SpringSaLaD: a spatial, particle-based biochemical simulation platform with excluded volume

P. J. Michalski and L. M. Loew

SUPPORTING MATERIAL

SUPPLEMENTARY SECTION I: Validation Simulations

Here we present the results of simulations used to validate various components of SpringSaLaD.

Diffusion

The Langevin dynamics algorithms underlying SpringSaLaD are guaranteed to reproduce diffusive motion with the given diffusion constant. We verified this with a model consisting of a single molecule species, A, composed of one inert site, with diameter 1 nm and diffusion constant $D = 1\ \mu m^2/s$. We calculated the mean-squared-displacement as a function of time and compared it to the expected short- and long-time behavior.

At short times we expect $msd(t) \approx 2nDt$, where n is the dimensionality of the system. To verify this behavior, we simulated 2000 A molecules diffusing in a cube of side length 2001 nm, which is effectively 2000 nm after accounting for the radius of *A*. The occupied volume fraction, ϕ, is

$$\phi = \frac{2000\ V_{particle}}{V_{system}} = 1.31 \text{ x } 10^{-7},$$

which allows us to ignore excluded volume effects and treat this as a system of non-interacting particles. Short time behavior is expected for $t \ll \frac{L^2}{\pi^2 D} = 0.4$ s, and we simulated diffusion for 10 μs, well-within the short-time regime. In Fig. 1 we plot the msd as a function of time, and use a linear fit to determine the "measured" diffusion constant, which gives $D_{\text{measured}} = 1.0008\ \mu m^2/s$, which is within 0.1% of the expected value.

At long times we expect to observe confined diffusion. In 1D, the long-time behavior is given by

$$msd(t) = \frac{L^2}{6} - \frac{16L^2}{\pi^4} exp\left(-\frac{D\pi^2 t}{L^2}\right), \quad (1)$$

where L is the length of the system. To verify this behavior, we simulated 200 *A* molecules diffusing in a cube of side length 101 nm for 6 ms. The occupied volume fraction is $\phi = 1.0 \text{ x } 10^{-4}$, so we can treat this as a system of non-interacting particles. Figure 2 shows a plot of the msd along the x-coordinate as a function of time, along with a best fit to equation (1). The fit gives $D_{\text{measured}} = 0.984\ \mu m^2/s$, which is within 2% of the expected value.

SpringSaLaD can also be used to investigate the effects of crowding on diffusional motion. For 2- and 3-dimensional motion crowding simply renormalizes the diffusion constant, but in 1D crowding also changes the scaling exponent such that at high densities $msd(t) \sim t^{1/2}$. We investigated if SpringSaLaD could reproduce this behavior by using a model consisting of 200 membrane-embedded molecules with a diameter of 4 nm and $D = 1\ \mu m^2/s$. The system has a width of 4.5 nm, which effectively restricts motion to one dimension. The

length was varied from 80000 nm to 1000 nm to study system densities from 0.01, where we expect free diffusion, to 0.8, where we hoped to see crowding behavior. The system was simulated for $20\ \mu s$.

Figure 3 shows a plot of the msd as a function of time for each of the simulated systems. As the system density increases the msd is reduced, and at large densities the relationship no longer appears to be linear. We quantified the effects of crowding by fitting $msd(t) = at^b$ for each curve, and Fig. 4 shows a plot of the exponent, b, as a function of system density. At low densities $b \approx 1$, as expected, and b decreases as the density increases. At a density of 0.7 we found $b = 0.54$, and at a density of 0.8 we found $b = 0.46$, both of which are in good agreement with the theoretical result.

Creation/Decay Reactions

SpringSaLaD supports zeroeth-order particle creation reactions defined by a rate constant k_c , with units $\mu M\ s^{-1}$, and particle decay reactions defined by a rate constant k_d, with units s^{-1}. These reactions were verified using a system with the same composition as that used to verify diffusion (single site molecule A, diameter 1 nm, $D = 1\ \mu m^2/s$). If we let $p_m(t)$ be the probability of finding m Amolecules at time t, then the chemical master equation is

$$\dot{p}_m = k_d[(m+1)p_{m+1} - mp_m] + r_c[p_{m-1} - p_m], \tag{2}$$

where $r_c$has units of s^{-1} and is given by $r_c = 602\ V\ k_c$, where V is the volume of the system in μm^3, and the factor of 602 comes from Avogadro's constant and unit conversion factors. The macroscopic quantity of interest is the average number of particles at time t,

$$\langle m(t) \rangle = \sum_{m=0}^{\infty} m p_m(t)\ . \tag{3}$$

Equation (2) can be solved exactly for certain limiting cases.

1) If $k_d = 0$ then the equation describes a Poisson process. Assuming the system begins with no particles, $p_m(t) = \delta_{m,0}$, then the complete solution is

$$p_m(t) = \frac{(r_c t)^m}{m!}\ e^{-r_c t}\ , \tag{4}$$

$$\langle m(t) \rangle = r_c t\ . \tag{5}$$

We ran 500 simulations with $k_c = 0.5\ \mu M\ s^{-1}$ in a cube of side length 2000 nm, giving $r_c = 2408\ s^{-1}$. Simulations were run for 10 ms. Figure 5(a) shows the average number of particles as a function of time compared to the prediction of Eq. (5), and the results are seen to be in excellent agreement. As a more stringent test, Fig. 5(b) shows the full distribution at 10 ms compared to the predicted distribution of Eq. (4). The distributions are statistically equivalent.

2) If $r_c = 0$ then the equation describes a simple decay process. Assuming the system begins with M particles at time t=0, then the complete solution is

$$p_m(t) = \binom{M}{m} \; e^{-m\,k_d t}(1 - e^{-k_d t})^{M-m}\,, \tag{6}$$

$$\langle m(t)\rangle = M e^{-k_d t}\,. \tag{7}$$

We ran 500 simulations with $k_d = 100\ s^{-1}$ in a cube of side length 2000 nm. Simulations were run for 100 ms. Figure 6(a) shows the average number of particles as a function of time compared to the prediction of Eq. (6), and the results are seen to be in excellent agreement. As a more stringent test, Fig. 6(b) shows the full distribution at 10 ms compared to the predicted distribution of Eq. (7). The distributions are statistically equivalent.

3) If both creation and decay processes are allowed, then we can solve for the stationary distribution and the time-dependent average. Assuming the system begins with no particles, we obtain

$$p_m = \frac{\lambda^m}{m!}\, e^{-\lambda}\,, \tag{8}$$

$$\langle m(t)\rangle = \lambda(1 - e^{-k_d t}), \tag{9}$$

where $\lambda = r_c/k_d$. We ran 500 simulations with $k_c = 1\ \mu M\ s^{-1}$ ($r_c = 4816\ s^{-1}$) and $k_d = 250\ s^{-1}$, giving $\lambda = 19.264$, in a cube of side length 2000 nm. Simulations were run for 50 ms, and equilibrium was achieved at approximately 15 ms. Figure 7(a) shows the average number of particles as a function of time compared to the prediction of Eq. (9), and the results are seen to be in excellent agreement. As a more stringent test, Fig. 7(b) shows the full equilibrium distribution (averaged over the distributions at 20, 30, 40, and 50 ms) compared to the expected distribution from Eq. (8). The distributions are statistically equivalent.

Bimolecular Reactions

SpringSaLaD supports bimolecular association/dissociation reactions which are defined by an on rate, k_{on}, with units $\mu M^{-1}s^{-1}$, and an off rate, k_{off}, with units s^{-1}.

Single Species Combination

Consider a molecular species, A, which can dimerize to form a complex, *C*. Mass conservation implies $A_{\text{tot}} = A(t) + 2C(t)$, which shows that the system is completely characterized by the concentration of *A*. Let m denote the total number of free *A* molecules at time *t*, and assume there are M total molecules of A in the system. If we let n be the total number of C molecules, then $n = \frac{1}{2}(M - m)$. The chemical master equation is

$$\dot{p}_m = \frac{1}{2}k_{\text{off}}[(M-m+2)p_{m-2} - (M-m)p_m] + \frac{(k_{\text{on}}/N_A)}{V}[(m+2)(m+1)p_{m+2} - m(m-1)p_m], \tag{10}$$

where N_A is Avogadro's constant. It is straightforward to show that the average number of particles obeys

$$\frac{d\langle m\rangle}{dt} = k_{\text{off}}(M - \langle m\rangle) - \frac{2k_{\text{on}}}{VN_A}[\langle m^2\rangle - \langle m\rangle]\,. \tag{11}$$

We try to recover the deterministic limit by letting $A_{\text{tot}} = \frac{M}{VN_A}$, $= \frac{\langle m\rangle}{VN_A}$, and assuming zero variance, $\langle m^2\rangle = \langle m\rangle^2$, and find

$$\frac{dA}{dt} = k_{\text{off}}(A_{\text{tot}} - A) - 2k_{\text{on}}A^2 + 2k_{\text{on}}\frac{A}{VN_A}, \tag{12}$$

which only agrees with the classical result when $V \to \infty$.

The steady-state solutions to the master equation are

$$p_m = \frac{\left(\frac{M-1}{2}\right)!}{\left(\frac{M-m}{2}\right)!\,m!}\,\frac{(2r)^{(m-1)/2}}{{}_1F_1\left(-\frac{M-1}{2},\frac{3}{2},-\frac{r}{2}\right)} \quad \text{for } M \text{ and } m \text{ odd}, \tag{13}$$

$$p_m = \frac{\left(\frac{M}{2}\right)!}{\left(\frac{M-m}{2}\right)!\,m!}\,\frac{(2r)^{m/2}}{{}_1F_1\left(-\frac{M}{2},\frac{1}{2},-\frac{r}{2}\right)} \quad \text{for } M \text{ and } m \text{ even}, \tag{14}$$

where $r = \frac{VN_Ak_{\text{off}}}{2k_{\text{on}}}$ is a dimensionless parameter and ${}_1F_1(a,b,c)$ is the Kummer confluent hypergeometric function.

The complete time dependent solution can be obtained in the deterministic limit. If we define $a(t) = 2A(t)/K_D$, where $K_D = k_{\text{off}}/k_{\text{on}}$ is the equilibrium dissociation constant, then we find

$$a(t) = -\frac{1}{2} + p\,\frac{(a_0+1/2)\cosh(pk_{\text{off}}t) + p\sinh(pk_{\text{off}}t)}{(a_0+1/2)\sinh(pk_{\text{off}}t) + p\cosh(pk_{\text{off}}t)}, \tag{15}$$

where $p = \sqrt{a_0 + 1/4}$.

We modeled *A* as a molecule with a single site of radius 1 nm and $D = 2\ \mu m^2/s$, which could undergo a dimerization reaction with $k_{\text{on}} = 25\ \mu M^{-1}s^{-1}$ and $k_{\text{off}} = 50\ s^{-1}$. The simulation volume was a cube with side length 500 nm ($V = 0.125\ \mu m^3$), giving $r = 75.25$. The stochastic result only differs from the classical deterministic result when $r \ll 1$, and thus we expect the average concentrations to follow the deterministic results. The simulations consisted of 150 initially free A molecules ($A = 1.99\ \mu M$), and simulations were run for 400 ms. Figure 8(a) shows the average number of particles as a function of time compared to the prediction of Eq. (15), and Fig. 8(b) shows the full equilibrium distribution (averaged over the distributions at 100, 150, 200, 250, 300, 350, and 400 ms) compared to the expected distribution from Eq. (14). The observations match the expected values, except the average number of particles is larger than expected at short time scales. This is a consequence of the fact that the on-rate can be calculated to match the initial time

course or the equilibrium values, but not both, and we choose to match the on-rate to the equilibrium value. A further discussion of this issue can be found in Supplementary Section II.

Two Species Combination

We now consider the situation where two species, $A$and B, dimerize to form a complex, C. Mass conservation implies $A_{tot} = A(t) + C(t)$ and $B_{tot} = B(t) + C(t)$. Letting $A(0) = A_0, B(0) = B_0,$ and $C(0) = C_0$, then we must have $A(t) - A_0 = B(t) - B_0 = C_0 - C(t)$, and we see that only one of these species can be taken as an independent variable. To simplify the analysis we assume $C_0 = 0$, giving $A_0 = A_{tot}\ and\ B_0 = B_{tot}$. Let $m$be the number of A molecules, and let m_0 and n_0 be the initial numbers of A and B, respectively. Then the chemical master equation is

$$\begin{aligned}\dot{p}_m = {} & k_{\text{off}}\big[(m_0 - (m-1))p_{m-1} - (m_0 - m)p_m\big] \\ & + \frac{(k_{\text{on}}/N_A)}{V}[(m+1)(m+1-m_0+n_0)p_{m+1} - m(m-m_0+n_0)p_m],\end{aligned} \tag{16}$$

and the average number of particles obeys

$$\frac{d\langle m\rangle}{dt} = k_{\text{off}}[m_0 - \langle m\rangle] - \frac{k_{\text{on}}}{N_A V}\left[\langle m^2\rangle - (m_0 - n_0)\langle m\rangle\right]. \tag{17}$$

The same relations used to go from Eq. (11) to Eq. (12) now give

$$\frac{dA}{dt} = k_{\text{off}}[A_0 - A] - k_{\text{on}}[A^2 - (A_0 - B_0)A]. \tag{18}$$

The steady state solutions to the master equation are

$$p_m = \binom{m_0}{m}\frac{(n_0 - m_0)!}{(n_0 - m_0 + m)!}\frac{r^m}{{}_1F_1(-m_0, n_0 - m_0 + 1, -r)} \quad \text{for } m_0 < n_0\ (0 \le m \le m_0), \tag{19}$$

$$p_m = \binom{n_0}{m - (m_0 - n_0)}\frac{(m_0 - n_0)!\ r^{m-(m_0-n_0)}}{m!\ {}_1F_1(-n_0, m_0 - n_0 + 1, -r)} \quad \text{for } m_0 > n_0\ (m_0 - n \le m \le m_0), \tag{20}$$

where $r = VN_Ak_{\text{off}}/k_{\text{on}}$.

The complete time dependent solution can be obtained in the deterministic limit. Define $a(t) = A(t)/K_D$, then

$$a(t) = \frac{a_0 - b_0}{2} - \frac{1}{2} + s\frac{(s+u)e^{2sk_{\text{off}}t} - (s-u)}{(s+u)e^{2sk_{\text{off}}t} + (s-u)}, \tag{21}$$

where $u = (a_0 + b_0 + 1)/2$ and $s^2 = \frac{(a_0-b_0)^2}{4} + \frac{1}{4} + \frac{a_0+b_0}{2}$.

We modeled both A and B as molecules with a single site of radius 1 nm and $D = 1\ \mu m^2/s$, which could undergo a dimerization reaction with $k_{\text{on}} = 25\ \mu M^{-1}s^{-1}$ and $k_{\text{off}} = 50\ s^{-1}$. The simulation volume was a cube with side length 500 nm ($V = 0.125\ \mu m^3$), giving $r = 150.5$. The stochastic result only differs from the classical deterministic result when $r \ll 1$, and thus we expect the average concentrations to follow the deterministic results. The simulations consisted of 75 initially free A molecules ($A = 0.997\ \mu M$) and 150 initially free B molecules ($B = 1.99\ \mu M$), and simulations were run for 400 ms. Figure 9(a) shows the average

number of particles as a function of time compared to the prediction of Eq. (21), and Fig. 9(b) shows the full equilibrium distribution (averaged over the distributions at 100, 150, 200, 250, 300, 350, and 400 ms) compared to the expected distribution from Eq. (19). The observations match the expected values, except the average number of particles is larger than expected at short time scales, as discussed above.

Transition Reactions

SpringSaLaD supports transition reactions between states for each site, characterized by a single rate constant r with units s^{-1}. A transition reaction could be used to model the phosphorylation of a site (conversion between an "unphosphorylated" and "phosphorylated" state), the activation of an enzyme (conversion between "off" and "on" state), or many other types of reactions. These reactions may be restricted such that they only occur under certain conditions, and SpringSaLaD currently supports four such conditions.

1) No condition. Transition can always occur.
2) Free. Transition can only occur if the site is not bound by another site. This is useful if binding tends to lock a molecule in a particular state and thus prevent the transition.
3) Bound. Transition can only occur if the site is bound by specific site. A phosphorylation reaction would use this condition to ensure phosphorylation only occurs when the site is bound by the kinase.
4) Allosteric. Transition can only occur if a specific site *in the same molecule* is in a particular state.

Allosteric Activation

We test our implementation of conditions (1) and (4) using a simple model of allosteric activation of an enzyme. Consider an enzyme with two sites: an allosteric site, A, which can be either "off" or "on", and an enzymatic site, E, which can also be either "off" or "on." We assume the allosteric site can spontaneously turn on, and the enzymatic site can only turn on once the allosteric site is activated. We assume all molecules begin as $A_{\text{off}}E_{\text{off}}$, and assume there are no back reactions, which keeps this model simple and analytically tractable. The full reaction diagram is thus

$$A_{\text{off}}E_{\text{off}} \xrightarrow{r_o} A_{\text{on}}E_{\text{off}} \xrightarrow{r_1} A_{\text{on}}E_{\text{on}},$$

and we will assume $r_0 \neq r_1$to simplify the discussion.

Let m be the number of $A_{\text{off}}E_{\text{off}}$, n be the number of $A_{\text{on}}E_{\text{off}}$, and $k = M - m - n$ be the number of $A_{\text{on}}E_{\text{on}}$ (only m and n are independent variables), where M is the total number of molecules in the system. The chemical master equation for the joint probability distribution is

$$\dot{p}_{m,n} = r_0\big[(m+1)p_{m+1,n-1} - mp_{m,n}\big] + r_1\big[(n+1)p_{m,n+1} - np_{m,n}\big], \tag{22}$$

with solution

$$p_{m,n}(t) = \frac{M!}{m!\,n!\,(M-m-n)!}[f(t)]^m[g(t)]^n[h(t)]^{M-m-n}, \tag{23}$$

where

$$f(t) = e^{-r_0 t}, \tag{24}$$

$$g(t) = \frac{r_0}{r_1 - r_0}(e^{-r_0 t} - e^{-r_1 t}), \tag{25}$$

$$h(t) = \frac{r_1}{r_1 - r_0}(1 - e^{-r_0 t}) - \frac{r_0}{r_1 - r_0}(1 - e^{-r_1 t}). \tag{26}$$

The individual probability distributions may be obtained by summing over the joint probability distribution, for example, $p_m(t) = \sum_{n=0}^{M-m} p_{mn}(t)$, gives the probability distribution for $A_{\text{off}}E_{\text{off}}$. We obtain

$$p_m(t) = \binom{M}{m}[f(t)]^m[1 - f(t)]^{M-m} \quad \text{for } A_{\text{off}}E_{\text{off}} \tag{27}$$

$$p_n(t) = \binom{M}{n}[g(t)]^n[1 - g(t)]^{M-n} \quad \text{for } A_{\text{on}}E_{\text{off}}, \tag{28}$$

$$p_k(t) = \binom{M}{k}[h(t)]^k[1 - h(t)]^{M-k} \quad \text{for } A_{\text{on}}E_{\text{on}}, \tag{29}$$

with average values

$$\langle m \rangle = f(t) \quad \text{for } A_{\text{off}}E_{\text{off}} \tag{30}$$

$$\langle n \rangle = g(t) \quad \text{for } A_{\text{on}}E_{\text{off}}, \tag{31}$$

$$\langle k \rangle = h(t) \quad \text{for } A_{\text{on}}E_{\text{on}}. \tag{32}$$

We simulated 200 molecules in a cubic container with side length 500 nm for a total of simulation time of 5 ms. The reaction rates were $r_0 = 500\ s^{-1}$ and $r_1 = 300\ s^{-1}$, and we ran 500 independent simulations. Figures 10(a) and 10(c) plot the time dependence of the average values of $A_{\text{off}}E_{\text{off}}$ and $A_{\text{on}}E_{\text{on}}$, respectively, and compares them to the predictions of Eqs. (30) and (32). The results are in excellent agreement with the expected values. Figures 10(b) and 10(d) show the full probability distributions at 1.5 ms and compare them to the predicted distributions from Eqs. (27) and (29). The distributions are statistically equivalent.

Enzymatic Activation

We test our implementation of conditions (2) and (3) with a simple model of ligand-mediated enzymatic activation. Consider a system with two species, a ligand which can be either inactive or active, and an enzyme which can be either off or on. Denote the ligand by L_k , for k="inact" or "act", and denote the enzyme by E_j , where j="off" or "on". Assume that the ligand can convert between its active and inactive states when it is not bound to the enzyme, and assume that the enzyme can turn on only when bound by the ligand. To simplify the system, assume that the ligand cannot dissociate from the active enzyme. Then the complete set of reactions is as follows:

1) Two reactions describing ligand activation and deactivation. We assume that the ligand can only switch states when it is free (unbound).

$$L_{inact} \xrightarrow{k\_L\_act} L_{act}$$

$$L_{act} \xrightarrow{k\_L\_inact} L_{inact}$$

2) Two reactions describing E activation and deactivation. We assume that E can only switch states while it is bound by ligand.

$$L_{act} \cdot E_{off} \xrightarrow{k\_E\_act} L_{act} \cdot E_{on}$$

$$L_{act} \cdot E_{on} \xrightarrow{k\_E\_inact} L_{act} \cdot E_{off}$$

3) Two reactions (association and dissociation) describing ligand binding and unbinding. We assume that binding is reversible when E is off, but dissociation is prevented when E is on.

$$L_{act} + E_{off} \xleftrightarrow{k_{on},\ k_{off}} L_{act} \cdot E_{off}$$

It is unlikely that analytic solutions exist for this system. Thus, we implemented this model in Virtual Cell as a non-spatial stochastic simulation. The VCell solvers have been well-validated and we take this result as the "expected" solution. The VCell model is publicly available as "Enzymatic-Activation" under username "pjmichal." The model parameters are listed in Table S1.

The simulation consisted of 60 ligand molecules initially in the inactive state, and 40 enzyme molecules initially in the off state. The system consisted of a cube with side length 200 nm, and the total simulation time was 0.2 seconds. Averages were taken over 500 independent simulations. Figures 11(a) and 11(c) plot the time dependence of L_{act} and E_{on}, respectively. The results are in excellent agreement with the VCell simulations. Figures 11(b) and 11(d) show the full equilibrium probability distributions averaged over the distributions at 40, 80, 120, 160, and 200 ms, along with the distribution calculated over 10,000 non-spatial stochastic VCell simulations. The distributions are statistically equivalent.

SUPPLEMENTARY SECTION II: Derivation of On-Rate Parameter

Reversible vs Irreversible Reactions: Discussion of the Smoluchowski Model

As we mentioned in Supplementary Section I, SpringSaLaD uses a relation between the macroscopic on-rate and the corresponding microscopic reaction rate which gives the expected equilibrium distribution but not the correct initial dynamics for a bimolecular reaction. Surprisingly, it is impossible to choose a microscopic rate which will recapitulate both the initial dynamics and equilibrium distributions. We will illustrate this result with the simple and well-known Smoluchowski model.

The Smoluchowski model assumes two spherical molecules, A and B, react immediately upon collision to form a complex. The aim of the model is to relate the reaction radii to the macroscopic on-rate. Let species A (B) have radius R_A (R_B) and diffusion constant D_A (D_B). We assume A and B are dilute, and choose a coordinate system centered on a single B molecule. Through straightforward rescaling we can consider an equivalent system where A is a point particle with diffusion constant $D = D_A + D_B$ and B is a sphere of radius $R = R_A + R_B$ fixed at the origin. If we let $c(r)$ be the concentration of A molecules, then $c(r)$ satisfies the diffusion equation, $\partial c/\partial t = D\nabla^2 c$. The surface of B is an absorbing boundary for A because the particles react immediately upon collision, and thus one boundary condition is $c(r = R) = 0$. The conventional Smoluchowski model assumes that the concentration of A goes to a constant at infinity, $c(r = \infty) = c_\infty$. With these boundary conditions the solution to the diffusion equation is

$$c(r) = c_\infty(1 - R/r).$$

The flux density of A at the surface of this single B molecule is $j_R = -D\left(\frac{\partial c}{\partial r}\right)_{r=R} = -c_\infty D/R$, and the total flux at a single B molecule is thus $J = \int d^2 r j_R = -4\pi DRc_\infty$. In the dilute limit each B molecule is assumed to be independent, and thus the total rate of destruction of A molecules is

$$\frac{dN_A}{dt} = JN_B = -4\pi DRN_B c_\infty.$$

We divide through by the volume of the system to convert from numbers to concentrations, and we associate $c_\infty = A$ to arrive at

$$\frac{dA}{dt} = -4\pi DR\ A\ B.$$

Comparison with the macroscopic rate equation, $\dot{A} = -k_{on}AB$ shows that the macroscopic on-rate is given by

$$k_{on} = 4\pi DR,$$

which is the classical Smoluchowski result.

Now consider the same system in equilibrium, where a dissociation reaction produces an A and B molecule from the complex. Assume, for simplicity, that upon dissociation the newly separated A and B molecules are exactly a distance $a > R$ apart. The important observation is that for $r > a$ the fluxes due to the forward and backward reactions must cancel in equilibrium, and thus $c(r) = c_0$ for $r \geq a$. This condition must replace the boundary condition at infinity which was used above, and the new solution to the diffusion equation is

$$c(r) = \frac{ac_0}{a-R}\left(1 - \frac{R}{r}\right).$$

Following the same steps as in the preceding paragraph to relate k_{on} and R , we find

$$k_{on} = \frac{4\pi DR}{1 - R/a}.$$

Thus, the reaction radius which reproduces the equilibrium (reversible) distribution is not the same as the reaction radius for the irreversible reaction considered in the preceding paragraph. In fact, we find

$$R_{rev} = \frac{R_{irr}}{1 + R_{irr}/a},$$

showing that the reversible reaction radius is always smaller than the irreversible reaction radius. If R_{rev} is used then the initial rate of reaction will be slower than expected from the macroscopic k_{on}, but the equilibrium distribution will be correct. Conversely, if R_{irr} is used then there will be more complex than expected in "equilibrium," but the transient dynamics will be correct.

The decision between the two radii must be made based on how the macroscopic on-rate was measured. A review of the literature suggests that the on-rate is usually inferred by directly measuring k_{off} and K_D, and then using $k_{on} = \frac{k_{off}}{K_D}$. In this case the measured on-rate is an equilibrium rate, and it is correct to use R_{rev}.

SpringSaLaD relates the macroscopic on-rate to a first order reaction rate, λ, instead of a reaction radius, but the same principle applies, namely, there is a different relationship between k_{on} and λ for irreversible and reversible reactions. SpringSaLaD uses the relationship which matches equilibrium, and this is the reason the transient dynamics show a lower reaction rate than might be expected, as shown in 8(a) and 9(a).

Derivation of SpringSaLaD bimolecular on-rate

SpringSaLaD enforces excluded volume, and to allow this feature each site must be characterized by two radii, as illustrated in Figure 12: 1) a physical radius, ρ, which represents the excluded volume, and 2) a reaction radius, $R > \rho$, which determines the maximum distance between two sites such that they can undergo a binding reaction. Specifically, if the reaction radii of two binding partners overlap, then those sites can undergo a binding reaction with probability $P = \lambda\, dt$, where λ is a first-order reaction rate. The goal here is to relate λ to the macroscopic on-rate, k_{on}.

We make the same transformations as in the Smoluchowski model, which allows us to consider A as a point particle with diffusion constant $D = D_A + D_B$, and B as fixed at the origin with physical radius $\rho = \rho_A + \rho_B$ and reaction radius $R = R_A + R_B$. The diffusion equation is

$$\frac{\partial c}{\partial t} = D\nabla^2 c - \lambda c \quad \text{for} \quad r \le R,$$

$$\frac{\partial c}{\partial t} = D\nabla^2 c \quad \text{for} \quad r > R,$$

with the following four boundary conditions:

1) A reflecting boundary at $r = \rho$ to account for excluded volume, $\left(\partial c/\partial r\right)_{r=\rho} = 0$.
2) The equilibrium condition forces no flux beyond the dissociation distance, and thus $c(r = a) = c_0$, and for simplicity we assume $a \geq R$. (It is also possible to solve the system with $\rho \leq a \leq R$, but SpringSaLaD uses $a = R$ and thus only the former solution is required.)
3) Continuity of concentration at $r = R$, $c(R_+) = c(R_-)$, where $R_\pm = R \pm \varepsilon$ as $\varepsilon \to 0$.
4) Continuity of flux at $r = R$, $\left(\partial c/\partial r\right)_{R_+} = \left(\partial c/\partial r\right)_{R_-}$.

The general solution to the diffusion equation is

$$c(r) = \begin{cases} \frac{b_0}{r}\sinh\left(\frac{r-\rho}{r_0}\right) + \frac{b_1}{r}\cosh\left(\frac{r-\rho}{r_0}\right) & \text{for } \rho \leq r \leq R \\ b_2 + \frac{b_3}{r} & \text{for } R < r \leq a \end{cases},$$

where $r_0^2 = D/\lambda$ and the b_i are constants to be fit by the boundary conditions.

Boundary conditions (1) and (2) can be used to eliminate b_1 and b_2, giving

$$c(r) = \begin{cases} \frac{b_0}{r}\left[\sinh\left(\frac{r-\rho}{r_0}\right) + \frac{\rho}{r_0}\cosh\left(\frac{r-\rho}{r_0}\right)\right] & \text{for } \rho \leq r \leq R \\ c_0 + \frac{b_3}{a}\left(1 - \frac{a}{r}\right) & \text{for } R < r \leq a \end{cases}.$$

Boundary condition (3) gives

$$\frac{b_0}{R}\left[\sinh\left(\frac{R-\rho}{r_0}\right) + \frac{\rho}{r_0}\cosh\left(\frac{R-\rho}{r_0}\right)\right] = c_0 + \frac{b_3}{a}\left(1 - \frac{a}{R}\right),$$

while boundary condition (4) gives

$$b_3 = -b_0\left\{\sinh\left(\frac{R-\rho}{r_0}\right) + \frac{\rho}{r_0}\cosh\left(\frac{R-\rho}{r_0}\right) - \frac{R}{r_0}\left[\cosh\left(\frac{R-\rho}{r_0}\right) + \frac{\rho}{r_0}\sinh\left(\frac{R-\rho}{r_0}\right)\right]\right\}.$$

These can be solved to eliminate b_0 and b_3, and we finally arrive at

$$c(r) = \begin{cases} c_0\frac{a}{rQ_1}\left[r_0\sinh\left(\frac{r-\rho}{r_0}\right) + \rho\cosh\left(\frac{r-\rho}{r_0}\right)\right] & \text{for } \rho \leq r \leq R \\ c_0\left[1 - \frac{Q_0}{Q_1}\left(1 - \frac{a}{r}\right)\right] & \text{for } R < r \leq a \end{cases},$$

where

$$Q_0 = r_0\sinh\left(\frac{R-\rho}{r_0}\right) + \rho\cosh\left(\frac{R-\rho}{r_0}\right) - \frac{R}{r_0}\left[\rho\sinh\left(\frac{R-\rho}{r_0}\right) + r_0\cosh\left(\frac{R-\rho}{r_0}\right)\right],$$

and

$$Q_1 = r_0\sinh\left(\frac{R-\rho}{r_0}\right) + \rho\cosh\left(\frac{R-\rho}{r_0}\right) - \left(\frac{R}{r_0}\right)\left(1 - \frac{a}{R}\right)\left[\rho\sinh\left(\frac{R-\rho}{r_0}\right) + r_0\cosh\left(\frac{R-\rho}{r_0}\right)\right].$$

We now proceed as in the Smoluchowski case: calculate the total flux at the surface of B, write an expression for the rate of change of the concentration of A, and compare that to the macroscopic kinetic equation. We find

$$k_{on} = \frac{4\pi RD\phi}{1 - \frac{R}{a}\phi},$$

where

$$\phi = 1 - \left(\frac{r_0}{R}\right)\frac{r_0 \sinh\left(\frac{R-\rho}{r_0}\right) + \rho \cosh\left(\frac{R-\rho}{r_0}\right)}{\rho \sinh\left(\frac{R-\rho}{r_0}\right) + r_0 \cosh\left(\frac{R-\rho}{r_0}\right)}.$$

We can check that this reduces to the Smoluchowski result and Erban-Chapman results in the appropriate limits. The Smoluchowski model is recovered by letting $\lambda \to \infty$, in which case the particles react immediately upon collision of their reaction radii. This corresponds to taking $r_0 \to 0$, whereupon $\varphi \to 1$ and we find

$$k_{on,Smol} = \lim_{\phi \to 1} k_{on} = \frac{4\pi RD}{1 - \frac{R}{a}},$$

which is indeed the Smoluchowski result derived above. The Erban-Chapman limit corresponds to letting $\rho \to 0$, which gives

$$\phi = 1 - \frac{r_0}{R}\tanh\left(\frac{R}{r_0}\right),$$

which is the expected result.

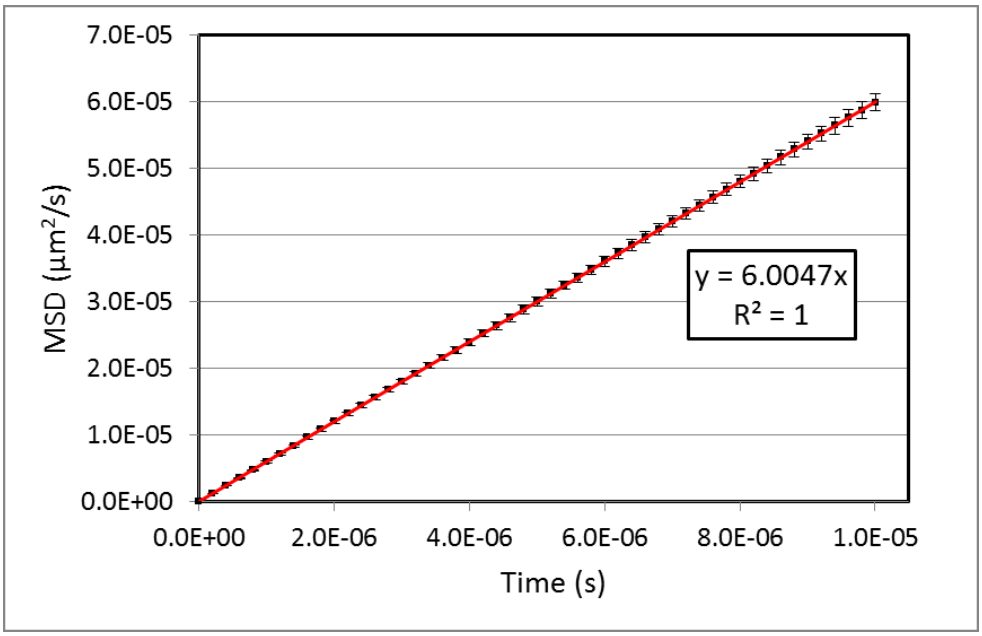

Figure 1: Validation of short-time diffusion. A plot of MSD vs time for the system described in the text. The plot shows the average and standard deviation from 100 simulations.

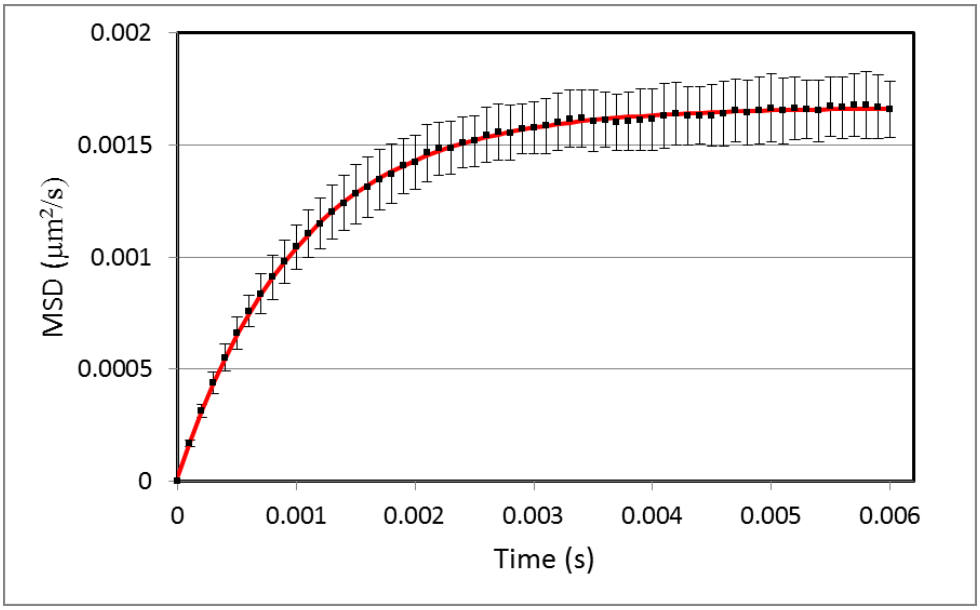

Figure 2: Validation of long-time diffusion. A plot of MSD vs time for the system described in the text. The plot shows the average and standard deviation from 100 simulations.

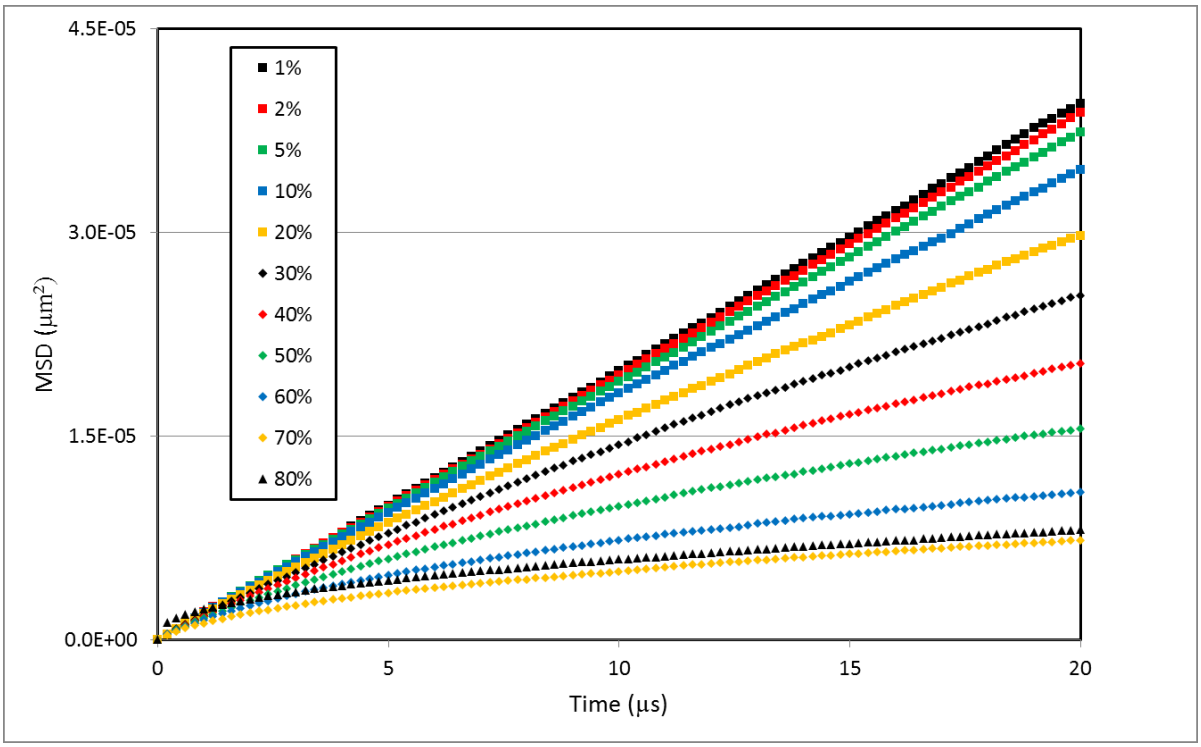

Figure 3: Effects of crowding on 1D diffusion. A plot of MSD vs time for the system described in the text as the density is varied, as indicated in the legend.

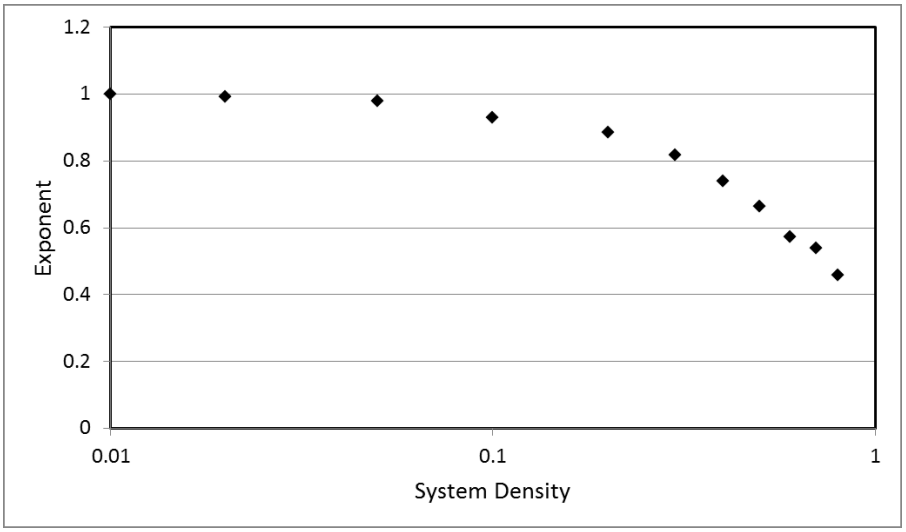

Figure 4: Effects of crowding on the scaling exponent. Each curve in Fig. 3 was fit to $msd(t) = at^b$, and here we plot b as a function of system density. At low densities b matches the exponent for free diffusion, while at large densities b approaches the theoretical result for crowded diffusion.

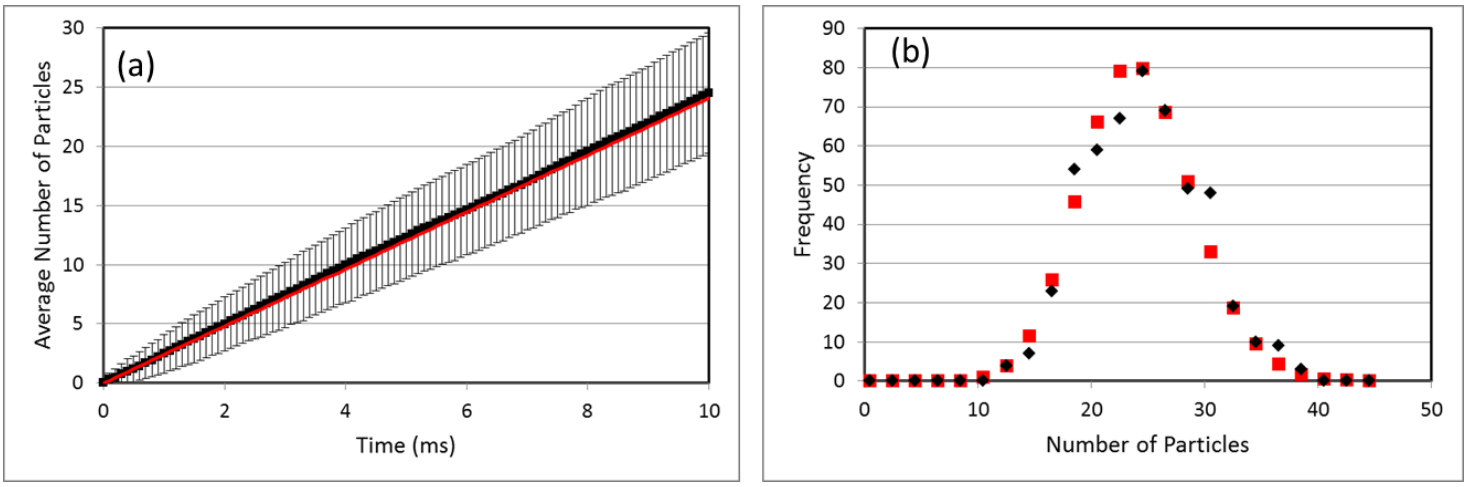

Figure 5: Validation of creation reactions. (a) Average number of particles as a function of time for the system described in the text. Average is taken over 500 independent simulations, error bars are standard deviation. The red line is the expected solution, Eq. (5). (b) The full distribution at 10 ms calculated over 500 runs (black diamonds) and the expected distribution according to Eq. (4) (red squares). The two distributions are statistically equivalent ($\chi^2 = 18.52, p = 0.07 > 0.05$).

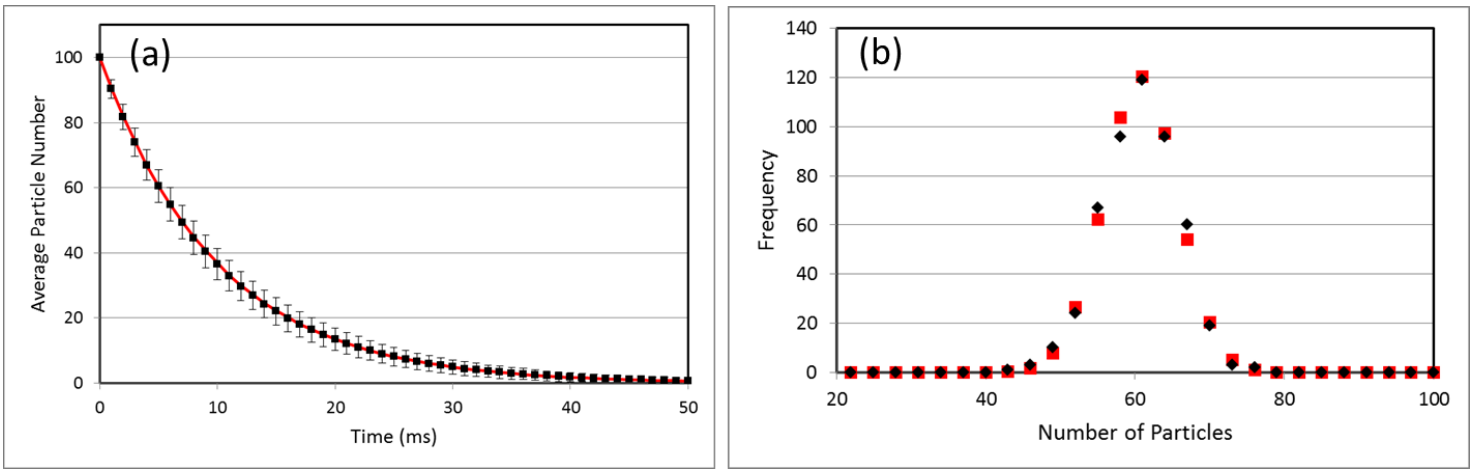

Figure 6: Validation of decay reactions. (a) Average number of particles as a function of time for the system described in the text. Average is taken over 500 independent simulations, error bars are standard deviations. The red line is the expected solution, Eq. (7). (b) The full distribution at 5 ms calculated over 500 run (black diamonds) and the expected distribution according to Eq. (6) (red squares). The two distributions are statistically equivalent ($\chi^2 = 3.30, p = 0.86 > 0.05$).

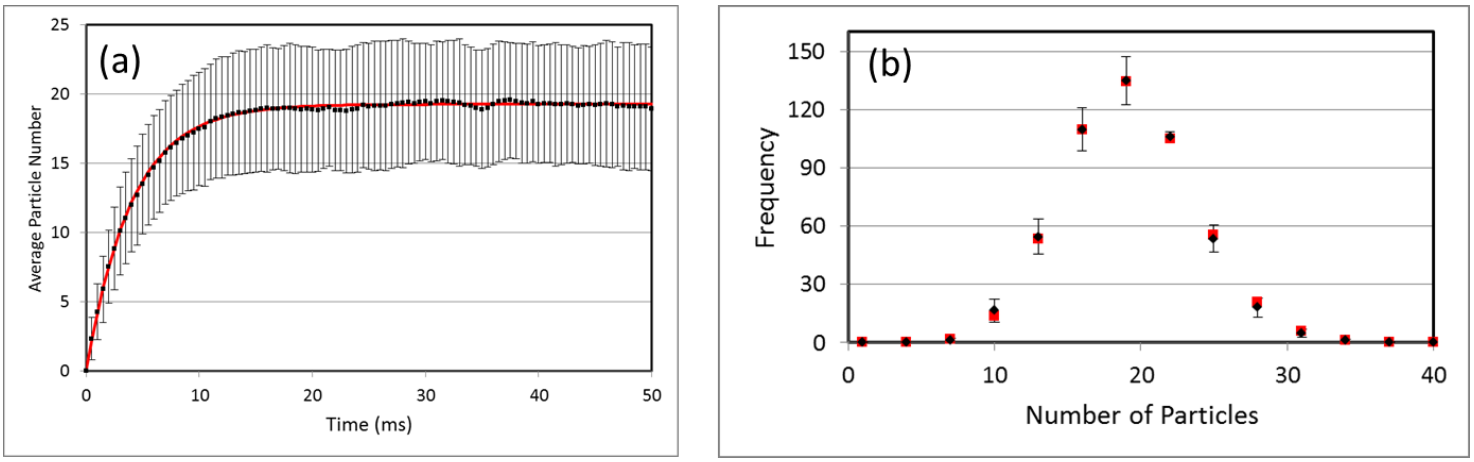

Figure 7: Validation of creation and decay reactions. (a) Average number of particles as a function of time for the system described in the text. Average is taken over 500 independent simulations, error bars are standard deviation. The red line is the expected solution, Eq. (9). (b) The full equilibrium distribution (black diamonds) compared to the expected distribution (red squares) from Eq. (8). The plot shows the average and standard deviation of the full distributions at 20, 30, 40, and 50 ms. The two distributions are statistically equivalent ($\chi^2 = 1.20,\ p = 0.98 > 0.05$).

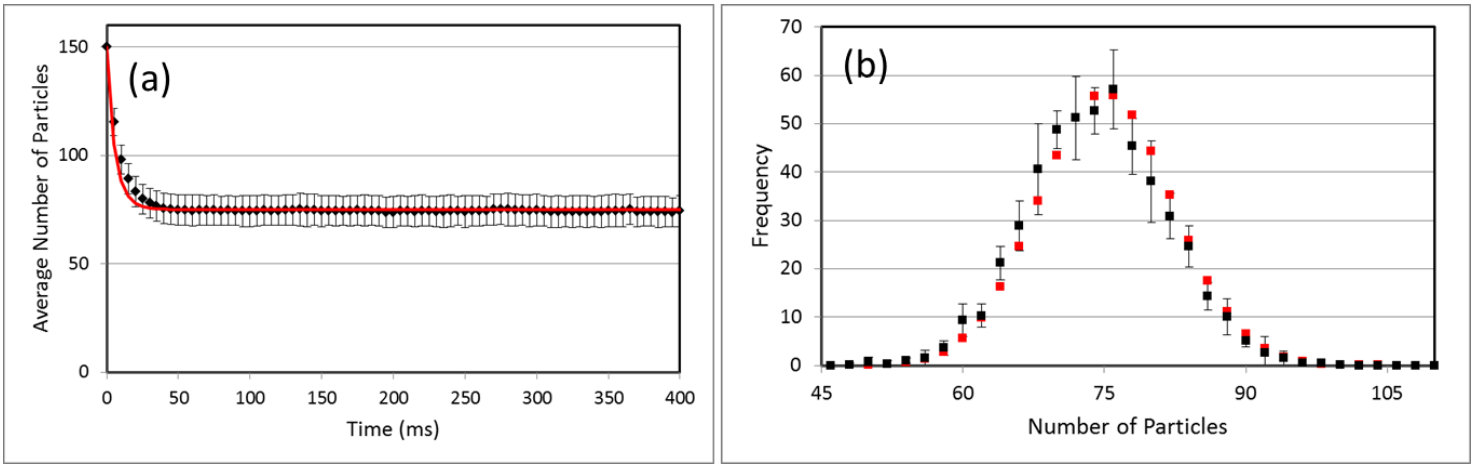

Figure 8: Validation of single species bimolecular reactions. (a) Average number of particles as a function of time for the system described in the text. Average is taken over 500 independent simulations, error bars are standard deviation. The red line is the expected solution, Eq. (15). (b) The full equilibrium distribution (black squares) compared to the expected distribution (red squares) from Eq. (14). The plot shows the average and standard deviation of the full distribution at 100, 150, 200, 250, 300, 350, and 400 ms. The two distributions are statistically equivalent ($\chi^2 = 10.21,\ p = 0.81 > 0.05$).

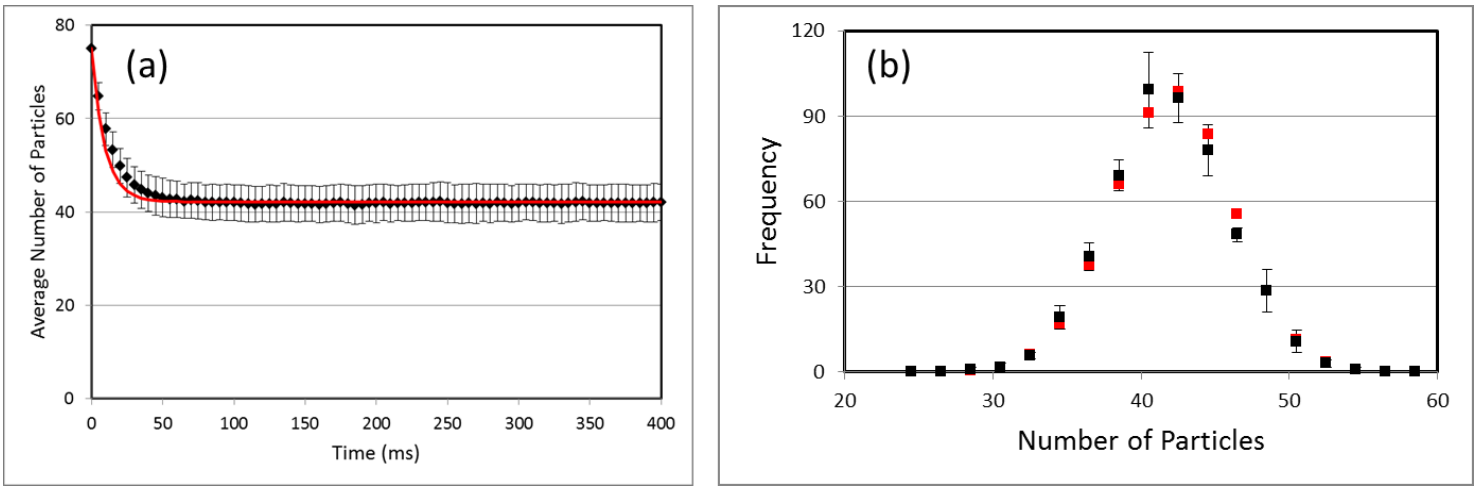

Figure 9: Validation of two species bimolecular reactions. (a) Average number of A particles as a function of time for the system described in the text. Average is taken over 500 independent simulations, error bars are standard deviation. The red line is the expected solution, Eq. (21). (b) The full equilibrium distribution (black squares) compared to the expected distribution (red squares) from Eq. (19). The plot shows the average and standard deviation of the full distribution at 100, 150, 200, 250, 300, 350, and 400 ms. The two distributions are statistically equivalent ($\chi^2 = 2.92,\ p = 0.98 > 0.05$).

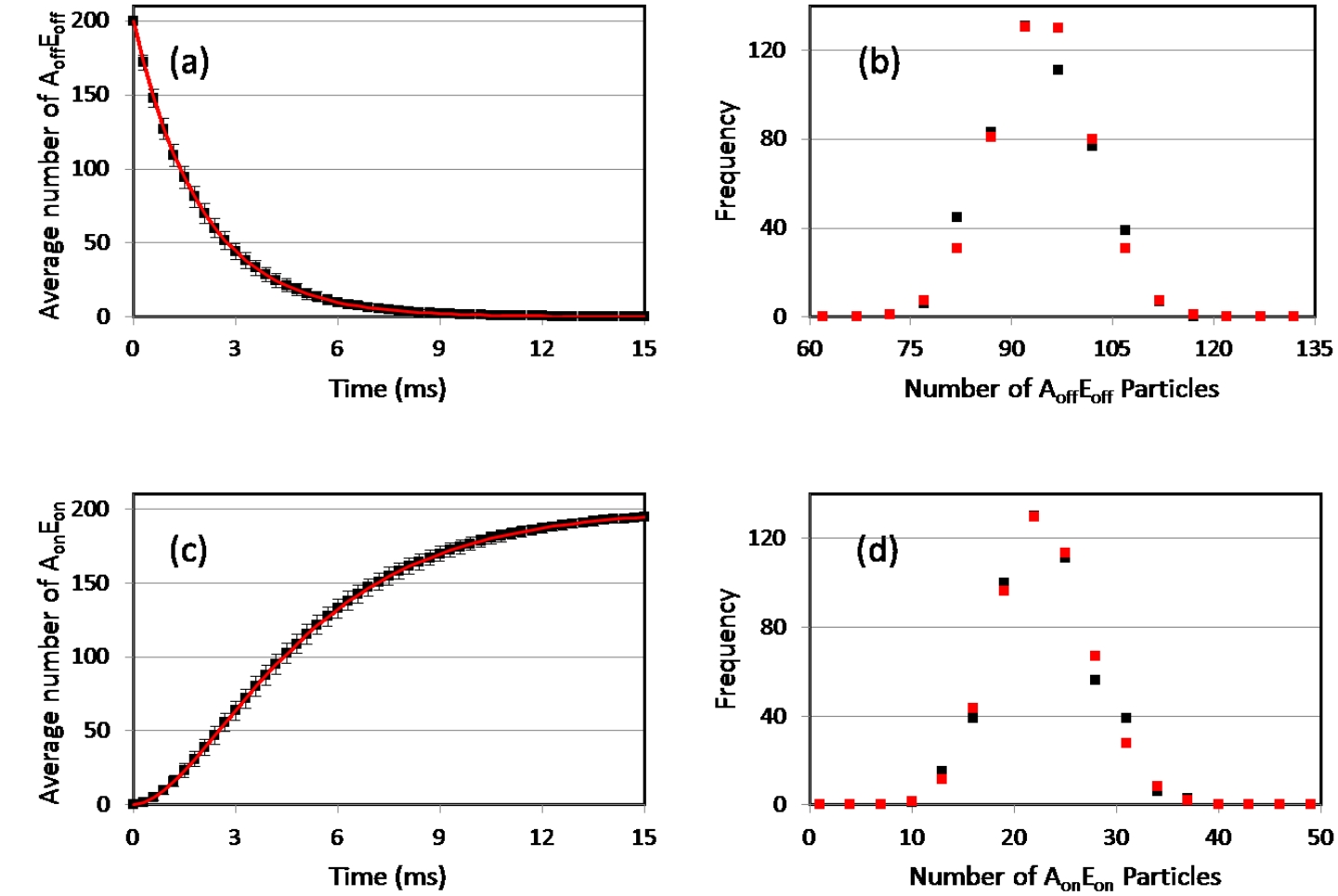

Figure 10: Validation of allosteric transition reactions. (a,c) Average number of $A_{\text{off}}E_{\text{off}}$ (a) and $A_{\text{on}}E_{\text{on}}$ (c) molecules as a function of time, for the system described in the text. The average is taken over 500 independent simulations, and the error bars are standard deviation. The red lines are the predictions from Eqs. (30) and (32). (b,d) The full distributions at 1.5 ms for $A_{\text{off}}E_{\text{off}}$ (b) and $A_{\text{on}}E_{\text{on}}$(d) (black squares) compared to the expected distributions (red squares) from Eqs. (27) and (29). Both observed distributions are statistically equivalent to their expected distributions ($A_{\text{off}}E_{\text{off}}$: $\chi^2 = 11.80,\ p = 0.11 > 0.05$, $A_{\text{on}}E_{\text{on}}$: $\chi^2 = 9.15,\ p = 0.24 > 0.05$).

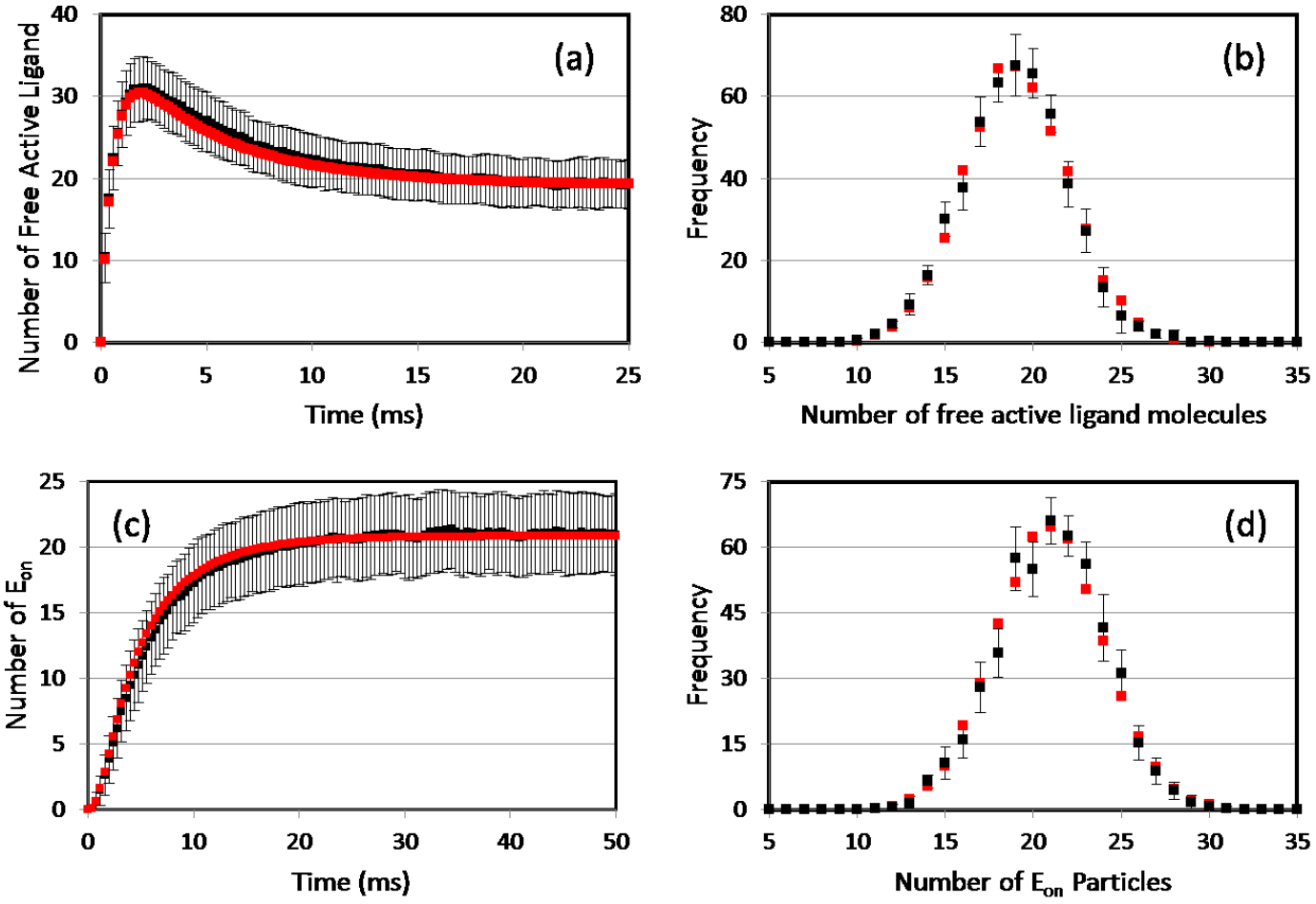

Figure 11: Validation of enzymatic activation reactions. (a,c) Average number of L_{act} (a) and E_{on} (c) molecules as a function of time, for the system described in the text. The average is taken over 500 independent simulations, and the error bars are standard deviation. The red lines are the results of deterministic VCell simulations. (b,d) The full equilibrium distribution for L_{act} (b) and E_{on}(d) (black squares) compared to the expected distributions. The expected distributions represent the results of 10000 VCell non-spatial stochastic simulations. The observed distributions are the average over the distributions at 40, 80, 120, 160, and 200 ms. Both observed distributions are statistically equivalent to their expected distributions (L_{act}: $\chi^2 = 3.90,\ p = 0.99 > 0.05$, E_{on}: $\chi^2 = 5.55,\ p = 0.96 > 0.05$).

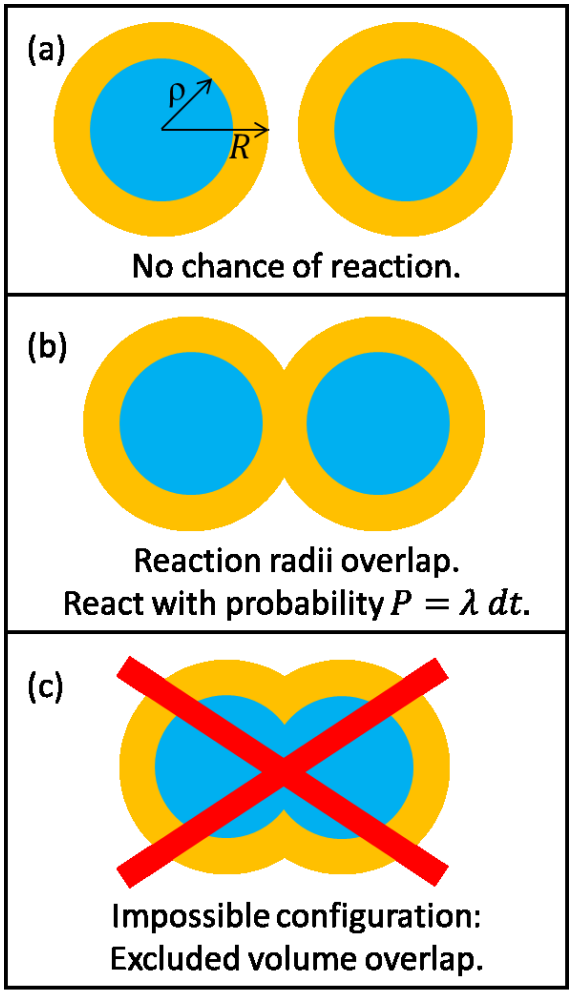

Figure 12: Schematic illustration of the relationship between the physical radius, ρ, and reaction radius, R, used in defining the bimolecular reaction rate λ.

Parameter Name	Value
k_L_act	$1000\ s^{-1}$
k_L_inact	$500\ s^{-1}$
k_E_act	$1000\ s^{-1}$
k_E_inact	$500\ s^{-1}$
k_{on}	$30\ \mu M^{-1}s^{-1}$
k_{off}	$100\ s^{-1}$

Table S1: Parameters used in the "Enzymatic Activation" model, which is used to validate the transition reactions. Identical parameters were used in both the VCell and SpringSaLaD models.