Matzavinos, A., and H. G. Othmer, 2007. A stochastic analysis of actin polymerization in the presence of twinfilin and gelsolin. *J. Theor. Biol.* 249:723–736.

27. Perelson, A. S., 1980. Receptor clustering on a cell surface. II. Theory of receptor cross-linking by ligands bearing two chemically distinct functional groups. *Math. Biosci.* 49:87–110.

28. Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992. *Numerical Recipes in C,* 2nd ed. Cambridge Univ. Press.

29. Holowka, D., D. Sil, C. Torigoe, and B. Baird, 2007. Insights into immunoglobulin E receptor signaling from structurally defined ligands. *Immun. Rev.* 217:269–279.

30. Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein, 2001. *Introduction to Algorithms,* 2nd ed. MIT Press, Cambridge.

31. Faeder, J. R., W. S. Hlavacek, I. Reischl, M. L. Blinov, H. Metzger, A. Redondo, C. Wofsy, and B. Goldstein, 2003. Investigation of early events in FcεRI-mediated signaling using a detailed mathematical model. *J. Immunol.* 170:3769–3781.

32. Holowka, D., and H. Metzger, 1982. Further characterization of the beta-component of the receptor for immunoglobulin E. *Mol. Immunol.* 19:219–227.

33. Barsumian, E. L., C. Isersky, M. G. Petrino, and R. P. Siraganian, 1981. IgE-induced histamine release from rat basophilic leukemia cell lines: isolation of releasing and nonreleasing clones. *Eur. J. Immunol.* 11:317–323.

34. Posner, R. G., J. M. Paar, A. Licht, I. P. D. H. Conrad, and W. S. Hlavacek, 2004. Interaction of a monoclonal IgE-specific antibody with cell-surface IgE-FcεRI: characterization of equilibrium binding and secretory response. *Biochemistry* 43:11352–11360.

references are out of order

Figure Legends

Figure 1.

Schematic representations of (a) a ligand and a receptor, (b) a chain-like receptor aggregate, and (c) a tree-like branched aggregate. (d) Examples of cyclic aggregates that are not considered. In all panels, circles represent binding sites.

Figure 2.

Reaction scheme of the trivalent ligand – bivalent receptor system. Missing or variable parts of complexes are indicated by dotted lines. (a) A ligand from solution is captured by a cell-surface receptor with single-sie rate constant k_{+1}. (b) A tethered ligand crosslinks two receptors with single-site rate constant k_{+2}. (c) In the extended TLBR model, a hexagonal receptor aggregate can form with rate j_{+6} when free ligand and receptor sites belonging to the same aggregate are positioned next to each other. In all cases, dissociation is a context-independent reaction and occurs with single-site rate constant k_{off}.

Figure 3.

Configurations of molecules in the extended TLBR model. (a) Ligands and receptors can have six orientations. Enumeration of sites is clockwise. (b) A binding reaction requires adjustment of angles and positions of the two interacting molecules, l and r. (c) Representation of an association reaction on a lattice. Coordinate (0,0) is assigned to receptor r with a free binding site rotated by angle $4\pi/3$. Another complex, which can bind through ligand l, is positioned so that its free binding site is rotated by $\pi/3$. Coordinate (-1,-1) is assigned to ligand l. Spatial positions of the rest of the molecules of both complexes are retrieved from stored information about site connectivities and rotation angles.

Figure 4.

Fit of the Goldstein-Perelson or the equivalent-site TLBR model to flow cytometric binding data. The y-axis indicates the normalized amount of ligand bound to cell-surface receptors, λ, at different ligand doses. Dots represent scaled measurements of average cell-associated fluorescence from the Alexa-488-labeled ligand. Parameter values for the best-fit binding curve (solid line) are $K_1 = 4.7 \times 10^8$ M^{-1},

$K_2 = 8.7 \times 10^{10}$ M^{-1}, and $\alpha = 0.82$. Other parameter values used in simulations of the TLBR model are $k_{\text{off}} = 0.01$ s^{-1}, $N_R^* = 300$ and volume $V^* = 10^{-12}$ L.

Figure 5.

Parametric uncertainty plot showing the region of acceptable agreement between the model and experimental data. The fitting was done by using the Goldstein-Perelson model combined with the Levenberg-Marquardt non-linear least squares algorithm. For every combination of K_1 and K_2 values, an optimization routine was run to find a minimum of the root mean square (RMS) deviation between the experimental data and the model. Dots denote local minima at which RMS< 0.02. Values of k_{+1} and k_{+2} indicated in this plot are based on $k_{\text{off}} = 0.01$ s^{-1}, $N_R^* = 300$, $V^* = 10^{-12}$ L.

Figure 6.

Simulation of Alexa-488-labeled trivalent DNP ligand binding to bivalent anti-DNP IgE-FcϵRI complexes at equilibrium using the equivalent-site TLBR model. Extensive aggregation is observed for ligand concentrations from 0.2 nM to 30 nM. The simulation results showing the fraction of receptors in the gel phase f_g^s (solid line) agree with f_g given by the Goldstein-Perelson model (dashed line). Parameter values used to calculate the equilibrium crosslinking curve (solid line) are $K_1 = 4.7 \times 10^8$ M^{-1}, $K_2 = 8.7 \times 10^{10}$ M^{-1}, and $\alpha = 0.82$. Additional parameter values used in simulations of the TLBR model are $k_{\text{off}} = 0.01$ s^{-1}, $N_R^* = 300$ and volume $V^* = 10^{-12}$ L.

Figure 7.

Percolation transition (PT) in the TLBR and extended TLBR models. (a) PT diagram in the parameter space of normalized forward rate constants $c = 3k_{+1}N_{L,\infty}/k_{\text{off}}$ and $\beta = k_{+2}N_R/k_{\text{off}}$. The PT boundary obtained from the Goldstein-Perelson model (solid line) agrees well with TLBR model simulation results (not shown). An approximate PT boundary obtained using the extended TLBR model (dashed line) with $j_{+6} = 100$ s^{-1}, along which $f_g^s = 0.05$, shows that the region of the gel phase extends significantly, which is due to formation of stable cyclic (hexagonal) structures. (b-d) Typical aggregate size distributions obtained with the TLBR and extended TLBR models at combinations of parameter values indicated by labeled dots in (a). Fixed parameters: $k_{\text{off}} = 0.01$ s^{-1}, $N_R^* = 300$, $N_L^* = 4200$, $\chi = 10$.

Figure 8.

PT modified by finite-size effects and steric constraints. The y-axis in each panel indicates the fraction of receptors in the gel phase. The x-axis in each panel indicates the value of the dimensionless crosslinking rate constant β. (a) Finite-size effects in the equivalent-site and extended TLBR models without rings ($j_{+6} = 0$) depend on the volumetric scaling factor χ. Parameter values are scaled by χ as follows: $N_R = \chi N_R^*$, $N_L = \chi N_L^*$, $k_{+1} = \chi^{-1} k_{+1}^*$, and $k_{+2} = \chi^{-1} k_{+2}^*$. The value of f_g given by the Goldstein-Perelson model is indicated by the solid line. For $\chi = 0.1$, 1, 10 and 100, calculations based on the equivalent-site TLBR model (dotted lines) and the extended TLBR model (dashed lines) are shown. The equivalent-site TLBR model approaches the continuum model as $\chi \to \infty$. (b) Effect of j_{+6} on the PT according to the extended TLBR model with rings. Increase of j_{+6} above 100 does not have a significant effect on the PT. In all calculations, we used the following parameter values: $N_R^* = 300$, $N_L^* = 4200$, $c = 0.36$ ($k_{+1}^* = 3 \times 10^{-7}$ molecules^{-1}s^{-1} and $N_{L,\infty} \approx N_L$), $k_{+2}^* = \beta k_{\text{off}} / N_R^*$ s^{-1} and $k_{\text{off}} = 0.01$ s^{-1}.

Figure 9.

Effect of j_{+6} on cyclic receptor aggregates. (a) f_g^s is used as a measure of the aggregate density. At larger system volume (e.g., at $\chi = 10$ vs. $\chi = 0.1$), the difference in values of f_g^s while j_{+6} is varied is steeper because of finite-size effects. (b,c) Snap-shots of simulations showing two fragments of aggregates: (b) a low-density structure (obtained with $j_{+6} = 0$), and (c) a high-density structure (obtained with $j_{+6} = 100$ s^{-1}). In all calculations, we used the following parameter values: $N_R^* = 300$, $N_L^* = 4200$, $c = 0.36$ ($k_{+1}^* = 3 \times 10^{-7}$ molecules^{-1}s^{-1}, $N_{L,\infty} \approx N_L$), $\beta = 90$ ($k_{+2}^* = 3 \times 10^{-3}$ s^{-1}) and $k_{\text{off}} = 0.01$ s^{-1}.

Figure 10.

Kinetics of receptor aggregation according to the extended TLBR model without ring formation ($j_{+6} = 0$). (a) Fraction of receptors in aggregates containing 1, 3, 5, 7 or 9 receptors or the largest aggregate as a function of time at a low ligand concentration of 0.33 nM ($c = 0.11$). (b) The same curves as in panel (a) but for a high ligand concentration of 8.3 nM ($c = 2.7$). In all calculations, $\beta = 16.8$, $k_{\text{off}} = 0.01$ s^{-1}, $N_R^* = 300$, and $\chi = 1$.

Figure 11.

Transient behavior of the mean aggregate size at the same conditions as shown in Fig. 10. The time courses were obtained using the following models: the equivalent-site TLBR model (for curves a and b), the extended TLBR model without rings ($j_{+6} = 0$) (for curves c and d), and the extended TLBR model with rings ($j_{+6} = 100\ \mathrm{s}^{-1}$) (for curves e and f). Solid lines correspond to $c = 2.7$ (8.3 nM), and dashed lines correspond to $c = 0.11$ (0.33 nM). In all calculations, $\beta = 16.8$, $k_{\mathrm{off}} = 0.01\ \mathrm{s}^{-1}$, $N_R^* = 300$, and $\chi = 1$.

Figure 12.

(a) Empirical probability of successful binding events, $P(s_l, s_r)$, as a function of aggregate size. (b,c) Comparison of the simulation results obtained with the extended TLBR model without rings and the equivalent-site TLBR model in which binding rates are modified by the empirical function $P(s_l, s_r)$. Different threshold values of maximum aggregate sizes were used: $s_{max} = 100$ (b) and 200 (c). In all calculations, $N_R^* = 300$, $N_L^* = 4200$, $c = 0.36$ ($k_{+1}^* = 3 \times 10^{-7}$ molecules^{-1}s^{-1} and $N_{L,\infty} \approx N_L$), $\beta = 90$ ($k_{+2}^* = 3 \times 10^{-3}\ \mathrm{s}^{-1}$) and $k_{\mathrm{off}} = 0.01\ \mathrm{s}^{-1}$.

Table 1: Reference values of parameters used in simulations.

Parameter	Value	Units
V^*	10^{-12}	L
N_R^*	300	molecules
N_L^*	4200	molecules
k_{off}	0.01	s^{-1}

Table 2: Best-fit values for the equilibrium binding constants and scaling factor α.

Parameters	Mean	Lower limit	Upper limit
K_1 (nM^{-1})	0.467	0.111	0.767
K_2 (nM^{-1})	87.03	31.6	128.1
α	0.816	0.758	0.881

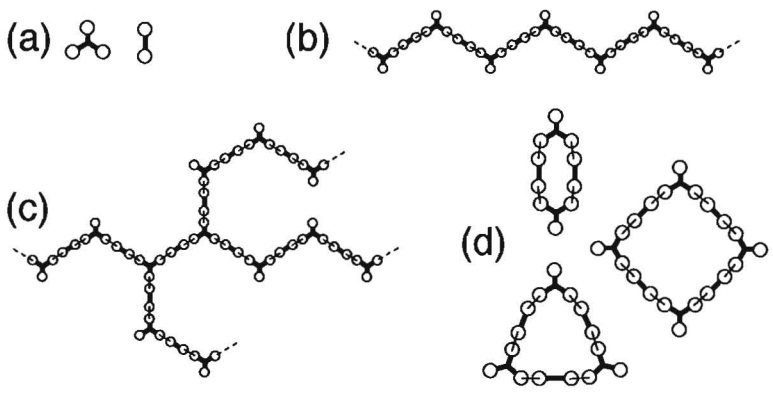

Figure 1:

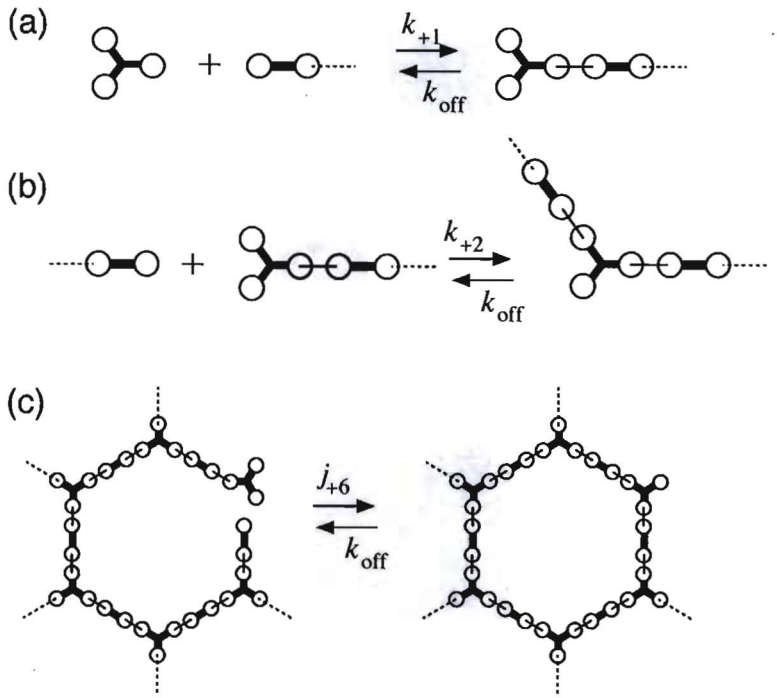

Figure 2:

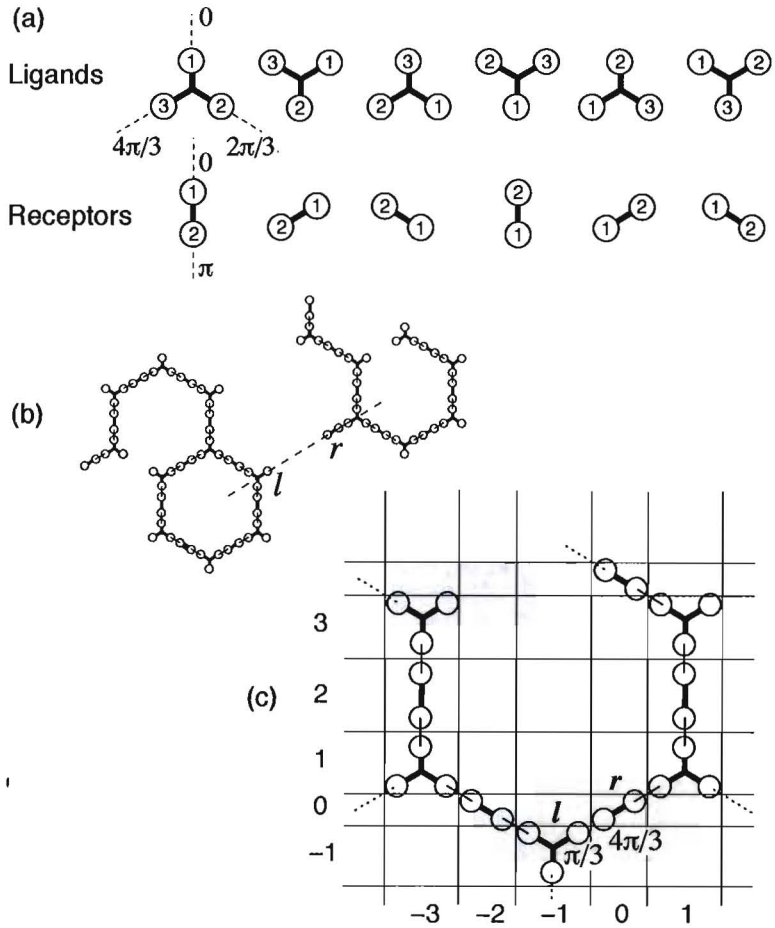

Figure 3:

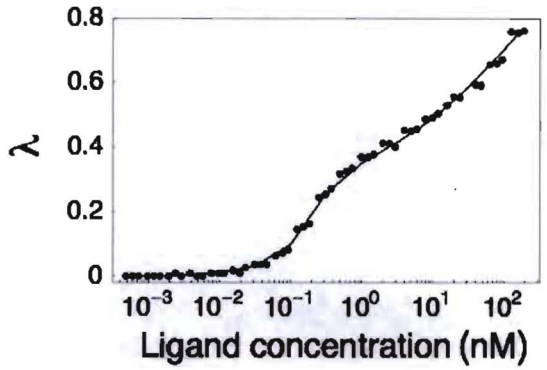

Figure 4:

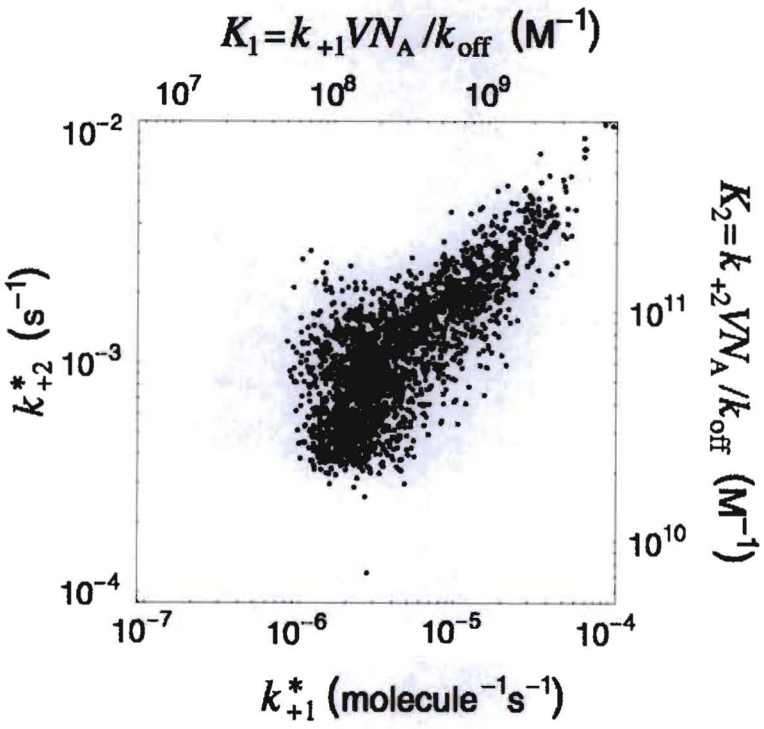
$K_1 = k_{+1} V N_A / k_{off}$ (M^{-1})
10^7
10^8
10^9
10^{-2}
10^{-3}
10^{-4}
k^*_{+2} (s^{-1})
10^{11}
10^{10}
$K_2 = k_{+2} V N_A / k_{off}$ (M^{-1})
10^{-7}
10^{-6}
10^{-5}
10^{-4}
k^*_{+1} (molecule^{-1}s^{-1})

Figure 5:

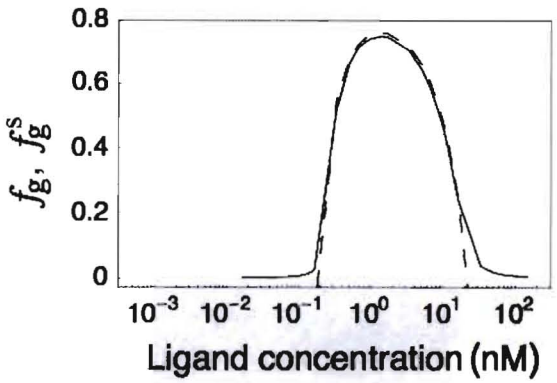
0.8
0.6
0.4
0.2
0
f_g, f_g^S
10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1} 10^{2}
Ligand concentration (nM)

Figure 6:

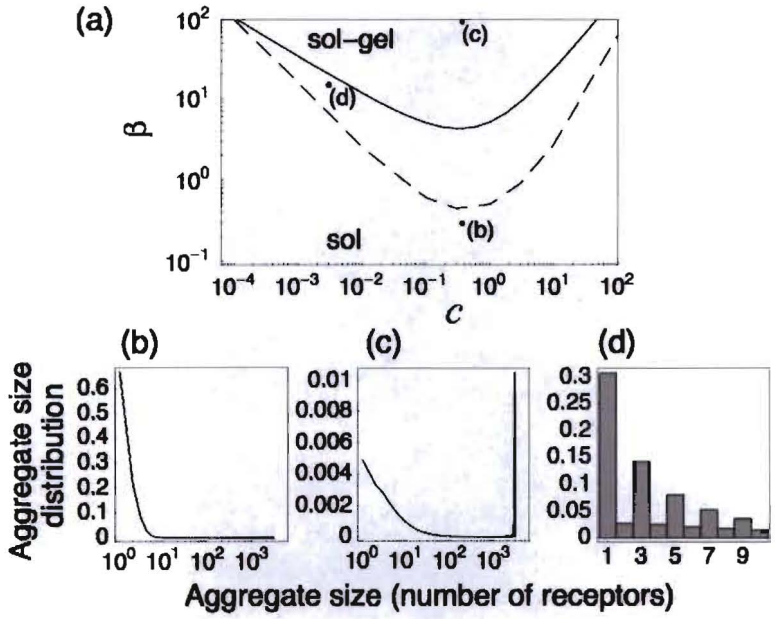
(a)
sol–gel
(c)
(d)
(b)
sol
β
c
(b)
(c)
(d)
Aggregate size distribution
Aggregate size (number of receptors)

Figure 7:

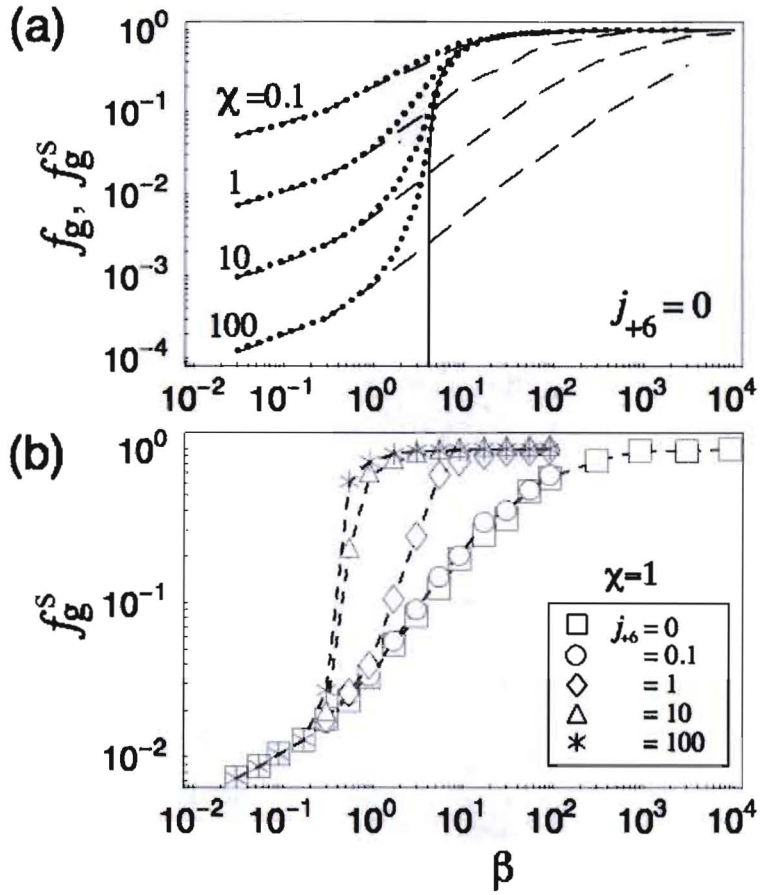
(a)
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
f_g, f_g^S
χ=0.1
1
10
100
$j_{+6}=0$
10^{-2} 10^{-1} 10^0 10^1 10^2 10^3 10^4
(b)
10^0
10^{-1}
10^{-2}
f_g^S
χ=1
$j_{+6}=0$
=0.1
=1
=10
=100
10^{-2} 10^{-1} 10^0 10^1 10^2 10^3 10^4
β

Figure 8:

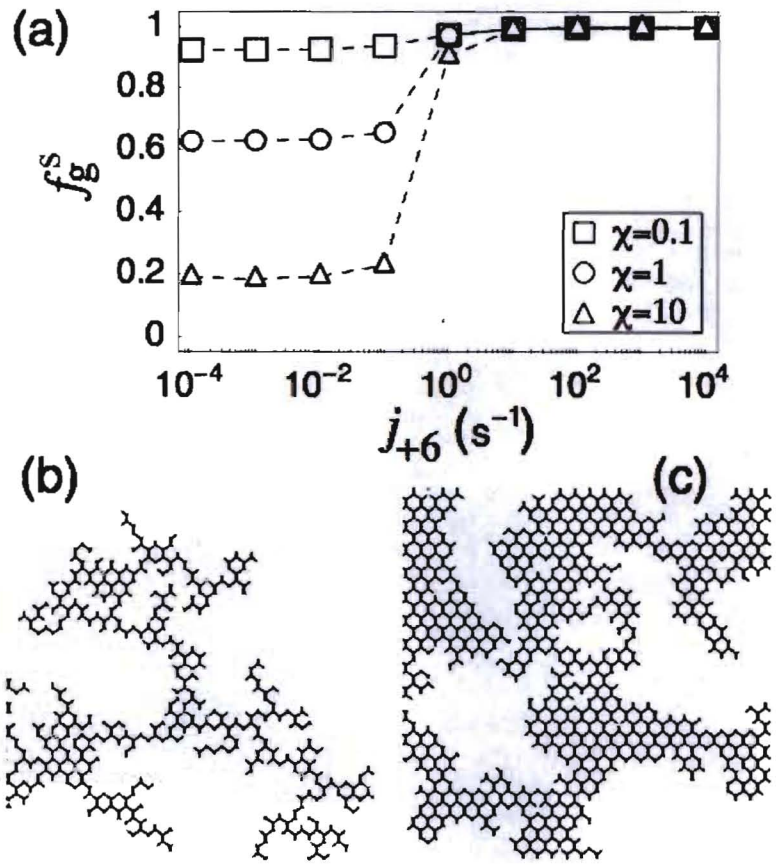

Figure 9:

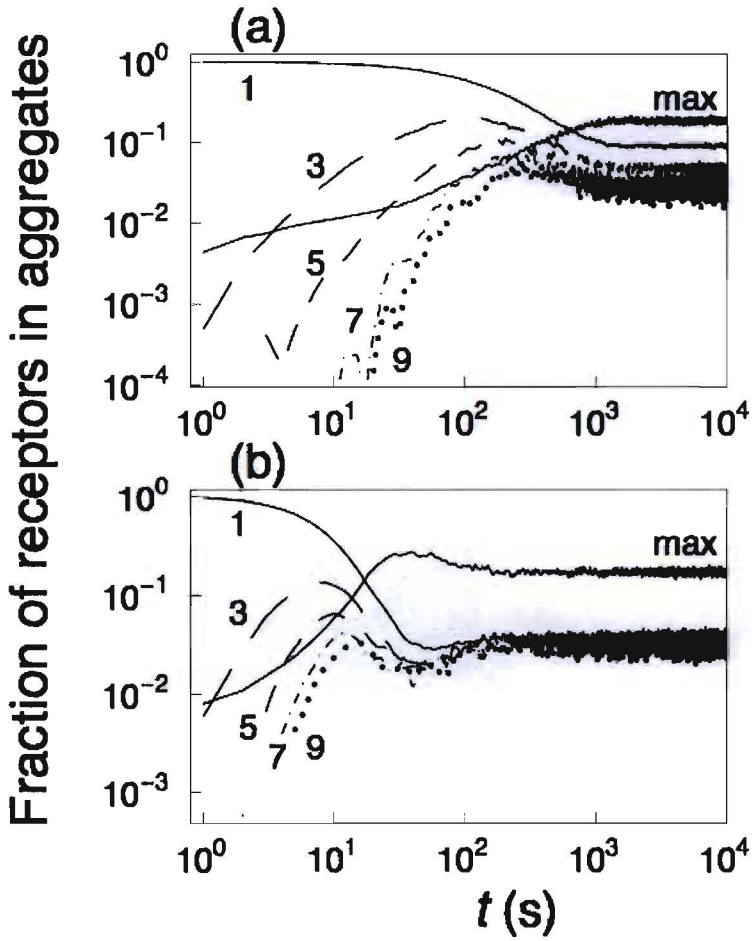
(a)
(b)
Fraction of receptors in aggregates
t (s)
max
1
3
5
7
9
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^1
10^2
10^3
10^4

Figure 10:

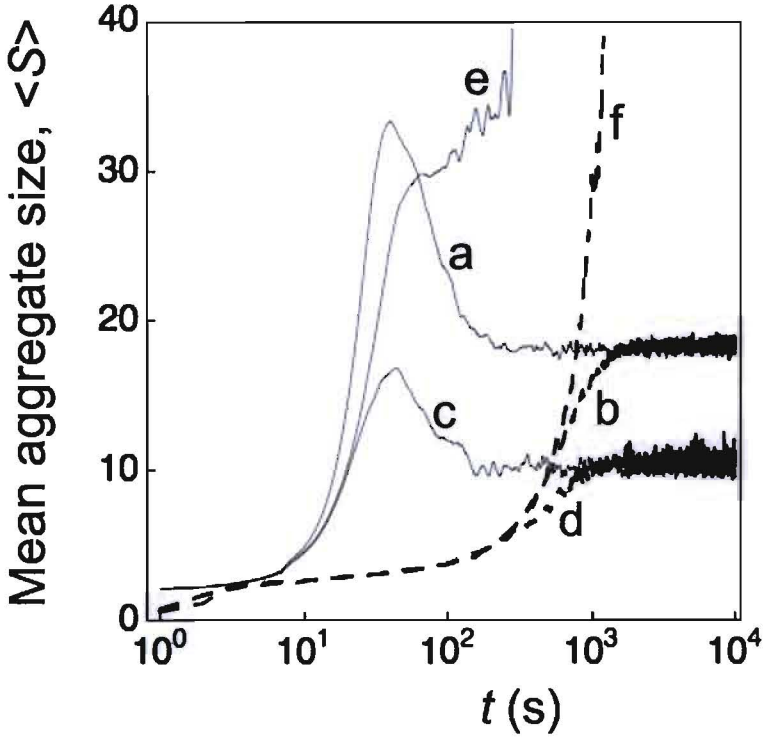
Mean aggregate size, <S>
40
30
20
10
0
10^0
10^1
10^2
10^3
10^4
t (s)
a
b
c
d
e
f

Figure 11:

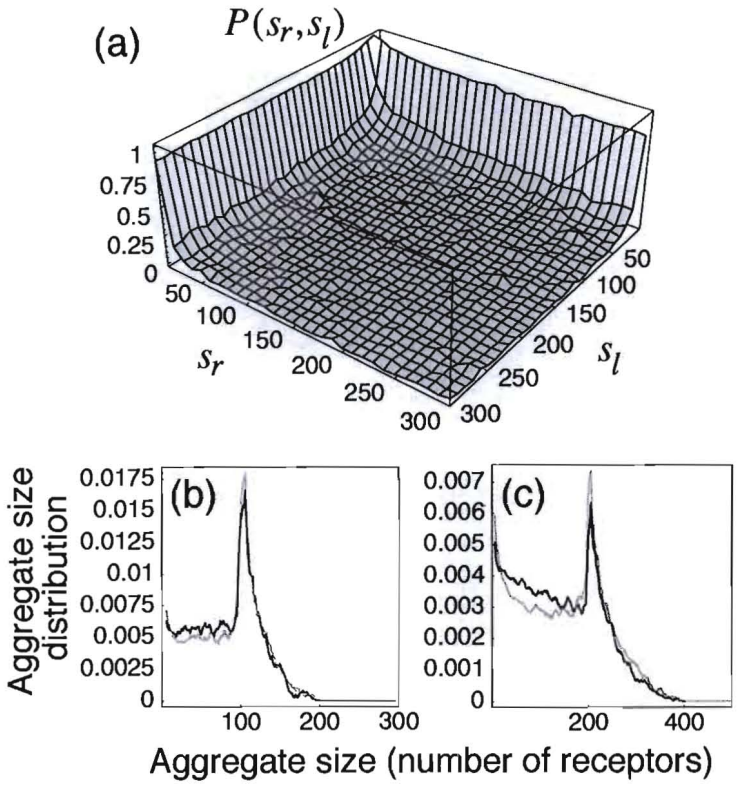
(a)
$P(s_r, s_l)$
1
0.75
0.5
0.25
0
50
100
150
200
250
300
s_r
s_l
(b)
(c)
Aggregate size distribution
0.0175
0.015
0.0125
0.01
0.0075
0.005
0.0025
0.007
0.006
0.005
0.004
0.003
0.002
0.001
Aggregate size (number of receptors)

Figure 12: