Tutor/s

Dr. Sergio Madurga
Dr. Francesc Mas Pujadas
Departament de Ciència de Materials i Química Física, Universitat de Barcelona

EUROPEAN MASTER IN THEORETICAL CHEMISTRY AND COMPUTATIONAL MODELLING

Master's Thesis

Development of a Brownian Dynamics simulation code for enzymatic reaction-diffusion processes in crowded intracellular environments.

Desenvolupament d'un codi de Dinàmica Browniana per a processos de reacció-difusió en medis intracel·lulars obstruits.

Mireia Via Nadal

June 2016

UNIVERSITAT DE BARCELONA

ACKNOWLEDGMENTS

M'agradaria expressar els meus agraïments, usant la meva llengua,

Al grup de BioPhysicalChemistry, per haver-me introduït en el món de la recerca i haver-me ajudat a créixer com a investigadora científica, i sobretot per l'amabilitat de tots els components del grup durant els tres anys que hem treballat junts;

A mis compañeros de máster TCCM, tanto nacionales como internacionales, pues la experiencia vivida durante estos dos años ha sido única y los intensivos compartidos son unos recuerdos que jamás olvidaré;

A la meva família, que des de ben petita m'han motivat i recolzat les meves pròpies decisions;

Als meus amics, la meva segona família, que m'han aportat consells i suport quan ho he necessitat;

Al Sergi, per animar-me en els moments de desesperació i estar-hi sempre;

Però sobretot a vosaltres, Alba, Toni, Efrem, Alex i Martín, perquè sense vosaltres aquest màster no hagués sigut el mateix.

Dedicat a tu, papa.

CONTENTS

RESUM

Tant l'interior d'una cèl·lula com la membrana que l'envolta són medis altament estructurats i amb una elevada concentració de macromolècules que, a través d'interaccions no específiques, afecten notablement els processos de difusió i reacció que hi tenen lloc. Això fa que els models teòrics que descriuen correctament aquests processos en medis homogenis deixin de ser vàlids en aquestes condicions. Per tant, és necessari conèixer bé el mecanisme d'aquests processos en medis altament obstruïts per tal de poder interpretar els resultats experimentals i proposar així nous models.

Aquests medis poden ser estudiats tant en experiments *in vivo-like* com en experiments *in-silico*. En aquest projecte, es desenvolupa un codi *off-lattice* de reacció-difusió per Dinàmica Browniana en tres dimensions en el llenguatge C++, el qual anomenem RK3D (*R*eaction *K*inetics in *3-D*imensions). D'aquesta manera, es pretén estudiar l'efecte del volum exclòs per qualsevol sistema enzimàtic que segueixi el mecanisme de Michaelis-Menten. El moviment Brownià de les partícules és descrit mitjançant l'equació de Langevin, mentre que les reaccions es consideren processos de Poisson (criteri Monte Carlo) on les probabilitats de reacció, directament relacionades amb les constants cinètiques de reacció, es descriuen com a la probabilitat de trobar, al menys, un esdeveniment reactiu en un pas iteratiu de la simulació. Els diferents algoritmes i subrutines implementats al codi s'exposen en aquesta memòria, juntament amb el seu fonament teòric.

La precisió i l'exactitud del codi RK3D s'avaluen mitjançant estudis comparatius de difusió i de reacció. El sistema model escollit és el de la proteïna α-quimotripsina, un sistema enzimàtic que el nostre grup n'ha estudiat prèviament la seva difusió i reactivitat tant experimentalment com computacionalment (*on-lattice*).

SUMMARY

Both the inner part of a cell and its membrane are very dense media, with huge concentrations of biomolecules that, by means of non-specific interactions, produce a considerable effect in physico-chemical processes like diffusion and reactivity. Therefore, theoretical models that describe these processes in homogeneous media are no longer valid in crowded conditions. Due to that, diffusion and reactivity ought to be studied in these conditions in order to obtain more reliable and realistic results.

These media can be studied by both *in vivo-like* and *in-silico* experiments. In this project, an off-lattice, 3-dymensional reaction-diffusion Brownian Dynamics simulation code written in C++ language, which we named RK3D (*R*eaction *K*inetics in *3-D*imensions), is developed in order to study the effect of the excluded volume for any enzymatic system that follows the Michaelis-Menten mechanism. The particles' Brownian motion is described by the Langevin equation, whereas reactions are considered as Poisson processes (Monte Carlo criterion) and the reaction probabilities, directly related to the reaction constants, are described as the probability to find, at least, a reactive event in a given time step. The different algorithms and subroutines implemented in the code, along with their corresponding theoretical framework, are explained in this project.

The accuracy and precision of the RK3D code are evaluated by means of reaction-diffusion comparative studies using different initial conditions of enzyme-substrate concentration and excluded volume. The system chosen for such experiments is α-chymotrypsin, an enzymatic system that our group has previously studied its diffusion and reactivity both experimentally and computationally (on-lattice*).*

1. INTRODUCTION

1.1 THE CELLULAR ENVIRONMENT: MACROMOLECULAR CROWDING

Biomacromolecules work within intracellular environments that are highly occupied by other macromolecules, that is, environments in which macromolecules occupy a considerable fraction of the total volume of the cell's cytosol. A great amount of species are present in the cellular lumen (from 5% to 40% of the total volume in the cell), but in small concentrations [1, 2], since no single species of macromolecule is necessarily present at high concentrations. That is, the total concentration of different macromolecules is extremely big, but the concentration of a single species is rather small. Thereby, it is more convenient to say that the intracellular medium is rather *crowded* or *volume-occupied* by biopolymers rather than *concentrated* [3-5].

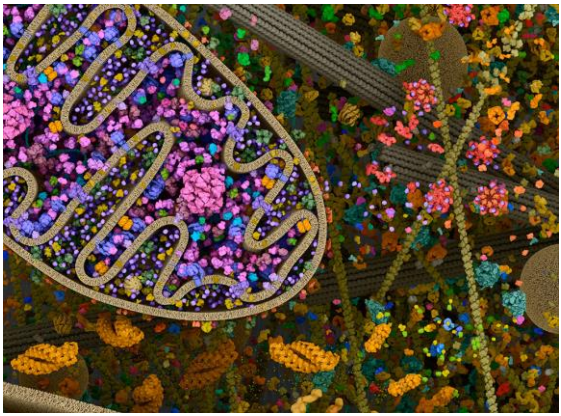

Figure 1. Illustration of a cross-section through a dilute cytoplasm. (Image by David S. Goodsell)

Such cellular species represent an obstacle for the macromolecules that need to move throughout the cytoplasm. These obstacles, or *crowders*, hinder the diffusion of the solutes in the solution and reduce the available solvent volume, which increases the effective concentration of all the species present in the system. This hindered diffusion can influence the protein assembly, transport and other equilibrium and dynamical properties that, due to macromolecular crowding, may be very different from their ideal solution counterparts.

Thereby, macromolecular crowding is a phenomenon that is based on some alterations of the physicochemical properties of macromolecules due to the existence of high concentrations of unrelated macromolecules in the solution. It connotes the non-specific influence of steric exclusion repulsions on specific reactions and processes that occur in highly volume-occupied media. This non-specific steric repulsion is intrinsic to the cellular cytosol and cannot be avoided. Since the effect of crowding on thermodynamic activity is exerted by macromolecules on other macromolecules, such effect is called *macromolecular crowding* [5].

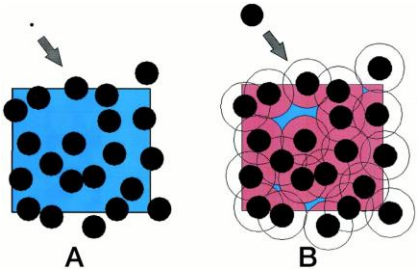

Figure 2. The importance of size in volume exclusion. If a particle of infinitesimal size is introduced in a solution of macromolecules, it is free to occupy the entire volume in between the rest of macromolecules (all the blue space depicted in A). If a particle of a comparable size to the rest of the macromolecules is introduced, the available volume to dissolve becomes incredibly reduced (B). Blue zones represent accessible space; pink zones represent the excluded volume. (Image taken from *Minton et al*, ref. 6)

Macromolecular crowding is more accurately termed *the excluded volume effect* due to the fact that solutes have mutual impenetrability [6]. The size of a given solute is determining in the contribution of such particle in the excluded volume effect. Excluded volume ϕ is the fraction of volume occupied by solutes with respect to the total volume. A more representative explanation can be extracted by looking at figure 2: The volume available to a molecule that has to be solvated is defined as the fraction that can be occupied by the centre of that solute. If the introduced molecule is small relative to the macromolecules already dissolved in the system, the volume that is accessible to it is almost all the volume that is not occupied (the rest of the volume not occupied by other particles). However, if the introduced molecule is large (comparable to that of the other solutes), like an enzyme, the available volume is considerably smaller because the centre of that molecule can approach the centre of the other macromolecules to no less than the distance at which the surfaces of the two molecules meet. The mutual impenetrability of the species causes an excluded volume effect in the reaction media. Reactants are spatially constrained on the microscopic level by steric repulsions and

attractive interactions which occur between molecules. The contribution of steric repulsion to reduced entropy and increased free energy is greater in the last case [4, 6, 7].

Because of the excluded volume effect, *in vitro* experiments may not reproduce the actual cellular conditions at all (50-400 mg mL^{-1} of macromolecules), since these experiments are performed under quasi-ideal conditions (1-10 mg mL^{-1}) [5, 8], where non-specific interactions with other particles in the medium are not accounted for. Big part of the volume is physically occupied and as a consequence the thermodynamics and kinetics of the processes taking place in the intracellular medium become affected – this includes diffusive and reactivity properties.

It is widely known that macromolecular crowding affects protein assembly [5, 7] and folding [9-11], structural organization of DNA [12], reversibility of biochemical reactions and large quantitative effects on both rates and equilibria involving macromolecules [1, 2, 13-20], which is the main focus of this work. The rate of any biochemical process that is diffusion-limited will be affected by crowding, since diffusion directly relates to the probability of encounters between two molecules in a finite space. If the reaction is limited by the encounter of the substrates (in bimolecular or higher order reactions), crowding may cause two opposite effects: by one side, it may increase reactivity due to confinement. Usually, when both reactants are smaller than the

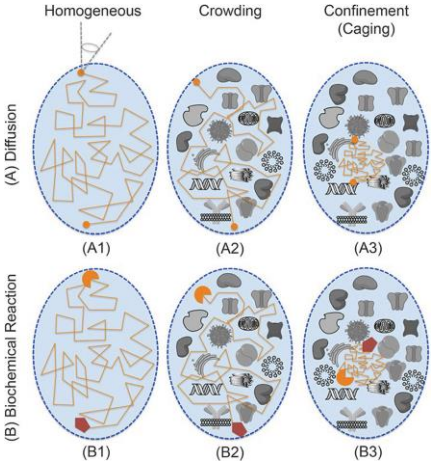

Figure 3. Illustration of macromolecular crowding effects on (A) diffusion and (B) biochemical reactions. In A), diffusion processes are sensitive to association and confinement in crowded environments. In B), the phenomenological reaction rate between two biomolecules depends on the corresponding diffusion coefficient of each reactant, crowding, and confinement in cages made of macromolecules. Due to confinement, reactants can find each other easily; on the contrary, if reactants find themselves in distant environments, it will require more diffusion time to encounter each other. (Image taken from *Ahmed A. Heikal, SPIE Newsroom,* ***2015***)

crowding agents themselves, the presence of large obstacles implies an overall division of the system in small volumes surrounded by these obstacles that, as a consequence, produce a smaller reaction media. These small spaces, that are usually called *confined volumes*, generate a confinement of the particles in the system that enhances the encounters between particles [19]. This effect is also known as cage effect, since the reactive particles are trapped by the crowders (see figure 3).

If the reaction is limited by the activity of the complex (reaction-limited), crowding tends to increase the rate, as it has been seen that crowding increases activity [8, 21, 22]. Thereby, due to macromolecular crowding, it is compulsory to study biochemical processes in nature-like (or *in vivo-like*) environments in order to approximately recreate the effect of these average high concentration of macromolecules in the intracellular medium. In that way, the reaction rate constants obtained in such crowded environments may provide results that would imply an update to kinetic databases, in which almost all the data gathered until now has been obtained in *in vitro* conditions, that is, quasi-ideal (dilute) conditions, where it is assumed that the reaction medium is homogeneous and well-stirred.

Figure 4. Visual representation of macromolecular crowding between *in vitro* systems (left) and *in vivo* situations (right), where the concentration of other species besides of the protein to study (yellow) cannot be underestimated. (Image taken from *Ellis & Hartl*, ref. 9)

1.2 Brownian Motion and Diffusion

The motion of chemical species in an enzymatic system is usually described by means of Brownian motion. This motion is the random movement of a *tracer*, i.e. a small particle with diameters compressed between 1nm to 1μm suspended or immersed in a fluid. In 1827, Robert Brown systematically investigated the vigorous, irregular and never-stopping (but with stationary velocity) motion of pollen grains and other inorganic species, and so this type of motion was named after him [23]. Since he showed that this motion was present in any suspension of fine particles, such as glass, coal, etc., he demonstrated that the origin of this motion was not due to

life. In 1905, Einstein provided a theoretical analysis of Brownian motion by describing it in terms of diffusion processes [24]. He proved (and simultaneously but independently by Smoluchowski in 1906 [25]) that the origin of this motion was due to the exceedingly frequent impacts on any small particles of the incessantly moving molecules of solvent in which they are suspended, but also due to thermal fluctuations.

To describe this motion, Einstein worked on the hypothesis that the motion of these molecules can only be described probabilistically in terms of exceedingly frequent statistically independent impacts. Each individual particle has a motion that is independent of the motions of all other particles, and the movements of one and the same particle in different time intervals are independent processes, as long as such time intervals are not chosen too small.

In Einstein's 1905 paper [24] it was proven that the diffusion coefficient D, a parameter that characterizes the migration of a given kind of particle at a given temperature in a given fluid, is related to its mobility (and to the friction it experiences) according to the Einstein relation, and verifies that Brownian motion is related to the thermal motion of molecules

$$D = \mu k_{\mathrm{B}} T = k_{\mathrm{B}} T / f \tag{1}$$

where μ is the mobility and f is the friction coefficient of a given particle in a given fluid. The friction coefficient resulting from a translational motion for a spherical particle in a Newtonian liquid is described by the Stokes friction factor, $f = 6\pi\eta r$, where r is the radius and η is the liquid's viscosity. With this consideration, the definition of the diffusion coefficient turns into the Stokes-Einstein equation (S-E equation):

$$D = \frac{k_{\mathrm{B}} T}{6\pi\eta \mathrm{r}} \tag{2}$$

Some years later, in 1916, J. Perrin [26, 27] was able to experimentally measure the Avogadro constant by measuring the mean squared displacement (MSD) of colloids by following their trajectory while suspended in water (see figure 5). The relation used to measure the MSD was also provided by Einstein in 1905, and is known as the Einstein-Smoluchowski relation (E-S relation):

$$\langle r^2(t) \rangle = \langle r^2(t) - r^2(t_0) \rangle = (2\mathrm{d}) D \mathrm{t} \tag{3}$$

where d represents the dimensionality of the system. This measure allows studying the spreading of the particles with respect to their initial position. According to equation 3, the spreading of the particle increases linearly with time.

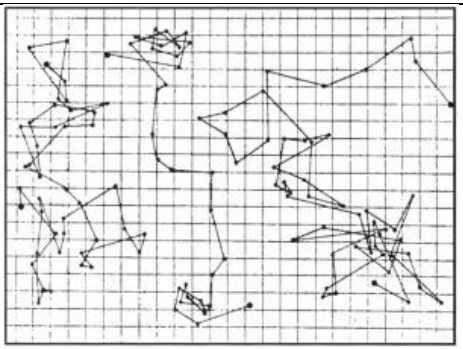

Figure 5. Brownian motion observed by Perrin for mastic spheres (radius 0.53 μm) in water, where particle positions were marked every 30 seconds. The side of a square in A is about 3 μm. (Image taken from *J. Perrin*, ref. 29)

1.2.1. Anomalous diffusion

In a crowded environment, diffusion becomes hindered by non-reacting species and also by species involved in the reaction (auto-crowding phenomenon). The E-S relation is expressed in general terms as [15, 28, 29]:

$$\langle r^2(t)\rangle = (2\mathrm{d})\Gamma t^{\alpha} = (2\mathrm{d})D_{\phi}(t)t \qquad (4)$$

where α is the anomalous diffusion coefficient or exponent ($0 < \alpha < 1$) and the diffusion coefficient is expressed as a generalised transport parameter denoted as Γ, and is related to a time-dependent diffusion function $D_{\phi}(t)$ as:

$$D_{\phi}(t) = \frac{1}{(2\mathrm{d})}\frac{<r^2>}{t} = \Gamma t^{\alpha-1} \qquad (5)$$

When the anomalous diffusion exponent is equal to 1, diffusion is constant at all times; when it is lower than 1, the system behaves with a subdiffusive trend (usual in a crowded environment), and superdiffusion otherwise. The regime where particles move in a subdiffusive motion is known as the anomalous diffusion region.

To study anomalous diffusion it is more appropriate to analyse the MSD in a logarithmic representation (log-log plot of diffusion over time [15, 29]:

$$\log\left(\frac{<r^2>}{t}\right) = \log\left[(2d)D_{\phi}(t)\right] = (\alpha - 1)\log(t) + \log[(2\mathrm{d})\Gamma] \qquad (6)$$

In this way, we can differentiate a succession of three differentiated regions (see figure 6). The initial diffusion times (region A) are non-time dependent, where particles diffuse as if they

were in homogeneous, dilute solution. The particles have still not recognized the dense environment and, thus, diffusion is not yet affected by macromolecular crowding. The diffusion coefficient in a homogeneous dilute-like solution is usually labelled as D_0 or D_{short}(for short diffusion times). The following region (region B) is known as anomalous diffusion region, and shows how the particles explore the whole volume of the solution and start to notice how crowded the environment is. After exploring the whole volume (large diffusion times, region C), the diffusion coefficient reaches a constant value once again, which reflects diffusion in a homogeneous dense medium. The diffusion coefficient is obviously lower than the corresponding to the initial diffusion coefficient D_0. This diffusion coefficient is named as limiting diffusion coefficient D_∞, also labelled as D_{long} [15, 30]. The shifting from the normal diffusion regime to the anomalous one is characterized by the crossover time τ, and it defines the time after which the diffusion completely "feels" the crowding effects.

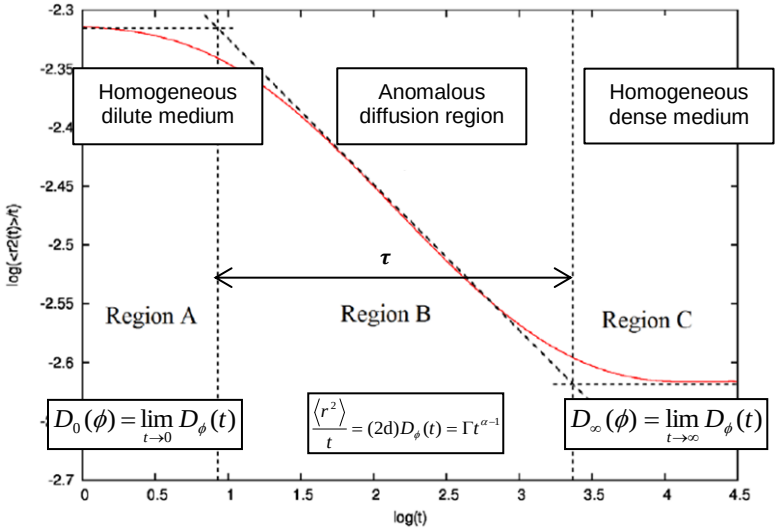

Figure 6. Logarithmic representation of the mean-squared distance versus the logarithm of time for a Brownian particle. There are three differentiated zones: first, a region where the particle feels no obstruction to the movement at all (region A), a time-dependent region that shows how the Brownian particle is exploring the whole solution and starts to feel its movement hindered (region B), and a final region that depicts how the tracer has already felt the obstruction by all the crowding agents (region C). (Image taken from *E. Vilaseca*, ref. 28).

1.2.2 The Langevin equation

The time evolution of the position of a Brownian particle is best predicted using the Langevin equation, the simplest and most widely used mathematical model for Brownian motion. The derivation of this fundamental equation helps to understand how the stochasticity of this motion is described mathematically [31, 32]. Recalling the classic Newton's equation of motion, for a single Brownian particle

$$m\frac{\mathrm{d}v(t)}{\mathrm{d}t} = F(t) \tag{7}$$

it shows that a particle experiences an instantaneous force $F(t)$ at any time t. This force is due to the interaction of the Brownian particle with the surrounding medium, i.e. the solvent molecules. According to Langevin, this total force can be split in, at least, two forces:

- A frictional force, or viscous drag, that represents the dynamical friction experienced by the particle with the solvent molecules. The friction coefficient γ is given by the Stokes law:

$$F_{fric}(t) = -fv(t) = -6\pi\eta r v(t) \tag{8}$$

- A fluctuating force, which represents the collisions of the molecules of the solvent on a Brownian particle. This random force is a stochastic variable that gives the effect of background noise due to the fluid on the Brownian particle. This noise is known to be a Gaussian white-noise (GWN), referring that the process is ergodic and random with zero mean in which the values for different times of the sample are identically distributed in a normal distribution and statistically independent no matter how close they are in time. This noise is represented as [33]:

$$F_{fluc}(t) = \xi(t) \tag{9}$$

The fluctuation force tries to reproduce the effect of the occasional impacts between fluid molecules and the other diffusing particles. It is important to remark the Markov property of this force: the impacts occur in short intervals of time. The timescales between the collisions and diffusion of particles are much separated. At a given time interval, around 10^7 collisions can occur with the fluid molecules. Since these collisions are frequent, any memory between forces at different times will be lost (Markov property), leaving an infinitely short correlation time. Then, it is said that this random force is δ-correlated: there is no correlation between impacts in any distinct time intervals. This behaviour can expressed in the autocorrelation function

$$\langle \xi(t_1)\xi(t_2)\rangle = 2D\delta(t_1 - t_2) \tag{10.1}$$

where the delta function depicts the loss of correlation between collisions in different times t_1 and t_2. Also, another property of the fluctuating force or *noise* can be summarized as

$$\langle \xi(t) \rangle = 0 \tag{10.2}$$

After decomposing the stochasticity of the fluctuating force of the Langevin equation, the whole equation of motion (EOM) for a Brownian particle is written like

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -\frac{6\pi\eta r\, v(t)}{m} + \frac{1}{m}\xi(t) \tag{11.1}$$

$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = v(t) \tag{11.2}$$

However, for Brownian Dynamics simulations it is more convenient to use the over-damped Langevin equation with isotropic diffusion [34, 35]:

$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = -\frac{\nabla V(R(t))}{f} + \sqrt{\frac{2k_\mathrm{B}T}{f}}\frac{\mathrm{d}\xi(t)}{\mathrm{d}t} \tag{12}$$

where the first term in the right-side of the equality of equation 12 represents the deterministic force, governed by a gradient of potential and friction forces, and the second term depicts the stochastic force.

The fluctuation-dissipation theorem relates both frictional and temperature via the diffusion constant:

$$D = \frac{k_\mathrm{B}T}{f} \tag{13}$$

and henceforth the previous Langevin equation is rewritten as

$$\frac{\mathrm{d}\boldsymbol{R}(t)}{\mathrm{d}t} = -D\frac{\nabla V(\boldsymbol{R}(t))}{k_\mathrm{B}T} + \sqrt{2D}\frac{\mathrm{d}\xi(t)}{\mathrm{d}t} \tag{14}$$

The integration form of the Langevin equation with constant time step Δt is used as the EOM followed by Brownian particles, and provides a sequence of positions for a given time step.

$$\boldsymbol{R}(t+\Delta t) = \boldsymbol{R}(t) - \Delta t D\frac{\nabla V(\boldsymbol{R}(t))}{k_\mathrm{B}T} + \sqrt{2D\Delta t}\xi(t) \tag{15}$$

1.3 Enzyme Kinetics: The Michaelis-Menten mechanism

Enzymes are proteins that catalyse biological reactions by making them kinetically favourable, and are the main focus of this project. An usual mechanism to study enzyme kinetics, and the one we are following in this study, is the irreversible Michaelis-Menten (MM) model of enzymatic reactions, proposed by Leonor Michaelis and Maud Menten in 1913 [36]. Such model decomposes the enzymatic reaction as the following

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} \mathrm{E} + \mathrm{P} \tag{16}$$

The first part of the whole reaction is considered reversible. The substrate (S) binds to the active centre of the enzyme (E) by specific interactions and form a reversible complex (C) in a fast step governed by a bimolecular reaction rate constan7

t k_1, and can reconvert into the reactive species by means of a much smaller inverse constant k_{-1}. The second part of the global reaction represents the actual reaction itself, and in the original MM scheme it is considered to be irreversible. The complexed enzyme catalyses the conversion of the substrate into product (P), and finally, both species unbind, recovering the enzyme in its "free" (not complexed) form. This second step is much slower than the first part, being the limiting step of the whole reaction, and is governed by the catalytic constant k_2, also known as turnover number.

1.3.2 The species' governing equations

The populations for each chemical species in the system evolve in time according to a set of coupled ordinary differential equations, or ODEs, when the system is considered to be thermally equilibrated and well-stirred:

$$\frac{dX_i}{dt} = f_i(X_1, \cdots, X_N) \tag{17}$$

where the index i accounts for the chemical specie, X_i denotes the type of the specie, and f_i are functions that are deduced from the specifics of the various reactions and indicate how the populations change with respect to time, and for chemical reactions such constants are the reaction rate constants [37]. By applying the law of mass action to the MM reaction mechanism, the ODEs governing the kinetics of the system are:

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = k_{-1}[C] - k_1[E][S] \tag{18.1}$$

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = (k_{-1} + k_2)[C] - k_1[E][S] \tag{18.2}$$

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1[E][S] - (k_{-1} + k_2)[C] \tag{18.3}$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_2[C] \tag{18.4}$$

The steady-state approximation assumes complex concentration to remain constant in time. Two important parameters in enzyme kinetics can be extracted from such approximation. The first one is the definition of the Michaelis constant:

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1[E][S] - (k_{-1} + k_2)[C] = 0 \quad (19.1)$$

$$\frac{[E][S]}{[C]} = \frac{k_{-1}+k_2}{k_1}$$

$$K_\mathrm{M} = \frac{k_{-1}+k_2}{k_1} \quad (19.2)$$

The Michaelis constant (equation 19.2) is, thus, a dissociation constant, independent of enzyme and substrate concentrations. Nevertheless, it cannot always be considered the dissociation constant K_d of the complex. Otherwise, K_M can be considered the dissociation constant K_d when the value of the turnover number k_2 is much smaller than the inverse reaction rate constant k_{-1} ($k_2 \lll k_{-1}$) [38]:

$$\text{if } k_2 \ll k_{-1} \rightarrow \frac{k_{-1}}{k_1} = K_\mathrm{M} \equiv \frac{[E][S]}{[C]} = K_d \quad (20)$$

and, consequently, K_M becomes a measure of the strength of the complex (high K_M values indicate weak binding, and strong binding otherwise).

The second implication resulting from the steady-state approximation (equation 19.1) is the asymptotic value of complex concentration. By considering the enzyme mass balance, $[E]_0 = [E] + [C]$, and assuming that substrate concentration is in great excess with respect to the enzyme concentration, it is easy to find that the complex concentration is equal to:

$$[C]^{0^{\mathrm{th}}\mathrm{app}} \approx \frac{[E]_0[S]_0}{K_\mathrm{M}+[S]_0} \quad (21)$$

which is the 0th order approximation to the complex concentration for large reaction times. This expression, when substituted in equation 18.3, leads to the reaction rate of the whole reaction:

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = v = \frac{k_2[E]_{tot}[S]_{tot}}{K_\mathrm{M}+[S]_{tot}} = \frac{v_{max}[S]_{tot}}{K_\mathrm{M}+[S]_{tot}} \quad (22)$$

The first order approximation arises from integrating once the differential equation that describes complex variation with respect to time. Like in the 0th order approximation, substrate concentration is assumed to keep stationary due to its excess regarding enzyme concentration. After a straightforward integration, the time-dependent approximation leads to a complex concentration behaviour like:

$$[C]^{1^{\mathrm{st}}\mathrm{app}} \approx \frac{[E]_0[S]_0}{K_\mathrm{M}+[S]_0}\{1 - exp[-(k_1[S]_0 + k_{-1} + k_2)t]\} \quad (23)$$

Nevertheless, instead of using approximations (whichever 0th or 1st order), one would like to consider the resolution of the ODEs and have a deterministic overlook to the variation of the concentrations with respect to time. By imposing again the enzyme mass balance and the substrate mass balance ($[S]_0 = [S] + [C] + [P]$), the system of equations 18.1–4 can be expressed as a pair of coupled equations:

$$\frac{d[S]}{dt} = k_{-1}[C] - k_1[S]([E]_0 - [C]) \quad (24.1)$$

$$\frac{d[C]}{dt} = k_1[S]([E]_0 - [C]) - (k_{-1} + k_2)[C] \quad (24.2)$$

which can be solved numerically by any mathematical method.

However, these equations consider the chemical reaction to be homogeneous and well-stirred (ideal mixture). In that sense, the law of mass action does not provide an accurate description of the temporal population evolution of the species in systems that are obstructed. These crowded environments may provide faster population consumption/production when the reactants are confined or slower when these species are found in different sub-regions of the reacting environment. This implies an important change in the bimolecular rate constant, since E-S encounters may become enhanced (if confined) or hindered (if crowded). Nonetheless, first-order reaction constants are unaffected by constraints on diffusion and mixing generated by physical obstacles, and thus their kinetics are not affected by macromolecular crowding.

The rate v of an enzymatic reaction is described by the MM mechanism as follows

$$v_0 = \frac{v_{max}[S]_0}{K_M + [S]_0} = k_2[C] \quad (25.1)$$

$$v_{max} = k_2[E]_0 \quad (25.2)$$

which would be the Michaelis-Menten equation. The enzyme's maximum rate is attained when the catalytic sites of the enzyme are saturated with substrate.

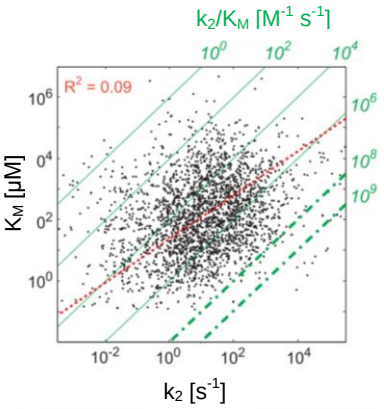

Figure 7. The most common Michaelis and catalytic constant values and catalytic efficiencies (green lines) in biological enzymes. Bold, dashed lines represent the diffusion limit for small metabolites interacting with proteins. Each dot represents the k_2 and the K_M of a given enzymatic reaction with its corresponding substrate. The red line represents the correlation between K_M and k_2 across some reactions studied by the authors. (Image taken from *A Bar-Even*, ref. 37).

In the literature, it is more usual to find the measure of the catalytic efficiency (k_2/K_M), where the average enzyme exhibits values of k_2/K_M ~ 10^5 M^{-1} s^{-1} (see figure 7). The turnover number is in the range of 1 – 100 s^{-1} (~ 10^{-8} ns^{-1}), and its theoretical limit tells that it is impossible to reach k_2 values higher than 10^6 – 10^7 s^{-1}. Thereby, K_M constants take values of 10^{-4} M, or analogously ~10^{-4} nm^3 per particle (see figure 7 for a detailed representation of the most abundant k_2 and K_M values in biological systems). Typical k_1 constants tend to be equal or greater than the catalytic efficiency,

$$k_1 \geq k_2/K_M \tag{26}$$

which implies a range of values compressed between 10^5 and 10^9 M^{-1} s^{-1} (being the largest limit much less common than the smaller limit). These typical rate constant numbers depict that the most common enzyme kinetics are usually much below the diffusion limit; in contrast, most of the enzymes have moderate catalytic efficiencies, far from the kinetic perfection [39].

The MM mechanism (equation 16), however, is only valid at initial reaction times: it assumes that the initial conditions must have a large excess of substrate with respect to the enzyme concentration, so the substrate can be considered saturated (usually the initial substrate concentration is set to be five or six orders of magnitude greater than the enzyme concentration). Moreover, the decomposition of the complex is assumed to be irreversible (a fact that is not always true, since there is a wide extension of enzymes with negative cooperativity to the direct reaction due to the formation of the product [40]).

1.3.1 Fractal kinetics

Some computational studies have demonstrated that, when the reaction media is diffusion-controlled, the time required for any two reactants to interact and react increases and k_1 becomes time-dependent, showing a decaying trend over time [15, 41, 42]. When this happens, the system is said to follow fractal kinetics.

The time dependency of the rate constant takes the form

$$k_1(t) = k_0\left(1 + \frac{t}{\tau}\right)^{-h} \tag{27}$$

where the use of the fractal-like kinetic parameter h, always positive, becomes primordial to quantify the importance of such type of kinetics. In equation 19, k_0 represents the rate constant at initial reaction times, where the constant is still time-independent, and the τ parameter takes the same meaning as in diffusion (*crossover time*). Notice that when h is 0 the system does not follow fractal kinetics anymore.

It is also important to remark that h is greater than 0 in diffusion-limited conditions, which implies that the binding probability of S and E (the probability related to k_1, which we label as p_1) is approximately equal to 1.

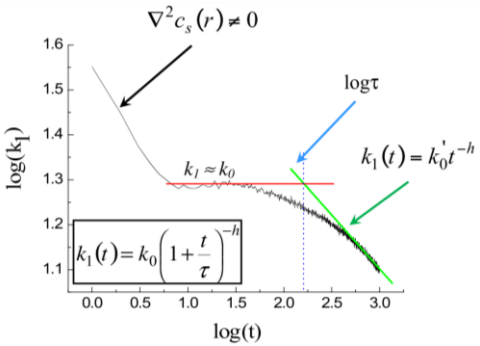

Figure 8. Log-log plot for the simulated forward rate constant k_1 as a function over time in the Michaelis-Menten mechanism. Diffusion stationary state is not reached until log(t) ≈ 1.0. Long-time behaviour shows fractal-like kinetics characteristics, i.e. linear decay of k_1. k_0' is a dimensionless parameter that relates k_0 as $k_0' = k_0\tau^h$ (Image taken from *L. Pitulice et al.*, ref. 39).

1.4 Brownian Dynamics

Diffusion and reactivity in crowded media can both be studied by means of experimental and computational techniques. *In vivo-like* experiments allow to recreate the crowded cytosol by means of inert and neutral polymers in the reaction cell [5]. However, experiments usually have only a few types of crowding agents at hand that accomplish such statements. On the other hand, *in silico* experiments can be used to gain insight into the behaviour and influence of crowding on both diffusion and reactivity at different levels of modelling detail.

Since the system to study (the cellular environment) is small enough, the set of deterministic ODEs (equations 18.1–4) may not depict accurately the system's true behaviour with respect to the population evolutions. In the present case, the molecular populations of at least some of the reactant species are not too many orders of magnitude larger than one (which is often the case in intracellular environments), so discreteness and stochasticity may play important roles. Thus, simulation stochastic algorithms become an adequate approach that, by performing a great amount of repetitions of the same simulation, the average of the properties from these different repetitions become really close to the analytic result of the deterministic approach of the case (deterministic limit) [43].

Having already discarded deterministic models, then, different paths can be followed in order to approach the problem stochastically. Stochastic simulation algorithms (SSA), such as

the Gillespie SSA [37, 44], or other variations also used in our group [29], are Monte Carlo (MC) procedures for numerically generating time trajectories of the molecular populations. However, in our study we also want to take into consideration the actual effect of diffusion upon reactivity, and the effect of macromolecular crowding onto diffusion itself. Brownian Dynamics (BD) is the stochastic dynamics scheme usually used to simulate such biochemical networks [45, 46]. In this framework, the solvent is treated implicitly and only the solutes are described explicitly. The collisions between solute and solvent particles are simulated by means of the fluctuation force of the equation of motion. This force is a random element and is part of the net force of the system, in addition to contributions from interactions with other particles.

Reactivity is studied by means of specific probabilities that try to reproduce the correct reaction kinetics of the system [47]. BD allows to study all types of reactions (unimolecular, bimolecular, …), but their reaction probabilities must accomplish all the requisites that allow a reaction to take place, such as particle collisions and particle orientation, amongst others.

Nowadays, many reaction-diffusion algorithms incorporating the BD framework have been developed with many levels of detail. The basic level of detail, which is the fundament of any BD-based algorithm, implies free diffusion in a continuous space in a simulation box, and particles react with a certain probability when become close to each other. Implementation of membranes or channels (confined diffusion), the addition of other non-reacting particles in the simulation box to increase the excluded volume of the system (macromolecular crowding) and the inclusion hydrodynamic interactions and other particle-particle potentials (a level that becomes very close to Molecular Dynamics simulations) imply higher levels of detail that try to represent biological systems as accurate as possible.

1.4.1 Hydrodynamic interactions

Hydrodynamic interactions are usually added to BD codes in order to improve and obtain a more realistic approach on Brownian motion [48, 49]. These interactions arise when a Brownian particle collides with solvent particles which, in turn, also collide with other macromolecules. This correlation of the particle's movement through the liquid induces a change in the dynamical properties of the system. One of the affected parameters is the diffusion coefficient, which experiments its decrease with respect to the dilute solution diffusion coefficient D_0.

The simplest approach to treat the solvent in BD is by means of the stochastic term of the Langevin equation. To improve the level of detail, one has to take into consideration the effect of these hydrodynamic interactions. The Tokuyama model [50] is a mean-field method and is the

easiest approach to include the effect of hydrodynamic interactions in a BD simulation. The dilute diffusion coefficient D_0 for hard-sphere suspensions is corrected according to the excluded volume present in the system by means of the following equation:

$$D_0^T(\phi) = \frac{D_0}{1+H(\phi)} \tag{28.1}$$

$$H(\phi) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)} \tag{28.2}$$

$$\text{with } b = (9\phi/8)^{1/2} \text{ and } c = 11\phi/16 \tag{28.3}$$

In a Brownian Dynamics algorithm, thus, the diffusion coefficient in dilute conditions is corrected with the short-time Tokuyama diffusion coefficient, $D_0 = D_0^T(\phi)$. This model also counts with an analytic equation that allows predicting the diffusion coefficient after long simulation times, D_∞

$$D_\infty^T(\phi) = \frac{D_0^T(\phi)}{1 + \kappa \frac{D_0^T(\phi)}{D_0} \left(\frac{\phi}{\phi_c}\right) \left(1 - \frac{\phi}{\phi_c}\right)^{-2}} \tag{28.4}$$

where κ and ϕ_c are analytical parameters. For BD algorithms, $\kappa = 2.0$ and $\phi_c = 1.09$.

1.4.2 Brownian Dynamics simulation codes

There exists a large catalogue of BD codes. To name some of them, the most known and used tools include Cell++ [51], Smoldyn [52], ChemCell [53], Mcell [54], Ridgway [55], and ReaDDy [56], amongst others (see ref. [57] for a review article). Many improvements, variations and implementations for the BD framework have been proposed in the past years [13, 47, 49, 56-61], including different treatments for the biomolecules in the system, for instance by increasing the level of the coarse-graining of the proteins; adding more interactions between the species; studying confined environments with membranes and ionic channels; etc.

All in all, one has the freedom to choose the code that suits more the needs of the research. However, most of these codes are black boxes that do not allow a modification on its internal code (and some of them present non open-source features). Due to this fact, our group has decided to tailor a BD code that suits the demands of our research, and allows an easy modification of the code whilst progressing throughout the research. This includes the modification of the interaction forces between particles, the inclusion of a detailed description of hydrodynamic forces, the study of the reversible Michaelis-Menten mechanism, etc. To do so, the ReaDDy [56] software package has been used as a basis and guideline to develop our BD code, yet in the end they result very different reaction-diffusion simulation codes.

2. OBJECTIVES

The main objective of this master thesis is to develop a reaction-diffusion Brownian Dynamics simulation code in C++ programming language. Our algorithm must simulate enzymatic systems that follow the Michaelis-Menten mechanism and allow the study of the effect of macromolecular crowding. As a reference, we use previous Brownian Dynamics algorithms as a reference.

The code's performance must be validated in both diffusion and reaction terms. Diffusion validations should be tested by time-step consistency studies; Reactivity ought to be tested by comparing the specie variation throughout the whole simulation with the numerical solution of each specie's respective ordinary differential equation, which determine the theoretical variation of specie concentration with respect to time.

After validation, we perform diffusion and reactivity studies and compare our results with previous studies performed by our working group, both *in vivo*-like and *in silico* (on-lattice) experiments.

3. METHODOLOGY

In order to simulate the intracellular environment at the biologically relevant length and time scales, proteins and substrates are modelled explicitly as individual, non-overlapping coarse-grained hard spheres that are transported via stochastic dynamics and undergo reactions with other chemical species when certain criteria are met. The algorithm and the theoretical background of the different subroutines of the code, that we named RK3D (Reaction Kinetics in 3 Dimensions) are explained on the following sections.

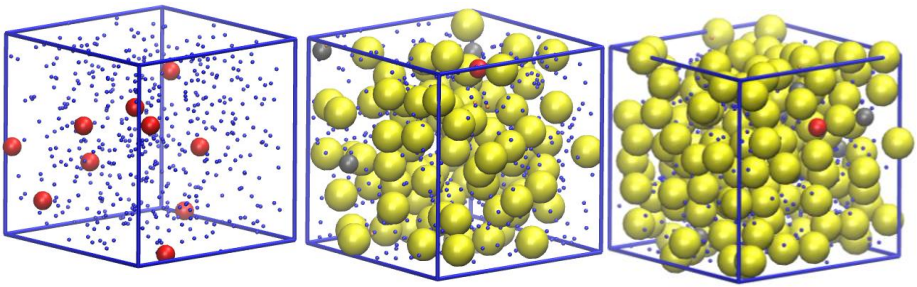

Figure 9. Example of a simulation box with enzyme (red), substrate (blue), complex (grey), product (orange) and obstacle (yellow) particles in dilute conditions (left), with 20% of excluded volume (centre) and with 40% (right).

3.1 ALGORITHM OUTLINE

The code follows an iterative process that works over all the particles in the system.

1) Generate initial positions for all the particles in the system in the available volume by means of three different random numbers (one for each spatial dimension).
2) Equilibration step (if chosen to be non-zero).
    a. Calculation of the particle-particle interactions.
    b. Update of the positions of the particles by means of the integrated form of the Langevin equation.
3) Evolution of the system and calculation of properties
    a. Calculation of the particle-particle interactions.

b. Update of the positions of the particles by means of the integrated form of the Langevin equation.

c. Determine whether a reactive event occurs by means of the Monte Carlo algorithm.

Particle populations, mean squared displacements and the instantaneous reaction rate constants are calculated at each iterative step and are printed with a frequency previously given by the user.

The random number generator implemented in the code is called `ran2(idum)` and can be found in Numerical Recipes [62]. It generates random numbers uniformly distributed in [0,1].

3.2 INITIALIZATION OF THE SYSTEM

The input of the code requires the specification of the length of the sides of the simulation box, the number of substrate (S), enzyme (E) and obstacle (two different types with different physical properties, Obs1 and Obs2) particles to simulate, the radii of the spheres and their diffusion coefficient (if chosen, such coefficient can be calculated by means of equation 2). The reaction probabilities must be specified, along with the number of iterative steps, the number of equilibration iterative steps, the value of the time-step, and the temperature and a repulsive parameter that depicts the repulsive interaction of the particles.

The code saves the number of particles according to their nature in a vector. Since all the substrates can potentially become into products, the vector saves as many P particles as S particles indicated in the input file, and the same happens for the E and C particles. The initial coordinates of the initial particles (S, E, Obs1 and Obs2) are generated randomly by means of uniformly distributed random numbers. The P and C particles are saved as dummies until their "counterparts" react: each of these S-P and E-C groups of particles are paired to each other, in the sense that when an E particle reacts, its coordinates are swapped to the C particle that has the same particle index to the reacting E (its pair); since the coordinates are swapped, the E is now saved as a dummy until a reactive event occurs again.

Since the initial coordinates are generated randomly, it is very probable that particles are overlapped. This would generate an erroneously big displacement for the particles at the following iterative step. To prevent that, the equilibration step is run (if chosen to be non-zero), where particles diffuse. Reactions do not take place in this equilibration phase.

3.3 IMPLEMENTATION OF THE LANGEVIN EQUATION

3.3.1 Deterministic force: the interaction potentials

The implementation of the deterministic force of the Langevin equation is straightforward. In our approach, a particle-particle repulsive potential is implemented to mimic non-specific interactions. This potential is governed by a quadratic, harmonic potential that acts on all pairs of particles i and j in the system.

$$V(x,y,z) = \begin{cases} \frac{1}{2} k_{pair} \left(d_{ij} - r_{ij}\right)^2 & d_{ij} < r_{ij} \\ 0 & d_{ij} \geq r_{ij} \end{cases} \quad (29)$$

where $d_{ij} = \left[\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2\right]^{1/2}$ is the pairwise distance between particles i and j (considering no self-interaction), and $r_{ij} = (r_i + r_j)$ are the sum of the collision distances of the respective particles. k_{pair} is an inter-particle repulsion force constant, which is chosen to be 10 kJ mol^{-1} nm^{-2} [56].

Hence, this repulsive potential is non-zero when two particles are in contact with each other or become overlapped after an iterative step in the simulation (see figure 10). Besides, this potential is only calculated for particles i and j that have not reacted within themselves in that given iteration.

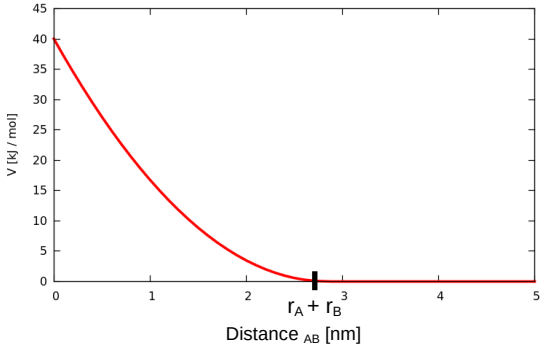

Figure 10. Quadratic potential (in one dimension) according to equation 29 felt when a particle *A* of radius r_A = 2.33 nm and a particle B of radius r_B = 0.5 nm find themselves at different distances.

The Langevin equation moves the particles of the system by means of forces, so the gradient of this quadratic potential is the one used in the code. This gradient has the form of:

$$\mathrm{d}V_x = \frac{\mathrm{d}V(x,y,z)}{\mathrm{d}x} = \begin{cases} \frac{k_{pair}(x_i - x_j)(d_{ij} - r_{ij})}{d_{ij}} & d_{ij} < r_{ij} \\ 0 & d_{ij} \geq r_{ij} \end{cases} \quad (30)$$

and the other two equivalent equations for the other two dimensions of the three-dimensional space. This force is the one that governs the non-stochastic part of the EOM.

3.3.2 The stochastic force

To calculate the stochastic term of the Langevin equation, the code needs to pick a random point from a Gaussian distribution. Since the random number generator implemented in the code provides uniformly distributed points, the polar form of the Box-Muller transformation [63] is applied to transform these set of variables from the uniform distribution into a new set of variables that are normally distributed. The basic form of this algorithm is a bit slow and sometimes haves numerical stability problems, whereas the polar form is both faster and more robust, and is widely implemented in stochastic modelling.

The algorithm for the polar form of the Box-Muller transformation is included in Numerical Recipes [62]:

```
do {
        x1 = 2.0 * ran2() - 1.0;
        x2 = 2.0 * ran2() - 1.0;
        wy = x1 * x1 + x2 * x2;
} while ( wy >= 1.0 )||( wy == 0.0 ));
wy = sqrt( (-2.0 * log( wy ) ) / wy );
r1 = x1 * wy;
```

where `ran2`, as stated previously, generates a random number uniformly distributed in [0,1], and `r1` is a random number following a Gaussian distribution with $\sigma = 1$.

3.4 DISPLACEMENT MODELLING

The code generates a trail movement at each time interval, in both equilibration and production processes. Particle displacements are described by the Langevin equation, which generates a 3D random vector that is related to the diffusion coefficient of the particle, according to its nature (enzyme, substrate, complex, product or obstacle). Our code uses periodic (or cyclic) boundary conditions (PBC), so the system is virtually replicated infinite times by just modifying the particle coordinates each time they are supposed to leave the simulation box. The pseudocode of PBC is (for just one dimension, can be extrapolated to the three dimensions)

```
if x_new[i] > L_box
    x_new[i] = x_new[i] – L_box
```

```
else if x_new[i] < L_box

   x_new[i] = x_new[i] + L_box
```

Particle positions are saved at each iterative step to obtain the mean-squared displacement for every type of particle (Einstein-Smoluchowski relation, equation 3). This magnitude is not calculated during the equilibration phase, being the initial coordinates the ones belonging to the first production iterative step.

3.5 REACTION-DIFFUSION MODELLING

Reactivity in the Brownian Dynamics framework implies the study of reactive collisions or events leading to the formation of the enzyme-substrate complex and the subsequent dissociation of the complex to form either substrate or product. The Monte Carlo (MC) algorithm is used in order to decide whether a collision is a reactive event or not (for bimolecular reactions) or whether a reactive event occurs or not (for unimolecular reactions). For a bimolecular reaction, the collision between the two reactive particles leads to products if a random number is greater or equal to the bimolecular reaction probability, which is a value compressed between 0 and 1 (where 1 indicates maximum reaction probability).

For the Michaelis-Menten mechanism (equation 16), to study the complex formation rate one has to take into account the following scheme:

$$E + S \rightleftharpoons E \cdot S \rightleftharpoons C \rightarrow E + P \tag{31}$$

To convert the reaction constants into probabilities, we follow the procedure used by the ReaDDy developers [56]. The relation between the microscopic constants (all of them are expressed in time^{-1} units) and their macroscopic counterparts (the ones that are measured experimentally) depend on the order of the reaction. For unimolecular reactions, the relationship is $k_{unim}^{micro} = k_{unim}^{macro} = k_{unim}$, since the reaction rate constant depicts the inverse mean time needed for the reactant to decay into a product.

On the other hand, bimolecular reactions require more detail to obtain an accurate relationship. The first event in the decomposed MM mechanism represents the formation of the edduct or the encounter complex after a collision between an enzyme and a substrate particle. In BD, a collision takes place when two particles are close to each other and a trial displacement of one of the reactants leads to an overlap with the other particle. This encounter rate is parametrized by the Smoluchowski rate constant k_{enc}, which is defined for bimolecular

reactions in three-dimensions homogeneous media with spherical particles in infinite dilute conditions [64-66]. This constant reflects reactions that are diffusion-limited:

$$k_{bim}^{enc} = 4\pi D_{ES} r_{ES} \quad (33)$$

with $D_{ES} = D_E + D_S$ is the sum of the diffusion coefficients of the reactive species and $r_{ES} = r_E + r_S$ is the sum of their respective radii. Equation 33 is expressed per particle units. If it wants to be expressed in the IS units, such equation has to be multiplied by the Avogadro constant.

However, the bimolecular reaction into complex will happen according to a kinetic-limited microscopic rate constant that takes into consideration the time spent by the two reactants achieving a proper orientation that allows the reaction to take place (the orientation of the substrate in the active space of the enzyme by means of electrostatic interactions, steric hindrances, rotations, etc.). The equation used in this work to represent this process is

$$k_{bim}^{macro} = 4\pi D_{ES}\left(r_{ES} - \sqrt{\frac{D_{ES}}{k_{bim}^{micro}}}\tanh\left(\sqrt{\frac{k_{bim}^{micro}}{D_{ES}}}\right)\right) \quad (34)$$

which would be a reduced form of the Smoluchowski encounter constant that takes into consideration the orientation of the particles and the probability of the substrate to bind into the enzyme's active space [56, 58]. Equation 34 becomes equation 33 when the microscopic constant tends to infinite, and so equation 34 reaches the asymptotic value of k_{enc}. This latter equation provides the relationship between the microscopic and macroscopic bimolecular reaction rate constants. When the microscopic constant is much smaller than $D_{ES}/(r_{ES})^2$, by means of a Taylor expansion it is easy to demonstrate that equation 34 can be simplified to a linear expression like $k_{bim}^{macro} \approx k_{bim}^{micro}(4/3)\pi {r_{ES}}^3$, where both constants are related by the spherical volume in which the reaction takes place [58].

The reaction probabilities are directly related with the microscopic reaction rate constants if we consider such probabilities to be Poisson-distributed. A Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events (in this case, reactive events) in a time interval. Then, the probability of accepting a reactive event is obtained from the Poisson probability of finding at least one reactive event in a time window Δt [47, 56, 66, 67]

$$p(\Delta t) = 1 - exp(-\Delta t k_{micro}) \approx k_{micro}\Delta t \quad \text{if } k_{micro}\Delta t \ll 1 \quad (35)$$

In a dynamics simulation, the time window Δt would become the time step between each iterative step. These probabilities represent the probability of finding, at least, a reactive event during an iterative step.

Hence, at each iterative move, a trial move for a particle is chosen according to the equation of motion of the system, which is the memoryless Langevin equation described in section 1.2.2. Particle collisions occur when these trial moves lead to particle overlaps, and will lead to the respective product (complex) according to the Poisson probability. For mononuclear reactions, each iterative step can potentially lead to a reactive event according to the reaction probability of the first-order reaction. By means of the Monte-Carlo criterion, if the random number obtained is greater or equal than the reaction probability, such reactive event takes place.

3.5.1 Enzyme-complex formation

The enzyme and substrate diffuse throughout the simulation box until they encounter each other at a distance equal or smaller the sum of their radii. The reactive collision occurs with a probability p_1, directly related to the microscopic rate constant $k_{1,micro}$ according to equation 35. The reaction leads to a complex molecule if the random number picked is smaller or equal to p_1. The S molecule is saved outside the simulation box, and the E particle is converted into a C particle.

3.5.2 Dissociation of the enzyme-substrate complex

Once a complex particle is formed, it may either dissociate to release the substrate with a probability p_{-1} or the catalytic reaction can occur, producing the main product of the reaction with a probability p_2, always following the MC criterion. These probabilities are directly obtained by means of equation 35. The C particle is swapped into an E particle, and the newly formed X particles (where X can be either S or P molecules) appear at a distance $r_d = r_E + r_X + 0.1(r_E + r_X)$ from the late C particle. The latter term ensures that both E and X do not directly touch each other after dissociating from C. Otherwise, if X is S, the bimolecular reaction would be more proper to occur again, providing an incorrectly enhanced reactivity for the direct reaction.

3.5.2.1 *Coordinates for new formed particles*

The spatial coordinates of the newly formed S or P particles should have a uniform probability to appear around the E particle at a r_d distance. Since the simulation environment is 3-dymensional, spherical coordinates are used to generate the random new position for these X particles.

The spherical coordinate system uses two angles, the polar angle θ, measured from a fixed zenith direction, compresses angles between -π and π; the azimuthal angle φ moves on the reference plane, orthogonal to the zenith, and compresses angles between 0 and 2π.

Figure 11. (Left) Sphere section showing the spherical coordinates: the polar angle θ, the azimuthal angle φ and the radius r. (Right) Sphere wireframe. The area elements closer to the poles are smaller than the ones around the equator.

To pick a random point on the surface of a unit sphere, it is not as straightforward as generating two random numbers and select the spherical angles θ ∈ [0, π) and φ ∈ [0, 2π). The area element of a sphere dΩ is a function of the azimuthal angle, $d\Omega = r^2 \sin(\varphi) d\theta d\varphi$ (see figure 11). If done as explained, an uneven distribution would be obtained, since it is more probable to generate a point closer to the poles of the sphere. For instance, if the polar angle is 0 (the equator of the sphere), the random points generated will be sparse around the equator depending on the azimuthal angle. However, if the polar angle is π (the pole), whatever the value of the azimuthal angle is, there exists a clustering of points on the pole (see figure 12).

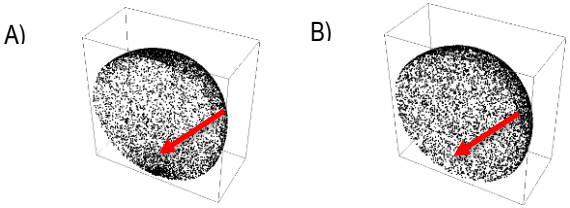

Figure 12. Section and top views of the pole of a sphere and with A) an uneven distribution and B) a uniform distribution of points on their surface. A clustering of points can be seen at the poles of B) sphere.

The correct manner of picking a point from the surface of a sphere with a uniform probability is by generating both angles according to:

$$\varphi = 2\pi\xi_1$$

$$\theta = \arccos(1 - 2\xi_2)$$

The azimuthal angle keeps compressing angles between 0 and 2π, but, by imposing a greater probability density around the equator, the cosine of the polar angle must be compressed

between -1 and 1. From these random angles, the Cartesian coordinates for the newly formed S or P particles are calculated according to the Cartesian-polar coordinate's relation.

3.5.3 Reaction rate constants

According to reaction kinetics and the law of mass action, the rates of the reactions being studied are (recall the Michaelis-Menten mechanism, equation 16):

$$\frac{\mathrm{d}[C]_{k_1}}{\mathrm{d}t} = k_1^{macro}[E][S] \tag{36.1}$$

$$\frac{\mathrm{d}[C]_{k_{-1}}}{\mathrm{d}t} = -k_{-1}[C] \tag{36.2}$$

$$\frac{\mathrm{d}[C]_{k_2}}{\mathrm{d}t} = -k_2[C] \tag{36.3}$$

where equation 36.1 indicates that the association step is a bimolecular reaction, whereas the other two reactions belong to the unimolecular cases. Even though these three rates are expressed in terms of the variation of the concentration of complex with respect to time, such terms do not represent the same: for the bimolecular reaction (equation 36.1), $\mathrm{d}[C]_{k_1}/\mathrm{d}t$ refers to the amount of complex generated due to the association reaction governed by k_1, without taking into consideration the dissociation and catalytic reactions. Since the simulation times for the calculations are very short (t~10^4 ns), the instantaneous concentrations of each species are used in the previous equations.

To determine the macromolecular reaction rate constants, the previous expressions are implemented in the code [13, 15, 68]:

$$k_1 = \frac{[C]\ generated\ due\ to\ association\ reaction}{[E][S]}\frac{1}{\mathrm{d}t} = \frac{C_{k_1}}{E\ S}V\frac{1}{\mathrm{d}t} \tag{37.1}$$

$$k_{-1} = \frac{[C]\ vanished\ due\ to\ dissociation\ reaction}{[C]}\frac{1}{\mathrm{d}t} = \frac{C_{k_{-1}}}{C}\frac{1}{\mathrm{d}t} \tag{37.2}$$

$$k_2 = \frac{[C]\ vanished\ due\ to\ catalytic\ reaction}{[C]}\frac{1}{\mathrm{d}t} = \frac{C_{k_2}}{C}\frac{1}{\mathrm{d}t} \tag{37.3}$$

where E, S and C_j ($j = k_1, k_2$ or k_{-1}) are the number of enzyme, substrate and complex particles, respectively; and V is the volume where the reaction takes place. C_j could as well be referred as a reaction counter, where C_{k_1} counts the number of effective collisions between an enzyme and a substrate particle, and $C_{k_{-1}}$ and C_{k_2} count the number of dissociations of the complex into enzyme and substrate or product, respectively.

By means of these constants, the Michaelis constant can be calculated in terms of both equilibrium's dissociation constant (equation 20), considering the steady state approximation in equation 19.1, and according to the definition of the Michaelis constant (equation 19.2).

4. RESULTS

Since RK3D is a reaction-diffusion code, validation is performed by both diffusion and reactivity. We use as a reference in diffusion studies the α-chymotrypsin enzymatic system, since it has been previously been studied by our group [17], and we make use of previous on-lattice studies for comparing both diffusion and reaction studies for a α-chymotrypsin-like enzymatic system [29, 41]. Once the code is validated, we study the excluded volume effect on both diffusion and reactivity.

4.1 DIFFUSION

4.1.1 Diffusion validation: the time-step consistency

The trajectories of the particles are a continuous function in space with respect to time. Such function must be discretised when performing computational simulations, thereby a time step Δt must be specified, and it stablishes the discretisation level of a simulation. Hence, BD algorithms require relatively small time steps in order to accurately simulate the dynamics of diffusion and reaction kinetics. This makes the selection of the time-step not trivial. Since the total simulation time is related to the time step by $t_{simulation} = \Delta t * iterations$, the higher the value of the time step, the lower the amount of iterative steps required in order to arrive to a given total simulation time. For instance, if one wants to simulate an experiment of 1,000 ns with a time step of 10 ns, only 100 iterations have to be run. The problem, however, relies in that Δt needs to be chosen shorter than the smallest timescale of the system, and it depends on the stiffness of the potential and on the diffusion constants (that is, the time step must belong to the Brownian motion time-scale). Otherwise, particle trajectories lose precision and the particle movement at each time step is incorrectly high. These over-estimated moves may affect incredibly when predicting particle collisions (the number of collisions normally decreases, so diffusion coefficients become erroneously estimated as well as their reactivity [66]).

The most appropriate time step for our simulations must be close, then, to the Brownian limit. This limit is calculated by means of equation 3, in which the mean squared displacement MSD cannot be greater than the radius of the largest particle in the system, which is the enzyme

(the tracer particle). In most of our simulations we make use as a reference the kinetic parameters of α-chymotrypsin (r = 2.33 nm, $D_{Stokes\text{-}Einstein,\ 293.15K}$ = 0.09197 nm^2 ns^{-1}, $D_{Stokes\text{-}Einstein,\ 298.15K}$ = 0.10531 nm^2 ns^{-1}), so the upper limit of the time step is, approximately, $\Delta t \leq 10$ ns.

Different simulations of a system with a fixed concentration of enzymes (one with a single tracer particle and two others using a tracer density of 0.0002 particles nm^{-3}) were tested with seven Δt values: 0.001, 0.01, 0.1, 1.0, 2.0, 5.0 and 10 ns. The diffusion coefficients obtained for the tracer are compared between them and with the theoretical value, calculated according Stokes-Einstein equation. These studies were performed at two different temperatures, at 298.15 K and 293.15 K. The total simulation time run is 100 ns, and each simulation is repeated 1,000 times with k_{pair} = 10 kJ nm^{-2} mol^{-1}. Simulations with more than a single particle were performed using two different equilibration process (using 10 and 100 iterative steps).

For a single tracer particle, a non-realistic situation, the diffusion coefficient is expected to be independent of the value of the time step introduced, as there are no other particles present in the system to collide with (thus the tracer will not feel any repulsive interaction during all the diffusion process), and hence its movement should always be dependent only on its diffusion coefficient. This behaviour is obtained through the simulation with just a single enzyme particle (see figure 13.a).

For denser simulations, figures 13.b—c show that the Δt = 10 ns (the time step limit) provide biased diffusion coefficients to higher values. The displacement of the tracer cannot exceed the value of its radius, as this is the maximum displacement we have imposed previously in order to find the upper limit of the time-step. In that sense, these simulations generated movements greater than the enzyme radius at almost each iterative step. These displacements are erroneously big (a warning signal pops up when this occurs during a simulation), and produce diffusion coefficients larger than expected. The coefficient obtained with the simulation run at 293.15 K using 100 equilibration steps does not show this biased trend, but we believe that this is due to random collisions of the tracers with themselves or simple fluctuations that, in the end, have compensated the erroneous displacements.

Radius [nm]	**Diffusion coefficient 293.15 K** [nm^2 ns^{-1}]	**Diffusion coefficient 298.15 K** [nm^2 ns^{-1}]
2.33	0.09197	0.10531

Table 1. Diffusion coefficients for α-chymotrypsin at two different temperatures calculated using the Stokes-Einstein equation (equation 2).

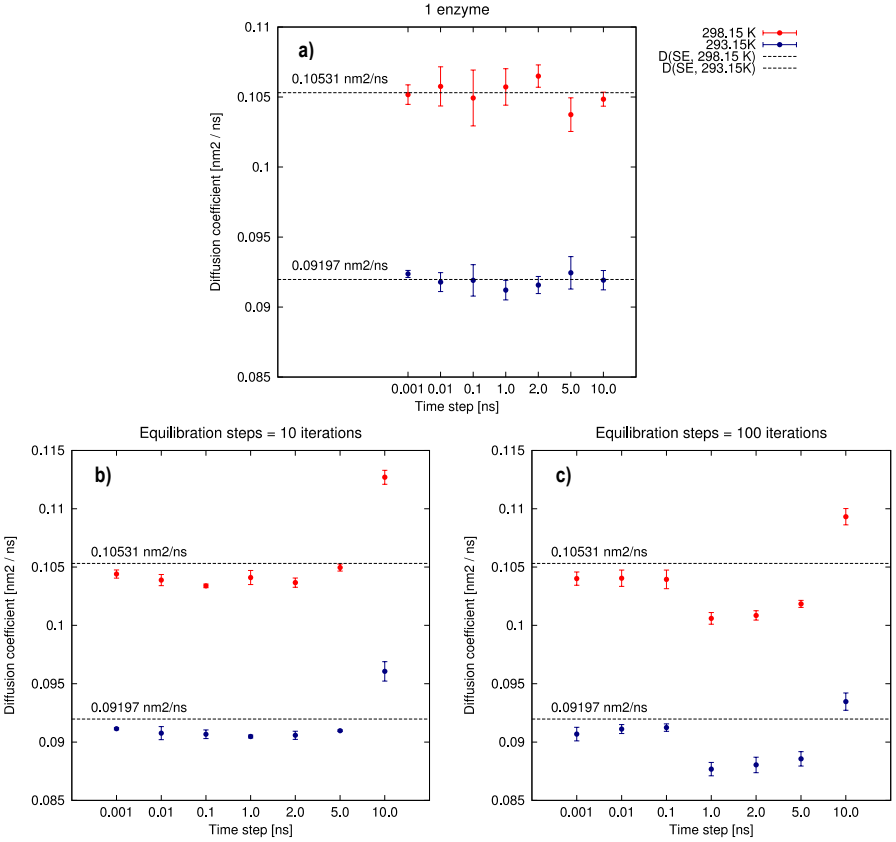

Figure 13. Diffusion coefficients calculated for particles of 2.33 nm of radius obtained using different time step Δt values and two different temperatures, at 25ºC (red dots) and 20ºC (blue dots). Plot a) is for a single tracer particle moving freely in the simulation box; Plots b) and c) have a fixed density of 0.0002 particles/nm^3, and have been previously equilibrated with 10 iterations (b) or 100 iterations (c). The dashed lines represent Stokes-Einstein coefficient.

Time steps Δt = 1.0, 2.0 and 5.0 ns do not show consistency throughout simulations with different equilibration steps. The equilibration phase should relax the system to a particle distribution in the volume such that the interaction potentials are smaller with respect to the initial configuration of the system. These time steps are so big that the particles cannot find a stable distribution: the greater the number of equilibration steps, the greater the number of large displacements is, and thereby it is more probable for particles to find closer to each other at almost every time step. To understand better this phenomenon, one could compare the tracer with a car. The maneuvers to get the car out of its parking place should not provide big

displacements; otherwise it would crash with its surroundings (an unstable situation). This problem is inexistent in the simulations of one enzyme because there are no other particles to collide with throughout the simulation.

In contrast, small displacements allow refining the interactions between particles (the movements with the car) and find a stable situation (taking the car without crashing). Simulations with time steps Δt = 0.001, 0.01 and 0.1 ns are small enough to provide adequate displacements, and are stable independently of the number of equilibration steps used. Since we need to find a compromise between accuracy and computational time consumption, the most adequate time step is chosen to be 0.1 ns (if Δt = 0.001 ns, we would require 10^8 iterative steps to generate 10^5 ns of simulation; instead, if Δt = 0.1 ns, only 10^6 iterative steps are needed.

4.1.2 The excluded volume effect on diffusion

Diffusion studies with different values of excluded volume are performed with an enzyme concentration of 0.001 particles/nm^3, using crowding agents with a radius of 2.33 nm, as the tracer particle, and calculating the dilute diffusion coefficients D_0 with the Stokes-Einstein equation. The enzyme radius selection is chosen in order to perform further comparisons with experimental results for α-chymotrypsin. Figure 14 shows the logarithmic representation of the diffusion coefficient versus time in order to validate the macromolecular crowding effects on diffusion.

Before studying diffusion under the influence of the excluded volume effect, a single enzyme particle is simulated in the same volume. Its diffusion coefficient remains constant throughout the whole simulation, $D \neq D(t)$, regardless the fluctuations due to a modest number of repetitions. This behaviour is expected, since the tracer does not interact with any other species in the simulation box (as seen in the previous section with the time-step consistency). On the other hand, the presence of other species makes the diffusion coefficient a time-dependent variable with a decreasing trend. Even when no crowding agents are present, the tracers themselves hinder their movement through the simulation box (autocrowding). In that sense, in realistic systems the diffusion coefficient is never time-independent.

As the presence of non-reactive species in the system increases, the steeper the slope of the log-log plot becomes (see equation 4). This implies a decrease of the anomalous diffusion exponent α, revealing that particles experience anomalous motion due to crowding agents, in

such a way that subdiffusion becomes a focal point in crowded environments. Consequently, this yields to a smaller limiting diffusion coefficient D_∞. Notice that the crossover time τ diminishes as the excluded volume increases.

Thereby, the trend observed throughout these experiments relates an increase of excluded volume with a decrease of both characteristic parameters of the diffusion process.

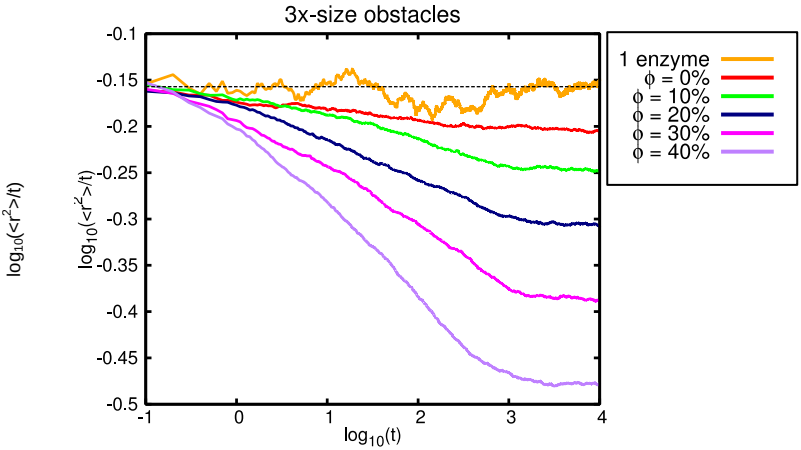

Figure 14. Logarithmic representation of the mean-squared displacements $\langle r^2 \rangle$ versus time for simulations with 0.001 tracer nm^{-3} and different values of excluded volume (Φ). Each simulation has been repeated 900 times, using 10^5 productive iterations with a time-step $\Delta t = 0.1$ ns.

4.1.2.1 *Obstacle size effect on diffusion*

The effect of macromolecular crowding on diffusion has been studied with different obstacle sizes, using the same radius as the tracer (that we denote as 1x-size), three times bigger (3x-size), five (5x-size) and seven (7x-size) times, and also with Φ = 10, 20, 30 and 40%. Figure 15 reveals the effect on the anomalous diffusion exponent α and the limiting diffusion coefficient D_∞. In figure 15.b we represent this coefficient with respect to the diffusion coefficient in dilute conditions, i.e. we represent the relative diffusion coefficient $D^* = D_\infty / D_0$.

The main observation from these experiments is that an increase of the crowding agents' concentration implies a bigger hindrance to diffusion, as more particles are present in the medium. However, the effect of crowding size may not be so obvious. As the excluded volume increases, big obstacles cause less obstruction to particles' movement than small obstacles.

For a given excluded volume, the number of small obstacles required to exclude that volume is higher than for bigger obstacles. Because of that, small obstacles produce a greater

fragmentation of the total volume, implying a higher number of collisions between the crowding agents and the diffusing particles. Hence, the bigger the concentration of small obstacles the more hindered becomes diffusion, leading to bigger subdiffusive motions. The limiting diffusion constant denotes the same trend, where obstacles of the same size as the tracer reduce its diffusivity up to the 50% from its value in dilute conditions when $\Phi = 0.3$.

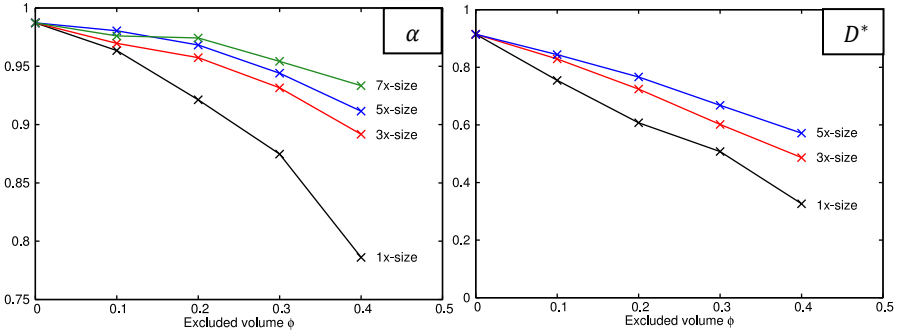

Figure 15. Anomalous diffusion exponent (α) and limiting diffusion coefficient ($D^* = D_\infty/D_0$) behaviour with excluded volume using different sizes of obstacles: the same size as the enzyme (black), three times bigger (red), five (blue) and seven times bigger (green). D^* values for obstacles seven times bigger the enzyme could not be calculated, since the mean-squared displacement curves presented bumps that did not allow its accurate measurement.

The effect of immobile obstacles is also studied within the same tracer conditions. Fixed crowders produce an even more intense effect on both diffusion parameters α and D^* than mobile obstacles. In other words, diffusion becomes more obstructed when the obstacles are forced to stay fixed. The effect of fixed obstacles becomes more important when the tracers are of the same size of the enzyme (1x1x1) rather than other sizes (results not shown). Again, for a fixed value of excluded volume, the number of small obstacles required to generate such volume is greater than for bigger obstacles, as the latter occupy more volume. In that sense, a large number of fixed molecules obstruct the diffusing media, in such a way that channel-like pathways are created, forcing the tracer particles to find the appropriate path to move through (the less hindered path).

The diffusive behaviour of α-chymotrypsin has been previously studied by our group with both *in vivo-like* and other *in-silico* methodologies. Fluorescence Recovery After Photobleaching (FRAP) experiments were performed by marking the enzyme with fluorescence to follow its motion, using Dextrans as crowding agents [17]. On the other hand, our group also performed an on-lattice study of an α-chymotrypsin-like enzyme's diffusive behaviour with different

obstacle sizes [29]. The on-lattice approach also studied diffusion under the effect of a mobility reduction factor, which reduces the mobility of tracers when they are close to obstacles.

A comparison between diffusion simulations using obstacles three times bigger the simulated enzyme and its analogue in the experimental case with Dextran D50 (M_w = 50 kDa) is performed. We see that the off-lattice approach provides closer results to the experimental data than the Monte Carlo simulations for the diffusive parameters anomalous diffusion exponent α and the relative limiting diffusion coefficient D^*, as it can be observed in figure 16. The diffusion coefficients obtained by our BD code, however, are in quite better agreement with the experimental results rather than the values obtained for the anomalous diffusion coefficient.

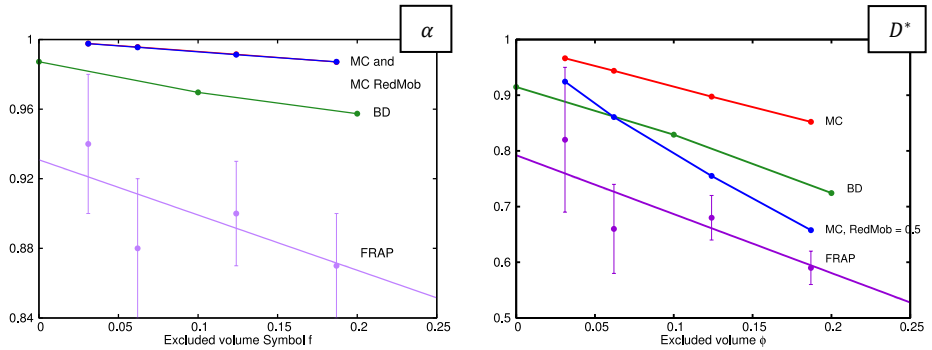

Figure 16. Comparison between *in vivo-like* FRAP experiments (violet dots) using Dextran D50 as crowding agent and *in silico* experiments, using an on-lattice Monte Carlo (MC) approach (red and blue dots) and an off-lattice Brownian Dynamics (BD) approach (green dots) on reproducing the diffusion parameters $\boldsymbol{\alpha}$ and $\boldsymbol{D^*}$ using obstacles three times bigger than the enzyme (r_{Obs} = 6.99 nm).

Monte Carlo simulations that include the mobility reduction factor provided D^* values that are in better agreement with the experimental results. This mobility reduction could be treated as a consequence of hydrodynamic interactions (HI), an effect that was not considered in the BD simulations shown in figures 15 and 16. Nevertheless, the comparison of obstacle size between the *in vivo*-like and computational experiments is not so feasible, since Dextrans are extremely branched polysaccharides and, therefore, approximating its shape into a sphere is not accurate enough. Because of that, in the following section we expose studied that allow to model Dextrans in BD simulations.

4.1.2.1 *Hydrodynamic interactions effect on diffusion*

A side master thesis from our workgroup is currently working in different approaches on implementing these HI on our BD code [69]. In there it is studied how to model Dextrans in BD.

Dextrans are polymers usually used in crowding studies because of their properties (non-reactive, random-coiled…). These species are extremely branched polysaccharides that, when found in big concentrations, are able to compact themselves. Because of this behaviour, their modelling in BD becomes a bit complex. In dilute conditions, their hydrodynamic radius r_h describes its shape accurately, and hence Dextrans could be modelled as hard spheres using r_h. On the other hand, in concentrate (crowded) conditions, the use of a smaller radius becomes more suitable to model its compact shape.

To find a more precise way for modelling Dextrans, experimental data is used to find a power law that relates the molecular weight of a Dextran with its radius:

$$r = K M_w{}^{\alpha} \tag{38}$$

where K and α are the constants of the power law. Different models (and thus different constants) are studied in order to find the most accurate polymer's radius. The effective radius model is chosen to be the one that represents the most Dextrans' radius in solution ($K = 0.045$ nm Da$^{-\alpha}$ and $\alpha = 0.387$). This model provides an effective radius compressed between r_h and the compact radius, representing an intermediate description of their shape. Detailed information about this study can be found in reference 69.

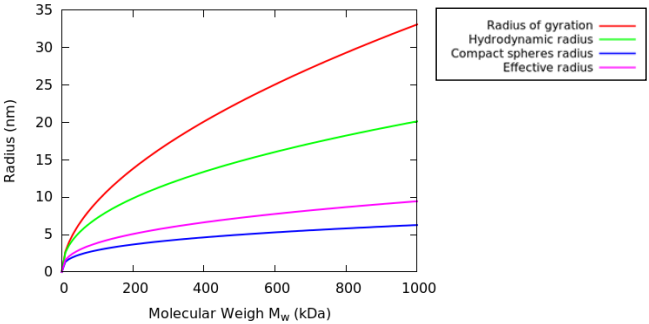

Figure 17. Relationship between the Dextran radius and its molecular weight according to the power law shown in equation 38, using different models proposed in the literature. For more information see ref. 69.

Diffusion coefficients at long simulation times D_∞ computed by means of different HI models are compared with the FRAP experimental values. Note that, due to the Dextran modelling studies, the obstacle sizes used in this current simulations are not the same as the ones used in the previous section. In figure 18 we present a comparison of relative values of D_∞ (in the plots labelled as D_{long}) computed by means of the Tokuyama model and the Rotne-Prager-Yamakawa diffusion tensor (RPY approximation) for BD algorithms [70, 71] for an α-chymotrypsin-like

diffusing system. The RPY diffusion tensor is a far-field approximation that updates the diffusion coefficient during the simulation by means of a 3N x 3N symmetric matrix.

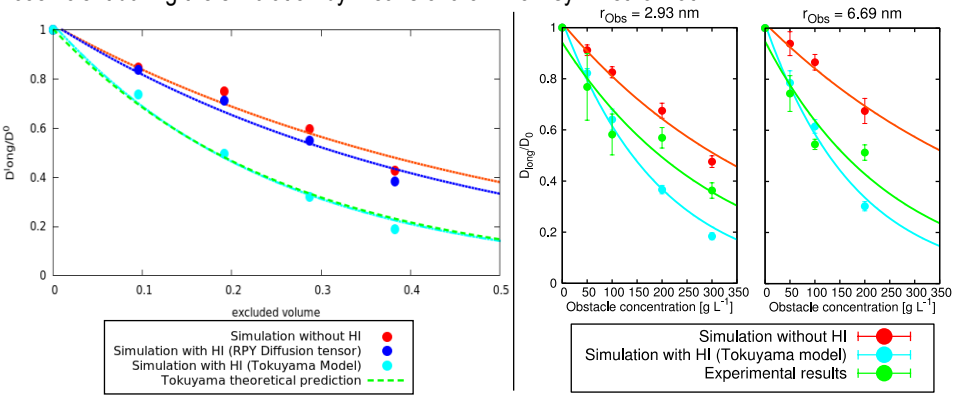

Figure 18. Relative diffusion coefficients at long simulation times D_{long} or D_∞ against excluded volume Φ for a homogeneous suspension of equal-sized spheres (α-chymotrypsin-like system). On the left, simulations are computed using three hydrodynamic interaction treatments. Red dots depict the diffusive behaviour against excluded volume for simulations that do not take into account the HI effect. Blue dots are for simulations that include the HI effect: marine blue dots are for simulations that included the RPY diffusion tensor; sky blue dots are for the ones computed according to the Tokuyama model. The green dotted line shows the theoretical prediction of D_∞ according to the Tokuyama model (equation 28.4). On the right, D_∞ is compared with two sizes of crowding agents for simulations using HI (sky blue dots) and without considering them (red dots), and are compared to FRAP experimental values (green dots).

Results reveal that diffusion becomes affected when HI interactions are considered, but the effect obtained by means of the RPY tensor is not as notable as the one caused by the Tokuyama model. When comparing the Tokuyama diffusion coefficients with FRAP experimental data it is legitimate to claim that using the Tokuyama model is more worth than implementing the RPY tensor in the code, since it recalculates the diffusion coefficient at almost every iterative step, scaling the computational cost greatly for almost no improvement at all on diffusion simulations. The RPY tensor calculates long-range HI accurately, but short-range HI are not well described. This becomes a problem for crowded environments, since long-range HI are not responsible for the decrease on D_∞. The Tokuyama model, on the other hand, produces diffusion coefficients that are in quite good agreement with the experimental results, since it provides a good description of the short-range HI. The differences with respect to the *in vivo-like* experiments are due to the fact that the Tokuyama model is deduced for equal-sized spheres. This implies that polydisperse systems such as crowded media are not well described, so a more rigorous short HI method is needed for crowded systems. Nonetheless, the Tokuyama model provides a good compromise between accuracy and computational cost (which implies

no extra computational cost), so the effect of HI on reactivity will be studied using the Tokuyama model.

4.2 REACTIVITY

In order to validate reactivity, a set different values of k_1 are selected to validate and study reactivity in diffusion control. Our criterion is to choose microscopic constants that accomplish $k_{-1} = 0.02k_1$ and $k_2 = 0.04k_1$. This relation has been chosen in previous studies [41]. The range selected for macroscopic k_1 covers values from 10^9 M^{-1} s^{-1} to 10^5 M^{-1} s^{-1}, according to the typical enzyme limits, seen in section 1.3. From now on, the label for each experiment is labelled according to the macroscopic k_1 constant in the I.S. units, i.e. 1E9 for experiments with $k_1 = 10^9$ M^{-1} s^{-1}, and analogously for the rest of the cases. The 1E5 kinetic would be the closest one to *in-vivo* experiments, where $k_2 \sim 100$ s^{-1}, and 1E9 would simulate the case of the fastest – and less usual – catalytic enzymes known.

4.2.1. Set of initial conditions

The k_1 scan study is performed using a range of concentrations that intends to embrace the usual experimental concentrations - which, of course, involve cellular cytosol-like concentrations, where the usual enzyme concentration oscillates around 10 μM. Another important issue is the substrate:enzyme (S:E) ratio. The usual concentrations used to study enzyme kinetics in the Michaelis-Menten scheme cover ratios from 5:1 to 70:1, being 50:1 the most usual relation.

However, the results that will be exposed in this thesis embrace enzyme concentrations of [E] = 133 μM, ten times bigger than biological conditions. Simulations using [E] = 10 μM were performed but due to a lack of compromise between accuracy and computational cost (when studying systems with crowding agents) we do not expose the results in the present work. In these experiments, a total of 3 enzyme particles were used to simulate the system, so simulations could be run with a very small computational cost. Obviously, this number of particles is not big enough to provide accurate results. To increase the number of enzyme particles, one has to increase the size of the simulation box, and this implies to increase the number of crowder particles in the system. Then, this enlargement of the simulation box yields a huge scaling on the computational time required, since the number of particles present in the system is also big.

Thereby, the enzyme concentration used is 133 μM, using a simulation box of 50 x 50 x 50 nm^3 and we use three different S:E ratios: 20:1, 50:1 and 70:1. These experiments are performed in dilute and crowded conditions, using four different values of excluded volume: 10, 20, 30 and 40%.

The physicochemical parameters used in the experiments are summarized in table 2, and table 3 indicates the used reaction constants in detail –both macroscopic and microscopic–, along with their respective reaction probabilities. Simulations that have the largest k_1 constants (1E9 and 1E8) have been run with ~10^5 production steps; the rest of simulations have been run with 10^6 production steps, since it is less likely for them to provide reactive events. All the simulations have been run with 100 equilibration steps at 298.15 K.

Specie	**Radius** [nm]	**Diffusion constant**(a) [nm^2/ns]
Enzyme	2.33	0.1053
Substrate	0.5	0.4360
Obstacle	4.0	0.0311

(a) Calculated from the Stokes-Einstein equation (equation 2).

Table 2. Physicochemical properties of the particles in the simulation.

Label	k_1^{macro} [M^{-1} s^{-1}]	k_1^{macro} [nm^3 ns^{-1}]	k_1^{micro}(a) [ns^{-1}]	k_{-1}(b) [ns^{-1}]	k_2(b) [ns^{-1}]	p_1(c)	p_{-1}(c)	p_2(c)
1E9	10^9	1.66	1.95×10^{-2}	3.9×10^{-4}	7.8×10^{-4}	1.95×10^{-3}	3.9×10^{-5}	7.8×10^{-5}
1E8	10^8	1.66×10^{-1}	1.77×10^{-3}	3.5×10^{-5}	7.1×10^{-5}	1.77×10^{-4}	3.5×10^{-6}	7.1×10^{-6}
1E7	10^7	1.66×10^{-2}	1.75×10^{-4}	3.5×10^{-6}	7.0×10^{-6}	1.75×10^{-5}	3.5×10^{-7}	7.0×10^{-7}
1E6	10^6	$1.66\cdot10^{-3}$	$1.75\cdot10^{-5}$	$3.5\cdot10^{-7}$	$7.0\cdot10^{-7}$	$1.75\cdot10^{-6}$	$3.5\cdot10^{-8}$	$7.0\cdot10^{-8}$
1E5	10^5	$1.66\cdot10^{-4}$	$1.75\cdot10^{-6}$	$3.5\cdot10^{-8}$	$7.0\cdot10^{-8}$	$1.75\cdot10^{-7}$	$3.5\cdot10^{-9}$	$7.0\cdot10^{-9}$

(a) Relationship between macroscopic and microscopic reaction constants is shown in equation 34.

(b) The unimolecular macroscopic rate constant equals its microscopic form, see section 3.5.

(c) Poissonian probabilities calculated according to equation 35.

Table 3. Reaction rate constants along with their respective probabilities for the association (k_1 or p_1), the inverse k_{-1}or p_{-1}) and the catalytic (k_2 or p_2) reactions used in the reactivity validation scan.

4.2.2 Reactivity validation

To validate our reaction-kinetics code, we analyse the particle variations with respect to simulation time and compare them with their respective numerical solution of the coupled ODEs,

expressed in equations 24.1 and 24.2. These ODEs suppose that the reaction container if found in dilute conditions and that it is well-stirred during all the reaction time.

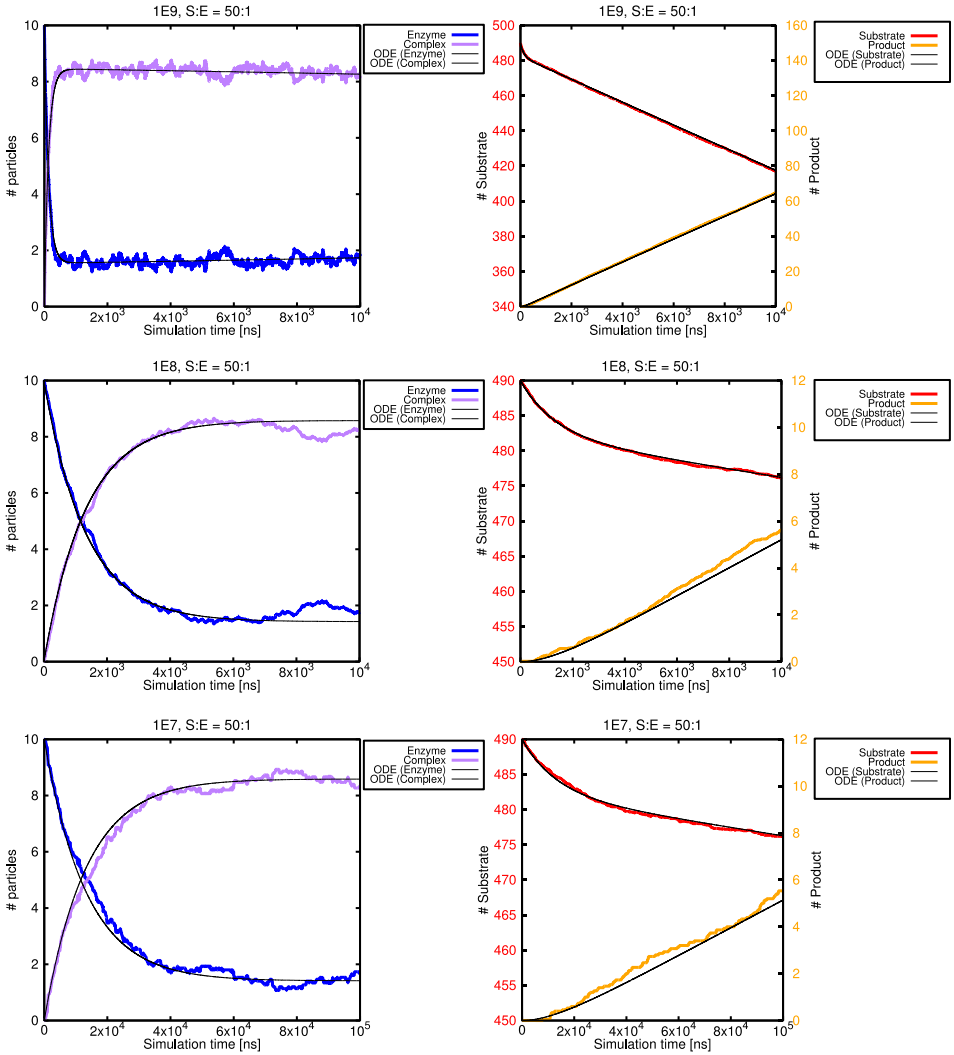

Figure 19. Temporal evolution of particle populations for experiments with S:E = 50:1 and concentrate enzyme and substrate conditions. The plots on the left column depict the enzyme (blue) and complex (purple) curves; plots on the right depict the substrate (red) and product (orange) curves. The black lines show the numerical solution of the coupled ordinary differential equations for each simulation.

In figure 19 we show a representative example of these comparisons for experiments that used a substrate-enzyme ratio S:E = 50:1. The plots for 1E5 and 1E6 experiments are not shown since hardly any reactive event occurred in 10^6 ns of simulation time, and as a

consequence the rate constants from this experiment could not be estimated (since they were almost equal to 0).

The specie variation curves without non-reacting obstacles simulated with our code perfectly fit the ODE solutions; thus, our stochastic RK3D code reaches the deterministic limit when a considerable number of repetitions of the same simulation are performed. Our code follows the law of mass action in systems that do not have hindering particles.

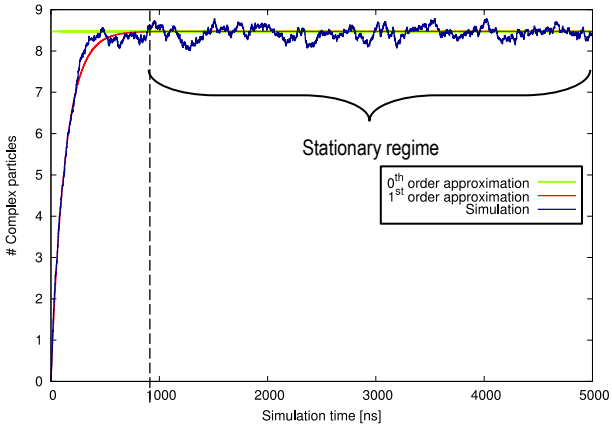

Figure 20. Complex particle population evolution with time of a simulation using k_1=1E9, [E] = 133 μM and S:E = 50:1 (blue line) compared with the complex variation according to the 0th (lime line) and 1st (red line) order approximations described in equations 21 and 23, respectively. The stationary state is reached after ~1,000 ns of simulation time.

The steady state approximations (both order 0th and 1st) to complex concentration variation with time are used as well to validate our calculations. In figure 20, experiment 1E9 S:E = 50:1 is compared with both approximations. Our code provides data that also accomplishes the approximations, even the initial production of complex described by the 1st order approximation, and reaching the asymptotic complex concentration. We take this occasion to remark that, in this particular case, the stationary regime is not reached after ~1,000 ns of simulation, when the complex concentration does not change with time anymore and remains constant. In that sense, the information found before the stationary regime should not be used for the calculation of the rate constants, since they are defined as *equilibrium* constants (where complex concentration remains constant).

Hence, the initial 1,000 ns correspond to a rapid reaction taking place before the achievement of the stationary profile of complex concentration, favoured by a homogeneous distribution of the reactants due to the random arrangement of the molecules.

4.2.3 Constant rate analysis methodology studies

For each simulation, our code provides an estimation of the three reaction constants according the Michaelis model. They are calculated at each iterative step by counting the number of reactive events (equations 37.1–3). The analysis of these constants show that the unimolecular rate constants k_{-1} and k_2 always provide an oscillating average value around the introduced parameter, even with experiments with excluded volume (see figure 21 for a representative case). This behaviour is expected, since these constants represent the mononuclear reactions, which do not depend on diffusion: a complex particle disproportionate into whichever substrate or product and enzyme with a given probability, yet it does not need to collide with any other chemical specie. Therefore, they are not crowding-dependent – this also implies that they are not affected by autocrowding. From this point, no plots regarding k_{-1} nor k_2 are shown, as they remain constant throughout the experiments, even in crowded environments, when calculated according the number of reactive events. $k_1$calculation using this methodology, however, becomes more troublesome. Due to some lack of consistency between experiments using the "counting" method (equations 37.1–3), a different methodology was tested in order to discern which one was the most accurate average for the bimolecular constant. In the following paragraphs, a discussion between both methods is done.

The "*counting* method" consists in, analogously to the mononuclear constants case, counting the number of reactive collisions throughout the simulation. Even though this method seems trivial and straightforward, the simulation time and the number of repetitions of each experiment may not seem the adequate to obtain enough data to provide an accurate and precise result. Each replica's data was analysed individually by means of the block average method (data was a bit correlated), and then averaged for all the replicas of the experiment. Due to lack of time and computational capacity and availability, the number of replicas of a given simulation has not been sufficient. This implies a small number of reactive collisions in the simulations, which leads to a small population of instantaneous $k_{1,i}$ data, which result sufficient to obtain nor accuracy neither precision in the statistics, and this leads to average values deviated from the expected value.

The alternative path to analyse the bimolecular constant is to fit each experiment's substrate population curve into an ODE by means of the weighted least squares fitting method. $k_{-1}$and k_2 values are kept fixed in the fitting, since they should not change whichever the medium conditions are. We expect the fitted k_1 values to be more reliable, since the number of particles

in the system is bigger than the number of reactive collisions throughout the simulation (more data population). The errors obtained for the fitting are very small compared to the fitted k_1 value, i.e. four orders of magnitude smaller. This error does not reflect the error due to lack of sample repetitions, but the fluctuation error on the fitting.

Figure 21 compares the relative reaction constants calculated according equations 37.1–3 (including the unimolecular constants) with the k_1 parameter obtained with the weighted least-squares fitting. Both k_1 values are quite similar in some cases, but in general it is not clear which value should be the most accurate.

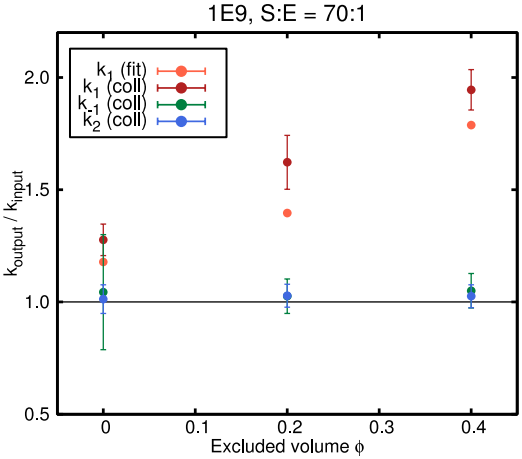

Figure 21. Relative reaction rate constants in three different environments of excluded volume for a system with E (r_E = 2.33 nm), S (r_S = 0.5 nm) and Obs (r_{Obs} = 4.0 nm). The unimolecular constants k_{-1} (green dots) and k_2 (blue dots) do not change as the presence of crowding agents in the medium increases, since they are not diffusion-dependent. On the contrary, the bimolecular rate constant (red and orange dots) does change with the excluded volume. Red dots represent k_1 values obtained from the "reactive collisions counter" method (k_1(coll)) and orange dots represent k_1 values obtained by the weighted least-squares fitting into the substrate ODE curve

To discern which the most precise methodology is, the coupled ODEs were solved for some systems with excluded volume using the *fitted* bimolecular constant (for the constant calculated with the weighted least squares fitting method) and the *collision* constant (for the constant calculated according to the number of reactive collisions). Figure 22 compares these ODEs, which we name ODE(fit) for the ODE solved with the fitted k_1 and ODE(coll) for the collision k_1 (using the calculated unimolecular constants) , with the simulated particle curves of two systems with different values of excluded volume and the theoretical ODE (dilute conditions).

It can clearly be seen that the most suitable curve i.e. the one that better fits the simulation results is the ODE(fit). Thereby, the bimolecular reaction constants k_1 are analysed according to the k_1 fitting into the substrate ODE of each case. The values that will be shown henceforth are calculated, then, by means of the weighted least square fitting method.

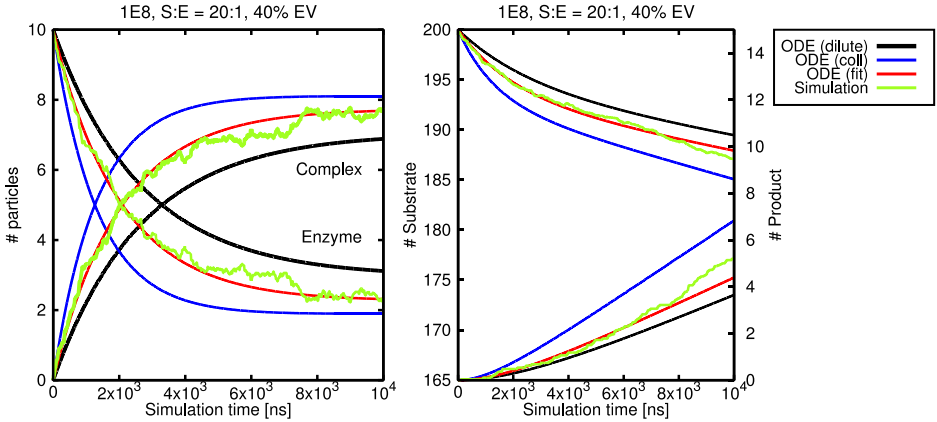

Figure 22. Comparison between the three ODE curves for the four varying species present in the simulations with a representative case. The simulated curves (green line) depict the temporal evolution of each specie in the presence of crowing agents. This line never follows the ODE(dil) curve that descrives the particle evolution for the species in dilute conditions (black line), and neither the ODE(coll) curve calculated according the averaged constants provided by the program (blue line). Instead, the curves obtained from the simulations do follow the ODE(fit) curves that were solved by means of the fitted k_1 (red line).

4.2.4 Discussion on dilute conditions

As stated in the reactivity validation section, neither 1E5 nor 1E6 kinetic experiments produce hardly any product particle in any of the simulations. The reason relies in that their reaction probabilities are extremely small, and thereby the reaction time simulated is not long enough to let any reactive event to take place. To obtain a quantifiable number of reactions, the simulations should reach the second, *s,* units, which would imply simulations with 10^3 or 10^4 iterations more than to the ones shown in this thesis. Doing that would require a large amount of computational resources.

These two former specific cases pretended to simulate the closest enzyme kinetics to what is found in intracellular media. Nevertheless, usual enzymes have a catalytic efficiency that is found below the diffusion limit; their rates are so small that their kinetics are limited by the

enzyme's activity (reactivity control), and not by diffusion. That is, these systems depict a well-mixed, homogeneous system in which diffusion is faster than reactivity itself.

1E7, 1E8 and 1E9 cases depict systems that are controlled by diffusion or they might also be controlled by both diffusion and enzyme activation (see figure 18). In any case, the substrate concentration (or number of particles) does not completely vanish within the simulation time range studied, not even with the 20:1 systems (picture not shown), where the substrate concentration is lower; nevertheless, as expected, the substrate concentration decreases with a slower rate as its governing constant, the bimolecular rate constant k_1, decreases. Also notice that it requires more simulation time to the complex to reach the steady state as k_1 decreases.

At this point, we have concluded that the 1E5 and 1E6 kinetic cases are no longer diffusion-controlled systems, but reaction (or activation)–controlled; diffusion is no longer the limiting step of the reaction. We also consider that the 1E7 and 1E8 cases are found in a mixed control by both diffusion and activation. As seen in the introductory section, the most abundant enzymes are found below the diffusion limit; their rates are so small that their kinetics are controlled by the enzyme's activity (reactivity control), and not by diffusion.

Figure 23 shows the three constants variation with time calculated using the reactive event counting method for the 1E9 experiment using a substrate:enzyme relation of 50:1 (analogous behaviour is obtained for the other cases). The presence of fluctuations is due to the small probability values of the system: for the bimolecular case, since not all the collisions are reactive, the instantaneous constant equals to 0 in almost all the iterative steps. This profile can be observed clearer when analysing the mononuclear constants, since their probabilities are 0.02 or 0.04 times smaller than the bimolecular constant. None of the reaction constants show any time dependency, which implies no fractal kinetics in the systems proposed. The experiments 1E9 and 1E8 have diffusion-limited kinetics, since their reaction constants are very big. However, their respective reaction probabilities are far from 1 (recall that the cases where fractal kinetics are observed is when the reaction probabilities are close to the unit, $p_1 \sim 1$). For instance, for experiment 1E9, i.e. the one that has the fastest kinetics, the bimolecular rate constant $k_1 = 1.6603$ nm^3 ns^{-1} implies a Poissonian probability of $p_1 = 1.95 \cdot 10^{-3}$, which is 100 times smaller than the unit. Thereby, even if the reaction is diffusion-controlled, the system will not show fractal-like kinetics unless their reaction probabilities are extremely fast. Since the rest of the experiments (from 1E8 to 1E5) suppose no larger k_1 values, none of our systems present a time-dependent k_1.

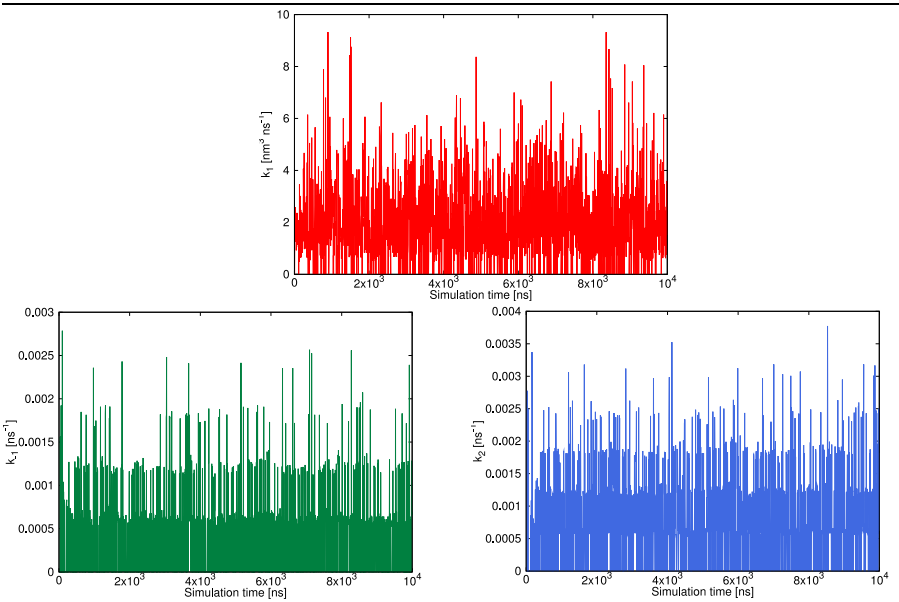

Figure 23. Average variation with respect to simulation time of the macroscopic reaction rate constants k_1 (red), k_{-1} (green) and k_2 (blue) for the c1E9 S:E = 50:1 experiments. Average values are calculated once the steady state is reached, which in this case is after 700 ns.

The relative average bimolecular rate constants k_1 in dilute conditions (Φ = 0) present oscillating values around theoretical constant throughout all the experiments (figure 24), suggesting that the performance of our code is acceptable since it provides precise and quite accurate constant values.

Bigger deviations from the input constant are detected for the 1E7 kinetic case for small S:E ratios. In acutely dilute conditions (S:E = 20:1 and 50:1), the activation part of the reaction control plays an important role, since the number of effective collisions between reactants is small regardless the number of encounters may be big. Instead, when the number of particles present in the system is sufficiently large (S:E = 70:1), the reaction rate related to k_1 increases, reaching the same relative value as the other kinetic cases. In this case, the number of collisions between both reactive species becomes more probable, which in turn increases the number of reactive collisions. This leaves the activation-control part of the reaction in the background. We conclude, then, that the 1E7 reaction is the only one that has an important mixed-control, whereas 1E8 and 1E9 kinetics are principally diffusion-controlled.

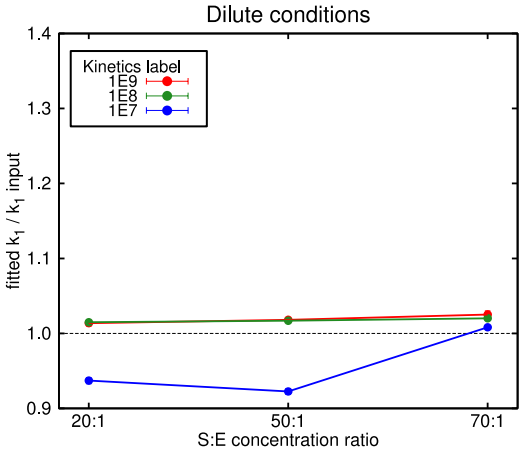

Figure 24. Relative bimolecular reaction rate constants for three substrate:enzyme concentration ratios when there is no presence of non-reacting obstacles or crowding agents (ϕ = 0%), using three kinetic systems: 1E9 (red dots), 1E8 (green dots) and 1E7 (blue dots).

The Michaelis constant K_M is calculated using the two definitions, according to its definition (equation 19.2) and considering the dissociation constant approximation K_d (equation 20), described in section 1.3. For the first definition, we use the value of the rate constants at each iterative step, i.e. we use the "instantaneous rate constants", which we will name as $K_{M,kt}$; for the second definition, we use the particle concentration at each point in time as well, labelled $K_{M,conc}$. In figure 25 we represent the variation of K_M with time calculated according the two definitions for the 1E9 S:E=50:1 case. We consider the K_M theoretical value the one resulting from the input reaction constants, calculated using equation 19.2. For instance, for the experiment shown in figure 25,

$$K_{M,theor} = \frac{\left(3.9 \times 10^{-4} + 7.8 \times 10^{-4}\right) \text{ns}^{-1}}{1.6603 \text{ nm}^3\text{ns}^{-1}} = 7.04 \times 10^{-4}\ nm^{-3}$$

The value of K_M is not stabilized until the steady state is reached, i.e. after ~500 ns of simulation. $K_{M,kt}$ presents a big amount of fluctuations due to the fluctuations obtained in the individual rate constants (see figure 23), but its average value is in perfect agreement with the theoretical value, as well as $K_{M,conc}$ after reaching equilibration. This behaviour is also obtained for the rest of the experiments (results not shown). In that sense, it is legitimate to claim that, in the conditions studied, the Michaelis constant *is* the equilibrium's dissociation constant because the system is in equilibrium:

$$K_\text{M} = \frac{k_2 + k_{-1}}{k_1} = \frac{[E][S]}{[C]} = K_\text{d}$$

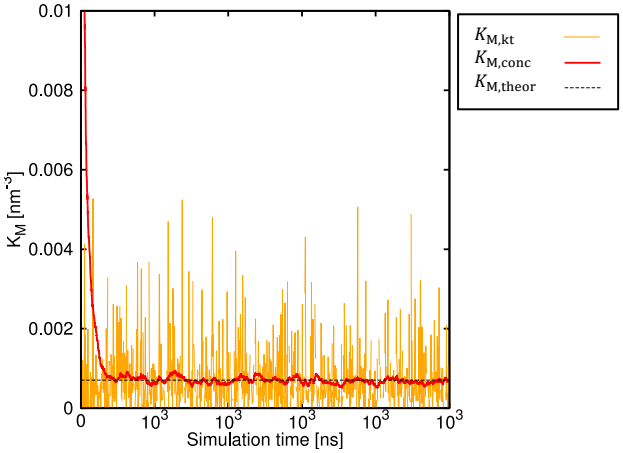

Figure 25. Michaelis constant K_M calculated according to its definition ($K_{M,kt}$, equation 19.2, orange line) and approximating it as the dissociation constant ($K_{M,conc}$, equation 20, red line), compared with the theoretical, input value (black dashed line) for 1E9, S:E=50:1 experiment.

4.2.5. Crowded conditions

The kinetic systems studied in dilute conditions are now simulated with different environments of macromolecular crowding. In this section, the size of the obstacles throughout all the experiments is of 4 nm (see table 1 on the previous section).

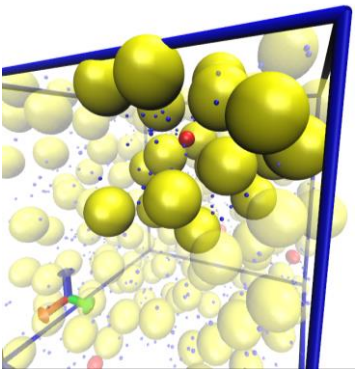

Figure 26. Representation of the first iterative step of a simulation with Φ = 40%, shot from a trajectory visualised with VMD (Visual Molecular Dynamics). The obstacles (yellow spheres) trap the substrates (blue spheres) and enzymes (red spheres) in small volumes or cages, which favours the encounters between reactants, consequently yielding an increase on the initial reaction rate.

In all the experiments we notice an increase of the reaction rate as the excluded volume increased, suggesting a confinement of the reactants in small effective volumes or reaction cages (see figure 26 for a representative case), which suppose a higher relative concentration of reactive species. As a consequence, the enzyme and substrate consumption becomes boosted as the presence of crowding agents rises (figure 27). A raise of the excluded volume yields a decrease on the confined volume. Undoubtedly, this implies an increase of the bimolecular rate constant with respect to the usual rate in dilute conditions, since the encounters between E and S become favoured.

That is, macromolecular crowding stimulates reaction because both E and S are confined and the volume they need to explore until finding other reactants is small. Needless to say, the unimolecular constants remain invariant (results not shown), as stated in previous sections.

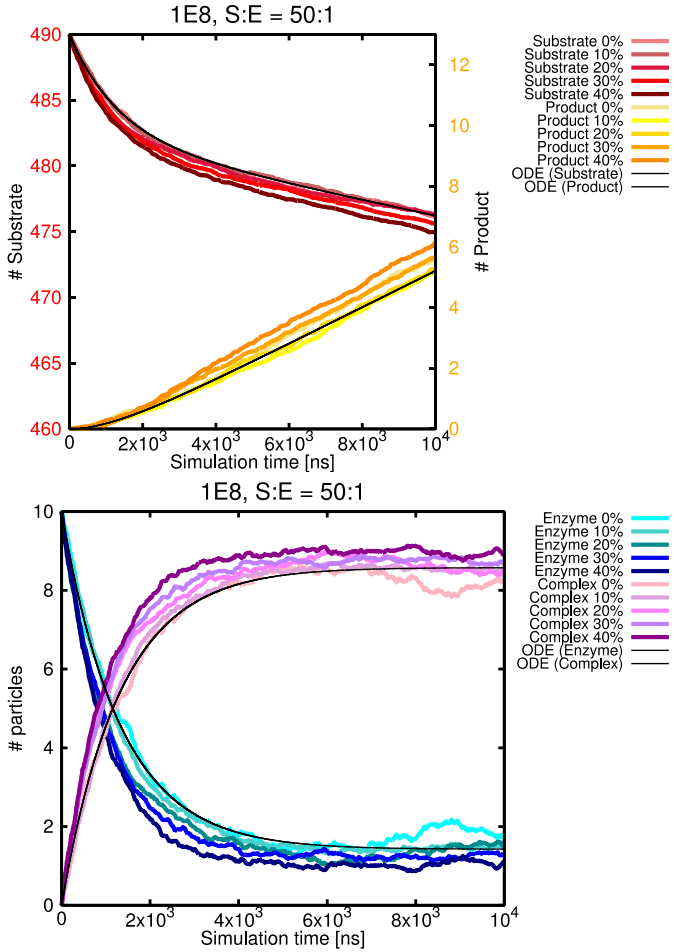

Figure 27. Representative example of the temporal evolution of substrate (red lines), enzyme (blue lines), complex (purple lines) and product (yellow lines) populations in presence of non-reacting macromolecules. The black lines show the numerical solution of the coupled ordinary differential equations (ODEs) for each simulation in dilute conditions.

Figure 28 collects all the bimolecular rate constants for all the excluded volume studies performed retrieved from the weighted least squares fitting of the substrate curves.

The first outcome one can notice is that the k_1 fittings obtained are not quite accurate. Even though the behaviour on k_1 always corresponds to a rising trend along with the increase of

excluded volume, the trend obtained is not smooth. The effect of S:E is not congruent, since in some values of Φ a given S:E relation experiences a high increase on its k_1, and in the consecutive Φ point analysed the behaviour is reversed (see 1E9 with S:E 70:1 and 50:1 for an example).

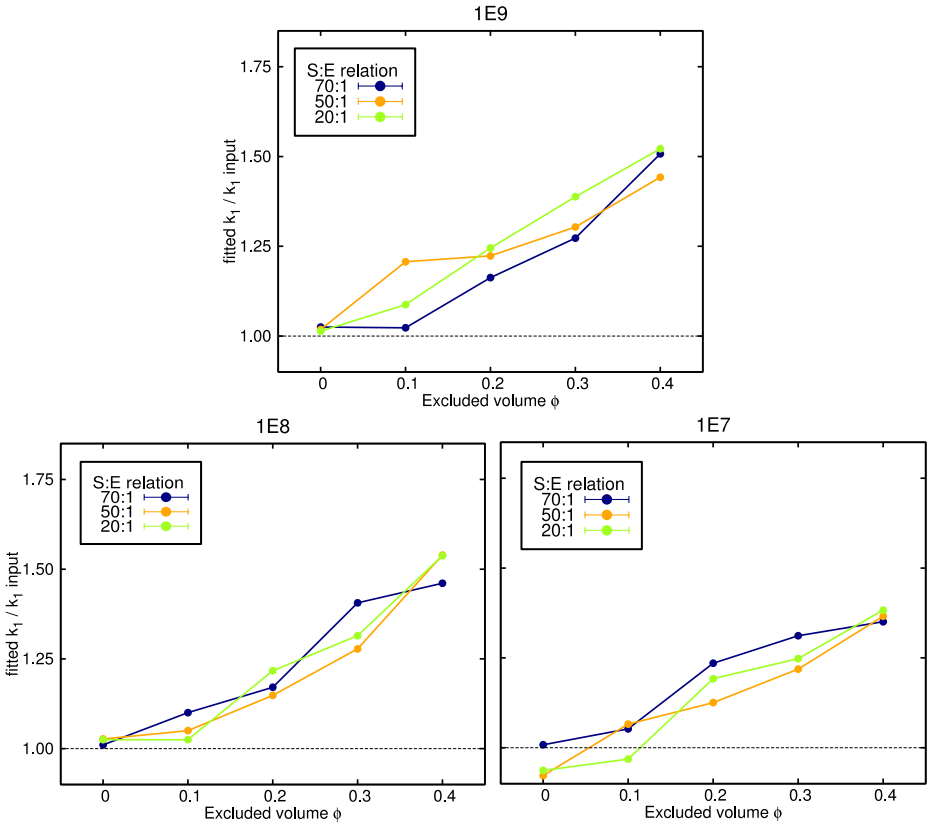

Figure 28. Relative bimolecular reaction rate constant k_1 in different crowded environments using three different substrate:enzyme concentration (S:E) ratios: 20:1 (green dots), 50:1 (orange dots) and 70:1 (blue dots) in three different kinetic systems: 1E9 (above), 1E8 (bottom left) and 1E7 (bottom right).

We see that simulations with S:E = 70:1 experience a gentler increase on their reaction rate in comparison with the other S:E relations, which is rougher. In these situations, we assume that the system experiences a mild autocrowding, so the reaction medium is not only confined by the obstacles but also crowded by the substrate species. Nevertheless, this latter effect is not as abrupt as confinement. In contrast, systems with S:E = 20:1 always experience the highest increase on their k_1 when Φ = 0.4, which, in comparison with the autocrowding experienced by

the denser systems, suggests that in these more dilute environments autocrowding is not that important. Nevertheless, this trend is only experienced for the most obstacle-concentrated situations. Further improved studies should be continued in that line to determine the paper of the autocrowding effect in these conditions.

Another important feature is that the magnitude on the rate increase is related with the kinetics of the reaction. For instance, the maximum increase on 1E7 system is $k_{1,simulation}/k_{1,input}$ ~ 1.3, whilst the 1E9 kinetic cases can easily reach increments of ~ 1.5. This observation suggests that the feature that controls the reactivity becomes an important factor when studying the macromolecular crowding effect. For instance, let us consider the kinetic systems 1E7 and 1E9. Reactivity on 1E9 experiments is clearly diffusion-controlled, whereas 1E7 cases are mixed-controlled. So, even though the macromolecular confinement is the same for both 1E7 and 1E9 kinetics, it is straightforward to see that 1E7 systems are not as affected by such confinement than the 1E9 cases. In that sense, the activation-control part of 1E7 kinetics is important, whilst 1E9 becomes utterly affected by such confinement.

All in all, an increasing change on the bimolecular constant is always observed when the excluded volume increases. Macromolecular confinement is the reason for this rate enhancement, and it occurs for any enzyme:substrate concentration relation and for any kinetic. In the Monte Carlo reaction simulation studies performed by our group [41], it is demonstrated that the radius of the confined volume r_{cv} defines the value of the k_1 parameter in crowded conditions:

$$k_1^{\Phi} = k_1^{enc}\left(\frac{r_{cv}}{r_{cv}-r_{ES}}\right) \tag{39}$$

where $r_{ES} = r_E + r_S$ is the distance of maximum approach between the reactant molecules for a reactive collision, k_1^{Φ} is the bimolecular constant for a given value of excluded volume, and k_1^{enc} is the Smoluchowski rate constant (equation 33), which would describe the constant in infinite dilute conditions.

However, the dependence of r_{cv} is not so easy to express in a simple equation. The obstacle size and the number of obstacles present in the medium (the excluded volume Φ) are important factors that define the size of the reaction cages. This volume also depends on the crowder distribution in the reaction box, and different treatments can be proposed according to their mobility. Nevertheless, even though the equation does not provide quantitative information, it provides a qualitative explanation why enzyme kinetics becomes enhanced due to small confined volumes.

The increasing trend on k_1 in enzyme kinetics that follow the Michaelis-Menten mechanism has also been reported previously by these on-lattice simulations [41]. There, we concluded that enzymes smaller than the crowding agents suffered an increase to the reactivity's velocity as the excluded volume increased. Nevertheless, we cannot perform a quantitative comparison, since the reaction probabilities chosen in those studies were chosen to study fractal kinetics (the on-lattice approach considered reaction probabilities of $p_1 = 1$, $p_{-1} = 0.02$ and $p_2 = 0.04$).

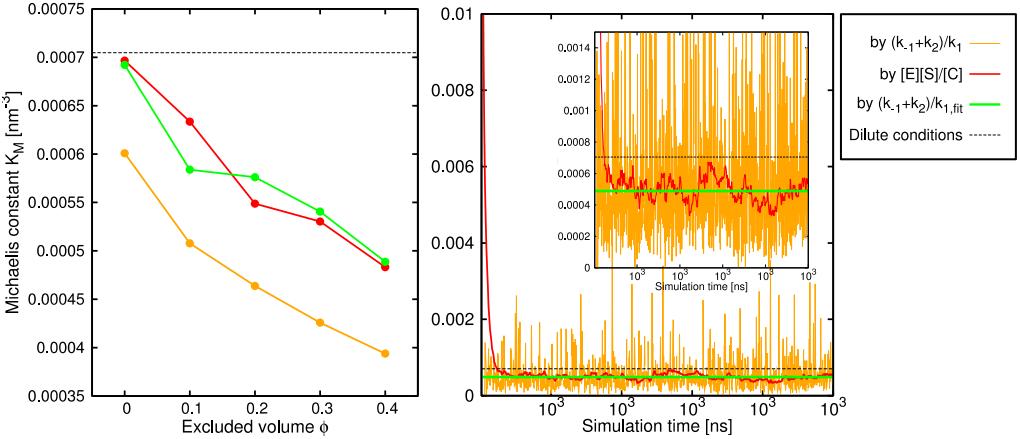

Figure 29. Experiments 1E9, S:E = 50:1. (Left) Michaelis constant K_{M} in different crowded environments Φ calculated according to its definition (equation 19.2), using $k_{1,fit}$ (green dots) and the "instantaneous" constants using the reaction counter (orange dots), and considering the steady state approximation (equation 20, red dots), compared to the input value, supposing dilute conditions (dashed black line). (Right) Michaelis constant K_{M} calculated according to its definition ($K_{M,kt}$, equation 19.2, orange line, using the reaction counter) and approximating it as the dissociation constant ($K_{\mathrm{M,conc}}$, equation 20, red line), compared with the input value (black dashed line) for Φ=0.4 experiment. Zoomed, the fitted k_1 value (green line) provides a Michaelis constant that is in agreement with the constants obtained by other calculation methods for Φ=0.4 experiment.

An increasing drift on k_1 supposes a decreasing trend on the Michaelis constant K_{M}. As in the previous section, we have calculated K_{M} using the same two definitions, and in both cases the constant suffers a decrease with respect to the theoretical value. In crowded situations, the Michaelis constant can still be considered the equilibrium's dissociation constant K_{d} (see left plot in figure 29). The small deviations obtained are due to the fitting error commented in the previous paragraphs; nevertheless we see in the right graph in figure 29 that the average values computed using both descriptions are in agreement between themselves. Zoomed, we show how the Michaelis constant calculated using the fitted $k_{1,fit}$ value is congruent with the constant values calculated according the other methods, $K_{M,kt}$ and $K_{M,conc}$.

In this case, though, no comparison with experimental data can be performed with the *in vivo-like* approach. α-chymotrypsin is an enzyme that may follow a more complicated mechanism rather than the MM one. The experimental studies performed in our group [72] suggest that the hydrolysis of N-succinyl-L-phenyl-Ala-p-nitroanilide catalysed by this small enzyme may experience inhibition by product, which implies that the catalytic reaction in the MM approach is no longer irreversible. This inhibition plays an important role in the kinetic description, since it has been reported that the reaction rate of the hydrolysis reaction decreases as the excluded volume increases, which is the inverse behaviour that we have seen in these experiments.

4.2.5.1 *Obstacle mobility effect*

In section 4.1.2 we discussed that the effect of fixed crowding agents in the simulation box provided smaller long-time diffusion coefficients D_∞ than simulations with mobile obstacles. This behaviour was detected in both on-lattice and off-lattice studies. Hence, the obstacle mobility effect on reactivity is studied using the 1E8 kinetic system with a substrate:enzyme ratio of 50:1. What we've seen is that, while obstacle mobility produced a quantitative difference in diffusion, the effect on reactivity is almost unnoticeable, leaving the excluded volume effect the main protagonist in the reaction rate alteration.

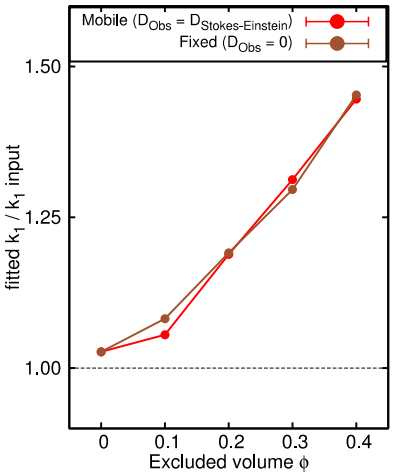

Figure 30. Relative bimolecular rate constant for the 1E8 S:E=50:1 experiment when obstacles are fixed (brown dots) and when their diffusion coefficient is the one obtained from the Stokes-Einstein equation (red dots).

4.2.5.2 *Obstacle size effect*

In the previous sections, we studied the effect of macromolecular crowding using obstacles with a fixed size (4.0 nm). As stated in the introductory section, the obstacle size causes a notable effect on particles' diffusion, so therefore the association reaction must feel in some way the effect of the obstacle size. Moreover, it has been previously reported in the literature that the k_1 dependence with excluded volume changes according to the obstacle size [15, 41]. For

instance, obstacles that are notably bigger than the enzyme itself cause an increasing trend of k_1 as the excluded volume increases (as we've seen in the previous sections). On the other hand, obstacles smaller or equal to the enzyme size tend to decrease the constant value when the excluded volume increases or may remain invariable, becoming non-dependent of the excluded volume effect [41, 73].

To study this effect, experiments using obstacle sizes of 1.0 nm, 2.33 nm (the same size as the enzyme) and 6.00 nm are run using the 1E8 kinetic case with a substrate-enzyme relation of 50:1.

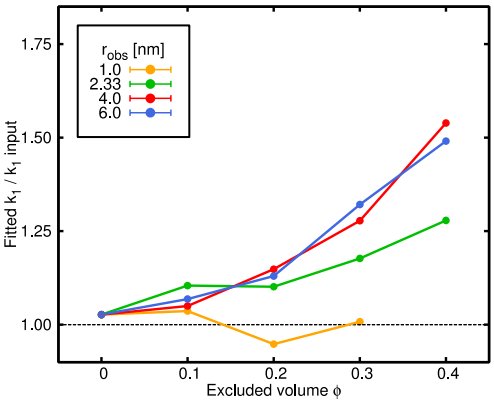

Figure 31. Relative bimolecular constant calculated using different sizes of crowding particles. The radii used are 1 nm (orange dots); the same radius as the enzyme, 2.33 nm (green dots); the obstacles used in the previous sections, 4 nm (red dots), and 6 nm (blue dots).

As expected form the literature, crowding agents smaller than the enzyme itself leave the reaction kinetics almost unaffected, sometimes provoking a decrease on the reaction constant (figure 31). The reason is probably due to that the probability of colliding with the small obstacles becomes lower due to their size, or it may also be caused due to a great proportion of obstacle clustering. In the latter case, the number of apparent obstacles decreases, giving a smaller reactive–cluster ratio and, therefore, decreasing the collision probability between both types of species.

On the contrary, obstacles of greater size (or equal) enhance the association reaction. The caging of the reactive species into small effective volumes becomes a more important issue as the size of the crowder increases. This leads to an increase of the reactive species' effective concentration that causes an increase to the number of effective collisions. Besides, an increase in the excluded volume implies a decrease in the confined volume, yielding an

increased bimolecular constant. Nevertheless, in our experiments we almost see no difference between obstacles of radii 4 and 6 nm. Finally, obstacles of equal size of the enzyme also enhance the reactivity, but in a lower magnitude. This trend is in agreement with equation 39, which relates the increase of k_1 with the size of the confined volume that, in turn, decreases with obstacle size.

One could discuss that the only important effect when the obstacle size is equal or greater to the size of the enzyme is confinement, whilst for smaller obstacles exists a competition between confinement (the reaction cages cause a greater number of reactive collisions, leading to an increase on k_1) with the diffusion hindrance (the obstructed medium makes reactive collisions less probable, leading to a decrease on k_1).

Our results are in good agreement with the Monte Carlo studies [41] performed by our group with excepting the case where the obstacle shares the same size as the enzyme (1x-size obstacles, r_{Obs} = 2.33 nm). For those cases Pitulice *et al.* obtained that k_1 decreases as the excluded volume increases, whilst in our simulations we obtained the inverse tendency. This discrepancy may emerge from the discretization of the lattice in the Monte Carlo approach, which may not reproduce the actual movement of particles, since they actually move freely in the three dimensions; and because such studies are performed using reaction probabilities higher than the ones considered in this project.

All things considered, throughout our studies we see that macromolecular confinement plays an important role. Since the model enzyme used for this study is relatively small, any obstacle that is equal or larger than the tracer generates reaction cages. The behaviour of the rate constant illustrates, then, that the enzyme-obstacle relative size is a factor that needs to be highlighted when studying enzyme kinetics.

4.2.5.3 *Hydrodynamic interactions effect on reactivity*

The effect of hydrodynamic interactions on reactivity is studied using the 1E8 kinetic system with S:E = 50:1 and the system that was the most affected with autocrowding, 1E9 with S:E = 70:1. Instead of using the Stokes-Einstein diffusion coefficient, we perform experiments using the short-time self-diffusion coefficient correction for each species in the system according to the Tokuyama model (equations 28.1–3).

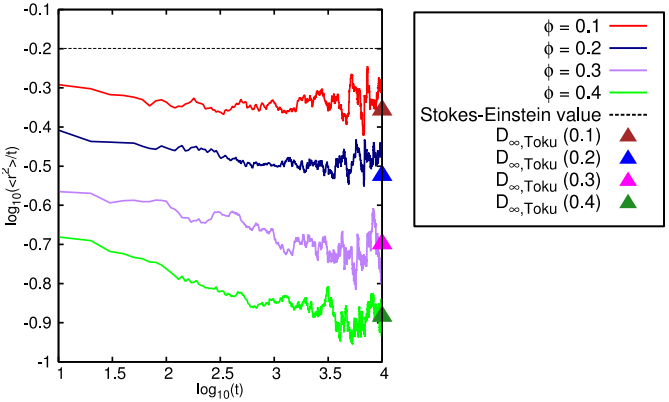

Figure 32. Log-log plot for the mean-squared displacement $\langle r^2 \rangle$ for the experiments performed using the Tokuyama model. Triangle values are the long-time diffusion coefficient D_∞ prediction values of the Tokuyama model (equation 28.4), which are in agreement with the experimental results obtained.

The first obvious consequence is the mean-squared displacement variation of the diffusing particles. According to the Tokuyama model, the introduced diffusion coefficient D^T is a correction with respect to the dilute conditions coefficient D_0, and therefore the initial diffusion coefficient for the particles in the system is lower than D_0. As a consequence, the MSD curves do not start from the same initial value, like the case studied in section 4.1.2. As figure 32 shows, the coefficient variation with respect to time becomes more pronounced – that is, the slope becomes steeper, so α decreases – as the excluded volume increases, seen as well in figure 14 in the Diffusion section. We have also seen that the Tokuyama prediction for long simulation times diffusion coefficient D_∞ is correct when studying reaction-diffusion processes (triangles in figure 32).

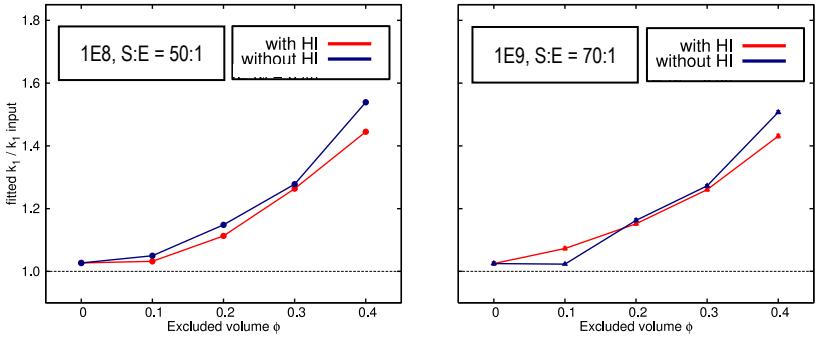

Figure 33. Relative k_1 values for the kinetic systems 1E8 with S:E = 50:1 (left, dots) and 1E9 with S:E = 70:1 (right, triangles) obtained with simulations that considered HI (red) and not (blue).

While the effect of including hydrodynamic interactions in the system seems quite important in diffusion, the effect on the temporal evolution of specie populations is not that notable. Equivalent specie population curves are obtained for simulations that include the description of HI and the ones that do not (figure 34). Nevertheless, the association reaction experiences a smooth decrease on its rate when HI are considered in the simulations, as we can see in figure 33. The magnitude of the interactions becomes more important as the number of crowding species present in the reaction environment increases, a fact that seems quite obvious: a high number of molecules in the reaction media imply a higher number of collisions with all the particles in the system by part of the solvent.

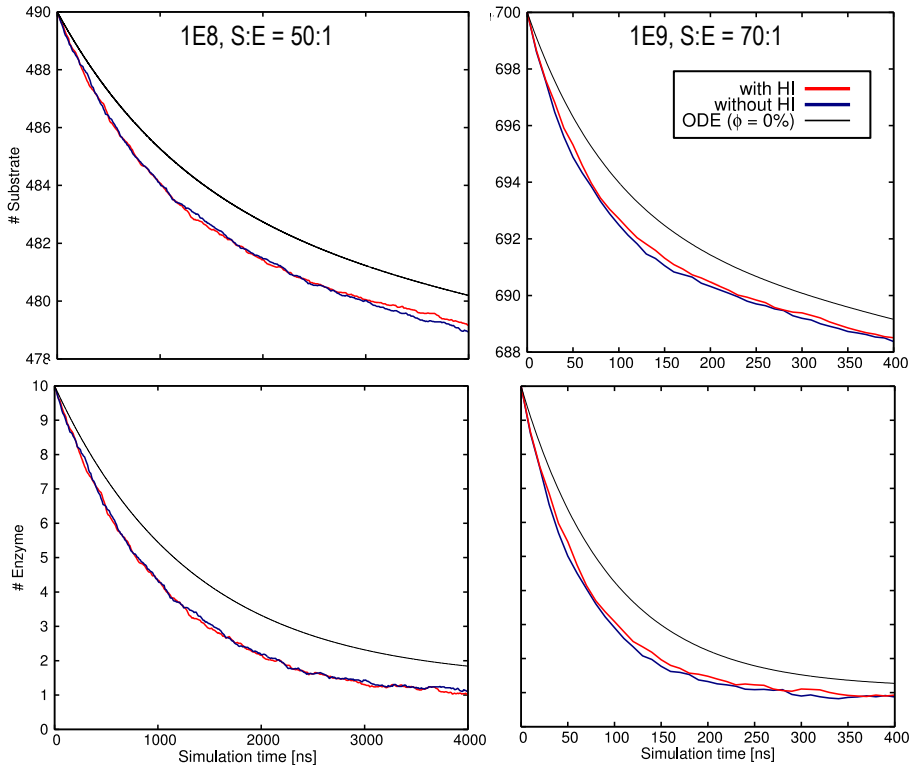

Figure 34. Number of substrate (above) and enzyme (below) particles variation with respect to simulation time in the kinetic systems 1E8, S:E = 50:1 (left) and 1E9, S:E = 70:1 (right) when HI are considered (red lines) and not (blue lines) for Φ = 0.4.

The substrate and enzyme variation curves under an excluded volume of Φ = 0.4 are represented in figure 34. Such particle variations show almost no difference, if any, with respect

to the simulations performed using the Stokes-Einstein coefficient (i.e. no HI considered), but a smooth decrease on reactant consumption when the HI are considered. Nevertheless, such mild deviations may be fruit of fluctuations throughout the simulations, and thus the reaction constant differences seen in figure 33 may be miscalculated and untrue.

All in all, the effect of HI on reactivity is not as meaningful as the excluded volume effect, yet its repercussion on reactivity may become emphasized as the number of species present in the system increases.

5. CONCLUSIONS

We have developed a Brownian Dynamics reaction-diffusion code that allows the simulation of enzymatic systems that follow the Michaelis-Menten mechanism in crowded environments.The efficiency and accuracy of the code has been tested in both diffusion and reaction simulations, and has demonstrated that its performance is satisfactory. Different methodologies have been analysed to determine in an accurate way the value of the bimolecular rate constant k_1 by means of the data provided by our code. We concluded that fitting the bimolecular constant into an ordinary differential equation (ODE) by means of the weighted least squares fitting method was the most appropriate of the strategies tested.

Regarding diffusion studies, we observed that diffusion throughout the simulation medium became disfavoured as the concentration of crowding agents increased, provoking a subdiffusive motion to the Brownian particles. The diffusion coefficient changes with time during a period of anomalous diffusion until it reaches a time-independent value, corresponding to a coefficient in a denser homogeneous medium. We characterized our simulations by analysing the anomalous diffusion exponent and the limiting diffusion constant five different environments of excluded volume, using four sizes of crowding agents. We observed that diffusion became more anomalous as the size of the crowding agents decreased, in a sense that the obstacle-enzyme relative size must be taken into consideration when studying diffusion in crowded environments.

Reactivity studies revealed that crowded systems with obstacles equal or bigger than the enzyme yield a rising on the bimolecular constant k_1 value as the excluded volume increases, implying the inverse trend on the Michaelis constant. The reason relies on the generation of small confined volumes that favours the enzyme-substrate encounters at the initial reaction times (cage effect). Besides, an increase on the obstacle size implies a confined volume diminishment, and consequently an increase on the reaction rate. The behaviour of the rate constant illustrates that the enzyme-obstacle relative size is a very important factor in enzymatic reactions. We also noticed that crowders smaller than the enzyme leave the reaction rate unaffected. We associate the reason to this later behaviour to the net effect resulting from a

competition between the cage effect, which produces an increase on the reaction rate, and obstacle hindrance to diffusion, that causes a decrease on the reaction rate. We also checked that, even though obstacle mobility has a considerable effect on diffusion, reactivity remains unaltered, providing similar results when simulations were run with both fixed and mobile crowding agents.

Finally, we attempted a first approach of introducing hydrodynamic interactions in the Brownian Dynamics algorithm by means of the Tokuyama model, a mean-field method that corrects the dilute diffusion coefficient. In reactivity studies, we detected that the hydrodynamic interactions effect is not as significant as the excluded volume effect, yet it becomes more important as the number of crowding species present in the system increases. However, we need to introduce more complex models to study such effects, since the Tokuyama model is deduced for equal-sized systems. In that sense, further reactivity studies ought to be performed to obtain a more accurate description of the short-range hydrodynamic interactions.

To improve the accuracy of our results, longer simulations ought to be performed. In that way, the bimolecular constant can be calculated according to the mass balance-derived expressions, implemented in the code, as the number of reactive events will be sufficiently big and, thus, more trustful. Fluctuations also suppose a problem, so more replicas of each experiment should be performed to reach the deterministic limit from our stochastic kinetics.

6. REFERENCES AND NOTES

1. Zimmerman, S.B., Trach, S. O., *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli.* J. Mol. Biol., 1991. **222**: p. 599-620.
2. Hall, D., Minton, A. P., *Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges.* Biochim. Biophys. Acta., 2003. **1649**(2): p. 127-139.
3. Fulton, A.B., *How crowded is the cytoplasm?* Cell, 1982. **30**: p. 345-347.
4. Record, T.M.J., Courtenay, E. S., Cayley, S., Guttman, H. J., *Biophysical compensation mechanisms buffering E. coli protein–nucleic acid interactions against changing environments.* Trend. Biochem. Sci, 1998. **23**: p. 190-194.
5. Ellis, R.J., *Macromolecular crowding, obvious but underappreciated.* Trend. Biochem. Sci., 2001. **26**(10): p. 597-604.
6. Minton, A.P., *The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media.* J. Biol. Chem., 2001. **276**: p. 10577-10580.
7. Minton, A.P., *Implications of macromolecular crowding for protein assembly.* Curr. Opin. Struct. Biol., 2000. **10**: p. 34-39.
8. Ralston, G.B., *The effect of crowding in protein solutions.* J. Chem. Educ., 1990. **67**: p. 857-860.
9. Ellis, R.J., Hartl, F. U., *Principles of protein folding in the cellular environment.* Curr. Opin. Struct. Biol., 1999. **10**: p. 34-39.
10. Tokuriki, N., Kinjo, M., Negi, S., Hoshino, M., Goto, Y., Urabe, I., Yomo, T., *Protein Folding by the Effects of Macromolecular Crowding.* Protein Sci., 2004. **13**(1): p. 125-133.
11. Christiansen, A., Wang, Q., Cheung, S., Wittung-Stafshede, P., *Effects of macromolecular crowding agents on protein folding in vitro and in silico.* Biophys. Rev., 2013. **5**(2): p. 137-145.
12. Zimmerman, S.B., *M acromolecular crowding effects on macromolecular interactions: Some implications for genome structure and function.* Biochim. Biophys. Acta, 1993. **1216**: p. 175-185.
13. Schnell, S., Turner, T. E., *Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws.* BProg. Biophys. Mol. Biol., 2004. **85**: p. 235-260.
14. Zimmerman, S.B., Minton, A. P., *Macromolecular crowding: biochemical, biophysical, and physiological consequences.* Annu. Rev. Biophys. Biomol. Struct., 1993. **22**: p. 27-25.
15. Mourão, M., Kreitman, D., Schnell, S., *Unravelling the impact of obstacles in diffusion and kinetics of an enzyme catalysed reaction.* Phys. Chem. Chem. Phys., 2014. **16**: p. 4492-4503.

16. Minton, A.P., *Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion.* Methods. Enzymol., 1998. **295**: p. 127-149.
17. Pastor, I., Vilaseca, E., Madurga, S., Garcés, J. L., Cascante, M., Mas, F. , *Diffusion of alpha-chymotrypsin in solution-crowded media. A fluorescence recovery after photobleaching study.* J. Phys. Chem. B, 2010. **114**(11): p. 4028-4034.
18. Pitulice, L., Pastor, I., Vilaseca, E., Madurga, S., Isvoran, A., Cascante, M., Mas, F. , *Influence of macromolecular crowding on the oxidation of ABTS by hydrogen peroxcide catalyzed by HRP.* J. Biocatal. Biotransform., 2013. **2**: p. 1.
19. Balcells, C., Pastor, I., Pitulice, L., Hernández, C., Via, M., Garcés, J. L., Madurga, S., Vilaseca, E., Isvoran, A., Cascante, M., Mas, F., *Macromolecular Crowding upon in vivo-like enzyme-kinetics: effect of enzyme-obstacle size ratio.* New Front. Chem., 2015. **24**(1): p. 3-16.
20. Pastor, I., Pitulice, L., Balcells, C., Vilaseca, E., Madurga, S., Isvoran, A., Mas, F., *Effect of crowding by Dextrans in enzymatic reactions.* Biophys. Chem., 2014. **185**: p. 8-13.
21. Laurent, T.C., *The interaction between polysaccharides and other molecules: the solubility of proteins in the presence of dextran.* Biochem. J., 1963. **89**: p. 253-257.
22. Minton, A.P., *Influence of excluded volume upon macromolecular structure and associations in crowded media.* Curr. Opin. Struct. Biol., 1997. **8**: p. 65-69.
23. Brown, R., *A brief account of microscopical observations made in the months of June, July, and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies.* Edinburgh New Philosophical Journal, 1828. **5**(371-385).
24. Einstein, A., *Über die von der molekularkinetischen Theorie dew Wärme gefordete Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen (Investigations on the Theory of the Brownian Movement).* Ann. Phys., 1905. **17**: p. 549-560.
25. Smoluchowski, M., *Essai d'une théorie cinétique du mouvement Brownien et des milieux troubles (Outline of the kinetic theory of Brownian motion of suspensions.* Bulletin International de l'Académie des Sciences de Cracovie, 1906: p. 577-602.
26. Perrin, J., *Mouvement brownien et réalité moléculaire.* Annales de Chimie et de Physique, 1909. **18**: p. 1-114.
27. Perrin, J., *Les Atomes*. 1913.
28. Ben-Avraham, D., Havlin, S., *Diffusion and reactions in fractals and disordered systems*. 2000.
29. Vilaseca, E., Isvoran, A., Madurga, S., Pastor, I., Garcés, 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., 2011. **13**: p. 7396-7407.
30. Saxton, M.J., *A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model.* Biophys. J., 2007. **92**: p. 1178-1191.
31. Gillespie, D.T., *The multivariate Langevin and Fokker-Planck equations.* Am. J. Phys., 1996. **64**: p. 1246-1257.
32. Hänggi, P., *Generalized Langevin Equations: A useful tool for the perplexed modeller of nonequilibrium fluctuations?* Stochastic Dynamics, Lecture Notes in Physics. Vol. 484. 1997.

33. Marmarelis, V.Z., *Nonlinear Dynamic Modeling of Physiological Systems (Appendix II)*, ed. W.O. Library. 2004.
34. Zwanzig, R., *Nonlinear generalized Langevin Equations.* J. Stat. Phys., 1973. **9**: p. 215-220.
35. Mori, H., *Transport, Collective Motion, and Brownian Motion.* Prog. Theo. Phys., 1965. **33**: p. 423-455.
36. Michaelis, L., Menten, M.L, *Die Kinetik der Invertinwirkung.* Biochem. Z., 1913. **49**: p. 333-369.
37. Gillespie, D.T., *Stochastic Simulation of Chemical Kinetics.* Annu. Rev. Phys. Chem., 2007. **58**: p. 35-55.
38. Berg, J.M., Tymoczko, J. L., Stryer, L., *Biochemistry, Section 8.4*. 5th Edition ed. 2002, New York.
39. Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., Milo, R., *The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters.* Biochemistry, 2011. **50**: p. 4402-4410.
40. Olsthoorn, A.J., Otsuki, T., Duine, J. A, *Negative cooperativity in the steady-state kinetics of sugar oxidation by soluble quinoprotein glucose dehydrogenase from Acinetobacter calcoaceticus.* Eur. J. Biochem., 1998. **255**(1): p. 255-261.
41. Pitulice, L., Vilaseca, E., Pastor, I., Madurga, S., Garcés, J. L., Isvoran, A., Mas, F., *Monte Carlo simulations of enzymatic reactions in crowded media. Effect of the enzyme-obstacle relative size.* Math. Biosci., 2014. **251**: p. 72-82.
42. Kopelman, R., *Fractal reaction kinetics.* Science, 1988. **241**: p. 1620-1626.
43. McAdams, H.H., Arkin, A. P., *Stochastic mechanisms in gene expression.* Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 814-819.
44. Erban, R., Chapman, S. J., Maini, P. K., *A practical guide to stochastic simulations of reaction-diffusion processes.* arXiv:0704.1908v2 [q-bio.SC], 2007.
45. Ermak, D.L., McCammon, J. A., *Brownian Dynamics with hydrodynamic interactions.* j. Chem. Phys., 1978. **69**: p. 1352-1360.
46. Nortrup, S., Erickson, H., *Kinetics of protein-protein association explained by Brownian Dynamics computer simulation.* Proc. Natl. Acad. Sci. USA, 1992. **89**: p. 3338-3342.
47. Morelli, M.J., ten Wolde, P. R., *Reaction Brownian Dynamics and the effect of spatial fluctuations on the gain of a push-pull network.* J. Chem. Phys., 2008. **129**(5): p. 054112.
48. Garcia de la Torre, J., Bloomfield, V. A., *Hydrodynamic Properties of Macromolecular Complexes. I. Translation.* Biopolymers, 1977. **16**: p. 1747-1763.
49. Zuk, P.J., Wajnryb, E., Mizerski, K. A., Szymczak, P., *Rotne-Prager-Yamakawa approximation for different-sized particles in application to macromolecular bead models.* J. Fluid Mech., 2014. **741**(R5): p. 1-13.
50. Tokuyama, M., Moriki, T., Kimura, Y., *Self-diffusion of biomolecules in solution.* Phys. Rev. E, 2011. **83**(051402): p. 1-8.
51. Sanford, C.Y., M. White, C., Parkinson, J., *Cell++ - Simulating biochemical pathways.* Bioinformatics, 2006. **22**: p. 2918-2925.
52. Andrews, S., Bray, D., *Stochastic simulation of chemical reactions with spatial resolution and single molecule detail.* Phys. Biol., 2004. **1**: p. 137-151.

53. Plimpton, S., Slepoy, A., *ChemCell: A particle-based model of protein chemistry and diffusion in microbial cells.* Sandia National Laboratories, 2003: p. 1-29.
54. Kerr, R.A., Bartol, T.M., Kaminsky, B., Dittrich, M., Chang, J. C. J., et al., *Fast Monte Carlo simulation methods for bioogical reaction-diffusion systems in solution and on surfaces.* SIAM Journal on Scientific Computing, 2008. **30**: p. 3126.
55. Ridgway, D., Broderick, G., Lopez-Campistrous, A., Ru'aini, M., Winter, P., Hamilton, M., Boulanger, P., Kovalenko, A., Ellison, M. J., *Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm.* Biophys. J., 2008. **94**(10): p. 3748-3759.
56. Schöneberg, J., Noé, F., *ReaDDy: A software for particle-based reaction-diffusion dynamics in crowded cellular environments.* PLOS One, 2013. **8**(9): p. 1-14.
57. Schöneberg, J., Ullrich, A., Noé, F., *Simulation tools for particle-based reaction-diffusion dynamics in continuou space.* BMC. Byophysics, 2014. **7**(11): p. 1-10.
58. Erban, R., Chapman, S. J., *Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions.* Phys. Biol., 2009. **6**: p. 046001.
59. Berezhkovskii, A.M., Szabo, A., *Theory of Crowding Effects on Bimolecular Reaction Rates.* J. Phys. Chem. B, 2016. **Article ASAP**.
60. Chen, J.C., Kim, A. S., *Brownian Dynamics, Molecular Dynamics, and Monte Carlo modeling of coloidal systems.* Adv. Colloid Interface Sci., 2004. **112**: p. 159-173.
61. Ando, T.a.S., J. , *Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion.* Proc. Natl. Acad. Sci. USA, 2010. **107**: p. 18457-18462.
62. Press, W.H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. 2007: New York: Cambridge University Press.
63. Box, G.E.P., Muller, M. E. , *A Note on the Generation of Random Normal Deviates.* Ann. Math. Statist., 1958. **29**(2): p. 610-611.
64. Smoluchowski, M., *Attempt for a mathematical theory of kinetic coagulation of colloid solutions.* Z. Physik. Chem., 1917. **92**: p. 129-168.
65. Rice, S.A., *Diffusion-Limited Reactions*. Vol. 25. 1985, New York: Elsevier.
66. Frazier, Z., Alber, F., *A Computational Approach to Increase Time Scales in Brownian Dynamics-Based Reaction-Diffusion Modeling.* J. Comp. Biol., 2012. **19**(6): p. 606-618.
67. VanKampen, N., *Stochastic Processes in Physics and Chemistry*. 3rd ed. 2003: North-Holland personal library.
68. Berry, H., *Monte carlo simulations of enzyme reactions in two dimensions: fractal kinetics and spatial segregation.* Biophys. J., 2002. **83**: p. 1981-1901.
69. Blanco, P.M., *Master Thesis: Brownian Dynamics simulation of macromolecule diffusion in intracellular media*, in *Màster en Modelització Computacional Atomística i Multiescala en Física, Química i Bioquímica*. 2016, Universitat de Barcelona.
70. Rotne, J., Prager, S., *Variational Treatment of Hydrodynamic Interaction in Polymers.* J. Chem. Phys., 1969. **50**: p. 4831.
71. Yamakawa, H., *Transport Properties of Polymer Chains in Dilute Solution: Hydrodynamic Interaction.* J. Chem. Phys., 1970. **53**(436).
72. Pastor, I., Vilaseca, E., Madurga, S., Garcés, J. L., Cascante, M., Mas, F., *Effect of Crowding by Dextrans on the Hydrolysis of N-Succinyl-l-phenyl-Ala-p-nitroanilide Catalyzed by α-Chymotrypsin.* J. Phys. Chem. B, 2011. **115**(5): p. 1115-1121.

73. L. Homchaudhuri, N.S.a.R.S., *Effect of crowding by Dextrans and Ficolls on the rate of alkaline phosphatase-catalyzed hydrolysis: a size-dependent investigation.* Blopolymers, 2006. **83**(5): p. 477-486.

7. APPENDICES

APPENDIX 1. BIMOLECULAR CONSTANTS OBTAINED BY WEIGHTED LEAST SQUARES FITTING

1E7	**20:1**	**50:1**	**70:1**
Dilute	0.016 ± 2E-4	0.015 ± 2E-5	0.017 ± 6E-5
0.1	0.016 ± 2E-5	0.018 ± 3E-5	0.017 ± 5E-5
0.2	0.020 ± 2E-4	0.019 ± 9E-5	0.021 ± 2E-5
0.3	0.021 ± 6E-5	0.020 ± 5E-5	0.022 ± 4E-5
0.4	0.023 ± 1E-4	0.023 ± 9E-5	0.022 ± 9E-5
1E8	**20:1**	**50:1**	**70:1**
Dilute	0.17 ± 5E-4	0.17 ± 6E-5	0.17 ± 1.2E-4
0.1	0.17 ± 6E-4	0.18 ± 9E-4	0.18 ± 2E-3
0.2	0.20 ± 9E-4	0.20 ± 2E-3	0.19 ± 9E-4
0.3	0.22 ± 9E-4	0.22 ± 6E-4	0.23 ± 4E-3
0.4	0.26 ± 2E-3	0.24 ± 2E-3	0.24 ± 4E-3
1E9	**20:1**	**50:1**	**70:1**
Dilute	1.7 ± 5E-3	1.7 ± 9E-3	1. 7 ± 2.4E-2
0.1	1.8 ± 6E-3	2.00 ± 9E-3	1. 7 ± 2E-2
0.2	2.07 ± 9E-3	2.03 ± 9E-3	1.9 ± 2E-2
0.3	2.3 ± 9E-3	2.2 ± 2E-2	2.11 ± 2E-2
0.4	2.5 ± 3E-2	2.4 ± 2E-2	2.5 ± 2.4E-2

Fitting error values shown as 3σ.

APPENDIX 2. BIMOLECULAR CONSTANTS OBTAINED BY WEIGHTED LEAST SQUARES FITTING FOR HYDRODYNAMIC INTERACTIONS STUDIES

1E8	50:1	1E9	70:1
Dilute	0.17 ± 6E-5	Dilute	1.7 ± 3E-2
0.1	0.17 ± 5E-4	10%	1.8 ± 3E-2
0.2	0.18 ± 1.2E-3	20%	1.9 ± 3E-2
0.3	0.21 ± 2E-3	30%	2.1 ± 3E-2
0.4	0.24 ± 9E-4	40%	2.4 ± 3E-2

Fitting error values shown as 3σ.

APPENDIX 3. BIMOLECULAR CONSTANTS OBTAINED BY WEIGHTED LEAST SQUARES FITTING FOR OBSTACLE SIZE STUDIES

Obstacle radium [nm]	**0.1**	**0.2**	**0.3**	**0.4**
1.0	0.17 ± 2E-3	0.16 ± 9E-4	0.17 ± 3E-3	-
2.33	0.18 ± 1.2E-3	0.18 ± 3E-5	0.20 ± 1.2E-3	0.21 ± 2E-3
4.0	0.17 ± 9E-4	0.19 ± 2E-3	0.21 ± 5E-4	0.26 ± 2E-3
6.0	0.18 ± 3E-4	0.19 ± 5E-4	0.22 ± 2E-3	0.25 ± 1.2E-3

Calculations performed for the 1E8, S:E = 50:1 system Fitting error values shown as 3σ.

APPENDIX 4. ANOMALOUS DIFFUSION EXPONENTS AND LIMITING DIFFUSION COEFFICIENTS FOR DIFFUSION STUDIES

R_{Obs} = 2.33	α	D* [$nm^2\ ns^{-1}$]	r_{Obs} = 6.99	α	D* [$nm^2\ ns^{-1}$]
0	0.9872	0.1043	0	0.9872	0.1043
0.1	0.9635	0.0860	0.1	0.9696	0.0945
0.2	0.9213	0.0693	0.2	0.9574	0.0826
0.3	0.8747	0.0579	0.3	0.9316	0.0686
0.4	0.7861	0.0372	0.4	0.9817	0.0555
R_{Obs} = 11.65	α	D* [$nm^2\ ns^{-1}$]	r_{Obs} = 16.31	α	D* [$nm^2\ ns^{-1}$]
0	0.9872	0.1043	0	0.9872	0.1043
0.1	0.9804	0.0962	0.1	0.9760	-
0.2	0.9682	0.0874	0.2	0.9743	-
0.3	0.9440	0.0762	0.3	0.9542	-
0.4	0.9115	0.0651	0.4	0.9333	-