Università degli Studi di Cagliari

DOTTORATO DI RICERCA

FISICA

Ciclo XXVIII

TITOLO TESI

Enhanced sampling methods and their application in the study of molecular permeation in gram-negative bacteria

Settore/i scientifico disciplinari di afferenza

FIS 07 / Fisica Applicata

Presentata da: *Tommaso D'Agostino*

Coordinatore Dottorato *Prof. Alessandro De Falco*
Prof. Paolo Ruggerone

Tutor *Prof. Matteo Ceccarelli*

Esame finale anno accademico 2014 – 2015

UNIONE EUROPEA
Fondo sociale europeo

REGIONE AUTONOMA DELLA SARDEGNA

innovative medicines initiative

ENHANCED SAMPLING METHODS AND THEIR APPLICATION IN THE STUDY OF MOLECULAR PERMEATION IN GRAM-NEGATIVE BACTERIA

Tommaso D'Agostino

Università degli Studi di Cagliari

Facoltà di Scienze

Dipartimento di Fisica

Relatore: *Prof. Matteo Ceccarelli*

Coordinatori: *Prof. Alessandro De Falco*
Prof. Paolo Ruggerone

ABSTRACT

Antimicrobial resistance is inhibiting our ability to fight against pathogens. By selectively changing the composition and expression of influx water-filled proteins filling their outer membrane, gram-negative bacteria are able to reduce the rates at which specific polar compounds are able to permeate. A clear comprehension of the mechanism determining substrates diffusion through these pores is still missing. In this thesis, we show how biased computer simulations may offer a unique perspective in the study of molecular permeation through porins, overcoming the intrinsic limitations of both experimental techniques and standard molecular dynamics.

The first test-case is Acinetobacter baumannii's CarO. The use of substrates with varying charge and molecular weight, as well as the creation of a loop-less mutant missing the extracellular domain of the protein, allowed to determine the charge selectivity and the transition rates of polar molecules. We obtained good agreement with the results of liposome swelling assays experiments. Further, we compared the passage of two carbapenem antibiotics in a series of mutated proteins extracted from a patient undergoing long term hospital infection. We connected the mutation of few key residues to a drastic change in the internal electric field of the proteins, showing that the antibiotics follow the choreography of water molecules inside the channels. In the last section, we present a kinetic model that allows to determine for a molecule the relative probability of different conformations and the time required for the translocation through a pore. This approach allowed to connect the results of enhanced sampling MD methods with current blockages in single-channel experiments.

All these results together show that multiscale MD techniques can offer an exhaustive view on the mechanism of molecular diffusion through pores, helping to understand the most important characteristics that determine the rates of translocation of different compounds in gram-negative bacteria. We can use these data to complement experimental results and to design the next generation of antibiotics.

PUBLICATIONS

The results of this thesis have been published as follows:

[1] M. Zahn, **T. D'Agostino**, E. Eren, A. Baslé, M. Ceccarelli, B. van den Berg, Small-Molecule Transport by CarO, an Abudant Eight-Stranded β-Barrel Outer Membrane Protein from *Acinetobacter baumannii*, *J. Mol. Bio.* **427(14)**, 2329-2339 (2015)

[2] **T. D'Agostino**, S. Salis, M. Ceccarelli, A Kinetic Model for Molecular Diffusion through Pores, *BBA Biomemb.*, **DOI: 10.1016/j.bbamem.2016.01.004** (2016)

[3] M.A. Scorciapino, **T. D'Agostino**, S. Acosta-Gutierrez, I. Bodrenko, M. Ceccarelli, Internal Electric Field of GRAM- Unspecific Porins Directs the Choreography of Antibiotic Translocation (Abstract), *Biophys. Jou.*, **DOI: 10.1016/j.bpj.2015.11.674** (2016)

[4] M.A. Scorciapino*, **T. D'Agostino***, S. Acosta-Gutierrez, G. Malloci, I. Bodrenko, M. Ceccarelli, Exploiting the porin pathway for polar compounds delivery into Gram-negative bacteria, *SUBMITTED* (2016)

*) authors contributed equally to this work

ACKNOWLEDGMENTS

First of all, to my parents, R. D'Agostino and C. Greco, for their continous support during the years.

To the biophysics group here at the University of Cagliari and in particular to my supervisor, M. Ceccarelli. His guidance and comprehension allowed me to be here. Fruitful conversations with S. Acosta-Gutierrez helped in elucidating doubts during the last two years. I would also like to thank A. Emanuele, my supervisor during master's degree, for moral and scientific support.

To the project I've been part of, Innovative Medicine Initiative. This collaboration granted me the opportunity of working with an international team of experts that offered a multidisciplinary approach to the topic of antimicrobial resistance. A special mention has to be done towards the group of B. van den Berg at Newcastle University, that produced the experimental results shown in section 3.1. I would specially thank the coordinator of the translocation consortium M. Winterhalter, who has kindly referred this thesis.

A special acknowledgment has to be done to F. Reina, E. Gioè and L. Brucculeri, whose friendship wasn't affected by the distance that separated us.

Tommaso D'Agostino

CONTENTS

LIST OF ACRONYMS

MDR	Multiple Drug Resistant
IMI	Innovative Medicines Initiative
ND4BB	New Drugs for Bad Bugs
GN	Gram-negative
GP	Gram-positive
OM	outer membrane
OMP	outer membrane protein
MD	Molecular Dynamics
AMBER	Assisted Model Building with Energy Refinement
PBC	periodic boundary conditions
CV	collective variable
CM	centre of mass
FES	free energy surface
MWMeta	Multiple Walkers Metadynamics
WTMeta	Well Tempered Metadynamics
BEMeta	Bias Exchange Metadynamics
BEMeta	Minimum Mode Metadynamics
PTMeta	Parallel Tempering Metadynamics
WTEMeta	Well Tempered Ensemble Metadynamics
FG	Fine Grained
CG	Coarse Grained
KS	Kolmogorov-Smirnov
CarO	Carbapenem-resistance associated Outer membrane protein

Part I

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

1 INTRODUCTION

Contents

1.1 ANTIMICROBIAL RESISTANCE

One general law, leading to the advancement of all organic beings, namely, multiply, vary, let the strongest live and the weakest die

— Charles Darwin, The Origin of Species

1.1.1 *Advent of Resistance*

Antimicrobial resistance refers to the ability of bacteria to develop resistance to antibiotics once effective against them. It can be regarded as a direct effect of the survival of the fittest theory in the case of the bacteria. When a compound is used to fight pathogens, the elements that show a resistance against it will have an improved rate of survival with respect to the unresistant ones and hence the next generation of pathogens will have more resistant elements than the previous one (Figure 1).

This characteristic is now prevalent in a number of key gram-negative bacterial pathogens, including *Escherichia coli*, *Salmonella*, *Klebsiella*, *Enterobacter*, and *Pseudomonas*[2]. The use of antimicrobials has led, during the last decades, to the appearance of the so-called *superbugs*, that are multidrug-resistant (MDR) forms of the bacteria subsequent to prolonged antibiotic use[3].

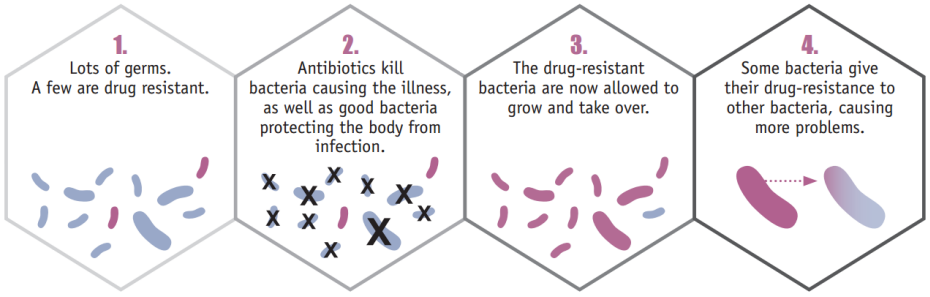

Figure 1: Schematic scheme of antimicrobial resistance. Reprinted from[1]

Four major mechanisms account for the evolution of bacterial resistance to antibiotics that correlates with the use of the drugs[4]:

1. emergence of "new" opportunistic pathogenic soil microorganisms that are often multiresistant to antibiotics, like *Acinetobacter baumannii*

2. emergence of "new" acquired resistance mechanisms. An example is given by *Enterococci* species, that has recently become resistant against glycopeptides

3. occurrence of mutations in genes located in the host chromosome or plasmid borne. An example is given by the appereance of β-lactamases

4. spread of "old" (i.e., already known) resistance genes into species that were previously uniformly susceptible

All these processes contribute to the natural selection of the bacterial strains that are more effective against drugs. On a local scale, the arouse of bacterial resistance can be very rapid (resistant isolates of *Enterobacter aerogenes* where isolated from patients undergoing an antibiotic therapy for less than 5 days[5]); globally, in the last decades this has led to a selection of resistant strains that are virtually unaffected by antimicrobial agents (Figure 2).

Together with the tightening of new drug approval regulations, antimicrobial resistance has led to a huge increase in the cost necessary for discovering new drugs (Figure 3): as a consequence, there has been a *discovery void* in the creation of new classes of antibiotics after 1980s[8].

The creation of new antibiotics effective against MDR pathogens is either done by slightly changing the characteristics of existing molecules or by discovery of completely new scaffolds of antimicrobials[10]. The latter method, albeit more sustanaible in the long

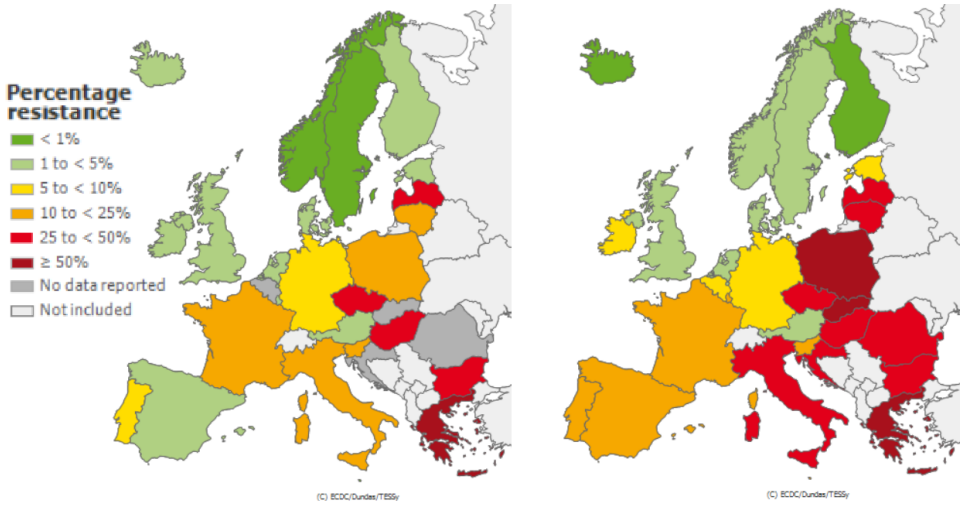

Figure 2: Percentage of *Klebsiella pneumoniae* showing multidrug-resistance towards third-generation Cephalosporins, Fluoroquinolones and Aminoglycosides in 2008 (left) and 2013 (right). Redrawn from [6,7]

term, is currently limited by a trial-error approach at the early stage (target to hit, hit to lead), followed by detailed analysis on the best performing molecules (leads), that are successively optimized and tested in preclinical and clinical trials[9]. During the last years, *Innovative Medicines Initiative* (IMI) has led the development of new medicines in Europe by supporting collaborative research projects between public and private companies[11]. New Drugs for Bad Bugs (ND4BB), the project this thesys is supported by, aims to find an new, rational way to project the next generation of antibiotics, one that necessarly has to start from a better comprehension of the mecha-

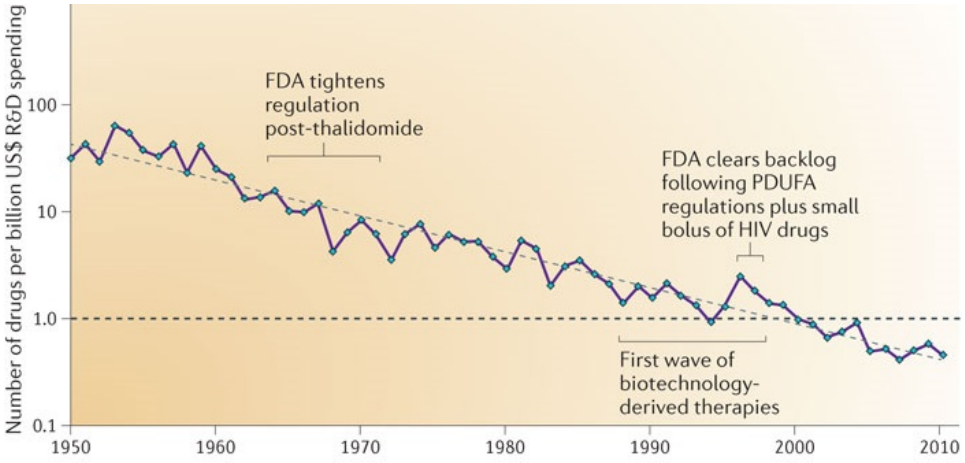

Figure 3: Number of new drugs discovered per billion dollar (inflaction-adjusted) as a function of time. Redrawn from [9]

nisms that determine the resistance that most pathogens are showing.

1.1.2 *Gram-negative Bacteria*

Superbugs are not the only microbial threats, but they are recognized as the most menacing with respect to morbidity and mortality worldwide[3]. *Gram-negative* (GN) bacteria are particularly concerning, as in the last decades an increasing percentage of infections as been related to them: indeed, while the proportion between infections related to *Gram-positive* (GP) and GN bacteria was 3:1 in the period 1989-1993, after 10 years GN were causing 50% of hospital acquired infections[12].

Recent studies suggest that possible resistance genes from GP microorganisms are transferred and readily expressed in GN bacteria[4], whereas the reverse is generally not true[13]. The search for the next generation of antibiotics is therefore strongly related to the comprehension of the mechanisms of resistance that GN bacteria possess.

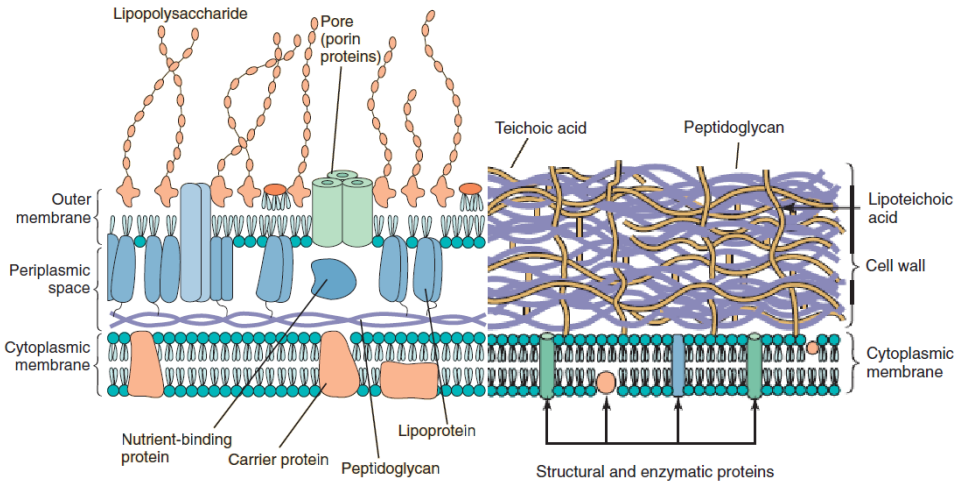

Figure 4: Cell wall of GN (left) and GP (right) bacteria. Redrawn from[14]

Both GP and GN bacteria share a cytoplasmic membrane composed of a symmetric lipid bilayer and filled with structural proteins, whose task range from electron transport to uptake of metabolites and enzymes. This *inner membrane* is surrounded by a multilayered peptidoglycanic cell wall that provides structural stability to the bacteria while being porous enough for allowing passage of nutrients[14].

With respect to their GP counterparts, GN bacteria possess a more complex cell wall (Figure 4). The presence of an *outer mem-*

brane (OM), formed by an additional asymmetrical lipid bilayer composed of lipopolysacharides and phospholipids, constitutes the first permeability barrier that protects the cells against environmental stresses including chemical, biophysical and biological attacks[15]. The OM is filled with large quantities (>100 copies per cell[16]) of influx proteins (*porins*) and efflux pumps (*transporters*). Porins mediate the transmembrane passage of ions, small molecules and peptides required for cell growth and function[17]: they are the major determinants for the permeability of the outer membrane for hydrophilic solutes. Nonspecific or general porins allow the diffusion of a broad range of charged and zwitterionic substances, ranging from nutrients[16,18] to antibiotics[19].

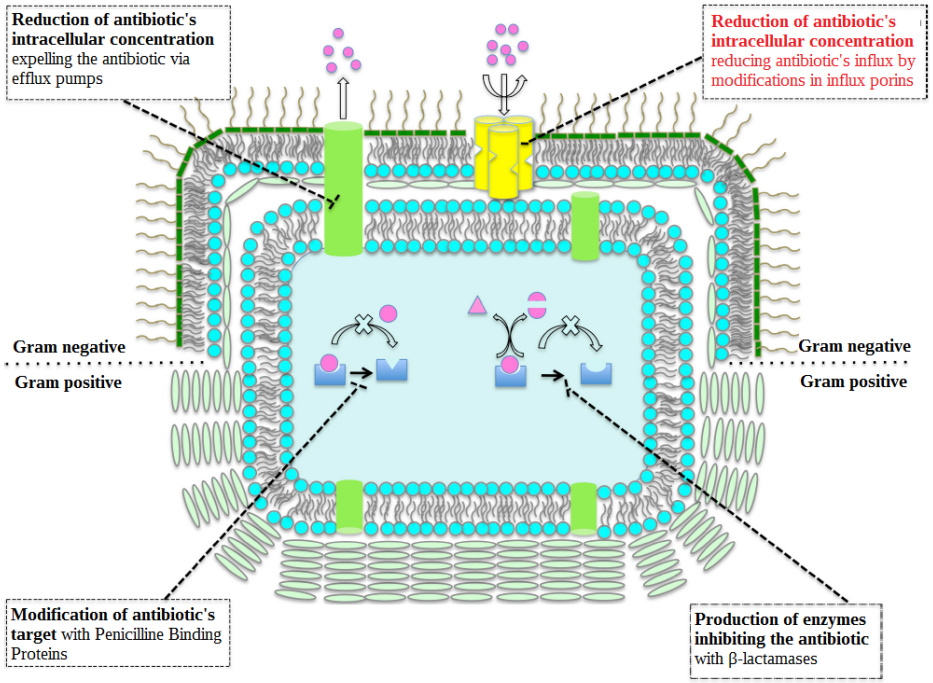

Figure 5: Resistance Mechanisms in Bacteria. Redrawn from[20]

Antibiotic translocation across membranes of Gram-negative bacteria is a key step for the activity on their specific intracellular targets. We can distinguish 4 types of antimicrobial based on different modes of action:

- Inhibition of OM syntesys[21] (cephalosporins, β-lactams)
- Inhibition of inner membrane syntesys (polypeptides)
- Inhibition of protein syntesis (aminoglycosides, tetracyclines)
- Inhibition of DNA syntesis (sulfamides, quinolones)

In this work, we have focused on the study of the influx of antibiotics of β-lactam family in GN Bacteria. GN and GP bacteria are able to combat the action of β-lactams by either overproducing the proteins that are targeted by the antibiotic (*penicilline binding proteins*) or by destroying the active group of the antibiotic itself (*β-lactamases*). In addition, GN bacteria control their membrane permeability as a first line of defense to protect themselves against external toxic compounds such as antibiotics and biocides[18] (Figure 5).

The concentration of harmful substances inside GN bacteria can be reduced by lowering the expression of influx proteins that mediate the transport of these compounds to the periplasmic space; conversely, an increased expression of efflux pumps allows to actively expel molecules that have already penetrated into the inner layers of the pathogen. Exchange in the type of influx proteins expressed, as well as mutations or modifications that impair the functional properties of the channels, act as additional defense mechanisms against harmful compounds[2,15,18].

1.1.3 *Outer Membrane Proteins*

Influx proteins may be divided in two classes, depending on their specificity (Figure 6):

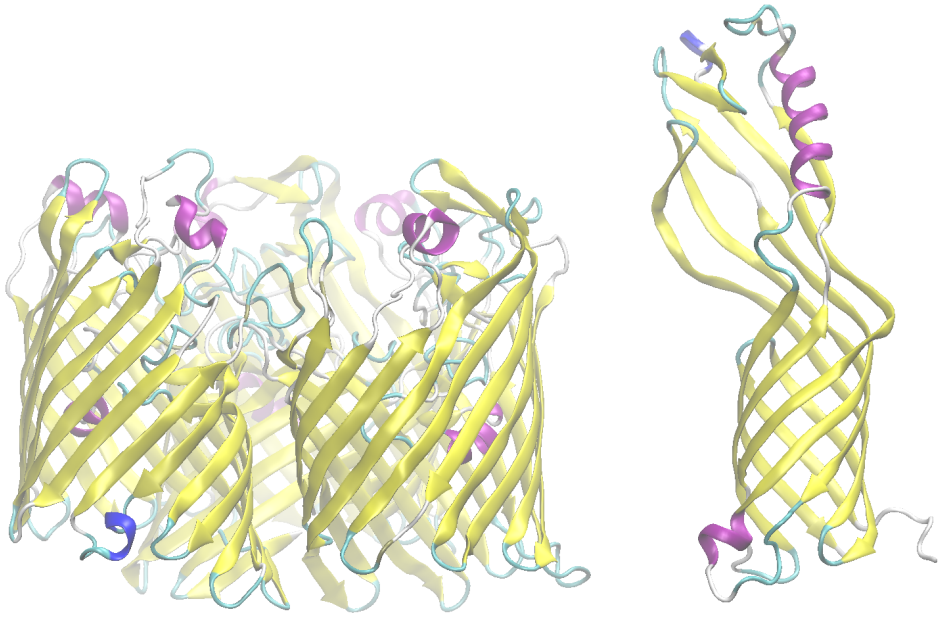

Figure 6: Example of Outer Membrane Proteins: A generic protein, OmpPst1 from *Providencia stuartii* (left) and a specific porin, CarO from *Acinetobacter baumannii* (right)

1. **General porins**: these proteins form water-filled channels through which hydrophilic solutes gain access into the bacterial cell[2,15]. They have usually a low degree of selectivity, mainly based on substrate's net charge rather than on particular chemical characteristics. The net flux of compounds through general porins is expected to be proportional to the concentration of the compounds themselves.

2. **Specific channels**: these proteins are characterized by a lower accessible area with respect to their general counterparts, and by the presence of strongly selective, high-affinity binding sites. The presence of a stable binding site reduces the rate at which compounds can penetrate through the channel, producing saturation-like kinetics.

While the tertiary structure of specific channels can vary greatly, we can make some general considerations regarding generic proteins. OMPs' structure and assembly are conserved throughout all species of Proteobacteria, while the corresponding genomes have been predicted to encode a variable number of OMPs[22]. This may not only reflect a common evolution but also very well-adapted pathways to rule the influx of nutrients and the efflux of intracellular noxious compounds[18]. Most of OMPs are formed by large homotrimers of β-barrels consisting of a conserved antiparallel 16-strand structure[21]. The fold of the monomer is very simple: the strands of the hollow β-barrel show short β-turn connections at the periplasmic rim (T1–T8) and long, irregular loops at the extracellular end (L1, L2, L4–L8)[19,21] (Figures 7A, 7B).

The large number and configuration of β-strands allows each monomer to form a separate central pore, which functions independently of the others in terms of solute permeation. Extracellular loop L2 has the additional function of contacting the adjacent porin monomer to stabilize the trimer structure[21]. Loop L3 departs from this simple scheme in that it extends along the inner side of the barrel wall and constricts the pore halfway through[19] which defines the size exclusion limit for permeation[21].

In *Escherichia coli*'s OmpF, one of the most expressed general porins of the pathogen, the constriction zone is marked by a cluster of acidic residues found on L3, which faces a cluster of basic residues on the adjacent β-barrel wall (Figure 7C). This gives rise to two effects: a reduction of the accessible area, that in the case of OmpF shrinks to approximately $7x10$Å[18], and the appearance of a strong electric field that is perpendicular with respect to the

axis of diffusion. The simultaneous presence of a ladder of positively charged residues (K16, R42, R82, R132) and of of negatively charged residues located in internal loop L3 (D113, E117) is thought to affect ion permeation and solute transfer[18,23–25].

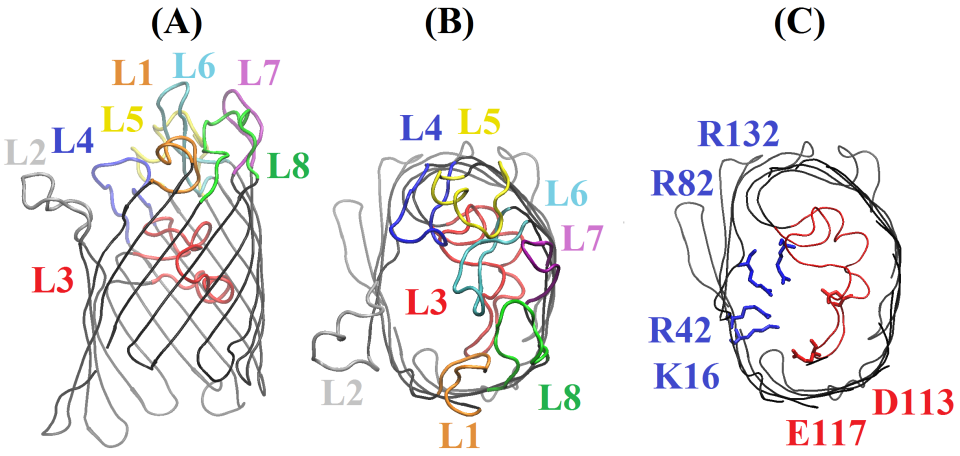

Figure 7: (A) Front view and (B) top view of OmpF monomer from *Escherichia coli*. Loops L1-L8 are highlighted. (C) The constriction zone of OmpF, with emphasis on the charged residues

Diffusion efficacy depends on multiple electrostatic interactions and hydrogen bonds that engage chemical groups on the substrate and amino acids of the porin[18]: a subtle balance of interactions, localized mainly at the channel constriction region[26], is thought of governing the diffusion of molecules through protein channels. This assumption is supported by the fact that mutations in porins of resistant clinical isolates exclusively and constantly target amino acids of the loop L3[2]. It remains unclear whether mutations alter the permeability of the OM because they change the electrostatic properties of the pore or because they alter the size of the channel: indeed, some clinical isolates often express porin with small channel size to decrease their antibiotic susceptibility[27].

The sequence of OmpC, the third major porin in *Escherichia coli*, is 60% identical to OmpF. The geometry of the two channels is virtually identical, so that replacement of an OmpC monomer with OmpF within a trimer does not produce any severe clashes[19]. Indeed, the formation of OmpF/OmpC heterotrimers in nature has been already reported[28]. OmpC is slightly more cation selective than OmpF and its pore is slightly smaller[19]; however, insertion and deletion in the external loops correspond to different fold of the protuberance at the entrance of the two channels. It has been argued whether the expression of OmpC in the place of OmpF is dependant on enviromental conditions[29,30], such as pH, tempera-

ture[31,32] and nutrient availability[18]. Most studies have implicated the osmolarity of the medium as the stimulus for the differential regulation: the presence of OmpF confers an advantage for the uptake of glucose, whereas OmpC is beneficial for the utilization of histidine as nitrogen source[19]. Long-term antibiotic treatment has been shown to decrease OmpC expression in the clinics[33] or to lead to the accumulation of amino acid substitutions that result in a smaller pore. Conversely, the presence of OmpF[34] as well as OmpC mutants selected for larger pore size exhibit increased antibiotic sensitivity[35].

These comparisons only allow to make partial hypotesis on the role and importance of single channels on the uptake of antibiotics and substrates in general by bacteria. A clear analysis of the role of the channels, based on a knowledge of the dynamics that molecules undergo during translocation, is currently missing. It seems that electrostatic pore potential and particular atomic details of the pore linings are the critical parameters that physiologically distinguish different OMPs[19].

1.2 LIMITATIONS OF EXPERIMENTS AND SIMULATIONS

> *If your experiment needs statistics, you ought to have done a better experiment*
>
> — Ernest Rutherford

Our understanding of antibiotic transport inside bacterial porins has been greatly enhanced by molecular-based multidisciplinary approaches, that combine *in vitro* and *in vivo* experiments and simulations[18]. Indeed, both experimental and simulation methods show their current disadvantages and prevent the accurate description of the complex process of antibiotic translocation[26]. Because of these limitations, an emerging strategy in the fight against GN bacteria is to pursue a multidisciplinary, rational and microscopically founded drug design that starts from the molecular knowledge of resistant mechanisms[26].

1.2.1 *Limitations in Currently Available Experimental Methods*

At present day, a direct and robust method for assessing the permeability of proteins towards different substrates by quantifying the flux is absent[36]. Current methods for determining the permeability of porins to different compounds, such as liposome swelling assays[17,37] and electrophysiology[38], rely on indirect measurements

to deduce the values of the binding constants for the system under study.

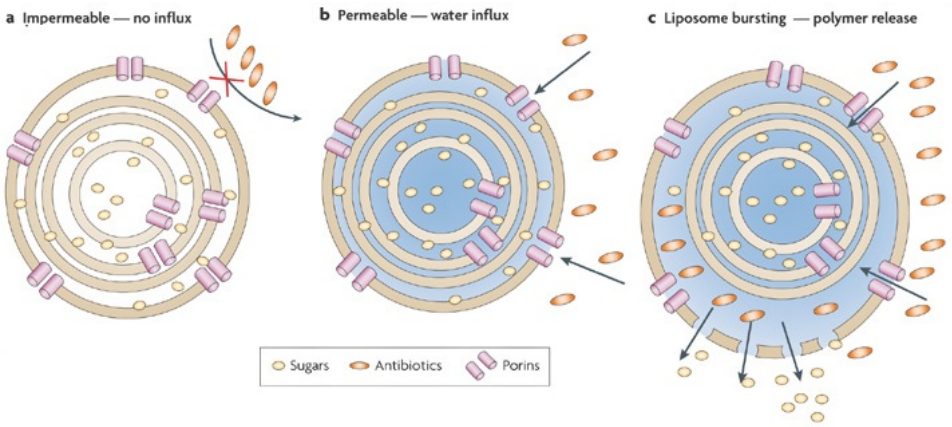

Figure 8: Scheme of a liposome swelling assays experiment. Proteins are added to a solution of lipids and sugars; the lipids will reconstitute into liposomes. the liposomes are then mixed in a solution with different optical density and with the same osmotic pressure. (a) If the proteins are impermeable to the antibiotic, no change in the turbidity of the solution is observed. If the antibiotics permeate in the liposomes (b), they will enlarge and eventually release the sugars, thus reducing the optical density of the medium. Reprinted from [2]

In liposome swelling assay experiments (Figure 8), proteins are reconstituted into liposomes and the permeability of the channels is assumed to be proportional to the variation of the optical density of these liposomes over time, in presence of a suspension of the substrate to study [2,37]. the method is limited to a comparison of different compounds because is strongly biased by the preparation of the suspension in which the liposomes are reconstituted and by the liposomes used, that are in general different from the constituent of the heterogeneous bilayer of the OM; moreover, it doesn't allow to assess the time required for the translocation of the substrates themselves.

In single-channel ion current experiments, permeation of a compound is deduced from the reduction of ionic current caused by blockage of the channel [16,38,39]. this electrophysiology method consists in the reconstitution of purified proteins into planar lipid bilayers; ion-current measurements provide structural and functional characteristics such as conductance (e.g. pore size), ion selectivity and voltage-dependent closure [18]. This method is affected by two

limitations, one intrinsic and one dependant on the experimental setup currently available.

The intrinsic problem is the inability in distinguishing between blockages of the channel caused by different effects: the method is unable to discriminate between reductions of the current caused by actual translocation of the substrates and the ones caused by a binding of the substrates to the channel that correspond to a successive rejection of the molecule[39,40].

The main limitation in this kind of experiment is the presently available time resolution in measurements of current interruption: indeed, the resolution limit of the technique restricts the detection of events occurring at timescales lower than $100\mu s$[41]. In the case of *Providencia stuartii*'s OmpPst1, an absence of current blockages has been observed for the case of imipenem (Figure 9); however, the passage of β-lactam antibiotics through the channel is supported by the fact that resistant derivatives of *P. stuartii* grown in the presence of β-lactam antibiotics showed a deficiency in the expression of OmpPst1 channel[27]. It is unclear whether this absence of blockages is due to lack of translocations or to translocations that are too fast to be seen by the time resolution of the experiment. Very recently, a method to address this problem and to enhance the time resolution of the method by doing a statistical analysis of the current fluctuations during the experiment has been suggested[42].

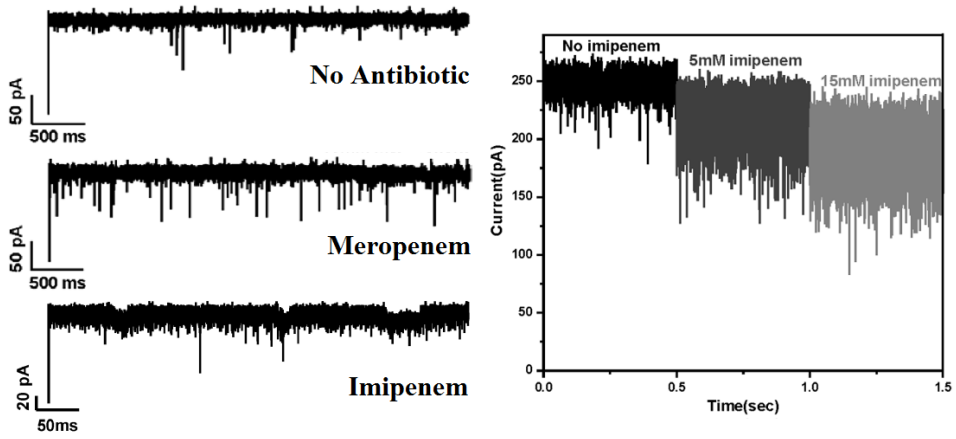

Figure 9: Single-channel ion current trace in *Providencia stuartii's* OmpPst1. Left: corrent interruption without antibiotics (top), in presence of Meropenem (center), in presence of Imipenem (bottom). Right: varying concentrations of imipenem result in change in overall conductance of the channel and not in single blockages. Redrawn from[41]

1.2.2 *Limits of Molecular Dynamics Simulations*

In the previous section, we have seen how presently available experimental methods either rely on indirect measures to deduce the kinetics that substrates undergo during translocation, or are affected by limitations that prevent to obtain a clear understanding of the process. Molecular Dynamics (MD) simulations, in contrast, would have the required microscopic accuracy to link structure and dynamics (of the drug and porin) to the rate of permeation[26]. Moreover, they can give an accurate description of the changes induced by mutations, either on the substrates (see Section 3.1) or on the pores (Section 3.2) on the transport properties. Such structural/dynamical information should be taken into account to complement and, especially, to rationalize experimental findings[43,44]. The main limitation of MD simulations is the computational effort required to solve the equations of motion (see Section 2.1) of the systems under study. This limits both the number of atoms that can constitute the unitary cell of the system and the time that can be simulated. In the last years, thanks to the advent of low-cost high-performance GPU computing and to the improvement of high-throughtput clusters like Anton[45], BSC[46] and Curie[47], the size of the systems that is possible to simulate is constantly increasing. At present time, we are able to run MD simulations of whole HIV capsids (Figure 10) for few nanoseconds or, conversely, to simulate systems of hundreds of thousands atoms for timescales of the order of tents of microseconds[17].

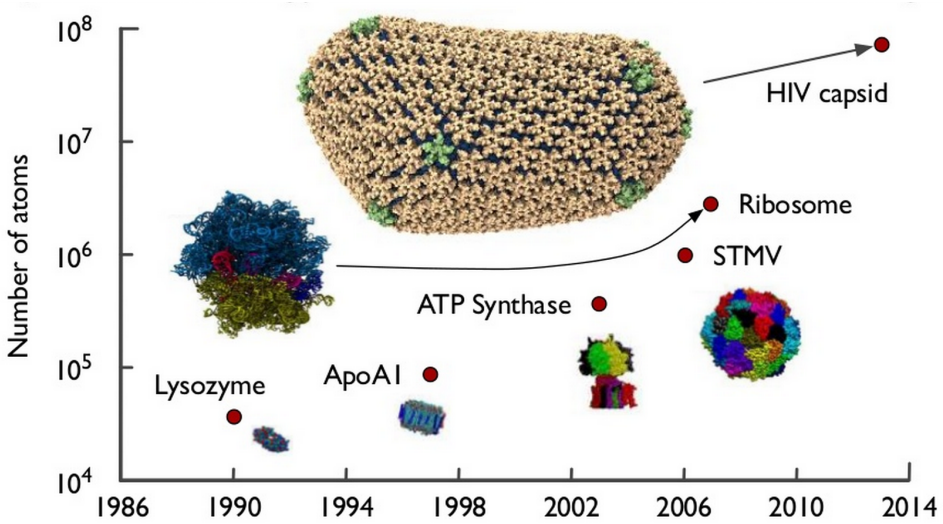

Figure 10: Increase in the size of the systems simulated through MD simulations during the years. Reprinted from[48]

The gap between the timescales accessible through experiments and the ones accessible through simulations is being filled, but the study of substrates translocation is influenced by the need of obtaining a statistics of the process under study. The free energy barrier associated to the translocation of antibiotics is of the order of $10 Kcal/mol$ [17,49] and, as predicted by Boltzmann statistics (see section 2.2), this corresponds to a characteristic time of 10 to 100 microseconds to observe a *single* translocation. It has been already pointed out that entropic effects may change the subtle balance of interactions that govern the diffusion of the antibiotic through OMPs[26]. To study the reactive pathway that antibiotics follow during passive diffusion through pores[38], we therefore need to increase the speed at which the process occurs during simulation by *biasing* the potential experienced by the system, thus altering its *dynamics* (see Sections 2.2.1,3.1). In addition, for a clear and exhaustive analysis of the results, we need to integrate the findings obtained through standard MD simulations with a model of the *kinetics* of the process (see Sections 2.3,3.3).

2

METHODS

Contents

2.1 STANDARD MD SIMULATIONS

Everything that living things do can be understood in terms of jigglings and wigglings of atoms

— Richard Feynman

Molecular Dynamics simulations describe the behaviour of a system with atomistic detail. The basic concept behind MD simulations is that the evolution of a deterministic system is purely given by the *initial conditions* and by solving the equations of motion of the system itself[50]:

$$\frac{d}{dt}\mathbf{q}_i = \frac{\partial H}{\partial \mathbf{p}_i} \qquad \frac{d}{dt}\mathbf{p}_i = -\frac{\partial H}{\partial \mathbf{q}_i} \tag{1}$$

Equations 1 are called Hamilton's equations of the system, and the initial conditions are the starting values of the generalized coordinates $\mathbf{q}_i$ and of their momenta $\mathbf{p}_i$[1]. For the cases we are interested in, the generalized coordinates are neither but the positions of the atoms we are simulating, and the momenta are the products

1 This is true for systems in which the energy of the system is held as fixed (microcanomical, or NVE ensemble); for NVT and NPT systems other parameters, such as the dimensions of the system or the external pressure acting on it, must be provided

$\mathbf{p}_i = m_i\mathbf{v}_i = m_i\partial\mathbf{v}_i/\partial t$. The index i identifies the particle we are considering: $i = 1..N$, N being the number of atoms of the system.

The initial coordinates are chosen by the user in order to represent the case we are interested in; the velocities of the particles are usually randomly generated through the statistical Maxwell distribution relative to a given temperature:

$$\rho(v) = 4\pi \left(\frac{m}{2\pi K_B T}\right)^{3/2} v^2 e^{-mv^2/2K_B T} \tag{2}$$

where $\rho(v)$ is the probability of having a particle possessing speed $v = ||\mathbf{v}||$. The randomicity that determines the velocities of the single particles is reflected on the fact that *statistical properties* of the system don't depend on the conditions of a single atom, but rather of the whole ensemble describing the macroscopic sample. The evolution of the system, however, is strongly determined by the characteristics of the *Hamiltonian* $H = V + T$ of the system, where $T = \sum_i (m_i v_i^2)/2$ is the kinetic energy and V is the potential.

2.1.1 *Stability, Force Field, Periodic Boundary Conditions*

Equations 1 allow to find the evolution of a system with given initial conditions. To calculate the evolution, however, the form of the potential through which the particles interact between themselves must be known. The potential V must, in general, take account for all the possible interactions of each of the particles (total number N) with the other $N-1$ elements of the system: therefore, equations 1 are a set of $2 * N * (N-1)$ coupled equations. Such a *multi body problem* doesn't have an analytical solution and is therefore solved numerically. Therefore, the evolution of the system at a time t_i is calculated from the conditions of the system at the time $t_{i-1} = t_i - \Delta t$. The numerical algorithm used for integrating differential equations 1 used in this work is the Verlet integrator[51]. Δt is called the *timestep* of the simulation, and its value is important to determine the stability of the simulation. A simulation is unstable if errors on the calculus of the quantities enlarge over successive iterations, making quantities such as the energy of the system divergent. In our simulations, thanks to M-SHAKE algorithm[52] and to hydrogen mass repartitioning[53], we were able to use a timestep ranging from 2 to 4 femtoseconds, that is, the same order of magnitude of the greatest frequency of the atomic vibrations.

The form and the parameters of V, on the other hand, must be chosen to obtain agreement with experimental data. The functional

form of *AMBER* (Assisted Model Building with Energy Refinement) force field is[54,55]:

$$V(\mathbf{r}^N) = [V_{stretches} + V_{torsions} + V_{dihedrals}] + V_{VdW} + V_{electrostatic} \quad (3)$$

Equation 3 considers the total potential exherted by a group of N atoms as the sum of two main contributions: the first one is given by the effect of *bonded interactions* and can be summarized as follows:

$$\begin{aligned} V_{bonded} &= [V_{stretches} + V_{torsions} + V_{dihedrals}] \\ &= \sum_{bonds} k_b(l - l_0)^2 \quad (4) \\ &\quad + \sum_{angles} k_a(\theta - \theta_0)^2 \quad (5) \\ &\quad + \sum_{proper\ torsions} \sum_n \frac{1}{2} V_n[1 + \cos(n\omega - \gamma)] \quad (6) \\ &\quad + \sum_{improper\ torsions} \sum_n k_{dih}(\omega - \omega_0)^2 \quad (7) \end{aligned}$$

The first part of previous formula (term 4) accounts for stretching of the bonds, that is, changes in the distance between the positions of two atoms (Figure 11A); in an analogous way, term 5 is relative to the torsion of a group of 3 atoms (Figure 11B). Both these terms mainly serve to reproduce the topology of the molecules and can be treated with a parabolic potential in the form $const * (q - q_0)^2$, where q_0 is the value of the variable q in equilibrium conditions. Proper (part 6) and improper (part 7) are due to the changes in the planes formed by groups of 3 atoms in a group of 4 atoms (Figures 11C and 11D). Improper term is used to stabilize planar rings in molecules and is therefore parametrized as an harmonic potential in a similar way of $V_{stretches}$ and $V_{torsions}$; the form of the potential relative to proper dihedral term 7 (n is an integer number > 0) allows to consider the multiplicity of the chemical bond described by a specific dihedral angle.

It is important to notice that the numerical value obtained with the use of V_{bonded} is heavily dependant on the values of both the force constants and the equilibrium values for the single terms: these values are taken by experimental data such as crystallographic and phase transition measurements and determine the overall goodness of the potential used. Indeed, AMBER force field contains the parameters for proteins only; an addition to include the cases of lipids and other molecules, called GAFF (Generalized Amber Force

Field)[56] has been extensively used to parametrize the lipid bilayer and the substrates used in our simulations. The parameters relative to the antibiotics have been improved through explicit simulations and Density Functional Theory and are reported in a recent work by our group[57].

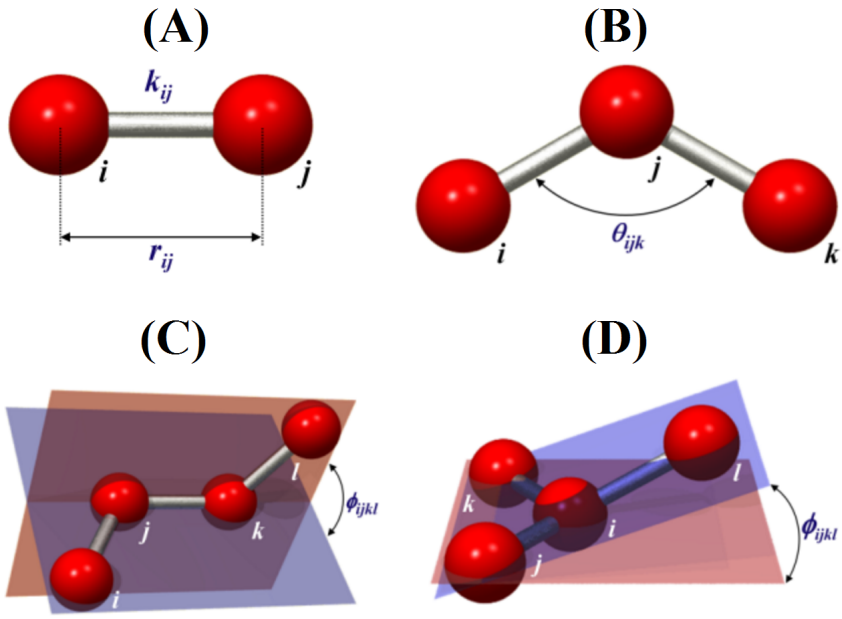

Figure 11: Types of interactions between bonded atoms. (A) bond stretch, (B) bond torsion, (C) proper dihedral and (D) improper dihedral

Terms 4-7 account only for the interactions that are present only in between atoms being part of the same molecule. It is easy to understand that, in addition, we have to consider the interactions between couples of atoms not necessarily connected by a chemical bond:

$$
\begin{aligned}
V_{unbonded} &= V_{VdW} + V_{electrostatic} \\
&= \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^{6} \right] \qquad (8) \\
&\quad + \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \frac{q_i q_j}{4\pi\varepsilon r_{ij}} \qquad (9)
\end{aligned}
$$

The first term (referred as 8) has the form of a Lennard-Jones potential and takes account to the Van den Waals forces between two generic atoms, that is, the repulsion that prevents two atoms from intersecting each other at distances that are lower than the dimensions of the atom itself; the second one (term 9) is the electrostatic

potential existent between two electrically charged atoms and has the form of a Coulomb interaction. In the force fields used in this thesys, polarization is not taken into account and water molecules are modelled as a simple, rigid molecule through the TIP3P parameters[58].

In most cases simulated with MD, the need to avoid artefacts at the boundaries of the unit cell suggests the use of *periodic boundary conditions* (PBC) to describe the system (Figure 12). This corresponds to suppose that there are infinite copies of the system in all the directions, thus avoiding the need to parametrize the interactions of the particles with the walls defining the dimensional boundaries of the system. This approximation is particularly acceptable in the case we are studying because, as stated in page 7, multiple copies of influx proteins are located in the OM of GN bacteria. In all the simulations reported in this work, we used *rectangular* PBC.

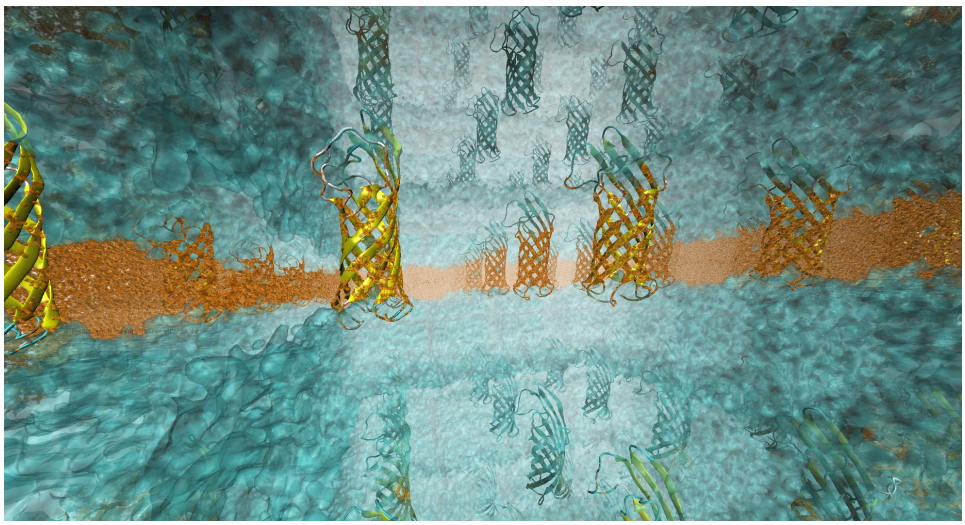

Figure 12: Example of a system represented with its periodic images

Although the form of $V_{unbounded}$ is simpler and requires less preliminary calculations than $V_{bounded}$ (the parameters for $V_{unbounded}$ have to be parametrize only once per new types of atoms), most of the computational effort in solving the value of $V(\mathbf{r}^N)$ for a given time comes from the unbounded term. This is the result of two simplifications: first, the number of atoms in the summations of *bonds, angles, proper torsions* and *improper torsions* terms is far lower than the number $N * (N-1)$ of particles to be considered for the all the pairwise interactions relative to the unbonded part of the potential. In addition, while the term V_{VdW} has the form of a 6-12 potential,

and can thus be regarded as a short range force, $V_{electrostatic}$ doesn't have this feature.

In practice, a cut-off (whose value is usually chosen at around 10Å) is considered for V_{VdW}, and the calculations for atoms having distances greater than the cut-off are avoided. The main limitation is that the dimensions of the system have to be greater than twice the cut-off value to avoid *self-interactions* of the system with its periodic images. The calculus of $V_{electrostatic}$, on the other hand, requires to consider the whole periodicity of the system. To reduce the computational effort, techniques like *Ewald summation technique*[59,60] and its numerical implementation for large systems, *Particle Mesh Ewald*[61], have been widely used.

2.1.2 *Equilibration Protocol*

As stated in the previous chapter, MD simulations are used to obtain a microscopic description of experimental set-ups. Because of this, they have been referred to as *in silico experiments* in the last years. Indeed, modeling and simulation live at the intersection between theory and experiment. As experimental works need to equilibrate the system under study before being able to run the actual measurements, so do the simulations. The creation of the system constituted by a protein inserted in the OM, represented in our simulations by a lipid bilayer, the addition of the solvent and of eventual substrates and the filling of the channel(s) are referred to as *equilibration*, a necessary step for the correctness of the successive computed quantities.

The protocol used for preparing all the systems reported in this work is given by:

1. Energy minimization of the protein in vacuum

2. Superimposition of the protein with a preequilibrated POPCs lipid bilayer and deletion of the lipids within a distance of 3Å from the protein (Figure 13A). The primary sequence and tertiary structure of the protein can be directly obtained from crystallographic data[17], taken from online databases[62] or deduced from similar proteins with homology modelling[63,64]

3. Relaxation of the lipids around the protein (Figure 13B). This is done with successive NVT[2] and NPT[3] simulations with increasing temperatures. The positions of the protein's $C_\alpha s$ is restrained in all the directions to avoid deformations of the tertiary structure; the lipids' phosphorous heads are restrained in the coordinate perpendicular to the plane of the membrane to held the width of the bilayer fixed while allowing for the lipids to pack around the protein.

4. Addition of water (Figure 13C) and relaxation of the water with successive NVT and NPT simulations up until the temperature we wish to simulate, namely 300K. The restraints on the protein's $C_\alpha s$ and on the POPCs' P-heads are removed only after the water has filled the channel (Figure 13D). The ratio between the dimension of the unit cell are allowed to change during the NPT equilibration, that is, the pressure bath is semi-isotropic

2.2 RARE EVENTS AND ENHANCED SAMPLING TECHNIQUES

The degree of slowness is directionally proportional to the intensity of memory. The degree of speed is directionally proportional to the intensity of forgetting

— Milan Kundera

According to Boltzmann statistics, the probability of an event in a canonical ensemble, that is, a system held at constant temperature, is given by[69]:

$$P(event) = \kappa e^{-\Delta E / K_B T} \tag{10}$$

that is dependant on the ratio between the energy required by the process ΔE and $K_B T$, that is a measure of the average available energy in the system. The prefactor κ is dependant on the specific reaction and is proportional to the ratio $\omega_0 / \sqrt{K_B T}$, where $\omega_0 = 1 ps^{-1}$.

The main implication of the above formula is that, at thermal equilibrium, processes that require a greater amount of activation

2 NVT means that the number of particles, the volume and the temperature of the system are held as constant. This result is obtained with the addition of a thermal bath[65–67] coupled with the system

3 In NPT simulations the number of particles, the pressure and the temperature of the system are held as constant. To do so, a pressure bath[65,68] is added to the thermal bath. The pressure can be either considered as isotropic (all the dimensions of the system vary by the same ratio), semi-isotropic (2 dimensions vary by the same ratio, the 3rd is independent) or anisotropic (all the 3 dimensions vary independently)

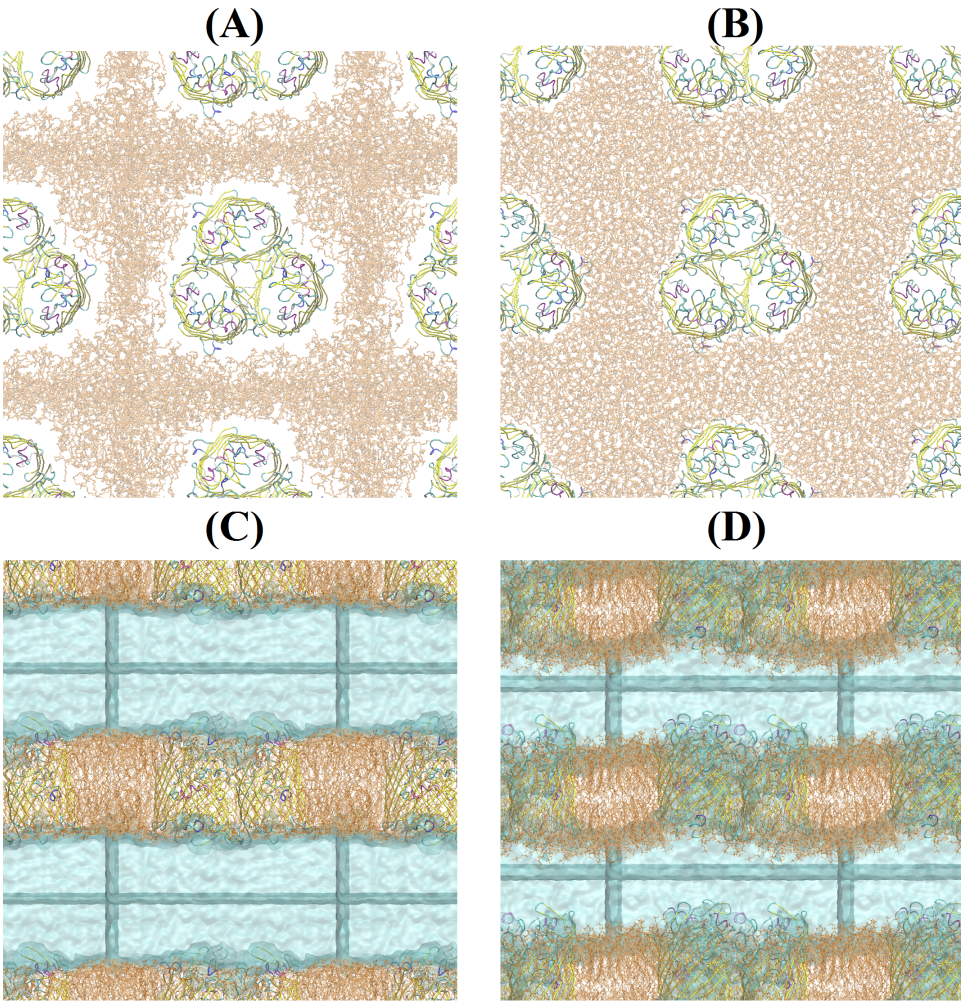

Figure 13: Schematic scheme of the protocol used for equilibrating the systems: (A) insertion of the protein into a pre-equilibrated lipid bilayer and deletion of intersecting lipids; (B) relaxation of the lipids around the protein; (C) addition of water; (D) relaxation of water and filling of the channel

energy ΔE will be less probable. In particular, at $T = 300K$ (sometimes referred to as room temperature), we obtain that the probability of the realization of a process that requires an energy of $10Kcal/mol$ is 0.023% of the probability of a process that requires an energy of $5Kcal/mol$. Indeed, the passage of antibiotics in outer membrane proteins is usually predicted to require an activation energy of $\approx 10Kcal/mol$[17,49]. This means that translocations of antibiotics are *rare events* in the time scales studied in our simulations. Practically, the time required for translocation can range from a few[42] to some hundreds[41] of μs.

The average accessible energy at room temperature is $K_B * T \approx 0.6 Kcal/mol$. The reason for which the average time required by translocation is still in the order of microseconds$= 10^{-6}s$ is that the average time in between energy exchanges is of the same order of atomic vibrations, that is, femtoseconds$= 10^{-15}s$. This corresponds to the average time between two consecutive hits in between particles. However, as stated in the previous section, we are forced to use a timestep that is of the same order of the atomic vibrations in order to keep the system stable and avoid divergent and unphysical behaviour. Therefore, even with the last innovations in computational power, we are currently able to simulate only times up to hundreds of microseconds. To obtain a *statistic* for the process under study, we need to increase the speed of the exploration.

2.2.1 *Metadynamics*

The translocation of substrates is a process during which some molecule goes from a low-energy condition (that is, being surrounded by bulk water) to another. The translocation itself is realized by passing through a series of local minima, usually realized when the substrate forms bonds with specific parts of the protein.

To obtain an *homogeneous* exploration of the phase space[4], an *external bias* can be added to the original potential depicting the system. This method consist in iteratively adding a potential that discourages the system from returning to the conformations already seen during the simulation[49,70–77]. This allows to increase the fluctuations in the free energy of the system to increase the speed at which transitions occur.

The method uses a a gaussian-shaped potential whose center coincides with the position of the system in the phase space. Notice, in accordance with Figure 14, that the added bias will flatten the free energy hypersurface in which the system moves: therefore, the percentage of the time spent in the bottlenecks of the process is increased. This allows to obtain an improved sampling of the conditions that inhibit the passage of compounds in a peculiar porin.

4 Also referred to as conformational space, the phase space of the system is the $2N$-dimensional space in which each point correspond to a specific value for all the components of the coordinates $\mathbf{r}_i$ and momenta $\mathbf{p}_i$. Systems with boundary conditions like *canonical ensemble* NVT have a number of freedom degrees lower than 2N and that corresponds to the system being constricted in an *hypersurface* of the whole space

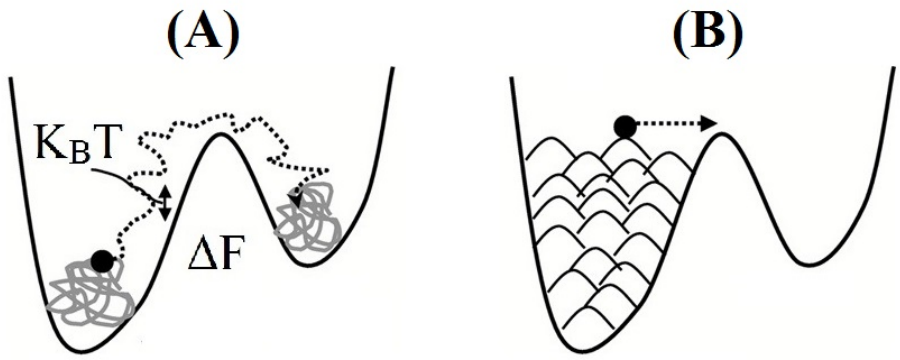

Figure 14: Schematic scheme of the biased potential used in Metadynamics method. A system may require free energy jumps much greater than the average fluctuations in energy at a given temperature (A); In Metadynamics[49,70–77], an history-dependant potential is added discouraging the system from visiting already explored conformations (B). Reprinted from[49]

The total bias that has been added in the simulation at a given time is:

$$V_G(S(\mathbf{q}^N, \mathbf{p}^N, t) = \\ = \omega \sum_{\substack{t' = \tau, 2\tau, \dots \\ t' < t}} \left[\exp\left(\sum_{\alpha=1}^{d} \frac{-[S_\alpha(\mathbf{q}^N(t), \mathbf{p}^N(t)) - S_\alpha(\mathbf{q}^N(t'), \mathbf{p}^N(t'))]^2}{2\Delta_\alpha^2} \right) \right] \quad (11) \\ \xrightarrow[t \to +\infty]{} -F(x)$$

where τ is the frequency of addition of the hills, ω is the height of the added potential and Δ_α its width. The index $\alpha = 1, .., d$ identifies the direction, in the phase space, in which we are adding a bias. In general, $S_\alpha(\mathbf{q}^N(t), \mathbf{p}^N(t))$ can be *any* linear combination of the generalized coordinates $\mathbf{q}^N(t)$ and of their momenta $\mathbf{p}^N(t)$ at a given time.

A previous knowledge of the free energy difference required for the transition, as well as of the location or the number of jumps, is not necessary; however, the convergence of the method strongly depends on the *choice* of the *d collective variables* (CVs) $S_\alpha(\mathbf{q}^N, \mathbf{p}^N)$. Indeed, if one or more collective variables that determine the rates of the transition are not biased, the speed-up in the exploration can be negligible, or the energy basins can be over-filled and unwanted events (e.g. break-up of the protein) can be observed[73]. The CVs should clearly distinguish between all the relevant states of the transition, as well as all the slow events that are relevant for the pro-

cess[73]. By actively enhancing the fluctuations on all the slow degrees of freedom that depict the process, the speed of the simulation is selectively increased in the direction required for the process to happen.

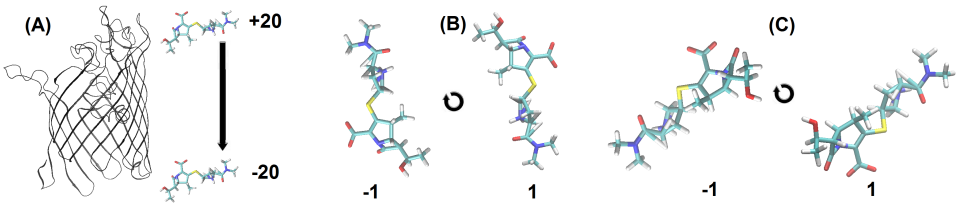

Figure 15: CVs used for exploration enhancement in the permeation of compounds in porins. (A) distance between che center of masses; (B) Projection of the orientation of the substrate on the direction perpendicular to the OM; (C) Projection of the dipole moment of the substrate on the direction perpendicular to the OM

We define the *axis of diffusion* as the direction perpendicular to the plane of the membrane. The CVs used for our studies (Figure 15) are the projection on the axis of diffusion of the following quantities:

1. Distance between the centres of mass of the substrate and the one of the protein
2. Orientation of the substrate
3. Dipole moment of the substrate

The first choice is a natural way to describe the process of a molecule moving between two specific regions, namely the extracellular side and the periplasmic space. We therefore biased this CV in all our simulations. It has been already pointed out[73] that the efficiency of the method scales exponentially with the number of CVs biased. Indeed, the time required for filling the free energy well (Figure 14A) is proportional to $(1/\Delta^2)^d$, Δ being the d-dimensional width of the added bias at a given time. This drawback has forced us to restrict the exploration to 2 dimensions. The second variable has been chosen as either the *orientation* (Sections 3.2,3.3) or the *dipole moment* (Section 3.1) of the substrate. The definition of the orientation is strongly molecule-dependant; we can roughly say that the orientation is barely parallel to the main axis of inertia of the molecule. Biasing the orientation helps us in overcoming free energy barriers due to *entropic* effects, as it is a (rough) measure on how small the minimal projection area of the antibiotic will be. The alignment between the dipole moment of the molecule and the axis of diffusion,

on the other hand, has an *entalpic* character as it is the measure of the strength of the internal electric field of the protein. It is important to note, however, that the dipole moment of a molecule and its orientation cannot be considered as being independent.

The variation of the two collective variables can be pratically shown in a 2D heat map, where the values of the 2 biased CVs are plotted against the Free Energy of the system. This value can be obtained by taking the negative value of the bias added previously:

$$F(S_\alpha) = -K_B T \int e^{-V(S'_\alpha)/K_B T} \delta(S_\alpha - S'_\alpha)\, dS'_\alpha \tag{12}$$

Equation 12 can be easily related to the percentage of the time spent during the passage, that is the probability, through Boltzmann statistics:

$$P(S_\alpha) = \frac{e^{-F(S_\alpha)/K_B T}}{\int e^{-F(S'_\alpha)/K_B T} dS'_\alpha} \tag{13}$$

In the next sections we will refer of these 2D heat maps as *free energy surfaces* (FESs).

2.2.2 *Extensions of Metadynamics*

Independently on the improvements of the exploration, the speed of the simulation is limited by the computational cost of the simulation itself. Albeit already optimized for parallelization, the iterative algorithm necessary for calculating the evolution of a system from the previous timestep is intrinsically sequential: therefore, the loss in speedup can be severe with increasing number of processors (Figure 16).

A possible way to overcome the problem is represented by launching multiple single simulations (i.e. walkers) in different regions of the space. The ergotic hypotesis that ensemble and time averages are equivalent, i.e. that $\langle A \rangle_{ensemble} = \langle A \rangle_{time}$ for any observable A, allow to treat equivalently the sum of the simulations of the walkers and a unique, longer simulation. The need to explore uniformly the phase space suggests to restrict the movement of the single simulations in specific regions of the space, e.g. different depths of the channel. This can be done imposing bias walls for given values of the selected CVs. The presence of these walls can substantially alter the potential at the boundaries. To reconstruct a unique FES, techniques like WHAM (*Weighted Histogram Analysis Method*)[79–81] must be used in order to obtain a continuous FES (Figure 17).

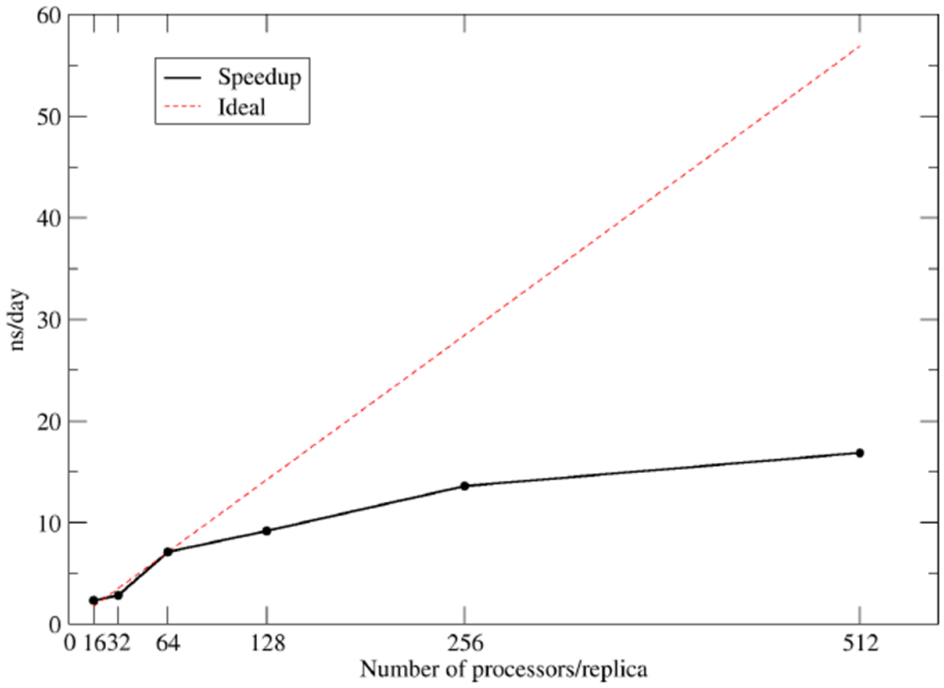

Figure 16: Comparison between ideal (red dashed curve) and real (black curve) speedups obtained with the use of Gromacs 4.6.5[78] + Plumed 2.0.4[73,77] on Curie cluster[47]

In *Multiple Walkers Metadynamics* (MWMeta)[71,73], multiple simulations contribute to the reconstruction of the same FES. The efficiency is improved by exchanging only informations on the added bias and by treating the walkers as single metadynamics simulations otherwise. The bias added by the other walkers discourages the simulations from exploring conformations that have been already seen, therefore we don't need to use walls that could produce artefacts in the phase space.

A metadynamics run reaches *convergence* when the free energy surface is completely explored. As stated earlier, the action of the bias flattens the experienced potential, therefore in principle we should get a flat potential in throughout the system (free diffusion); however, the constant addition of the bias will eventually overfill the region of the FES we want to observe. This will correspond to have an error on the calculus of free energy differences proportional to the added height of the hills w. *Well-Tempered Metadynamics* (WTMeta)[74] consists in decreasing exponentially the height of the added bias. This allows to improve the convergence of the algorithm. In figure 18, we see how the scaling of the potential helped in reaching a stable and flat FES in the case of the translocation of Imipenem

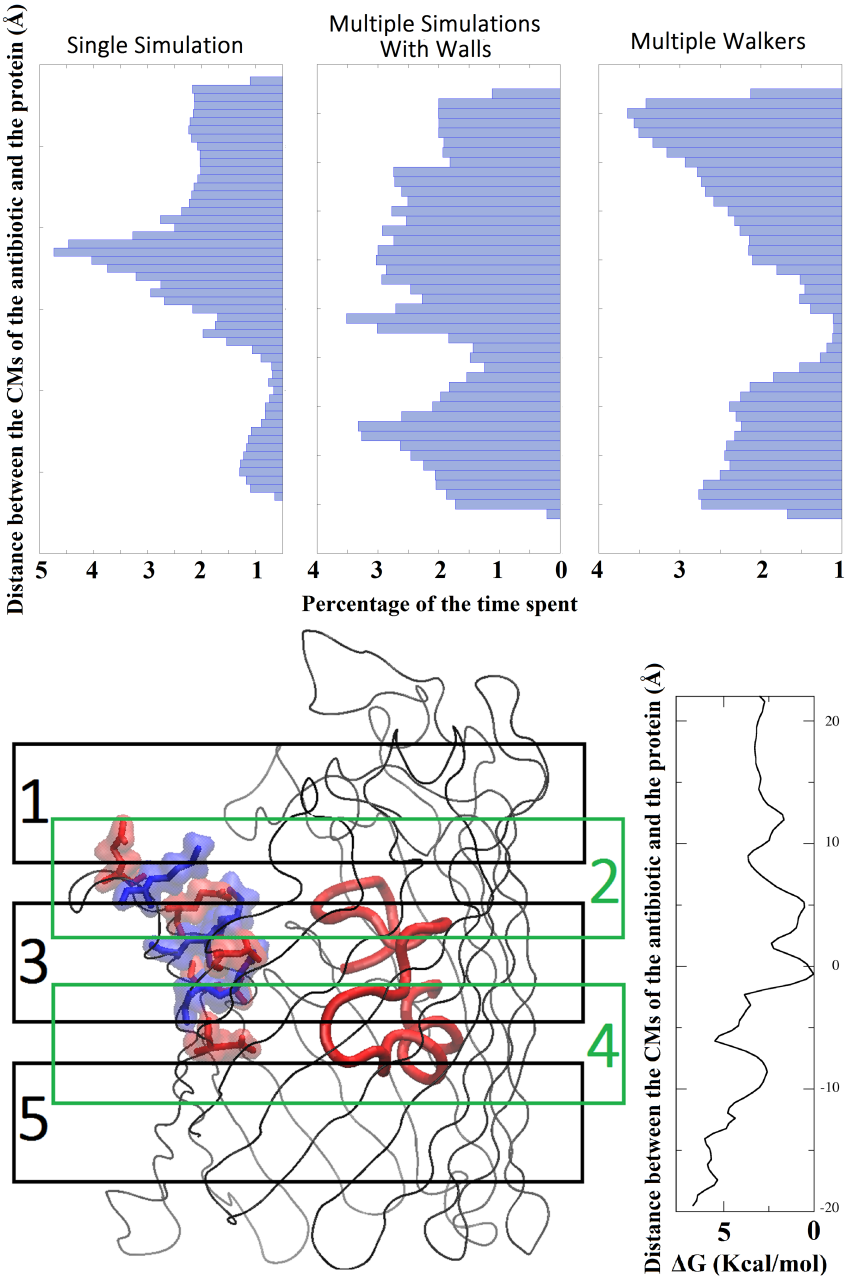

Figure 17: Comparison of the time spent at varying sections of the protein for different methods. Up: Percentages of the time spent at different sections of the protein for a single metadynamics run[70] (left), multiple metadynamics simulations with walls (centre) and MWMeta[71,73] (Right). Bottom: Position of the walls used, with emphasis on the electric charges of the protein in the constriction region (left); Final Free Energy profile obtained with MWMeta (right).

in *Escherichia coli*'s OmpC. All metadynamics simulations reported in this work make use of WTMeta.

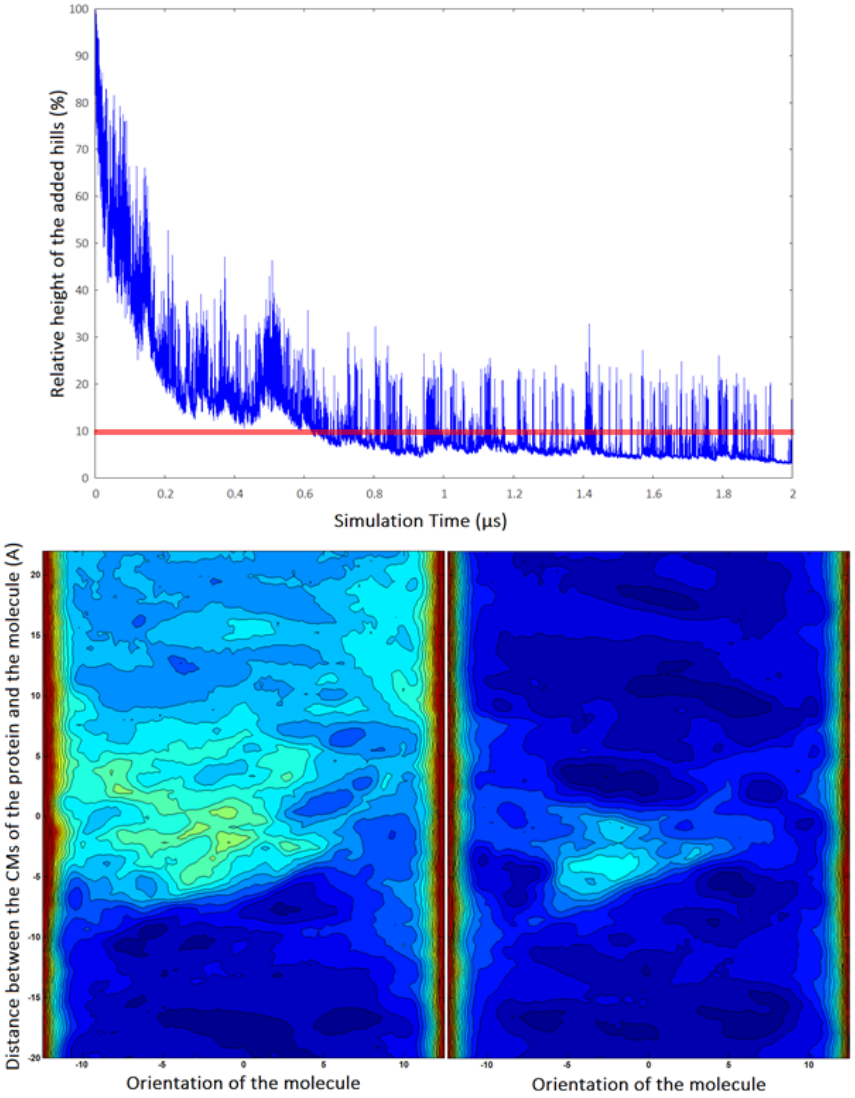

Figure 18: Convergence of the simulations with WTMeta[74]. Top: scaling of the bias as a function of the length of the simulation. Bottom: comparison between the obtained FESs at 1μs (left) and 2μs (right).

2.2.3 *Replica Exchange methods*

We refer to Replica Exchange Methods as those techniques in which two or more simulations are run in parallel and exchanges of the coordinates between different simulations are attempted with a *Monte Carlo* scheme. The acceptance criteria, that is the probability of two

replicas to swap the coordinates of their particles at the next step, is given by[82]:

$$P(i \leftrightarrow j) = min\left[1, exp\left(\frac{1}{K_B T_i} - \frac{1}{K_B T_j}\right)(U_i - U_j)\right] \qquad (14)$$

Where U_i and U_j are the values of the potential energy for the two replicas. If the exchange is done, the velocities are scaled by a factor T_i/T_j and simulations continue.

The choice of the CVs to bias is usually the most crucial parameter to set in a metadynamics simulation; since adding another CV enlarges the dimensionality of the FES to explore, exploiting more than 2 CVs at the same time is discouraged. To limit the exploration of the phase space to a subregion, while biasing a greater number of CVs, *Bias Exchange Metadynamics* (BEMeta)[73,83] can be used. This corresponds in taking $T_i = T_j$ in Equation 14. Since different simulations bias different degrees of freedom, multiple biases allow to distribute the computational power between the replicas; since not all the CVs are explored uniformly, the computational cost increases linearly with the number of CVs, rather than in an exponential measure. The main drawback of the technique is that it is not possible to give to the deposited bias the meaning of Free Energy Surface for obtaining the energy barriers encountered.

An improvement of the technique has been proposed as *minimum mode Metadynamics*[84] (MMMeta): in this case the potential is added linearly in the direction in which the gradient of the total potential shows the greatest variation. This allows to find a multidimensional trajectory in the phase space even in the case of a single replica.

Parallel tempering[85–87] consist in exchanging replicas that differ for their temperature T. The main advantage of the method is that increasing the temperature increases the speed of *all* the degrees of freedom because it corresponds to increasing the average speed of the particles within the simulation. If used in conjunction with metadynamics scheme, we speak of *Parallel Tempering Metadynamics* (PTMeta)[72,73].

In the last versions of Plumed algorithm[77], the possibility of biasing the whole potential energy of the system has been added. Therefore, it is possible to increase the flutuations of U_i and U_j in equation 14 and to obtain non-negligible exchange rates even for replicas at very different temperature. This tecnique is usually referred to as *Well Tempered Ensemble Metadynamics* (WTEMeta)[75,88].

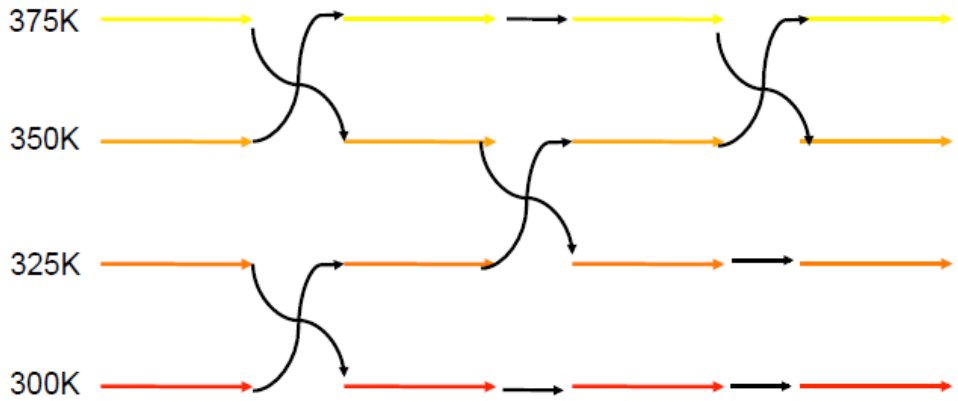

Figure 19: Example of parallel tempering scheme: multiple simulations at varying temperature run simoultaneously; at a given frequency, exchange between consecutive replicas are attempted based on criterion 14

2.2.4 *Coarse Grained Simulations*

All the methods introduced so far improve the exploration of the phase space by iteratively changing the potential felt by the atoms (Sections 2.2.1,2.2.2) or by connecting the system with thermal baths at higher temperatures (Section 2.2.3). Another possibility is to improve the speed of the simulation by enlarging the timestep Δt used for integrating the equations of motion (Equations 1,3). As stated in page 18, the main reason for limiting the timestep is given by stability purposes. Indeed, the timestep cannot exceed the frequency of the fastest atomic vibrations of the system or the system will become unstable and the simulation will explode.

We can reduce the detail of the representation of the system: in this case we speak of *Coarse Grained* (CG) models, in contrast with the *Fine Grained* (FG) representation in which each atom of the system is treated as a single particle. In *Martini* force field[89–92] small groups of atoms are parametrized as one particle only (Figure 20Top).

This allows to increase the timestep to $10 - 15$ femtoseconds; moreover, the reduction in the number of particles to simulate reduces the computational cost of CG simulations with respect to a full atomistic representation. The time that it is possible to simulate with CG simulations is therefore $2 - 3$ orders of magnitude larger compared to FG ones, providing a bridge between the atomistic and the mesoscopic scale[90]

The parametrization of the force field ensures that experimental distributions, phase transitions and macroscopical characteristics

are conserved in the new CG representation; once the system is simulated, it can be reconverted to the FG scale[93]. The lack of detail with respect to atomistic resolution may play an important role in depicting a process in which a subtle balance of interactions governs the passage[26]. In Figure 20Bottom, we show the comparison between the accessible area obtained with FG (Amber/GAFF) and CG (Martini) representations of *Providencia stuartii*'s OmpPst1. The reduction of the accessible area in the pore is almost absent because of the low level of detail in the representation of the protein. On the basis of these preliminary results, we have decided to restrict our simulations to FG models.

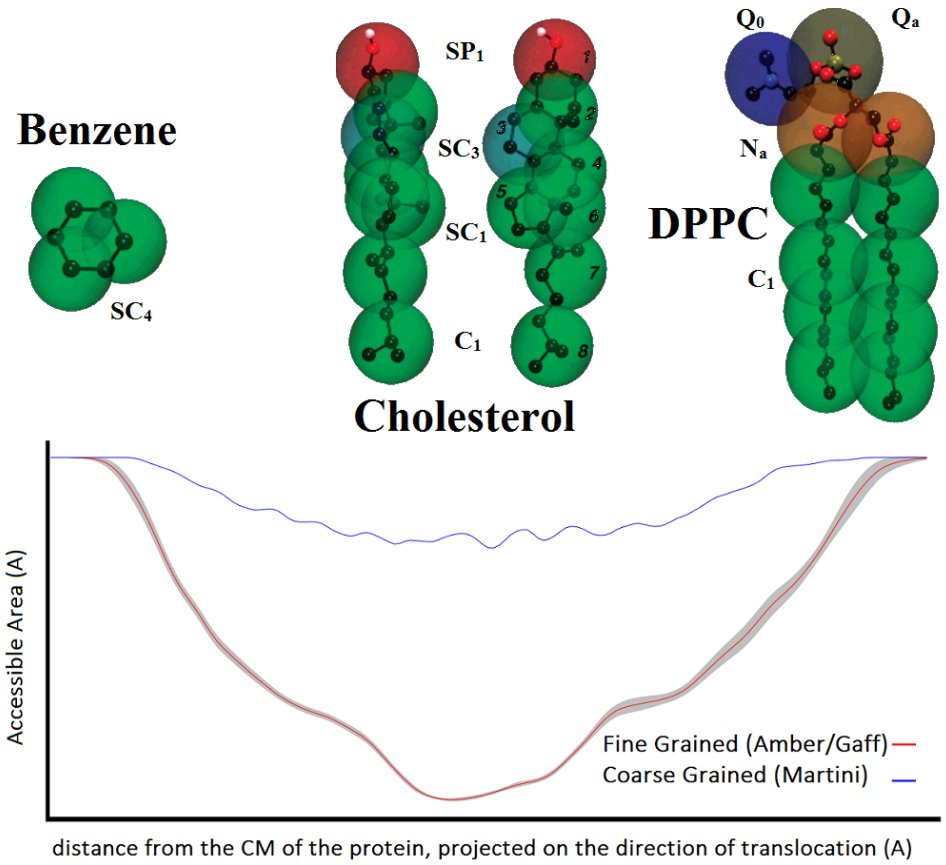

Figure 20: Top: mapping between the chemical structure and the Martini coarse grained model for benzene (left), cholesterol (centre) and DPPC (right). Redrawn from[89]. Bottom: comparison of the minimal accessible area obtained through fine grained (Amber/GAFF) and coarse grained (Martini) models for *Providencia stuartii*'s OmpPst1

2.3 MODELLING THE KINETICS OF THE PROCESS

Only in the cooperation of an enormously large number of atoms do statistical laws begin to operate and control the behaviour of these assemblies with an accuracy increasing as the number of atoms involved increases. It is in that way that the events acquire truly orderly features

— Erwin Schrödinger

In sections 2.2.1-2.2.3 we showed presently available computational methods that allow to increase the speed at which a specific process is realized. This result can be obtained either with an alteration of the potential felt by the particles (Sections 2.2.1,2.2.2) or by altering the temperature of the system (Section 2.2.3), directly

connected with the average speed of the atoms (that is, the average kinetic energy) through Maxwell-Boltzmann distribution.

One of the main drawbacks of these methods is that dynamical properties, such as transmission coefficients, cannot be estimated because of the different potential that the systems under study are experiencing. In order to compare simulation results directly with experiments, a kinetic model needs to be introduced with an a posteriori analysis of trajectories that is usually dependent on an initial guess of the diffusion constant of the system[94]. Moreover, the critical dependence on the free energy barrier encountered during the transition[95] may affect the calculation and, therefore, the goodness of the results[94].

Recently, the scope of Metadynamics has been extended. By applying an a posteriori analysis one can obtain rates of transitions and rate-limiting steps of the process under study[96,97], directly comparable with kinetic data extracted from electrophysiology experiments. The basic idea can be roughly summarized by treating the modification of the potential caused by the use of metadynamics as a perturbation: if this is the case, the alteration of the speed of the simulation can be properly depicted with a Taylor series expanded up until the first order.

$$\alpha(t) = \frac{1}{t}\int_0^t e^{-V[\mathbf{S}(\mathbf{q}(t'),\mathbf{p}(t'))]/K_BT}dt' \tag{15}$$

Equation 15 relates the *acceleration* of the time scale obtained in a metadynamics run with the bias added during the run itself. The *real* time τ of the process, that is the time that an unbiased simulation would require to get the same results obtained in an *accelerated* time t in a metadynamics run, is obtained with:

$$\tau(t) = \int_0^t \alpha(t')dt' \tag{16}$$

Therefore, in this case the reweighting procedure is straightforward, computationally insignificant and doesn't require the guess of intrinsic properties of the system such as the transmission coefficient of the process under study[94].

The assumption that allows the use of equations 15-16 is that the deposited bias doesn't substantially alter the speed of the transition itself: we speak of *quasistationarity* of the added bias with respect to the hypersurface $\hat{\mathbf{S}}(\mathbf{q}(t'),\mathbf{p}(t'))$ that determines the transition[96]. This is equivalent to say that no bias should be deposited near $\hat{\mathbf{S}}$. To obtain a correct sample of the hypersurface $\hat{\mathbf{S}}$, it is additionally expected that the chosen CVs allow to properly distinguish between all the relevant states of the process.

The procedure we have just described can be joined with a statistical approach by running multiple simulations for the same process[97,98]. To assess the validity of the assumption that the added bias is quasistationary, it has to be checked that the results obtained effectively follow a statistical distribution. This can be conveniently done by running the so-called Kolmogorov-Smirnov (KS) test[98]. Supposing that we are interested in the exit from a certain metastable state (as will be done in Section 3.3), KS test consist in checking that the times obtained for escaping from the state follow a Poissonian distribution[98,99]. For a Poisson process, the cumulative distribution function defines the probability of observing at least one transition by time t as:

$$P_{n\geq 1} = 1 - e^{-t/const} \tag{17}$$

We can fit this distribution with the *empirical* cumulative distribution function constituted by the obtained transition times, after the real times τ are obtained with the use of Equations 15-16. The probability that the theoretical and empirical distribution are compatible is called *p-value*[99]. If the p-value is higher than a certain threshold value, the obtained distribution is considered statistically relevant. Usually, and in the case described in Section 3.3, we chose a threshold value equal to 0.05.

Part II

RESULTS, DISCUSSION AND CONCLUSIONS

3

RESULTS

Contents

3.1 A COMPARISON BASED ON DIFFERENT SUBSTRATES: *acinetobacter baumannii*'S CARO

Wir müssen wissen.
Wir werden wissen.[1]

— David Hilbert

Within the outer membrane (OM) of Gram-negative bacteria, β-barrel proteins composed of 12-18 β-strands mediate the diffusional uptake of small molecules required for cell growth and function[17]. We can distinguish between the different diffusion dynamics of general and specific channels, that correspond to translocation dynamics that are either proportional to the concentration of substrates or, conversely, show saturation-like kinetics (see page 9). The smallest known transport channels are Tsx[100] and NanC[101], both showing a cylindrical structure composed of 12 β-strands.

Small OM proteins with barrels of 10 strands or less are abundant in nature, but are not known to transport small molecules and may instead have structural roles or be involved in cell adhesion and invasion. Due to the small diameters of those barrels, inward-pointing residues from both sides of the barrel interact with each other, such that none of these proteins forms an open channel through the OM. While several of these small barrels appear able to mediate the passage of monovalent ions based on electrophysiological data[102,103], none of them has so far been implicated in transport of small molecules such as amino acids.

1 We must know. We will know.

Acinetobacter baumannii is an opportunistic pathogen that, in the last years, has shown resistance towards multiple families of antibiotics[104]. The appearance of resistance has been linked[105] with the loss of the protein *CarO* (Carbapenem-resistance associated Outer membrane protein); however, a quantitative analysis of the importance of CarO in the passage of compounds was missing because of the lack of a tertiary structure of the protein. In a very recent article[17], we have collaborated with the group that has successfully crystallized all the 3 isoforms of CarO and have characterized the charge selectivity of the channel and the residues that are most important for the translocation of substrates with atomistic detail.

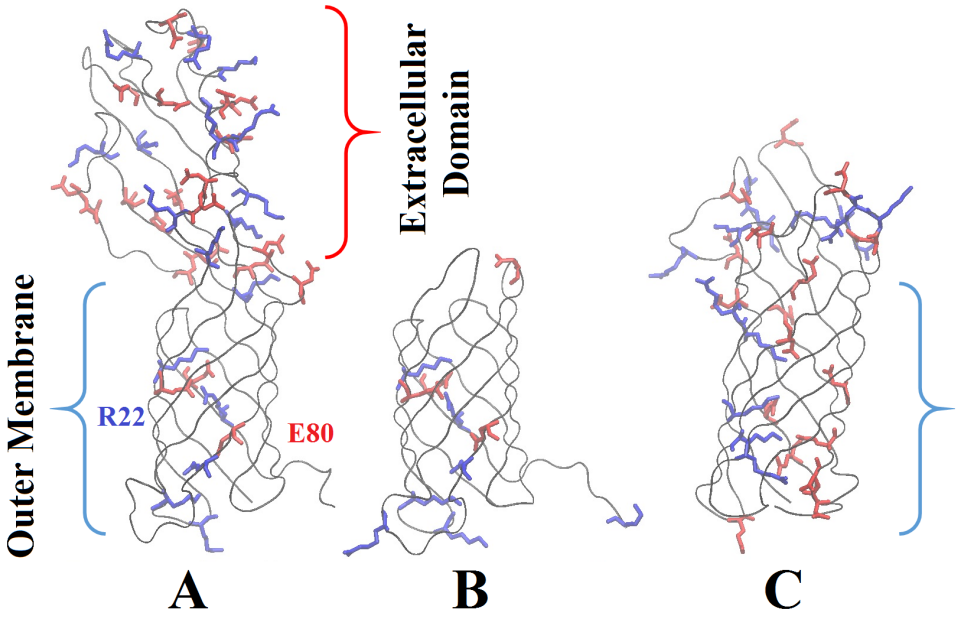

Figure 21: 3D structure for the proteins simulated in this section: CarO2 (A), CarO3Δ (B), OprF_N (C). Positively charged residues are shown in blue and negatively charged ones are shown in red

The shape of CarO (Figure 21A) presents some peculiarities when compared to other 8-stranded proteins such as OprF (Figure 21C), the structural protein OmpA ortholog of *Pseudomonas aeruginosa*[17]. As can be expected for an eight-stranded barrel formed by a relatively large protein (225 residues), three of the four extracellular loops (L2-L4) of CarO are very long, with ≈20-40 residues exposed on the extracellular side of the OM[17]. The three loops form an extended, antiparallel five-stranded β-sheet that is flanked on one side by the α-helix of loop L4. Together, the L2-L4 domain forms a striking structure that overhangs the extracellular entrance to the channel and resembles a baseball glove (Figure 21A). This glove contains the divergent residues between the different isoforms. Indeed,

the variable residues are located almost exclusively in the extracellular domain, especially in the tips of loops L3 and L4. Because of the drastic differences in mobility and structure between the higher and lower part of CarO proteins, they can be considered as *2-domain* OMPs.

We modelled a "loop-less" CarO mutant named CarO3Δ by replacing three out of four extracellular loops each with two glycine residues (Figure 21B). Since the barrel inserted in the OM is virtually invariant between the different CarO isoforms, CarO3Δ represents the truncated version of all three CarO proteins.

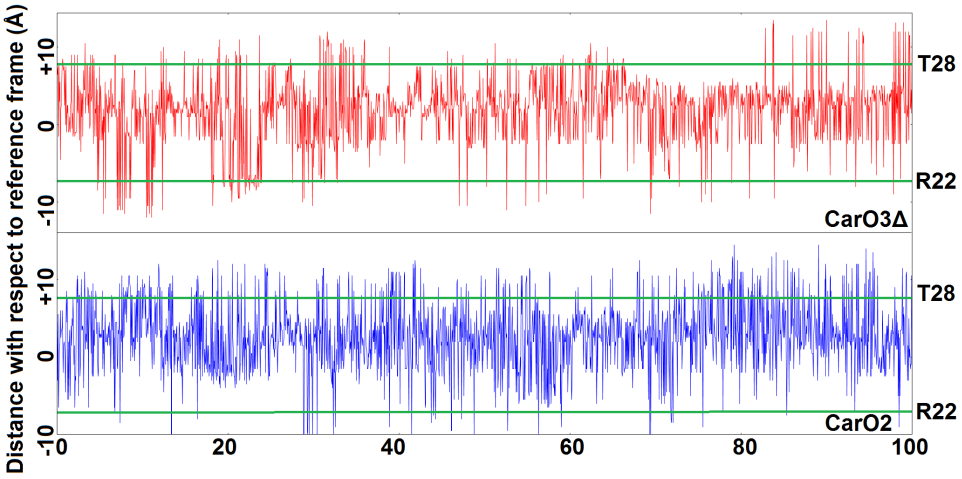

Figure 22: Minimal solvent accessible area location with respect to the first frame of the simulation for CarO3Δ (top) and CarO2 (bottom). Both proteins show considerable changes in the location of the minimal area during simulations, suggesting that multiple internal hydrogen bonds and salt bridges are formed over time, closing the channels. The boundaries of the constriction region, corresponding to the positions of residues T28 and R22, are shown as green lines.

None of the CarO isoforms forms an open channel; this consideration is true even for the loop-less mutant. There are two major areas that block a putative channel: the first is at the channel entrance with a patch of predominantly hydrophobic residues; the second area is approximately in the middle of the membrane. The formation of a salt bridge between R22 and E80 blocks the channel and prevents the formation of a water-filled porin. Ions and other putative transport substrates have to traverse the lumen of the barrel, since CarO is monomeric and the outside of the barrel is higly hydrophobic. We calculated the area accessible to the solvent as a function of time (Figure 22). The most striking feature is that the minimum is exactly zero, corresponding to a channel that is completely closed. Further-

more, the location where the area is zero varies considerably during time, and lies in a region between the hydrophobic channel entrance and the position of the salt bridge formed by R22 and E80. Therefore, despite being closed in the crystal structures, the CarO channel is highly dynamic.

Recent experimental works connect the expression of CarO proteins with *Acinetobacter baumannii*'s supsceptibility towards Carbapenems[105]. In addition, CarO was shown to be involved in the uptake of the amino acid ornithine, with the different isoforms appearing to vary in their ability to mediate passage of this compound[106]. Given the fact that CarO is a small OM protein of 225 residues in length, its purposed role as an uptake channel for amino acids and antibiotics is intriguing, warranting structural studies. To assess its role as an influx protein, we have tested the uptake for zwitterionic glycine, positively charged ornithine and negatively charged glutamic acid (Figure 23top). The use of Liposome Swelling Assays experiments[17,37] allowed to characterize quantitatively the charge selectiveness of the channel (see page 12). Surprisingly, the deletion mutant CarO3Δ expressed very well in the OM and could be used in the experiment. We also measured substrate uptake by the N-terminal barrel domain of OprF (OprF_N) (Figure 21C), a protein with structural role not involved in the translocation of compounds[17].

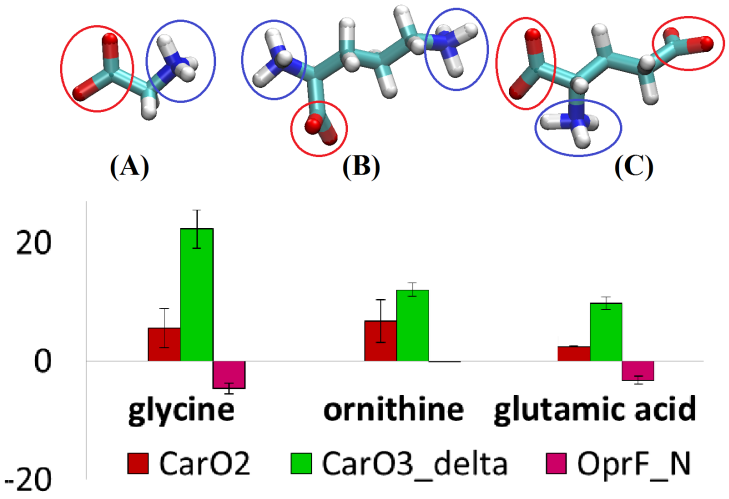

Figure 23: Top: chemical and 3D structure of substrates used in this section: neutral glycine (A), positive ornithine (B), negative glutamic acid (C). The location of positive (blue) and negative (red) groups is highlighted. Bottom: Lipid Swelling Assays results

Liposome swelling experiments (Figure 23bottom) demonstrate that full-length CarO as well as CarO3Δ mutant can mediate low

but significant levels of uptake for glycine and ornithine, but that wild type CarO don't contribute to any uptake for glutamic acid. The low levels of uptake are consistent with the small *in vitro* single channel ion conductance values for CarO of ≈ 20 pS in 1 M KCl solution[107]. Importantly, CarO3Δ displays similar or higher levels of uptake as the full-length proteins, indicating that the measured uptake activities are not due to a low abundance large-barrel form of full-length CarO generated by an alternative folding pathway. The measured activities of OprF_N are, on the other hand, indistinguishable from that of "empty" vesicles: this indicates that OprF_N does not take up any of the used substrates.

The experimental results are confirmed by biased Molecular Dynamics simulations, which allowed us to model quantitatively the transport of selected small molecules at the molecular level (Figure 24 and Table 1). The 3 substrates used for this study differ for both molecular weight, net charge and direction of the dipole moment; to improve the exploration of the phase space, we have chosen to bias (see Section 2.2):

1. The distance between the center of mass of the protein and the substrate projected along the direction of diffusion. A positive value corresponds to the extracellular side

2. The projection of the dipole moment of the substrate with respect to the direction of diffusion

The duration of the single metadynamics simulations, ranging from 500 to 1500 nanoseconds, was roughly proportional to the free energy jump required. We evaluated the barrier for translocation (Table 1) as the difference between the lowest minimum in the extracellular region and the lowest saddle point (Figure 24) with respect to the chosen CVs.

	glycine	ornithine	glutamic acid
CarO3Δ	6	14	16
CarO2	8	14	36
OprF_N	18	26	> 46

Table 1: Free energy barriers associated with the passage of the selected substrates in CarO3Δ, CarO2, OprF_N. The barriers, expressed in Kcal/mol, correspond to the difference between the minimum and maximum value of the free energy in the path with the minimum free energy jump in Figure 24

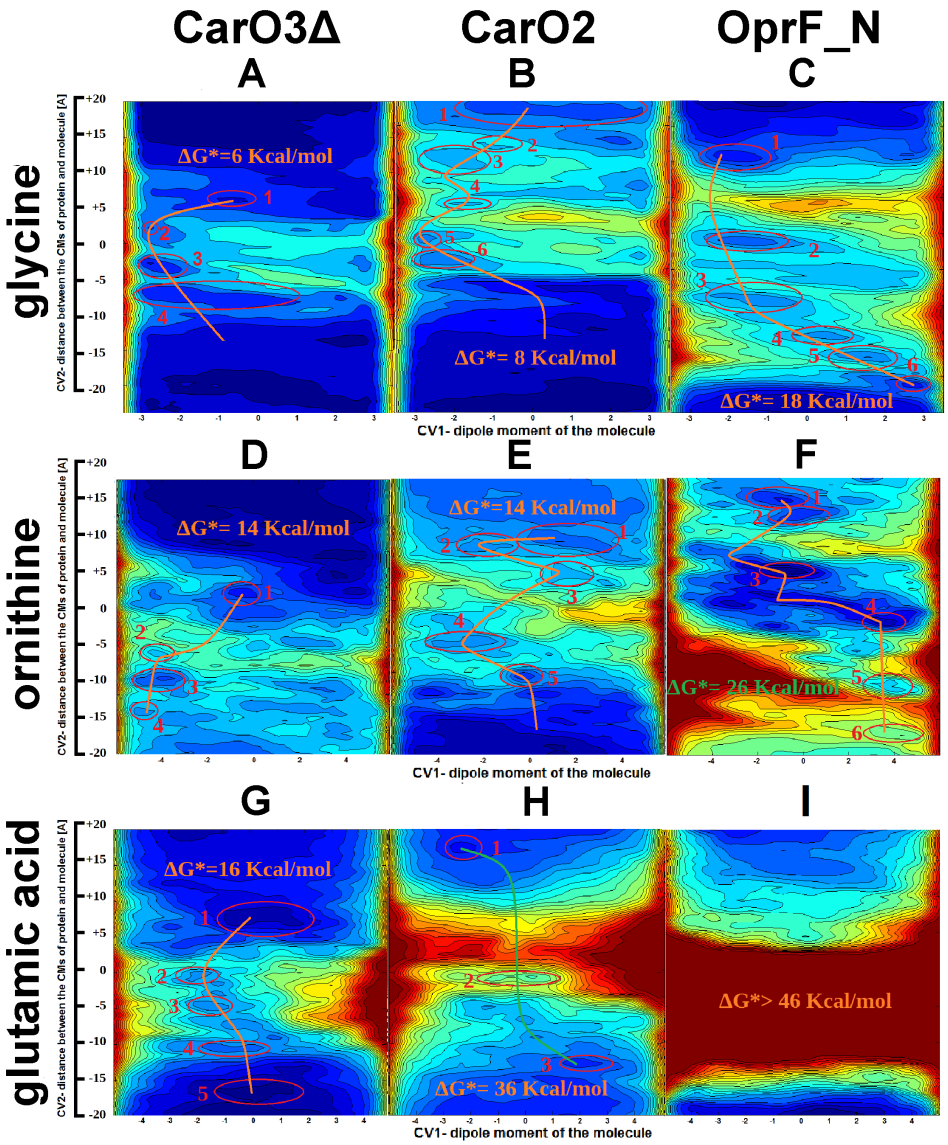

Figure 24: Comparison of small-molecule passage through 8-stranded barrels by MD. Free Energy Surfaces for CarO3Δ + glycine (A), CarO2 + glycine (B), OprF_N + glycine (C), CarO3Δ + ornithine (D), CarO2 + ornithine (E), OprF_N + ornithine (F), CarO3Δ + glutamic acid (G), CarO2 + glutamic acid (H), OprF_N + glutamic acid (I). Each contour line corresponds to a free energy difference of $2 Kcal/mol$. A line corresponding to the path with the minimal free energy jump is drawn for each system, and the corresponding free energy jump is shown

The free energy barrier in CarO2 associated with the uptake of the smallest substrate glycine is 8 kcal/mol. While it is not possible to correlate the exact value of the free energy barrier with uptake rates observed in e.g. liposome swelling experiments, the observed barrier for glycine is relatively small and comparable to the bar-

riers observed for other systems that are known to transport small molecules, e.g. 13 kcal/mol for the uptake of ampicillin in OmpF[26]. Interestingly, the relative values of the free energy barriers for substrate uptake for different 8-stranded barrels correspond reasonably well with our experimental liposome swelling data[17]. In the case of glycine, the smallest barrier for transport is obtained for CarO3Δ (6 kcal/mol), which also has the highest liposome swelling activity of the CarO proteins. Likewise, a high transport barrier of 18 kcal/mol for glycine passage through OprF_N correlates well with the fact that no uptake activity could be measured by liposome swelling. For glutamic acid we observed uptake with liposome swelling in the case of CarO3Δ only and obtained a very high free energy barrier of 36 Kcal/mol for CarO2; in the case of OprF_N, we weren't able to observe any translocation of the substrate and were therefore uniquely able to estimate lower limits for the energy barrier as 46 kcal/mol. The difference between the translocation dynamics in CarO2 and OprF_N seems to be the presence of an additional barrier located in the lower part of OprF_N, a feature that can be easily seen in the case of ornithine translocation; on the other hand, the barrier for CarO is mainly located in between the mouth of the channel and the internal salt-bridge formed by E80 and R22.

After reconstructing the free energy surface in the space of the 2 CVs, we performed 10ns NVT non-biased simulations for each of the single minima identified. The most stable hydrogen bonds formed between the substrate and the proteins were taken as the most important interactions for the selected minima (Figure 25, Table 2). With our MD simulations, we are able to assess the different roles of the two domains of CarO: in the case of glycine, the external glove of CarO helped the translocation of the substrate, bringing the glycine to the mouth of the channel. However, for larger compounds such as glutamic acid, interactions with the external glove led to rejection of the substrate, possibly as a consequence of the presence of many negative charges at the back of the glove. Indeed, the extracellular domain of CarO proteins contains most of the charged residues. It is mainly composed of negative charges: the net charge of CarO2 is -5 elementary charges, while CarO3Δ has a net charge of $+3$.

Comparing the behavior of the loop-less CarO with full-length CarO, we can suggest a double role for the glove:

- It causes a barrier for translocation at the entrance of the channel, by compressing a cluster of hydrophobic residues that are located at the interface between the two domains

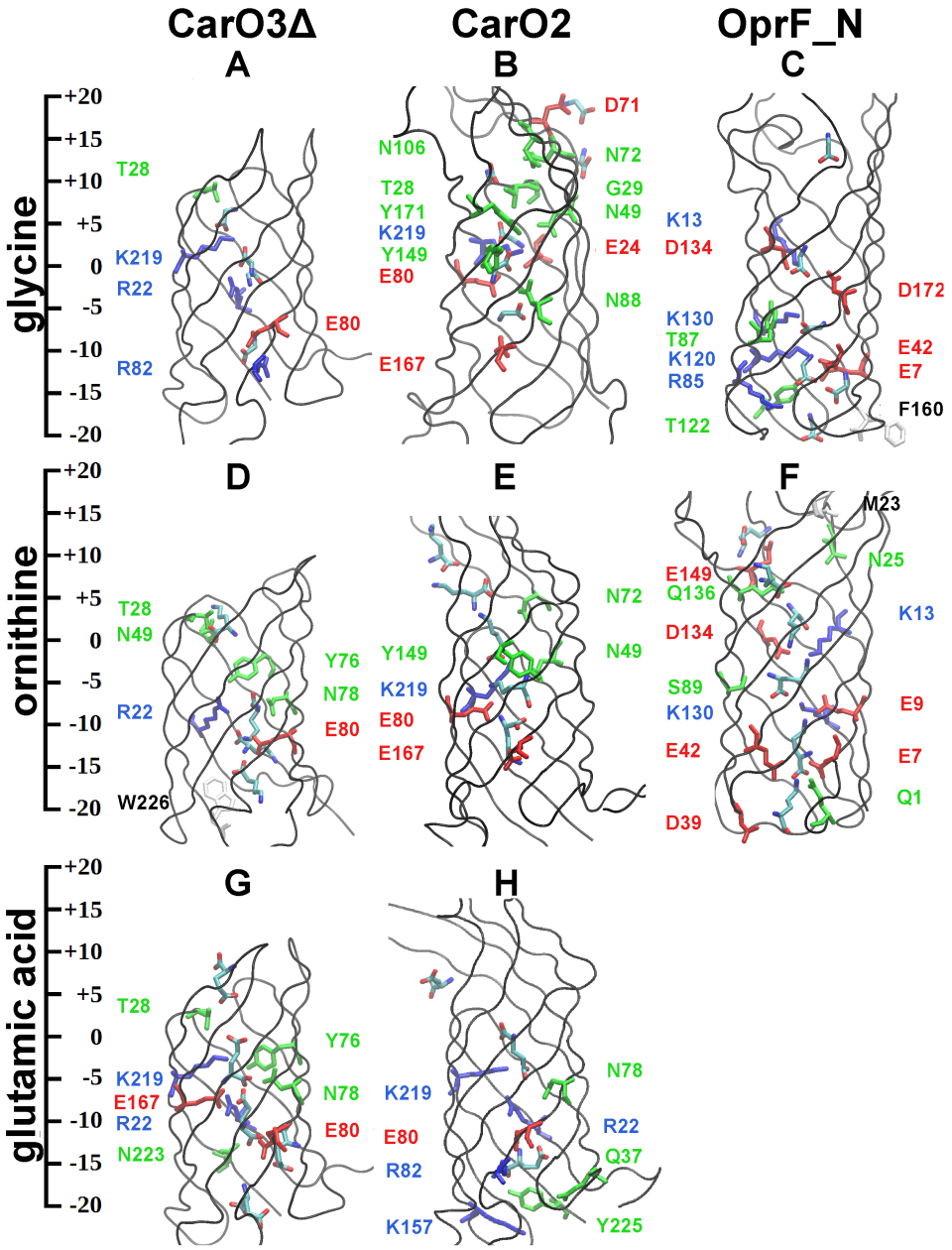

Figure 25: Key interactions for the passage of glycine in CarO3Δ (A), CarO2 (B), OprF_N (C); for the passage of ornithine in CarO3Δ (D), CarO2 (E), OprF_N (F); for the passage of glutamic acid in CarO3Δ (G), CarO2 (H). The positions of the substrates correspond to the minima shown in Figure 24

- It is able to select for positive against negative substrates

The first consideration is supported by the low but present difference between the free energy barriers encountered by glycine in the wild type porin and in CarO3Δ; the fact that the barrier is the same for ornithine can be explained as a counterbalance of the sec-

ond effect, that lowers the barrier necessary for the passage of ornithine. The second effect is also responsible for the huge difference in the energy barriers obtained for going from ornithine to glutamic acid in the cases of CarO3Δ and CarO2: we observe that the increase is equal to 2 kcal/mol for the first and to 20 kcal/mol for the latter. This charge selectiveness might be caused by a relatively large number of negatively charged residues located in the upper domain. Further, when compared with OprF_N, an OM protein with similar characteristics to CarO (with respect to the diameter of the barrel domain), the net effect of the CarO glove seems to be an overall decrease of the energy barrier observed for translocation. This might be a consequence of increased movement due to the external domain of CarO to facilitate a continuous exchange of the intra-molecular interactions observed within the channel constriction zone.

Our simulations indicate that the passage of substrates in CarO family protein can be modelled with a *network disturbance transport* model: the substrate interacts with the residues internal to the protein, disturbing their interactions and causing a partial opening of the channel. This would, in turn, allow for additional diffusion of the substrate and for further dismantlement of the successive internal bonds. It is important to note that the high energy barriers (or, conversely, low liposome swelling rates) associated with CarO proteins can however contribute substancially to the uptake of substrates even though the transport activities of individual channels are low, given the fact that CarO is the second most abundant OMP in *Acinetobacter baumannii*[108]. The relatively low free energy barriers associated with the transport of ornithine provide an explanation for the observed CarO dependence of the pathogen's growth on ornithine[105,106]. Our analysis suggests that CarO family proteins allow for slow and selective permeation of small substrates, with a strong preference for positively charged molecules.

CarO3Δ	glycine	ornithine	glutamic acid
ARG10	12.50	6.75	10.20
THR16	25.00	18.44	13.10
TYR43	0.00	19.62	9.45
ASN45	1.75	3.56	14.25
GLU47	5.31	35.44	6.30
LYS109	18.69	1.44	16.25

CarO2	glycine	ornithine	glutamic acid
ARG9	1.04	0.00	18.50
GLN24	0.00	0.00	11.67
ASN65	7.08	2.40	19.00
GLU67	5.42	5.60	18.42
ARG69	0.00	0.00	14.42
GLU154	15.08	14.60	0.00
TYR158	11.38	1.25	0.00
LYS206	8.58	21.10	0.00
TYR212	0.00	0.00	25.75

OprF_N	glycine	ornithine
GLN1	1.38	14.75
GLU7	20.21	11.17
GLU9	2.75	11.04
LYS13	17.67	16.62
GLU42	14.54	14.12
TYR122	12.62	0.00
LYS130	13.46	6.33
ASP134	7.26	10.71

Table 2: Comparison of the residues that are relevant for the translocation of glycine, ornithine and glutamic acid in CarO3Δ (top) , CarO2 (middle) and OprF_N (bottom). The score given to each residue is calculated as the percentage of time that an hydrogen bond was formed with the substrate, averaged over the minima of the specific passage. Multiple hydrogen bonds contribute separately. Only residues that form an hydrogen bond for more than 10% of the translocation time are shown

3.2 THE ROLE OF MUTATIONS AND THE IMPORTANCE OF THE DIPOLE MOMENT: *escherichia coli*'S OMPS

Details make perfection, and perfection is not a detail

— Leonardo da Vinci

In the previous section we have seen how the differences in charge and molecular weight of distinct substrates correspond to discrepancies in the rates at which the compound penetrate in the same protein; by creating an *in vitro* mutant of the protein under study, we were able to distinguish between the effects caused by the single domains of the channel. Finally, we showed that the variations in the free energy barriers encountered during metadynamics runs can be directly linked to the experimental results obtained though liposome swelling assays. Conversely, in this section we will compare the dynamics of the same substrate in proteins that differ by pinpoint mutations to extract a general rule for distinguishing between molecules that are more capable of translocating through OMPs.

Bacteria are able to change the permeability of the OM by modulating the expression of the influx porins that are used to get nutrients from the external environment. In *Escherichia coli*, OmpF and OmpC are among the most expressed porins of the OM, with more than 100 copies per cell[18]. They serve as general pathways for the influx of small molecules (e.g. molecular weight under 600 daltons) and present little charge selectivity. Both these water-filled channels present trimeric structures, with an hourglass shape. The constriction zone, located in the middle of the channels, is caused by the presence of a loop (L3) that folds internally, compressing the area accessible to the solvent (see Figure 7). With respect to OmpF, OmpC presents a smaller effective channel[37,109,110] and a lower conductance[27]; however, it has been already pointed out that the permeability of the two channels is similar, given the higher osmomolarity of the environments that increase the expression of OmpC over OmpF.

Often, small changes in the composition of the same protein are sufficient to drastically alter the rates at which a particular compound is able to penetrate within the pathogen: as an example, we consider the case of a patient with chronic *Eschericha coli* infection[?]. The bacteria extracted from the subject presented an higher expression of OmpC than expected, leading to increased resistance rates; moreover, over a two-year time period, a series of OmpC mutants were isolated[111?]. In this section, we compare the characteristics of

OmpF, OmpC and of the first and the last of these mutants, respectively named OmpC20 and OmpC33. In Figure 26, we report the primary sequence alignment of these 4 mutants.

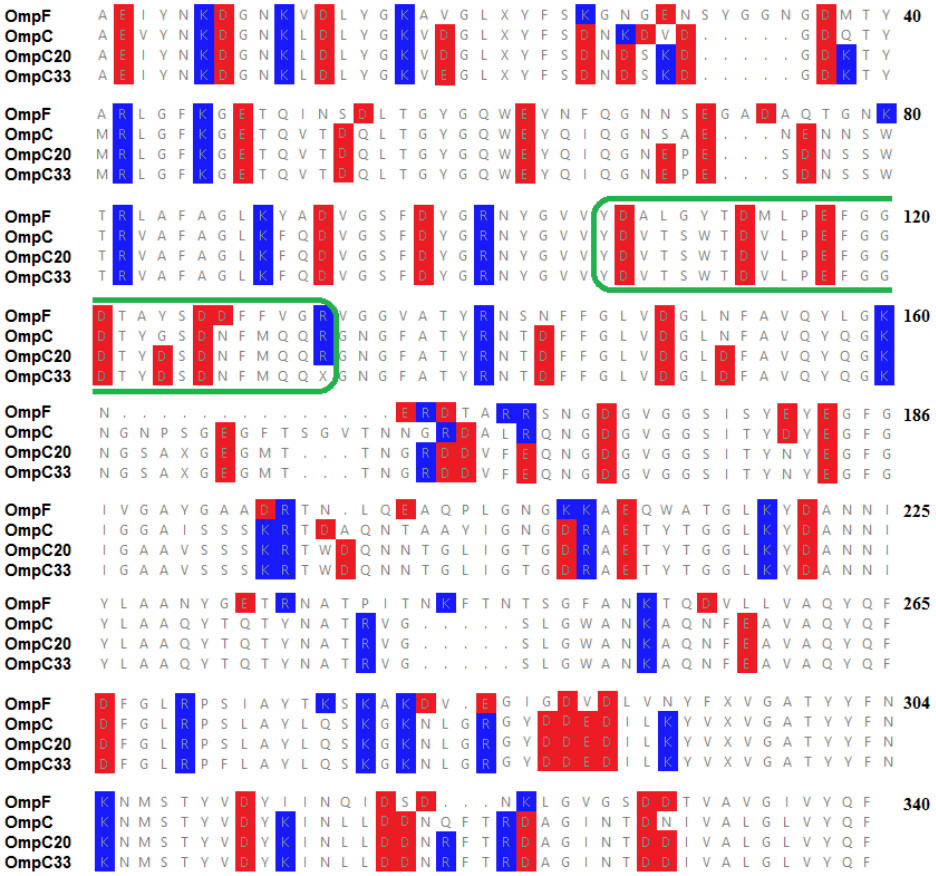

Figure 26: Primary sequence alignment between the proteins used in this section. The position of positively (blue) and negatively (red) charged residues, as well as the position of internal loop L3 (green), is highlighted

Interestingly, the structures of the OmpC mutants showed essentially an unchanged size[111]. This fact suggest that the increased antibiotic resistance in the clinical isolates is due to electrostatic, rather than steric hindrance, effects[112]. Both the negatively charged loop L3 and the basic ladder of positively charged residues facing it are highly conserved among the mutants: indeed, the constriction region is almost identical among the 4 proteins, while the majority of charged residues mutations and insertions are observed in the extracellular loops[112].

To explain how few mutated residues can drastically alter the permeability of the OM, we compared the translocation dynamics

of Imipenem, a zwitterionic β-lactam antibiotic with low molecular weight. We studied the dynamics of the antibiotic in function of the distance between the CMs of the antibiotic and the one of the protein (vertical axis, hereby denoted as Z_{CM}), and orientation, represented in the horizontal axis of the following figures. A positive value of the orientation corresponds to the positive charges of the antibiotic being nearer to the periplasmic side, analogously to Figure 15B.

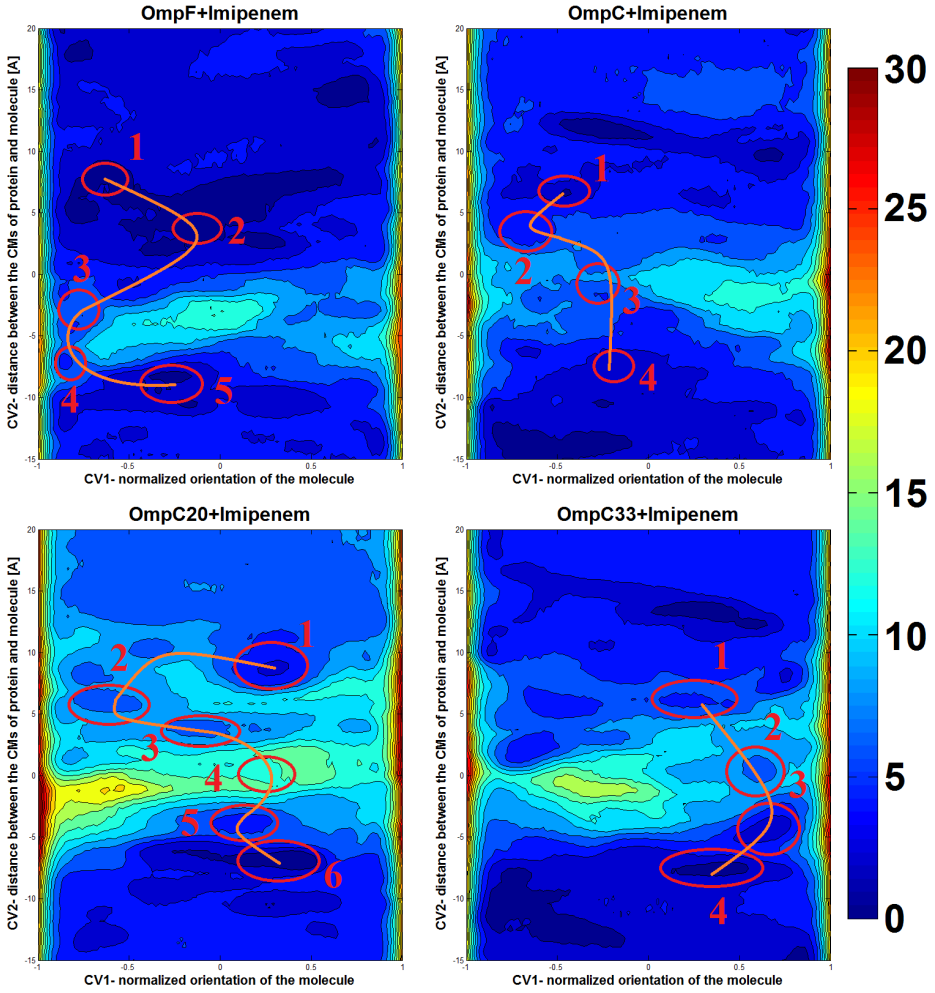

Figure 27: Translocation dynamics for Imipenem in OmpF, OmpC, OmpC20, OmpC33 in function of the selected CVs. The most stable conformation for different values of Z_{CM}, given by the lowest value of the free energy for a given depth, is reported. Each line corresponds to a 2 Kcal/mol free energy jump

By comparing the FESs for the 4 proteins, as shown in figure 27, we see that the most probable conformations at various values of Z_{CM} correspond to diverging values of the orientation assumed by

the antibiotic; the free energy jumps, however, are comparable between the proteins: indeed, no change in Imipenem sensitivity was observed in *in vivo* viability assays (see Figure S3 in[111]). Measurement of the minimum inhibitory concentrations for various antibiotics, on the other hand, revealed that the isolates had progressively greater antibiotic resistance[111].

To interpret the increased resistance associated with the mutated proteins, we can perform a statistical analysis on the different conformations that contribute to create the FESs shown in Figure 27: in a recent work from our group, it has already been suggested to look at the electric dipole moment of molecules, rather than to their mere zwitterionic character, to understand the ability of antibitiocs to successfully penetrate within the OM[112]. In Figure 28 we show the comparison between the projection of the average dipole moment of the antibiotic in the plane of the membrane for different values of Z_{CM}.

The 4 proteins show sensibly different characteristics for the orientation assumed by the dipole moment of Imipenem at various depths. Since Imipenem is a rather flexible molecule with a low molecular weight, these orientations can be seen as a direct effect of the varying direction adopted by the electric field of the proteins.

In OmpF and OmpC, the orientation of the dipole moment of Imipenem at the mouth of the channel, hereby denoted as preorientation region, is roughly preserved throughout the constriction zone. The preorientation region extends for values of Z_{CM} in between +10 and +0 Å. Between OmpF and OmpC, the only difference is represented by the extent of the orientation, that is a consequence of the varying strength of the electric field in the two proteins[27].

In OmpC20, the direction assumed by the dipole moment of Imipenem at the mouth of the channel is opposite with respect to the one in the constriction zone; this misalingment reduces to approximately 90 degrees at halfway of the preorientation region ($Z_{CM} = +5$). In OmpC20, therefore, the antibiotic is forced to reorient, thus reducing the probability of successfull translocation. In OmpC33, the adverse preorientation is retained even deeper inside the channel and the direction of the dipole is forced to change when closer to the constriction region. The lack of accessible area corresponds to an even greater increase in the free energy of the antibiotic during the translocation, that is, a lower probability of permeation.

The orientation of the mean dipole moment adopted by Imipenem in a specific region is a measure of the direction of the local electrostatic interactions inside the channel; the magnitude of this dipole moment, roughly represented by the length of the arrows

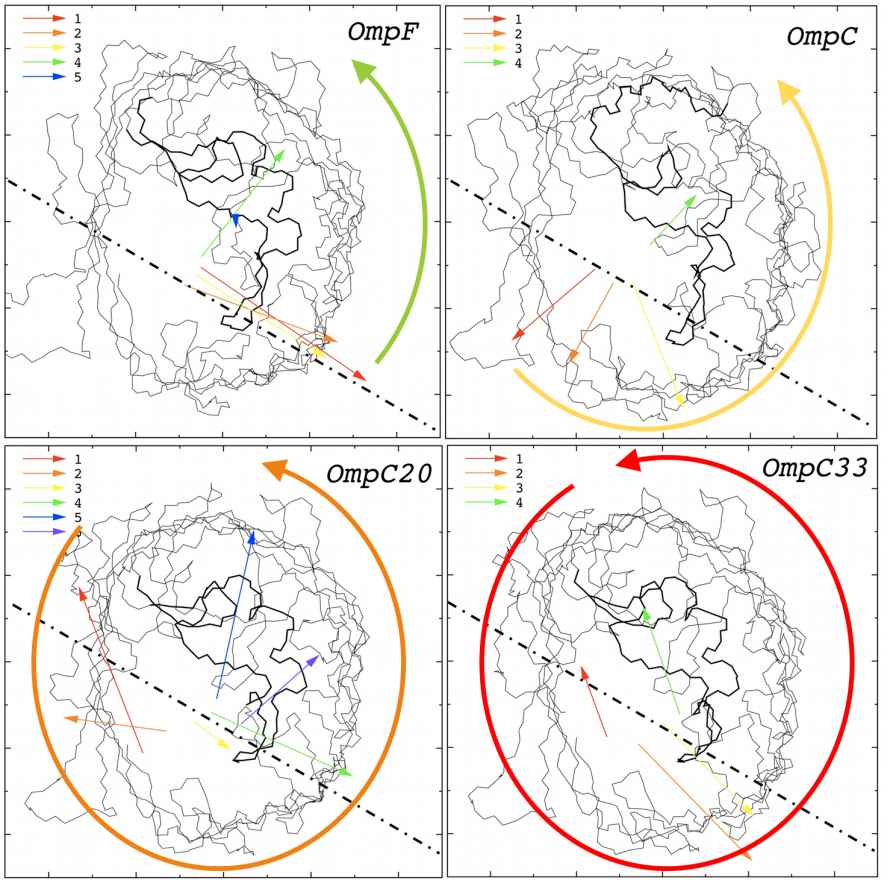

Figure 28: Projection of the dipole moment of Imipenem in the plane of the membrane for the conformations depicted Figure 27. The position of the protein is shown with a black line; the position of internal loop L3 is shown with a bold line. The circular lines correspond to the maximum extent of the reorientation experienced by the dipole moment of the antibiotic during the passage. A dashed line is drawn in correspondence of the best orientation inside the constriction region

in Figure 28, can be conversely seen as an index of the strength of the electric field itself: indeed, a lower mean value of the dipole assumed by the molecule corresponds to a lower ordering of the various conformations obtained at a specific depth. This effect is easily seen by looking at the mean orientation of water molecules at different values of Z_{CM} for the 4 proteins (Figure 29). We set the orientation of waters in the constriction region, that is the most conserved between the mutants, as 0 degrees.

In TIP3P model used in our simulations[58], waters are treated as rigid molecules, that therefore orient their static electric dipole accordingly to the direction and strength of the electrostatic interactions. Along the minimum energy path for translocation, the an-

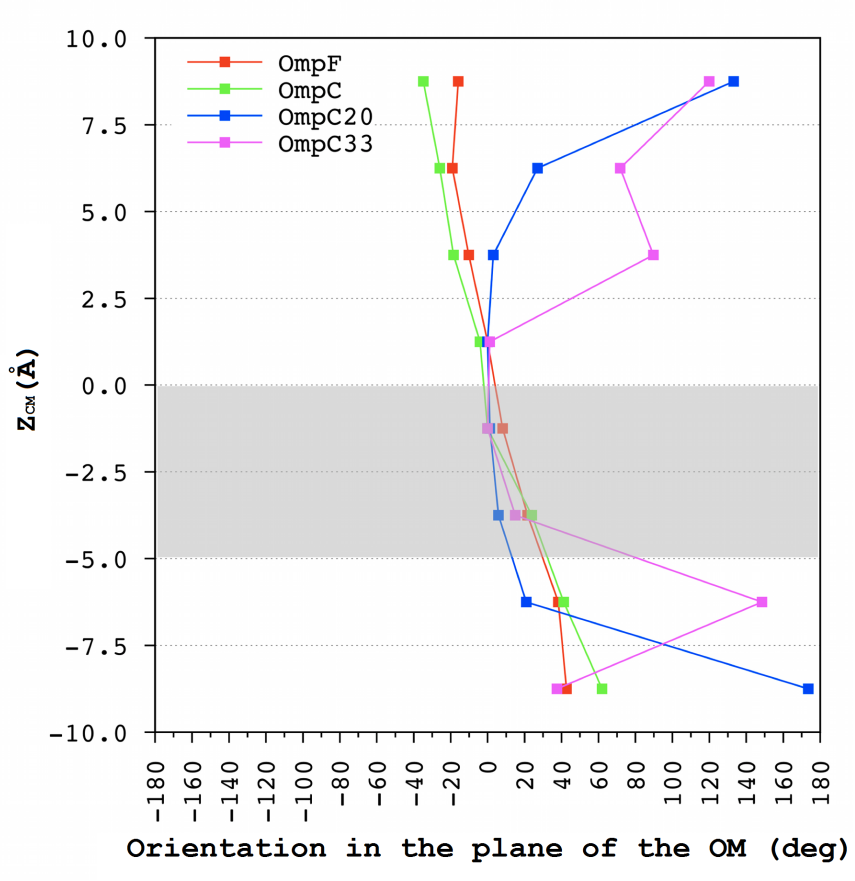

Figure 29: Orientation assumed by the dipole moment of water molecules, in the plane of the membrane, for different values of Z_{CM} and for the 4 proteins used in this study. The orientation in the constriction zone is set as 0 degrees. The constriction zone, located in the range [-5:0] of Z_{CM}, is shown in grey

tibiotic aligns its dipole to that of the water molecules at the same level inside the channel, thus following the *choreography* designed by electrostatics.

The differences obtained between the orientation of water molecues at different depths are neither but a consequence of the mutations of the single isolates, that show sensible modifications in the preorientation region. The emerging picture is that specific mutations have been introduced to tune the transversal electric field right at the entrance of the costriction region, while leaving the constriction zone itself almost unaffected[112].

The results shown in this section allow to affirm that the conformations assumed by a last resort antibiotic follow the dynamics of water molecules within the channel. A general mechanism for

translocation can be put forward: the drug is continuously forced to reorient along its path through the pore, in order to align its electric dipole to the channel transversal electric field. The azimuthal reorientations of the electric dipole appears fundamental in determining the translocation path and the overall free energy difference. With respect to water molecules, however, Imipenem presents a much higher molecular weight that can partially hinder the reorientation promoted by the electric field; this consideration allows to understand why the changes in the viability assays observed in the mutated strains are more promiment with respect to Gentamicin, an antibiotic with a greater molecular weight, than what observed for Imipenem[111].

The balance between the reduction in the free energy caused by successful alignment of the electric dipole[112,113] can be counterbalanced by the increase due to the reduction of the accessible area, that in turn causes a decrease of the conformational entropy of the substrate[112]. The subtle balance between these two terms is a promising parameter for determining the probability of successful translocation of substrates in the OM of GN bacteria; moreover, this mechanism allows to rationally interpret the effect that single-point mutations in both *in vivo* and *in vitro* bacterial strains will have with respect to antibiotic resistance.

3.3 MERGING THE GAP BETWEEN SIMULATIONS AND EXPERIMENTS: A KINETIC MODEL FOR MOLECULAR DIFFUSION THROUGH PORES

εντελεχεια[2]

— Aristotle

In the previous sections we have seen how the differences in the chemical structure of either the substrates (Section 3.1) or the proteins (Section 3.2) can be used to characterize the different dynamics that the substrates undergo during translocation in the OM of GN bacteria. The use of metadynamics has allowed us to qualitatively characterize the probability of diffusion; one of the major drawbacks of the method is that, however, dynamical quantities such as the average residence time in a specific conformation cannot be properly calculated by our accelerated simulations.

Recently, Tiwary and Parrinello have extended the scope of metadynamics with an *a posteriori* analysis that allows to calculate the time that an unbiased simulation would require to get the same results of a metadynamics run (see[96], Section 2.3). The technique has been applied to the study of unbinding kinetics of trypsine enzyme[97], but the application to the case of substrates' diffusion is affected by additional complications. The presence of multiple metastable states with similar affinities (that is, multiple minima with similar free energy depth in the FES) often corresponds to numerous available paths for the translocation in function of the selected CVs. To find the relative probability of the single states, and therefore to create a *kinetic scheme of the process*, we ran multiple simulations from each metastable state; furthermore, we applied the method reported in section 2.3[96,97] and obtained the real (Equation 16) characteristic time of the process.

The passage of Meropenem in *Escherichia coli*'s OmpF is particularly indicated for the analysis described: indeed, in a very recent article[42] the blockage time associated with the translocation of the antibiotic has been resolved and was found to be equal to $1.5 \pm 0.4\ \mu s$. This result was obtained with a Markovian analysis of the ion current interruption in a single-channel electrophysiology experiment (see page 13). Such a low residence time is barely in the range accessible through normal MD simulations (Sections 1.2.2,2.1): therefore, we don't need to bias excessively the evolution of the system and

2 entelecheia is a term used to identify the reality that has reached its full potential; it's usually counterposed to *ενεργεια* (energeia, the potentiality), but the meaning of the two words tends to converge.

we are in the conditions of quasistationarity required for the use of the analysis[96] (see page 36).

As a starting point, we reconstructed the free energy with respect to 2 CVs using a 2 μs long Metadynamics run in order to identify the most probable conformations. After 1.5 μs, we didn't observe any substantial change in the free energy surface and in the location of most stable states. We chose to bias the evolution of the system according to the projection distance between the CMs of Meropenem and OmpF along the axis of diffusion (vertical axis in the FESs, from now on denoted as Z_{cm}) and on the orientation of Meropenem (horizontal axis), as done in Section 3.2. The main FES was obtained with a single walker till the first translocation was accomplished, followed by the simultaneous exploration of 4 walkers[71]. Following the well-tempered scheme[74] (Section 2.2.2), we added a gaussian potential every 2 picoseconds with initial height of 1.2 Kcal/mol and modulated by a secondary temperature of 5000 K. We report the main FES in Figure 30A.

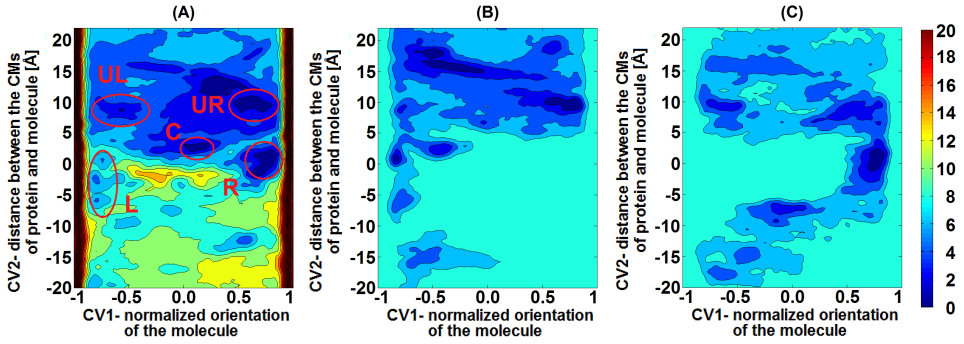

Figure 30: Free Energy surfaces of the system as a function of the selected collective variables. Each line represents a free energy difference of 2 Kcal/mol. (A) 2 μs FES. The main minima are highlighted; (B) representative FES of the left path; (C) representative FES of the right path. Simulations relative to FESs B and C have been stopped immediately after observing the first translocation

The region located in the range of Z_{cm} $[-5 : 2]$ Å shows on average high free energy: we see a central prohibited region and two minima, labeled L (left) and R (right), corresponding to Meropenem translocating with its principal axis parallel to the direction of diffusion, respectively with the negative and positive group pointing down. This is the constriction region, where steric hindrance plays an important role in the permeation, not having the Meropenem enough space to reorient. On the other hand, the presence of three minima, one just above the constriction region, C, and the other two

well above, labeled UL and UR, shows that above the constriction region Meropenem can assume any orientation. With a depth of 8 Kcal/mol, minimum R appears more stable than the opposite orientation represented by minimum L (4 Kcal/mol).

The lack of accessible area in the constriction region makes the two paths represented by the use of minimum L (from now on referred to as "left" path) and the one relative to minimum R ("right" path) as alternative for the diffusion through the constriction region. A first attempt to determine the relative probability of the two different paths was done by running 20 single-shot Metadynamics simulations, stopped immediately after the first passage through the constriction region (Figure 31). The two paths, in fact, appear to be alternative and we were able to divide the single-shot simulations in two groups relative to the passage through either the left or the right one. In this test, the right path appeared as to be slightly more probable than the left one, with a relative probability of 60% (12 simulations over 20) against the 40% (8/20) following a left path. We report a representative FES for the two groups in Figures 30B,30C.

To reduce the probability that the obtained probabilities for the two paths might be affected by the use of the external potential added through metadynamics, in single-shot simulations we lowered the initial height of the hills to 0.1 Kcal/mol with a secondary temperature of 3000 K and the inverse deposition rate to 5 picoseconds. From now on, we will refer to metadynamics run under these conditions as *low-biased* simulations.

In single channel electrophysiology experiments the passage of the compounds is associated with the reduction of the ion current: a comparison with the experimental current blockage requires therefore an analysis of the occlusion of the channel. We have analysed and compared the value of the minimal accessible area with and without the antibiotic for the three minima at or near the constriction region, respectively C, L, R (Table 3). In absence of the molecule, we obtain a minimal accessible radius of 3.1 ± 0.2 Å, a value compatible with previous findings[110]. Only in R the occlusion is almost 60%, probably causing a blockage to the passage of ions[26]. In any case, both minima L and R correspond to a vast reduction of the minimal accessible radius. The presence of a deep minimum located in the constriction region was observed for all the single-shot simulations depicted in Figure 31: this feature suggest that the translocation of Meropenem in OmpF corresponds effectively to a reduction of the ion current through the channel.

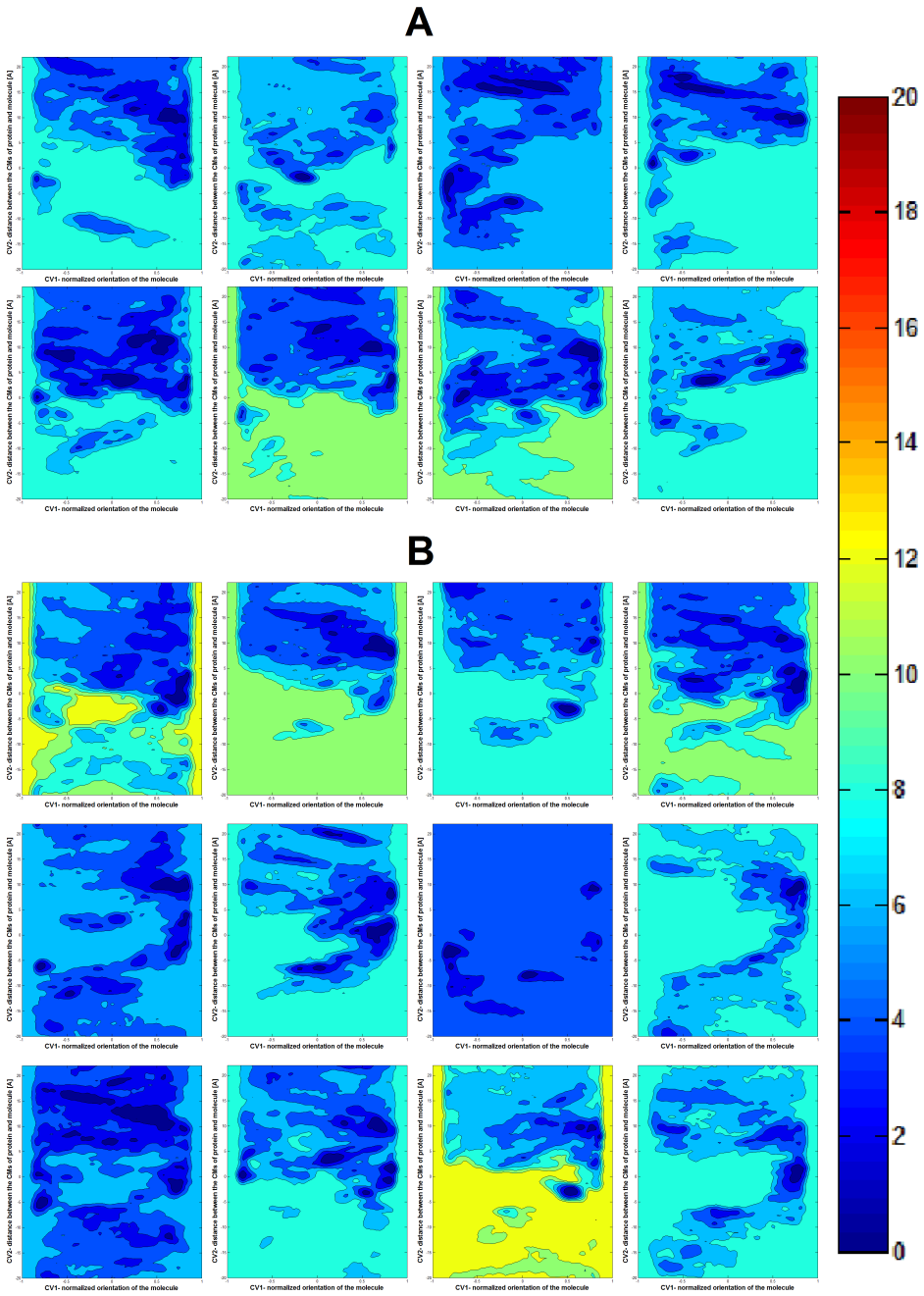

Figure 31: Single free energy surfaces relative to the passage of Meropenem in OmpF as a function of the projection of the distance between the center of masses of the antibiotic and the one of the protein (y axis) and of the projection on the direction of the translocation of the whole orientation of the antibiotic (x axis). (A) simulations in which a "left" path, corresponding to the antibiotic translocating at first with the negative part, was observed. (B) simulations in which a "right" path, complemental to the one observed in (A), was used by Meropenem

Structure	Minimum Radius (Å)	Location (Å)	Reduction of accessible area (%)
No Meropenem	3.1 ± 0.2	-2.5 ± 0.8	0
Minimum C	2.3 ± 0.3	0 ± 1	26
Minimum L	1.9 ± 0.3	-3 ± 1	39
Minimum R	1.3 ± 0.3	-3 ± 1	58

Table 3: Value and location of the minimal accessible radius of OmpF monomer in absence and presence of Meropenem in various minima of Figure 30A. The value reported as location is the distance with respect to the center of mass of the protein.

We have already pointed out in section 3.2 that the residues in the constriction zone of OMPs tend to be very well conserved among different proteins, as they seem to be fundamental for differentiating between different compounds; as a matter of fact, the mutation of these residues can be linked to variation in the susceptibility of OmpF to antibiotics. To assess quantitatively the role of the single residues towards the screening of Meropenem, we have analysed the frames of the main FES relative to minima C, L, R and we have selected the residues that formed an H-bond for more than 10% of the simulation time, as done in Section 3.1. We report these *key* residues, as well as the position of the antibiotic in the single minima, in Figure 32.

It's interesting to note that minima L (Figure 32A) and R (Figure 32B), corresponding to opposite orientations of the substrate, are using complementary parts of the ladder of positively-charged residues facing loop L3 (Figure 7C) for the passage. These results confirm the importance of electrostatic interactions in determining the permeability of the channel for zwitterionic and charged compounds[112].

We ran 40 independent, low-biased (page 60) simulations starting from each minimum identified along the FES shown in Figure 30A, as well as starting from the extracellular/intracellular mouth regions. With these simulations we evaluated the rates of transition among the various states and created a scheme of the process (Figure 33). The use of the low and infrequent bias allowed to conclude that the saddle zones of the single transitions are not considerably biased by the hills depositions: therefore, we conclude that the rates reported are the same we would have obtained without the addition of the external potential.

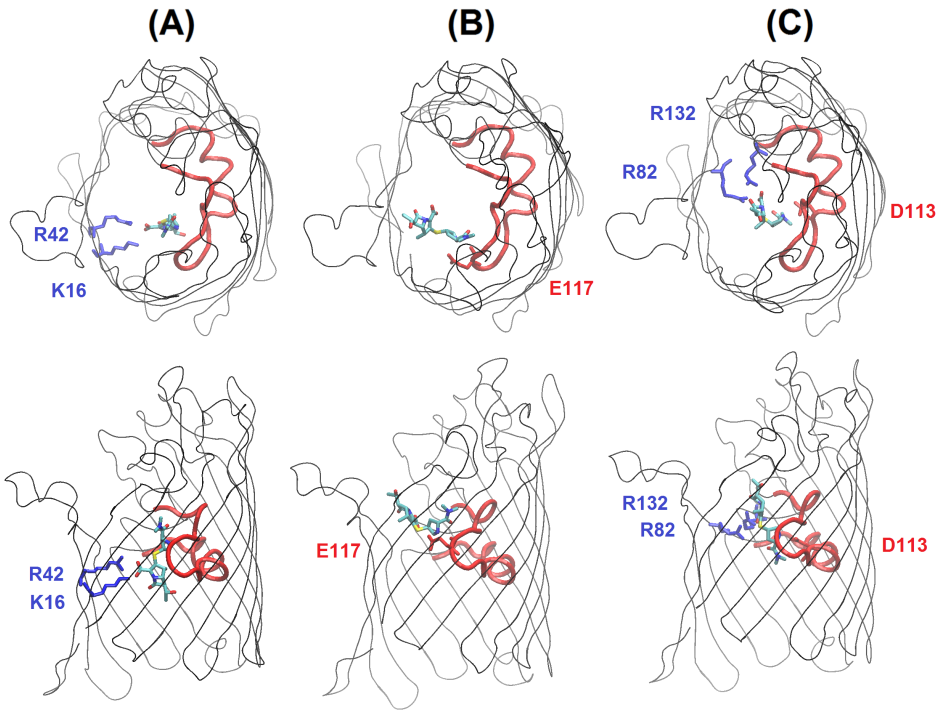

Figure 32: Top (top) and front (bottom) view of OmpF in complex with Meropenem in minima L (A), C (B) and R (C) of Figure 30A. The position of the antibiotic (with nitrogens colored in blue and oxygens colored in red), as well as the most stable interactions, are shown. In all the figures, the red tube highlights the position of internal loop L3.

It's particularly interesting to note that the probability of transition through minimum L, when starting from the extracellular region, is null. Indeed, the whole translocation process from the extracellular region (EX) to the periplasmic side (PE) has a 100% probability of using minimum R, that is equivalent to say that Meropenem always translocates with the positive charges followed by the negative ones. Because of the low number of simulations (40) started from each minimum, we can't exclude that there could be a low rate of translocations involving minimum L; our results, however, suggest that at least the vast majority of the translocations involve the antibiotic to enter the constriction zone with the positive charged group before. This consideration is also compatible with the rates observed for the transitions PE→L and escape rates from minimum L itself: this minimum appears to be the main pathway for the translocation from PE to EX, i.e., for efflux of Meropenem through OmpF.

A possible explanation for the apparent disagreement between the rates obtained with the simulations starting from the single

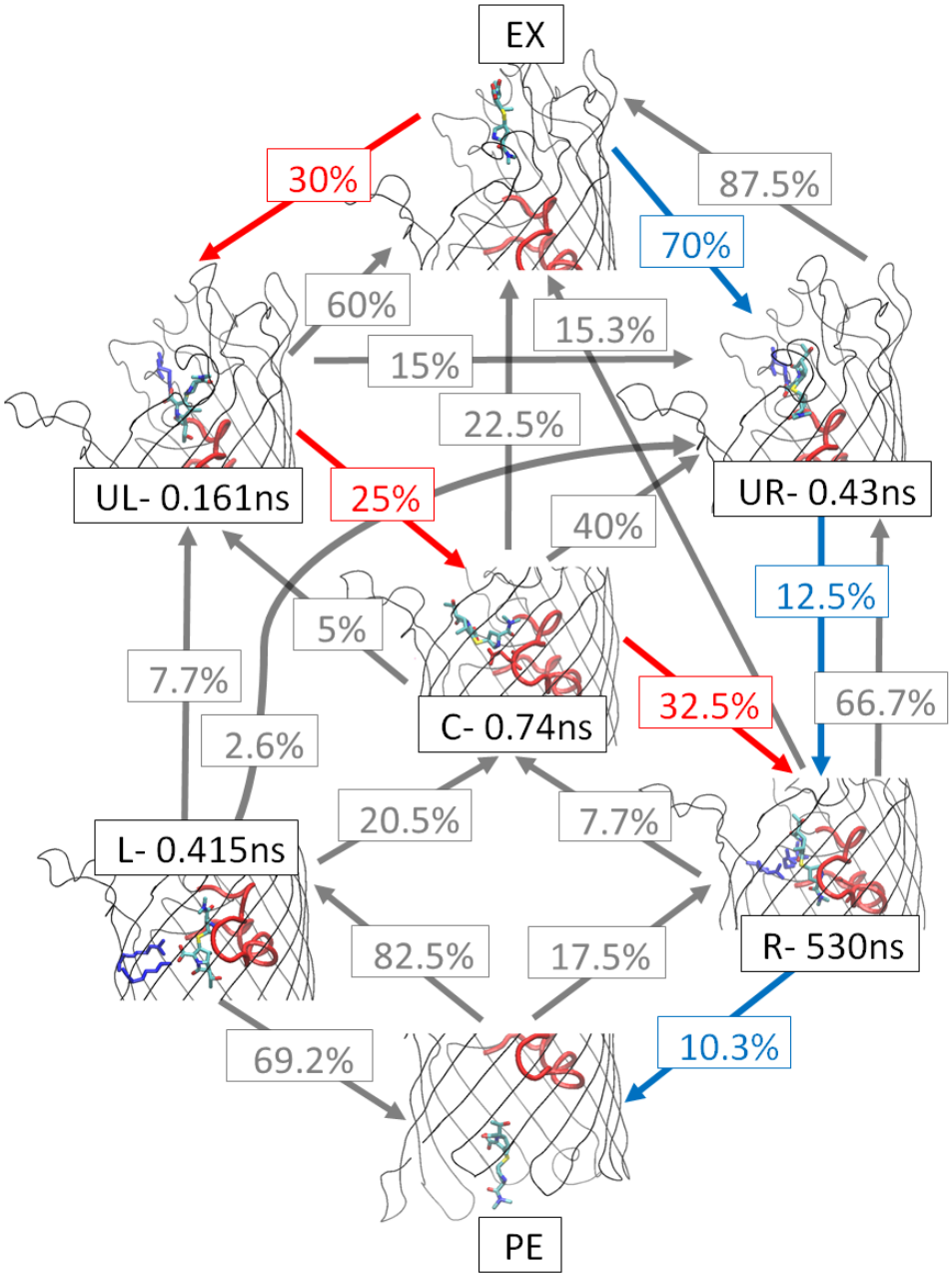

Figure 33: Kinetics of the translocation of the substrate. Starting from minima UL, UR, C, L, R, as well as from the external side of the membrane (EX) and the periplasmic space (PE), the transition rates have been calculated. For each minima, average residence time is reported. (blue) most probable path of translocation, EX→UR→R→PE, with probability 0.90%. (red) second most probable path of translocation, EX→UL→C→R→PE, with probability 0.25%.

minima (Figure 33) and the paths observed in one-shot simulations (Figure 31) is that, in the latter, the history-dependant potential was

maintained until the first translocation to the periplasmic side was observed. To avoid the escape of the molecule in the bulk water, we were forced to add two artificial potential walls placed at Z_{cm}=+25 Å and Z_{cm}=-20 Å, respectively; this can alter the exploration of the phase space when alternative paths are present: upon a first attempt to pass through the constriction region using the most probable path, the bias already added in conjunction with the artificial wall push the system toward the other path instead of exiting. This effect is absent in the simulations starting from the single minima of the main FES and therefore doesn't affect the rates involving escape from minima UL, UR, C, L, R. Simulations starting from either EX or PE, on the other hand, are influenced by these walls. The net result is that the probability of escaping from EX or PE to bulk water is neglected in the rates reported. The two most relevant paths obtained with this assumption contribute for a total probability of translocation of 1.15%, which can be set as an upper limit. The probability of successful translocation is, therefore, far lower than 1%.

The use of the method of Tiwary and Parrinello[96] allowed us to calculate the acceleration factor caused by the deposition of the external potential. We were consequently able to calculate the real residence time (Equation 16) for each minimum, that is the time required to escape from the specific conformation. By merging these times with the rates obtained through the multiple simulations, we built a kinetic scheme of the translocation process[96,97], from the extracellular region to the periplasmic space. In Table 4 we show the results of KS test run on the exit times from each of the minima of Figure 30A. The characteristic times have been obtained with an exponential fit with a Poisson distribution of the single transitions[96,98], as explained in page 37. The fits are reported in Figure 34.

Kolgomorov-Smirnov test allowed us to check the robustness of our results. While from L and R we were obliged to discard a single simulation (over 40) to pass the test with a minimum allowed p-value of 0.05, in the 3 upper minima the test was fine without any deletion (Table 4). These results confirm that our simulations can be treated as statistical ensembles, that is the low-biased conditions we chose for the single-shot simulations kept the kinetics of the escape process unaltered. Our results suggest, additionally, that the selected CVs are valid to explore the conformational space of the system under study. To further improve the results of KS test in the constriction region, we suggest the additional use of a coordinate

Starting conformation	p-value	Characteristic time (ns)
UL	0.42	0.161 ± 0.004
UR	0.65	0.43 ± 0.01
C	0.59	0.74 ± 0.02
L	0.031	0.44 ± 0.01
L*	0.25	0.42 ± 0.01
R	0.045	550 ± 20
R*	0.31	530 ± 10
Single-channel experiment[42]		1500 ± 400

Table 4: Results of Kolgomorov-Smirnov test on the times obtained for escaping from the minima of Figure 30A. Minima marked with an asterisk required to eliminate the biggest time from the distribution to pass the test with a p-value greater than 0.05.

related to the local interactions. The need to restrict to the use of 2 CVs only didn't allow us, however, to test this hypothesis.

Our results indicate quite different properties for these two minima. The residence time is 3 orders of magnitude larger for minima R (submicrosecond) compared to L (subnanoseconds), and the occlusion is larger for R (58%) compared to L (39%). This can be seen as a consequence of the presence of stable hydrogen bonds between the antibiotic and both the negative internal loop L3 (D113) and the ladder of positively-charged residues facing it (R82, R132). Indeed, several experimental works[44,114–116] link mutations on these residues to an increase in bacterial susceptibility to several lactam compounds. This feature supports the hypothesis that only a low number of key residues of the porin determine the rates of transitions for different molecules[117]. The value of the residence time to escape from R (submicroseconds) together with the high occlusion of the pore are compatible with the analysis of current noise in a single-channel experiment of OmpF with Meropenem recently published[42].While electrophysiology cannot distinguish between blockages of the channel, i.e. binding and rejection, from translocations[38,39], the blockage time of 1.5 μs seems to indicate that what observed experimentally is indeed relative to the real translocation of Meropenem. Our results suggest that the translocation process through the whole channel can be seen as a multistep process among quite different regions: the extracellular is characterized by

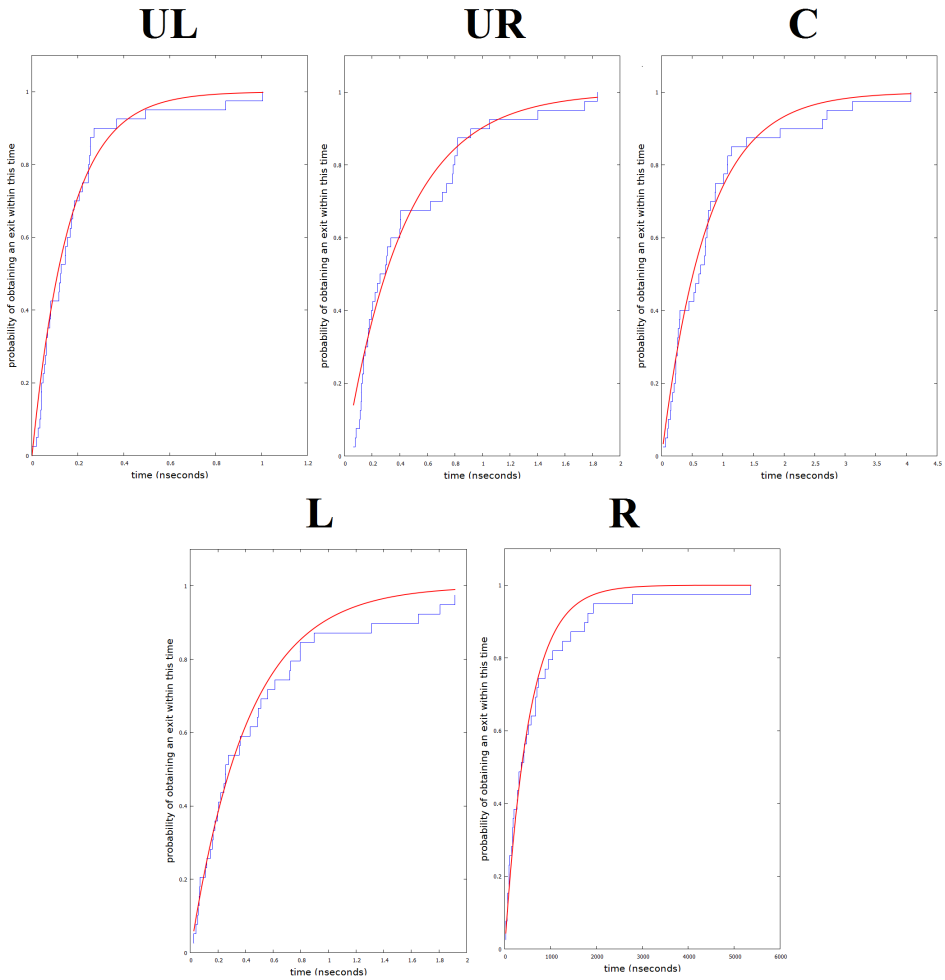

Figure 34: Results of the Kolgomorov-Smirnov test on the exit times from the minima of Figure 30A. Blue: probability of obtaining an escape from the selected minimum within the time t. Red: empirical cumulative distribution function fit. Fitted parameters are reported in Table 4

long loops; the periplasmic is quite large; the constriction region is characterized by strong electrostatic interactions. With the method presented, we where able to quantitatively assess the importance of the single interactions between the molecules and the pore; additionally, we characterized the relative importance of different conditions to the diffusion of Meropenem in OmpF.

4

CONCLUSIONS

Science was proving itself incapable to comprehend the causes of epidemies. And all the world, even in the most remote regions, seemed victim of that spell with no explanation.
- Almost all the world -, said Baldabiou slowly. - Almost -, spilling two drops of water in his Pernod.

— Alessandro Baricco

In this work, we have studied the dynamics that different molecules undergo during translocation in the outer membrane of gram-negative bacteria. We have shown how accelerated molecular dynamics (Sections 2.2.1,2.2.2) allow to properly distinguish between the most probable conformations that characterize the diffusion of compounds, supporting the results of experimental techniques (Section 3.1) and allowing to overcome the intrinsic limitations that affect presently available *in vivo* approaches (Section 3.3).

In the first chapter, we have introduced the problem represented by antibiotic resistance. We have seen how the advent of multiple drug resistant strains is hindering our ability to contrast the infections caused by pathogens (Section 1.1). Bacteria can reduce the efficacy of antimicrobials by tuning the permeability of their outer membrane, a wide layer composed of lipopolysacharides and phospholipids that constitutes the external cell wall. This barrier is filled with influx proteins that permit the passive diffusion of nutrients from the external membrane. By reducing the expression of a particular channel or by modifying its chemical structure with single-point mutations, pathogens are able to inhibit the passage of noxious compounds. A clear understanding of the mechanism that determines molecular translocation in pores is fundamental to design the next generation of antibiotics and to overcome the problem represented by antimicrobial resistance.

To study the permeability of the outer membrane, both *in vitro* and *in silico* methods have been used extensively: the limitations of both approaches have been elucidated in Section 1.2. Currently available experimental methods lack the precision and resolution necessary to the study of a process that has a time scale of the order of tents of microseconds. Moreover, presently used techniques aren't

able to distinguish between signals due to binding and rejection of the substrates from the ones caused by effective translocation. With standard molecular dynamics, conversely, the process can be studied in great detail and the consequences of single mutations can be evaluated; however, the computational resources required to simulate a single translocation don't allow to obtain a statistics of the process.

In the second chapter, we have reviewed the computational methods that have been used in this work. We firstly presented the main characteristics of standard molecular dynamics simulations (Section 2.1). By numerically integrate the equations of motion, the evolution of a system can be calculated. The proteins that we wish to simulate have to be inserted in a lipid bilayer and successively equilibrated, analogously to what is done in *in vitro* measurements of single-channel experiments; the process of translocation, however, is characterized by a free energy jump that drastically reduces the probability of observing an event and imposes the use of enhanced sampling techniques (Section 2.2). These methods rely on altering the potential experienced by the system to increase the speed at which a particular event is realized. To elucidate the role of entropic effects and to deduce the evolution that an unbiased system would have experienced, a kinetic model has to be introduced (Section 2.3).

We applied the techniques described in the methods section to the study of some test-cases. In Section 3.1, we compared the dynamics of substrates with different net charge and molecular weight during translocation in CarO, an 8-stranded β-barrel specific channel with low permeability that is abundant in *Acinetobacter baumannii*. The small or null accessible area increases the importance that pairwise interactions have in determining the diffusion rates of different compounds. The presence of a striking stucture overhanging the extracellular entrance to the porin and resembling a baseball glove suggests to treat the channel as a two-domain protein. We were able to distinguish the effect of the single domains by creating a loop-less mutant and comparing the diffusion rates of polar compounds in the two systems. We understood that the charge selectivity of CarO proteins is an effect of the net negative charge of the external glove, that improves the diffusion of positively charged ornithine by increasing the movement of the residues internal to the protein; conversely, the loop-less mutant showed higher permeabilty towards zwitterionic glycine and negatively charged glutamic acid. We envised a network disturbance model to describe the pas-

sage of compounds in the channel, where interactions between the substrates and the internal residues of the protein disturb the internal hidrogen bonds and salt bridges, allowing diffusion. Our findings allowed to interpret experimental rates obtained by liposome swelling assays experiments and illuminated the roles played by abundant small proteins within low-permeability outer membranes of certain Gram-negative bacteria.

Bacteria are able to change the permeability of the outer membrane by selectively changing some key residues of their influx channels: therefore, rationalizing the mutations observed in proteins is a key to enhance the efficacy of antibiotics. We have studied the differences in the translocation of Imipenem in the 2 major channels of *Escherichia coli* and in 2 mutated strains observed in a patient with long term infection over a two-year period (Section 3.2). By comparing the strength and the orientation of the mean transversal electric dipole assumed by the antibiotic at different stages of the translocation, we found that the mutations alter the preorientation assumed by Imipenem before reaching the bottleneck of the diffusion, represented by the region with the lowest accessible area. Successive mutations correspond to reducing the strength of the preorientation obtained by the antimicrobial, or even imposing a wrong orientation in the lumen of the channel. This effect is easily elucidated by studying the choreography of water molecules inside the channel. The internal electric field of the protein is altered to impose rotations that increase the free energy barrier encountered during the passage: the energy cost is proportional to the loss of conformational entropy caused by the absence of accessible area within the channel. To reduce the effect of mutations in the permeation of antimicrobial, the angle between the molecule's main axis of inertia and its electric dipole has to be properly adjusted.

The enthalpic contribute caused by a positive interaction with the internal electric field can be counterbalanced by entropic interactions due to excluded volume effects. To further assess the role of entropy, we combined enhanced sampling methods with statistical ensembles to build a kinetic model of molecular diffusion through bacterial pores (Section 3.3). We tested the validity of the approach investigating the translocation of Meropenem through OmpF, finding good agreement with the analysis of single channel ion current noise in electrophysiology experiments. The use of Metadynamics showed the presence of two minima in the constriction region, associated with opposite orientations of the antibiotic; with our kinetic model we discriminated between them, showing that only the path associated with Meropenem entering with the positive group first

results in actual translocation. The upper limit probability obtained for the translocation process suggests that only a low percentage of the antibiotics approaching the channel are effectively able to translocate. This entropic barrier seems compensated by formation of simultaneous hydrogen bonds with the negatively charged internal loop L3 and with the wall of positively charged residues facing it in the centre of the protein. The approach proposed in this section allows to obtain a clear and precise view of the passage of substrates even in the case of multiple possible paths for the translocation process, overcoming the intrinsic limitations of currently available simulation and experimental methods.

PERSPECTIVES

The use of enhanced sampling allowed us to overcome the limitations in the time scales accessible through simulations, but the approaches presented in this work are still too slow to enable the use of *in silico* methods for high throughtput screening of potential candidates for the next generation of antibiotics.

Our group is currently focusing on the study of coarse grained methods, in an attempt to create a model able to predict with minimal computational effort the rates of permeation of different compounds. This model has to be compared with the results of both fine grained simulations and in vitro experiments to obtain the optimal parameters that determine the overall efficacy of a molecule in permeation through pores.

The need to obtain results directly comparable with experimental data requires to reduce the biases imposed by the addition of an external potential, while preserving the increased speed in the exploration of the conformational space of the process under study. We are presently studying the performance and applicability of recently introduced schemes that permit to run multiple simulations in parallel and to retain only those that fulfil certain requirements, to reduce the exploration of the otherwise statistically more probable states that are irrelevant for the process of translocation.

Our studies advocate the use of rational schemes for the creation of the next generation of antibiotics to win the struggle against the increasing intrinsic resistance that many re-emergent pathogens are showing.

BIBLIOGRAPHY

[1] "Antibiotic resistance threats in the united states, 2013." Centers for Disease Control and Prevention (CDC), 2013. (Cited on page 4.)

[2] J. C. E. Pagès, Jean-Marie 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 3, 8, 9, 10, and 12.)

[3] J. Davies and D. Davies, "Origins and evolution of antibiotic resistance," *Microbiology and Molecular Biology Reviews*, vol. 74, no. 3, pp. 417–433, 2010. (Cited on pages 3 and 6.)

[4] P. Courvalin, "Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria.," *Antimicrobial agents and chemotherapy*, vol. 38, no. 7, p. 1447, 1994. (Cited on pages 4 and 6.)

[5] C. Bornet, A. Davin-Regli, C. Bosi, J.-M. Pagès, and C. Bollet, "Imipenem resistance of enterobacter aerogenes mediated by outer membrane permeability," *Journal of clinical microbiology*, vol. 38, no. 3, pp. 1048–1052, 2000. (Cited on page 4.)

[6] "Antimicrobial resistance surveillance in europe 2008. annual report of the european antimicrobial resistance surveillance network (ears-net)." (Cited on page 5.)

[7] "Antimicrobial resistance surveillance in europe 2013. annual report of the european antimicrobial resistance surveillance network (ears-net)." (Cited on page 5.)

[8] C. A. Arias and B. E. Murray, "Antibiotic-resistant bugs in the 21st century—a clinical super-challenge," *New England Journal of Medicine*, vol. 360, no. 5, pp. 439–443, 2009. (Cited on page 4.)

[9] J. W. Scannell, A. Blanckley, H. Boldon, and B. Warrington, "Diagnosing the decline in pharmaceutical r&d efficiency," *Nature reviews Drug discovery*, vol. 11, no. 3, pp. 191–200, 2012. (Cited on page 5.)

[10] M. A. Fischbach and C. T. Walsh, "Antibiotics for emerging pathogens," *Science*, vol. 325, no. 5944, pp. 1089–1093, 2009. (Cited on page 4.)

[11] M. Gunn, M. Lim, D. Cross, and M. Goldman, "Benchmarking the scientific output of the innovative medicines initiative," *Nature Biotechnology*, vol. 33, no. 8, pp. 811–812, 2015. (Cited on page 5.)

[12] A. Kumar, *Transport of antibiotics through bacterial porins: insights from atomistics simulations*. PhD thesis, Università degli studi di Cagliari, 2009. (Cited on page 6.)

[13] L. Lideman, "Factors limiting the gene expression of escherichia coli in the cells of bacilli," *Genetika*, vol. 19, no. 5, pp. 693–707, 1983. (Cited on page 6.)

[14] P. R. Murray, K. S. Rosenthal, and M. A. Pfaller, *Medical microbiology*. Mosby, 1994. (Cited on page 6.)

[15] 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 7, 8, and 9.)

[16] A. H. Delcour, "Outer membrane permeability and antibiotic resistance," *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, vol. 1794, no. 5, pp. 808–816, 2009. (Cited on pages 7 and 12.)

[17] M. Zahn, T. D'Agostino, E. Eren, A. Baslé, M. Ceccarelli, and B. van den Berg, "Small-molecule transport by caro, an abundant eight-stranded β-barrel outer membrane protein from acinetobacterbaumannii," *Journal of molecular biology*, 2015. (Cited on pages 7, 11, 14, 15, 22, 24, 41, 42, 44, and 47.)

[18] M. Masi and J.-M. Pagès, "Structure, function and regulation of outer membrane proteins involved in drug transport in enterobactericeae: the ompf/c–tolc case," *The open microbiology journal*, vol. 7, p. 22, 2013. (Cited on pages 7, 8, 9, 10, 11, 12, and 51.)

[19] A. Baslé, G. Rummel, P. Storici, J. P. Rosenbusch, and T. Schirmer, "Crystal structure of osmoporin ompc from e. coli at 2.0 å," *Journal of molecular biology*, vol. 362, no. 5, pp. 933–942, 2006. (Cited on pages 7, 9, 10, and 11.)

[20] C. Nasrallah, *Étude du rôle joué par les porines dans la persistance des infections par Providencia stuartii*. PhD thesis, Université de Grenoble, 2006. (Cited on page 7.)

[21] B. K. Ziervogel and B. Roux, "The binding of antibiotics in ompf porin," *Structure*, vol. 21, no. 1, pp. 76–87, 2013. (Cited on pages 7 and 9.)

[22] J.-M. Pagès, "Role of bacterial porins in antibiotic susceptibility of gram-negative bacteria," *Bacterial and Eukaryotic Porins*, pp. 41–59, 2004. (Cited on page 9.)

[23] W. Im and B. Roux, "Ion permeation and selectivity of ompf porin: a theoretical study based on molecular dynamics, brownian dynamics, and continuum electrodiffusion theory," *Journal of molecular biology*, vol. 322, no. 4, pp. 851–869, 2002. (Cited on page 10.)

[24] W. Im and B. Roux, "Ions and counterions in a biological channel: a molecular dynamics simulation of ompf porin from escherichia coli in an explicit membrane with 1m kcl aqueous salt solution," *Journal of molecular biology*, vol. 319, no. 5, pp. 1177–1197, 2002.

[25] S. Cowan, T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. Pauptit, J. Jansonius, and J. Rosenbusch, "Crystal structures explain functional properties of two e. coli porins," *Nature*, vol. 358, no. 6389, pp. 727–733, 1992. (Cited on page 10.)

[26] E. Hajjar, A. Bessonov, A. Molitor, A. Kumar, K. R. Mahendran, M. Winterhalter, J.-M. Pagès, P. Ruggerone, and M. Ceccarelli, "Toward screening for antibiotics with enhanced permeation properties through bacterial porins," *Biochemistry*, vol. 49, no. 32, pp. 6928–6935, 2010. (Cited on pages 10, 11, 14, 15, 34, 47, and 60.)

[27] Q.-T. Tran, K. R. Mahendran, E. Hajjar, M. Ceccarelli, A. Davin-Regli, M. Winterhalter, H. Weingart, and J.-M. Pagès, "Implication of porins in β-lactam resistance of providencia stuartii," *Journal of Biological Chemistry*, vol. 285, no. 42, pp. 32273–32281, 2010. (Cited on pages 10, 13, 51, and 54.)

[28] K. Gehring and H. Nikaido, "Existence and purification of porin heterotrimers of escherichia coli k12 ompc, ompf, and phoe proteins.," *Journal of Biological Chemistry*, vol. 264, no. 5, pp. 2810–2815, 1989. (Cited on page 10.)

[29] S. Begic and E. A. Worobec, "Regulation of serratia marcescens ompf and ompc porin genes in response to osmotic stress, salicylate, temperature and ph," *Microbiology*, vol. 152, no. 2, pp. 485–491, 2006. (Cited on page 10.)

[30] D. Walthers, A. Go, and L. J. Kenney, "Regulation of porin gene expression by the two-component regulatory system envz/ompr," *Bacterial and Eukaryotic Porins: Structure, Function, Mechanism*, pp. 1–24, 2004. (Cited on page 10.)

[31] X. Liu and T. Ferenci, "Regulation of porin-mediated outer membrane permeability by nutrient limitation in escherichia coli," *Journal of bacteriology*, vol. 180, no. 15, pp. 3917–3922, 1998. (Cited on page 11.)

[32] X. Liu and T. Ferenci, "An analysis of multifactorial influences on the transcriptional control of ompf and ompc porin expression under nutrient limitation," *Microbiology*, vol. 147, no. 11, pp. 2981–2989, 2001. (Cited on page 11.)

[33] A. S. Low, F. M. MacKenzie, I. M. Gould, and I. R. Booth, "Protected environments allow parallel evolution of a bacterial pathogen in a patient subjected to long-term antibiotic therapy," *Molecular microbiology*, vol. 42, no. 3, pp. 619–630, 2001. (Cited on page 11.)

[34] H. Nikaido, "Outer membrane barrier as a mechanism of antimicrobial resistance.," *Antimicrobial Agents and Chemotherapy*, vol. 33, no. 11, p. 1831, 1989. (Cited on page 11.)

[35] R. Misra and S. Benson, "Isolation and characterization of ompc porin mutants with altered pore properties.," *Journal of bacteriology*, vol. 170, no. 2, pp. 528–533, 1988. (Cited on page 11.)

[36] 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, pp. 63–67, 2015. (Cited on page 11.)

[37] H. Nikaido and E. Y. Rosenberg, "Porin channels in escherichia coli: studies with liposomes reconstituted from purified proteins.," *Journal of Bacteriology*, vol. 153, no. 1, pp. 241–252, 1983. (Cited on pages 11, 12, 44, and 51.)

[38] 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 11, 12, 15, and 66.)

[39] 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 66.)

[40] C. Danelon, E. M. Nestorovich, M. Winterhalter, M. Ceccarelli, and S. M. Bezrukov, "Interaction of zwitterionic penicillins with the ompf channel facilitates their translocation," *Biophysical journal*, vol. 90, no. 5, pp. 1617–1627, 2006. (Cited on page 13.)

[41] H. Bajaj, Q.-T. Tran, K. R. Mahendran, C. Nasrallah, J.-P. Colletier, A. Davin-Regli, J.-M. Bolla, J.-M. Pagès, and M. Winterhalter, "Antibiotic uptake through membrane channels: role of providencia stuartii ompst1 porin in carbapenem resistance," *Biochemistry*, vol. 51, no. 51, pp. 10244–10249, 2012. (Cited on pages 13 and 24.)

[42] I. Bodrenko, H. Bajaj, P. Ruggerone, M. Winterhalter, and M. Ceccarelli, "Analysis of fast channel blockage: revealing substrate binding in the microsecond range," *Analyst*, 2015. (Cited on pages 13, 24, 58, and 66.)

[43] D. Jeanteur, T. Schirmer, D. Fourel, V. Simonet, G. Rummel, C. Widmer, J. P. Rosenbusch, F. Pattus, and J.-M. Pagès, "Structural and functional alterations of a colicin-resistant mutant of ompf porin from escherichia coli," *Proceedings of the National Academy of Sciences*, vol. 91, no. 22, pp. 10675–10679, 1994. (Cited on page 14.)

[44] S. Vidal, J. Bredin, J.-M. Pagès, and J. Barbe, "β-lactam screening by specific residues of the ompf eyelet," *Journal of medicinal chemistry*, vol. 48, no. 5, pp. 1395–1400, 2005. (Cited on pages 14 and 66.)

[45] D. E. Shaw, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao, *et al.*, "Anton, a special-purpose machine for molecular dynamics simulation," *Communications of the ACM*, vol. 51, no. 7, pp. 91–97, 2008. (Cited on page 14.)

[46] "Barcelona supercomputer center." http://www.bsc.es/. (Cited on page 14.)

[47] "Tgcc curie supercomputer." http://www-hpc.cea.fr/en/complexe/tgcc-curie.htm. (Cited on pages 14 and 29.)

[48] J. E. Stone. (Cited on page 14.)

[49] M. Ceccarelli, C. Danelon, A. Laio, and M. Parrinello, "Microscopic mechanism of antibiotics translocation through a porin," *Biophysical journal*, vol. 87, no. 1, pp. 58–64, 2004. (Cited on pages 15, 24, 25, and 26.)

[50] L. Landau and E. Lifshitz, *Classical mechanics*. Pergamon Press, Oxford, 1960. (Cited on page 17.)

[51] L. Verlet, "Computer experiments on classical fluids. i. thermodynamical properties of lennard-jones molecules," *Physical review*, vol. 159, no. 1, p. 98, 1967. (Cited on page 18.)

[52] V. Kräutler, W. F. van Gunsteren, and P. H. Hünenberger, "A fast shake algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations," *Journal of computational chemistry*, vol. 22, no. 5, pp. 501–508, 2001. (Cited on page 18.)

[53] K. A. Feenstra, B. Hess, and H. J. Berendsen, "Improving e_ciency of large timescale molecular dynamics simulations of hydrogen-rich systems," *J Comput Chem*, vol. 20, pp. 786–798, 1999. (Cited on page 18.)

[54] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules," *Journal of the American Chemical Society*, vol. 117, no. 19, pp. 5179–5197, 1995. (Cited on page 19.)

[55] 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 page 19.)

[56] 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 20.)

[57] G. Malloci, A. V. Vargiu, G. Serra, A. Bosin, P. Ruggerone, and M. Ceccarelli, "A database of force-field parameters, dynamics, and properties of antimicrobial compounds," *Molecules*, vol. 20, no. 8, pp. 13997–14021, 2015. (Cited on page 20.)

[58] 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, no. 2, pp. 926–935, 1983. (Cited on pages 21 and 55.)

[59] P. P. Ewald, "Die berechnung optischer und elektrostatischer gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921. (Cited on page 22.)

[60] H. D. Herce, A. E. Garcia, and T. Darden, "The electrostatic surface term:(i) periodic systems," *The Journal of chemical physics*, vol. 126, no. 12, p. 124106, 2007. (Cited on page 22.)

[61] T. Darden, D. York, and L. Pedersen, "Particle mesh ewald: An n· log (n) method for ewald sums in large systems," *The Journal of chemical physics*, vol. 98, no. 12, pp. 10089–10092, 1993. (Cited on page 22.)

[62] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, "The protein data bank: a computer-based archival file for macromolecular structures," *Archives of biochemistry and biophysics*, vol. 185, no. 2, pp. 584–591, 1978. (Cited on page 22.)

[63] A. Šali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints," *Journal of molecular biology*, vol. 234, no. 3, pp. 779–815, 1993. (Cited on page 22.)

[64] N. Eswar, B. Webb, M. A. Marti-Renom, M. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper, and A. Sali, "Comparative protein structure modeling using modeller," *Current protocols in bioinformatics*, pp. 5–6, 2006. (Cited on page 22.)

[65] H. J. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of chemical physics*, vol. 81, no. 8, pp. 3684–3690, 1984. (Cited on page 23.)

[66] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of chemical physics*, vol. 81, no. 1, pp. 511–519, 1984.

[67] W. G. Hoover, "Canonical dynamics: equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, p. 1695, 1985. (Cited on page 23.)

[68] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," *The Journal of chemical physics*, vol. 72, no. 4, pp. 2384–2393, 1980. (Cited on page 23.)

[69] J. W. Gibbs, "Elementary principles in statistical mechanics, developed with especial reference to the rational foundations of thermodynamics," *Bull. Amer. Math. Soc. 12 (1906), 194-210 DOI: http://dx. doi. org/10.1090/S0002-9904-1906-01319-2 PII*, pp. 0002–9904, 1906. (Cited on page 23.)

[70] A. Laio and M. Parrinello, "Escaping free-energy minima," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12562–12566, 2002. (Cited on pages 25, 26, and 30.)

[71] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, "Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics," *The Journal of Physical Chemistry B*, vol. 110, no. 8, pp. 3533–3539, 2006. (Cited on pages 29, 30, and 59.)

[72] G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello, "Free-energy landscape for β hairpin folding from combined parallel tempering and metadynamics," *Journal of the American Chemical Society*, vol. 128, no. 41, pp. 13435–13441, 2006. (Cited on page 32.)

[73] A. Laio and F. L. Gervasio, "Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science," *Reports on Progress in Physics*, vol. 71, no. 12, p. 126601, 2008. (Cited on pages 26, 27, 29, 30, and 32.)

[74] A. Barducci, G. Bussi, and M. Parrinello, "Well-tempered metadynamics: A smoothly converging and tunable free-energy method," *Physical review letters*, vol. 100, no. 2, p. 020603, 2008. (Cited on pages 29, 31, and 59.)

[75] M. Bonomi and M. Parrinello, "Enhanced sampling in the well-tempered ensemble," *Physical review letters*, vol. 104, no. 19, p. 190601, 2010. (Cited on page 32.)

[76] A. Barducci, M. Bonomi, and M. Parrinello, "Metadynamics," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 1, no. 5, pp. 826–843, 2011.

[77] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "Plumed 2: New feathers for an old bird," *Computer Physics Communications*, vol. 185, no. 2, pp. 604–613, 2014. (Cited on pages 25, 26, 29, and 32.)

[78] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, "Gromacs: fast, flexible, and free," *Journal of computational chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005. (Cited on page 29.)

[79] J. M. Rosenbergl, "The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method," *Journal of computational chemistry*, vol. 13, no. 8, pp. 1011–1021, 1992. (Cited on page 28.)

[80] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "Multidimensional free-energy calculations using the weighted histogram analysis method," *Journal of Computational Chemistry*, vol. 16, no. 11, pp. 1339–1350, 1995.

[81] J. S. Hub, B. L. De Groot, and D. Van Der Spoel, "g_wham: A free weighted histogram analysis implementation including robust error and autocorrelation estimates," *Journal of Chemical Theory and Computation*, vol. 6, no. 12, pp. 3713–3720, 2010. (Cited on page 28.)

[82] E. L. B. H. M.J. Abraham, D. van der Spoel and the GROMACS development team, "Gromacs user manual version 5.0.7, www.gromacs.org," 2015. (Cited on page 32.)

[83] S. Piana and A. Laio, "A bias-exchange approach to protein folding," *The Journal of Physical Chemistry B*, vol. 111, no. 17, pp. 4553–4559, 2007. (Cited on page 32.)

[84] F. Marinelli, "Following easy slope paths on a free energy landscape: the case study of the trp-cage folding mechanism," *Biophysical journal*, vol. 105, no. 5, pp. 1236–1247, 2013. (Cited on page 32.)

[85] U. H. Hansmann, "Parallel tempering algorithm for conformational studies of biological molecules," *Chemical Physics Letters*, vol. 281, no. 1, pp. 140–150, 1997. (Cited on page 32.)

[86] D. J. Earl and M. W. Deem, "Parallel tempering: Theory, applications, and new perspectives," *Physical Chemistry Chemical Physics*, vol. 7, no. 23, pp. 3910–3916, 2005.

[87] H. Merlitz and W. Wenzel, "Comparison of stochastic optimization methods for receptor–ligand docking," *Chemical physics letters*, vol. 362, no. 3, pp. 271–277, 2002. (Cited on page 32.)

[88] M. Deighan, M. Bonomi, and J. Pfaendtner, "Efficient simulation of explicitly solvated proteins in the well-tempered ensemble," *Journal of Chemical Theory and Computation*, vol. 8, no. 7, pp. 2189–2192, 2012. (Cited on page 32.)

[89] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. De Vries, "The martini force field: coarse grained model for biomolecular simulations," *The Journal of Physical Chemistry B*, vol. 111, no. 27, pp. 7812–7824, 2007. (Cited on pages 33 and 35.)

[90] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S.-J. Marrink, "The martini coarse-grained force field: extension to proteins," *Journal of chemical theory and computation*, vol. 4, no. 5, pp. 819–834, 2008. (Cited on page 33.)

[91] D. H. de Jong, G. Singh, W. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schafer, X. Periole, D. P. Tieleman, and S. J. Marrink, "Improved parameters for the martini coarse-grained protein force field," *Journal of Chemical Theory and Computation*, vol. 9, no. 1, pp. 687–697, 2012.

[92] X. Periole and S.-J. Marrink, "The martini coarse-grained force field," in *Biomolecular Simulations*, pp. 533–565, Springer, 2013. (Cited on page 33.)

[93] T. A. Wassenaar, K. Pluhackova, R. A. Bo"ockmann, S. J. Marrink, and D. P. Tieleman, "Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models," *Journal of Chemical Theory and Computation*, vol. 10, no. 2, pp. 676–690, 2014. (Cited on page 34.)

[94] F. Marinelli, F. Pietrucci, A. Laio, and S. Piana, "A kinetic model of trp-cage folding from multiple biased molecular dynamics simulations," *PLoS Comput. Biol*, vol. 5, no. 8, p. e1000452, 2009. (Cited on page 36.)

[95] 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 36.)

[96] P. Tiwary and M. Parrinello, "From metadynamics to dynamics," *Physical review letters*, vol. 111, no. 23, p. 230602, 2013. (Cited on pages 36, 58, 59, and 65.)

[97] P. Tiwary, V. Limongelli, M. Salvalaglio, and M. Parrinello, "Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps," *Proceedings of the National Academy of Sciences*, vol. 112, no. 5, pp. E386–E391, 2015. (Cited on pages 36, 37, 58, and 65.)

[98] 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. 1420–1425, 2014. (Cited on pages 37 and 65.)

[99] S. M. Ross *et al.*, *Stochastic processes*, vol. 2. John Wiley & Sons New York, 1996. (Cited on page 37.)

[100] J. Ye and B. Van den Berg, "Crystal structure of the bacterial nucleoside transporter tsx," *The EMBO journal*, vol. 23, no. 16, pp. 3187–3195, 2004. (Cited on page 41.)

[101] C. Wirth, G. Condemine, C. Boiteux, S. Bernèche, T. Schirmer, and C. M. Peneff, "Nanc crystal structure, a model for outer-membrane channels of the acidic sugar-specific kdgm porin family," *Journal of molecular biology*, vol. 394, no. 4, pp. 718–731, 2009. (Cited on page 41.)

[102] H. Hong, D. R. Patel, L. K. Tamm, and B. van den Berg, "The outer membrane protein ompw forms an eight-stranded β-barrel with a hydrophobic channel," *Journal of biological chemistry*, vol. 281, no. 11, pp. 7568–7577, 2006. (Cited on page 41.)

[103] H. Hong, G. Szabo, and L. K. Tamm, "Electrostatic couplings in ompa ion-channel gating suggest a mechanism for pore opening," *Nature chemical biology*, vol. 2, no. 11, pp. 627–635, 2006. (Cited on page 41.)

[104] J. S. Esterly, C. L. Richardson, N. S. Eltoukhy, C. Qi, and M. H. Scheetz, "Genetic mechanisms of antimicrobial resistance of acinetobacter baumannii," *Annals of Pharmacotherapy*, vol. 45, no. 2, pp. 218–228, 2011. (Cited on page 42.)

[105] M. A. Mussi, A. S. Limansky, and A. M. Viale, "Acquisition of resistance to carbapenems in multidrug-resistant clinical strains of acinetobacter baumannii: natural insertional inactivation of a gene encoding a member of a novel family of β-barrel outer membrane proteins," *Antimicrobial agents and chemotherapy*, vol. 49, no. 4, pp. 1432–1440, 2005. (Cited on pages 42, 44, and 49.)

[106] M. A. Mussi, A. S. Limansky, V. Relling, P. Ravasi, A. Arakaki, L. A. Actis, and A. M. Viale, "Horizontal gene transfer/assortative recombination within the acinetobacter baumannii clinical population provides genetic diversity at the single caro gene encoding a major outer membrane protein channel.," *Journal of bacteriology*, pp. JB–01533, 2011. (Cited on pages 44 and 49.)

[107] M. Catel-Ferreira, G. Coadou, V. Molle, P. Mugnier, P. Nordmann, A. Siroy, T. Jouenne, and E. Dé, "Structure–function relationships of caro, the carbapenem resistance-associated outer membrane protein of acinetobacter baumannii," *Journal of antimicrobial chemotherapy*, vol. 66, no. 9, pp. 2053–2056, 2011. (Cited on page 45.)

[108] J. Vashist, V. Tiwari, A. Kapil, and M. R. Rajeswari, "Quantitative profiling and identification of outer membrane proteins of β-lactam resistant strain of acinetobacter baumannii," *Journal of proteome research*, vol. 9, no. 2, pp. 1121–1128, 2010. (Cited on page 49.)

[109] H. Nikaido, E. Rosenberg, and J. Foulds, "Porin channels in escherichia coli: studies with beta-lactams in intact cells.," *Journal of Bacteriology*, vol. 153, no. 1, pp. 232–240, 1983. (Cited on page 51.)

[110] A. Kumar, E. Hajjar, P. Ruggerone, and M. Ceccarelli, "Structural and dynamical properties of the porins ompf and ompc: insights from molecular simulations," *Journal of Physics: Condensed Matter*, vol. 22, no. 45, p. 454125, 2010. (Cited on pages 51 and 60.)

[111] H. Lou, M. Chen, S. S. Black, S. R. Bushell, M. Ceccarelli, T. Mach, K. Beis, A. S. Low, V. A. Bamford, I. R. Booth, *et al.*, "Altered antibiotic transport in ompc mutants isolated from a series of clinical strains of multi-drug resistant e. coli," *PLoS One*, vol. 6, no. 10, p. e25825, 2011. (Cited on pages 51, 52, 54, and 57.)

[112] S. Acosta Gutierrez, M. A. Scorciapino, I. Bodrenko, and M. Ceccarelli, "Filtering with electric field: The case of e. coli porins," *The Journal of Physical Chemistry Letters*, 2015. (Cited on pages 52, 54, 56, 57, and 62.)

[113] R. McNulty, J. P. Ulmschneider, H. Luecke, and M. B. Ulmschneider, "Mechanisms of molecular transport through the urea channel of helicobacter pylori," *Nature communications*, vol. 4, 2013. (Cited on page 57.)

[114] J. Bredin, N. Saint, M. Malléa, E. De, G. Molle, V. Simonet, *et al.*, "Alteration of pore properties of escherichia coli ompf induced by mutation of key residues in anti-loop 3 region," *Biochem. J*, vol. 363, pp. 521–528, 2002. (Cited on page 66.)

[115] J. Bredin, V. Simonet, R. Iyer, A. Delcour, *et al.*, "Colicins, spermine and cephalosporins: a competitive interaction with the ompf eyelet," *Biochem. J*, vol. 376, pp. 245–252, 2003.

[116] V. Simonet, M. Malléa, and J.-M. Pagès, "Substitutions in the eyelet region disrupt cefepime diffusion through the escherichia coli ompf channel," *Antimicrobial agents and chemotherapy*, vol. 44, no. 2, pp. 311–315, 2000. (Cited on page 66.)

[117] 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 page 66.)