Monte Carlo simulation of super-selective supramolecular polymers on cell membranes

Author: Adrià Terradellas Igual
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.**

Advisor: Ignacio Pagonabarraga and Co-advisor: Lorenzo Albertazzi

Abstract: Super-selectivity is a key challenge to design multivalent nano-entities that can bind ideally in a binary response "on-off"to selective receptor-concentration targeting. In this study, it is simulated a 2D canonical ensemble Monte Carlo to sample semiflexible polymer chains conformations with determined percentage of functionalized monomers, which are capable to be bound to a receptor. In addition, it is quantized the adsorption of polymers to cell membranes and analyzed their behaviour depending on potentially selective sources. In closing, it is simulated a novel self-assembled nanofiber to analyze its therapeutic potential for drug delivery.

I. MULTIVALENCY EFFECTS

One novel challenge in nano-medicine is the ability to design supramolecular entities that can bind selectively to pathogenic cells. Selectivity provides the capability of these entities to distinguish between substrates with different densities of binding sites, deducing from it a binary response to which the adsorption can be carried out or not depending on optimal conditions. In this "on-off"case, it is designated a super-selective regime which applications are broadly found in [1].

Recent experiments indicate that such selective behaviour can be obtained using multivalency [3]. Thus, it is necessary a theoretical analysis to describe the origin of the receptor-concentration threshold in the guest-host binding of multiligand nano-particles and the mechanism by which multivalency can lead to a super-selective regime. The analytical model presented by Francisco J. Martinez-Veracoechea and Daan Frenkel compare the targeting selectivity of monovalent and multivalent guest nano-particles [1]. Results obtained not only show that monovalent guests yield the *Langmuir isotherm*, ergo, adsorption θ depends linearly on the density of hosts n_R in $\log-\log$ form, but also that selectivity α is never larger than one (deduced by definition in Eq. 1).

$$\alpha \equiv \frac{d\log\theta}{d\log n_R} \tag{1}$$

In contrast, the parameter α for multivalent binding results to be extremely sensitive to changes in n_R, showing super-selective regimes over a broad range of concentrations. In Fig. 1 is displayed a comparison between two monovalent guests and a multivalent guest, which behaviours are simulated with same parameters but changing the considered origin of super-selectivity: the guest-host binding energy. Simulations, then, denote a significant difference in binding energies such that multivalent bonds are necessary to be weaker than monovalent bonds. Moreover, it is also suggested that this enhanced adsorption is owing to the overall contribution of weak bonds, which are capable to bind stronger than a strong bond. In fact, and this is explained in detail at section II, making a ligand-receptor bond stronger means making also the energy state of the system more negative (less probable). Thus, it is well-known that the increased value of α is not as a result of strong ligand-receptors interactions but sensitive to receptor-concentration [1].

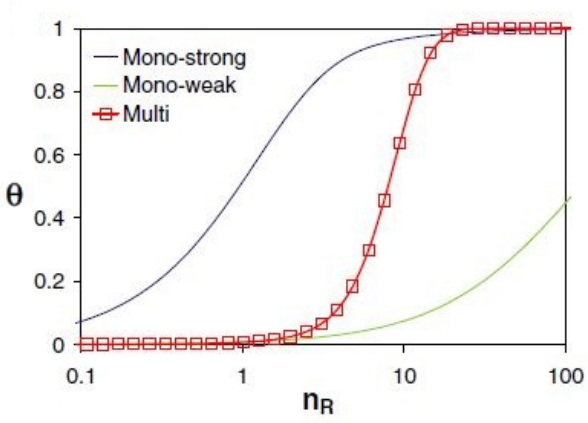

FIG. 1: Results of the analytical model from Francisco J. Martinez-Veracoechea and Daan Frenkel [1]. A direct comparison of two monovalent guests is distinguished although both responses are sigmoidal at different regimes. The multivalent guest presents a steep non-linear function typical from super-selective regimes.

A recent work developed by Tine Curk, Jure Dobnikar and Daan Frenkel presents a deep study on this emergent property of multivalency, which reveals that its origin is not only due to the sensitive binding energy but also remarks the entropic degeneracy pre-factor [6]. The degeneracy Ω, defined in Eq. 2, measures the number of ways in which i bonds can be formed in the flexible binding case between k ligands and n_R receptors. Thanks to the flexible capability of ligands, Ω becomes a very steep and non-linear function of k and n_R, a response super-selectivity does in a same way (Fig. 1).

*Electronic address: adria.terradellas@gmail.com

$$\Omega = \begin{pmatrix} n_R \\ i \end{pmatrix} \begin{pmatrix} k \\ i \end{pmatrix} i! = \frac{n_R!\ k!}{(n_R - i)!\ (k - i)!\ i!} \qquad (2)$$

Multivalent polymer adsorption is a good aspirant to validate super-selectivity described above, according to [6]. Thus, it is considered for this study semiflexible polymers with ligands randomly attached along the chain and mobile receptors with a single effective zone. This system requires an advanced Monte Carlo method based on the Rosenbluth algorithm, as outlined below.

II. CONFIGURATIONAL-BIAS MONTE CARLO

The nature of this study is, by definition, probabilistic. Consequently, and as it is no need to compute forces, a Monte Carlo simulation will generate random distributions in equilibrium that will depend on the details of the canonical ensemble *NVT*, that is to say, the number of particles, volume and temperature are constant. Then, it will be necessary to impose the condition of detailed balance between consecutive configurations (new and old) through variations of energy, as well as to determine the probabilities of generating a particular configuration and validate it through the Metropolis algorithm [4].

In the basic Metropolis method, a particle is moved to a new trial position by sampling the space around its old position randomly and uniformly. However, a polymer chain cannot be generated in a like manner. In this new case, it is more efficient to move a particle, in future a monomer, to a surrounding space where its potential is more negative and, therefore, the trial configuration has a high probability density [5]. This new technique is named *Configurational-Bias Monte Carlo method* (CBMC) and it will be used in order to know the probability map of the space surrounding the monomer in advance of the trial move.

Notice that, from a physical point of view, the CBMC method differ from the Boltzmann distribution. Hence, this method can be applied if and only if microscopic reversibility in the configurational bias algorithm is achieved [4, 5].

In the following subsections are developed in detail how all the elements of the system are constructed and their energetic effects to accept the trial configuration or not. Both fibres and receptors have in common the use of *Periodic Boundary Conditions* (PBCs) on a 2D lattice to approximate an infinite system by using a small part of it. In addition, both elements are also mobile, but as time dependent variables don't play any role, their mobility is simulated by generating new configurations such that each element of it can occupy different positions with different energies, but always in equilibrium.

A. Semiflexible polymer

By performing non-random sampling it is proposed a CBMC method based on the *Rosenbluth algorithm* to generate polymer chains. The Rosenbluth approach consists essentially in two steps: One step for generating a chain conformation with a bias that makes sure likely acceptable configurations with a high probability density, and another step to rectify the bias with a weight factor [4, 5]. Thus, the Rosenbluth scheme for a polymer chain is constructed segment by segment, up to l monomers, each of it with k possible directions, namely unoccupied nearest neighbours. The following strategy is specific for the treated system:

1. At the beginning of a trial configuration, a polymer n occupies its first monomer on a random position of the lattice. Neglecting the interaction energy between monomers and solvent atoms, the Boltzmann weight associated to the first monomer is only led by the coordination number of the lattice, $k = 4$.

2. All subsequent monomers will occupy positions r_2^n $\cdots \mathrm{r}_l^n$ determined by calculating, as in Eq. 3, the probability of occupying, from the previous monomer, i - 1, each available site j from a total of k.

$$\rho(r_{i,j}^n) = \frac{\exp\Big(-\beta H_i^n(r_{i,j}^n)\Big)}{\sum_{q=1}^{k} \exp\Big(-\beta H_i^n(r_{i,q}^n)\Big)} \qquad (3)$$

   Notice that $\beta = 1/\mathrm{k}_B T$, where k_B is the Boltzmann constant and T the absolute temperature. The Hamiltonian H describes the energy of the state which, in particular, is discussed below for a semiflexible polymer.

3. When a trial move is accepted, the monomer has a designated *Rosenbluth weight* (Eq. 4) in which contribute all Boltzmann weights from all k sites.

$$w_i^n = \frac{1}{k} \sum_{j=1}^{k} \exp\Big(-\beta H_i^n(r_{i,j})\Big) \qquad (4)$$

4. After the coordinate vector is fulfilled, it is required to calculate the *Overall Rosenbluth weight* for the whole polymer, which is the normalizing component of the transition matrix for this Markov process.

$$W_n = \prod_{i=1}^{l} w_i^n \qquad (5)$$

   Eq. 5 describes the CBMC method for an isolated chain and determines its acceptance.

5. In case of having a chain suspension, all the steps above are repeated until m fibers take place on the lattice. Then, it is applied Eq. 5 into the Metropolis algorithm, the acceptance ratio for consecutive polymer chains configurations defined in Eq. 6.

   If the ratio between the old and new configurations is greater or equal to 1, the new configuration is accepted by default, replacing the old with the new. If not, the new configuration has a determined probability to be accepted and, if it's not, the old configuration isn't replaced.

$$\min\left(1, \frac{\sum_{n=1}^{m} W_n^{new}}{\sum_{n=1}^{m} W_n^{old}}\right) \tag{6}$$

This method explained above is valid for every chain on a discrete lattice. However, it is needed to specify the type of conformation it's being treated, and the variable responsible to itemize it is the Hamiltonian. A semiflexible polymer behaviour can be modelled with the so-called *Worm-like chain model* (WLC), but this treatment is only valid for continuous lattices.

A discrete equivalent elastic energy is modelled in substitution to the WLC, based on a simulation developed by A.G. Bailey, C.P. Lowe, I. Pagonabarraga, and M. Cosentino Lagomarsino [2]:

$$U_e = \frac{B_s}{b} \sum_{i=1}^{n-1} [1 - \cos(\theta_i)] \tag{7}$$

Notice that B_s is the bending rigidity and b is the distance between two consecutive beads. But as semiflexible polymers on a discrete lattice have only two possible angles, namely 0 (straight) and $\pi/2$ (right and left sides), the bending Hamiltonian H_{bend} that is used in this study can be simplified. Nevertheless, note that the bending rigidity can be expressed also as $B_s = l_p k_B T$, where l_p is the persistence length. This polymer property is used not only to simplify even more Eq. 7 but also to deduce, as shown in Eq. 8, that H_{bend} makes the Boltzmann weight and, thus, all others derived from it, non-dependent on the temperature of the system (β's are cancelled).

$$H_{bend} = \begin{cases} 0, & \text{if } \theta = 0 \\ \frac{l_p}{\beta b}, & \text{if } \theta = \frac{\pi}{2} \end{cases} \tag{8}$$

Hence, Eq. 8 is applied into the Rosenbluth algorithm. Notice that, due to l_p and the exponential decay, trial moves on right and left sites are unlikely to be accepted. In fact, and this can be proved easily in simulations: the greater l_p, the straighter is the polymer.

B. Ligands

Ligands are the functional part of the polymer. Without them, fibres wouldn't be able to bind to receptors at all. In this study, ligands are randomly distributed along the fiber and aren't mobile. This is simulated by keeping their given monomer number at every generated configuration. While applying the Rosenbluth algorithm, its required treatment is particularly different. In this new case, the molecular spacer, which by definition provides a connection between the monomer and the ligand, has a significant role for the aim of the study. The importance of the flexibility of the spacer is detailed in section I.

At each trial position, the spacer of the monomer with ligand attached embraces an area approximated to an area of a square with length $d = 2L_s$, where L_s is the extension of the spacer. Then, it is chosen randomly a bead from it and, if and only if its position equals to an effective zone from the lattice of membrane receptors (subsection II C), the Hamiltonian will have an additive contribution due to the ligand-receptor binding energy E_{l-r}. If not, the binding energy is considered 0. Thereby, ligands are treated with a slightly different Hamiltonian, Eq. 9. Notice that, the more flexible the spacer, the more degeneracy can occur, meaning the number of effective zones to which can bind into a total of d^2 sites from the square area.

$$H_i^n(r_{i,j}^n) = H_{bend} + E_{l-r} \tag{9}$$

The binding energy can be attractive or repulsive. As mentioned in section I, a strong interaction doesn't make binding more probable due to the exponential decay in Eq. 3 and 4.

Fibers have as a variable the number of spacers per monomer N_s. In case $N_s > 1$, then, and if the spacer reach a receptor's effective zone, there is an extra difficulty for the ligand to be bound, namely a probability $1/N_s$ due to the assumption that each monomer can attach only a single ligand. This assumption is applied on the simulations taken for this study, in section III.

Examples of simulations with different types of fibers with ligands attached are shown in Fig. 2.

C. Receptors

Analogously to the lattice for fibers, another lattice is generated to represent same space but translated into a cell membrane, with receptors that are randomly distributed with determined surface coverage n_R.

A receptor, then, is simulated as a square, the length of which, namely its size S_R, is a variable of this study. However, it is remarked that, independently of S_R, all receptors have an unique effective zone located at the center of each one, meaning that if and only if the random position of the spacer equals to the center of a receptor, E_{l-r} can take place.

Thus, if a ligand is bound, the effective zone of the receptor involved is not functional anymore. To quantize correctly the surface coverage, simulations detect occupied sites due to receptors in order not to overlap them. Furthermore, notice that if $S_R = 1$ all receptors are made

of a single bead, namely the effective zone. By increasing S_R, as shown in Fig. 3, the number of effective zones it is being reduced by a factor $1/S_R^2$.

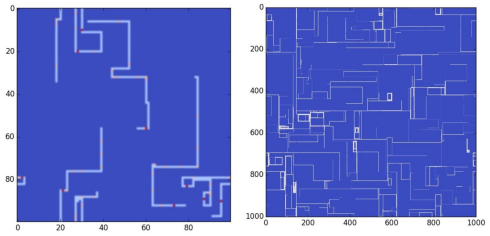

FIG. 2: Examples of two configurations for different lattices, which length is equal to the number of monomers per fiber, both with 5% of surface coverage of fibres. On the left, a lattice is reproduced with 5 fibres of 100 monomers each one. On the right, the lattice has 50 fibres, each of it with 1000 monomers.

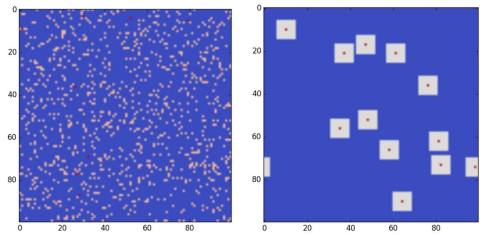

FIG. 3: Examples of two configurations on a lattice with length 100, simulating cell membranes with 10% of receptors with different sizes, known as S_R. On the left, $S_R = 1$, a single bead, meaning 1000 receptors on the lattice with only their effective zone. On the right, the size of the receptors is a length square of 9 beads, meaning 12 receptors on the lattice, each of it with an effective zone at its center. Notice that the number of effective zones are drastically reduced by a factor $1/S_R^2$ ($\frac{1000}{9^2} \approx 12$).

III. RESULTS AND DISCUSSION

First of all, series of simulations are carried out to analyze the computational cost for lattices with different sizes and different surface coverage of fibres. Simulations proved that more energetic contributions imply less percentage of accepted configurations. This could be solved by increasing the number of cycles per simulation, but time results to increase in this new case exponentially. Thus, to get more statistics in optimal time, the simulations described below are carried out with a lattice length equals to the number of monomers per fibre and 5% of surface coverage of fibres, as shown in Fig. 2.

All graphs represent adsorption (θ) vs. density of receptors (n_R) in semi-log plot, following the graph shown in Fig. 1. The n_R is normalized to 100, being 100 a density of receptors up to 50% of surface coverage. In the simulations a fiber is considered bound from just a single ligand-receptor interaction.

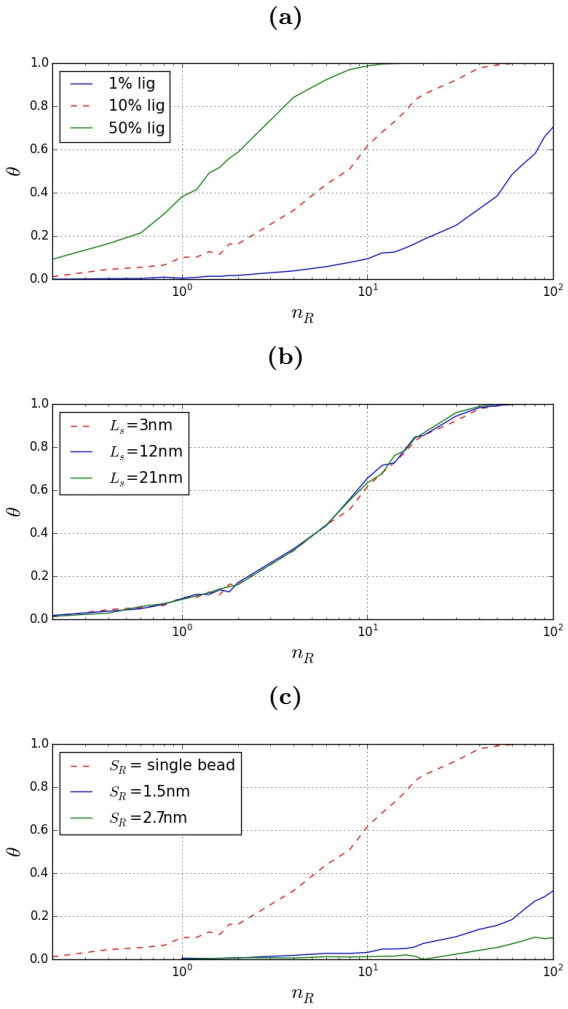

FIG. 4: θ vs. n_R of different simulations, which have in common the reference curve plotted in red dashes. Reference parameters are the following: $S_R = 1$, $L_s = 3nm$, $b = 0.3nm$, $N_s = 1$, $E_{l-r} = 3k_BT$, l_p is 1/6 of the total length, and 100 monomers (30 nm) per fibre with 10% of ligands. All simulations change only one variable, namely (a) the percentage of ligands, 1% and 50%; (b) the length of the spacer, rigid and flexible; (c) the size of the receptor, up to 2.7nm.

A. Potentially selective sources

A simulation has been carried out with variables defined in Fig. 4 and fitting a monovalent response. From here, the aim is to prove which variables change significantly the adsorption curve and verify super-selectivity proposed at section I. In same figure are exposed three different parameters which response are considered interesting to be discussed.

First of all, the number of ligands select different regions of concentrations. Although it's not shown a super-selective regime, they fit with the mono-weak and mono-strong behaviour from Fig. 1.

Another variable shown is the flexibility of the spacer, liable for super-selective behaviour. However, results conclude that, although L_s is almost the length of the fiber, there's no super-selectivity. In fact, this same simulation could be tested by increasing the other source of super-selectivity mentioned, E_{l-r}. Thus, from $3k_BT$ (weak) to $100k_BT$ (strong), no differences could be appreciated, opposing the multivalent property. A further continuation of this study would be to determine a minimum number of binding interactions to consider bound a fiber, such that weak interactions can bind stronger than strong interactions, and to observe super-selectivity due to flexible spacers. A single minimum interaction is not enough to reach a super-selective regime.

Finally, by increasing the receptor size, and as detailed in Fig. 3, the number of effective zones is being drastically reduced. Thus, it is very difficult to bind big receptors, specially if L_s acts rigidly, the cases of which break the typical sigmoidal response.

Other variables could be tested for their selective potential. By changing the persistence length no significant difference is appreciated. The number of spacers could be also tested and, although there's a clear tendency to soft the adsorption slope at a greater number, the difference is not appreciated to consider it a selective source.

B. Novel self-assembled nanofibers

The Nanoscopy for nanomedicine group, at IBEC, have synthesized a novel self-assembled fiber which monomers have an hydrophobic core and 3 hydrophilic spacers. These fibers are potentially candidates for drug delivery. It is necessary, then, understand their viability and determine if a super-selective regime can be carried out to get null adsorption into healthy cells, and full adsorption into cancer cells (which have much more receptor surface coverage). These kind of polymers have been observed by TEM and look semiflexible. Thus, and not by chance, this study can be applied on them.

In particular, these fibers have the following parameters: $S_R = 3nm$, $L_s = 3nm$, $b = 0.3nm$, $N_s = 3$ and l_p is 1/3 of the total length. Each fiber is made of 1000 monomers (300 nm) and has 1% of ligands. However, looking at the behaviour analyzed in the previous subsection, and specially due to big receptors and rigid spacers, even no monovalent response is observed and, same as Fig. 4c, the sigmoidal curve is broken. Experiments will conclude if a greater number of ligands doesn't cause destabilization of polymers.

A further study, then, is required to stand against the effect of a big receptor. Furthermore, this study opens the door to realize the impact of a 3D model compared to the 2D approximation used.

IV. CONCLUSIONS

In this study it has been presented, in a first place, the multivalent properties that lead to a super-selective response, namely the effect of the overall weak binding energies in contrast to a strong interaction, and the degeneracy of the spacer through its flexibility.

Secondly, the Rosenbluth algorithm has been proposed to simulate polymer chains. It has been detailed the role of the ligands and receptors on generated configurations and a bending Hamiltonian has been modelled for a semiflexible polymer on a discrete lattice, the Boltzmann weight of which results to be non-dependent on the temperature of the system.

Finally, simulations conclude that, from all candidates, the percentage of ligands is a selective source. The analysis of the receptor size promote the challenge to overcome the effect of big receptors, which are unlikely to be bound. This study has been found necessary to determine a minimum number of ligand-receptor interactions per fibre to prove both super-selective causes, since a single minimum interaction results to be not enough.

Acknowledgments

I would like to give a special thank to my advisor, Ignasi Pagonabarraga, for his continued guidance and for finding time in his busy schedule, even in Switzerland. Thank you very much also to Lorenzo Albertazzi for the opportunity to collaborate in this challenging project.

[1] Francisco J. Martinez-Veracoechea and Daan Frenkel, *Designing super selectivity in multivalent nano-particle binding*, PNAS 108 (27), 10963-10968 (2011)

[2] A.G. Bailey, C.P. Lowe, I. Pagonabarraga, and M. Cosentino Lagomarsino, *Accurate simulation dynamics of microscopic filaments using "caterpillar"Oseen hydrodynamics* Phys. Rev. E **180**, 046707 (2009)

[3] Lorenzo Albertazzi et al., *Spatiotemporal control and superselectivity in supramolecular polymers using multivalency*, PNAS 110 (30), 12203-12208 (2013)

[4] Daan Frenkel and Berend Smith, *Understanding Molecular Simulation: From Algorithms to Applications*, Chapter 13, pages 270-280, Academic Press (2002)

[5] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, pages 321-325

[6] Tine Curk, Jure Dobnikar and Daan Frenkel. *Design principles for super selectivity using multivalent interactions* (2016)