 4   **Working Groups**

4.1   Simulating macromolecular crowding with particle and lattice-based methods (Team 3)

Steven S. Andrews, Satya N. V. Arjunan, Gianfranco Balbo, Arne T. Bittig, Jerome Feret, Kazunari Kaizu, and Fei Liu

Abstract.   Many simulation algorithms have been developed to help model spatial structure in cellular systems, each of which is intended to represent reaction dynamics with high spatial resolution. In this study, we simulated the effects of macromolecular crowding on biochemical reaction rates to investigate which method actually performs best in practice. All 5 simulators investigated showed that diffusion-limited reaction rates decreased monotonically with the fractional crowder occupancy and activation-limited reactions exhibited an initial reaction rate increase with crowder occupancy (due to excluded volume effects). The eGFRD simulations were presumably highly accurate, but were too computationally intensive to be ideal for this problem. The Smoluchowski method as implemented in Smoldyn had simulation parameters that could be connected directly to physical parameters, and did not appear to exhibit simulation artifacts. The Smoluchowski method as implemented in NL-space produced qualitatively similar diffusion-limited results, but did not show a change of dynamics when changing to activation-limited conditions. Spatiocyte used a microscopic lattice, which enabled it to run very fast but introduced lattice artifacts in the results. Finally, we did not collect quantitative results with Kappa, but instead observed that Kappa can be used for this type of problem. Overall, this study showed that the detailed simulation methods substantially affect the results and that each of these simulators can still be improved.

4.1.1   Introduction

Many different biochemical simulation algorithms have been developed that are each intended to represent intracellular reaction dynamics with high spatial resolution and single-molecule precision [1]. These include Green's Function Reaction Dynamics [24], the Smoluchowski method as implemented in Smoldyn [4], the microscopic lattice method as implemented in Spatiocyte [5], and the particle-based method as implemented in NL-space [7, 8]. In addition, the next subvolume method [11] works at a slightly lower level of precision but is also intended to represent spatial detail accurately. It is straightforward to describe the differences between these methods, which we do below. However, the important question is how these algorithms actually perform in practice, which is not obvious from their descriptions. An understanding of this performance is clearly necessary for selecting the algorithm that is most appropriate for a specific modeling task. In this work, we investigated algorithm performance by comparing the abilities of the above-mentioned algorithms to accurately model bimolecular reaction rates in crowded spaces. This is a good test problem because crowded spaces are intrinsically difficult to simulate well.

Crowded spaces are also biologically important. In 1982, Fulton published that actively growing cells are about 17 to 26 percent protein and that red blood cells are about 35 percent protein [14]. These numbers are still commonly accepted. The vast majority of these proteins, and other macromolecular species such as RNAs and ribosomes, typically do not participate directly in any particular reaction that is of interest, but influence it indirectly through

volume exclusion, diffusion inhibition, and other effects. As such, cellular components can be classified as "reactants", which engage in the reaction of interest, and "crowders", which comprise everything else. Macromolecular crowding has several effects, including slowing diffusion, stabilizing protein folding, and accelerating bimolecular reactions, all of which have been reviewed extensively [26, 25, 16, 12]. Of these, our focus was on bimolecular reaction rates.

4.1.2   Description of simulation methods

Reaction models

The simulation methods that we investigated are based upon a couple of basic reaction models. Consider the generic irreversible chemical reaction A + B $\rightarrow$ C, which has reaction rate constant k. We define this reaction rate constant as the mass action reaction rate when the system is at steady-state. That is, k is defined from the mass action reaction kinetics equation $d[C]/dt = k[A][B]$, where square brackets represent chemical concentrations.

In the Smoluchowski model [21], A and B molecules diffuse according to mathematically ideal Brownian motion, meaning that molecules move with infinitely detailed random trajectories. These molecules do not interact with others of the same species, including through excluded volume effects. However, when A and B molecules collide together, where a collision is defined as their centers being separated by a distance equal to the sum of the two molecular radii, they react immediately to form a C molecule. From Smoluchowski's work [21], the steady-state reaction rate constant is

$$k = 4\pi\sigma D \tag{1}$$

where σ is the sum of the A and B radii and D is the sum of the A and B diffusion coefficients. This reaction rate is limited solely by diffusion, leading to its being called the diffusion-limited reaction rate.

The Collins and Kimball model [9] extends the Smoluchowski model by treating collisions between A and B molecules with the radiation boundary condition [10] rather than the absorbing boundary condition. In concept, this means that A and B molecules have a small probability of reaction at each collision, so they collide multiple times before they either react or diffuse apart without reacting. The assumption of mathematically ideal diffusion makes the actual model slightly more complicated than this because all dynamics need to be taken in the limit of small diffusive step sizes. In particular, it implies that any single collision between A and B molecules is essentially certain to be followed by an infinite number of collisions and that the reaction probability at each individual collision is infinitesimal. We refer the reader elsewhere for more thorough descriptions [20]. The result of this radiation boundary condition assumption is that the steady-state reaction rate arises from a combination of the diffusion-limited reaction rate, which gives the rate of initial collisions, and also the "intrinsic" reaction rate, which gives the rate of reaction after the first collision. The intrinsic reaction rate is also called the activation-limited reaction rate because it is the observed rate when the chemical reaction is strictly limited by molecules attaining sufficient activation energy to react, and not by the rate of diffusive collisions. In a form introduced by Noyes [18], the Collins and Kimball reaction rate constant is

$$\frac{1}{k} = \frac{1}{4\pi\sigma D} + \frac{1}{k_{int}} \tag{2}$$

where k_{int} is the intrinsic reaction rate constant.

The reaction-diffusion master equation model is yet a third model. It differs from the Smoluchowski and Collins and Kimball models in that it is based upon a macroscopic description rather than a microscopic description. It combines the assumptions of Fick's law for chemical diffusion [6] and mass action reaction kinetics for chemical reactions. For our particular example, the spatially-dependent concentrations of A, B, and C molecules change over time according to

$$[\dot{A}] = D_A \nabla^2 [A] - k[A][B]$$
$$[\dot{B}] = D_B \nabla^2 [B] - k[A][B] \tag{3}$$
$$[\dot{C}] = D_C \nabla^2 [C] + k[A][B]$$

where time and spatial dependencies are implied but not shown for the chemical concentrations.

Enhanced Green's Function Reaction Dynamics

Enhanced Green's Function Reaction Dynamics (eGFRD) is a particle-based method that simulates the Collins and Kimball reaction model exactly [23]. In it, non-overlapping spherical protective domains are drawn around each particle or pair of particles. Then, random times are drawn from the appropriate probability densities for the possible events that could happen, including particles diffusing to the edges of their domains, single particles reacting through unimolecular reactions, and pairs of particles reacting with each other. The smallest of these times is chosen and that particular event is performed. The system is then updated as necessary, which typically includes the computation of at least some new protective spheres and event times. Then, the next event in the queue is chosen, and so forth. Because the time is stepped from one reaction to the next, this is an event-based algorithm. See Takahashi and ten Wolde [23] for details.

Smoluchowski dynamics as implemented in Smoldyn

Smoldyn simulations perform a discrete-time version of the Smoluchowski model, using fixed time-steps [4]. At each time-step, Smoldyn displaces each molecule, on each spatial coordinate, by a value chosen from a Gaussian distributed probability density in order to simulate diffusion. It ignores all molecule interactions at this point. Next, Smoldyn performs surface interactions [2]. In the case of inert impermeable surfaces, such as those that we used in this work, it simulates reflection off of the surfaces using ballistic molecular trajectories. These are not based on the assumption that molecules in solution move with long straight-line trajectories, which they do not, but instead on the solution for the probability density of ideally diffusing molecules near planar surfaces, which is simulated exactly using ballistic trajectories [4]. Then, Smoldyn executes reactions for each A-B molecule pair that is separated by a "binding radius" or less. Smoldyn computes this binding radius before the simulation begins from the user's choices of reaction rate constant, the simulation time step, and the A and B diffusion coefficients so that the simulated steady-state reaction rate will be the same as the user's requested reaction rate constant. Although Smoldyn's reaction probability density upon collision is 1, as it is in the Smoluchowski model, Smoldyn actually simulates reaction dynamics in closer agreement with the Collins and Kimball model due to the fact that molecules can diffuse relatively long distances in each time step [4]. The choice of the simulation time step determines where the simulated reaction dynamics are on the continuum between being diffusion-limited and activation-limited.

Smoluchowski dynamics as implemented in ML-Space

ML-Space simulates reactions in a similar manner as Smoldyn. However, molecules in ML-Space have assigned radii. After ML-Space diffuses a molecule by Gaussian-distributed displacements, as in Smoldyn, it looks for molecule pairs with overlapping radii. If the molecules in the pair are non-reactive, the the displacement is reversed and another random displacement is attempted a (customizable) number of times before the last displacement is applied only partially such that the molecule at most touches another. Excluded volume effects are thus covered. If molecules are reactive (as specified in ML-Space's own attributed rule-based language), the respective changes (molecule property changes, replacement, consumption) are applied, the original collision is resolved, if necessary (i.e. if neither is consumed), by moving the colliding molecules apart such that they touch but not overlap, and to-be-produced entities, if present, are placed near the collision site without overlapping any present particles. Such a rule application may fail due to spatial constraints, i.e. non-resolvable collisions or no space for to-be-produced entities. The probability with which a reaction execution shall be attempted on collision can be taken from the ratio of the desired macroscopic rate constant and the theoretical diffusion-limited reaction rate arising from Smoluchowski's equation (1).

The main goal of ML-Space is to bring together individual-based simulation of larger spatial entities (large molecules or entire biological compartments) and population-, reaction-diffusion-based simulation of small particles as in the Next Subvolume Method [11]. However, we here focus on the purely continuous-space part, not the hybrid simulator.

Microscopic lattice method as implemented in Kappa

Kappa is a leading language for rule-based modeling (for defining the species and reactions that arise in the formation of multimeric complexes; [13]) and is also software for the same rule-based modeling. Even if spatial extension exists [22], in this work we developed a model in the core of Kappa, thanks to a non-spatial stochastic simulator that runs the Gillespie algorithm [15]. Our goal is to implement the Microscopic lattice method, primarily as an exercise to see whether this could be done.

More precisely, the simulation is done in three steps.

1. The first step is a self-assembling of the lattice of locations. Indeed, space is encoded as a rectangular box of agents, each agent denotes a location being connected to its six neighboring agents through some sites the name of which specifies the direction. So as to avoid border effects, each face of the cube is connected to its opposite one, so that particles can exit from one face and reenter through the opposite one.

2. The second step consists in spawning particles at random in the rectangular box. We assume that each location can contain at most one particle. We consider five kinds of particles: A, B, C, AB, and D. At the beginning, the system contains particles of kinds A, B, and C only.

3. The third step consists in diffusing the particles and letting them react according to the following reactions:

$$
\begin{aligned}
A + B &\rightarrow AB & @\, k_{AB} \\
AB &\rightarrow A + B & @\, k_{AB}^d \\
AB &\rightarrow B + C & @\, k_C
\end{aligned}
$$
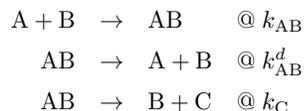
It is worth noticing that the particles of kind D do not react. The first reaction can apply only to adjacent particles. Moreover, the last two reactions require an adjacent location

to be free. Moreover, each particle diffuses to adjacent free locations at respective rates d_A, d_B, d_{AB}, d_C, and d_D along each of the six directions.

At each algorithm iteration, the next event (including diffusion of a molecule from one location to its neighboring one and reactions of molecules within adjacent locations) is selected according to its propensity, and the time between two consecutive events is randomly selected according to an exponential law the parameter of which is the overall amount of the propensities of all the potential events. Then, the algorithm repeats. Simulation stops when there are only 10 instances of As left in the system (either free or in AB).

Microscopic lattice method as implemented in Spatiocyte

Spatiocyte represents space using a fine hexagonal close-packed lattice, in which each lattice site can contain up to one molecule [5]. It performs events using a combination of event-driven and time-driven methods. For diffusion, all molecules that share a diffusion coefficient (e. g. those of the same species) are diffused periodically, at the frequency which produces the correct diffusion coefficient. Molecules cannot share lattice sites, so any non-reactive collisions result in molecules being put back to their starting locations. On the other hand, if two molecules collide and can react, then they react with a pre-determined probability that is calibrated to yield the correct reaction rate; if they don't react, then they are separated like other non-reactive collisions. Unimolecular reactions are performed with event-driven methods, using the Gillespie algorithm [15]. Spatiocyte chooses the event with the earliest time, which may be diffusive or unimolecular reactive, and executes it. Then, Spatiocyte updates the system and repeats.

4.1.3   Theory for crowding effects

Crowding affects irreversible association reactions in two primary ways. First, the crowders occupy volume, which reduces the volume available to the reactants and thus increases their effective concentrations. This increases reaction rates. Also, crowding slows diffusion, which reduces the rate at which reactants collide with each other. This decreases reaction rates. Although these qualitative effects have been well-known for many years, the actual amount by which crowding modifies bimolecular reaction rates is still an open question. Of particular note is recent modeling work by Kim and Yethiraj [17], who showed both the reaction acceleration and deceleration effects. However, their results were not based entirely on physical parameters, but instead were functions of their simulation parameters (their reaction probability upon collision), which limits their value.

The effects of crowding on reaction rates can be estimated some cases. Assume that reactions are irreversible, the crowders are stationary, and the reactants have sufficiently low concentrations that their excluded volume interactions can be ignored. In the activation-limited extreme, in which diffusion timescales are much faster than reaction timescales, the reactants are well-mixed throughout the available volume, meaning that which is not occupied by crowders. This volume is $V_{avail.} = V_{total}(1 - \phi)$, where V_{total} is the total system volume and ϕ is the fractional volume occupancy by crowders. From eq. 2, the reaction rate constant is simply k_{int}. Within the available volume, the reaction rate is

$$\frac{dn_C}{V_{avail.}dt} = k_{int}\frac{n_A}{V_{avail.}}\frac{n_B}{V_{avail.}}$$

where n_A, n_B, and n_C represent the numbers of A, B, and C molecules, respectively.

Substituting and simplifying leads to

$$\frac{dn_C}{dt} = \frac{k_{int} n_A n_B}{V_{total}(1 - \phi)}$$

$$\frac{dn_C}{V_{total} dt} = \frac{k_{int}}{1 - \phi} \frac{n_A}{V_{total}} \frac{n_B}{V_{total}}$$

$$\frac{d[\text{C}]}{dt} = \frac{k_{int}}{1 - \phi} [\text{A}][\text{B}]$$

$$k_{act.}(\phi) = \frac{k_{int}}{1 - \phi} \tag{4}$$

Thus, crowding causes activation-limited reactions to accelerate by the factor $1/(1 - \phi)$.

We are unable to solve for the diffusion-limited extreme, but offer a hypothesis instead. In this case, the available volume is still reduced by the same factor of $1 - \phi$, so it would make sense for reaction rates to be accelerated exactly as before. In addition, the diffusion coefficient is reduced from D to some crowding-dependent amount which we denote $D(\phi)$. This dependence varies depending on the precise crowding model. Combining these effects, our hypothesis is that the diffusion-limited reaction rate constant changes from eq. 1 to

$$k_{diff.}(\phi) = \frac{4\pi\sigma D(\phi)}{1 - \phi} \tag{5}$$

Although intuitively sensible, this derivation is not rigorous. In particular, the Smoluchowski reaction rate equation, eq. 1, is typically derived by computing the radial distribution function of B molecules around the A molecules. The presence of crowders likely changes this radial distribution function, although those effects were not accounted for here.

We are also unable to solve for the general diffusion-influenced reaction rate constant. However, we offer the hypothesis that the diffusion-limited and activation-limited reaction rates, in the presence of crowders, can be combined in the same way as in they are in the Collins and Kimball equation, eq. 2. This yields

$$k(\phi) = \left[\frac{1 - \phi}{4\pi\sigma D(\phi)} + \frac{1 - \phi}{k_{int}} \right]^{-1} \tag{6}$$

Below, we test these hypotheses with simulations.

4.1.4   Results and Discussion

Smoldyn

Smoldyn simulations were performed in a 50 x 50 x 50 nm^3 cube with periodic boundaries. Simulations ran for 10 µs in steps of 0.001 µs, and data were recorded every 0.01 µs. We generated crowders, using the SmolCrowd software, as randomly positioned non-overlapping spheres with 0.5 nm radii. These radii were then increased to 1 nm (which led to overlaps of up to 0.5 nm) as a simple way of accounting for radii of the A and B molecules that equaled 0.5 nm. This increase of the crowder radii enabled us to represent the A and B molecules as simple points, but for them to behave as though they had 0.5 nm radii. We computed the crowder volume fraction, ϕ, as the fraction of the simulation volume that was within at least one of these 1 nm radii crowder spheres.

Each simulation started with about 1000 randomly placed molecules for each of the three species, A, B, and tracers. All three species diffused with diffusion coefficients of $D_0 = 10$ nm^2/µs (equal to 10 µm^2/s, which is a typical, albeit slow, intracellular protein diffusion
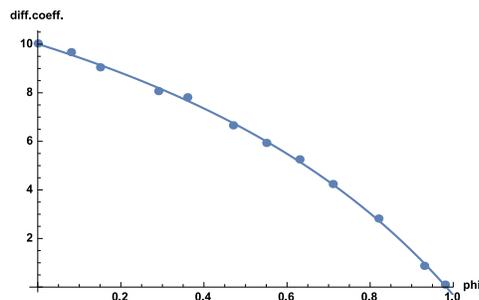
■ **Figure 1** Effective diffusion coefficient as a function of crowder volume fraction. Dots represent simulation data for tracer molecules and the line represents a least-squares best fit to the simulation data using a rational function, as described in the main text. These data were from simulations with $k_0 = 251.3$ nm^3/μs; the data for other simulations were essentially identical.

coefficient [3]). The tracer molecules did not participate in any reactions or interact with the A or B molecules. Instead, they simply diffused around the system, and we used their mean squared displacements at the end of each simulation to compute their effective diffusion coefficients and, by extension, the effective diffusion coefficients of the A and B molecules.

The A and B molecules reacted with each other with reaction rate constants (for uncrowded systems) of either $k_0 = 251.3$ or $k_0 = 25.13$ nm^3/μs (equal to 1.5×10^8 M^{-1}/s^{-1} and 1.5×10^7 M^{-1}/s^{-1}, both of which are extremely fast reaction rates). We chose the former rate constant because its binding radius in the Smoluchowski model, eq. 1, is 1 nm. We used it to investigate nearly diffusion-limited reactions and we used the latter rate constant to investigate more activation-limited reactions. Smoldyn reported that the effective activation-limited reaction rate constants for the two sets of simulations were 1238 and 33.83 nm^3/μs, respectively, which were computed from eq. 41 of Andrews and Bray [4]. From these and the k_0 values, the diffusion-limited reaction rate constants were 315.3 and 97.73 nm^3/μs, respectively. In contrast to the reactions introduced above, we used the reaction A + B → B here, so that the concentration of B stayed constant throughout the simulation. This simplified the reaction rate constant estimation, as described below. We ran each simulation 10 times and averaged the results for the 10 runs.

As expected, we found that effective diffusion coefficients decreased monotonically with the crowder occupancy, shown with dots in Figure 1. These data fit well to the rational function

$$D(\phi) = D_0 \frac{1 - a\phi}{1 - b\phi} \tag{7}$$

where a and b were fit parameters. The best fit, shown with the line in Figure 1, has $a = 1.02$ and $b = 0.48$. These fit parameters are sufficiently close to 1 and 1/2 to be suggestive of a theoretical basis to this fitting function, but we did not pursue it in this work. The percolation threshold, meaning the crowder occupancy where the effective diffusion coefficient becomes zero, is $\phi_{perc.} = 1/a = 0.98$.

To compute the steady-state reaction rate constant from simulation data, we first recorded the number of A molecules surviving as a function of time, with a typical example shown in Figure 2A. We then numerically differentiated these data according to the equation

$$k_i = -\frac{n_{A,i+1} - n_{A,i-1}}{(t_{i+1} - t_{i-1})n_{A,i}} \tag{8}$$

(a) Survival curve

(b) Rate coefficient
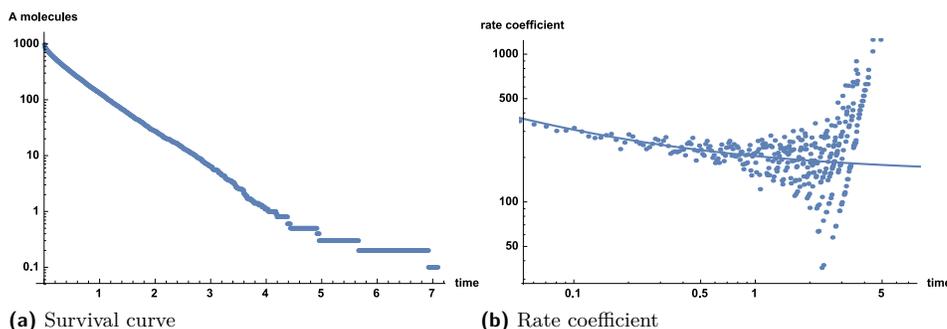
Figure 2 Data analysis for computing steady-state reaction rate constants. (A) Number of A molecules surviving as a function of time for the average of 10 simulations with $k_0 = 251.3 \text{ nm}^3/\mu s$ and $\phi = 0.47$; other data sets were qualitatively similar. (B) Points represent the reaction rate coefficient as a function of time, computed from the data shown in Panel A using eq. 8. The line is a best fit line to the points, using eq. 9

where k_i is the reaction rate at time point i, $n_{A,i}$ is the number of surviving A molecules at time point i, and t_i is the simulation time at time point i. This numerical derivative produced a very noisy reaction rate coefficient function, as shown in Figure 2B. Adding to the challenge of estimating the reaction rate constant, there is no sharp cut-off between the transient fast reaction rate coefficient at very short times and the steady-state reaction rate constant. Thus, we fit the reaction rate coefficient data with the following function, which has the form of the time-dependent reaction rate coefficient for both the Smoluchowski and Collins and Kimball models,

$$k(t) = c(1 + \frac{d}{\sqrt{t}}) \tag{9}$$

where c and d are fit parameters; c is also the steady-state reaction rate constant. This fit skipped the first 19 data points in order to reduce the effect of the short-time transient reaction rate. This fit also used the number of A molecules at each time point as a weighting parameter for the data points in order to give more weight to the less noisy data and less to the noisy data. As seen in Figure 2B, the resulting fits agreed with the data very well. Fitting to this function was possible because we kept the concentration of B molecules constant throughout a simulation.

Figure 3 shows the effect of the crowder volume occupancy on the steady-state reaction rate constant, for primarily diffusion-limited and primarily activation-limited situations. In both cases, the simulated reaction rate at zero crowder density, quantified with the process described above, agreed very closely with the input reaction rate constant (3.5% error for $k_0 = 251.3 \text{ nm}^3/\mu s$ and 0.1% error for $k_0 = 25.13 \text{ nm}^3/\mu s$), which gave us high confidence in our reaction rate quantification method. Both curves qualitatively agree with the predictions given above, in which diffusion-limited reactions are slowed down by crowders due to the slowed diffusion, and activation-limited reactions are accelerated by crowders due to the reduction of accessible volume. However, comparing the data points with the solid blue lines shows that the simulation data do not agree with our hypothesis. We computed these hypothesis curves from eq. 6, while using the empirical fit in eq. 7 for $D(\phi)$, the activation-limited reaction rate reported by Smoldyn for k_{int}, and the diffusion-limited reaction rate constants given above and eq. 1 to compute σ. Note that there are no adjustable parameters in this comparison.

(a) Nearly diffusion-limited

(b) Nearly activation-limited
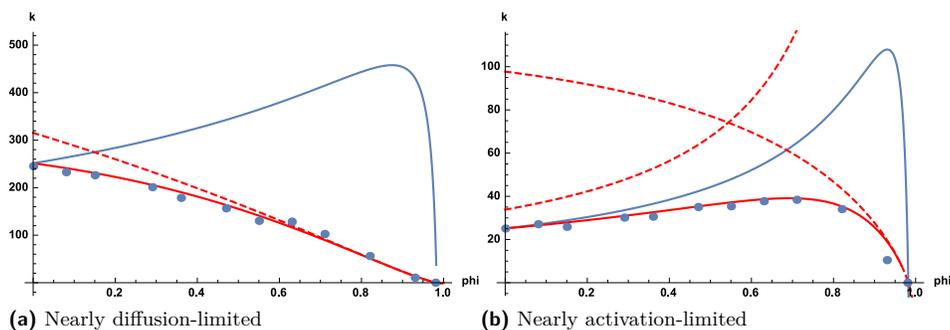
■ **Figure 3** Simulated reaction rates as functions of crowder volume occupancy. (A) Results for simulations with $k_0 = 251.3$ nm^3/μs, leading to nearly diffusion-limited reactions. (B) Results for simulations with $k_0 = 25.13$ nm^3/μs, leading to nearly activation-limited reactions at low crowder densities. In both panels, dots represent simulation data and the solid blue curves represent our hypothesis from eq. 6. The solid red curves represent our modified hypothesis from eq. 10 in which there is one fitting parameter. Dashed lines that tend downwards represent the diffusion-limited reaction rate component of our modified hypothesis, while the dashed line that tends upwards in Panel B represents the activation-limited reaction rate component of our modified hypothesis (the comparable line for Panel A is outside of the displayed plot range).

On the other hand, the solid red lines in Figure 3 show that the data agree well with a modified version of our hypothesis, given as

$$k(\phi) = \left[\frac{(1-\phi)^\gamma}{4\pi\sigma D(\phi)} + \frac{1-\phi}{k_{int}} \right]^{-1} \tag{10}$$

where γ is 1 in our hypothesis and is a fit parameter in this modified version. This modification only affects the diffusion-limited portion of the equation, which we were unable to derive rigorously. The nearly diffusion-limited reactions ($k_0 = 251.3$ nm^3/μs) fit well when γ was -0.3 and the nearly activation-limited reactions ($k_0 = 25.13$ nm^3/μs) fit well when γ was 0.27, both of which we fit by eye. The latter γ value is quite different from our hypothesis value of 1, but agrees with our intuition that the volume exclusion of crowders should accelerate reaction rates, even when reactions are strongly diffusion influenced. However, the former negative γ value is quite surprising. It shows that when reactions are diffusion-limited, the reaction rate decreases faster than the diffusion coefficient as the crowder density is increased. We do not have an explanation for this result.

Overall, we found that Smoldyn performed very well for simulating the effects of crowding on reaction rates. Simulated diffusion and reaction rate results agreed essentially perfectly with the respective input values when there were no crowders. Also, reaction rates in the activation-limited case increased in essentially perfect agreement with theory (low ϕ values in Figure 3B). In contrast, those for diffusion-limited situations differed substantially from those in our initial hypothesis. Because Smoldyn has been thoroughly tested in prior work and it agreed with the other results here, this discrepancy strongly suggests that our initial hypothesis was wrong.

eGFRD

eGFRD simulations were also performed in a 50 x 50 x 50 nm^3 cube with periodic boundaries. 100 A and B molecules were randomly positioned in the cube at the initialization. To keep
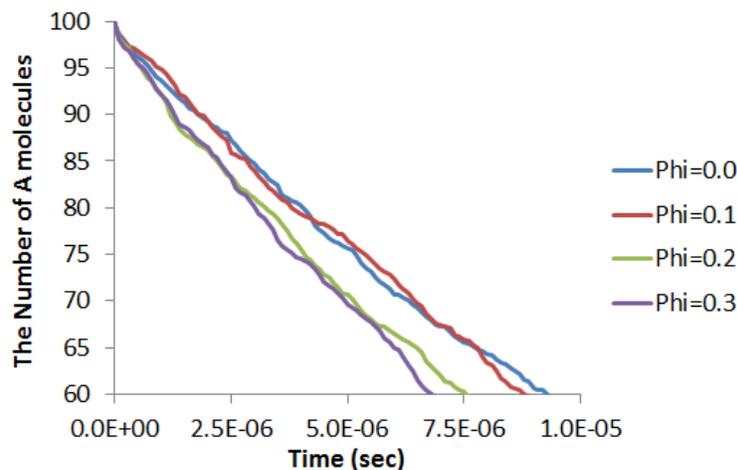
Figure 4 Simulation results. Number of A molecules surviving as a function of time for the average of 10 simulations with the volume fraction of 2 nm radii crowders, $\phi = 0, 0.1, 0.2$ and 0.3.

the total excluded volume fraction of molecules during the simulations, A and B molecules react and produce both B and non-reactive C molecules $(A + B \rightarrow C + B)$. Therefore, the concentration of B molecules was kept constant during the simulation too. A, B and C molecules were represented as 0.5 nm radii hard-body spheres. Simulations ran until a half of A and B molecules (50 molecules) reacted. With no crowders, it takes about 10 μs. All the exact time of reactions were recorded in the event-driven way. All three species diffused with diffusion coefficients of $D_0 = 10$ μm^2/s.

The reaction rate constant of A and B molecules, k_0, was $0.3382 \times k_D$ (corresponding to 85 nm^3/μs), where $k_D = 4\pi\sigma D$ in eq. 1. This kinetic rate gives nearly diffusion-limited situation. The Collins and Kimball equation, eq. 2, gives the effective reaction rate constant for the intrinsic rate k_0 as 63.52 nm^3/μs. The effective rate is four times slower than the perfectly diffusion-limited rate constant, $k_D = 251.3$ nm^3/μs. We ran each simulation 10 times. The number of A molecules were averaged every 0.1 μs for the 10 runs (figure 4).

First, we generated crowders with 2.1 nm radii (about 4 times larger than other three molecules). 320, 640 and 960 crowders were randomly placed with no overlap for the crowders volume fraction, $\phi = 0.1, 0.2$ and 0.3, respectively. All crowder molecules were fixed in place throughout the simulations.

The effective reaction rate constants in the crowded media were evaluated by numerically differentiating the time course data of the number of A molecules. (See eq. 8.) To get the steady state rate constant, we ignored the first few data $(t = 0 \ 0.3$ μs) and averaged data up to 7 μs. Normalizing with the effective rate constant in non-crowded medium expected by the Collins Kimball equation (63.52 nm^3/μs) and the constant concentration of B (about 1.3 M), we evaluated the effect of crowders on the rate constants, shown in Figure 5. With these large crowders, the effective rate constant was just affected by the excluded volume (the latter part of eq. 6), but not by the change in the diffusion rate (the former part of eq. 6). Therefore, the effective rate constants were simply given by $k(\phi) = k/(1 - \phi)$.

Next, to evaluate the condition with smaller crowders, we randomly placed 47748 molecules with 0.5 nm radii (ϕ=0.2). We ran the simulations in the same condition for other A and B molecules. However, we could not collect enough data for the analysis because the averaged
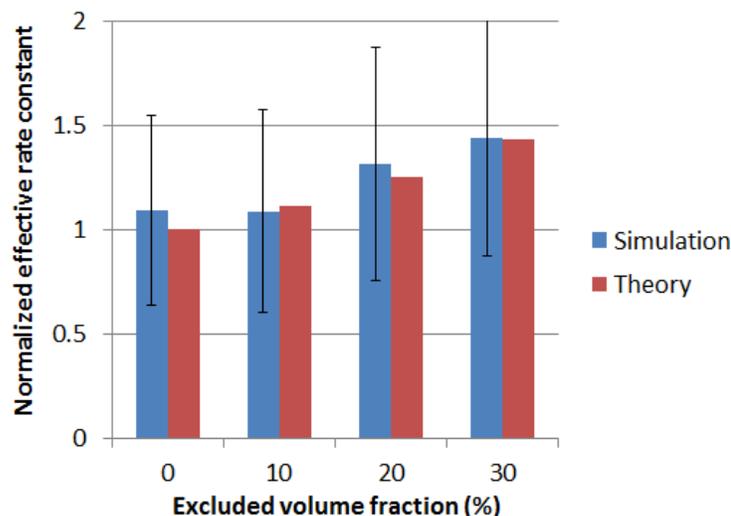
Figure 5 Effect of the excluded volume of 2 nm radii crowders on the effective steady-state rate constants ($\phi = 0$, 0.1, 0.2 and 0.3). Theoretical values were given by the equation: $1/(1 - \phi)$.

step size in the simulations was too small. The eGFRD method applies the Reaction Brownian Dynamics (RBD) method locally to each domain with more than two molecules, which is called a ÒMultiÓ domain. To guarantee the exactness and accuracy of a simulation, the step size in the ÒMultiÓ domain must be smaller than at least $10^{-5} \times \tau$, where τ is the averaged time to diffuse over the diameter of a molecule, $\sigma^2/(6D)$. In our simulations, the step size must be less than 10^7 μs. Thus, by using the larger step size, $10^{-1} \times \tau$, we could obtain simulation data for the condition. As a result, the effective rate constant in the crowded medium ($\phi = 0.2$) was about 20% of the rate in non-crowded media. With accounting for the effect on the excluded volume, $1/(1 - \phi)$, the slowed diffusion decreased the reaction rate down to 17.2%. Together with the former result, we observed the two contrary effects of molecular crowding on the effective reaction rate by using the eGFRD method. However, as mentioned above, the effective rate constant with small crowders was highly affected by the step size of simulations. To evaluate the theory in the quantitative way, we need much longer and more simulation runs.

ML-Space

ML-Space simulations were also performed in a 50 x 50 x 50 nm^3 cube, starting with 1000 molecules each of volume $\pi/6$ nm^3 (i. e. spheres of radius 0.5) for species A and B. These already occupy 0.84% of the available space ($\frac{\pi}{6} \frac{2 \cdot 1000}{50 \cdot 50 \cdot 50}$). Crowders of the same size were added in numbers such that the total volume of molecules corresponded to a desired ratio ϕ of the total space. Molecules were placed randomly in continuous space such that there was no pair of overlapping molecules. When this was not possible in a reasonable number of attempts (here generally for $\phi > 0.3$), a regular cubic grid of points with distance ≥ 1 was generated and molecules were placed consecutively with each center at a random, so far unoccupied grid point. This way, $\phi < \frac{\pi}{6} \approx 0.524$ (the density of a cubic lattice sphere packing). The simulation time steps were chosen such that the average traveled distance was 0.1, 0.2 or 0.4 nm. For a fixed diffusion coefficient ($D_0 = 10$ nm^2/μs here, too), each
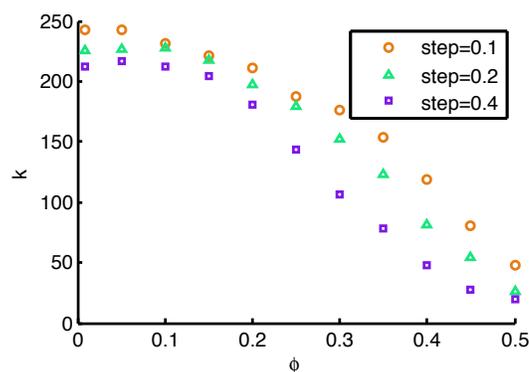
Figure 6 Simulated reaction rates as functions of crowder volume occupancy, depending on chosen step size. Results for simulations with reaction probability 1 (in case of collision), leading to a theoretical (diffusion-limited) reaction rate of $k_0 = 251.3$ $nm^3/\mu s$.

doubling of this step size increases the time step by a factor of 4. The reaction $A + B \rightarrow C$ was used, i. e. the product was not available for further reactions. The effective reaction rate was calculated from the number of reactions in an initial window of $0.32\mu s$ and averaged over 5 runs each.

Simulation results for the diffusion-limited case are in general agreement with the predictions. An initially observed increase in the effective reaction rate after the addition of the first few crowders (i. e. small increases in ϕ from its minimum) was not found to be significant. As activation-limited reactions would be incorporated in ML-Space by adjusting the probability of a reaction given an appropriate collision by a factor derived from the desired reaction rate and a calculated collision frequency, simulations of the activation-limited reaction case should only yield a scaled version of the same curve.

The results point to two main insights related to the chosen approach. First, the effective reaction rate decreases with higher crowding, but much faster than expected. We attribute this to several factors:

- When simulating spheres of the same size, the maximum possible crowding coefficient is $\phi \approx 0.74$, i. e. much smaller than 1 to begin with.
- With all molecules represented as hard spheres and with ML-Space "resolving" non-reactive collisions by retrying or partially applying the random position update, it should be harder for reactive molecules to get "past" crowders than in approaches that allow temporary partial overlap or treat some particles as points only.
- Our ad-hoc initialization using a cubic lattice may have "trapped" more potentially reactive molecules between crowders than another random initialization approach might have.

Second, we observe that a larger step size leads to a lower effective reaction rate, an effect that is especially pronounced for moderate crowding. Without crowding, the lower rate should arise from collisions not detected when molecules make large(r) jumps past each other. For moderate crowding, on the other hand, the higher step size may lead to more non-resolvable collisions and thus to a lower effective diffusion, eventually decreasing the chance of reactive molecules colliding.

These considerations indicate that while ML-Space's continuous-space simulator is in principle capable of simulating crowded environments, representing all entities by hard spheres can impede the realism of the results while at the same time the computational
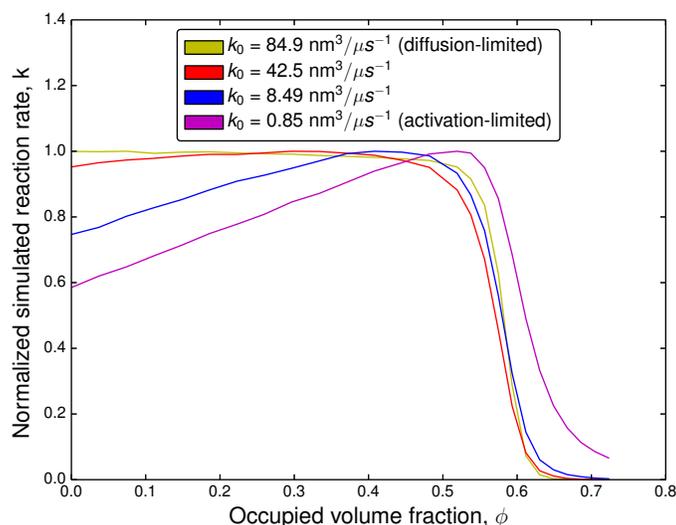
Figure 7 Normalized simulated reaction rates as a function of crowder volume occupancy with Spatiocyte.

costs rise significantly. It may be worthwhile to implement different methods for the initial placement of non-overlapping spheres in a suitably random manner and to investigate the effect of the collision resolution policy (e. g. number of retries for a move) on the effective diffusion coefficients.

Spatiocyte

In the case of Spatiocyte, the simulation model is made up of 50 x 50 x 50 nm^3 cube compartment with periodic boundaries. The radius of the lattice voxel is set to 0.5 nm. The reaction A + B → B was used to evaluate the changes in the effective reaction rate as a function of crowder volume occupancy. Initially, there were 100 A and 100 B molecules. The diffusion coefficients of A and B were set to 10 $\mu m^2/s$. Non-diffusing crowder species between 0 and 195000 molecules spread in 25 equal intervals were populated randomly in the compartment at initialization. We used four different reaction rates ($k_0 = 84.9$ $nm^3/\mu s$, $k_0 = 42.5$ $nm^3/\mu s$, $k_0 = 8.49$ $nm^3/\mu s$ and $k_0 = 0.85$ $nm^3/\mu s$) in the evaluations. Each model was run 100 times to obtain the average number of surviving A molecules. Therefore, in total, we ran 25 x 4 x 100 = 10000 simulations. We adopted the same approach employed by Smoldyn to calculate the steady-state reaction rate constant from the data. The results of our simulation are provided in Figure 7. Each curve representing the different reaction rates agrees well with our hypothesis.

Kappa

As written above, the goal of modeling the crowding effect in the core of Kappa was more about checking whether, or not, this kind of systems can be simulated efficiently. Thus, we have gone neither into the parameterisation process, nor into the back-end processing of the

■ **Table 1** Number of rules in the model written in Kappa.

Number of rules per simulation phase	
Self-assembling	18
Spawning of the particles	1
Reactions	18
Diffusion	30
Overall	67

■ **Table 2** Parameterisation of the model written in Kappa.

Size		Diffusion rates (cases/seconde)	
length	50 cases	A	1
width	50 cases	B	3
height	50 cases	AB	0.5
Number of particles (initial state)		C	1
A	1000	D	0
B	10000	Reaction rate (events/seconde)	
AB	10000	$A + B \rightarrow AB$	10000
C	0	$AB \rightarrow A + B$	2500
D	0	$AB \rightarrow B + C$	2500

results. Thus we have just performed a single simulation with arbitrary parameters, and we have reported the result of this simulation.

The model is made of 67 rules, which describes the self-assembling of the lattice of locations, the diffusion of particles and the chemical reactions. Table 1 details the number of rules for each phase of the simulation. The number of rules is quite large compared to the relative simplicity of the reaction networks. This is mainly due to the lack of supports for dealing with the symmetries of the lattice of locations. In particular, for the diffusion process, one copy of each diffusion rule had to be given for each of the 6 potential diffusion directions. The same way, 6 rules had to been given for the formation of the complex AB according to the relative position of the two reactants, and 6 rules had to be given for each of the unary reaction depending on which location the second product is spawned. The full model is available at the following url: `http://www.di.ens.fr/dagstuhl_14481/crowd_3d.ka`.

We have not computed the values of the parameters from a physical model. In particular, we have not converted continuous diffusion rates into discrete ones. The theory is well-known, but these computations require a careful handling of units and the approximation of 3D ideal Brownian motion into a discrete diffusion process within a finite lattice of locations. These conversions are available, once for all, in many formalisms (including Spatial Kappa [22] for Kappa). Thus, we have not been into these computations, but have used arbitrary parameters instead. See Table 2, for the values that we have assigned to parameters.

The result of the simulation is plotted in Fig. 8. In Fig. 8.(A), we show the survival curve of the particles of kind A, that is to say the sum between the number of instances of particles
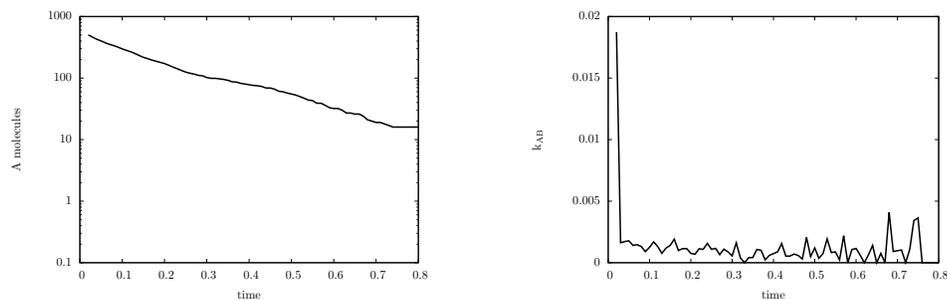
(a) Survival curve

(b) Observed k_{AB}

Figure 8 Data analysis for computing steady-state reaction rate constants. (A) Number of molecules (either free) or in a complex AB surviving as a function of time on a given simulation. (B) Reaction rate coefficient as a function of time. These data sets have been obtained with the parameters given in Table 2. Observed rates have been sampled over intervals of 0.01s and computed as the effective number of reaction applications, divided by product, for each reactant, of the middle value between the minimum number of instances and the maximum number of instances over the sampling interval.

Table 3 Benchmark for the model written in Kappa. Obtained on Dell LATITUDE E6430s, Proc intel® Core™ i7-3549M CPU @ 3.00 GHz × 4 with 8 Gio RAM, under UBUNTU 14.04 LTS.

Phase	Number of events	CPU time (seconds)	Simulation speed (events/CPU second)
Self-assembling	250148	37.61	6651
Spawning of the particles	21000	11.24	1868
Diffusion and reaction events	137176	55.02	2493
Overall	408324	103.87	3931

A and the number of instances of complexes AB. In Fig. 8.(B), we show the observed rate of association between particles A and particles B. This rate is sampled over 0.01 s intervals of time. During each sampling interval, the minimum and the maximum number of instances of particles A and of particles B are integrated, as well as the number of effective associations between particles of kind A and particles of kind B; the observed rate is then computed as the quotient of the number of associations between As and Bs by the product of the median number of instances of each reactant. This simulation has been obtained with the KASIM simulator [19] version `4.0-refactoring` with the random seed 24602700.

In Table 3, we give the computation time for the different phases of the simulation on a personal laptop Dell LATITUDE E6430s, Proc intel® Core™ i7-3549M CPU @ 3.00 GHz × 4 with 8 Gio RAM, under UBUNTU 14.04 LTS.

As a conclusion, we have, through this case study, identified three main kinds of difficulty:
1. The lack of supports to deal with symmetries. For instance, one needs 6 rules to describe the diffusion of the particles of kind A, because one has to provide one rule per potential direction. This is the same for the reactions which have to been duplicated according to the relative position of the reactants and/or where the new product is released. Thus, the lack of supports for dealing with the symmetries of the lattice space is quite cumbersome.

2. The lack of support for computing diffusion rates. Even if it is well known in theory how to convert rate constants from a continuous model of space to a discrete one. It is always quite tricky to make these computations on paper. Thus, having these computations done once for all, at the language level, is highly convenient.

3. Lastly, it is quite uneasy to describe a soup of particles in each spatial unit in the core language. This is why we have followed the microscopic lattice method. Indeed, the consequence of encoding each spatial location as an agent and the topology by the means of bond, is that the fact that a given particle is in a given location has to encoded by a bond between this particle and the agent that models this location. Then encoding soup of particles per location would require the use of complex data-structures such as hyper-links or double lists (with additional reactions to shuffle the element of these lists arbitrary). An alternative to use of bond to encode the location, is to encode the location of a particle as an internal state. Yet, internal states lack of algebraic structure, thus this alternative would require the duplication of reactions for each location, which is OK for the simulation engine, since the time complexity of an event simulation depends only logarithmically on the overall number of rules.

We notice that the two last points are handled conveniently in Spatial Kappa [22], in which all required conversions are done once for all at the language level; and in which locations are described as the internal state of a specific site for each particle and rules are macro-expanded accordingly. Yet, Spatial Kappa can only deal with regular lattices of locations such as arrays, rectangles, and rectangular boxes, with no periodic interpretation of the coordinates (but this could be implemented quite easily). Conversely, the use of bonds to model locations allows for the description of arbitrary, and even, dynamical topologies of locations.

4.1.5 Conclusions

In this work, we investigated the abilities of several simulators to model the effects of macromolecular crowders on chemical reaction rates. These simulators were an eGFRD simulator, Smoldyn, ML-space, Kappa, and Spatiocyte. Each of these treat space and molecular dynamics in subtly different ways. All of the quantitative data that were directly comparable with each other showed qualitatively similar results. In particular, diffusion-limited reaction rates decreased monotonically with the fractional crowder occupancy, while activation-limited reaction rates exhibited an initial reaction rate increase with crowder occupancy. These results also agreed qualitatively with our hypothesis.

The eGFRD simulations used the most accurate algorithm, so their results are presumably the most accurate. In practice, they agreed well with the theory for activation-limited reactions and reasonably large crowders. However, these simulations proved to be too computationally intensive for further analysis in this work.

The Smoldyn simulation method was better adapted to this investigation because it was still reasonably accurate but it ran much faster. The Smoldyn simulation parameters could be connected directly to physical parameters, which enabled us to verify that the simulated reaction rates closely matched theoretical ones for the cases where we knew the exact theory. This also enabled us to see that our initial hypothesis about the effect of crowding on reaction rates is incorrect. However, a modified hypothesis, which includes one fitting parameter, is able to fit the simulation data very well.

The ML-space results show a monotonic decrease of reaction rates with increasing crowding density. This agrees with the results that Smoldyn found for diffusion-limited reactions,

although the results were quantitatively different. These results had some puzzling aspects, such as the fact that they were time-step dependent, and that they are predicted to arise independent of whether reactions are diffusion-limited or activation-limited.

Spatiocyte was the fastest running simulator of those tested, which enabled it to generate the most result curves, each with the least noisy data. These results show a monotonic decrease of reaction rates with crowder occupancy for diffusion-limited reactions, and an initial reaction rate rise for activation-limited reactions, both of which agree with our hypotheses and with the Smoldyn simulations. Again though, the results are quantitatively different. The differences undoubtedly arise from the differences between continuous-space (Smoldyn) and lattice models (Spatiocyte).

We did not collect quantitative results with Kappa. Instead, we discovered in this investigation that Kappa can be used to successfully simulate reaction rates in crowded volumes, despite being far beyond the initial design goals for Kappa.

Two major conclusions can be drawn from these results. First, the detailed simulation algorithms can have a very large effect on the quantitative results. This includes the exact methods by which simulators treat excluded volume interactions and the use of lattice or continuous space. Second, all of these simulators could be improved upon. The results given here help illustrate the current limitations, and hence suggest areas for improvements.

References

  1 SS Andrews, T Dinh, AP Arkin, and RA Meyers. Stochastic models of biological processes. *Encyclopedia of Complexity and System Science*, 9:8730–8749, 2009.
  2 Steven S Andrews. Accurate particle-based simulation of adsorption, desorption and partial transmission. *Physical biology*, 6(4):046015, 2009.
  3 Steven S Andrews. Spatial and stochastic cellular modeling with the smoldyn simulator. In *Bacterial Molecular Networks*, pages 519–542. Springer, 2012.
  4 Steven S Andrews and Dennis Bray. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical biology*, 1(3):137, 2004.
  5 Satya Nanda Vel Arjunan and Masaru Tomita. A new multicompartmental reaction-diffusion modeling method links transient membrane attachment of e. coli mine to e-ring formation. *Systems and synthetic biology*, 4(1):35–53, 2010.
  6 Howard C Berg. *Random walks in biology*. Princeton University Press, 1993.
  7 Arne T. Bittig, Fiete Haack, Carsten Maus, and Adelinde M. Uhrmacher. Adapting rule-based model descriptions for simulating in continuous and hybrid space. In *Proceedings of the 9th International Conference on Computational Methods in Systems Biology*, CMSB '11, pages 161–170, New York, NY, USA, 2011. ACM.
  8 Arne T. Bittig, Claudia Matschegewski, J. Barbara Nebe, Susanne Stählke, and Adelinde M. Uhrmacher. Membrane related dynamics and the formation of actin in cells growing on micro-topographies: a spatial computational model. *BMC Systems Biology*, 8:106+, 2014.
  9 Frank C Collins and George E Kimball. Diffusion-controlled reaction rates. *Journal of colloid science*, 4(4):425–437, 1949.
  10 John Crank. *The Mathematics of Diffusion*. Clarendon Press, Oxford, second edition, 1975.
  11 Johan Elf and Måns Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Systems biology*, 1(2):230–236, 2004.
  12 R John Ellis. Macromolecular crowding: obvious but underappreciated. *Trends in biochemical sciences*, 26(10):597–604, 2001.

13 Jérôme Feret, Vincent Danos, Jean Krivine, Russ Harmer, and Walter Fontana. Internal coarse-graining of molecular systems. *Proceedings of the National Academy of Sciences*, 106(16):6453–6458, 2009.

14 Alice B Fulton. How crowded is the cytoplasm? *Cell*, 30(2):345–347, 1982.

15 Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

16 Damien Hall and Allen P Minton. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1649(2):127–139, 2003.

17 Jun Soo Kim and Arun Yethiraj. Effect of macromolecular crowding on reaction rates: a computational and theoretical study. *Biophysical journal*, 96(4):1333–1340, 2009.

18 Richard M Noyes. Effects of diffusion rates on chemical kinetics. *Prog. React. Kinet*, 1:129–160, 1961.

19 Jean Krivine Pierre Boutillier, Jérôme Feret. Kasim, 2010-present. available at https://github.com/Kappa-Dev/KaSim.

20 Stephen A Rice. *Diffusion-limited reactions*. Elsevier, 1985.

21 M von Smoluchowski et al. Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen. *Z. phys. Chem*, 92(129–168):9, 1917.

22 Donal Stewart. Spatial biomodelling. Master's thesis, University of Edinburgh, 2010.

23 Koichi Takahashi, Sorin Tănase-Nicola, and Pieter Rein Ten Wolde. Spatio-temporal correlations can drastically change the response of a mapk pathway. *Proceedings of the National Academy of Sciences*, 107(6):2473–2478, 2010.

24 Jeroen S van Zon and Pieter Rein Ten Wolde. Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Physical review letters*, 94(12):128103, 2005.

25 Huan-Xiang Zhou, Germán Rivas, and Allen P Minton. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annual review of biophysics*, 37:375, 2008.

26 Steven B Zimmerman and Allen P Minton. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual review of biophysics and biomolecular structure*, 22(1):27–65, 1993.

4.2 Multiscale modeling of S1P metabolism, secretion and signaling (Team 4)

Francesca Cordero, Anna Gambin, Andrzej Kierzek, Guillaume Madelaine, Joachim Niehren, Christian Rohr, and Weronika Wronowska

4.2.1 Background

Sphingolipids (SL) are a class of complex lipids with a sphingoid base (Sph). Modifications of this basic structure that consist in the addition of an amide- linked fatty acid or phosphorylation lead to the formation of bioactive sphingolipids such as ceramide (CER), ceramide-1-phosphate (C1P), sphingosine- 1-phosphate (S1P) or sphingomyelin (SM). For a long time, sphingolipids were believed to serve mainly structural purposes and have only been recognized as important messengers in cellular signaling pathways in the last two decades. A