Brownian Dynamics simulation of macromolecule diffusion in intracellular media

Author: Pablo M. Blanco Andrés.* and Advisor: Dr. Sergio Madurga Díez
Secció de Química Física and IQTCUB, Facultat de Química, Universitat de Barcelona, Barcelona, Spain.

Abstract: Cellular environment is a complex solution with a high concentration of macromolecules (this is called macromolecular crowding). Macromolecules can interact between them by means of non-specific interactions which have a large quantitative effect on their dynamic and equilibrium properties. In the present work, a Brownian Dynamics code is developed in order to study the effect of non-specific interactions, obstacle size and hydrodynamic interactions in the diffusion coefficient and the anomalous coefficient of a macromolecule diffusing in a crowded media.

I. INTRODUCTION

The study of reaction and diffusion processes in biological media has been a challenging topic in research. Although single macromolecules have low concentration, there is a high total concentration of macromolecules (such as proteins, polysaccharides...) in celular environments.

In this context, macromolecular crowding can be defined as *"macromolecular cosolutes that are nominally inert with respect to the reaction of interest"* [1]. In general, cell citosol has a occupied volume fraction of 20-30% (Fig. 1) which means an approximate macromolecule concentration of 200-300 g/L. Moreover, macromolecular crowding is also relevant outside cells *i.e.* blood plasma has a non-negligible 80 g/L protein concentration [2]. Crowding has a large quantitative effect on both equilibrium and kinetic properties of the biomacromolecules.

Since *in vitro* experiments are usually carried out at low concentrations (1-10 g/L), more information is needed in order to properly understand the kinetic and thermodynamic properties of biomacromolecules in realistic media.

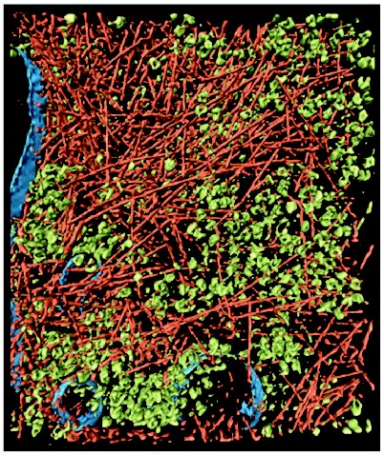

FIG. 1: An example of a crowded media. Representation of the cytoplasm of a *Dictyostelium discoideum* cell [3].

*Electronic address: pmblanco@ub.edu

Thus, different approaches have been done to get some insights on the effect of crowding. Experimentally, high concentrations of crowding agents, (usually Dextran or Ficoll macromolecules) which are considered to interact only by means of non-specific interactions (*i.e.* excluded volume effects), have been used to mimic *in vivo* environment (usually called *in vivo-like* media)[4].

Several computational studies have been performed in order to understand the effect of macromolecular crowding [5]. These studies apply different approaches as *on-lattice* Monte Carlo simulations [6] and *off-lattice* Brownian Dynamics (BD) simulations [7].

One of the main objectives of our research group is to develop a *off-lattice* BD code able to reproduce the reaction-diffusion processes taking place in *in vivo-like* environments. Therefore, this program has to describe properly macromolecule diffusion in confined media.

In dilute solution, the mean square displacement ($< r^2 >$) is lineal over time and follows the well-known Einstein-Smoluchowski equation:

$$< r^2 >= (2d)Dt. \tag{1}$$

Where D is the diffusion coefficient of the particle and d is the dimensionality of the system. However, in crowded media this equation is no longer valid since Brownian particle diffusion usually has three regimes over time (Fig. 2) [8].

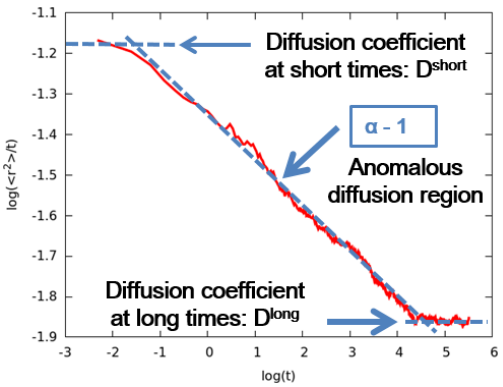

FIG. 2: Time evolution of the diffusion coefficients obtained using our BD code with a excluded volume fraction of Φ= 40%.

At short times, the particles of the system have not collided yet and therefore the diffusion coefficient remains constant over time. As the particles start to hit each other, the diffusion becomes anomalous since the ($< r^2 >$) is not lineal over time [9]:

$$< r^2 >= (2d)\Gamma t^{\alpha}. \quad (2)$$

Furthermore, the diffusion coefficient becomes not constant over time:

$$D(t) = \frac{\Gamma t^{\alpha-1}}{2d}. \quad (3)$$

Where α is the so-called anomalous coefficient and Γ is a generalized transport coefficient (also known as anomalous diffusion coefficient). Finally, at long times the system reaches equilibrium and the diffusion coefficient of the particles become constant over time again.

Another key factor in macromolecules dynamics in solution is the effect of the Hydrodynamic Interactions (HI) [10]. These interactions emerge from the fact that the motion of two particles in solution are correlated by means of solvent interactions. Since in BD the solvent is not simulated explicitly, HI need to be included in the algorithm. Two different descriptions of the HI have been used in order to do this: the Rotne-Prager-Yamakaya(RPY) diffusion tensor and the Tokuyama Model. The first one is a far-field approximation which calculates the HI between a pair of particles and the second one is a mean field approximation which calculates an effective diffusion coefficient which accounts for the HI.

In this study, a BD C++ code has been developed and properly tested. A new coarse-grained model for Dextran macromolecules is proposed in order to simulate different experimental conditions of protein diffusion studies in crowded media. Also, different procedures to include HI in the BD algorithm are discussed. Finally, the effect of non-specific interactions, obstacle size and HI in the diffusion coefficient and the anomalous coefficient is studied.

II. METHODOLOGY

The large amount of solvent molecules in a macromolecule solution makes all-atom Molecular Dynamics simulations computationally impracticable. In this context, the Langevin equation [11] of motion provides a suitable procedure since the solvent is simulated implicitly adding a stochastic force ($\mathbf{F}^{\mathrm{R}}(t)$) in the classic Newton's equation of motion which accounts for the collides of the Brownian particles with the solvent:

$$m\frac{d\mathbf{v}(t)}{dt} = \mathbf{F}_{\mathrm{T}}(t) = \mathbf{F}^{\mathrm{fric}}(t) + \mathbf{F}^{\mathrm{R}}(t). \quad (4)$$

Where $\mathbf{F}_{\mathrm{T}}(t)$ is the total instantaneous force acting in the system, $\mathbf{F}^{\mathrm{fric}}(t)$ is the force due to the dynamical friction of the particles with the solvent and m is the mass of the particles. Furthermore, in order to ensure proper statistics the stochastic force must fulfil:

$$< \mathbf{F}^{R}(t) >= 0 \qquad < \mathbf{F}^{\mathrm{R}}(t)\mathbf{F}^{\mathrm{R}}(t+\Delta t) >= 2D\Delta t. \quad (5)$$

However, for BD simulations it is more convenient to apply the over-damped limit $\left(\frac{d\mathbf{v}(t)}{dt} = 0\right)$ with isotropic diffusion [12] which, using the Fluctuation-Dissipation theorem, can be written as :

$$\frac{d\mathbf{r}(t)}{dt} = -\frac{\Delta t D}{k_B T}\nabla\mathbf{V}(\mathbf{r},t) + \sqrt{2D}\boldsymbol{\xi}(t). \quad (6)$$

This equation can be integrated at a constant time step (Δt) which gives the equation of motion of BD:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) - \frac{D\Delta t}{k_B T}\nabla\mathbf{V}(\mathbf{r},t) + \sqrt{2D\Delta t}\boldsymbol{\xi}(t). \quad (7)$$

Where $\boldsymbol{\xi}$ is a vector of $3N$ Gaussian random numbers with zero mean and unit variance and N is the number of particles of the system.

Furthermore, a coarse-grained approach has been employed because, even if all solvent atoms are removed, simulating all the atoms of a highly concentrated macromolecule solution would be too computationally expensive. Each macromolecule is modelled as a single sphere using a proper effective radius (Fig. 3).

In order to avoid overlapping, an harmonic pairwise repulsion is applied which acts when the distance between two macromolecules is smaller than their sum of radii:

$$\mathbf{V}_{ij}(\mathbf{r}_i,\mathbf{r}_j) = \begin{cases} \frac{1}{2}k(d_{ij}-R)^2 & d_{ij} < R \\ 0 & d_{ij} \geq R \end{cases} \quad (8)$$

Where R is the sum of the radii of particles i and j and k is a constant parameter set to $k = 50000$ J mol^{-1}nm^{-2}.

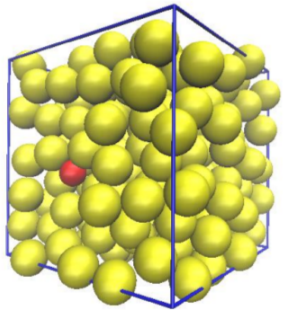

FIG. 3: Snapshot of one of dynamics performed. A protein diffuses (in red) between Dextran molecules (in yellow) which act as crowding agents.

The general procedure used to perform the BD simulations is shown in the following scheme (Fig. 4). The simulation box is a cubic one where Periodic Boundary Conditions (PBC) are applied in its frontiers. The Mean Square Displacement of the particles is only calculated after a thermalisation time of 10 ns to avoid the effect of initial high displacements due to the random initial configuration.

All the simulations performed have a length of 1000 ns and a time step of 0.1 ns. System temperature is 298.15 K. For each of the systems proposed in section IV, a minimum of 400 different BD simulations are averaged to calculate the final results.

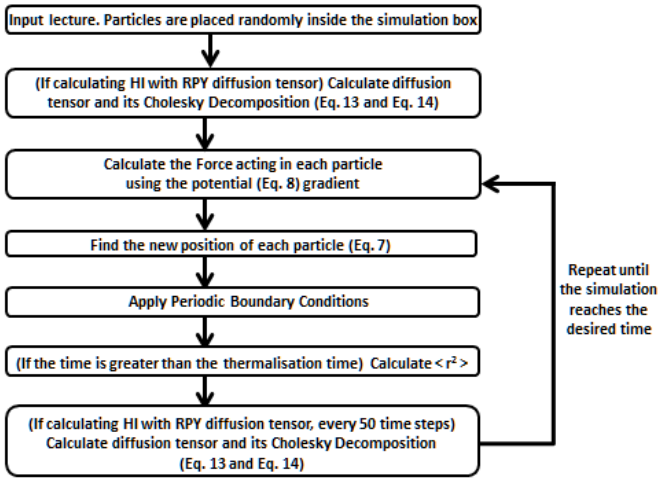

FIG. 4: Scheme illustrating the general algorithm implemented in the BD code.

A. Model

One of the main objectives of this study is to propose a model able to reproduce the experimental results of protein diffusion in crowded media. There are two main experimental techniques used in the study of macromolecule diffusion: Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Recovery After Photobleaching (FRAP).

In FCS, the fluctuations in the intensity of fluorescence in a small region of the sample are used to calculate the temporal auto-correlation function which allows to calculate the diffusion coefficient and the anomalous coefficient of the fluorescent molecule.

On the other hand, in FRAP a small volume of the sample is lighted with a laser beam. By doing so, all the molecules in that region become bleached and therefore there is no fluorescence. The diffusion coefficient and the anomalous coefficient of the fluorescent molecule can be estimated by means of the velocity of fluorescence recovery in the bleached region.

Two different experimental studies have been selected and reproduced in the present work. The first one is a previous work done by our research group [13]. It is a FRAP study of α-Chymiotrypsin protein diffusion at different concentrations of three different-sized Dextran macromolecules which act as crowding agents (Table I). This protein was chosen because it has not any known interaction with Dextran molecules and also due to its size ($R_H = 2.33$ nm), which is intermediate between the three different crowding agents.

The second, is a experimental study done by Banks *et al* [14] using FCS of Streptavidin protein diffusion at different concentrations of six distinct Dextran macromolecules. This protein is of bigger size ($R_H = 4.90$ nm) than α-Chymiotrypsin. In order to be able to compare both experimental studies only the experiments done using the Dextran D5, D50 and D400 (Table I) have been analysed. The use of Dextran as obstacles in experiments is usual since it is an inert macromolecule. Thus, it is able to perform a good mimic of a true intracellular environment in *in vitro* conditions.

As mentioned before, a coarse-grained approach is employed where each macromolecule is represented by a single sphere. In a dilute solution, the effective radius of a sphere that diffuses at the same rate as the macromolecule modelled should be its hydrodynamic radius(R_H) found experimentally. However, at high concentrations, this approximation is no longer valid since macromolecules have a flexible structure which can become more compact by means of steric compression. In the case of proteins, since they usually have a pretty rigid structure, this effect is not considered relevant.

But in Dextran macromolecules this effect is crucial because they have a very complex and branched structure which can be compacted conveniently. This means that if one calculates the occupied volume fraction of a 150-300 g/L solution of Dextran using the approximation that they are spheres of radius equal to their hydrodynamic radius this leads towards occupied volume fractions greater than 1. At this scope, a preliminary study was necessary to find a proper effective radius (R_{eff}) to model Dextran at high concentrations.

With this aim in mind, we have fitted the experimental data from [15] (hydrodynamic radius data) and [16] (radius of gyration (R_g) data) studies to a power law which relates the Dextran molecular weight with its radius:

Dextran	M_{W}(kDa)	R_{G}(nm)	R_{H}(nm)	R_{c}(nm)	R_{eff}(nm)
D5	5.2	1.7	1.9	1.1	1.2
D50	48.6	5.8	5.2	2.3	2.7
D400	409.8	17	13.5	4.7	6.1

TABLE I: Main characteristics of the Dextran molecules chosen as obstacles where M_{W} is the molecular weigh, R_{G} is the radius of gyration, R_{H} is the hydrodynamic radius, R_{c} is compact spheres radius and finally R_{eff} is the chosen effective radius.

	R_G	R_H	R_c	R_{eff}
K (nmDa$^{-\gamma}$)	0.018	0.043	0.063	0.045
γ	0.544	0.445	0.333	0.387

TABLE II: Parameters obtained fitting the experimental data of the different radius to the power law shown in Eq. (9)

$$R_x = KM_w^{\gamma}. \tag{9}$$

Where R_x can be R_G, R_H, R_c or R_{eff}. In addition, K and γ are fitting parameters. Another possible approach to calculate the excluded volume at high Dextran concentrations is to use an average Dextran specific volume (0.625 cm^3/g) [17]. This is equivalent to consider Dextran molecules as compact spheres (this is the reason why we call it compact radius (R_c)). In Fig. 5, the results of the fitting and the compact radius are shown. The effective radius we are looking for has to be in some point between the hydrodynamic and the compact radius behaviour because the radius or gyration is even a bigger overestimation than the hydrodynamic.

Since the systems we are simulating are at really high excluded volume fractions, Dextran behaviour should be closer to the compact radius one than to the one in dilute solution (hydrodynamic). Therefore, the parameters chosen for the effective radius are closer to the compact radius than to the hydrodynamic radius. The results of the fitting and the parameters chosen can be consulted in Table II. The behaviour of this effective radius as a function of the Dextran molecular weight can be also seen in Fig. 5.

Summing up, to model the tracer protein we have employed its hydrodynamic radius while to model Dextran an effective radius calculated by means of its molecular weight is used.

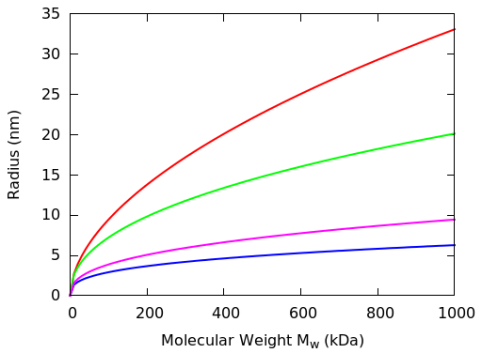

FIG. 5: Increase of the different radii studied as a function of the macromolecule molecular weight where R_G is in red, R_H is in green, R_c is in blue and R_{eff} is in purple.

B. Hydrodynamic Interactions algorithms

The inclusion of HI could be crucial in simulations where the solvent is implicitly included. These interactions originate when a Brownian particle collides with solvent particles which, in turn, also collide with other solvent molecules.

This cause particle motion to be correlated with the others even in absence of other interactions between them. Therefore, the inclusion of HI is crucial to properly model dynamic properties as particle diffusion.

One of the most appropriate procedures to include HI in BD is the algorithm proposed by Ermak and McCammon [18]:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) - \frac{\Delta t \mathbf{D}}{k_B T}\nabla \mathbf{V}(\mathbf{r},t) + \sqrt{2\Delta t}\mathbf{B}\boldsymbol{\xi}(t). \tag{10}$$

Where $\mathbf{D}$ is no longer an scalar but a diffusion tensor which accounts for the HI between all the particles of the system. $\mathbf{B}$ is a 3Nx3N matrix resulting from the factorization of $\mathbf{D}$, as required by the Fluctuation-Dissipation theorem:

$$\mathbf{D} = \mathbf{B}\cdot\mathbf{B}^{\mathbf{T}}. \tag{11}$$

One of the most popular approximations applied to calculate the diffusion tensor is the one proposed by Rotne, Prager and Yamakawa [19, 20] usually called RPY diffusion tensor. This is a far-field approximation approach involving pairs of particles, which means that long range HI are accurately calculated but short range HI are not well described.

The RPY diffusion tensor is a $3N$x$3N$ symmetric matrix containing 3x3 blocks $\mathbf{D}_{ij}$ which account for the HI between particles i and j. An illustrative example of how the RPY diffusion tensor looks like for a two-particle system is shown next:

$$\mathbf{D} = \begin{pmatrix} D_1^{SE} & 0 & 0 & D_{12,xx} & D_{12,xy} & D_{12,xz} \\ 0 & D_1^{SE} & 0 & D_{12,yx} & D_{12,yy} & D_{12,yz} \\ 0 & 0 & D_1^{SE} & D_{12,zx} & D_{12,zy} & D_{12,zz} \\ D_{21,xx} & D_{21,xy} & D_{21,xz} & D_2^{SE} & 0 & 0 \\ D_{21,yx} & D_{21,yy} & D_{21,yz} & 0 & D_2^{SE} & 0 \\ D_{21,zx} & D_{21,zy} & D_{21,zz} & 0 & 0 & D_2^{SE} \end{pmatrix} \tag{12}$$

The blocks $\mathbf{D}_{ij}$ with $i = j$ hold for the diffusion coefficient of particle i in ideal conditions (referred as $\mathbf{D}_i$). They are diagonal matrices where the diagonal elements are the Stokes-Einstein value for the diffusion coefficient:

$$\mathbf{D}_i = \frac{\mathrm{k_B}\,T}{6\pi R_i \eta}\mathbf{I}. \tag{13}$$

Where R_i is the radii of the particle i (usually its hydrodynamic radius but for Dextran in our case will be

its effective radius) and η is the dynamic viscosity of the medium. The interaction blocks $\mathbf{D}_{ij}$ with $i \neq j$ are calculated using a generalization for different-sized particles [21] of the RPY diffusion tensor:

$$\mathbf{D}_{ij} = \begin{cases} \frac{\mathrm{k_B}T}{8\pi\eta d_{ij}} \left[\left(1 + \frac{R_i^2+R_j^2}{3d_{ij}^2} \right) \mathbf{I} \right. \\ \left. + \left(1 - \frac{R_i^2+R_j^2}{d_{ij}^2} \right) \frac{\mathbf{d}_{ij} \times \mathbf{d}_{ij}}{d_{ij}^2} \right] & (R_i + R_j) < d_{ij} \\ \frac{\mathrm{k_B}T}{6\pi\eta R_i R_j} \left[\left(\frac{16d_{ij}^2(R_i+R_j) - ((R_i+R_j))^2 + 3d_{ij}^2)^2}{32d_{ij}^3} \right) \mathbf{I} \right. \\ \left. + \left(\frac{3(R_i^2+R_j^2)^2}{32d_{ij}^3} \right) \frac{\mathbf{d}_{ij} \times \mathbf{d}_{ij}}{d_{ij}^2} \right] & (R_> - R_<) < d_{ij} \leq (R_i + R_j) \\ \frac{\mathrm{k_B}T}{6\pi\eta R_>} \mathbf{I} & d_{ij} \leq (R_> - R_<) \end{cases} \tag{14}$$

Where $R_>$ and $R_<$ are the radius of the greatest and the smallest particle respectively. When $R_>$ is equal to $R_<$, the original RPY diffusion tensor is recovered.

Since $\mathbf{D}_{12}$ is always equal to $\mathbf{D}_{21}$ the RPY diffusion tensor is a symmetric matrix. Also this tensor is always a positive-definite matrix which is a crucial property because it allows to calculate the factorization of $\mathbf{D}$ (the $\mathbf{B}$ matrix mentioned in Eq. (10)) using a Cholesky decomposition. In order to include the Cholesky decomposition in our BD code the "choldc" subroutine for C++ in Numerical Recipes [22] was used.

The calculation of the diffusion tensor and its factorization is a computationally expensive procedure as it scales as N^3. Since fluctuations in HI are slow is not necessary to recalculate the diffusion tensor at every time step. Therefore, in our BD code the diffusion tensor and its factorization are only actualized every 50 time steps.

There are several approximations done in order to speed up the BD simulation using RPY diffusion tensor [23]. Mainly, these procedures are based on avoiding the Cholesky decomposition of the diffusion tensor since it is the bottleneck of the simulation. One of the most popular approaches is the one proposed by Fixman [24] where the square root of $\mathbf{D}$ is approximated using a Chebyshev polynomial expansion reducing the computational cost to $N^{2.25}$.

Recently, another promising procedure has been developed by Winter *et al* [25] based on describing the random displacements by an effective truncated *ansatz* formally equal to the contribution of the deterministic forces. This algorithm scales as N^2 without significant loss of accuracy.

However, at the present work these procedures have not been implemented and the factorization of the diffusion tensor is done using a Cholesky decomposition. Instead, we have compared the results obtained with the RPY diffusion tensor using another totally different description of the HI which is the Tokuyama model [26].

This model is a mean-field approximation for equal-sized soft core spheres to describe the self-diffusion of biomolecules in solution. Therefore, this model provides a better description of short range HI than the RPY diffusion tensor. Thus, comparison between these two methods provide some insights of the relevance of the different contributions of the HI.

Tokuyama model allows to make a prediction of the diffusion coefficient of the particle at short times (D^{short}) only as a function of the excluded volume fraction of the system (ϕ) and the diffusion coefficient at dilute solution (D_0) of the particle:

$$D^{\mathrm{short}}(\phi) = \frac{D_0}{[1 + H(\phi)]}. \tag{15}$$

$H(\phi)$ is the contribution due to the static HI valid in the short-time regime:

$$H(\phi) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)}. \tag{16}$$

Where $b = \sqrt{\frac{9}{8}\phi}$ and $c = \frac{11}{16}\phi$. This prediction of the diffusion coefficient at short time is used then in Eq. (7) to update particle position:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) - \frac{\Delta t D^{\mathrm{short}}}{k_B T} \nabla \mathbf{V}(\mathbf{r}, t) + \sqrt{2D^{\mathrm{short}} \Delta t} \boldsymbol{\xi}(t). \tag{17}$$

This procedure is starting to be widely used in BD simulations in crowded media [27, 28] because it is computationally cheaper than the RPY diffusion tensor method since it allows to introduce the HI contributions without calculating the diffusion tensor. Furthermore, the Tokuyama model allows to predict the diffusion coefficient at long times (D^{long}):

$$D^{\mathrm{long}}(\phi) = \frac{D^{\mathrm{short}}(\phi)}{\left[1 + \kappa \frac{D^{\mathrm{long}}(\phi)}{D_0} \left(\frac{\phi}{\phi_c} \right) \left(1 - \frac{\phi}{\phi_c} \right)^{-2} \right]}. \tag{18}$$

Where κ and ϕ_c are parameters that for BD simulations are set to $\kappa = 2.0$ and $\phi_c = 1.09$.

III. PREVIOUS WORK AND CODE TESTING

A. Preliminary study

The first step before starting to develop our own BD code was to check there was not any available program able to work with reaction-diffusion processes in crowded environments.

The software package ReaDDy (Reaction and Diffusion Dynamics) [29] was an interesting choice. It allows

to model reaction-diffusion processes with particle resolution using BD for dynamics and a Monte Carlo criterion for reactions. In contrast to other particle-based reaction kinetics programs, ReaDDy supports particle interaction potentials, which allows the user to modify at will. This has made ReaDDy an attractive choice for similar studies to the ones performed in our research group [30].

In order to get familiar with ReaDDy software, a preliminary study was done comparing the results obtained in a previous work of the group [31] which was done using a Monte Carlo *on lattice* algorithm (Fig. 6).

The results obtained were in qualitatively good agreement with the tendencies previously observed: Diffusion coefficient of the tracer particle decreases as the concentration of obstacles increases and it is enhanced if obstacles are mobile since particles finds easier a pathway to diffuse between them.

However, after analysing ReaDDy algorithm rigorously, we found that it does not provide the better possible description of macromolecule diffusion. This is because it applies an harmonic potential in the boundaries of the simulation box (instead of Periodic Boundary Conditions) which affects particle diffusion on the box frontiers. Also, ReaDDy does not take into account the HI in their BD algorithm which, as mentioned before, have been found to play a crucial role in macromolecule diffusion in crowded media.

Since ReaDDy source code is not available, we decided to make our own BD code including Periodic Boundary Conditions and a better description of the HI.

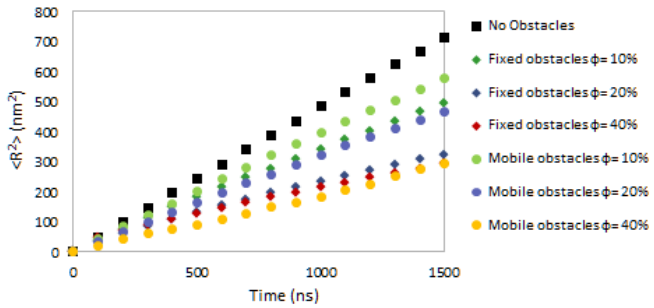

FIG. 6: Study done with ReaDDy package of the evolution of the Mean Square Displacement of a tracer particle at different excluded volume fractions (Φ) of mobile and fixed obstacles. Squares hold for diffusion without obstacles, diamonds for diffusion with fixed obstacles and circles for diffusion with mobile obstacles.

B. Code testing

In any programming work, the rigorous testing of the code implemented is of capital relevance in order to be able to achieve meaningful results.

One of the first studies done was to find the adequate working range for the time step. With this aim in mind, a study of the evolution of the $< r^2 >$ over time at different time steps was performed (Fig. 7).

This study was done at a high occupied fraction, which is the more critical condition, in order to ensure that the program does proper dynamics in all the other smaller excluded volume fractions. The time step (Δt) has to be small enough to capture the shape of the potential energy surface. If the time step chosen is too big, in a single time step the particles can overlap to much which causes the motion of the particles to become unrealistic.

In Fig. 7 is observed that simulations done with time steps from 0.01 ns to 0.1 ns show the same behaviour. On the other hand, for time steps greater than 0.5 ns, diffusion at short times is not well described since particles movement is to quick to do a proper integration of their motion. Therefore, a time step of 0.1 ns was chosen in order to achieve a compromise between the simulation accuracy and computational efficiency.

Another basic test performed was to verify that if when one particle motion (*i.e.* without any obstacle) is simulated its ideal diffusion coefficient is well described. The diffusion coefficient obtained, which is calculated using the root mean square displacement averaged of 10000 dynamics of 1000 ns each one, of one particle is statistically identical to its ideal diffusion coefficient with a relative error of a 0.5%. This means that the program is able to reproduce the Brownian motion of the particle correctly.

One of the most challenging parts of this project was to implement the RPY diffusion tensor algorithm (Eq. (13), Eq. (14) and Eq. (12)) explained in section II-B. In order to ensure that HI matrix (Eq. (12)) was properly calculated, one of the figures shown in [21] was reproduced (Fig. 8). The study used as reference was done using normalised magnitudes in terms of the mobility tensor, which can be easily related to the diffusion tensor using the Dissipation-Fluctuation theorem ($\mathbf{D} = k_B T \boldsymbol{\mu}$). The perfect agreement between our results and the reference ones validates the code developed.

The Cholesky decomposition algorithm implemented from [22] was tested by comparing the result of a diffusion

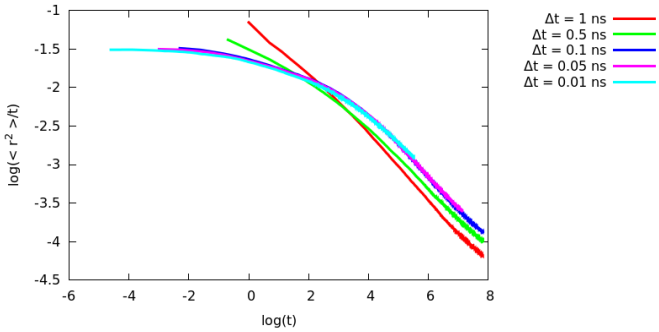

FIG. 7: Dependence of the $< r^2 >$ evolution over time as a function of the time step of the simulation. Study done in a homogeneous system with an excluded volume fraction of $\Phi = 55\%$.

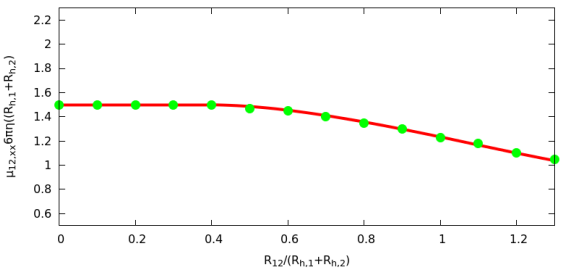

FIG. 8: Decreasing contribution of the HI between two particles as a function of their distance. Where $R_{h,1}$ is the hydrodynamic radius of one particle, $R_{h,2}$ is the hydrodynamic radius of the second particle, R_{12} is the distance between them, η is the medium kinematic viscosity and $\mu_{12,xx}$ is the first contribution of the HI between this two particles to the mobility tensor. The red line holds for the results obtained using our RPY diffusion tensor algorithm and the green dots for the reference values of [21].

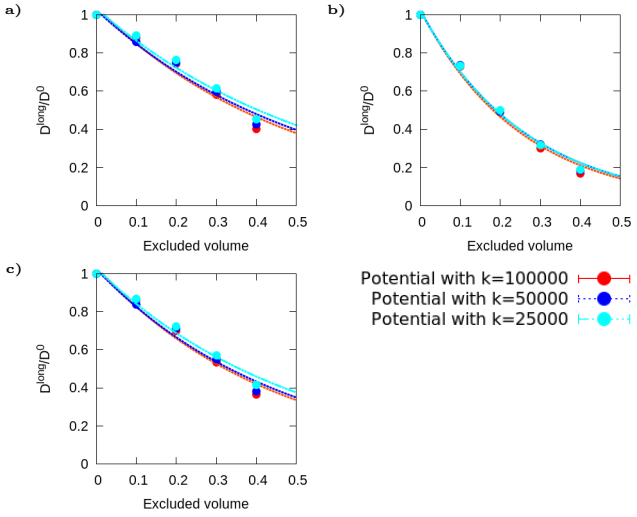

FIG. 9: Decrease of the normalised diffusion coefficient at long times (D^{long}) with the excluded volume fraction of the system at three different values of the constant k of the interaction potential. This have been done using BD without HI (a), BD with HI using Tokuyama model (b) and BD with HI using RPY diffusion tensor (c).

tensor decomposition done using our program with the result obtained using the well-tested Octave package [32]. Both codes give almost the same results, which means that the Cholesky decomposition algorithm was correctly implemented.

Finally, a study of the effect of increasing the stiffness of the interaction potential (Eq. (8)) was performed to ensure that the constant selected ($k = 50000$ J mol^{-1}nm^{-2}) was enough to prevent particles to overlap (Fig. 9). This study was done using the three different methods employed in this work at different excluded volume fractions. No significant difference was observed between the three constants proposed which means that the potential was already rigid enough to avoid overlapping. As expected, there is no effect of the method employed in this comparison as the interaction potential is equal for all of them.

IV. RESULTS AND DISCUSSION

In order to study how diffusion changes in the different systems proposed in the next subsections we have focused in two magnitudes: the diffusion coefficient at long times (D^{long}) and the anomalous coefficient (α).

As mentioned before, diffusion has three different regimes over time (Fig. 2). To compare the results obtained and see the effect of macromolecular crowding in the system, it is necessary to calculate the diffusion coefficient in the third regime (at long times) when it remains constant. The three-dimensional Einstein-Smoluchowski equation allows to calculate the D^{long} using the slope of the $< r^2 >$ vs time profile at long simulation times:

$$< r^2 >= 6D^{\text{long}}t. \tag{19}$$

In order to compare the results easily, D^{long} is always normalized using the diffusion coefficient at infinite dilution (D^0) of the particle.

The anomalous coefficient is obtained from the second diffusion regime (Fig. 2) using a linearisation of Eq. (2):

$$\log\left(\frac{< r^2 >}{t}\right) = (\alpha - 1)\log(t) + \log(6\Gamma). \tag{20}$$

Therefore, α can be obtained from the slope in the anomalous regime region. In order to avoid having to find where is the anomalous regime in every simulation done, the $< r^2 >$ vs time profile was fitted to a six order polynomial. By doing so, the slope can be computed at any point of the profile (and therefore also α) using the derivative of this function. Since the inflexion point of the $< r^2 >$ profile is the more representative point of the anomalous diffusion regime, the anomalous coefficient is calculated in that point.

A. Homogeneous system study

We have selected a simple system to compare and check the different methods. Thus, this study was done with an homogeneous system using equal sized spheres at different excluded volume fractions (Fig. 10).

This results give several insights about the methods we have implemented. First, simulations without HI and with HI using the RPY diffusion tensor description give a really similar behaviour.

This is in agreement with the results obtained by Ando *et al* [33] since in high density systems long range HI

become screened. Instead, short range HI (usually called lubrication forces) are found to be crucial to describe the diffusion coefficient reduction found experimentally. Thus, the far field approximation of the RPY is not valid in the systems we are interested.

Simulation results with HI using Tokuyama model give a stronger decay of the diffusion coefficient this means that they model better the short range HI than RPY diffusion tensor in a highly concentrated media.

Finally, is important to notice that there is an almost perfect agreement between the Tokoyama model theoretical prediction (Eq. (18)) and the results obtained in the BD simulations with HI using this model. This gives confidence in stating that the BD method using Tokuyama model has been properly implemented.

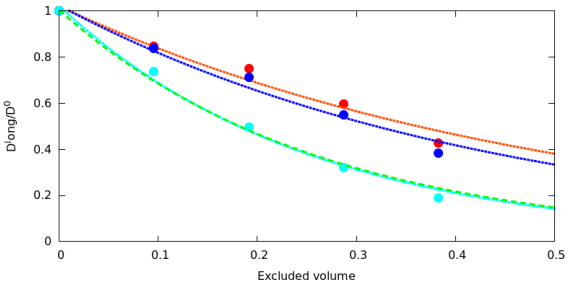

FIG. 10: Comparison of the decay of the diffusion coefficient calculated using BD without HI (red), BD with HI using RPY diffusion tensor (blue), BD with HI using Tokuyama model (cyan) and Tokuyama model theoretical prediction (green). Dashed lines are obtained by fitting the data to an exponential function.

B. Simulation of experimental systems

One of the best procedures to ensure a model is giving a proper description of reality is to compare the results obtained with the ones found experimentally.

As mentioned in section II-A, the results of two independent experimental studies have been simulated using the different methods previously discussed.

These studies include two different-sized proteins which are Streptavidin ($R_H = 4.90$ nm) and α-Chymiotrypsin ($R_H = 2.33$ nm). The diffusion of these tracer proteins are studied at different concentration of three different Dextran macromolecules (Table I) which act as obstacles.

In (Fig. 11), for Streptavidin protein diffusion, different insights of the effect of the macromolecular crowding in protein diffusion can be deduced.

First, D^{long} always decays as the concentration of obstacle increases independently of obstacle size. This was expected since particle diffusion become more hindered as the particles collide more frequently.

However, the downgrade of this decay does depend on the obstacle size. The bigger the obstacles, the more abrupt the decay of D^{long}. On the other hand, the results shown in (Fig. 12) have a different behaviour. This means that the degree of decay of D^{long} depends on the ratio between the size of the tracer particle and the obstacles $\left(\frac{R_{\text{tracer}}}{R_{\text{obstacle}}}\right)$.

The tendency shows that the decrease is higher as this ratio increases. This could be explained because for two systems with the same excluded volume fraction, if the volume is distributed among more particles, the tracer particle has more collisions with the obstacles.

Since BD simulations with HI using RPY diffusion tensor are computationally expensive and the results discussed in the previous subsection are not quite promising, these simulations were only carried out for two different Dextran sizes (D50 and D400). The results obtained using this method are indeed pretty close to the ones obtained in the description without HI and in general far from the experimental results.

Nevertheless, the results obtained in the BD simulations with HI using the Tokuyama model are in general closer to the ones observed experimentally. This means that the two main contributions to the D^{long} reduction in crowded media are excluded volume effects and short range HI.

Although the qualitative behaviour is properly modelled using the Tokuyama model description, in some cases the results obtained by simulation are still far from the experimental ones. This could be due to two different reasons.

The first one is because the Tokuyama model was deduced for equal-sized spheres and our systems are heterogeneous. Therefore, the accuracy of this method decreases as more different are the tracer protein and obsta-

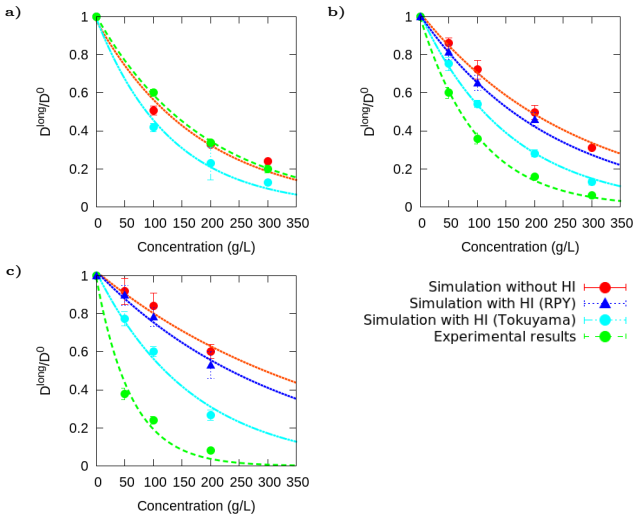

FIG. 11: Decay of the D^{long} of Streptavidin protein with Dextran concentration. Study done using three different sized Dextran as obstacles: D5(a), D50(b) and D400(c). Dashed lines are the result of fitting the data to an exponential function.

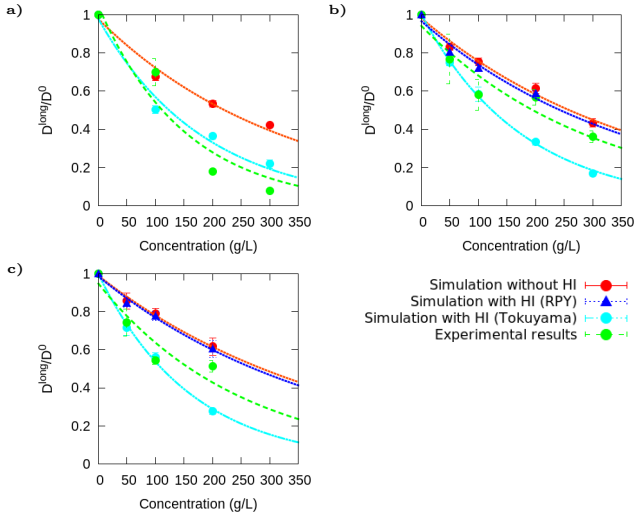

FIG. 12: α-Chymiotrypsin protein D^{long} decay with Dextran concentration. Study done using three different sized Dextran as obstacles: D5(a), D50(b) and D400(c). Dashed lines are the result of fitting the data to an exponential function.

cles particles. A generalization of the Tokuyama model for different spheres is not available yet and it would be a promising procedure to improve the results obtained. Another possibility is to improve the RPY diffusion tensor description of the many-body and near-field HI using the Durlofsky-Brady-Bossis approach [34, 35]. An interesting procedure to do this is to take advantage of the HYDROLIB [36] library for FORTRAN.

The second factor which can be responsible of the difference between the simulation results and the experimental ones is the procedure used to model Dextran.

This method does not take into account that Dextran macromolecules get more compressed as their concentration increases since it only uses an average volume for all the concentrations.

A promising possibility to enhance the Dextran description is to change the harmonic potential for a shouldered well potential similar to the one used in [37]. This potential can be tuned to have the deeper well at a distance corresponding to the Dextran hydrodynamic radius at dilute solution and the higher well at a distance corresponding to the Dextran compact radius. By doing so, Dextran volume should decrease as the concentration increases, which gives a more realistic description.

Finally, the anomalous diffusion coefficient α is studied for the same systems as the previously shown (Fig. 13) and (Fig. 14)). The α coefficient also decays with Dextran concentration as D^{long}.

In addition, the results obtained using the different methods are in general similar. In particular, the α calculated using the RPY diffusion tensor are usually higher than the ones obtained with the other methods.

The results obtained (except (Fig. 13 a)) are in general close to the experimental ones. In this context, the role of HI does not seem to be as crucial as in the D^{long} description. However, future studies with a better description of the short range HI should be carried out (*e.g.* using the Durlofsky-Brady-Bossis approach or the Tokuyama model generalization mentioned before) to ensure if them really have or not an effect on α calculation.

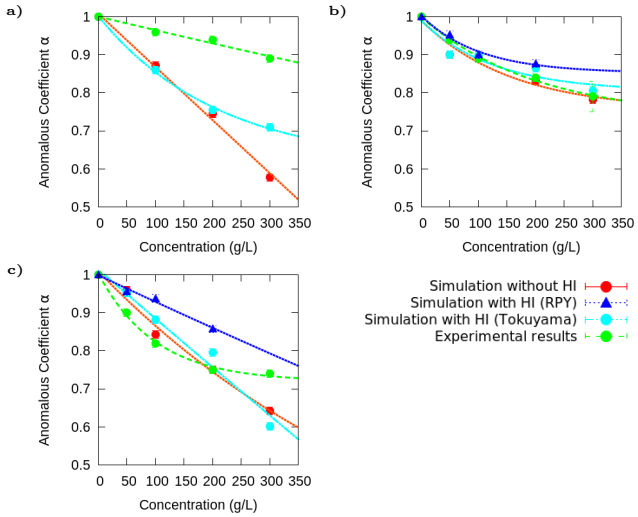

FIG. 13: Anomalous coefficient α of Streptavidin decay with Dextran concentration. Study done using three different sized Dextran as obstacles: D5(a), D50(b) and D400(c). Dashed lines are the result of fitting the data to an exponential function with an asymptotic decay.

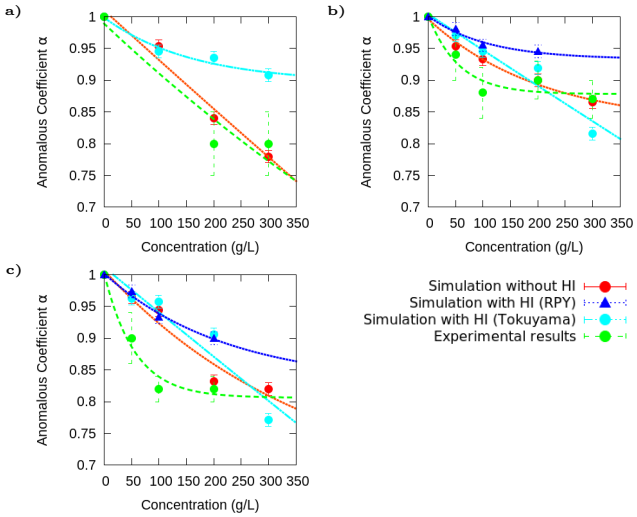

FIG. 14: α-Chymiotrypsin protein anomalous coefficient α decay with Dextran concentration. Study done using three different sized Dextran as obstacles: D5(a), D50(b) and D400(c). Dashed lines are the result of fitting the data to an exponential function with an asymptotic decay.

V. CONCLUSIONS

The BD code developed is able to reproduce qualitatively the experimental behaviour of diffusion in *in vivo-like* systems using the Tokuyama model description of HI.

Within these HI, the short range HI have been found to play a crucial role in the D^{long} description. However, since no relevant contribution of the HI methods in α calculation have been noticed, their importance in the anomalous diffusion regime remains unclear. Thus, a better description of short range HI is needed to ensure if HI really have or not an effect on α calculation.

Although the results obtained are in general close to the experimental data, the program can still be improved to achieve better quantitative results. Several new procedures have been proposed in order to reach this aim. The first one is to include a new model for Dextran macromolecules using a shouldered well potential. The second one is to improve the short range HI description by means of a Tokuyama Model generalization or including the Durlofsky-Brady-Bossis approach in the RPY diffusion tensor procedure.

Finally, the results obtained give some insights into the effect of the particles size in the diffusion coefficient reduction. In particular, the decay of the D^{long} has been found to be bigger as the ratio between the size of the tracer particle and the obstacles increases.

Acknowledgments

There are few projects that can be carried out only by a single person, and this work is not an exception. I would like to thank Dr. Sergio Madurga and Dr. Francesc Mas for their dedication, kindness and guide not only during this work elaboration but all the time we have been working together. Also I would like to thank to the other members of the BioPhysChem research group (especially to Nuria Vilaplana, Mireia Via and Enric Fortin) for making our study a warm place with their affection and jokes.

Finally, but not less important, I would like to thank to my family, Clara Baucells and my close friends for encouraging me in the darkest moments of this year and for their constant trust on me. This work would have never been possible without all of you.

[1] Zhou, H.X.; Rivas, G.; Minton, A.P. "Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences". Annu Rev Biophys **37**: 375 (2008).

[2] Ellis, R.J. "Macromolecular crowding: obvious but underappreciated". TRENDS in BioChem.Sci. **26**: 597 (2001).

[3] Zimmerman, S.B. "Macromolecular crowding effects on macromolecular interactions: some implications for genome structure and function". Biochim.Biophys. *Acta* **1216**: 175 (1993).

[4] Pastor, I.; Pitulice, L.; Balcells, C.; Vilaseca, E.; Madurga, S.; Isvoran, A.;Mas, F. "Effect of crowding by Dextrans in enzymatic reactions". Biophys.Chem. **185**: 8 (2014).

[5] Elcock, A.H. "Models of macromolecular crowding effects and the need of quantitative comparisons with experiment". **20**: 196 (2010).

[6] Pitulice, L.; Vilaseca, E.; Pastor, I.; Madurga, S.; Garcés, J.L.; Isovran, A.; Mas, F. "Monte Carlo simulations of enzymatic reactions in crowded media. Effect of the enzyme-obstacle relative size". Mathematical Biosciences **251**: 72 (2014).

[7] Hasnain, S.; McClendon, C.L.; Hsu, M.T.; Jacobson, M.P.; Bandyopadhyay, P. "A New Coarse-Grained Model for E. coli Cytoplasm: Accurate Calculation of the Diffusion Coefficient of Proteins and Observation of Anomalous Diffusion". PLoS ONE **9**: 1 (2014).

[8] Saxton, M.J. "Anomalous Diffusion Due to Obstacles: A Monte Carlo Study". Biophys.J. **66**: 394 (1994).

[9] Bauchaud, J.P. and Georges, A. "Anomalous diffusion in disordered media: Statistical mechanics, model and physical application". Phys.Rep. **185**: 127 (1990).

[10] Elcock, A. H. "Molecule-Centered Method for Accelerating the Calculation of Hydrodynamic Interactions in Brownian Dynamics Simulations Containing Many Flexible Biomolecules". J.Chem.Theory Comput. **9**: 3224 (2013).

[11] Gillespie, D.T. "The multivariate Langevin and Fokker-Planck equations" Am. J. Phys., **64**: 1246 (1996).

[12] Zwanzig, R. "Transport, Collective Motion and Brownian Motion" J.Stats.Phys. **9**: 215 (1973).

[13] Pastor, I.; Vilaseca, E.; Madurga, S.; Garces, J. L.; Cascante, M.; Mas, F. "Diffusion of α-Chymiotrypsin in Solution-Crowded Media. A Flourescence Recovery after Photobleaching Study". J.Phys.Chem. **114**: 4028 (2010; *ibid* **114**: 12152 (2010).

[14] Banks, D.S. and Fradin, C. "Anomalous Diffusion of Proteins Due to Molecular Crowding". Biophys.J. **89**: 2960 (2005).

[15] Armstrong, J.K.; Wenby, B.; H.J. Meiselman and Fisher, T.C. "The Hydrodynamic Radii of Macromolecules and Their Effect on Red Blood Cell Aggregation". Biophysical Journal **87**: 4259 (2004).

[16] Fundueanu, G. ; Nastruzzi, C.; Carpov, A.; Desbrieres, J. and Rinaudo, M. Biomaterials "Physico-chemical characterization of Ca-alginate microparticles produced with different methods". **20**: 1427 (1999).

[17] Schaefer, D.W. "A unified model for the structure of polymers in semidilute solution." Polym. **25**: 387 (1984).

[18] Ermak, D.L. and McCammon, J.A. "Brownian dynamics with hydrodynamic interactions". J.Chem. Phys. **69**: 1352 (1978).

[19] Rotne, J. and Prager, S. "Variational treatment of hy-

drodynamic interaction in polymers". J.Chem. Phys. **50**: 4831 (1969).

[20] Yamakawa, H. J. "Transport properties of polymer chains in dilute solution: hydrodynamic interaction". Chem. Phys. **53**: 436 (1970).

[21] Zuk, P.J.; Wajnryb, E.; Mizerski, K. A. and Szymczak, P. "Rotne-Prager-Yamakawa approximation for different-sized particles in application to macromolecular bead model". J.Fluid. Mech. **741**: R5 1 (2014).

[22] Press, W. H.; Teukolsky, S.A.; Vetterling, W.T. and Flannery, B.P. "Numerical Recipes in C The Art of Scientific Computing". 2nd ed. (Cambridge University Press, 1988)

[23] Rodríguez, R.; Hernández, J.G.; García, J. "Comparison of Brownian dynamics algorithms with hydrodynamic interactions". J.Chem.Phys. **135**: 084116 (2011).

[24] Fixman, M. "Implicit algorithm for Brownian Dynamics of Polymers". Macromolecules **19**: 1195 (1986).

[25] Geyer, T. and Winter, U. "An O(N^2) approximation for hydrodynamic interactions in Brownian dynamics simulations". J.Chem.Phys **130**: 114905 (2009).

[26] Tokuyama, M.; Moriki, T.; Kimura, Y. "Self-diffusion of biomolecules in solution". Physical Review E **83**: 081402-1 (2011).

[27] Sun, J. and Weinstein, H. "Toward realistic modeling of dynamic processes in cell signaling: Quantification of macromolecular crowding effects" J.Chem.Phys. **127**: 155105 (2007).

[28] Mereghetti, P. and Wade, R.C. "Atomic Detail Brownian Dynamics Simulations of Concentrated Protein Solutions with a Mean Field Treatment of Hydrodynamic Interactions". J.Chem.Phys. **116**: 8523 (2012).

[29] Schneberg, J. and Noé, F. "ReaDDy - A Software for Particle-Based Reaction-Diffusion Dynamics in Crowded Cellular Environments". PLoS ONE **8**: 9 (2013).

[30] Gomez, D. and Klumpp, S. "Biochemical reactions in crowded environments: revisiting the effects of volume exclusion with simulations". Front.Phys. **3**: 45 (2015).

[31] Vilaseca, E.; Isvoran,A.; Madurga, S.; Pastor, I.; Garceés, J.L.; Mas, F. "New insights into diffusion in 3D crowded media by Monte Carlo simulations: effect of size, mobility and spatial distribution of obstacles". Phys.Chem.Chem.Phys. **13**: 7396 (2011).

[32] Eaton, J.W.; Bateman, D. and Hauberg, S. "GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations". CreateSpace Independent Publishing Platform. ISBN 1441413006 (2009).

[33] Ando, T.; Chow, E.; Skolnick, J. "Dynamic simulation of concentrated macromolecular solutions with screened long-range hydrodynamic interactions: Algorithm and limitations". J.Chem.Phys. **139**: 121922 (2013).

[34] Durlofsky, L.; Brady, J.F. and Bossis, G. "Dynamic simulation of hydrodynamically interacting particles" J. Fluid Mech. **180**: 21 (1987).

[35] Brady, J.F. and Bossis, G. "Stokesian Dynamics". Annu. Rev. Fluid Mech. **20**: 111 (1998).

[36] Hinsen, K. "HYDROLIB: a library for the evaluation of hydrodynamic interactions in colloidal suspensions" Comp. Phys. Comm. **88**: 327 (1995).

[37] Vilaseca, P. and Franzese, G. "Softness dependence of the anomalies for the continuous shouldered well potential" J.Chem.Phys. **133**: 084507 (2010).