Università degli Studi di Cagliari

DOTTORATO DI RICERCA

FISICA

Ciclo XXVIII

TITOLO TESI

Computational methods for transport properties

Settore/i scientifico disciplinari di afferenza

FIS 07 / Fisica applicata

Presentata da: *Samuele Salis*

Coordinatore Dottorato *Prof. Alessandro De Falco*

Tutor *Prof. Paolo Ruggerone*

Esame finale anno accademico 2015 – 2016

Tesi discussa nella sessione d'esame marzo – aprile 2017

Computational methods for transport properties

SAMUELE SALIS

Università degli Studi di Cagliari
Facoltà di Scienze
Dipartimento di Fisica

Relatore:
Prof. Paolo Ruggerone

Coordinatore:
Prof. Alessandro De Falco

ABSTRACT

Through antimicrobial resistance[1] many bacteria can survive to an ever larger number of antibiotics. This is true in particular for a category of bacteria classified as gram–negative. These kinds of bacteria differ from the other ones by the presence of an outer membrane, which is able to protect them from the fast access (and consequently the action) of any antibiotics. The increasing capability of antibiotics to survive to many kinds of drugs has given rise to the Multiple Drug Resistance (MDR)[2]. New antibiotics could help to mitigate the MDR problem, but the poor understanding of permeability through outer membranes[3] has given an ever littler number of new patented antibiotics. This is due to a lack of experimental methods which are able to explain with a sufficient detail the permeation and, on the other side, to the difficulty in reaching the typical time scales (ms or even more[4]) of these processes. The category of antibiotics studied in this thesis can permeate the membrane crossing some porins[5] (beta barrel proteins nestled in bacterial outer membrane) so the permeation happens when we observe a transport of the antibiotic through a porin.

In this thesis we will focus on some computational methods, which are suitable to increase our understanding of transport processes. We will start with a post elaboration algorithm[6], that can be used to extract from an electrophysiology time series transport events apparently lower than the experimental device temporal sensitivity, continuing with another post elaboration algorithm that allows to extract the real transition time from a metadynamics simulation[7], skipping in this way the timescale problem in computer simulations, and we will finish with an ultra coarse grained model, that can be used to study the transport properties through a bacterial channel. Finally we will list the results obtained using the three aforementioned methods and we will summarise this thesis with the conclusions.

PUBLICATIONS

This thesis has contributed to the following articles:

[1] T. D'Agostino, **S. Salis**, M. Ceccarelli, A Kinetic Model for Molecular Diffusion through Pores, *BBA Biomemb.*, **1858(7B), 1772–1777** (2016)

[2] I. V. Bodrenko, **S. Salis**, S. Acosta–Gutierrez, G. Serra, A. Bosin, M. Ceccarelli, Small molecule transport through fluctuating nanochannels: the origin of the entropic barrier, *To be submitted*

[3] I. V. Bodrenko, J. Wang, **S. Salis**, M. Winterhalter, M. Ceccarelli, Sensing single molecules penetration into nanopores: pushing the time resolution to the diffusion limit, *Submitted*

ACKNOWLEDGMENTS

Of course thanks to the Lord, without which I would never write even a single word of this thesis.

To my family. To my wife Kezia and my children Cleofe and Giordano, that support me during the years of my PhD school with endless patience for a husband and parent often busy.

To my thesis supervisors Matteo Ceccarelli and Paolo Ruggerone for helpful advices. In particular to Matteo Ceccarelli that does not appear as a supervisor in the frontispiece of this thesis for a purely bureaucratic reason. A special thanks to Igor Bodrenko too, for priceless support, and to the other colleagues in the biophysics group.

Thanks to my thesis referees Mathias Winterhalter and Pratyush Tiwary that have accepted to read and judge this manuscript.

CONTENTS

LIST OF ACRONYMS

WHO	World Health Organization
MDR	Multiple Drug Resistant
FDA	Food and Drug Administration
IMI	Innovative Medicines Initiative
ND4BB	New Drugs for Bad Bugs
GN	Gram-negative
GP	Gram-positive
MD	Molecular Dynamics
CV	collective variable
MT	Metadynamics
WTMT	Well Tempered Metadynamics
MSM	Markov State Models
KMC	Kinetic Monte Carlo
FMTD	From Metadynamics to Dynamics
HOD	1–H–3–hydroxy–4–oxoquinaldine–2,4–dioxygenase
Omp	Outer membrane protein
AMBER	Assisted Model Building with Energy Refinement
CM	Centre of Mass
FES	Free Energy Surface
PSD	Power Spectral Density
KS	Kolmogorov-Smirnov

Part I

INTRODUCTION, SCOPE OF THE RESEARCH, STATE OF THE ART AND METHODS

1 INTRODUCTION

Contents

1.1 ANTIMICROBIAL RESISTANCE

According to World Health Organization (WHO) "Antimicrobial resistance happens when microorganisms (such as bacteria, fungi, viruses, and parasites) change when they are exposed to antimicrobial drugs (such as antibiotics, antifungals, antivirals, antimalarials, and anthelmintics). Microorganisms that develop antimicrobial resistance are sometimes referred to as *superbugs*".

This phenomena, in particular concerning Multiple Drug Resistance (MDR)[2] related to antibiotics, has shown an increasing trend in the last decades[8] and has reached alarming values as shown *e. g.* in Figure 1

The ways a bacterium can develop to become multidrug resistant are multiple[2].

1. **Mutational alteration of the target protein.** Every antibiotic must reach a target, that is a protein fundamental for the live cycle of the bacterium, and destroy it to accomplish its goal to kill the bacterium. The antibiotic is a chemical compound intrinsically static, so, if some bacterium develops a mutational alteration of the target protein, it can easily survive to that kind of antibiotic.

2. **Enzymatic inactivation of the drug.** The antibiotic is, by its nature, recognised as extraneous by the bacterium. If it develop some enzymatic defense sufficiently good, the antibiotic is destroyed before it can reach its target protein.

3. **Acquisition of genes for less susceptible target proteins from other species.** Some bacterium has already developed muta-

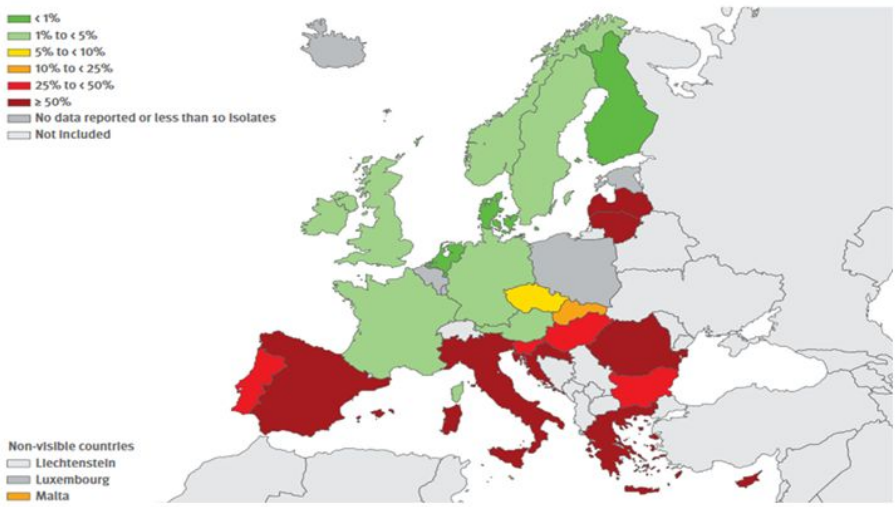

Figure 4. (Antimicrobial resistance surveillance in Europe, Surveillance report, ECDC, 2014). *Acinetobacter spp*. Percentage (%) of invasive isolates with combined resistance to fluoroquinolones, aminoglycosides and carbapenens, by country, EU/EEA countries, 2014.

Figure 1: Percentage (%) of invasive isolates with combined resistance to fluoroquinolones, aminoglycosides and carbapenens, by country, EU/EEA countries, 2014. Reprinted from [9].

tion of its target in a version less susceptible to the antibiotic action. Instead of develop again by scratch its own mutation, a bacterium can "import" DNA fragments from another kind of bacterium, and in this way it becomes antibiotic resistant in a shorter time.

4. **Bypassing of the target.** Specially in a hospital environment it is possible to develop bacteria strains multidrug resistant, giving rise to infections difficult to treat in patients.

5. **Preventing drug access to targets.** Antibiotics usually must overcome the bacterial membrane to reach its target located in the inner part. Bacteria have developed some special proteins (porins) nestled in the bacterial membrane and able to diffuse molecules in the inner part and some other (efflux pumps) devoted to actively extrude some molecules from the bacterium. Focusing on antibiotics, if the activity of porins is decreased (less antibiotics enter) or alternatively the action of efflux pumps is increased (more antibiotics exit), in these cases the action of antibiotics is suppressed and bacteria can survive.

The permeability of antibiotics through bacterial membranes is still poorly understood, and this lack of knowledge has given rise to an always littler production of new patented antibiotics (Figure 2) able to mitigate the aforementioned ability of bacteria to develop defense mechanisms against drugs.

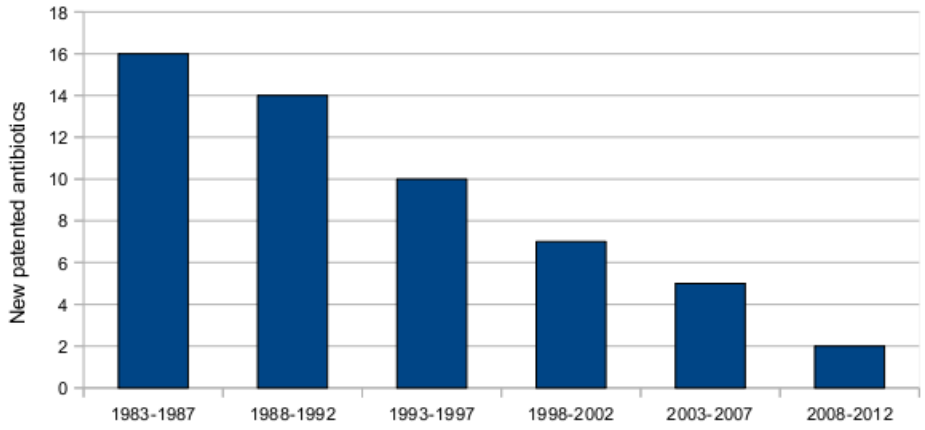

Figure 2: Number of new antibiotics patented by Food and Drug Administration (FDA) every five years from 1982 to 2012.[10].

The typical short duration of antibiotic therapy in patients affected by bacterial infection combined with the difficulties at finding new antimicrobial drugs have produced a poor interest of pharmaceutical companies in the research of new antibiotics. To cope with the increasing hazards for the populations and to the increasing governmental expenses in hospitalization of patients connected to the MDR and other kind of diseases, the international community has started promoting partnerships between public and private companies. An example in Europe is the *Innovative Medicines Initiative* (IMI) project able to develop new medicines[11]. Within the IMI project the New Drugs for Bad Bags (ND4BB) programme focus in particular on the research of new antibiotics and this thesis is involved through the academic group I have worked with.

In this thesis we will focus in particular on aspects connected to the study and comprehension of permeability of antibiotics through bacterial membranes of gram–negative bacteria. This aspect, as stated before, has a crucial importance in the action and efficiency of this kind of drugs.

1.2 UPTAKE OF ANTIBIOTICS IN GRAM–NEGATIVE BACTERIA

Bacteria can be divided in two categories: Gram–Positive (GP) and Gram–Negative (GN). For the purpose of this thesis the main difference between these types of bacteria is the presence (in GN) of an external membrane that does not exist in GP (Figure 3).

The outer membrane can be considered as being made of lypopolisaccharides (the main constituent of the membrane) that

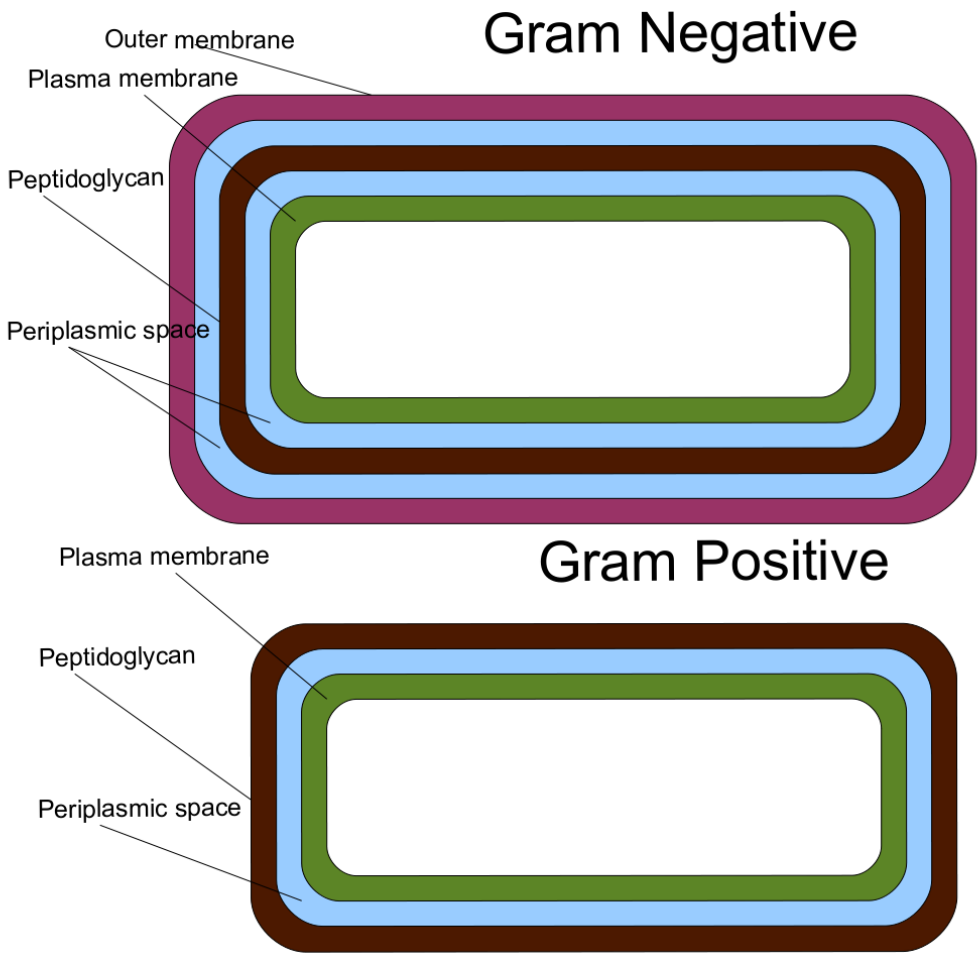

Figure 3: Schematic representation of GP and GN bacteria

forms a bilayer with a plethora of proteins implanted in it. Between these proteins there are the porins that in particular have been shown to be important[5] for the uptake of a category of antibiotics named β–lactams. The porins are organized in a secondary structure forming β–barrels, with beta sheets connected by loops, and can organise again in a trimeric (Figure 4) or monomeric (Figure 5) form. The porins are usually water filled (Figure 6) and are responsible of the passive penetration of hydrophilic molecules[5,12] that move through a diffusive motion.

Porins are often shaped to form an hourglass (Figure 6) with typical height and average radius of the order of 50 Å and 10 Å respectively. The central part is known as the constriction region, characterised by the minimal hollow cross section.

As already mentioned the porins usually transport many kind of molecules through the outer membrane and the transport has a typical diffusive nature[13]. Often the available cross section in the constriction region closely matches with the one of the molecule that

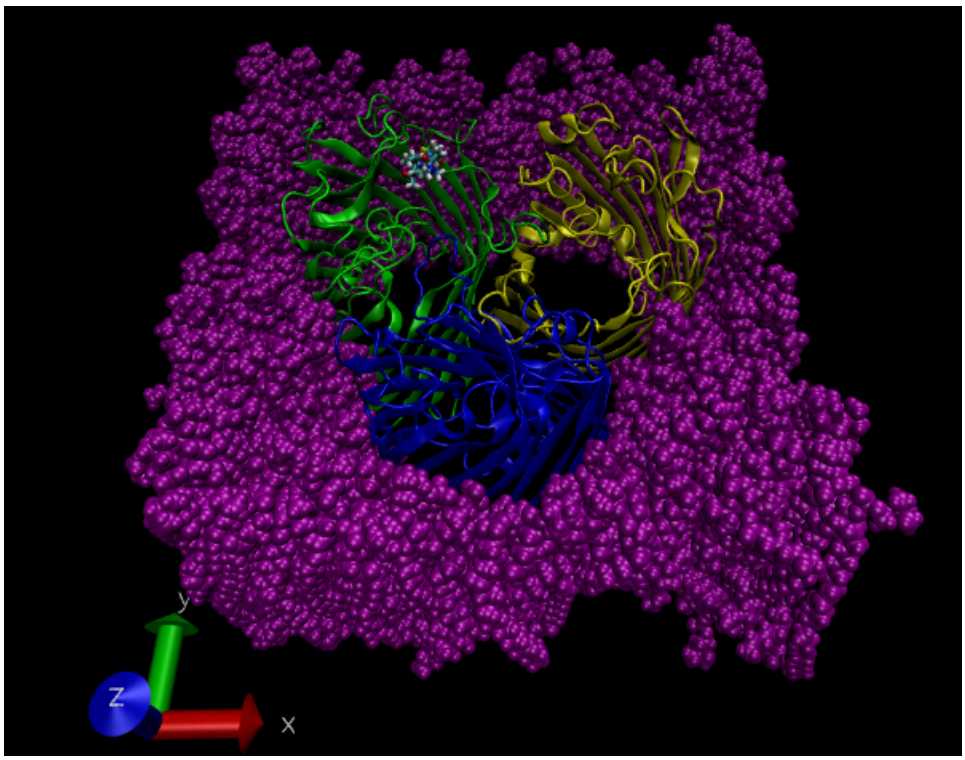

Figure 4: Representation (top view) of a typical simulated box with a trimer (in this figure an OmpF (Outer membrane protein F) constituted by three monomers in green, yellow and blue, a portion of lipid bilayer in magenta and an antibiotic (Meropenem) molecule close to the green monomer.

have to be translocated. For this reason size, shape, hydrophobicity or charge of the transported solute and of the residues composing in particular the constriction region can often influence the effective transport properties giving rise to porins more or less specific for a certain kind of molecules[12,13] mainly if a binding site for a specific molecule is present in the porin.

In our thesis we will focus particularly on transport of β–lactam antibiotics through porins, because is known[5] they use porins to cross the outer membrane. Because of the typical little space available for the antibiotic close to the constriction region, the times needed for the translocation increases and can reach times in the range of milliseconds in this particular case[14] or in comparable cases[4].

In the following part we will describe two approaches used to study translocations of antibiotics through bacterial porins: the experimental one and the computational. We will focus in particular in the experimental method called electrophysiology[15,16] and widely used also to study transport properties of antibiotics[5,14] and re-

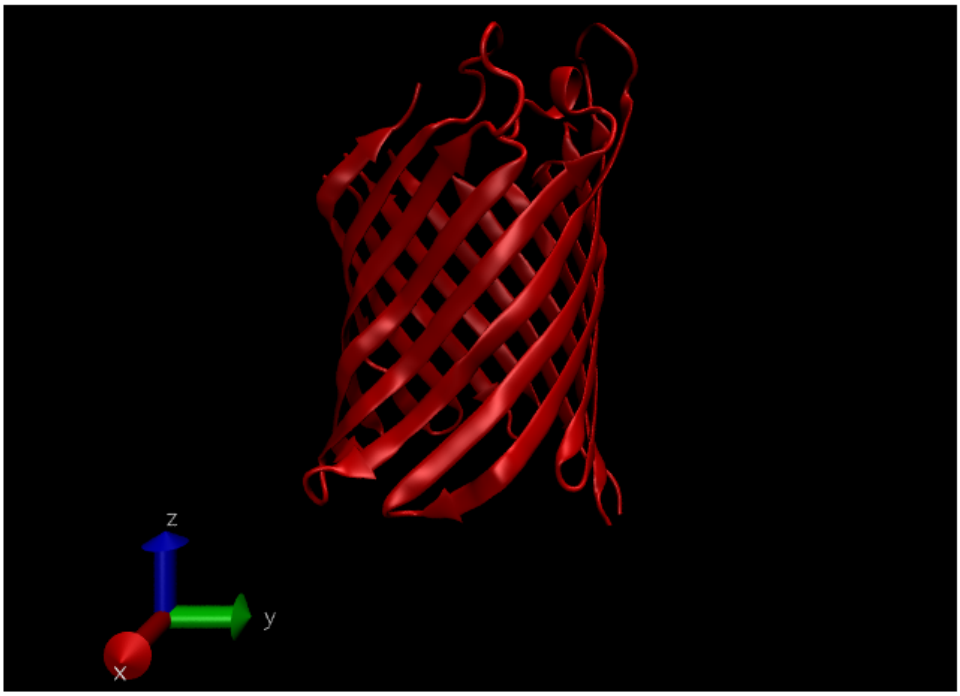

Figure 5: Example (side view) of a porin constituted only by a monomer. In this case is represented a Outer Membrane Protein G (OmpG).

cently postelaborated[6] to extend the time range analysable and in computational methods, focusing on the ones usable to skip the problem of the milliseconds time range, impossible to afford with standard computational calculations and power for systems like our composed of hundred of thousands of atoms. For every experimental and computational method we will point out positive and negative aspects.

1.3 EXPERIMENTAL STUDY OF TRANSLOCATIONS

Many experimental methods have been developed to study the translocation of solutes through porins[3]. Every method try to quantify only the flux of antibiotics into the bacteria, but can not give information about the microscopical details connected to the uptake of solutes. Between these methods it is interesting the one based on fluorescence microscopy[17], because it shows the best experimental detail in this field and it is able to detect the presence of fluorescent antibiotics with a single cell detail, but even in this most favourable case it is not available a microscopic detail and this method needs often the presence of a label applied to the antibiotic.

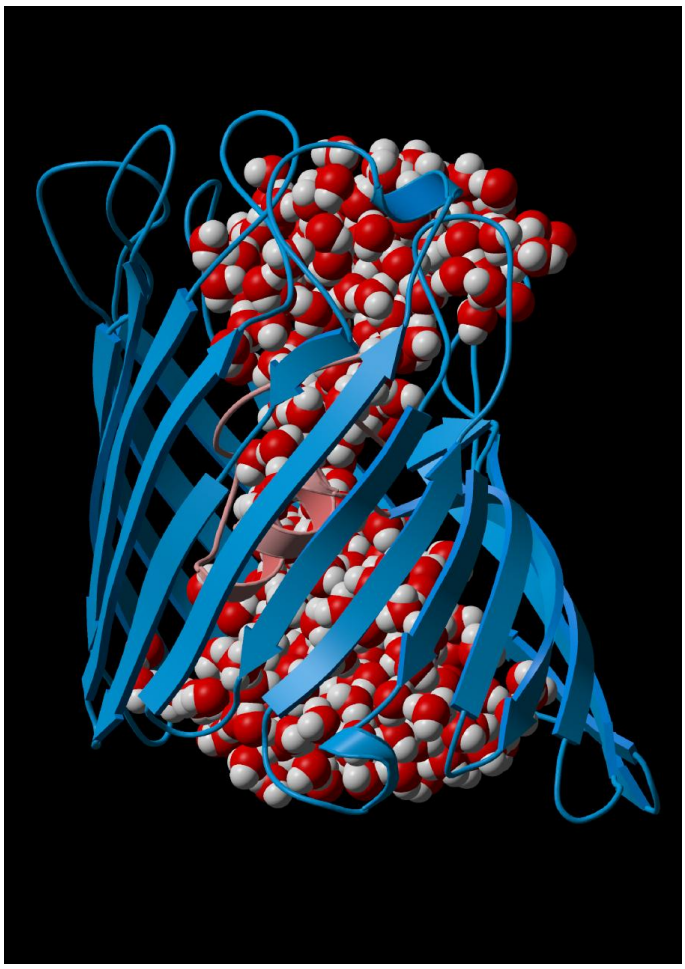

Figure 6: A monomer (side view) of a water filled OmpF. The distribution of the water into the porin displays the typical hourglass shape.

Another method widely used is the one based on electrophysiology, where a single porin is implanted in a planar lipid bilayer merged in an electrolytic solution, allowing to study conductance, selectivity and voltage gating of the channel[15,16] and allowing to characterise interactions between substrates and outer membrane channels[14,15]. Applying a voltage between the two sides of this solution separated by the lipid bilayer the ions will flow through the porin causing a tiny electric current. Finally inserting antibiotic molecules in the solution it is observed a drop in the current when an solute drug enters the porin (Figure: 7) reducing remarkably the ions flux for a sufficiently long time.

To know how long the current drop must last in order to be detected, it is useful to note the necessity to amplify the ionic current, to get a sufficiently big signal, and because of this amplification and other source of noise[18] the final current signal must be filtered usually first with a low–pass Bessel filter (4–th order or even more) at

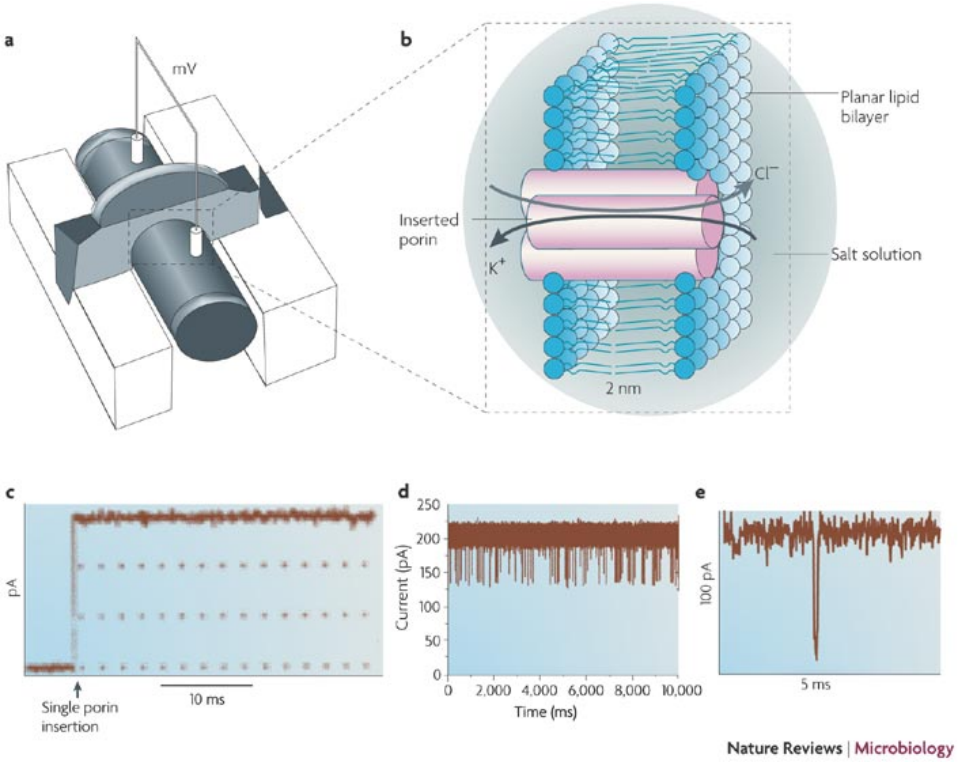

Figure 7: **a**. Scheme of the experimental apparatus used to perform electrophysiology measurements. **b**. Detail of the trimeric porin inserted in the planar lipid bilayer. **c**. Appearance of the ionic current through the porin, when its insertion in the lipid bilayer happens. **d**. Typical drops in former current when an antibiotic enters the porin. **e**. Zoom in a current drop. Reprinted from[5]

10 kHz and afterwards with a linear filter[6]. With this configuration a lower limit for this time (*i.e.* the lowest times detectable with this kind of instruments) was 50 μs[6,19] and it was a big drawback of this method because of the strong limit in detecting fast blockages, usually more interesting because connected with a poor interaction in the constriction region and a consequent potentially high antibiotic flux. Another minus of this method is the lack of microscopic detail and a simulated system often in environmental conditions different from the biological ones (*e.g.*applied voltage, high salt concentration, lipid bilayer instead of outer membrane) and the difficulty to connect the current drop to the real flux of the substrate through the channel[20].

One of the biggest problems connected to the electrophysiology method (the aforementioned minimum resolvable time around 50 μs) has been overcome recently[6], allowing the time resolution to be in the μ–seconds time range (this method will be described more in detail in the Section 2.1). Increasing the experimental time resolu-

tion has, in our opinion, a big practical influence, because it permits to study a set of translocation phenomena with a double approach: the electrophysiology and the computational methods. Without access to to the best computational facilities, this is possible only by simulating bigger time–scales through coarse grained calculations or other computational methods, that are able to increase the explored time range.

1.4 COMPUTATIONAL STUDY OF TRANSLOCATIONS

Another widely used approach to study drug discovery or, in particular, translocations of antibiotics through bacterial channels is the computational one[3,21]. Performing all–atoms simulations with explicit solvent of a typical system, that contains hundred of thousands of atoms as in the one described in Figure 4, it is possible to simulate up to the time range of 100 μ–seconds using the best available supercomputers dedicated to protein simulations[21] and classical Molecular Dynamics (MD) methods. This huge increase in computer power has helped to test with increasingly demanding benchmarks the force fields available for protein simulations, highlighting how performant these force field must be to manage all the events ranging from a typical femptosecond MD time step even to the milliseconds scale[22], with an even tiny imprecision likely to be amplified when a calculation is repeated millions or billions of times.

The translocations we simulate has often a typical time–scale of 100 μs or even more[4,14], making it practically impossible to simulate these processes even if we had a perfect force field and the most performant supercomputers or computational architectures[23,24]. Moreover a transport event often involves a molecule stuck in some free–energy basin[5]. It is necessary to cross the free–energy barrier to cross the bacterial channel and for this reason, in a MD simulation, we waste much of the computational time exploring states close to the free-energy minimum that have a really little interest compared to the ones close to the transition zone (saddle point). Speaking in a more mathematical mode, we can say that a translocation can be considered as a Poissonian event, that is a substantially stochastic event. Considering this last aspect, even having a few translocations simulated through MD, they could have a little impact in our studies because we need a number of crossings adequate to make a sufficiently accurate statistic of the event. To cope with all the needs of a large (dozens) number of translocations to

be obtained in a reasonable time (months), we have to choose some accelerated computational method.

One possible way out is to perform MD simulations using an increased time step. This is possible, but like every compromise has plus and minus. The increased speed in the simulation leads to an increased instability of the simulation with difficulties to preserve the energy conservation. On the other hand a bigger time–step cut all the events occurring in a time–scale littler than the integration step chosen, but this is not a wise choice, because we do not know *a priori* if the events removed have or not a role in the translocation we want to study. Moreover the strong dependence of the antibiotics flux on little changes in the constriction region[19] could suggest the need to preserve a little integration step to keep an high detail level.

Another method widely used especially to explore the free–energy surface, and for this reason suitable also for the problem of translocations, is the metadynamics[25] (MT). In this method is performed a MD simulation adding periodically to the system Hamiltonian a bias potential over a few (generally not more than two or three) Collective Variables (CV). This bias potential allows the system to escape easily from free–energy minima, exploring quickly and efficiently (if compared with other classical methods) the free–energy surface, but still keeping the MD detail of the simulated Hamiltonian. Using the Well Tempered Metadynamics[26] the convergence in this exploration is still faster and more efficient. One drawback of this method is the absence of any direct connection between the simulated MD time and the real time. To skip this problem, once available a free–energy profile of the phase space, it is possible to determine a set of metastable states and simulate massively (*e. g.* with a Kinetic Monte Carlo (KMC) scheme[27,28]) the system allowing to have in a short time a big number of translocation events. Unfortunately, following the Arrhenius equation, the times obtained for every transition between metastable states, and consequently for the whole translocation, are affected by big errors, because they are exponentially dependent on the height of the free–energy barriers estimated through the MT or WTMT, that still are affected by considerable errors[29]. Another way to obtain real transition times between metastable states is to use MT simulations and to post–elaborate a WTMT trajectory that made a transition over a saddle–point. Using a recent method[7] it is possible to obtain the real transition time from the otherwise useless MT time (from metadynamics to dynamics, as stated in the article title). Collecting a sufficiently big (at least around twenty) number of transitions and making some statistical tests on this times[30] it is possible to as-

sess the reliability of the obtained results under the hypothesis that every transition time follow a Poisson distribution. These methods, having basically the MT simulation as a core, give the possibility to obtain kinetic models useful to study the dynamics of translocation events[28,31].

Another approach which can be used to skip the time scale problem consists in implementing ultra coarse grained models able to point out the main physical parameters believed to influence the translocation. These models can be tested with *in silico* simulations to check the goodness of the physical hypothesis behind them. An essentially thermodynamic model, developed by our group following the basic idea here explained, will be described in more detail in the Section 2.3.

2

METHODS

Contents

2.1 ANALYSIS OF FAST CHANNEL BLOCKAGES

In electrophysiology experiments (see Section 1.3), we observe a typical generally on average constant current. Many kinds of disturbance cause noisy fluctuations of this current. When we have or have not a substrate into the electrophysiology solution, we observe a different behaviour of this current as shown in Figure 8. What we observe is a lower (if we focus on absolute values) average current (from $j_{av} = -147.3$ pA to $j_{av} = -144.3$ pA) when we insert a 5mM concentration of Cefepime and, at the same time, an increase of the standard deviation from $\sigma = 2.3$ pA to $\sigma = 4.8$ pA.

2.1.1 *Two–state Markov model for the current*

If we consider the channel monomer as a two–states Markov model with the monomer alternatively open or closed and we set two

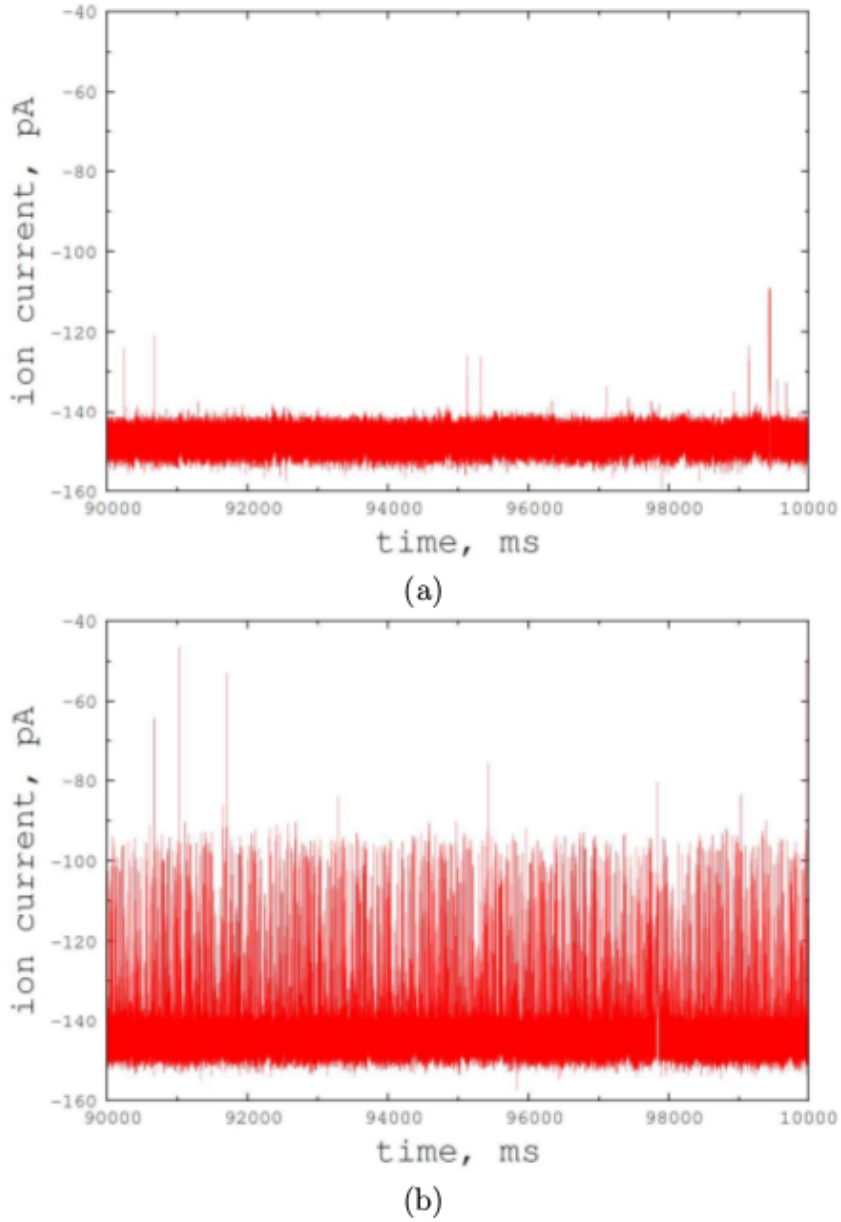

Figure 8: An example of ion current time series (a) without substrate or (b) with Cefepime 5mM.

values j_0 (open channel) and j_1 (occupied channel) for its current, we can calculate the average current of a channel constituted of N_c monomers at substrate concentration c as

$$j_{av}(c) = N_c[j_1 P(c) + j_0 \cdot (1 - P(c))] \tag{1}$$

where $P(c)$ represents the equilibrium probability of occupation for a monomer as a function of substrate concentration and for $c = 0$,

$j_{av}(0) = N_c j_0$. It is possible to express the occupation probability $P(c)$ in terms of the transition rates[4,32] obtaining

$$\langle \Delta j \rangle = |j_{av}(c) - j_{av}(0)| = \frac{N_c c k_{on} \Delta j}{k_{off} + c k_{on}} \tag{2}$$

with $\Delta j = |\langle j_1 \rangle - \langle j_0 \rangle|$ for a certain applied potential. In the last formula k_{off} and k_{on} represent respectively the dissociation and association rates. If the channel is completely blocked, $\Delta j = |j_{av}(0)|/N_c = |\langle j_0 \rangle|$, otherwise, for a partial monomer occupation, $\Delta j < |\langle j_0 \rangle|$.

The power spectral density (PSD) of the ion current fluctuations due to the blockages by blockages by the substrate in the two–state Markov model is[14]

$$G(f) = \frac{a}{1 + (2\pi f \tau_c)^2} \tag{3}$$

with

$$a = 4 N_c (\Delta j)^2 c k_{on} k_{off} \tau_c^3 \quad \text{and} \tag{4}$$

$$\tau_c = \frac{1}{k_{off} + c k_{on}}. \tag{5}$$

where a is the Lorentzian factor and τ_c is the correlation time of the blockage process.

When $k_{off} \gg k_{on}$, the average current drop $\langle \Delta j \rangle$ and the Lorentzian factor a can be expanded in a Taylor series up to the second order obtaining

$$\langle \Delta j \rangle (c) = N_c \Delta j \frac{c k_{on}}{k_{off}} \left(1 - \frac{c k_{on}}{k_{off}} \right) + o\left(\left(\frac{c k_{on}}{k_{off}} \right)^2 \right), \tag{6}$$

$$a(c) = \frac{4 N_c (\Delta j)^2}{k_{off}} \frac{c k_{on}}{k_{off}} \left(1 - 3 \frac{c k_{on}}{k_{off}} \right) + o\left(\left(\frac{c k_{on}}{k_{off}} \right)^2 \right). \tag{7}$$

2.1.2 *Current filtering*

Due to the large amount of noise in electrophysiology experiments[18], the ion–current signal is filtered by an analogous low–pass linear filter (usually a 4–order (4–pole) low–pass Bessel filter) and then sampled at a specific rate $f_s = 1/\Delta t$. Practically the second sampling is done averaging the signal for a period Δt and, for this reason, this second sampling can be considered a linear filter following the first analogous one. The average ion current j_{av} is not affected by the linear filter, but the dispersion, the auto–correlation

function and the PSD are modified by both filters. Calling $H_a(\omega)$ the transfer function of the analogous filter and $H_s(\omega)$ the one of the sampling filter, then the PSD of the original ion current in Equation 3 becomes, after filtering, the following

$$G_m(f) = \frac{a}{1 + (2\pi f \tau_c)^2} |H_a(2\pi f)|^2 |H_s(2\pi f)|^2 \tag{8}$$

At small frequencies both transfer function are close to 1 and do not modify the PSD. At large frequencies ($f \gg f_c$, with f_c representing the low–pass filter cut–off frequency, or $f \gg 1/\Delta t$), the original PSD is suppressed by the filters. From Equation 3 we can see that if $\tau_c \lessapprox \Delta t$ or $\tau_c \lessapprox f_c^{-1}$, the filters start to influence strongly the PSD.

2.1.3 *Determining k_{on} and k_{off}*

If we fit the filtered PSD in Equation 8 to the observed PSD, we can obtain a and τ_c as a function of c and after get k_{on} and k_{off} from Equation 4 and 5. Unfortunately the calculation of τ_c and $\langle \Delta j \rangle$ is not always possible and we can depict different scenarios.

2.1.3.1 *Case 1*

In the best case (τ_c and a obtained from the fitting of PSD in Equation 8 and Δj obtained from current distribution, for example) we can use Equation 4 and 5 to obtain

$$k_{off} = \frac{1}{2\tau_c}\left(1 + \sqrt{1 - \frac{a}{N_c \tau_c (\Delta j)^2}}\right) \tag{9}$$

$$k_{on} = \frac{1}{2\tau_c c}\left(1 - \sqrt{1 - \frac{a}{N_c \tau_c (\Delta j)^2}}\right) \tag{10}$$

2.1.3.2 *Case 2*

If it is not possible to calculate directly Δj, but it is still possible to calculate τ_c, then one can use Equation 5 to obtain k_{on} and k_{off}. Having these two values it is possible to obtain Δj using Equation 4

$$\Delta j = \sqrt{\frac{a}{4 N_c c k_{on} k_{off} \tau_c^3}} \tag{11}$$

2.1.3.3 *Case 3*

If τ_c is very small and can not be fitted using Equation 8, we are in the case where we can make the approximations that lead to

Equation 6 and 7, then we can fit the dependence of these equations on c with the following simplified equations:

$$\langle \Delta j \rangle(c) = b_j(1 - rc) \tag{12}$$

$$a(c) = b_a(1 - 3rc) \tag{13}$$

where $r = K = k_{on}/k_{off}$ is the equilibrium or stability constant of the solute/channel interaction. Comparing the fitting parameters b_a and b_j with physical values through Equation 6 and 7, it is possible to get k_{on}, k_{off} and Δj in the following way

$$k_{on} = \frac{4b_j^2}{N_c b_a}, \tag{14}$$

$$k_{off} = \frac{k_{on}}{r}, \tag{15}$$

$$\Delta j = \frac{b_j}{N_c r}. \tag{16}$$

2.1.3.4 *Case 4*

In the last case, if τ_c is so small to make impossible to find its dependence on c, but it is possible to obtain Δj from direct measures, then one can obtain k_{on} and k_{off} basically through Equation 6 and 7.

First we obtain r fitting to the experimental average current the following equation

$$\langle \Delta j \rangle(c) = N_c \Delta j(1 - rc)rc. \tag{17}$$

After, fitting Equation 13 with a fixed value of r, we obtain

$$k_{off} = \frac{4N_c(\Delta j)^2 r}{b_a} \quad \text{and} \tag{18}$$

$$k_{on} = rk_{off}. \tag{19}$$

2.1.4 *Detection of sub–microsecond channel gating*

With the method used in Section 2.1.1 has been obtained residence times t_{res} around one micro–second (see Section 3.1.2). This method is based essentially on the manipulation of PSD and average ion current, however a different approach can be used considering, instead of PSD and average ion current, two observables: average ion current and current variance. This method allows to set a lower limit for τ_{res}, that can produce $\tau_{res} \approx 400$ ns with a standard electrophysiology equipment.

2.1.4.1 *Theory*

Considering as in Section 2.1.1 the two filters (analogous low–pass and linear filter due to sampling), the average value of the ion–current signal is not affected by the linear filter, but the variance, the auto–correlation function and the PSD are modified by the both filters. If the transfer function (its Fourier transform) of the analogous filter is $H_a(\omega)$ and that of the sampling filter is $H_s(\omega)$, then the relation between the PSD of the original (the one before the filters) ion current, $g_0(\omega)$, and the one of the signal after successive application of the two filters is, similarly to how obtained before in Section 2.1.1,

$$g(\omega) \;=\; g_0(\omega)|H_{\mathrm{a}}(\omega)|^2|H_{\mathrm{s}}(\omega)|^2. \tag{20}$$

The variance of the ion-current signal is equal to the un–normalised auto–correlation function at zero time shift ($\sigma_j^2 = R(0)$ or $R(0,T)$ using the notation hereafter specified), if we define

$$R(\tau,T) = \frac{1}{N_T - M_\tau} \sum_{k=1}^{N_T - M_\tau} [j_k - j_{av}(T)][j_{k+M_\tau} - j_{av}(\tau,T)], \tag{21}$$

with Δt meaning the sampling interval, $T = N_T \Delta t$ representing the duration of the sampling, $j_{av}(T) = \frac{1}{N_T}\sum_{k=1}^{N_T} j_k$ and $j_{av}(\tau,T) = \frac{1}{N_T - M_\tau}\sum_{k=1}^{N_T - M_\tau} j_{k+M_\tau}$.

Then, as the power spectral density of the signal is the Fourier transform of the auto-correlation function,

$$g(\omega) \;=\; \frac{1}{\pi}\int_0^\infty R(\tau)\cos(\omega\tau)\mathrm{d}\tau\ , \tag{22}$$

and by using Equation (20) can be obtained

$$\sigma_j^2 = \int_{-\infty}^{\infty} g(\omega)\mathrm{d}\omega \;=\; \int_{-\infty}^{\infty} g_0(\omega)|H_{\mathrm{a}}(\omega)|^2|H_{\mathrm{s}}(\omega)|^2\mathrm{d}\omega. \tag{23}$$

This last equation relates the observed ion–current variance to the PSD of the original unfiltered ion current.

If the correlation time of the original (unfiltered) ion current is much smaller than the inverse of the cutoff frequency (f_c) of the analogue filter, then function $g_0(\omega)$ is much wider then transfer functions of the filter. It is also possible to consider that the sampling frequency f_s is sufficiently larger than f_c as this is the instrumental requirement to avoid aliasing effects. Therefore, one may take $g_0(\omega)$ and $|H_s(\omega)|^2$ out of the integration in Equation 23 and obtain

$$\sigma_j^2 \;\approx\; g_0(0)\int_{-\infty}^{\infty}|H_{\mathrm{a}}(\omega)|^2\mathrm{d}\omega. \tag{24}$$

Here, we have taken into account that $|H_s(0)|^2 = 1$. By using an integral definition of the correlation time of the original signal

$$\tau_c \equiv \frac{\int_0^\infty R_0(\tau)d\tau}{R_0(0)} = \frac{\pi g_0(0)}{\int_{-\infty}^{\infty} g_0(\omega)d\omega}, \tag{25}$$

from Equation (22) one obtains

$$g_0(0) = \frac{\tau_c \sigma_0^2}{\pi}, \tag{26}$$

where $R_0(\tau)$ and $\sigma_0^2 = R_0(0)$ are the autocorrelation function and the variance of the original (unfiltered) signal, correspondingly.

A linear low–pass filter is characterized by a single dimensional constant, the cutoff frequency $\omega_c = 2\pi f_c$, defined as $|H_a(\omega_c)|^2 = 1/2$. Therefore, the dimensionless transfer function must be a function of ratio ω/ω_c, *i. e.* $H_a(\omega) = f(\omega/\omega_c)$. Then it is possible to write

$$\int_{-\infty}^{\infty} |H_a(\omega)|^2 \mathrm{d}\omega = 2\pi f_c \xi_f, \tag{27}$$

where ξ_f is a dimensionless numerical constant depending on filter type.

For the 4–order Bessel filter, typical for the patch clamp electrophysiology, one finds $\xi_f \approx 2.093$. And by combining equations (Equation 24-Equation27) one obtains

$$\sigma_j^2 = 2\xi_f f_c \tau_c \sigma_0^2. \tag{28}$$

The last relation represents the linear low–pass filter correction to the variance of the signal at conditions $f_c\tau_c \ll 1$ and $f_c \ll f_s$. No major assumptions have been made about the statistical nature of the original signal except of its stationarity and of the existence of the correlation time in the sense of Equation 25.

In the following development, we will assume, as in Section 2.1.1, that the substrate–induced ion–current fluctuations are caused by a 2–state Markov process. The probability density of the ion current values are $f(j) = f_{op}(j)p_{op} + f_{cl}(j)p_{cl}$. Here, p_{op} and $f_{op}(j)$ are the equilibrium probability and the current probability density of the open channel, p_{cl} and $f_{cl}(j)$ are the corresponding quantities of the "closed" channel, *i. e.* , with the substrate molecule inside. We also have assumed, as justified in Section 2.3.4, that only a single substrate molecule at a time may be in the channel, at least, in the current blocking state. Then, the average current through the channel and its variance before filtering become, respectively,

$\bar{j} = \bar{j}_{op}p_{op} + \bar{j}_{cl}p_{cl}$ and $\sigma_0^2 = \sigma_{op}'^2 p_{op} + \sigma_{cl}'^2 p_{cl} + (\bar{j}_{op} - \bar{j}_{cl})^2 p_{op}p_{cl}$, where $\bar{j}_p$ and $\sigma_{op}'^2$ are the average current and its variance in the open state before filtering, while $\bar{j}_{cl}$ and $\sigma_{cl}'^2$ are the quantities for the closed state. If the ion–current fluctuations in the open channel state and those in the close state are independent and much faster (have much smaller correlation time) than the substrate–binding Markov process, then the observed (filtered) variance may be written as

$$\sigma_j^2 = \sigma_{op}^2 p_{op} + \sigma_{cl}^2 p_{cl} + 2\xi_f f_c \tau_c (\Delta j)^2 p_{op} p_{cl}. \tag{29}$$

Here $\Delta_j = |\bar{j}_{op} - \bar{j}_{cl}|$, while σ_{op}^2 and σ_{cl}^2 are the filter–corrected variances of the ion–current in the open and in the closed channel state, respectively. The values ξ_f and f_c are the filter constant and the cutoff frequency, as they were discussed for Equation 28. Finally τ_c is the correlation time of the substrate binding process.

Further, we suggest that substrate molecules are added from one side of the channel with bulk concentration c. By using the association rate constant, k_{on}, so that the association frequency equals $\nu_{\mathrm{on}} = ck_{\mathrm{on}}$, the dissociation rate constant k_{off} and the binding constant $r = k_{on}/k_{off}$, then the equilibrium probabilities of the open and the closed channel become, correspondingly, $p_{\mathrm{op}} = 1/(1+cr)$ and $p_{\mathrm{cl}} = cr/(1+cr)$, as can be argued from Equation 5. Always from the last equation we know that the correlation time is $\tau_c = (k_{off} + ck_{on})^{-1}$. Now we can define the model statistical observables for N_c independent active monomers channels, that is the average substrate–induced current shift $\delta j(c) = N_c|\bar{j} - \bar{j}_{op}|$ and the substrate–induced variance $\sigma_s^2(c) = N_c(\sigma_j^2 - \sigma_{op}^2)$,

$$\delta j(c) = N_c \Delta j \frac{cr}{1+cr}, \tag{30}$$

$$\sigma_s^2(c) = 2\xi_f N_c (\Delta j)^2 \frac{f_c}{k_{off}} \frac{cr}{(1+cr)^3} + \frac{N_c(\sigma_{cl}^2 - \sigma_{op}^2)cr}{1+cr}. \tag{31}$$

The second term in Equation 31 comes from the difference in the magnitude of ion–current fluctuations in the open channel state and that in the close state. In many cases this term may be neglected compared with the first one originating from the transitions between the states having significantly different conductance. Then, one may use a simplified equation for the substrate induced variance.

$$\sigma_s^2(c) = 2\xi_f N_c (\Delta j)^2 \frac{f_c}{k_{off}} \frac{cr}{(1+cr)^3}. \tag{32}$$

We will focus on this case below hereafter.

As the instrumental filter's properties and the substrate concentration are known from the experimental setup, there are 4 parameters in the two Equations 30 and 32, namely N_c, Δj, r and k_{off}. The number of active monomers may be determined from the experiment by measuring the conductance without the substrate and the remaining three parameters are not known in general case. They must be obtained by fitting Equations 30 and 32 to the observables calculated from the experimental ion–current traces, in the way similar to the that used in Section 2.1.3.3 for the PSD fitting. There are conditions for the substrate concentration when this complete 3–parameter data analysis is possible. If the substrate concentration is very small, so that $cr \ll 1$, then both the observables ($\delta j(c)$ and $\sigma_s^2(c)$) are directly proportional to the concentration. In this linear regime all the three unknown parameters are multiplicative and they may not be determined simultaneously. Thus, the largest substrate concentration, c_{max} used in the measurements should obey the conditions $cr \geq 0.1$, *i. e.* when the deviation from the linear concentration dependence in $\sigma_s^2(c)$ is clearly detectable. Because, in general, the r value is not known in advance, one need to perform the electrophysiology experiment increasing the substrate concentration until the non–linear dependence becomes quantifiable.

It is not always possible experimentally to achieve the necessarily high substrate concentration due to the limited solubility. In that case, one can still use $\delta j(c)$ and $\sigma_s^2(c)$ to determine the substrate binding kinetics, if *a priori* information on Δj is available. For example, if there is an experimental evidence or a theoretical argument that the monomer channel is completely blocked for the ion current when the substrate in bound ($\bar{j}_{\mathrm{cl}} = 0$), then $\Delta j = |j_{op}|$. The average current for the open monomer is determined directly from the experiments, if the number of active channels is also known. Then, the substrate–binding kinetic parameters are readily available from Equations 30 and 32 as follows

$$r = \frac{1}{c}\frac{\delta j}{(N_c \Delta j - \delta_j)}, \tag{33}$$

$$k_{off} = 2\xi_f f_c \frac{\Delta j \delta j}{\sigma_s^2}\left(1 - \frac{\delta j}{N_c \Delta j}\right). \tag{34}$$

The kinetic parameters r and k_{off} may be obtained by using Equations 33 and 34 from the ion–current trace measured at a single concentration c. However, a direct 2–parameter fit of Equations 30 and 32 to the experimental concentration–dependent $\delta j(c)$ and $\sigma_s^2(c)$ values might be preferred to improve the accuracy of the extracted

parameters. The ability of the model to reproduce the substrate concentration dependence of both $\delta j(c)$ and $\sigma_s^2(c)$ provides an additional consistency test.

The signal from the channel blockages due to the substrate may be detected over the background (*i. e.* , ion–current through the open channel) if its variance σ_s^2 exceeds or at least is similar to the variance of the background (σ_{op}^2), *i. e.* $\eta = \sigma_s^2/\sigma_{op}^2 \geq 1$. From Equation 32, by assuming condition $cr \ll 1$ (the linear substrate concentration dependence regime) one obtains

$$\eta \quad = \quad 2\xi_f N_c cr \frac{f_c}{k_{off}} \frac{(\Delta j)^2}{\sigma_{op}^2} \geq 1. \tag{35}$$

The latter equation is the condition for substrate–induced channel–blockages signal to be detectable. Formula 35 shows that one can improve the signal detection by increasing the number of active monomer channels in the experimental setup or by increasing the substrate concentration. Moreover, one may optimize the applied electric potential to maximize ratio $(\Delta j)^2/\sigma_{op}^2)$. However the usability of this way is limited, as the kinetic parameters may also be voltage dependent. A higher filter cutoff frequency might also improve the detection accuracy, but usually it is not possible to change continuously the filter cutoff frequency f_c, due to the instrumental limitations. Moreover, the increase of f_c is accompanied by the increase of the background noise, σ_{op}^2. If we take, for example, a typical 4–order Bessel filter with 10 kHz cutoff frequency, assume 3 active channels, $cr = 0.1$, $\Delta j = 50$ pA and $\sigma_{op}^2 = 10$ pA^2, then the detectable k_{off}, according to Equation 35, may be as high as 2.5×10^6 1/s. The latter corresponds to the residence time of 400 ns.

2.2 COMPUTATIONAL METHODS

Apart from standard MD[33], MT[25] and WTMT[26] simulations performed with the ACEMD[34] software with a PLUMED plugin version 1.3[35] and with interactions based on the AMBER 99SB–ILDN force field[36], the not standard method mentioned here is the procedure[7,30] that allows us to obtain transition times from WTMT simulations.

2.2.1 *From metadynamics to dynamics*

Metadynamics is a technique used to calculate static properties and its main use is connected to the exploration of the FES, in particular

when it has a rough profile and many basins could be not explored using standard MD simulations, due to the large energy barrier to be passed.

Metadynamics is based on MD simulations performed using a biased Hamiltonian. The biasing is obtained typically with some Gaussians periodically added around some coordinates named collective variables (CVs). If the CVs represent some relevant physical quantity for the system, this biasing allows a fast transition through the free–energy barriers, leading to a fast exploration of the FES. Due to the MD calculations involved in MT simulations a "MT time" is still present, but in the classical version of the method there are not connections between this time, basically considered like nothing more than an index, and the real time.

The method from metadynamics to dynamics (FMTD) permit to get a connection between the MT time and the real time needed to perform a transition through a saddle point in a free–energy basin. The basic assumptions are the presence of a, let suppose one–dimensional, free-energy basin that is passed when the value of $\lambda > \lambda^*$, with λ generally dependent, maybe through some CVs named s, on the system coordinates $\mathbf{R}$. If we assume a local equilibrium at all times, we can guess a Boltzmann distribution and, if the time spent in the saddle point is much less than the one spend in the basin, we can write the average time τ needed to make a transition through the bottleneck in the saddle point

$$\tau = \frac{1}{\omega\kappa}\frac{Z_0}{Z_0^*} = \frac{1}{\omega\kappa}\frac{\int_{\lambda\leq\lambda^*} e^{-\beta U(\mathrm{R})}\mathrm{d}\mathbf{R}}{\int_{\lambda=\lambda^*} e^{-\beta U(\mathrm{R})}\mathrm{d}\mathbf{R}}, \tag{36}$$

with ω a normalisation constant (intuitively the frequency of attempt to cross the barrier) and κ a transmission coefficient, whose value is not important in our case, because we assume a fast crossing of the saddle point.

If we assume the ability to perform MT simulations with good CVs that do not deposit bias near to the transition zone, then we can write the equivalent of the Equation 36 as

$$\tau_M(t) = \frac{1}{\omega\kappa_M}\frac{Z_M}{Z_M^*} \tag{37}$$

that is a function of the MT time t.

Assuming we do not bias the transition zone in our MT simulation, we can say that $\kappa_M \approx \kappa$ and, supposing good CVs that sample well the free–energy basin, we can assume $Z_M^* \approx Z_0^*$. At this point,

having the ratio between Equation 36 and 37, we can write down $\alpha = \tau/\tau_M$ as

$$\alpha(t) \approx \frac{Z_0}{Z_M} = \langle e^{\beta[V(s((R),t))]} \rangle_M \tag{38}$$

The factor α is simply the average of the bias potential in MT simulations calculated until the system stays in the basin in conditions $\lambda \leq \lambda^*$. Knowing this value we can easily get the real transition time as

$$\tau = \alpha\tau_M. \tag{39}$$

This formula means that the real time transition time is obtained multiplying the MT time by the α factor, named acceleration factor.

To know when the transition occurs is easy observing the trend of $\alpha(t)$, because when a transition occurs, the system visit a zone with zero bias $V(s((R),t))$ causing a typical drop in his value as shown in Figure 9

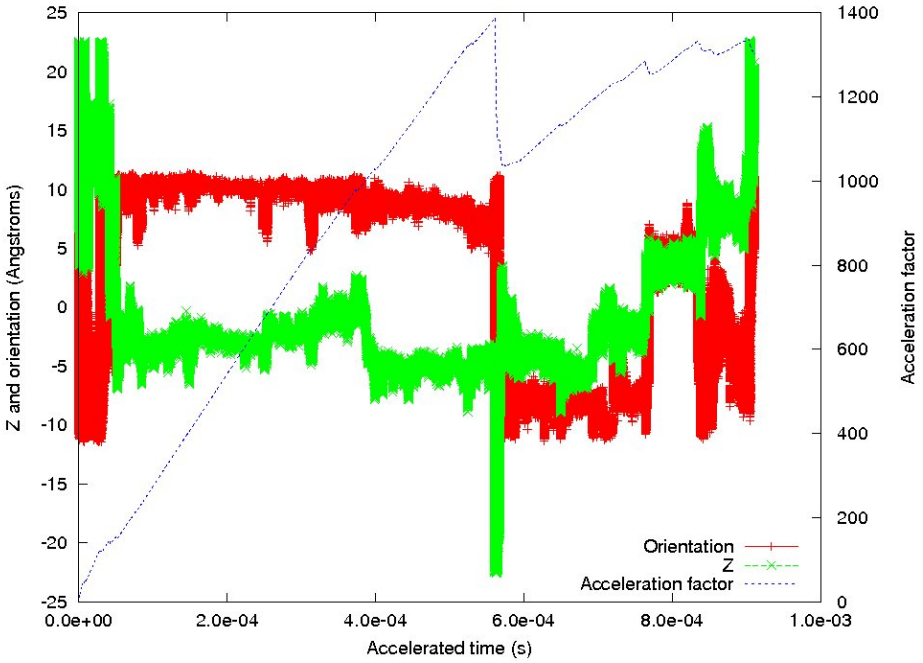

Figure 9: An example of drop in the acceleration factor for Meropenem in OmpF when a transition occurs.

To make sure if the real transition times obtained from the FMTD method are reliable, we can perform some standard statistical tests[30] to check if these times follow, as they should, a Poisson distribution. These test, to be sufficiently trusted, have to be made over a sufficiently big number of transition events (let say around 20). The fail of the test can be due typically to two factors: biasing in the transition region or wrong CVs choice. If the case is the second

one, what we observe is and overestimation of the obtained transition times, because much of the bias of WTMT simulations has not produced (or has produced only partially) an enhanced probability to visit the transition zone. For this reason we can say, as a rule of thumb, that, if we consider the faster transition times obtained with the FMTD method, these values (apart from the statistical fluctuation of the Poisson distribution of times) are more trusted than the slower ones.

2.3 TRANSLOCATION MODEL

The model used to simulate the translocation is a basically thermodynamic model that calculates the free–energy available for a particle given the geometrical properties of the particle and of the bacterial channel. Some modifications of the model are introduced, in order to take into account the vibrations of the porin.

2.3.1 *Particle probability density and average particle density*

Considering $\mathcal{N}$ identical independent particles confined in the configuration space Ω (multidimensional in general), the probability density for a particle to occupy the vicinity of the point $x \in \Omega$ is $p(x)$, with $\int_\Omega p(x)\mathrm{d}x = 1$. Given a subset $\mathcal{V} \subset \Omega$, the probability density to stay in $\mathcal{V}$ is $p_\mathcal{V} = \int_\mathcal{V} p(x)\mathrm{d}x$. Considering the independence of the particles, the probability of n particle of $\mathcal{N}$ to occupy $\mathcal{V}$ follows a binomial distribution and is

$$P^n_\mathcal{N}(\mathcal{V}) = \binom{\mathcal{N}}{n} (p_\mathcal{V})^n (1 - p_\mathcal{V})^{\mathcal{N}-n} \tag{40}$$

Considering an average number of particles $\overline{n}(\mathcal{V}) = p_\mathcal{V}\mathcal{N}$, in the hypothesis of a large number of particle that gives $\mathcal{N} \gg n$ and $\mathcal{N} \gg \overline{n}(\mathcal{V})$, we have $\binom{\mathcal{N}}{n} \approx \frac{\mathcal{N}^n}{n!}$ and the distribution in Equation 40 becomes

$$P_n(\overline{n}(\mathcal{V})) \approx \frac{(\overline{n}(\mathcal{V}))^n}{n!} e^{-\overline{n}(\mathcal{V})} \tag{41}$$

The probability to find zero or one particles in $\mathcal{V}$ is

$$P_0(\overline{n}(\mathcal{V})) = e^{-\overline{n}(\mathcal{V})} \quad \text{and} \quad P_1(\overline{n}(\mathcal{V})) = \overline{n}(\mathcal{V}) e^{-\overline{n}(\mathcal{V})}. \tag{42}$$

With the hypothesis of a subset $\mathcal{V}$ sufficiently little to assume $\overline{n}(\mathcal{V}) \ll 1$, the probabilities of Equation 42 becomes

$$\begin{aligned} P_0(\overline{n}(\mathcal{V})) &= 1 - \overline{n}(\mathcal{V}) + o(\overline{n}^2(\mathcal{V})) \quad \text{and} \\ P_1(\overline{n}(\mathcal{V})) &= \overline{n}(\mathcal{V}) + o(\overline{n}^2(\mathcal{V})). \end{aligned} \tag{43}$$

For an infinitesimal subset $\delta\mathcal{V} \subset \mathcal{V}$ with a volume $dV \equiv dx$ one obtains $\overline{n}(\delta\mathcal{V}) = c(x)dV + o(dV^2)$, where $c(x) = \mathcal{N}p(x)$ is the particle density in x. With these considerations, the probability density to find one particle in x is

$$\phi_1(x) = \frac{P_1(\overline{n}(\delta\mathcal{V}))}{dV} = c(x) \tag{44}$$

$$\int_{\mathcal{V}} \phi_1(x)\mathrm{d}x = P_1(\overline{n}(\mathcal{V})) = \overline{n}(\mathcal{V}) \tag{45}$$

2.3.2 *1 D diffusion model*

We start in our model from the one–dimensional Smoluchowski equation, with direction x corresponding to the main diffusion direction, that is the axis of the porin perpendicular to the bacterial membrane,

$$\frac{\partial\phi(x,t)}{\partial t} = \frac{\partial}{\partial x}D(x)\left(\frac{\partial\phi(x,t)}{\partial x} + \frac{\phi(x,t)}{kT}\frac{\partial U(x)}{\partial x}\right) \tag{46}$$

where $\phi(x,t)$ is the probability density, $D(x)$ the diffusion coefficient, $U(x)$ the free energy of the molecule, k the Boltzmann constant and T the absolute temperature.

In this equation the diffusion flux can be written as

$$J(x,t) = -D(x)\left(\frac{\partial\phi(x,t)}{\partial x} + \frac{\phi(x,t)}{kT}\frac{\partial U(x)}{\partial x}\right) \tag{47}$$

that leads to the continuity equation

$$\frac{\partial\phi(x,t)}{\partial t} + \frac{\partial}{\partial x}J(x,t) = 0 \tag{48}$$

2.3.3 *Steady state*

Due to the nature of our system (really little translocations through the porins) we can consider a quasi–stationary equilibrium. For this reason the probability density is time–independent and $\frac{\partial\phi(x,t)}{\partial t} = 0$. From Equation 48 we can state $\frac{\partial J(x,t)}{\partial x} = 0$ and consider $J(x,t) = J_0$ =const. In this situation the steady state probability density $\phi(x)$ satisfies the time–independent diffusion equation

$$-D(x)\left(\frac{\partial\phi(x)}{\partial x} + \frac{\phi(x)}{kT}\frac{\partial U(x)}{\partial x}\right) = J_0. \tag{49}$$

Substituting $\phi(x) = \psi(x)\exp(-U(x)/kT)$ from Equation 49 follows

$$\frac{d\psi(x)}{dx} = -J_0 \frac{\exp\left(\frac{U(x)}{kT}\right)}{D(x)} \tag{50}$$

and solving this differential equation, we get

$$\psi(x) = \psi_0 - J_0 \int_0^x \frac{\exp\left(\frac{U(x')}{kT}\right)}{D(x')} dx', \tag{51}$$

with $\psi_0 = \psi(0)$ representing an integration constant. Coming back to the probability density, we obtain

$$\phi(x) = \psi_0 e^{-\frac{U(x)}{kT}} - J_0 e^{-\frac{U(x)}{kT}} \int_0^x \frac{\exp\left(\frac{U(x')}{kT}\right)}{D(x')} dx'. \tag{52}$$

The values of J_0 and ψ_0 in Equation 52 can be calculated imposing boundary conditions to our system.

2.3.4 *Fixed concentration boundary conditions*

We suppose our bacterial channel starting at $x = 0$ and finishing at $x = L$ and we fix the molecular concentration to the value c_0 in $x = 0$ and c_L in $x = L$. Assuming a sufficiently little value of c_0 and c_L, we can consider both $\phi(x)$ and $c(x)$ as a single–particle probability density and we can assume the boundary conditions

$$\phi(0) = c_0 \quad \text{and} \quad \phi(L) = c_L \tag{53}$$

and obtain from Equation 52 the two values

$$\psi_0 = c_0 e^{\frac{U_0}{kT}} \text{ and} \tag{54}$$

$$J_0 = \frac{c_0 e^{\frac{U(0)}{kT}} - c_L e^{\frac{U(L)}{kT}}}{\int_0^L \frac{\exp\left(\frac{U(x)}{kT}\right)}{D(x)} dx}. \tag{55}$$

Substituting the ψ_0 value from Equation 54 into Equation 52 we obtain

$$\phi(x) = e^{-\frac{U(x)-U(0)}{kT}} \left(c_0 - J_0 \int_0^x \frac{\exp\left(\frac{U(x')-U(0)}{kT}\right)}{D(x')} dx' \right). \tag{56}$$

If $J_0 = 0$, we basically re–obtain the Boltzmann distribution with his famous factor, but in a quasi–equilibrium steady state ($J_0 \neq 0$),

the second term within brackets in Equation 56 give us the deviation from a Boltzmann distribution, that depends from the local diffusion too.

With our hypothesis of sufficiently little molecular concentrations the probabilities to find one (p_1) or zero (p_0) particles into the bacterial channel are respectively

$$p_1 = \int_0^L \phi(x)\mathrm{d}x \quad \text{and} \quad p_0 = 1 - p_1. \tag{57}$$

To have an *a posteriori* filling of the goodness of our hypothesis of little molecular concentrations, we have to be quite sure that the concentration guarantees a negligible probability to have two particles in the bacterial channel. Because our molecules are assumed to be in a large number, independent between them and we assume the theoretical same occupation probability in every state of the phase space, we can assume a Poisson distribution as stated in Section 2.3.1 and for this reason the probability to have two molecule contemporaneously present in the bacterial channel is $p_2 = 0.5p_1^2 \exp(-p_1) = 0.5p_1^2 + o(p_1^2)$. If our molecular concentration leads to a value of p_1 such that $p_2 \ll p_1$ and we can safely ignore p_2, we can consider $\phi(x)$ as the one–particle probability density.

Assuming the same free energy on each side of the channel ($U(0) = U(L) = U_0$) and independence of the diffusion constant on x ($D(x) = D$ =const), we can write

$$J_0 = -\frac{D \cdot \Delta c}{l(L)} \tag{58}$$

$$l(x) = \int_0^x \exp\left(\frac{U(x') - U(0)}{kT}\right)\mathrm{d}x' \tag{59}$$

with $\Delta c = c_L - c_0$. The function $l(x)$ may be called effective length of the channel and

- if $U(x) = U(0)$ =const, then $l(x) = x$,
- if $U(x) < U(0)$ (binding), then $l(x) < L$ and
- if $U(x) > U(0)$ (barrier), then $l(x) > L$.

With these hypothesis the probability density reads

$$\phi(x) = e^{-\frac{U(x)-U(0)}{kT}}\left(c_0 + \Delta c\frac{l(x)}{l(L)}\right) \tag{60}$$

and in this case we note an independence on the diffusion constant.

Assuming an uniform free–energy along all the channel ($U(x) = U_0$ =const), we get respectively from Equation 58, Equation 60 and finally Equations 57 and 62

$$J_0 = -\frac{D \cdot \Delta c}{L}, \tag{61}$$

$$\phi(x) = c_0 + \Delta c \cdot \frac{x}{L}, \tag{62}$$

$$p_1 = \frac{c_0 + c_L}{2} L. \tag{63}$$

Considering a slightly more complicated free–energy profile with a rectangular profile of magnitude ΔU and width Δl starting from the position $x = X_0$, then the effective free–energy and channel length assume the form

$$U(x) = \Delta U \cdot \text{rect}\left(\frac{x - x_0}{\Delta l}\right) \tag{64}$$

$$l(L) = e^{\frac{\Delta U}{kT}} \Delta l, \tag{65}$$

where $\text{rect}(x) = \theta(x) - \theta(x-1)$ assumes the value one for $x \in [0, 1]$, being $\theta(x)$ the Heaviside step function. We can see from Equation 65 an exponential increase or decrease of the effective length if we have a rectangular barrier ($\Delta U > 0$) or well ($\Delta U < 0$). Having a barrier composed of one rectangular well followed by one barrier

$$U(x) = -|\Delta U_-| \cdot \text{rect}\left(\frac{x - x_-}{\Delta l_-}\right) + |\Delta U_+| \cdot \text{rect}\left(\frac{x - x_+}{\Delta l_+}\right), \tag{66}$$

we obtain the following effective length

$$l(L) = e^{-\frac{|\Delta U_-|}{kT}} \Delta l_- + e^{\frac{|\Delta U_+|}{kT}} \Delta l_+. \tag{67}$$

From this equation we can see that, if $\Delta U_+ \gg kT$, the largely dominant term in the effective length is the barrier and, having a sequence of barrier, the dominant one is the one with the highest value.

2.3.5 *Models to calculate free–energy profiles*

In the following part we considerate a way to calculate, under certain hypothesis, the free–energy profile through our model. The basic assumption is a diffusion current null ($J_0 = 0$), that can be reached at equilibrium. Following Equation 56, the new probability density becomes

$$\phi(x) = e^{-\frac{U(x) - U(0)}{kT}} c_0. \tag{68}$$

Considering the different hypothesis on the joined effect of the finite channel and molecule size, we obtain different cases.

2.3.5.1 *Point–like molecule.*

Assuming a molecule size much smaller than the one of the channel, then the concentration of the molecules (and for what stated in Sections 2.3.1 and 2.3.4 the probability density too) is proportional to the cross section $S(x)$ of the channel, with

$$\phi(x) = c_V S(x), \tag{69}$$

and c_v representing the volume concentration assumed constant in the channel. By Equations 68 and 69 and assuming $c_0 = c_V S(0)$, we obtain

$$U(x) - U(0) = -kT \ln\left(\frac{S(x)}{S(0)}\right). \tag{70}$$

This last equation shows clearly the appearance of a steric barrier when the section of the channel restricts.

2.3.5.2 *Finite–sized small spherical molecule.*

If the size of the molecule is much smaller than that of the channel, but still finite, then one may model it by a ball with the average cross section S_m. In this case we replace in Equation 70 the cross section $S(x)$ with the available area, that is $A(x) = S(x) - S_m$. In this case the new free–energy will be

$$U(x) - U(0) = -kT \ln\left(\frac{A(x)}{A(0)}\right). \tag{71}$$

2.3.5.3 *Channel/molecule Gaussian size fluctuations.*

When the molecule cross section becomes comparable to the one of the channel (this is particularly true near to the constriction region assumed located in $x = x_c$), the fluctuations in the channel and molecule cross sections start to be important to allow the molecule to cross the channel.

If we assume a fluctuation in the channel cross section, with the corresponding probability density in position x given by $f_c(s_c, x)$ and the fluctuations of the molecule due to non–spherical form and conformational movements characterised by the probability density $f_m(s_m)$, then the probability to find the available area at position x in the channels becomes

$$\begin{aligned} f_a(s_a, x) &= \iint f_c(s_c, x) f_m(s_m) \delta(s_c - s_m - s_a) \mathrm{d}s_c \mathrm{d}s_m = \\ &= \int f_c(s_m + s_a, x) f_m(s_m) \mathrm{d}s_m. \end{aligned} \tag{72}$$

Since only positive values of instant available area make sense, the average available area in position x is

$$A(x) = \int_0^\infty s_a f_a(s_a, x) \mathrm{d}s_a. \tag{73}$$

Having a new available area, the steric contribution to the free–energy can still be calculated through Equation 71.

If the channel and molecular fluctuations are Gaussian and have average values of $S_c(x)$ and S_m and dispersion of $\sigma_c(x)$ and σ_m, then $f_a(s_a, x)$ is also Gaussian, with mean value $S_a(x) = S_c(x) - S_m$ and dispersion $\sigma_a^2(x) = \sigma_c^2(x) + \sigma_m^2$. The average available area is

$$A(x) = \int_0^\infty s_a \frac{\exp\left(-\frac{(s_a - S_a(x))^2}{2\sigma_a^2(x)}\right)}{\sqrt{2\pi\sigma_a^2(x)}} \mathrm{d}s_a. \tag{74}$$

If the channel is much wider than the molecule ($S_c(x) >> S_m$), then $A(x) \approx S_c(x) - S_m$ and we have the result of Section 2.3.5.2. In the other limit of a narrow channel and a thick molecule, we have $S_c(x) < S_m$ and $(S_m - S_c(x))/\sigma_a(x) \gg 1$ and we can limit the integration in Equation 74 only around zero in a small interval of the order of magnitude of $\sigma_a^2(x)/S_a(x)$ where

$$\exp\left(-\frac{(s_a - S_a(x))^2}{2\sigma_a^2(x)}\right) \approx \exp\left(-\frac{s_a|S_a(x)|}{\sigma_a^2(x)}\right) \exp\left(-\frac{S_a^2(x)}{2\sigma_a^2(x)}\right) \tag{75}$$

and we obtain the following estimate

$$A(x) \approx \frac{1}{\sqrt{2\pi}} \frac{\sigma_a^3(x)}{S_a^2(x)} \exp\left(-\frac{S_a^2(x)}{2\sigma_a^2(x)}\right) \tag{76}$$

Substituting Equation 76 in Equation 74, we obtain

$$\begin{aligned} U(x) - U(0) &\approx kT\left(\frac{\ln(2\pi)}{2} + \ln\left(\frac{S_a^2(x)}{\sigma_a^2(x)}\right) + \frac{S_a^2(x)}{2\sigma_a^2(x)} + \ln\left(\frac{S_a(0)}{\sigma_a(x)}\right)\right) \\ &\approx kT\left(1 + \frac{S_a^2(x)}{2\sigma_a^2(x)} + \ln\left(\frac{S_a(0)}{\sigma_a(x)}\right)\right), \end{aligned} \tag{77}$$

where the last approximation follows from $\ln(2\pi)/2 \approx 1$ and from the little assumed by the second logarithmic term compared to the third linear one.

2.3.5.4 *Asymmetric channel cross section fluctuations.*

If we suppose a small asymmetry of the channel cross section distribution, it can be considered in a generalised Gaussian distribution with a probability distribution function

$$a(u, \gamma, \eta) = \frac{e^{\phi(u,\gamma,\eta)}}{N(\gamma, \eta)} \tag{78}$$

with

$$\phi(u,\gamma,\eta) = -\frac{1}{2}u^2 + \frac{\gamma}{6}u^3 - \eta\frac{\gamma^2}{64}u^4 \quad \text{and} \tag{79}$$

$$N(\gamma,\eta) = \int e^{\phi(u,\gamma,\eta)}\mathrm{d}u. \tag{80}$$

The probability distribution function $a(u,\gamma,\eta)$ is an integrable function at $\eta > 0$, has a single maximum at $u = 0$ if $\eta \geq 1$ and has a concave phase ($\frac{\mathrm{d}\phi(u,\gamma,\eta)}{\mathrm{d}u} \leq 0$) if $\eta \geq 4/3$.

For small γ values, one obtains for the norm $N(\gamma,\eta)$ and for the lower raw moments valid up to the second order on γ these values

$$N(\gamma,\eta) = \sqrt{2\pi}\left(1 + \gamma^2\frac{3}{8}\left(\frac{5}{9} - \frac{\eta}{8}\right)\right) + o(\gamma^2), \tag{81}$$

$$\alpha_1 = \int a(u,\gamma,\eta)u\mathrm{d}u = \frac{\gamma}{2} + o(\gamma^2), \tag{82}$$

$$\alpha_2 = \int a(u,\gamma,\eta)u^2\mathrm{d}u = 1 + \gamma^2\frac{3}{8}\left(\frac{10}{3} - \frac{\eta}{2}\right) + o(\gamma^2), \tag{83}$$

$$\alpha_3 = \int a(u,\gamma,\eta)u^3\mathrm{d}u = \frac{5\gamma}{2} + o(\gamma^2), \tag{84}$$

$$\alpha_4 = \int a(u,\gamma,\eta)u^4\mathrm{d}u = 3 + \gamma^2\frac{3}{8}\left(\frac{100}{3} - 4\eta\right) + o(\gamma^2). \tag{85}$$

When $\eta \geq 1$, we can use $a(u,\gamma,\eta)$ as an uni–modal probability density function with a single maximum at $u = 0$ and respectively the following asymptotic expressions for the average, the variance, the skewness and the excess kurtosis

$$\overline{u} = \alpha_1 = \frac{\gamma}{2} + o(\gamma^2), \tag{86}$$

$$\sigma^2 = \overline{(u-\overline{u})^2} = 1 + \gamma^2\left(1 - \frac{3}{16}\eta\right) + o(\gamma^2), \tag{87}$$

$$\gamma_1 = \frac{\overline{(u-\overline{u})^3}}{\sigma^3} = \gamma + o(\gamma^2), \tag{88}$$

$$\gamma_2 = \frac{\overline{(u-\overline{u})^4}}{\sigma^4} - 3 = 3\gamma^2\left(1 - \frac{1}{8}\eta\right) + o(\gamma^2), \tag{89}$$

2.3.5.4.1LAW OF BIG NUMBERS CONSIDERATIONS. If the distribution of interest comes from the sum of n equivalent random variables, then

$$\gamma_1 = \frac{\gamma_1^0}{\sqrt{n}} \quad \text{and} \quad \gamma_2 = \frac{\gamma_2^0}{n}, \tag{90}$$

with γ_1^0 and γ_2^0 representing the skewness and excess kurtosis of the distribution of a single variable, respectively. Thus, γ_1 and γ_2 are small numbers at large n and one concludes

$$\frac{\gamma_2}{\gamma_1^2} = \frac{\gamma_2^0}{(\gamma_1^0)^2} = 3\left(1 - \frac{1}{8}\eta\right) + o\left(\frac{1}{n^2}\right). \tag{91}$$

Therefore, the parameter η is determined by the ratio $\frac{\gamma_2^0}{(\gamma_1^0)^2}$.

2.3.5.4.2EXTENDED DIFFUSION APPROXIMATION CONSIDERATIONS. At $\eta = 8$ we have $\gamma_2 = 0$ and therefore the forth cumulant of the distribution is zero: $\kappa_4 = 0$. This may be attributed to the extended diffusion approximation with retains the first three cumulants (κ_1, κ_2 and κ_3), but neglects all the others.

2.3.5.4.3MAXIMUM ENTROPY CONSIDERATIONS. The entropy of the distribution $a(u, \gamma, \eta)$ is

$$\begin{aligned} H(\gamma, \eta) &= -\int a(u) \ln(a(u)) \mathrm{d}u = \frac{1}{2}\alpha_2 - \frac{\gamma}{6}\alpha_3 + \frac{\eta\gamma^2}{64}\alpha_4 + \ln(N) \\ &= \frac{1}{2}\ln(2\pi e) + \gamma^2\left(\frac{25}{24} - \eta\frac{3}{16}\right) + o(\gamma^2). \end{aligned} \tag{92}$$

At the end one concludes that asymptotically, at small γ, the entropy increase linearly with decreasing η. Thus the asymptotically maximum–entropy distribution with a given skewness and concave phase ($\frac{\mathrm{d}\phi(u,\gamma,\eta)}{\mathrm{d}u} \leq 0$) is located at $\eta = 4/3$.

This distribution is

$$b(u, \gamma) = a(u, \gamma, 4/3) = \frac{e^{\phi(u,\gamma)}}{N(\gamma)}, \tag{93}$$

with the following asymptotic conditions

$$\phi(u, \gamma) = -\frac{1}{2}u^2 + \frac{\gamma}{6}u^3 - \frac{\gamma^2}{48}u^4, \tag{94}$$

$$N(\gamma) = \sqrt{2\pi}\left(1 + \frac{7}{48}\gamma^2\right) + o(\gamma^2), \tag{95}$$

$$\overline{u} = \frac{\gamma}{2} + o(\gamma^2), \tag{96}$$

$$\sigma^2 = 1 + \frac{3}{4}\gamma^2 + o(\gamma^2), \tag{97}$$

$$\gamma_1 = \gamma + o(\gamma^2), \tag{98}$$

$$\gamma_2 = \frac{5}{2}\gamma^2 + o(\gamma^2), \tag{99}$$

$$H(\gamma) = \frac{1}{2}\ln(2\pi e) + \frac{19}{24}\gamma^2 + o(\gamma^2). \tag{100}$$

Because the phase $\phi(u,\gamma)$ has a single zero of the second derivative at $u = 2/\gamma$, *i. e.* $\phi(2/\gamma,\gamma) = 0$, the distribution function $b(u)$ may be presented in an alternative form

$$\begin{aligned} b(u,\gamma) &= -\frac{e^{-\frac{1}{\gamma^2}-\frac{1}{3}\frac{2}{\gamma}(u-\frac{2}{\gamma})-\frac{\gamma^2}{48}(u-\frac{2}{\gamma})^4}}{N(\gamma)} \\ &= -\frac{e^{-\frac{1}{\gamma^2}-\frac{1}{3}\frac{2}{\gamma}(u-\frac{2}{\gamma})}}{N(\gamma)}\left(1+O\left(\left(u-\frac{2}{\gamma}\right)^4\right)\right). \end{aligned} \tag{101}$$

From last equation we conclude that

- function $b(u,\gamma)$ has an exponential long tail at $u \approx 2/\gamma$,
- at $u > 2/\gamma$ the function $b(u)$ decays as the forth power in the exponent ($b \sim e^{-\gamma^2 u^4/48}$) and
- around the maximum $u \approx 0$ the function $b(u)$ is Gaussian.

Finally, the asymmetric distribution of the channel cross section area may be expressed as

$$f(s_c,x) = \frac{1}{\sigma_0(x)}b\left(\frac{s_c - S_0(x)}{\sigma_0(x)},\gamma(x)\right), \tag{102}$$

where the parameters may be calculated by the measured average $S_c(x)$, the variance $\sigma_c^2(x)$ and the skewness $\gamma_1(x)$, by using the following formulas

$$S_0(x) = S_c(x) - \frac{\gamma_1(x)\sigma_c(x)}{2}, \tag{103}$$

$$\sigma_0^2(x) = \frac{\sigma_c^2(x)}{1+\frac{3}{4}\gamma_1^2(x)}, \tag{104}$$

$$\gamma(x) = \gamma_1(x). \tag{105}$$

Note that the above relations assume small γ values. It is probably reasonable to assume $\gamma \leq 0.5$. If $\gamma_1(x) > 0.5$ for some x values, then one may probably fix $\gamma = 0.5$ having probably still a better result than with a pure Gaussian.

Part II

RESULTS, DISCUSSION AND CONCLUSIONS

3

RESULTS

Contents

3.1 MICROSECOND AND SUB–MICROSECOND CHANNEL GATING

As a case study of the methods described in Section 2.1 we analysed the electrophysiology data referred to Meropenem in OmpF. These data are particularly suitable to check the post–elaboration method, because the classical methods fail with the couple Meropenem/OmpF because the blockages are under their resolution.

3.1.1 *Experimental configuration*

We used data obtained from an experimental configuration having a 1M KCl concentration, a pH=6 and a -50mV voltage. The filter applied is a 4–pole (4–order) low–pass Bessel filter at cut–off frequency

Set ID	c (mM)	f_s(kHz)
1	0;2.5;5.0;7.5;9.0	50
2	0;2.5;5.0;7.5;10.0	50
3	0;5.0;10.0;12.5;15.0;20.0;23.0	50,250
4	0;5.0;10.0;12.5;15.0;20.0;24.0	50,250
5	0;5.0;10.0;12.5;15.0;20.0;24.0	50,250

Table 1: List of parameters (ID for the set, concentration c of Meropenem and sampling frequency f_s) used for the five independent measures of Meropenem in OmpF.

$f_c = 10$ kHz ($\omega_0 \approx \pi f_c$) with the following absolute square of the transfer function

$$|H_a(\omega)|^2 = \frac{11025}{\left(\frac{\omega}{\omega_0}\right)^8 + 10\left(\frac{\omega}{\omega_0}\right)^6 + 135\left(\frac{\omega}{\omega_0}\right)^4 + 1575\left(\frac{\omega}{\omega_0}\right)^2 + 11025} \tag{106}$$

There was five set of measures described in Table 1. With different values of OmpF concentration c and sampling frequency f_s.

3.1.2 *Results*

The available sets of data described in Table 1 were divided in three independent data records named I, II and III. In the case of $f_s = 50$ kHz each record lasts 5 seconds, when $f_s = 250$ kHz each record lasts 4 seconds for a total number of values respectively of $2.5 \cdot 10^5$ and $1.0 \cdot 10^6$ events respectively.

If we analyse the behaviour of the current histogram when Meropenem is inserted, we can observe (Figure 10) the absence of a secondary peak that sometimes happens around the value of the dropped current (monomer occupied by the drug).

What can be seen is only a widening of the current distribution, but not evidence of the appearance of a second peak can be stated at first glance.

The autocorrelation function is calculated for τ up to $\tau_{max} = 20$ ms for $f_s = 50$ kHz and up to $\tau_{max} = 4$ ms for $f_s = 250$ kHz.

If we calculate $G_s(f)$, that is the signal PSD, subtracting from the PSD of the signal the one of the background (signal recorded before inserting the antibiotic, *i. e.* for $c = 0$), using the data set number 3 in Table 1 we obtain the results shown for data records I, II and III and for $f_c = 50$ kHz and $f_c = 250$ kHz in Figure 11 panel A and B.

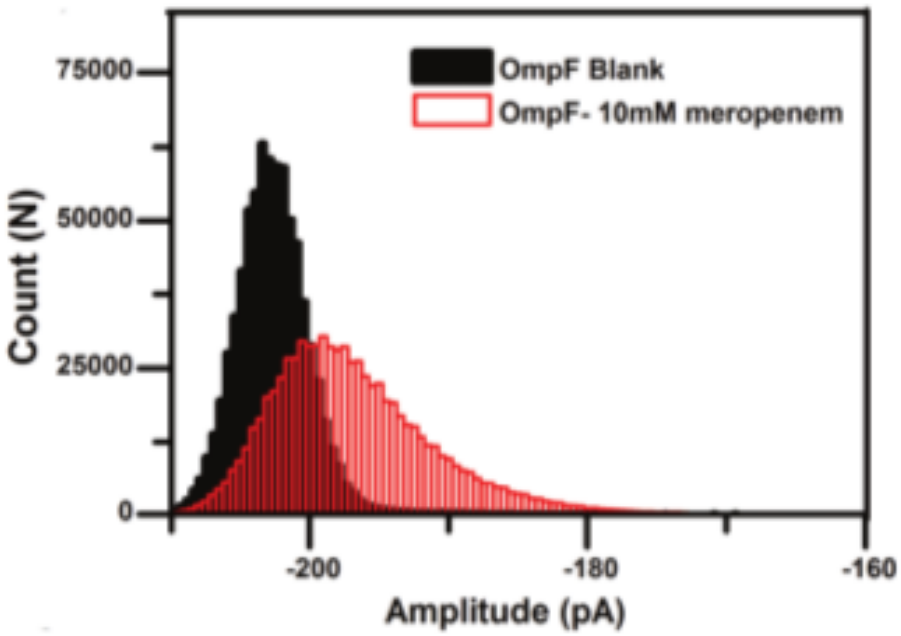

Figure 10: Histogram of the current distribution with or without a 10 mM Meropenem concentration.

In Figure 11 we can see the PSD of each of the three records. If we fit these data with respect to the PSD in Equation 8, using the range of $f \in [1, 25]$ kHz to avoid the large fluctuations observed below 1 kHz, we can obtain only an upper limit for τ_c (respectively $\tau_c < 0.01$ ms for $f_c = 50$ kHz and $\tau_c < 0.005$ ms for $f_c = 250$ kHz). For this reason, if we fix $\tau_c = 0$, we obtain from the fitting procedure $a = 2.6$ pA^2 kHz^{-1} and $a = 2.7$ pA^2 kHz^{-1} in the two cases, with a comparable dispersion for the records I, II and III. From the figure we can see how the sampling filtering assume a littler weight when the sampling frequency increases, because $|H_s(\omega)|^2 = \left(\frac{\sin(0.5\omega\Delta t)}{0.5\omega\Delta t}\right)^2$ with Δt representing the sampling interval. In the panel B is shown also the result we would have without applying any filter. We can see immediately a difference with the real values in the figure. In this case (we calculate this values only to compare them with the ones obtained with the new procedure) the fitted parameters are $a = 3.5$ pA^2 kHz^{-1} and $\tau_c = 0,03$ ms.

In Figure 12 we can see a pretty linear trend of a and Δj as a function of c (the error bars shown are the maximum between the fitting error and the error due to averaging between the three records I, II and III). For this reason we can assume the approximations in Equation 6 and 7 (Section 2.1) are good and we can apply the method number 3 described in Section 2.1.3.3 to the set number 3 in Table 1.

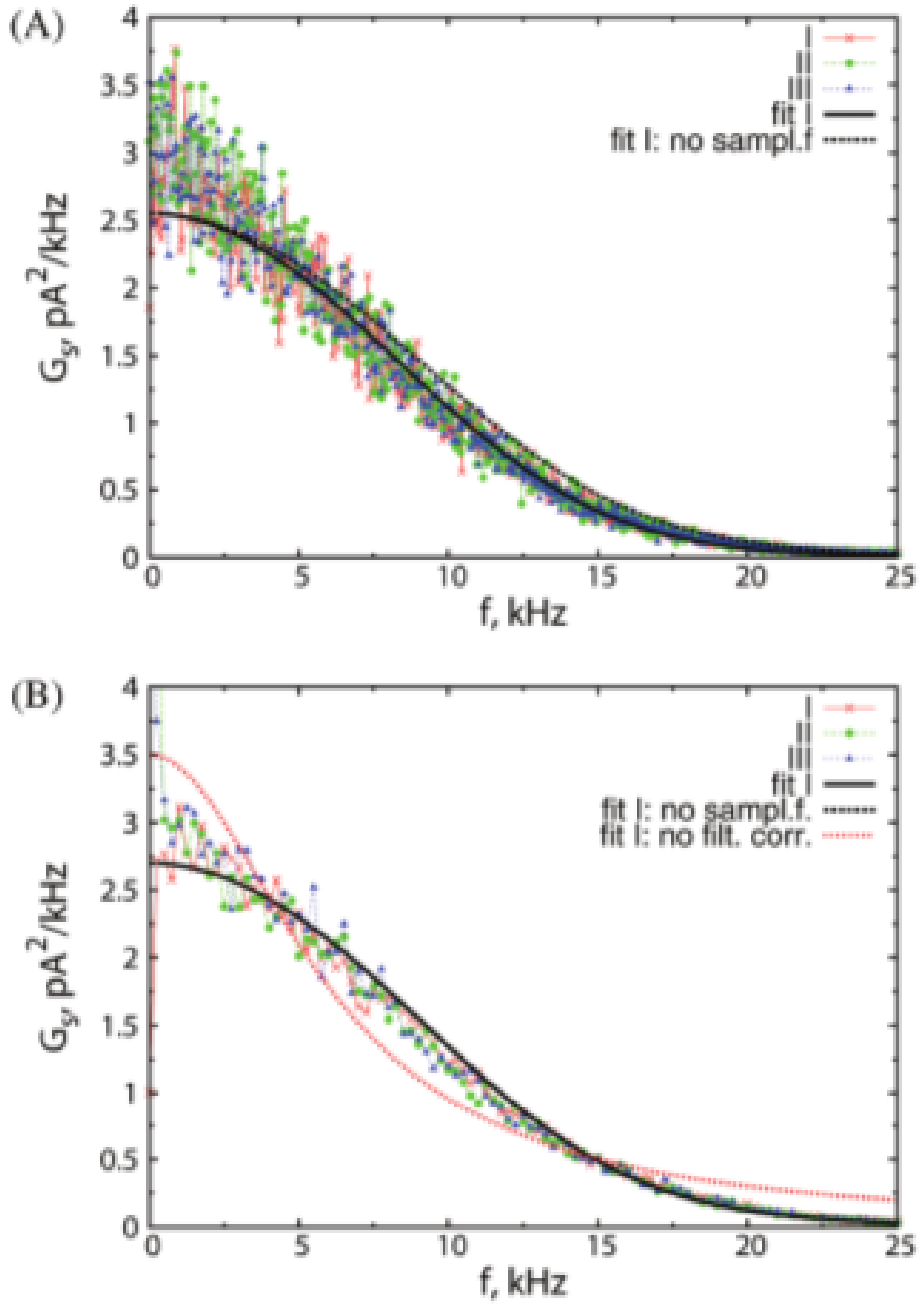

Figure 11: PSD of the ion current fluctuations sampled (A) at 50 kHz and (B) at 250 kHz. The thick solid lines represent the fitting obtained using the formula in Equation 8. The black dotted line is the fitting calculated without the sampling filtering and, only for panel B, the red dotted line is the result of the fitting computed without any filter. The data set used is the number 3 in Table 1 for $c = 10$ mM.

The fitted values are displayed in Table 2 and the kinetic parameters

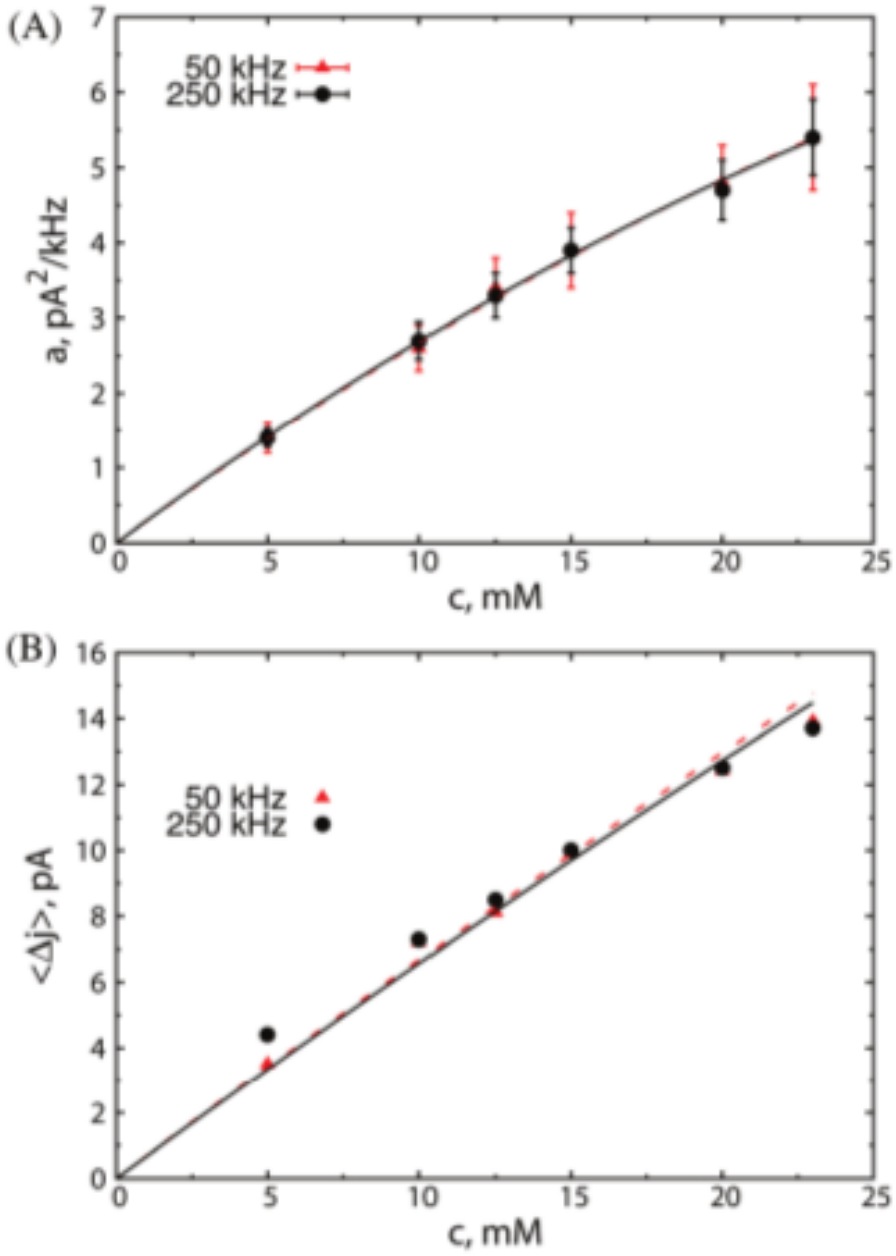

Figure 12: Lorentzian factor (panel A) and average ion current shift (panel B) as a function of Meropenem concentration. In every panel the red triangles refer to the 50 kHz data and the black circles to the 250 kHz data. The fitting parameters are obtained using the method number 3 in Section 2.1.3.3 and the results are shown in Table 2. The data set used is the number 3 in Table 1 for $c = 10$ mM.

in Table 3. Finally the same kinetic parameters are shown for all the sets in Table 4, but only for the sampling at 50 kHz.

$f_s(kHz)$	b_a	r	b_j
50	0.293 (0.019)	0.0029 (0.0011)	0.688 (0.015)
250	0.297 (0.013)	0.0031 (0.0008)	0.678 (0.020)

Table 2: Fitting parameters for ion–current fluctuations (data set 3 in Table 1) obtained with the method 3 in Section 2.1.3.3. In round brackets are displayed absolute error estimates.

$f_s(kHz)$	$\Delta j\ (pA)$	$k_{off}\ (ms^{-1})$	$k_{on}\ (ms^{-1}M^{-1})$
50	80 (30)	700 (250)	2030 (140)
250	72 (16)	660 (150)	2060 (140)

Table 3: Kinetic parameters for ion–current fluctuations (data set 3 in Table 1) obtained with the method 3 in Section 2.1.3.3. In round brackets are displayed absolute error estimates.

Set ID	$\Delta j\ (pA)$	$k_{off}\ (ms^{-1})$	$k_{on}\ (ms^{-1}M^{-1})$
1	NA	NA	800 (500)
2	NA	NA	800 (400)
3	80 (30)	700 (250)	2030 (140)
4	140 (70)	1200 (600)	1500 (100)
5	80 (20)	750 (200)	2030 (140)

Table 4: Kinetic parameters for ion–current fluctuations (data set 3 in Table 1) obtained with the method 3 in Section 2.1.3.3. In round brackets are displayed absolute error estimates. All data set are sampled at 50 kHz.

Set ID	Δj (pA)	k_{off} (ms^{-1})	k_{on} ($ms^{-1}M^{-1}$)	K (M^{-1})	K_{cp} (M^{-1})
1	73 (2)	370 (100)	800 (400)	2.2	NA
2	70 (2)	340 (80)	800 (400)	2.3	NA
3	66 (2)	600 (40)	2030 (120)	3.4	3.6
4	66 (2)	540 (40)	1430 (110)	2.6	2.1
5	68 (2)	580 (50)	2000 (200)	3.4	3.3

Table 5: Kinetic parameters for ion–current fluctuations (data set 3 in Table 1) obtained with the method 4 in Section 2.1.3.3. In round brackets are displayed absolute error estimates. All data set are sampled at 50 kHz and K_{cp} are the equilibrium binding constants obtained from the conductance probe method[15].

From Table 4 we can see that kinetics parameters for sets 1 and 2 are not available. This is due to the big uncertainty (almost 100%) in fitting parameters. Using the results of sets 3, 4 and 5 for Δj one can infer that, within the uncertainty error, Δj equals 1/3 of the open state trimer ion current, supposing each monomer blocks completely the ionic current. Taking this assumptions we can use the method number 4 described in Section 2.1.3.4 and re–analyse all data sets of Table 1. The results are listed in Table 5. In this table we can see how from the results obtained it is possible to get equilibrium binding constants K in good agreement with the one obtainable with the conductance probe method[15].

Within this thesis we are mostly interested on the residence times τ_{res} obtainable from this method. Using the data sampled with a frequency of 50 kHz or 250 kHz, we obtain in both cases $\tau_{res} = (1.5 \pm 0.4)\ \mu$s, that is around 1/50 of the one obtainable with standard current analysis[19].

3.1.3 *Sub–microsecond events.*

Using the method in Section 2.1.4 and in particular basing our calculations on Equation 35, we already saw that for a typical 4–order Bessel filter with 10 kHz cutoff frequency, assuming 3 active channels, $cr = 0.1$, $\Delta j = 50$ pA and $\sigma^2_{op} = 10\ \text{pA}^2$, then the detectable k_{off} may be as high as 2.5×10^6 1/s, giving a residence time τ_{res}=400 ns. This result is important because it give us, with a standard electrophysiology equipment, one detectable time more than two orders of magnitude littler than f_c^{-1}.

ID	N	$\langle\tau\rangle$ (s)	$\frac{\langle\tau\rangle\cdot\ln 2}{\tau_{median}}$	$\frac{\langle\tau\rangle}{\sigma}$	p–value
Oxygen1	50	$5.51\cdot10^{-10}$	0.968	0.989	0.789
Oxygen2	50	$2.61\cdot10^{-9}$	1.07	1.09	0.988
Water1	29	$2.15\cdot10^{-10}$	0.693	1.68	0.0679
Water2	29	$1.15\cdot10^{-8}$	0.772	1.42	0.0119

Table 6: Results of statistical analysis of time series from different exits from the HOD catalytic box. ID represents the code identifying every series of simulations. N, $\langle\tau\rangle$, τ_{median} and σ represent respectively the number, the arithmetic average, the median and the standard deviation of data composing every series. Finally p–value is referred to a KS test to check the correspondence of every series with a Poisson distribution.

To make sure about this method, we applied it again to the same set of measures as in Section 3.1.2 obtaining the same results. This confirms the goodness of this method that, moreover, can easily give a really good lower limit for detectable times.

3.2 TRANSITION TIMES FROM METADYNAMICS

To check the method[7,30] described in Section 2.2.1 we performed some simulations. We wanted first to have a feeling on the presence of events classifiable as Poissonian in a biological system and know with which level of confidence the KS test can be passed. For this reason we started analysing the times obtained in two events occurring in times allowing to be simulated through classical MD simulations: the case studies are the exit time of a molecule of Oxygen first and water later from a catalytic box located in the inner part of the protein 1–H–3–hydroxy–4–oxoquinaldine–2,4–dioxygenase (HOD)[37]. In Figure 13 is shown a representation of HOD and some residues of the catalytic box.

3.2.1 *Exit times from the catalytic box in HOD*

In the case of oxygen and water exit from the HOD catalytic box we performed two series of fifty and twenty–nine classical MD simulations respectively and recorded the times. The simulations started with different initial conditions for the oxygen and the water. Afterwards we applied the KS test obtaining the results listed in Table 6.

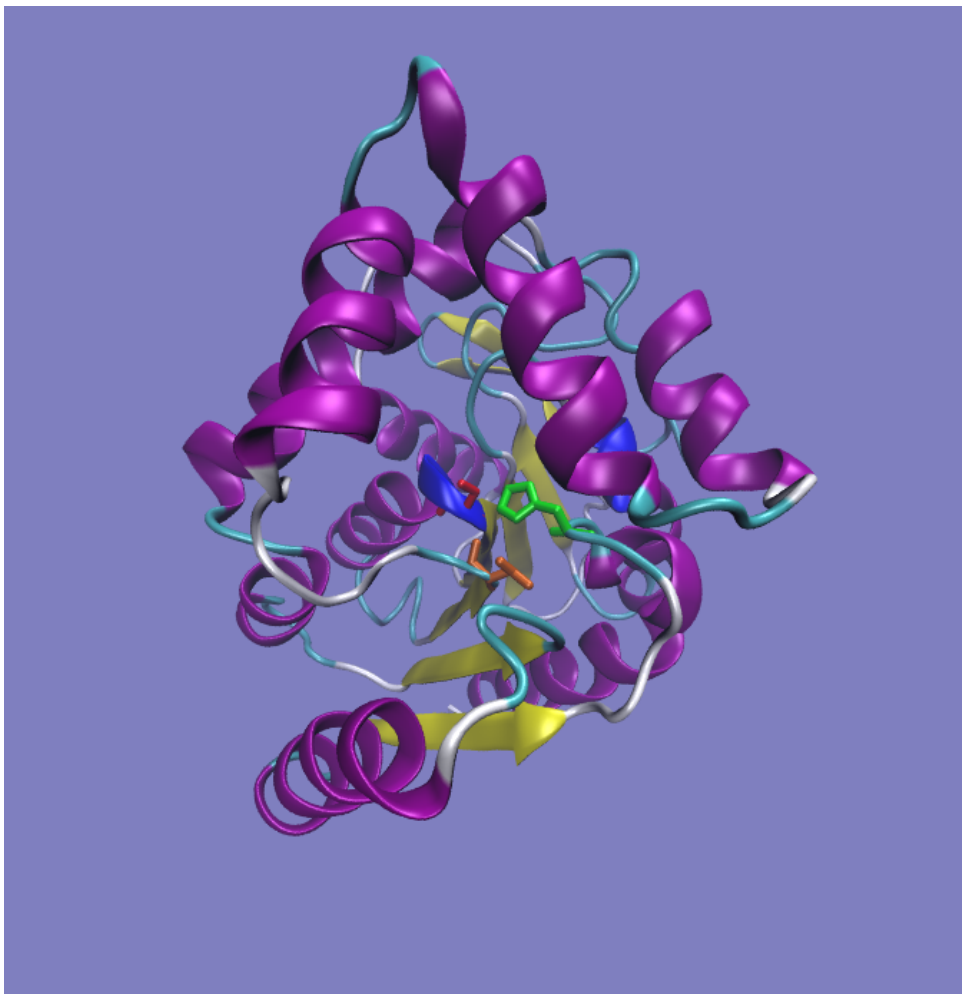

Figure 13: Representation of the protein HOD (cartoon representation) and the catalytic triad in its inner part (Histidine 251 in green, Aspartic acid 126 in orange and Ser 101 in red).

From table 6 we can see that despite the apparent similarity of the system to be simulated in the cases named Oxygen1 and Oxygen2 all the test are passed (with really good results in the case Oxygen2), but, in practice[1], not one case involving water passed the statistical tests, despite the relatively simple event analysed (water leaving a reactive box).

These results let us understand how complex can be a complete translocation of an antibiotic through a porin. During the translocation an antibiotic often meets very different conditions in terms of electric fields, available space, flexible loops and so on, that can strongly influence with various effects the translocation seen as a

1 The case Water1 has a p–value of 0.0679. This value is bigger than the minimum admissible value of 0.05, but, considering the low p–value obtained by Water2, we can say that both cases did not pass the KS test.

whole. For this reason it is quite obvious to do not expect to see very good results in statistical tests applied to translocation times obtained with the method FMTD (Section 2.2.1).

3.2.2 *Translocations of Meropenem and Meropenenum through OmpF*

To test the applicability of the FMTD method to translocations, we simulated the transit of antibiotics (Meropenem and Meropenenum) through bacterial channels (OmpF) inserted in a lipid bilayer (Figure 4) in a explicit solvent simulation. The system is simulated with the ACEMD[34] software with a PLUMED plugin version 1.3[35], the interactions are based on the AMBER 99SB–ILDN force field[36] for the OmpF, tip3p for water[38] and GAFF for lipids and substrates[39]. Visualization and plotting of molecular interactions, as well as measurement of the most stable interactions, are done with VMD software[40].

We chose to simulate Meropenem and Meropenenum because they are isomers differing only for their geometrical position of elements. Despite the little differences between them (Figure 14) they show a big clinical difference with Meropenem, which is commonly used to treat bacterial infections, and with Meropenenum, that shows instead a really low antibacterial effectiveness.

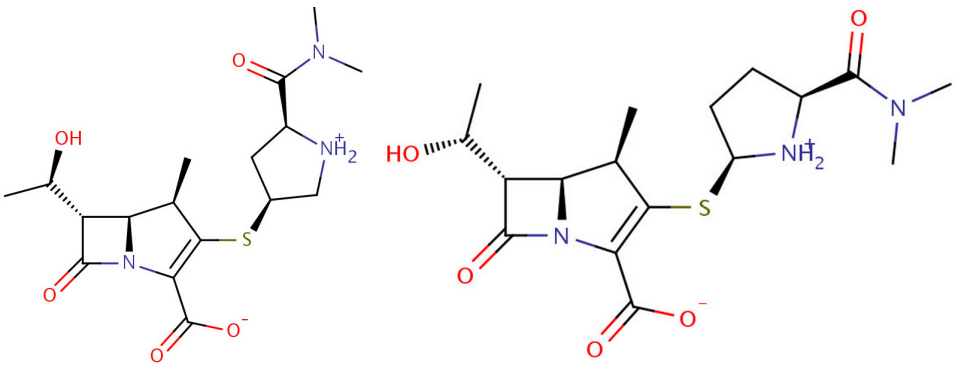

Figure 14: Comparison of the chemical structure of Meropenem and Meropenenum displaying only little geometrical differences between them. Note the dipole momentum of the two drugs.

The target of these antibiotics is located in both cases into the bacterium, so we expect a possible difference between the permeability of Meropenem and Meropenenum through OmpF to explain their different clinical effectiveness. A difference in permeability is directly related to a difference in the times needed for these drugs to enter the porin (in this case OmpF) and cross it. Because of almost identical electrical properties of these two drugs, we guess a similar

Drug	E_{bias} (kcal/mol)	f_{bias} (ps)	T_{bias} (K)	T_{sim} (K)	N
Meropenem	50	5	3000	300	17
Meropenenum	50	5	3000	300	17

Table 7: List of the parameters which have been used in WTMT simulations for Meropenem and Meropenenum.

rate of entrance into the porin[28]. For this reason we simulated the transport starting with the molecules already inside the porin, in the extracellular part, with a wall impeding them to exit again from OmpF. In MT simulations we biased two CVs. The first biased CV is the difference between the position of the center of mass (CM) of Meropenem/Meropenenum and the CM of the monomer where the antibiotics were captured (in both cases always the monomer number one). To take into account the dipole moment of the drugs, the second CV biased is the projection along the Z axis (parallel to the OmpF translocation direction) of the distance between the two atoms displayed (only for Meropenenum, because the atoms are the same also for Meropenem) in Figure 15.

The parameters that have been used in the simulation are summarised in Table 7.

3.2.2.1 *Meropenenum*

The KS test was initially not passed for all the 17 simulations of Meropenenum. To explain this difference we noted a correlation between the anomalously big times and the distance between the ASP121CG and TYR124OH atom in OmpF. Namely all the 10 faster translocations have a distance ASP121CG–TYR124OH around 4 Å and with little fluctuations when Meropenenum crosses the constriction region (this is the bottleneck in the transport process) and between the remaining 7 slower translocation show a much wider fluctuation around 5 Å for the same distance ASP121CG–TYR124O (Figure 16).

We believe this regularity can not be fortuitous, because, as explained in Section 3.2, a wrong CV for a trajectory will produce a perturbation in the Hamiltonian not corresponding to a real help to translocate. For this reason the acceleration factor calculated with the FMTD method will be bigger and will produce an anomalously big time. Instead little times are more likely to be correctly helped by the chosen bias. To confirm this idea we performed the KS test to the series of all the times and the series obtained selecting only the

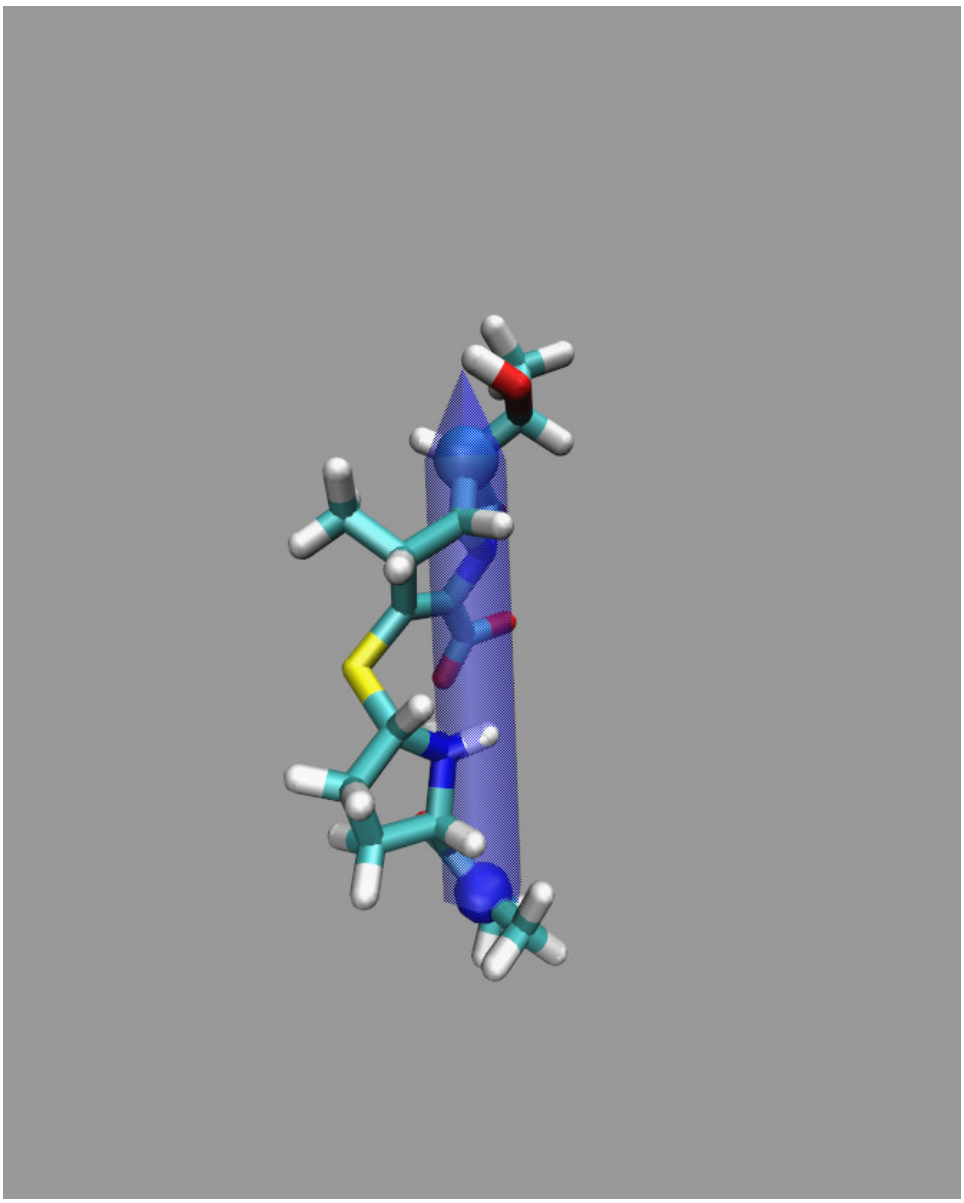

Figure 15: Representation of Meropenenum with a superimposed oriented distance between the two atoms (displayed as big spheres) chosen in the second CV biased in MT simulations.

transition occurring in a configuration "fast" referred to Figure 16. The results of the whole statistical tests[30] are shown in Table 8).

As we can see from these results, if we consider the whole set of simulated times, not one statistical test is passed, while considering only the trajectories showing the "fast" behaviour of the distance ASP121CG–TYR124OH, as described in Figure 16, every statistical test can be considered passed, in spite of the little number of data, considering the complexity of the simulated system. Based on these consideration we can say that the two CVs, which are used for bias-

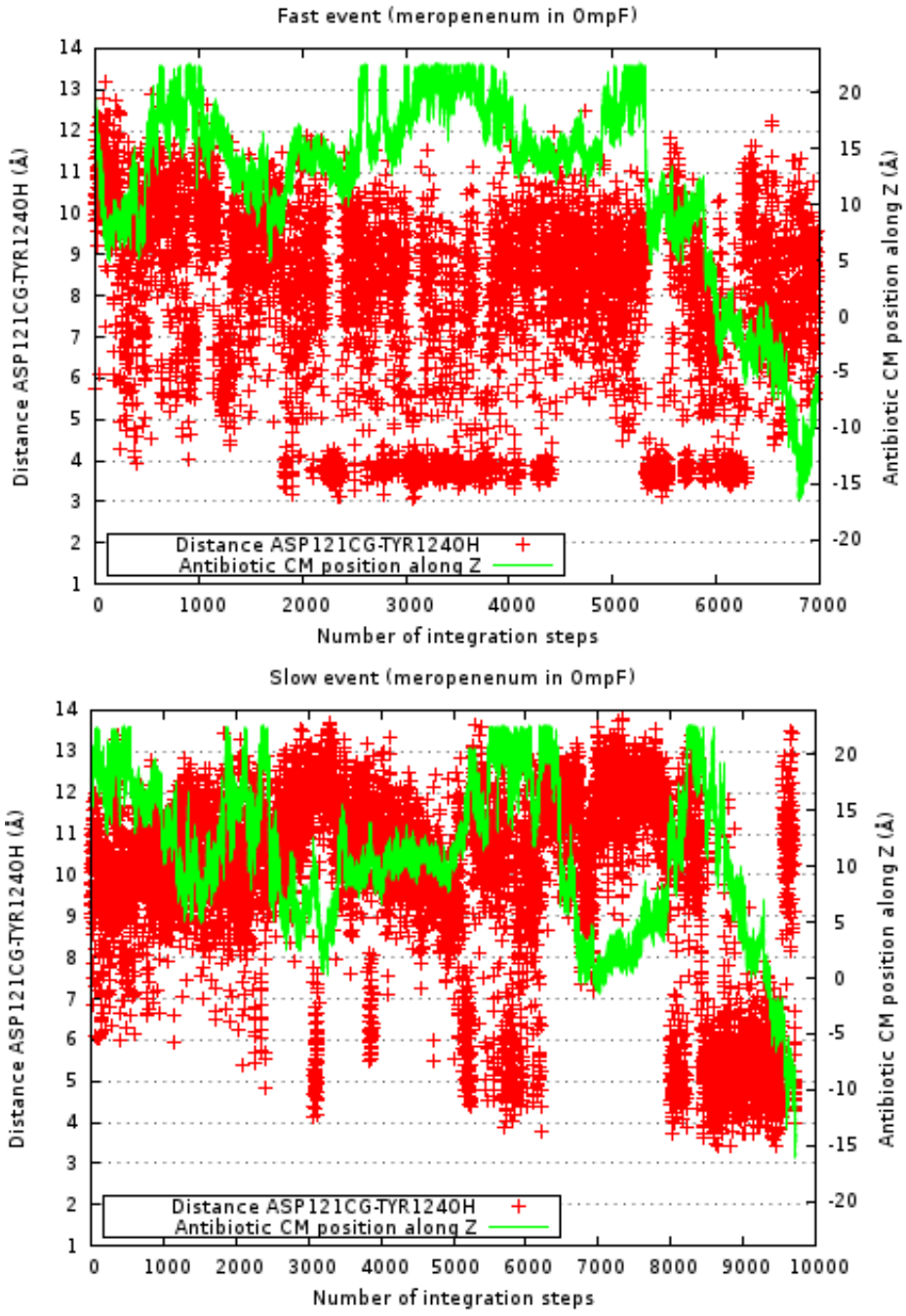

Figure 16: Comparison of the distance between the atom ASP121CG and TYR124OH during a fast ($\sim 10^{-4}$ s) and slow ($\sim 10^{-1}$ s) translocation event of Meropenenum through OmpF. Note the different behaviour of this distance when the antibiotic goes quickly to negative values crossing the porin.

ing in MT simulations, are relevant for Meropenenum to let it cross the OmpF when the system shows a "fast" behaviour of the distance ASP121CG–TYR124O. For sure we can not state that these two CVs

ID	N	$\langle\tau\rangle$ (s)	$\frac{\langle\tau\rangle\cdot\ln 2}{\tau_{median}}$	$\frac{\langle\tau\rangle}{\sigma}$	p–value
Meropenenum (all)	17	$2.6\cdot10^{-1}$	84.5	0.367	0.000
Meropenenum ("fast")	11	$1.9\cdot10^{-3}$	0.769	0.831	0.660

Table 8: Results of statistical analysis of time series transition times in the case of Meropenenum in OmpF. The results obtained using all the trajectories do not show a Poissonian behaviour, but using only the "fast" translocations (*i. e.* the events showing a "fast" distance ASP121CG–TYR124OH as depicted in Figure 16) statistical tests are passed with good results.

ID	N	$\langle\tau\rangle$ (s)	$\frac{\langle\tau\rangle\cdot\ln 2}{\tau_{median}}$	$\frac{\langle\tau\rangle}{\sigma}$	p–value
Meropenem (all)	17	$2.6\cdot10^{-2}$	19.8	0.293	0.000
Meropenem*	15	$2.9\cdot10^{-3}$	2.36	0.421	0.080

Table 9: Results of statistical analysis of translocation times obtained with the FMTD method. Considering all simulations the test is not passed. To pass the test, but still with really poor results, we must eliminate arbitrarily the two slowest translocation times (results shown in Meropenem*).

are the only ones influencing the translocation in this case, but from results they seem to be the most relevant. In the general case ("slow" events) these CVs play a role, but some other CVs, not considered, influence the translocation and determine the increased biasing in the Hamiltonian and the consequent anomalously big times.

3.2.2.2 *Meropenem*

Simulating Meropenem we performed 17 simulations, but we did not observe a correlation between the distance ASP121CG–TYR124O and the time obtained with the FMTD method (namely sorting the events by time it is not possible to note any correlation between transition times and distance ASP121CG–TYR124O). More or less all the simulations show always a similar behaviour with a distance between these two atoms typically littler than in the case of Meropenenum and having two states: around 5 Å and around 7 Å (see Figure 17).

Applying the FMTD method to all 17 simulations we obtained the results listed in Table 9.

These last results show statistical tests barely passed eliminating arbitrarily some of the slowest translocation times. The CVs cho-

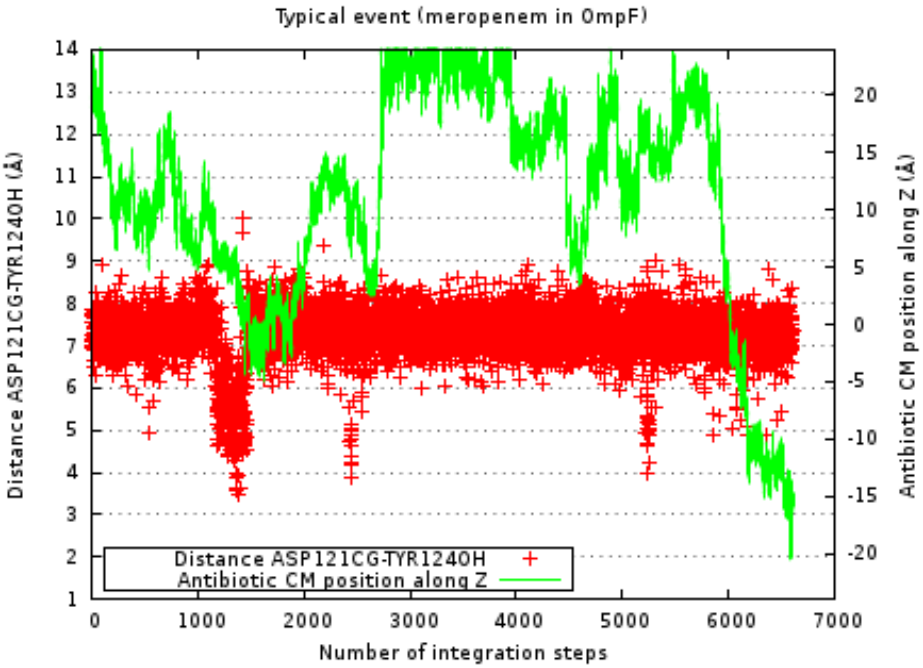

Figure 17: A typical behaviour of the distance between the atoms ASP121CG and TYR124OH during a translocation. This behaviour can be observed both in trajectories having transition times locate before or after the 50th percentile of translocation times. There is not evidence of a correlation between the distance ASP121CG–TYR124O and the translocation time.

sen seem to have some influence in translocation, because they still show times roughly distributed as a Poissonian, but for sure some relevant CVs are missed and cause the poor results in our tests. As before it seems that our CVs accumulate biasing in the Hamiltonian without producing a real help to translocate. With the FMTD method this case produce apparently big transition times realistically having not a correspondence in a translocation occurring in nature.

3.2.2.3 *Meropenem vs. Meropenenum*

Coming back to our original goal, that is to justify the different effectiveness of the considered antibiotics through eventually faster translocations in Meropenem than in Meropenenum as described in Section 3.2.2, we can compare the average time obtained using all the times obtained with Meropenem and Meropenenum (see Table 9 and 8). We can see at first glance a much littler average translocation time in Meropenem ($2.6 \cdot 10^{-2}$ s) compared to Meropenenum ($2.6 \cdot 10^{-1}$ s).

Meropenem (s)	Meropenenum (s)
$4.8 \cdot 10^{-7}$	$7.2 \cdot 10^{-5}$
$1.1 \cdot 10^{-5}$	$2.5 \cdot 10^{-4}$
$4.6 \cdot 10^{-5}$	$5.6 \cdot 10^{-4}$
$9.9 \cdot 10^{-5}$	$8.8 \cdot 10^{-4}$
$1.4 \cdot 10^{-4}$	$1.1 \cdot 10^{-3}$
$1.9 \cdot 10^{-4}$	$1.7 \cdot 10^{-3}$
$3.7 \cdot 10^{-4}$	$1.7 \cdot 10^{-3}$
$8.6 \cdot 10^{-4}$	$1.8 \cdot 10^{-3}$
$9.0 \cdot 10^{-4}$	$2.2 \cdot 10^{-3}$
$1.1 \cdot 10^{-3}$	$2.3 \cdot 10^{-3}$
$1.5 \cdot 10^{-3}$	$5.9 \cdot 10^{-3}$
$1.8 \cdot 10^{-3}$	$6.9 \cdot 10^{-3}$
$3.9 \cdot 10^{-3}$	$8.4 \cdot 10^{-3}$
$5.7 \cdot 10^{-3}$	$5.5 \cdot 10^{-2}$
$2.7 \cdot 10^{-2}$	$3.5 \cdot 10^{-1}$
$2.9 \cdot 10^{-2}$	1.3
$3.6 \cdot 10^{-1}$	2.8

Table 10: Comparison of sorted translocation times obtained with the FMTD method for Meropenem and Meropenenum. See the systematically shorter times of Meropenem with respect to Meropenenum. This trend is particularly clear for events littler than the median.

Moreover, sorting the time series obtained with the FMTD method for our two antibiotics and comparing them *one by one* with the element in the same ranking (Table 10) we can stress systematically shorter times for Meropenem with respect to Meropenenum, with a particular evidence roughly for times shorter than median. If we consider, as already stated in Section 3.2, that the FMTD method, practically, never underestimate the transition times, we can argue that, at least for the faster trajectories, where no extra bias seems to be deposited in MT, Meropenem result always faster than Meropenenum. We can connect this difference with the bigger clinical effectiveness of Meropenem, thanks to the argumentation given in Section 3.2.2. Considering the faster events recorded for Meropenem, they can be considered as possible real translocation times impossible to detect in electrophysiology without the post–elaboration method described in Section 2.1.

Summarising, the CVs which have been used in our MT simulations seems to play a role in translocation, but, given the complexity of the systems simulated, they can not give a full explanation of the whole translocation process except in some particular case (*e. g.* "fast" events in Meropenenum). Nevertheless it is still possible to have good hints of different effectiveness from the translocation times of Meropenem and Meropenenum as described in this Section. Thanks to these preliminary studies, we addressed ourselves to a different approach to the translocation events, similar to the MSM[41] one. Highlighting the presence of a series of metastable states in FESs obtained from translocation MD simulations and calculating through the FMTD method the transition times between these metastable states, it is possible[31] to have a more flexible and detailed picture of the whole transport process.

3.3 VERIFICATION OF A TRANSLOCATION MODEL

To check our model we focused mainly on Section 2.3.5. In particular we want to compare the free–energy obtained from the model with the ones obtained from WTMT simulations. This approach leads to evaluate the goodness of the hypothesis of our model, but, as the molecule to translocate becomes more complex, arise doubts about the reliability of WTMT simulations as a "true" free–energy, due to the problems evidenced in Section 3.2. The basic formula to be used is Equation 71 and for this reason we need to estimate the available area for the molecule, that is $A(z) = S(z) - S_m$, with $S(z)$ and S_m representing respectively the cross section of the channel and of the molecule studied.

3.3.1 *Estimation of cross section of OmpF*

In our case we consider as a channel the OmpF bacterial porin. We calculated (with a resolution of 0.5 Å for z) the cross section $S(z)$ numerically estimating for every monomer composing the OmpF trimer the free space available. We did that (for details about simulation see *e. g.*[42]) simulating via classical MD a system containing OmpF with explicit solvent and a lipid bilayer as bacterial membrane and considering the average free space in the inner part of a right circular cylinder having its axis parallel to the axis through the center of the monomer and perpendicular to the bacterial membrane. We varied the radius of this cylinder increasing it by one Angstrom from 14 Å to 18 Å. What we get is the general result depicted in Figure 18. From this figure we note a general increase of

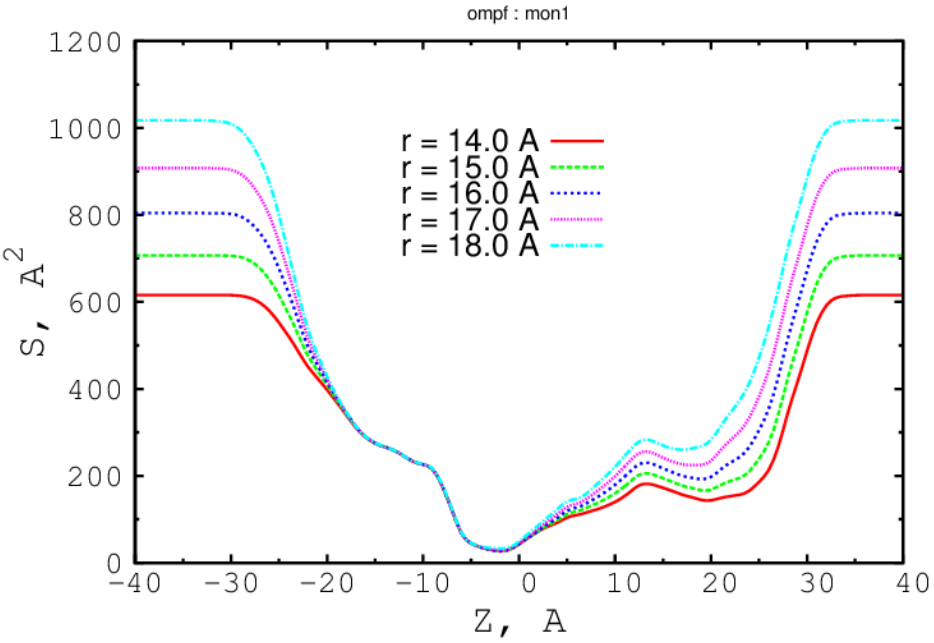

Figure 18: Average free cross section in first OmpF monomer for different radii of the used cylinder in the first monomer.

the available cross section when the cylinder radius becomes larger and brings more space in the larger zones of OmpF near the its outer part. This should suggest a even larger radius to take in account the whole free cross section even in the outer parts, but if we have a look in a detail on the available area close to the constriction region (Figure 19), we can note an increase of the cross section for cylinder radius bigger than 16 Å. This phenomenon have a strong influence in translocations because the constriction region is the one with bigger steric increase of free–energy. The explanation of this event resides

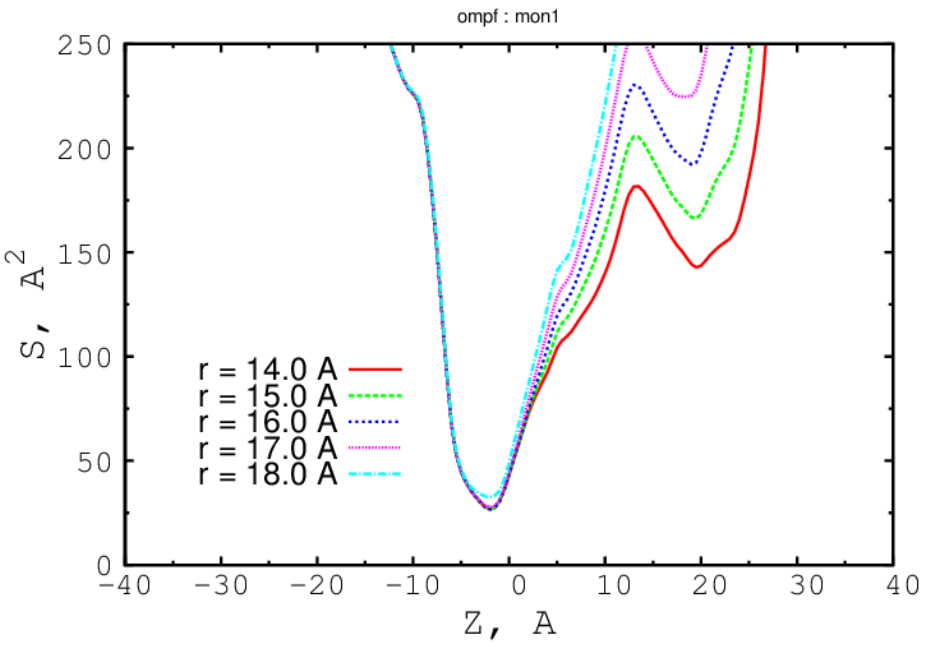

Figure 19: Detail of the free cross section in first OmpF monomer for different radius of the used cylinder in the first monomer. We can not an increase of the cross section near to the constriction region for cylinder radius greater than 16 Å.

in the vacuum space available between every monomer (Figure 4): when the radius increase to values bigger than 16 Å, some of the free vacuum space between each monomer starts to be erroneously computed as free available space for the translocation. Taking into account these arguments, we choose as best representation of the cross section the one coming out from the cylinder with a radius of 16 Å.

Considering this radius we can see a similar trend for cross section in all monomers, with only little differences in monomer three (Figure 20).

With the same set of MD simulations described before in this section, we calculate also along the z axis the variance and the skewness of OmpF's cross section, that is useful to estimate the distribution of fluctuations in this area.

3.3.2 *Free–energy calculation from the model*

Given a particle that try to cross the channel, the calculation of the free–energy from the model can be performed in different ways depending on the hypothesis done.

The available cases that can be compared are three:

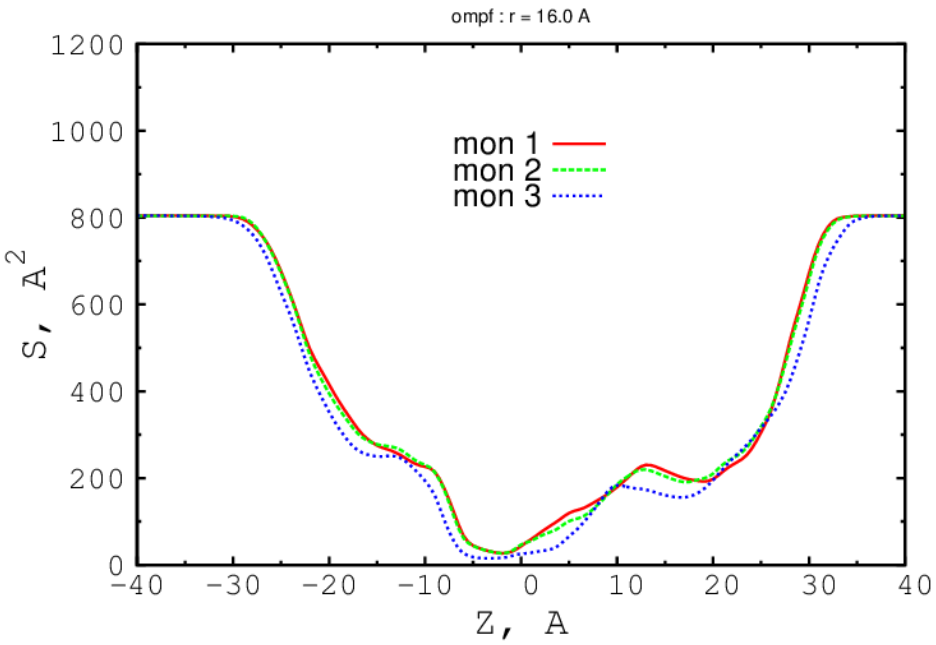

Figure 20: Comparison of cross section of each of the three monomers in OmpF for a cylinder radius of 16 Å.

- the simple geometrical case obtained from Equation 71 considering $A(x) = \langle S(x) \rangle - \langle S_m \rangle$, with $\langle S(x) \rangle$ and $\langle S_m \rangle$ representing the average value, obtained from MD simulation, of the cross section of the channel and the molecule respectively,
- the Gaussian case, where we use the calculated average and variance of both channel and molecule cross section and after we use them to obtain through Equation 74 the available area to be substituted in Equation 71 to get the new free–energy values from model,
- the case of a generalised asymmetric Gaussian, obtained using the asymmetric distribution of the channel cross section derived from Equation 102 to calculate through Equation 73 the available area to insert in Equation 71.

In Figure 21, 22 and 23, we show how these different ways to calculate free-energy perform changing the Van der Waals radius of the spherical atoms supposed to translocate through the channel.

We can see from this sequence of graphs that the way we calculate $A(x)$ in Equation 71 influence the free–energy calculated with our model. In Figure 21 the radius is still little (0.5 Å) compared to the approximated radius of the constriction region (more or less 3 Å). For this reason the geometric and Gaussian model perform bet-

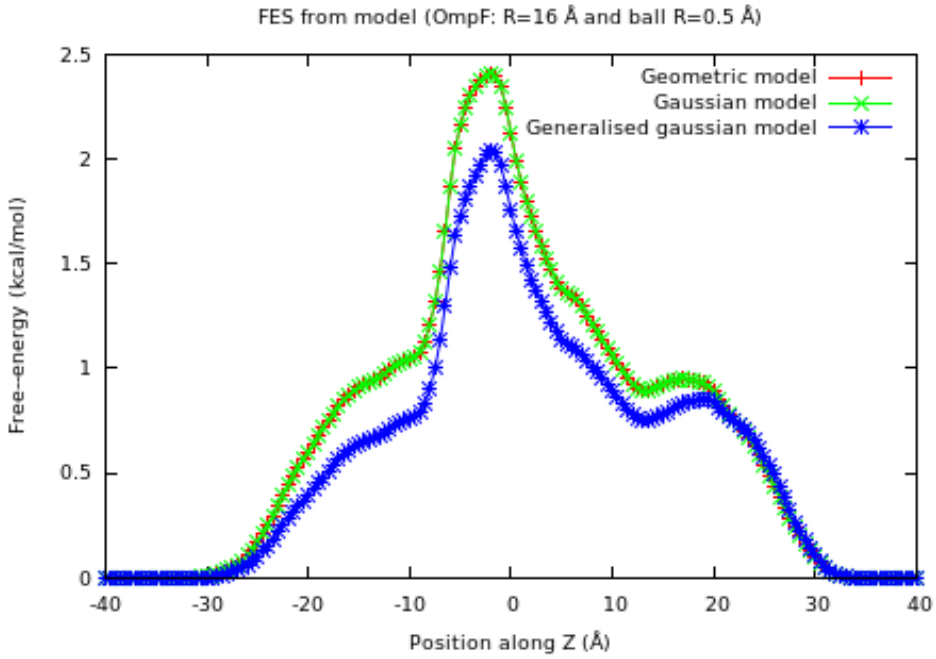

Figure 21: Comparison of the free energy values obtained using different ways (geometrical, Gaussian and Gaussian generalised) to calculate $A(x)$ in Equation 71 (the channel parameters are calculated as stated before using a cylinder radius of 16 Å). In this case a spherical atom having a 0.5 Å radius translocate through OmpF.

ter than the asymmetric Gaussian one, because the approximations used in the last model still do not match well with this case. When (Figure 22) the cross section of the atom that have to translocate (spherical atom with radius 3.5 Å) starts to be slightly bigger than the constriction region cross section, the geometrical model fails to calculate the free–energy in the smallest channel region and we can see a straight line in the Figure, due to not available (logarithm of negative quantities in Equation 71) data. The remaining two models show instead a good match in this smallest region and this means an almost purely Gaussian fluctuation of the channel for small fluctuations around the average cross section value. The last case (Figure 23) is obtained with a spherical atom (radius 5.5 Å) having a cross section much bigger than the one in the constriction region. In this case (apart from the always more evident failure of the geometric model) we see smaller values with the Gaussian generalised model than with the simple Gaussian model. We believe the values obtained with the Gaussian generalised model are more reliable because in our MD simulations used to study the fluctuations of the OmpF we observed a not negligible skewness and excess kurtosis, that can not be taken into account with the purely Gaussian model.

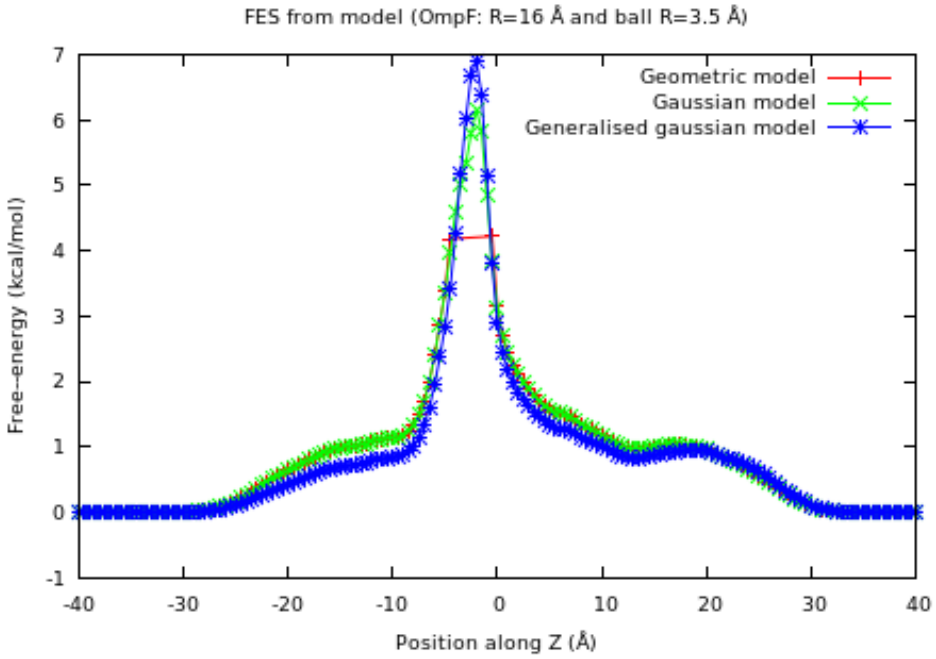

Figure 22: Comparison of the free energy values obtained using different ways (geometrical, Gaussian and Gaussian generalised) to calculate $A(x)$ in Equation 71 (the channel parameters are calculated as stated before using a cylinder radius of 16 Å). In this case a spherical atom having a 3.5 Å radius translocate through OmpF.

The translocation in this case occurs when we observe a large fluctuation of the constriction region size and this can be adequately generalised with an asymmetric Gaussian (like our) having parameters depending on skewness and excess kurtosis.

3.3.3 *Model vs. simulated free-energy*

The final goal of our model is the comparison with other, likely more trusted, values. To do that we used as reference values the ones simulated through WTMT simulations obtained using the classical system with explicit solvent of Figure 4, but having not a Meropenem as a molecule to transport but a simple spherical ball of various radius. The two CVs used are the distance along the Z and XY directions of the molecule with respect to the OmpF CM, with Z representing the direction of the OmpF's axis of symmetry. We used a CV temperature of 3000 K.

We simulated spherical atoms with different characteristics: neutral or charged with ± 1 e^-, with different Van der Waals radii. Fi-

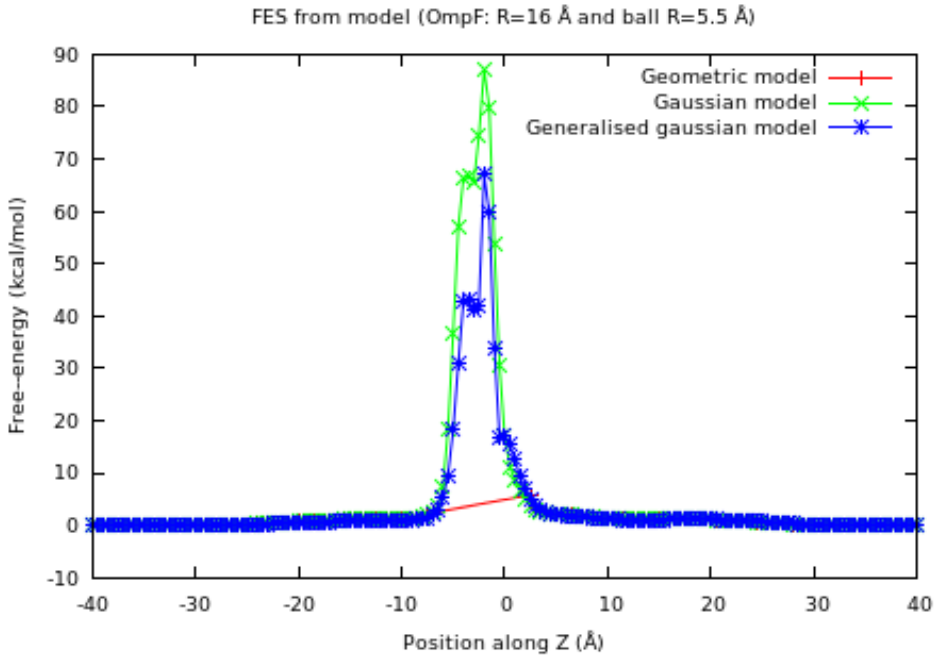

Figure 23: Comparison of the free energy values obtained using different ways (geometrical, Gaussian and Gaussian generalised) to calculate $A(x)$ in Equation 71 (the channel parameters are calculated as stated before using a cylinder radius of 16 Å). In this case a spherical atom having a 5.5 Å radius translocate through OmpF.

nally we performed this computation putting our atom in each of OmpF monomers.

When we start with the smallest spherical atoms, *e. g.* R=1.5 Å, we observe (Figure 24, 25 and 26) the absence of the steric barrier in the constriction region. This is due to the little size of the atom that allows it to be trapped in the many little cavities on OmpF channel, that are not taken into account in the model. We can see from Figure 24 the neutral atom case, that show two minima around 0 and 10 Å in correspondence of some of these cavities. The influence of the presence of charged residues in OmpF can be seen when we simulate an atom with same radius but charged. In these cases the two minima in 0 and 10 Å are alternatively (Figure 25 and 26) deepened (from around 1.5 kcal/mol to around 3.0 kcal/cal) or flattened depending on the presence near to these cavities of charges opposite or equal to the one of the spherical atom.

Increasing the spherical atom radius up to 2.5 Å, the presence of minima in free–energy due to the cavities in OmpF disappear, because the atom is too big to be trapped. In Figure 27 we can see a good agreement between the model and the WTMT simulations.

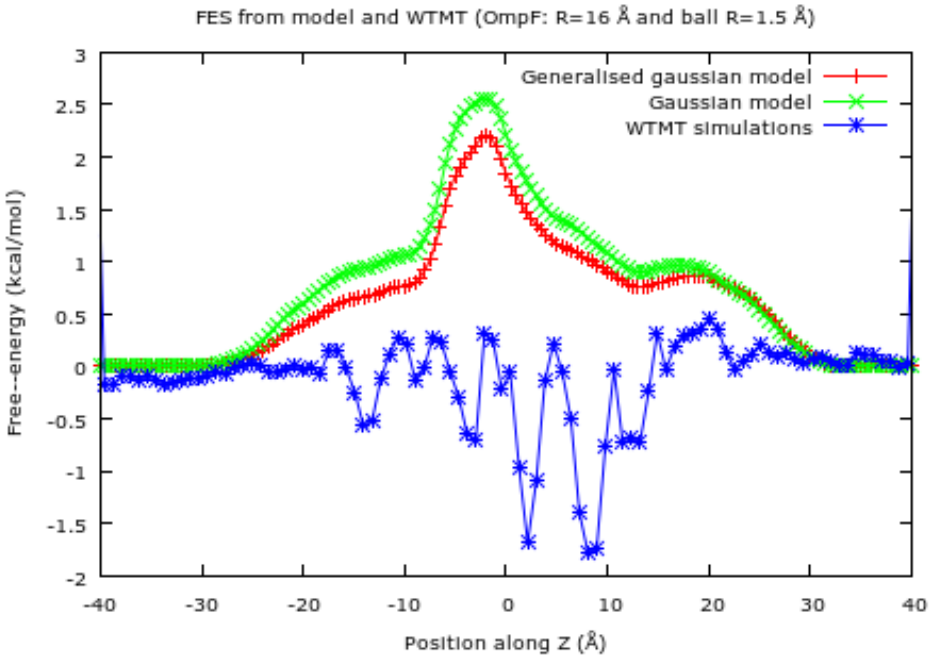

Figure 24: Comparison of the free energy values obtained using the Gaussian and Gaussian generalised (asymmetric) model, with the ones obtained from WTMT simulations. The channel parameters are calculated as stated before using a cylinder radius of 16 Å and a spherical neutral atom having a 1.5 Å radius translocate through OmpF.

In this last case the cross section of the atom is still littler than the one in constriction region, so we can not say anything about the hypothesis of channel vibration involved in translocation. To explore this guess, we increase the radius of the simulated atom until 4.0 Å (in this case the cross section of the atom is larger than the average one in constriction region), but even in this case we can see (Figure 28) a good agreement between model and WTMT simulations. In all the cases analysed (except the case of really small spherical atoms, like in Figure 24) we have observed smaller values of free–energy in the model, if compared with the WTMT simulations. This behaviour can be explained, because we performed the MD simulations that gave us the parameters (average channel cross section, variance, skewness and excess kurtosis) at equilibrium without any antibiotic in the channel. It is possible that in slightly different situations (equilibrium not reached and particle to be translocated in the channel) the porin reacts differently having less vibrations and giving a larger value for free–energy.

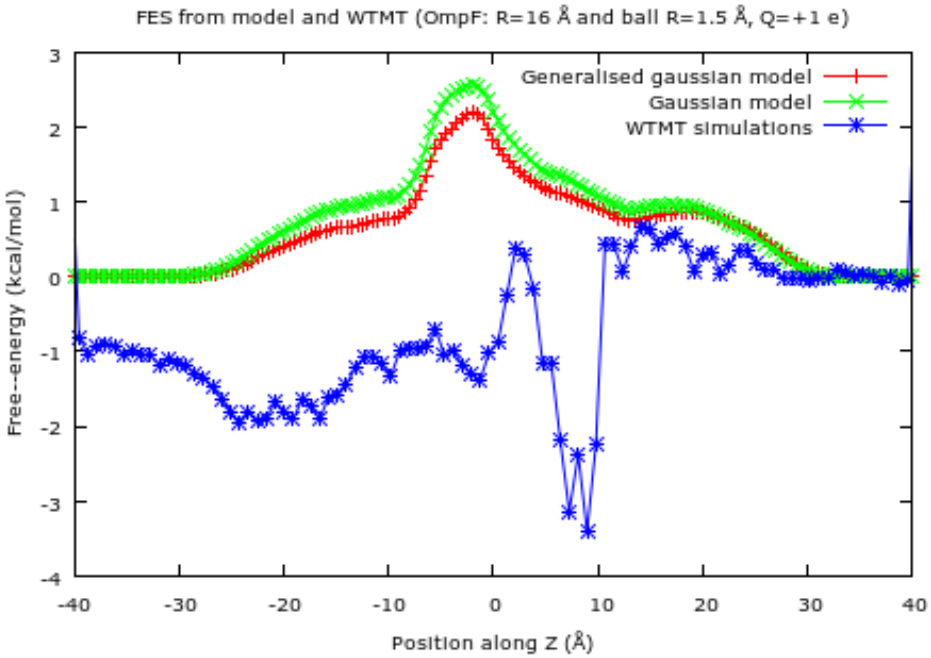

Figure 25: Comparison of the free energy values obtained using the Gaussian and Gaussian generalised (asymmetric) model, with the ones obtained from WTMT simulations. The channel parameters are calculated as stated before using a cylinder radius of 16 Å and a spherical atom having a 1.5 Å radius and a +1 e^- charge translocate through OmpF.

Finally we have not observed differences when we simulate the atom in monomer number one (all pictures here shown refer to this monomer) or in one of the other monomers.

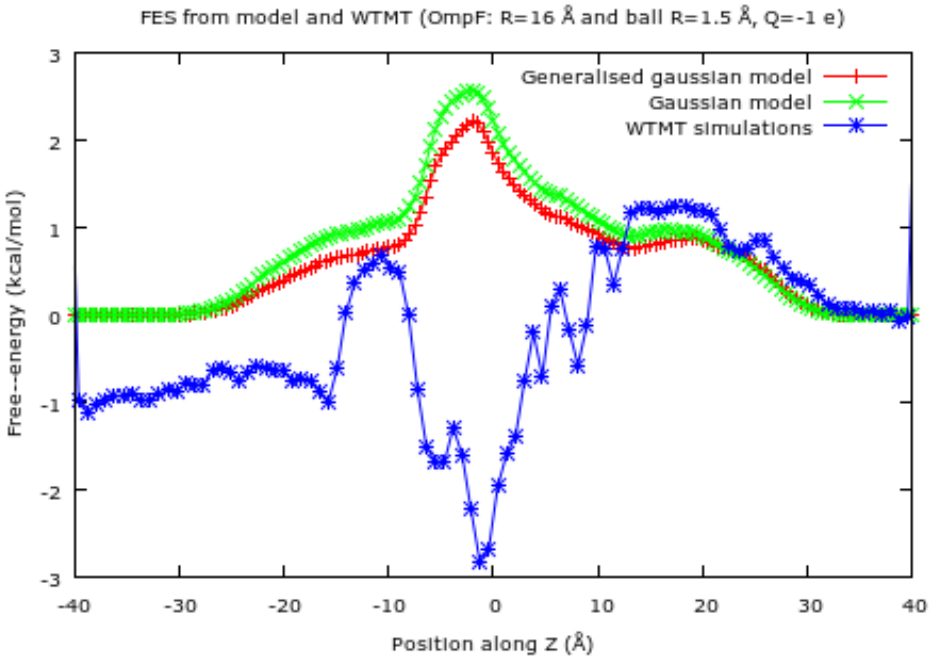

Figure 26: Comparison of the free energy values obtained using the Gaussian and Gaussian generalised (asymmetric) model, with the ones obtained from WTMT simulations. The channel parameters are calculated as stated before using a cylinder radius of 16 Å and a spherical atom having a 1.5 Å radius and a -1 e^- charge translocate through OmpF.

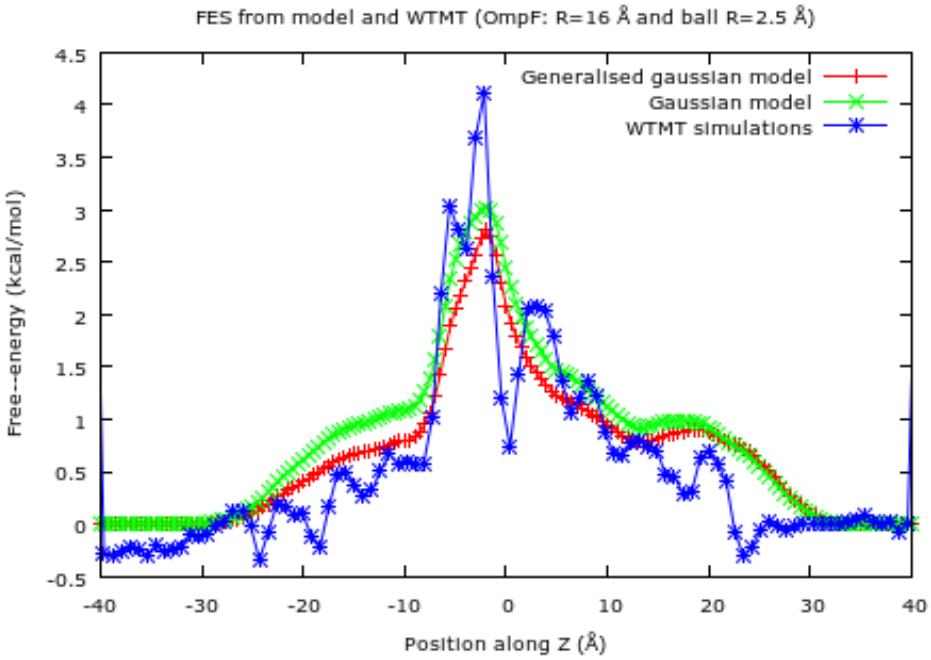

Figure 27: Comparison of the free energy values obtained using the Gaussian and Gaussian generalised (asymmetric) model, with the ones obtained from WTMT simulations. The channel parameters are calculated as stated before using a cylinder radius of 16 Å and a spherical neutral atom having a 2.5 Å radius translocate through OmpF.

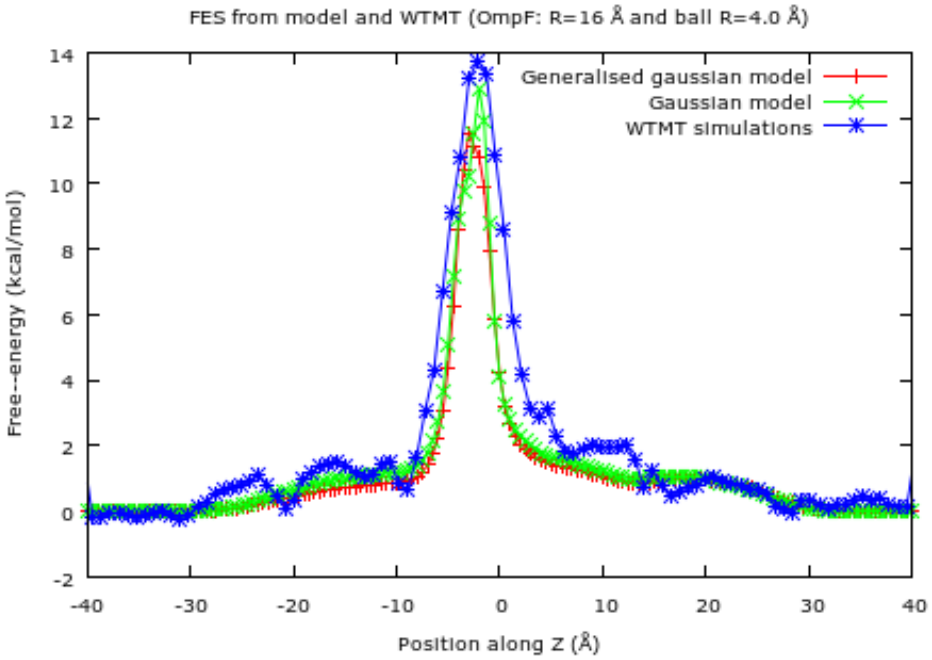

Figure 28: Comparison of the free energy values obtained using the Gaussian and Gaussian generalised (asymmetric) model, with the ones obtained from WTMT simulations. The channel parameters are calculated as stated before using a cylinder radius of 16 Å and a spherical neutral atom having a 4.0 Å radius translocate through OmpF.

4

CONCLUSIONS

Contents

In this thesis we afford the problem of the translocation of antibiotics through bacterial channels in a threefold way comparing results from different kinds of approaches: the experimental one through electrophysiology, the *in silico* approach through MD and WTMT simulations applying in particular the FMTD method and finally the more theoretical approach developing a model to study the properties of the translocation assuming basic hypothesis. This apparently confuse way of proceeding is actually almost mandatory due the complexity of the phenomena studied, that not allow us to obtain completely satisfactory results using, let say, a "pure" line.

4.1 MICROSECONDS AND SUB–MICROSECONDS CHANNEL GATING

We have first analysed the experimental data focusing on the widely used electrophysiology approach. The bigger limit in this method resides on the apparently (due to the typical 10 kHz low–pass filter) too big residence time detectable. At first glance it should seem that only $\tau_{res} \sim 10^{-4}$ s can be detectable, but we have described two methods that allow us to detect easily $\tau_{res} \sim 1\mu$s and give us a really good lower limit of $\tau_{res} \approx 0.5\mu$s still having a classic electrophysiology equipment. Moreover these methods, although quite sophisticated, have shown a good agreement with experimental results and for this reason can be safely used instead of the classical methods used to get kinetic parameters from electrophysiology.

The possibility to detect such a little time (microsecond or sub–microsecond time scale) is really important because give the opportunity to reduce considerably the distance between the time range of experimental and computational results.

4.2 TRANSITION TIMES FROM METADYNAMICS

Unfortunately, even with this still big improvement, it is not possible to cope with the time–range of systems to be simulated as big as our without super–computing facilities. This problem leads to necessarily coarse–grained computational methods to try to reach at least the μs time scale with almost standard computers. Our choice has been a recent method FMTD[7] that using WTMT simulations permit to obtain real transition times. We have opted for this method even because in our group there is a good level of competence in MT methods. The results obtained with the FMTD method permitted us to obtain good results[31] comparable with electrophysiology experiments.

Moreover in this thesis we have suggested an additional possible use of this method to study the whole translocation process in particular cases (*e. g.* "fast" events in Meropenenum in Section 3.2.2.1) or as a first screening of the effectiveness of an antibiotic as in Section 3.2.2.3. Although the FMTD method has not passed the statistical tests to assert the Poissonian nature of the event, from considerations about this method made in Section 3.2.2.3 and the results of Table 10 give a clear evidence of a faster translocation in Meropenem with respect to Meropenenum that can be likely the origin of the different effectiveness of the two antibiotics.

Analysing a quite simple case as the exit times of oxygen and water from the HOD catalytic box in Section 3.2.1, we have seen how, even with standard MD simulations, the transition times can or can not pass the statistical tests even with systems apparently really similar. Due the complexity of our system is not surprising to obtain poor results in statistical tests and, given these arguments, we can take in a bigger consideration even the results in Section 3.2.2.

4.3 TRANSLOCATION MODEL

Considering the poor understanding in general of the basic parameters influencing the permeability of antibiotics through bacterial channels, an approach based on a model can be useful. Starting from simple hypothesis we can make gradually more complex our model checking the effects of single physical variables on translocation and comparing the results with experimental or simulated properties.

In our case we have analysed in Section 3.3.3 the effects of dimensions of neutral spherical balls and we have seen that the model

fails for little ($\lessapprox$ 1.5 Å) radii (Figure 24). This is due to the intrinsic hypothesis of the model that consider average cross sections of the model without considering the many little cavities on OmpF channel, but fortunately we observed good agreement between model free–energy and WTMT simulated free–energy increasing the dimensions of the translocating atom to values near to the ones that characterise standard antibiotics. For spherical atoms having dimensions of 2.5 and 4.0 Å the results are quite good (Figure 27 and 28) and give us a good confidence on the basic assumptions of the model.

4.4 PERSPECTIVES AND GENERAL CONSIDERATIONS

The analysis of microseconds and sub–microseconds channel gating are in our opinion ready to be used in electrophysiology experiments and can help to boost them and integrate it more strongly with simulations giving a microscopical interpretation of blockages.

The FMTD method can be used with profit even to study translocations of antibiotics through bacterial channels, despite the complexity of the system to be simulated. The increased time–range available thanks to this method can easily allow us to compare experimental and simulated times, giving us at the same time a hint about the most relevant CVs (the ones that obtain better results in statistical tests and transition times closer to the experimental ones).

The translocation model, still being in an early stage, already shows good results and can be developed to take into account the effects, for example, of the shape of the molecule to be translocated or his dipole momentum or other physical parameters helping to focus on the most relevant parameters in drug design.

BIBLIOGRAPHY

[1] M. H. Kollef and V. J. Fraser, "Antibiotic resistance in the intensive care unit," *Annals of Internal Medicine*, vol. 134, no. 4, pp. 298–314, 2001. (Cited on page iii.)

[2] H. Nikaido, "Multidrug resistance in bacteria," *Annual Review of Biochemistry*, vol. 78, pp. 119–146, 2009. (Cited on pages iii and 3.)

[3] M. Winterhalter and M. Ceccarelli, "Physical methods to quantify small antibiotic molecules uptake into gram-negative bacteria," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 95, no. A, pp. 63–67, 2015. (Cited on pages iii, 8, and 11.)

[4] G. Schwarz, J.-M. Pagès, and M. Winterhalter, "On translocation through a membrane channel via an internal binding site: Kinetics and voltage dependence," *Biophysical Journal*, vol. 84, no. 5, pp. 2990–2998, 2003. (Cited on pages iii, 7, 11, and 17.)

[5] J.-M. Pagès, C. E. James, and M. Winterhalter, "The porin and the permeating antibiotic: a selective diffusion barrier in gram-negative bacteria," *Nature Reviews Microbiology*, vol. 6, no. 12, pp. 893–903, 2008. (Cited on pages iii, 6, 7, 10, and 11.)

[6] I. Bodrenko, H. Bajaj, P. Ruggerone, M. Winterhalter, and M. Ceccarelli, "Analysis of fast channel blockage: revealing substrate binding in the microsecond range," *Analyst*, vol. 140, no. 14, pp. 4820–4827, 2015. (Cited on pages iii, 8, and 10.)

[7] P. Tiwary and M. Parrinello, "From metadynamics to dynamics," *Physical Review Letters*, vol. 111, pp. 230602–230606, 2013. (Cited on pages iii, 12, 24, 46, and 68.)

[8] V. Oliveira, F. Rubio, M. Almeida, M. Noqueira, and A. Pignatari, "Trends of 9416 multidrug–resistant gram–negative bacteria.," *Revista da Associacão Médica Brasileira*, vol. 61, no. 3, pp. 244–249, 2016. (Cited on page 3.)

[9] W. Ricciardi, G. Giubbini, and P. Laurenti, "Surveillance and control of antibiotic resistance in the mediterranean region," *Mediterranean Journal of Hematology and Infectious Diseases*, vol. 8, no. 1, p. e2016036, 2016. (Cited on page 4.)

[10] "Food and drug administration." http://www.fda.gov, 2016. (Cited on page 5.)

[11] M. Gunn, M. Lim, D. Cross, and M. Goldman, "Benchmarking the scientific output of the innovative medicines initiative," *Nature Biotechnology*, vol. 33, pp. 811–812, 2015. (Cited on page 5.)

[12] H. Nikaido, "Molecular basis of bacterial outer membrane permeability revisited," *Microbiology and molecular biology reviews*, vol. 67, no. 4, pp. 593–656, 2003. (Cited on pages 6 and 7.)

[13] H. Nikaido, "Porins and specific channels of bacterial outer membranes," *Molecular Microbiology*, vol. 6, no. 4, pp. 435–442, 1992. (Cited on pages 6 and 7.)

[14] E. M. Nestorovich, C. Danelon, M. Winterhalter, and S. M. Bezrukov, "Designed to penetrate: time-resolved interaction of single antibiotic molecules with bacterial pores," *Proceedings of the National Academy of Sciences*, vol. 99, no. 15, pp. 9789–9794, 2002. (Cited on pages 7, 9, 11, and 17.)

[15] R. Benz, A. Schmid, and G. H. Vos-Scheperkeuter, "Mechanism of sugar transport through the sugar-specific lamb channel of escherichia coli outer membran," *The journal of membrane biology*, vol. 100, no. 1, pp. 21–29, 1987. (Cited on pages 7, 9, and 45.)

[16] A. H. Delcour, "Electrophysiology of bacteria," *Annual Review of Microbiology*, vol. 67, pp. 179–197, 2013. (Cited on pages 7 and 9.)

[17] J.-M. Pagès, S. Kascàkovà, L. Maigre, A. Allam, M. Alimi, J. Chevalier, E. Galardon, M. Réfrégiers, and I. Artaud, "New peptide-based antimicrobials for tackling drug resistance in bacteria: Single–cell fluorescence imaging," *ACS Medicinal Chemistry Letters*, vol. 4, no. 6, pp. 556–559, 2013. (Cited on page 8.)

[18] "The axon cns guide: a guide to electrophysiology and biophysics laboratory techniques." Chapter 12: noise in electrophysiological measurements. (Cited on pages 9 and 17.)

[19] E. Hajjar, K. R. Mahendran, A. Kumar, A. Bessonov, M. Petrescu, H. Weingart, P. Ruggerone, M. Winterhalter, and M. Ceccarelli, "Bridging timescales and length scales: From macroscopic flux to the molecular mechanism of antibiotic diffusion

through porins," *Biophysical Journal*, vol. 98, no. 4, pp. 569–575, 2010. (Cited on pages 10, 12, and 45.)

[20] K. R. Mahendran, E. Hajjar, T. Mach, M. Lovelle, A. Kumar, I. Sousa, E. Spiga, H. Weingart, P. Gameiro, M. Winterhalter, and M. Ceccarelli, "Molecular basis of enrofloxacin translocation through ompf, an outer membrane channel of escherichia coli - when binding does not imply translocation," *The Journal of Physical Chemistry B*, vol. 114, no. 15, pp. 5170–5179, 2010. (Cited on page 10.)

[21] D. W. Borhani and D. E. Shaw, "The future of molecular dynamics simulations in drug discovery," *Journal of Computer-Aided Molecular Design*, vol. 26, no. 1, pp. 15–26, 2012. (Cited on page 11.)

[22] J. A. Lemkul, J. Huang, B. Roux, and A. D. MacKerell Jr., "An empirical polarizable force field based on the classical drude oscillator model: Development history and recent applications," *Chemical Reviews*, vol. 116, no. 9, pp. 4983–5013, 2016. (Cited on page 11.)

[23] D. E. Shav, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, K. Salmon, John, C. Young, B. Batson, K. J. Bowers, J. C. Chao, M. P. Eastwood, J. Gagliardo, J. P. Grossman, C. R. Ho, D. J. Ierardi, I. Kolossvàry, J. L. Klepeis, T. Layman, C. McLeavey, M. A. Moraes, R. Mueller, E. C. Priest, Y. Shan, J. Spengler, M. Theobald, B. Towles, and S. C. Wang, "Anton, a special–purpose machine for molecular dynamics simulation," *Communications of the ACM*, vol. 51, no. 7, pp. 91–97, 2007. (Cited on page 11.)

[24] A. L. Beberg, D. L. Ensign, G. Jayachandran, S. Kaliq, and V. S. Pande, "Folding@home: Lessons from eight years of volunteer distributed computing," *IEEE International Symposium on Parallel and Distributed Processing*, pp. 1–8, 2009. (Cited on page 11.)

[25] A. Laio and M. Parrinello, "Escaping free–energy minima," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 20, pp. 12562–12562, 2002. (Cited on pages 12 and 24.)

[26] A. Barducci, G. Bussi, and M. Parrinello, "Well-tempered metadynamics: A smoothly converging and tunable free-energy method," *Physical Review Letters*, vol. 100, p. 020603, 2008. (Cited on pages 12 and 24.)

[27] K. A. Fichthorn and W. H. Weinberg, "Theoretical foundations of dynamical monte carlo simulations," *The Journal of Chemical Physics*, vol. 95, no. 2, pp. 1090–1096, 1991. (Cited on page 12.)

[28] M. Ceccarelli, A. V. Vargiu, and P. Ruggerone, "A kinetic monte carlo approach to investigate antibiotic translocation through bacterial porins," *Journal of Physics: Condensed Matter*, vol. 24, no. 10, p. 104012, 2012. (Cited on pages 12, 13, and 49.)

[29] A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, "Assessing the accuracy of metadynamics," *The Journal of Physical Chemistry B*, vol. 109, no. 14, pp. 6714–6721, 2005. (Cited on page 12.)

[30] M. Salvalaglio, P. Tiwary, and M. Parrinello, "Assessing the reliability of the dynamics reconstructed from metadynamics," *Journal of Chemical Theory and Computation*, vol. 10, no. 4, pp. 1420–145, 2014. (Cited on pages 12, 24, 26, 46, and 50.)

[31] T. D'Agostino, S. Salis, and M. Ceccarelli, "A kinetic model for molecular diffusion through pores," *Biochimica et Biophysica Acta*, vol. 1858, no. 7B, pp. 1772–1777, 2016. (Cited on pages 13, 55, and 68.)

[32] S. Nekolla, C. Andersen, and R. Benz, "Noise analysis of ion current through the open and the sugar-induced closed state of the lamb channel of escherichia coli outer membrane: evaluation of the sugar binding kinetics to the channel interior.," *Biophysical Journal*, vol. 66, no. 5, pp. 1388–1397, 1994. (Cited on page 17.)

[33] "Molecular dynamics simulation: Elementary methods." J. M. Haile, 1992. (Cited on page 24.)

[34] M. J. Harvey, G. Giupponi, and G. De Fabritiis, "Acemd: Accelerating biomolecular dynamics in the microsecond time scale," *Journal of Chemical Theory and Computation*, vol. 5, no. 6, pp. 1632–1639, 2009. (Cited on pages 24 and 48.)

[35] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, "Plumed: A portable plugin for free-energy calculations with molecular dynamics," *Computer Physics Communications*, vol. 180, no. 10, pp. 1961–1972, 2009. (Cited on pages 24 and 48.)

[36] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the amber ff99sb protein force field.," *Proteins: Structure, Function, and Bioinformatics*, vol. 78, no. 8, pp. 1950–1958, 2010. (Cited on pages 24 and 48.)

[37] R. A. Steiner, H. J. Janssen, P. Roversi, A. J. Oakley, and S. Fetzner, "Structural basis for cofactor–independent dioxygenation of n–heteroaromatic compounds at the α/β–hydrolase fold," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 2, pp. 657–662, 2005. (Cited on page 46.)

[38] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, pp. 926–935, 1983. (Cited on page 48.)

[39] C. J. Dickson, L. Rosso, R. M. Betz, R. C. Walker, and I. R. Gould, "Gafflipid: a general amber force field for the accurate molecular dynamics simulation of phospholipid," *Soft Matter*, vol. 8, no. 37, pp. 9617–9627, 2012. (Cited on page 48.)

[40] W. Humphrey, A. Dalke, and K. Schulte, "Vmd: Visual molecular dynamics," *Journal of Molecular Graphic*, vol. 14, no. 1, pp. 33–38, 1996. (Cited on page 48.)

[41] V. S. Pande, K. Beauchamp, and G. R. Bowman, "Everything you wanted to know about markov state models but were afraid to ask," *Methods*, vol. 52, no. 1, pp. 99–105, 2010. (Cited on page 55.)

[42] S. Acosta Gutiérrez, I. Bodrenko, M. A. Scorciapino, and M. Ceccarelli, "Macroscopic electric field inside water–filled biological nanopores," *Physical Chemistry Chemical Physics*, vol. 18, pp. 8855–8864, 2016. (Cited on page 56.)