Master Erasmus Mundus in

Theoretical Chemistry and Computational Modelling

European Master in

Theoretical Chemistry and Computational Modelling

Computation of the Free-Energy Profile along the Transport of Carboplatin across a Dioleoylphosphocholine Membrane

Cristina Gil Herrero

Director: Juan José Nogueira Pérez
Place where the project was carried out: Autonomous University of Madrid, Faculty of Science, Department of Chemistry

Master Thesis. Course 2020-2021

Abstract

Carboplatin is one of the leading chemotherapeutic drugs, however, its working principles in biological cells are not fully clarified. In this work we tried to shine light on the permeation process of this compound across a model lipid membrane (DOPC), which is presumably a crucial step of its cytotoxic activity, by the computation and analysis of its free-energy profile. For such a purpose, this project entails a combination of the umbrella sampling and WHAM approaches, along with steered MD trajectories. The asymmetry found in the free energy profiles and their erratic trend with the temperature revealed that the parameters employed in the simulations were not appropriate to provide accurate energies. Different force constant values for the bias potential and different number of windows were used during the simulations with the aim of obtaining higher quality profiles. However, all the attempts were unsuccessful. Overall, the results revealed that the diffusion of this compound through the membrane may be a favourable process, though further work is required to unveil the permeation mechanism.

Resumen

El carboplatino es, sin duda, uno de los agentes quimioterapéuticos más importantes. Sin embargo, el mecanismo de acción en las células no ha sido totalmente clarificado hasta la fecha. En este trabajo, pretendemos arrojar luz sobre el proceso de transporte de esta molécula a través de una membrana lipídica modelo (DOPC), el cuál es un paso clave para su actividad citotóxica, mediante la estimación y análisis de su perfil de energía libre. Para dicho propósito, en este proyecto se ha empleado una combinación de los métodos umbrella sampling, WHAM y steered MD. La asimetría encontrada en el perfil de energía libre y su comportamiento errático con la temperatura revelaron que los parámetros empleados en las simulaciones no fueron adecuados para obtener energías exactas. Diferentes constantes de fuerza para el potential artificial y número de ventanas han sido usados durante las simulaciones con el objetivo de obtener perfiles energéticos de mayor calidad. Sin embargo, todos los intentos fueron infructuosos. Finalmente, los resultados indican que el proceso de difusión de dicho compuesto a través de la membrana puede darse de forma favorable, aunque cálculos adicionales son necesarios para confirmar la veracidad de este fenómeno y el mecanismo de transporte.

Acknowledgements

Agradezco el tiempo computacional en el Centro de Computación Científica de la Universidad Autónoma de Madrid (CCC-UAM).
Además, me gustaría agradecer el tiempo y la dedicación de Gustavo y de Juanjo, mi paciente director de TFM.
Me gustaría agradecer todo el apoyo incondicional de mi familia, que a pesar de todos los momentos de frustracción siempre han sabido estar ahí de la forma en la que les necesitaba. En especial a mi hermana, fiel compañera de estudio durante este complicado año.
No podría olvidarme de mis amigos, en especial Sara y Álvaro, cuya preocupación constante y quedadas semanales me han ayudado a sobrellevar el estrés de la mejor manera posible.
Pero a quién más me gustaría agradecer su apoyo, ayuda y consejos es a Miguel, quién a pesar de la distancia ha sabido estar ahí de la mejor de las maneras.

Contents

Chapter 1

Introduction.

1.1 Platinating agents

The second-leading cause of death in the world are cancer diseases, consisting on an abnormal cell-growth which tends to proliferate in an uncontrolled way destroying normal body tissue. There are more than a hundred of distinctive cancer diseases.[1, 2]

The accidental discovery of the clinical applications of cisplatin, the first metal-based anticancer drug, played a key role in the cure of certain cancers such as the testicular one, whose probabilities of overcoming the disease raised till 95%.[3] Platinum-based compounds have been one of the most widely employed chemotherapeutic agents along the last 40 years for the therapy of various tumours.[4, 5] The three leading compounds approved worldwide are cisplatin, carboplatin and oxaliplatin whose structures are found respectively in black, blue and red in figure 1.1.

Cisplatin (cis-diamminedichloroplatinum(II)) is considered as the first generation and one of the most potent platinum chemotherapeutics. Nonetheless, the cisplatin treatment has to be complemented with a wide variety of drugs to alleviate several side problems related to immunity, gastrointestinal problems or kidney diseases among others. It is frequently used in bladder, head and neck, lung, ovarian and testicular cancers, in addition to sarcomas and lymphomas.[6, 7]

Another platinum complex considered as the second generation was carboplatin (cis-diammine (1,1-cyclobutanecarboxylate)platinum(II)), that was developed with the aim of limiting chemotherapy side effects providing analogous results. The carboplatin molecule presents a similar mode of action to the cisplatin despite its bulkier structure, where the two chlorines of cisplatin are

substituted by a bidentate ligand. This newer drug differs in the toxicity profile as it rarely produces nephrotoxicity, neurotoxicity and ototoxicity becoming more suitable when higher doses are required. It is employed in the treatment of patients with cell lung cancer, cell carcinomas of the head and neck and specially in ovarian cancer.[8, 9]

Oxaliplatin (oxalate(trans-l-1,2-diamminocyclohexane)platinum(II)) belongs to the third generation of these chemotherapy drugs. Its development was motivated by the appearance of resistance to cisplatin or carboplatin treatments in cells. In this complex, the chlorine ligands of the cisplatin are substituted by the oxalato bidentate ligand, and moreover, in place of the two monodentate ammine ligands a 1,2-diaminocyclohexane is found composing an even bulkier molecule than carboplatin. These bidentate ligands difficult its recognition by the repair machinery of the DNA leading to a higher probability to cell death than the previous compounds Therefore, it achieves its purpose of being an alternative when cells have already acquired resistance to cisplatin or carboplatin, whose resistance mechanisms are very similar. It is mainly applied in patients diagnosed with colorectal tumors.[10, 11, 12]

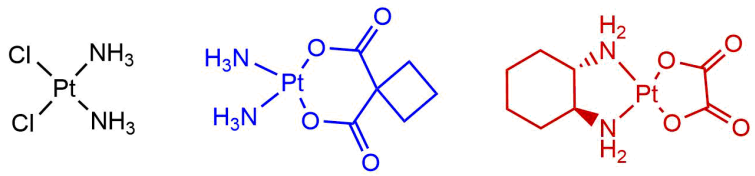

FIGURE 1.1: Pt(II) drugs: Cisplatin in black, carboplatin in blue and oxaliplatin in red.

Apart from these three compounds approved worldwide in which the oxidation state of platinum is II and the complex structure is square-planar, platinum(IV) complexes can also be an option in chemotherapy. They tend to adopt an octahedral geometry that contributes to avoid previous undesired reactions with molecules different from DNA. Nevertheless, they are called platinum(IV) prodrugs as they suffer from an irreversible reduction of the central metal releasing the two axial ligands from their coordination sphere becoming analogues to the previous platinum(II) agents. This is a crucial step for the activation of these compounds since, though they are able to react with DNA, the complete cytotoxic mechanism would last weeks while these compounds are eliminated from the body in hours. Some of the most studied platinum(IV) agents are tetraplatin, iproplatin and satraplatin whose structures are shown in figure 1.2.[3]

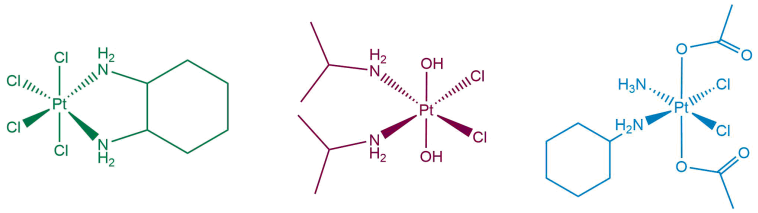

FIGURE 1.2: Pt(IV) drugs: tetraplatin in green, iproplatin in dark red and satraplatin in light blue.

1.2 Mechanism of action

The general mechanism of action of these drugs that is represented in figure 1.3 consists on entering the cell either by passive transport diffusing through the plasma membrane or by active transport. The proteins involved in this uptake mechanism are the cell transporters: the copper transport proteins, as the copper transporter-1 (CTR1), and/or organic cation transporters (OCTs). Although CTRs have generally been claimed to be highly selective, the analogous binding of copper(I) and Pt(II) to the sulfur ligand has been proved to be responsible for the role of CTRs in platinum-drugs cell uptake.[10, 13]

Although nowadays some details of this step are still not clearly established, studies have evidenced that the entrance of cisplatin into the cell is produced mainly by passive diffusion, where the interactions of the molecule are limited to the lipids in the membrane. Nevertheless, this importance of passive diffusion verified in cisplatin, cannot be generalised for all the platinum complexes as it strongly depends on the molecule and the membrane of the cell under study.[14, 15]

Experiments conducted using an overexpression of organic cation transporters (OCTs) revealed its great relevance in the oxaliplatin mechanism in contrast with its possibly very little contribution in the cisplatin or carboplatin transport. [13] The amount of these drugs that enter the cell determines the drug accumulation, being directly related to the cytotoxicity of these pharmacological compounds.

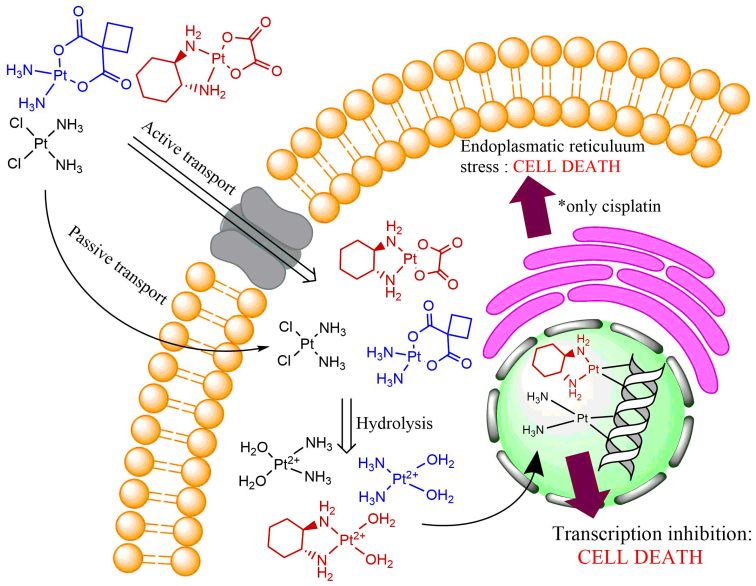

FIGURE 1.3: Mechanism of action of cisplatin, carboplatin and oxaliplatin when reaching the cell. The entrance into the molecule, the aquation process, the penetration to the nucleus and binding to the DNA along with the alternative cisplatin pathway.

Upon entering the cell, all platinating agents become aquated loosing the chlorine or oxalate ions that are substituted by two water molecules. The aquation process occurs in the cytoplasm due to a drop in some salt concentrations as the chlorine ion, which is lower than 20mM with respect to the bloodstream where it is approximately 100mM. This step takes place in shorter periods for the cisplatin than for the carboplatin or oxaliplatin molecules due to their chelating ligands. Afterwards, being positively charged, they are able to enter the nucleus and interact with nucleophilic molecules within the cell, with special consideration for the negatively charged DNA molecule which is agreed to be its preferential target.[3]

The covalent binding of these compounds with the DNA is performed on the N7 position of the purine bases; guanin and adenin, producing monoadducts, intrastrand crosslinks or interstrand crosslinks. The lesion referred to as monoadducts is initially formed when one molecule of water is lost from aquated platinating agents, nevertheless, in more than 90% of the formed monoadducts the drug, then, react again to form crosslinks. Furthermore, as it can be inferred from its name, the intrastrand crosslinks are formed when both of the water molecules are lost in the platinum complex, being able to establish two covalent bonds within the same DNA strand. In addition, the interstrand crosslinks are similar to the latter except for the DNA binding, which is formed with the two

strands. Although all crosslinks result in a distortion of the DNA helix, the interstrand ones tend to produce a higher bending. The inability of repairing the DNA damage ends up with the apoptosis, being therefore responsible of the platinum complexes cytotoxicity.[16]

Once the DNA alteration has been produced, the cell-cycle checkpoints are able to detect it producing an interruption of the normal cell cycle for its reparation before the cell division takes place, otherwise, the cancer cell proliferation will continue. The cell will activate the nucleotide repair system considered as the major pathway responsible for the removal of platinum-adducts. It mainly includes the high-mobility group (HMG) box proteins, repair proteins and transcription factors, among others, to attempt the removal of this damage. If the repair mechanisms succeed, it could lead to the acquirement of resistance to platinum compounds by the cells. Nevertheless, if not achieved, the HMG box protein HMGB1 appears totally altered in the damaged cell, becoming a signal of tissue affected. In addition, signal transduction pathways are modified frequently inhibiting the transcription process, and increasing proapoptotic signals which entails the initialization of the apoptotic cascade.[17, 15]

On the other hand, cell death is not only produced by disrupting the DNA structure, an alternative target of cisplatin when penetrating the cell was found, the endoplasmic reticulum (ER). Cisplatin induces ER stress, resulting on inhibition of the transport and folding of proteins which might be caused by alterations in calcium homeostasis.[18, 19]

1.3 Membrane composition

The bilayer composition is thought to play a major role in the permeability of the membrane to drugs and small molecules, producing an increment of the cellular uptake. The lipid cell membrane is composed of polar head groups that cover the surface of the outer leaflet, and non-polar fatty acids hydrocarbon chains that belong to the inner leaflet. The variability in lipid bilayers is mainly determined by the types of lipids involved and their relative abundances, along with the different proteins and steroids, specially cholesterol molecules. In addition to the organic molecules bonded to the phosphate group in the polar heads while the non-polar tails may differ in length as well as in number and position of double bonds.

The composition of the plasma membrane of cancer cells might differ from the non-malignant profile, nevertheless, it can also vary among different types of tumours in the same manner that it does for the diverse tissues. In addition, their composition could fluctuate in time according to the physiological conditions.[20] This entails a non selective distribution of platinum drugs in normal and cancer cells, representing one of the major problems, the lack of tumour specificity. [21]

Phosphatidylcholine (PC) is a neutral membrane lipid and one of the most abundant components along with cholesterol (CL), followed by phosphoetanolamine (PE), sphingomyelin (SM), phosphatidylserine (PS) and phosphatidylinositol (PI). It is frequently observed that the inner and outer leaflets of the bilayer present some variations in their composition.

Despite the fact that single-component membranes are not the most realistic ones, for the sake of simplicity the model membrane composition employed during the simulation was dioleoyl-phosphatidylcholine (DOPC). This symmetric single lipid component bilayer has been successfully used in a wide variety of studies as it involves phosphatidylcholine(PC), one of the major components of biological membranes.[22]

Chapter 2

Methodology

The birth and implementation of Molecular Dynamics (MD) simulations, along with quantum and computational chemistry, have been widely employed in different research fields. Its applicability has provided contributions to the pharmacological area by allowing for the study of the structural, functional and dynamical properties of drugs and proteins at an atomistic scale, consequently leading to improvements. Furthermore, it can be used for simplifying the evaluation of the binding energetics and kinetics helping in the search of more appropriate compounds.[23]

2.1 Classical molecular dynamics

MD is a simulation method that aims to describe the time evolution of the system and determine macroscopic properties that depend on this evolution, for example kinetics and thermodynamic properties. MD is useful in the investigation of biological systems among others, in this particular case, this systems are composed by biomolecules which, for the accuracy of the simulation, are usually surrounded by solvent molecules to correctly consider the influence of the environment.[24, 25]

These simulations pretend to be as similar as possible to a realistic situation, for that objective, initial conditions are firstly set, providing initial positions and velocities of the particles, this step is followed by a required equilibration period before starting with the production stage from which properties are computed. Nevertheless, it has to be considered that there are some parameters that need to be properly chosen along with the mathematical approximations employed which can lead to some problems or low inaccuracy.[26]

2.1.1 Equations of motion

Classical MD simulations are governed by the Newton's laws of motion, equivalent to the Hamilton's equations of motion if the system is conservative, which are integrated in order to simulate the dynamics of the system by acquiring new positions for the particles with their respective new velocities.[27, 28]

Considering Newton's second law ($F = \frac{\delta p}{\delta t}$) and that the forces acting on the particles are conservative ($F = -\frac{\delta V}{\delta r}$) for the first equation of motion and the derivative of the kinetic energy expression with respect to the linear momentum ($\frac{\delta T}{\delta p} = \frac{p^2/2m}{\delta p}$). The Hamilton's equations of motion result the following:

$$\frac{\delta p_i}{\delta t} = -\frac{\delta V}{\delta r_i} \tag{2.1}$$

$$\frac{\delta r_i}{\delta t} = -\frac{\delta T}{\delta p_i} \tag{2.2}$$

where p_i is the linear momentum of a particle that is the product of the mass and velocity of the particle ($p = mv$), r_i is the position of the particle, and V and T are the potential and kinetic energies respectively.[29, 30]

2.1.2 Integrators

A simple scheme of a simulation consists on: given the initial positions and velocities of a number of particles, forces acting on them and thus acceleration are calculated, then by reproducing the movement of the system, the new positions and velocities are acquired, repeating these steps till completing the overall simulation time.[26]

Integrators are algorithms whose purpose is to propagate the particle positions and velocities along time every certain finite time difference, known as time step (δt) in accordance to Hamilton's equations of motion. The methods presented below employ a third-order Taylor expansion for their purpose. The time step is a determinant factor in the accuracy of the calculations: if it is not small enough, the simulation can suffer from energy drifts, while if its value is extremely low, in order to simulate the overall simulation time the computational cost will be unreasonable. [28]

Verlet algorithm

The Verlet integrator is commonly apply due to its simplicity along with its accuracy. It is very easy to implement and does not employs much computer memory. The position of each particles at time $t + \delta t$ is computed as follows:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + a(t)\delta t^2 \tag{2.3}$$

where r stands for the position of the particle, and a represents the acceleration which is equivalent to the force underwent by the particle divided by its mass ($a = F/m$).

The main problem of this algorithm lies in the lack of the velocities, which although they are not required for the new positions of the particles, are key for the computation of the kinetic energy or instantaneous temperatures. They can be estimated by using the position terms: $r(t + \delta t)$ and $r(t - \delta t)$:

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \tag{2.4}$$

However, as the calculation of velocities is one step behind the positions their accuracy drops considerably with respect to the positions. If the kinetic energy is no correctly calculated, the evaluation of the total energy conservation fails, not achieving the aim of checking that a MD simulation is running properly.

Leap-frog algorithm

The Leap-frog integrator is another commonly applied algorithm in MD simulations. This algorithm computes positions and velocities but at different steps: firstly it obtains velocities at what is called mid-step ($t + \frac{\delta t}{2}$), and then utilises this velocities for the computation of the positions at what is called full step ($t + \delta t$).

Furthermore, it also introduces a problem when calculating the variables at different times, as the kinetic and potential energies are acquired at different moments, thus the total energy at each time step ($t + \delta t$) cannot be directly computed.[26]

$$r\left(t + \delta t\right) = r(t) + v\left(t + \frac{\delta t}{2}\right)\delta t \tag{2.5}$$

$$v\left(t + \frac{\delta t}{2}\right) = v\left(t - \frac{\delta t}{2}\right) + a(t)\delta t \tag{2.6}$$

Velocity verlet algorithm

This algorithm improves comparing to standard Verlet integrator due to the calculation of velocities at the same step as positions, which avoids storing the positions of the last three steps for that. In addition, it does not present the leap-frog algorithm problem in the computation of the total energy, for the same reason. For each integration cycle, the particle positions are recalculated, then the velocity is obtained at mid-step, before acquiring the velocity of the second half an update of the value of the forces is made, modifying the acceleration.[26]

$$r\left(t+\delta t\right) = r(t) + \delta t v(t) + \frac{1}{2}a(t)\left(\delta t\right)^2 \tag{2.7}$$

$$v\left(t+\frac{\delta t}{2}\right) = v(t) + \frac{1}{2}a(t)\delta t \tag{2.8}$$

$$a(t+\delta t) \quad is \quad calculated \tag{2.9}$$

$$v\left(t+\delta t\right) = v\left(t+\frac{\delta t}{2}\right) + \frac{1}{2}a(t+\delta t)\delta t \tag{2.10}$$

2.1.3 Force fields

The forces computed by the integrators can be obtained by applying a force field, or solving the Schrödinger equation, giving rise to classical and ab initio MD, respectively. The aim of force fields is to reproduce as accurately as possible the interactions among atoms and molecules. They are employed for the calculation of the potential energy of a system. Force Fields are constituted by a set of parameters and functions.

To describe the interactions there are two types of potentials; bonded and non bonded. In general, the bonded interactions are split into bond potentials (V_{bond}), angle potentials (V_{angle}) and dihedral potentials (V_{dihed}), while the non bonded energy is composed of two terms; Coulomb and van der Waals energy term that can be split into attractive and repulsion interactions. The overall energy is thus computed as:

$$V_{total} = V_{bond} + V_{angle} + V_{dihedral} + V_{Coulomb} + V_{vdW} \tag{2.11}$$

The bond potentials describe the energy variation between two atoms placed at a bond length or of the three atoms involved in the angle. It is often defined

by harmonic potentials that have the following expression:

$$V_{bond} = \frac{1}{2}k(x - x_0)^2 \tag{2.12}$$

where $x - x_0$ is the distance difference from the equilibrium and k is a force field parameter in both equations.

The angle potentials describe the energy variation between three atoms describing an angle. It is also usually defined by a harmonic potential:

$$V_{angle} = \frac{1}{2}k(\alpha - \alpha_0)^2 \tag{2.13}$$

where $\alpha - \alpha_0$ is the angle variation from the equilibrium angle and k is the force field parameter.

On the other hand, dihedral angle or torsion potentials, which are composed by four atoms, are usually represented by a Fourier series:

$$V_{dihedral} = k\,(1 + cos(n\omega - \gamma)) \tag{2.14}$$

where the ω term is an angle difference between the planes composed by the first three atoms and the last three, while k, n and γ are the parameters of the forcefield.

Regarding the non bonded interactions, the Coulombic electrostatic interactions potential presents the expression below:

$$V_{coulomb} = \frac{q_i q_j}{4\pi\epsilon r_{ij}} \tag{2.15}$$

where q_i and q_j are the charges from the pair of atoms and ϵ is the vacuum permitivity, and r_{ij} is the distance between the atoms.

The van de Waals and the repulsion potentials are frequently characterize by a Lennard-Jones potential:

$$V_{LJ}(r_{ij}) = 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{2.16}$$

where ϵ is the depth of the potential, σ is the equilibrium distance, r_{ij} the distance between the two particles involved, $\left(\frac{\sigma}{r_{ij}}\right)^{12}$ represents the repulsive force due to the Pauli repulsion while $\left(\frac{\sigma}{r_{ij}}\right)^{6}$ characterizes the attractive term of the van der Waals interactions.[30, 31]

The objective of the application of force fields is to reproduce properties of molecules extracted from experimental data. As a consequence, it exists a wide variety of force fields that can be applied to any atom or molecule. However, the excessive generalization of some force fields is achieved at the cost of a limited computational accuracy on the results.

On the other hand, specialized force fields can be developed for a more precise description of the simulated molecule such as it is done for the carboplatin in this project. Nevertheless, it has to be considered that their transferability to other compounds could be extremely limited.[32]

2.1.4 Simulation conditions

Thermostats

In a MD simulation of an isolated system an energy-conserving trajectory gives rise to the microcanonical ensemble (NVE) in which the number of particles, the volume of the system and its total energy are kept constant. Nevertheless, it is often found more interesting to apply a thermal ensemble or canonical ensemble (NVT) in which instead of constant total energy, it is the temperature the one that remains invariant during the simulation.As a result from the NVT ensemble, the equilibrium of the particles obeys the Maxwell-Boltzmann distribution.

For the purpose of controlling the temperature of a simulation algorithms denominated "thermostats" are employed. This algorithms can be divided in global and local thermostats; while the former is coupled to the total kinetic energy, depending on the temperature defined for the system, leading to a modification of the velocities of all particles, the latter acts in a localized manner by decreasing or increasing the energy provided to just certain particles.

The simplest global thermostats are Velocity Rescaling and Berendsen thermostat On the other hand, for the algorithms that belong to the local type of

thermostats : Andersen, Langevin and Nosé-Hoover-Andersen are examples of this category.[33]

The Velocity Rescaling algorithm is the simplest thermostat. It acts redefining the velocities of all the particles at every step or every certain number of steps in order for the kinetic energy of the system to be equal to the target one. However, it has to be contemplated that if this thermostat is apply at every step the kinetic energy remains constant resulting in not achieving a proper NVT ensemble.[34]

The Berendsen thermostat acts in a similar way to the Velocity Rescaling differing in the definition of the factor that rescales the velocities. It supposes that the system is weakly coupled to a thermal bath, nevertheless, the same difficulty for reproducing the NVT ensemble is encountered.[35]

The Langevin thermostat is widely applied, it acts by adding at each time step to all particles a random force simulating collisions and lowering their velocities with a constant friction coefficient. The random forces are extracted from a Gaussian distribution in which the temperature is considered. This method achieves the balance variation-dissipation, ensuring that the NVT ensemble is satisfied. This is the thermostat applied in all the calculations developed along this project for the correct simulation of the cell environment by avoiding temperature fluctuations.[33]

When referring to the Andersen thermostat, its main idea is to select a random particle and extract its velocity from the Gaussian distribution. This algorithm may introduce discontinuities in the trajectories caused by the random velocities leading to poor efficiency which makes it a not very commonly applied thermostat.[34]

The Nosé-Hoover thermostat accomplishes the NVT ensemble without managing random forces or velocities. It controls the velocities of the particles by a friction factor that actually scales the velocity including the mass, an additional degree of freedom.[36]

Pressure regulation or barostats

Similarly to thermostats, with the aim of controlling the pressure, certain algorithms called ”barostats” are used. They are specially relevant when the

simulations are desired to be performed under constant pressure conditions such as in isobaric-isothermal (NPT) ensemble, which is convenient to replicate the conditions commonly found in empirical experiments. Three simulation modes achieves the aim of maintaining the pressure constant; isotropically, semi-isotropically and anisotropically. In the isotropically coupling the pressure is scaled for the three dimensions in an equivalent form while in the semi-isotropic coupling the scaling is performed on the z axis, in this case normal to the membrane or to a surface, and the x-y plane separately. In addition, the anisotropic coupling consists on scale every dimension (x,y,z) independently.[37]

For a bilayer, it is important to maintain the pressure constant along the simulations for correctly mimicked cell conditions which allows the system to fluctuate in size and shape of the box. These possibly slight variations are disregarded and 1 bar of lateral and normal pressure is established. The simulations are performed anisotropically during the second heating and semi-isotropically for the rest of them.[38]

The Berendsen Barostat employs an external bath modifying the global pressure in order to be equal to the pressure bath by modifying the volume. The main advantage of this algorithm is that it is very simple to include into an existing program.[35][39]

Boundary Conditions

When setting up the conditions of a simulation, one of the key aspects is the environment of the reproduced system. This is defined by the boundary conditions, and it is a determining factor for the consideration of the interactions in the limits of our defined cell.

One of the most employed types of boundary conditions is the periodic boundary condition (PBC), which consists on infinitely replicating periodic images of the specified simulation box. Along the simulation, the properties of the original cell are propagated among the other adjacent images allowing unit cells to interact with particles from others. It is translated in the simultaneous appearance and disappearance of particles at opposites boundaries of the cell. The shape of this unit systems is not required to be square, but it can present different polyhedral frames. All in all, the most common one is a parallelepiped defined by the three different vectors.

Furthermore, this replications can be applied in one, two or three dimensions. On the one hand, in order to simulate the behaviour of gases, liquids or solids the three dimensional PBCs provide the best approximation. On the other hand, if the system to simulate entails planar surfaces, the PBCs should be established in two dimensions, which is also designated as slab boundary conditions. Moreover, if the PBCs are only used in one of the directions they are considered as wire configurations.

2.2 Enhanced Sampling Methods

The Enhanced Sampling methods can also be denominated Free Energy methods as their main goal is to determine the free energies of complex systems.[40] Enhanced Sampling methods are required either for accelerating the slow motion of simulated reactions that would not be computationally affordable or for overcoming high energy barriers in certain reactions that would rarely be sampled.

One of the most applied approaches is to add a bias potential to the Hamiltonian altering the potential energy surface (PES), this provides an energy boost to the system that allows for populating high energy regions that would otherwise be impossible to explore.[41]

2.2.1 Steered molecular dynamics

Steered molecular dynamics (SMD) also referred to as force probe simulations, is a method consisting of the acceleration of the dynamics of the simulated processes. It is considered as a non-equilibrium sampling technique due to the fact that for a correct simulation of an equilibrium process in which a perturbation such as a potential is involved, the application of this perturbation should be performed infinitely slow, which does not tend to occur in practise.[42]

The SMD approach employs time-dependent external forces to facilitate the corresponding process to occur. In this project the external force is applied as an harmonic potential to retrain the carboplatin ligand to a certain point in the space, shifting this point along the reaction coordinate pulling the carboplatin until the end of the reaction coordinate. For the correct application of the force, the direction should be specified, being a straight line often defined as the pathway to follow.[43] The external harmonic potential (U) has the following

expression:

$$U = \frac{1}{2}K(RC - RC_0)^2 \tag{2.17}$$

were RC_0 is the reference coordinate point at which the restraint is applied that varies for covering the entire profile, RC is the dynamically changing reaction coordinate that is pushed to RC_0. Thus, the external force can be defined as:

$$F = K(RC_0 + vt - RC) \tag{2.18}$$

being v the constant velocity and t standing for the time.

Not only qualitative information can be extracted from this method, quantitative information such as the thermodynamic potential can also be deduced from the work performed by the system along the reaction coordinate.

The identity derived by Jarzynski 2.19 directly relates the average of an exponential of the total work applied on the system during the simulation to the free energy difference.[44] In order to be able to apply this expression several simulations have to be performed starting from different configurations and then, its averaged is calculated.

$$\langle exp\,[-W/k_BT]\rangle = exp\,[-\Delta G/k_BT] \tag{2.19}$$

where k_B is the Boltzmann constant, W stands for the work performed, T represents the temperature and ΔG is the potential of mean force, that is the free energy as a function of a coordinate. [44]

2.2.2 Umbrella sampling

Umbrella sampling (US) is the first developed enhanced sampling collective variable (CV) approach whose main objective is to overcome the potential barrier allowing to sample inaccessible configurations. The collective variable (also denominated as reaction coordinate) is a continuous parameter that represents the progress along the reaction pathway.

It consists of adding a bias or an additional energy term as in the previous method, which avoids having non-populated regions along the reaction coordinate. This is a method that samples the system in equilibrium by splitting the path into windows, each of them covering a small region of the reaction

coordinate. One simulation is run in every window separately, nevertheless, it is required that the nearby window distributions overlap.[45]

The bias potential applied in each window is defined as $\omega_i(\xi)$, which is usually described by a harmonic potential as written in equation 2.21. Regarding the elements found in the equation, k is the force constant parameter that maintains the system, whose dynamical reaction coordinate value is described by ξ, in the vicinities of the coordinate of the window in which the simulation is run, ξ_0. The force constant is extremely important in this simulations as it has to be large enough to allow the system to go across high energy regions while not too high, as it can originate tight probability distributions of the windows avoiding the overlap among them. However, a good strategy that is adopted whenever the reaction under study requires the force constant to be high, is to simulate a larger amount of windows in order to achieve the overlap among their distributions.

Once the simulation is completed acquiring the unbiased free energy differences (ΔG) is the main objective of the calculation. This free-energy difference or Gibbs free energy is related to the Helmholtz free energy (ΔA) by the canonical partition function, resulting to be equivalent in the most relevant applications. In the NPT ensemble where the pressure, temperature and number of particles are constant the free energy calculated corresponds to the Gibbs free energy, while in the NVT one, where the number of particles, volume and temperature are conserved as in the performed simulations, is the Helmholtz free energy the computed one.

For the calculation of the Helmholtz energy ($A^{unbias}(\xi)$),the unbiased probability distribution is needed (2.22). The probability distribution terms are the normalized variables that measure the frequency of a system being around certain value of the reaction coordinate. The unbiased probability distribution is related to the biased one by equation 2.23. This biased probability distribution is extracted from the US simulation, and with the force expression 2.24, the final equation for the Helmholtz free energy ($A^{unbias}(\xi)$) is equation 2.25. [46, 45]

$$V^{\,bias}(r) = V(r) + \omega_i(\xi) \tag{2.20}$$

$$\omega_i(\xi) = \frac{1}{2}\, k\, (\xi - \xi_0)^2 \tag{2.21}$$

$$A_i^{\,unbias}(\xi) = -(1/\beta)\; ln\; P_i^{\,unbias}(\xi) \tag{2.22}$$

$$P_i^{\,unbias} = P_i^{\,bias}\; e^{\,\beta\; \omega_i(\xi)}\; \langle e^{-\beta\; \omega_i(\xi)} \rangle) \tag{2.23}$$

$$F_i = -(1/\beta)\; ln\; \langle e^{-\beta\; \omega_i(\xi)} \rangle) \tag{2.24}$$

$$A_i^{\,unbias}(\xi) = -(1/\beta)\; ln\; P_i^{\,bias}(\xi) - \omega_i(\xi) + F_i \tag{2.25}$$

where $\beta = 1/k_B T$ and i stands for the number of windows.

2.2.3 WHAM

The Weighed Histogram Analysis Method (WHAM) is a widely utilized tool for reconstructing the potential of mean force (PMF) along the reaction coordinate from the simulated windows. Thus, it is presented as an extension of the Umbrella Sampling approach, it is utilized for the calculation of F and $P^{\,unbias}$ in the equation 2.25. This method reduces to the maximum the statistical error of the unbiased probability distribution ($P^{\,unbias}(\xi)$).

The unbias probability distribution is obtained as the sum of the distributions of the windows multiplied by its weights, which are chosen for having the minimum error with the condition of the sum of all the window weights has to be equal to 1. The values of $P^{\,unbias}$ and F_i are solved by following a self consistent procedure using the equations below.[47, 45]

$$P^{\,unbias}(\xi) = \sum_i^N p_i(\xi)\; P_i^{\,unbias}(\xi) \tag{2.26}$$

$$p_i = \frac{a_i}{\sum_j a_j}; \quad a_i = N_i\; e^{-\beta\; \omega_i(\xi) + \beta\; F_i} \tag{2.27}$$

$$e^{-\beta\; F_i} = \int P^{\,unbias}(\xi)\; e^{-\beta\; \omega_i(\xi)} \delta\xi \tag{2.28}$$

Chapter 3

Results and discussion

3.1 Computational details

For the aim of simulating the carboplatin permeation throughout a DOPC lipid bilayer, and in order to compute its free energy profile, the umbrella sampling technique was used, which has been successfully employed to study the diffusion process of a wide variety of molecules across biological membranes.[48, 49]

Setup of the system

Initially, the setup of the system was performed starting from the construction of the membrane, for which the program CHARMM-GUI [50] was employed. The bilayer is composed of 128 DOPC lipids, 64 of them belong to the upper leaflet and the other 64 ones to the lower leaflet, constituting an 8 x 8 grid. The membrane was placed in a rectangular box that is reproduced periodically along the X, Y and Z dimensions (PBC) as explained in 2.1.4. Furthermore, in order to mimic cell conditions, the membrane was solvated including 37 water molecules per lipid along with an ion concentration of 0.15 M of KCl.

The dynamic calculations in this work were performed using the AMBER18 suite [51] along with the pmemd CUDA implementation [52], where the employed force field for lipids was Lipid 14[53], TIP3P [54] for water along with the Joung/Cheatham ion parameters [55], while for the carboplatin a in-house parametrized force field was selected, whose values can be found in table 3.1. The procedure of the parametrization of this force field consists on extracting the bond and angle parameters from the Cartesian Hessian matrix through the Seminario method employing the MCPB.py module of AmberTools19. However, the dihedral potentials were disregarded because metal-ligand torsion barriers are

usually below the thermal energy. The Lennard-Jones parameters along with the atomic charges were obtained from previous simulations.

TABLE 3.1: Carboplatin parameters of the in-house computed force field.

	Bond Parameters		
Bond	k_b (kcal/mol Å^2)	R_0 (Å)	
Pt-O1	176.8	1.9976	
Pt-O2	176.9	1.9957	
Pt-N1	104.5	2.1158	
Pt-N2	104.0	2.1171	
	Angle Parameters		
Angle	k_a (kcal/mol rad^2)	θ_0 (degrees)	
Pt-O1-C	173.16	122.77	
Pt-O2-C	173.14	122.57	
Pt-N1-H	60.01	110.39	
Pt-N2-H	60.18	110.39	
O2-Pt-O1	180.70	96.23	
N1-Pt-O1	181.64	175.50	
N1-Pt-O2	184.84	79.78	
N2-Pt-O1	188.29	79.67	
N2-Pt-O2	178.37	175.30	
N2-Pt-N1	189.05	104.22	
	Intermolecular Parameters		
Atom	q (a.u.)	$R_{min}/2$ (Å)	ϵ (kcal/mol)
Pt	0.152798	1.2660	0.0030764200
O1	-0.497677	1.6612	0.2100
O2	-0.497677	1.6612	0.2100
N1	-0.666393	1.8240	0.1700
N2	-0.666393	1.8240	0.1700

Equilibration of the DOPC membrane

The bilayer structure was equilibrated by performing a series of MD simulations consisting on a minimization, two heatings and a production step. The minimization step was run for 10000 cycles, the first 5000 steps of them using the steepest descent method, while for the other 5000 the conjugate gradient

method was employed. Furthermore, the initial heating calculation was performed in the canonical ensemble (NVT), applying the Langevin thermostat with a collision frequency of 1 ps $^{-1}$ along 2500 steps with a time step of 0.002 ps, in which the temperature was firstly raised from 0 to an intermediate one of 100 K. Then, a second heating calculation of 50000 steps with a time step of 0.002 ps was computed till reaching the target temperature of 303 K, controlled by the Langevin thermostat, but with constant pressure of 1 bar applying the Berendsen barostat, corresponding to a NPT ensemble. The relative positions of the lipids were restrained by applying a force constant of 10 kcal$\cdot$ mol^{-1} $\cdot$ Å^{-2} along the overall heating process. Afterwards, when the system was already heated and minimized, a production period of 100 ns was run maintaining similar conditions to the aforementioned ones. During all these steps the Coulomb interactions where calculated using the Particle Mesh Ewald method [56]. A cutoff of 10.0 Å for the PME and for the van der Waals interactions was chosen. When considering the hydrogen atoms, their bond distances were restrained by applying the SHAKE algorithm [57].

Umbrella sampling

For performing the umbrella sampling calculations: the carboplatin molecule was initially placed at around 10 Å above the surface of the bilayer, which corresponds to 32.00 Å from the centre of mass of the membrane, in the water phase. The reaction coordinate for the umbrella sampling simulation was defined as the distance between the center of mass of the carboplatin and the centre of mass of the membrane along the z axis that is normal to the membrane as shown in figure 3.1. This reaction coordinate was initially divided into 65 windows with a separation of 0.50 Å. However, these parameters were altered in further simulations as explained below.

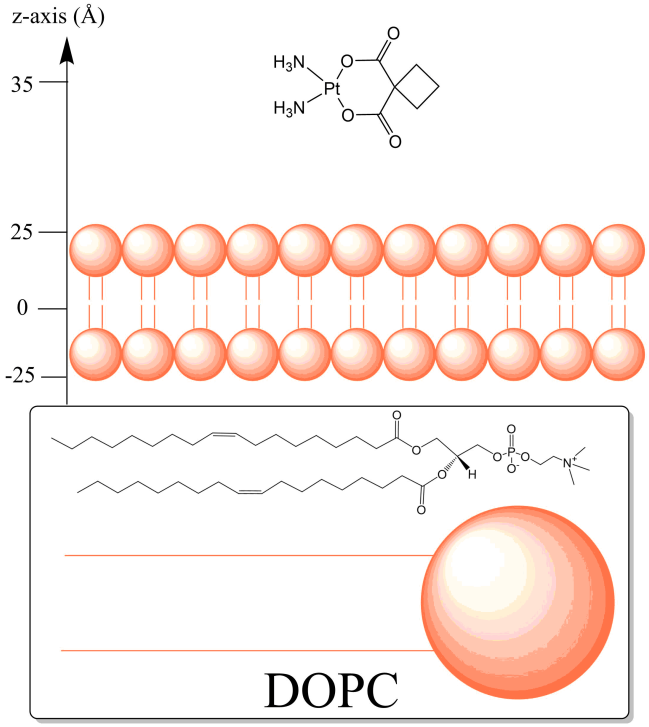

FIGURE 3.1: Schematic representation of the initial placement of the carboplatin with respect to the membrane, displaying the z vector where the reaction coordinate values are written.

Once the carboplatin was placed at 32.00 Å from the bilayer, a new equilibration stage, as the one defined above, was evolved: minimization, heating in two steps and a short equilibration including the carboplatin molecule.

For the sake of efficiency, the following procedure was then implemented: first a steered MD step of 32 ns was performed in which carboplatin is pulled towards the membrane with a force constant for the bias potential of 1.1 kcal· mol^{-1} · $Å^{-2}$. From this pull, a series of snapshots were extracted to initiate each of the window simulations of the umbrella sampling dynamics, and taking in each case the geometry that is closest to the center of the corresponding window as the initial geometry for the simulation. Then, a bias harmonic potential with a force constant of 2.5 kcal· mol^{-1} · $Å^{-2}$ was applied on each of the windows till completing a simulation time of 15 ns in the NPT ensemble with analogous technical parameters to the production of the membrane stabilization. This simulation time should be long enough to ensure a good convergence in the computed profiles.

In order to obtain a complete profile of the permeation of carboplatin through the lipid membrane, the same procedure mentioned above was performed considering the opposite leaflet of the lipid membrane. This accounts for the negative part of the reaction coordinate (specifically, the interval [-32,0] Å), which results in a total of 129 windows for the overall permeation process.

Once all the windows simulations were completed, the WHAM approach previously explained in section 2.2.3 was utilized for the computation of the free-energy profile.

Steered molecular dynamics

The steered MD simulations were employed to obtain starting snapshots for subsequent umbrella sampling trajectories, as explained above, but also in a separately manner to compute the free energy profile from them. This method was applied as follows: two pulls were performed, one starting from 32.00 Å to the center of the lipid membrane (0.00 Å) and the other one starting from -32.00 Å, to sample both leaflets of the membrane. Each pull was run for 50 ns, so that for each energy profile an overall 100 ns simulation was performed at a velocity of 0.64 Å$\cdot$ ns^{-1} , maintaining the same technical parameters as for the production simulations explained above.

The computation of the potential of mean force (PMF) or free energy profile was carried out according to Jarzynski's equality 2.19 as explained in chapter 2.2.1, in which the average of the external work of an ensemble of slow simulations is equal to the free energy difference.

3.2 Umbrella sampling 288, 303, 318 K

This biological process was studied at three different temperatures; 288, 303 and 318 K, which would provide us the possibility of gaining insights about the thermodynamic properties governing the dynamics of the system. Specifically, the enthalpic and entropic terms could thus be obtained, being able to analyse the contributions of each of them. As far as enthalpy is concerned, the interactions of carboplatin with the bilayer are fundamental, while regarding entropy, it is the desolvation of carboplatin when reaching the membrane along with the order of the membrane the most important factors.[58, 59]

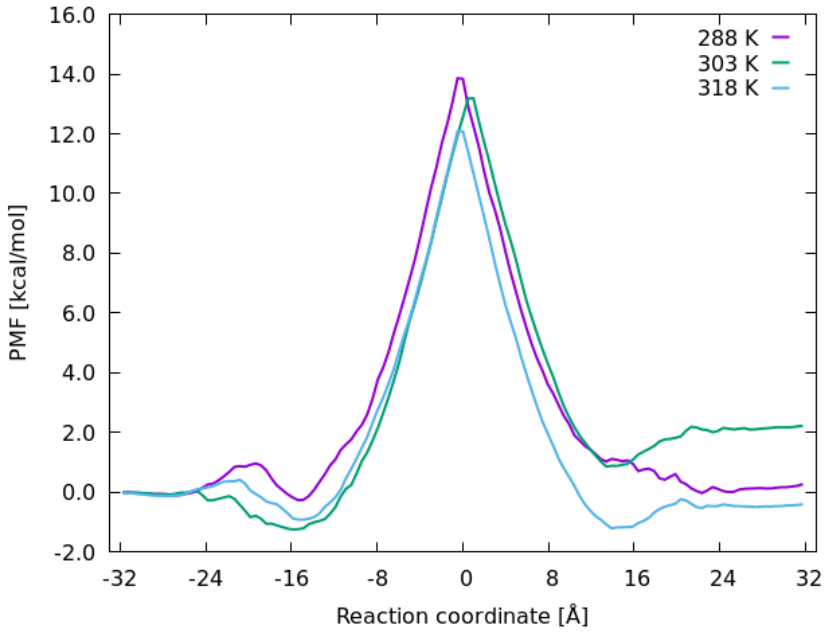

FIGURE 3.2: PMF profiles at three different temperatures; 288 K, 303 K and 318 K.

The calculated PMF profile at the three temperatures are set to 0 kcal/mol at the -32.00 Å reaction coordinate value in all of them for the sake of comparison. When observing the profiles together in figure 3.2, the most easily appreciated detail is the asymmetry presented by all of them. This was not expected considering that the membrane is completely symmetric and the simulation procedures performed have been analogous at each side of the bilayer. Furthermore, the most relevant fact is the incoherent variation of the energy barrier with the temperature: the lowest value belongs to the 318 K temperature, the intermediate one corresponded to the 288 K, and the highest barrier is found at 303 K. It has to be considered that the free energy should vary consistently with the temperature, either increasing or decreasing, but following the same trend.

This inconsistency lead us to think that the initial simulation time of each window could have affected the overall profile as the system might require some time to adapt itself to the reaction coordinate in which it is run. In the subsequent figures 3.3, 3.4 and 3.5 the free energy profiles achieved at 288 K, 303 K and 318 K are respectively represented. In each case, the convergence of the free energy profile is studied by truncating initial intervals of simulation ranging from 0 to 5 ns. This is done in order to determine the initial time that should be considered as the equilibration time and, thus, be disregarded from the simulation to compute the PMF.

Thus, three different profiles are represented in which the total simulation time has been considered (from 0 to 15 ns) and where different initial time intervals have been disregarded for the construction of the profiles. When observing the different profiles of the three figures, it is clearly appreciated how the one for which the total simulation time (0-15 ns) is evaluated presents a different trend along with the one that includes the second nanosecond (1-15 ns). In addition, at 303 K the PMF computed from 2 to 15 ns shows a still not converged profile, therefore, the first three nanoseconds of every window were disregarded in further analysis.

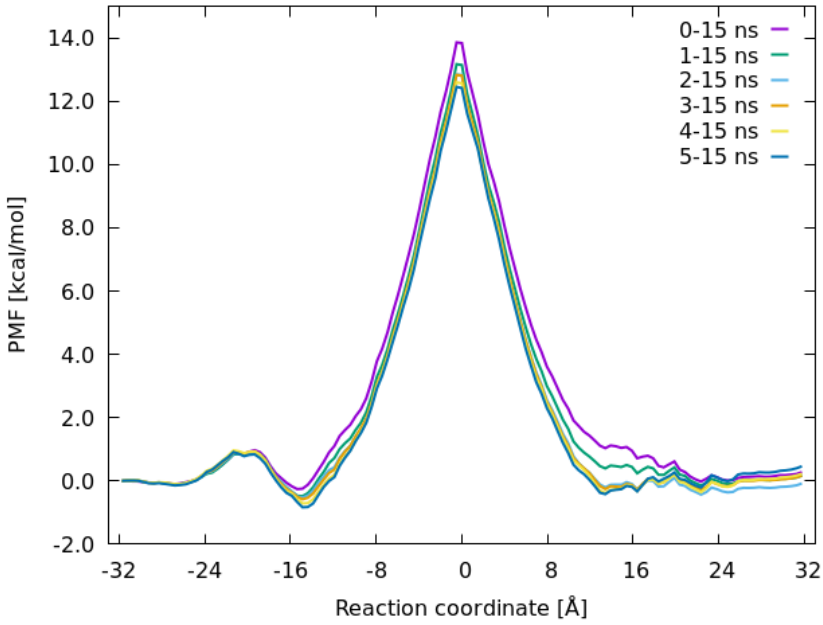

FIGURE 3.3: Convergence analysis of the potential of mean force (PMF) profile at 288 K by disregarding the initial nanoseconds.

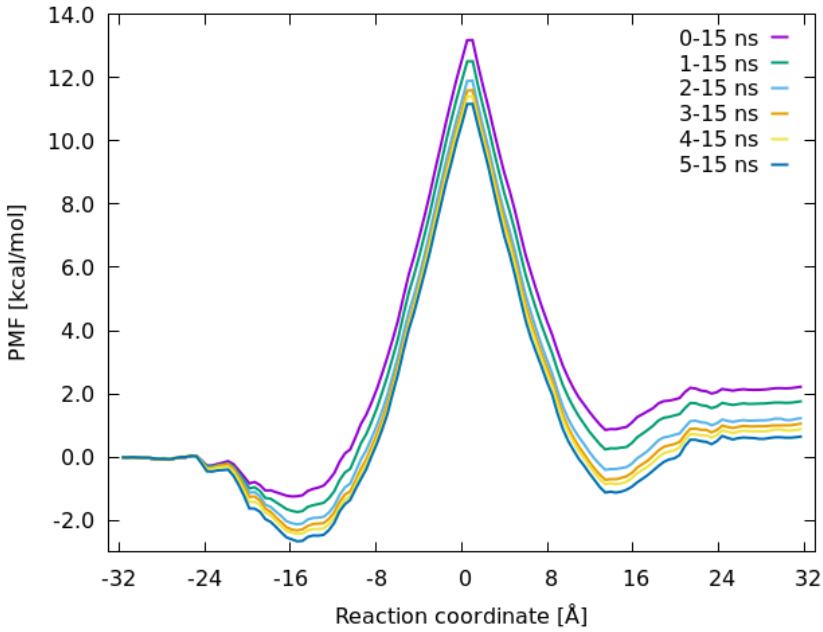

FIGURE 3.4: Convergence analysis of the potential of mean force (PMF) profile at 303 K by disregarding the initial nanoseconds.

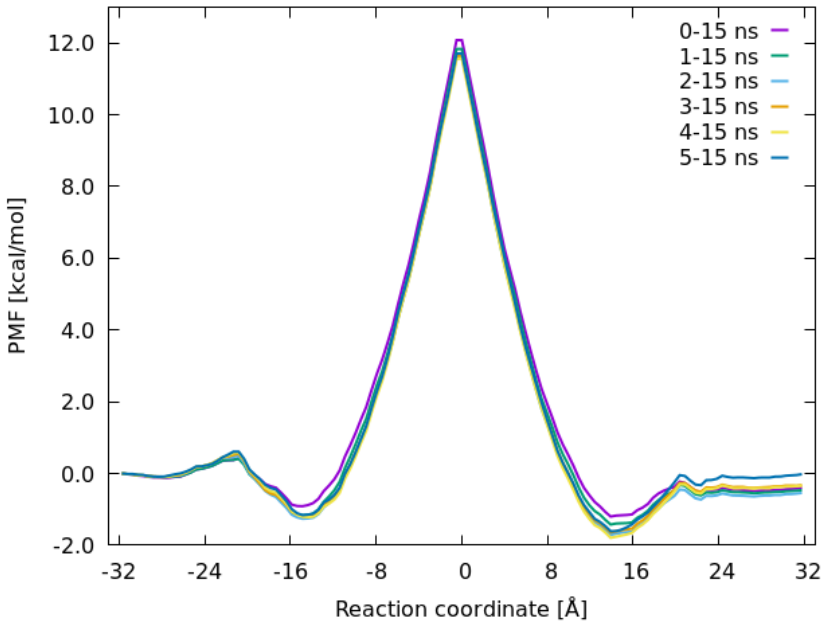

FIGURE 3.5: Convergence analysis of the potential of mean force (PMF) profile at 318 K by disregarding the initial nanoseconds.

Once this equilibration period is removed for the free energy profile computation, the three temperatures are again represented in figure 3.6. Nevertheless, as far as the energy variation with temperature is concerned, all reaction barriers are found to be in between 13 and 14 kcal/mol. They exhibited no clear trends with the temperature, which is in partial disagreement with chemical intuition

and may arise as a result of the limited accuracy of the model. In addition, the free-energy profiles are still not fully symmetric, especially the one computed at 303 K, indicating the presence of artifacts or mistakes in the dynamics protocol. Thus, the following analyses will focus on the computation of the profile at the intermediate temperature of 303 K, that is, the one with the highest asymmetry and that corresponds to a reasonable temperature for standard biological systems. The goal of the following analyses is to determine the factors that influence the shape of the PMF profile, and to obtain appropriate simulations parameters for the system under investigation.

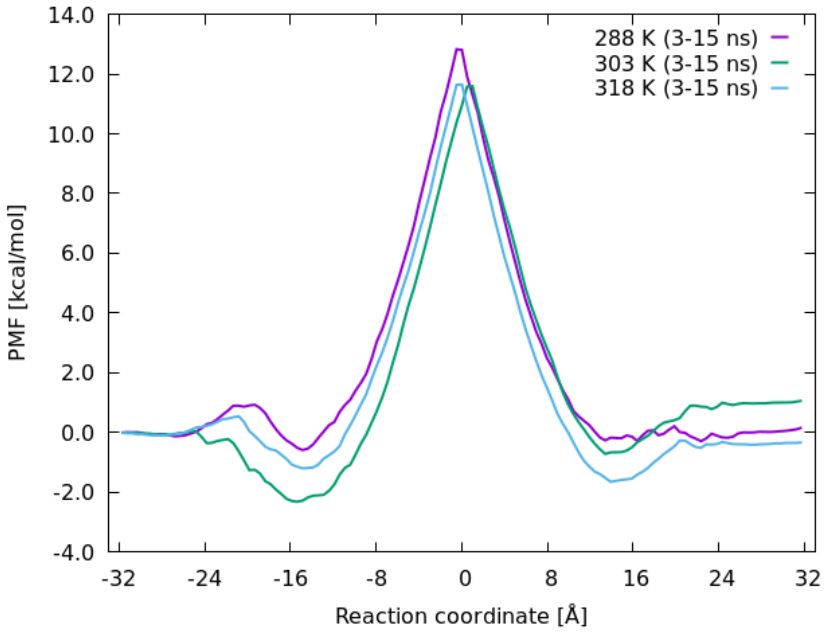

FIGURE 3.6: PMF profiles at three different temperatures; 288 K, 303 K and 318 K.

Due to the asymmetry of the obtained energy profiles, specially at 303 K, with respect to the centre of the membrane, the histograms that represent the probability distribution of the reaction coordinate at 303 K were thoroughly analyzed. The central range of the histograms is shown in figure 3.7. In all the windows the probability distribution suffered from certain displacement from the desired value, that is, the centre of the probability distribution was displaced with respect to the central value of the window, indicated with the blue vertical lines in the plot, where the peak of the distributions should be found.

Although, these deviations are generally encountered for all of them, it is worth remarking the gap found between 0.50 and 1.0 Å, which may be responsible for the more elevated energies on the positive reaction coordinates side

with respect to the negative ones in figure 3.6.

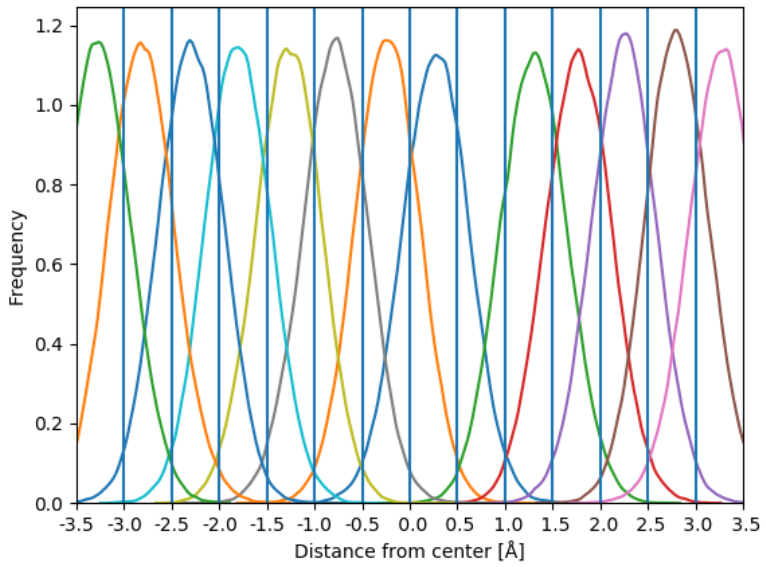

FIGURE 3.7: Histogram of the central windows from -3.00 Å to 3.00 Å at a temperature of 303 K. Vertical lines are added where the centre of the probability distribution of each window should be placed.

In addition to the deviations presented as in figure 3.7 for the central windows, the reaction coordinate deviation from the desired value of every window is shown in figure 3.8. It was computed as the average of the difference between the z value of each time step of the simulation and the central value of the reaction coordinate of the window. Once again, figure (3.8) reveals certain asymmetry in the deviations: for the negative values of the reaction coordinate, errors higher than 0.30 Å are found for the windows located between -8.00 Å and 0.00 Å, while the positive reaction coordinates windows present the highest deviations more localized, close to the centre of the bilayer. Nevertheless, the most relevant feature of the graph is found in the size of the deviation of every window, which is larger than 0.26 Å in all of them, exceeding the 0.40 Å in some central windows. Considering that the separation distance between windows is 0.50 Å , those deviations values are excessive, leading us to conclude that the force constant of the bias potential applied to keep the system in each window during the simulations was not high enough.

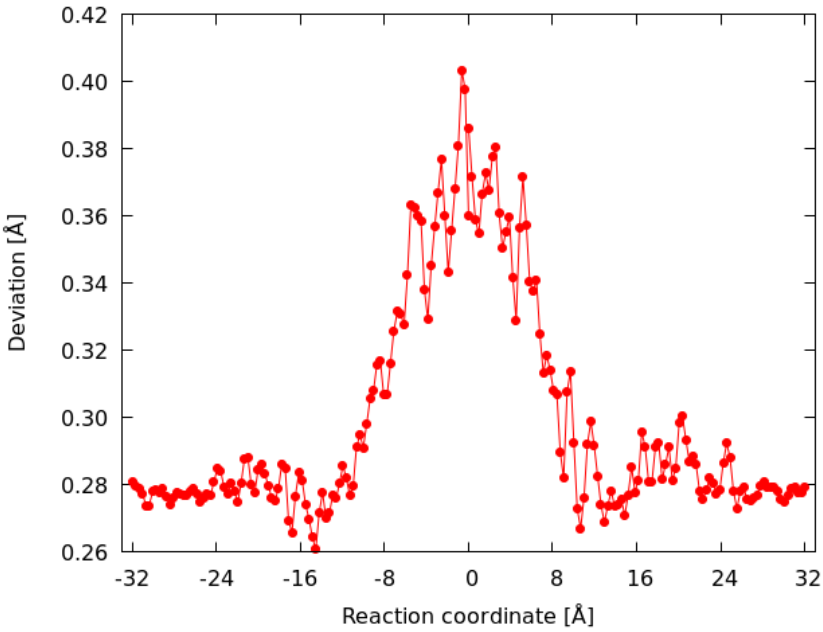

FIGURE 3.8: Average deviation (Å) of every window computed as the average of the difference between the z value of the step and the central value of the reaction coordinate of the window.

3.3 Selecting the force constant for each window

Due to the strong deviations suffered in the simulated windows, a different approach was attempted: an individual modification of the force constant for each window was determined according to a series of short simulations. By performing a one nanosecond simulation on each window, the sampled reaction coordinate values along the window simulation were extracted and their difference with respect to the central value of their respective window was calculated. The average of this differences was obtained for each window, and was considered to be acceptable under 0.10 Å in absolute value. One has to keep in mind that the separation between windows is 0.50 Å, therefore the maximum acceptable error is 20% of the size window.

Once the first umbrella sampling simulation was run covering all the windows, if the resultant average fluctuation was larger than the threshold established of 0.10 Åfor a particular window, the force constant value is raised in 2.5 kcal· mol^{-1} · $Å^{-2}$, and a one ns simulation is repeated with the new force constant value. Contrary, if the average deviation for a window is lower than the threshold, the calculation for that window stopped, saving the last applied force

constant as the correct one for that individual window. This process is repeated until the error of every window is below 0.10 Å. This procedure was initialized with the same force constant value that was applied in previous analysis, 2.5 kcal· mol^{-1} · $Å^{-2}$, which corresponds to iteration number one in figure 3.9. As can be observed, nine iterations were required, appreciating that already at iteration number eight the vast majority of the windows had reached the threshold and, thus, stopped. As a result, the optimum value of the force constant for most of the windows was found to be 20.0 kcal· mol^{-1} · $Å^{-2}$.

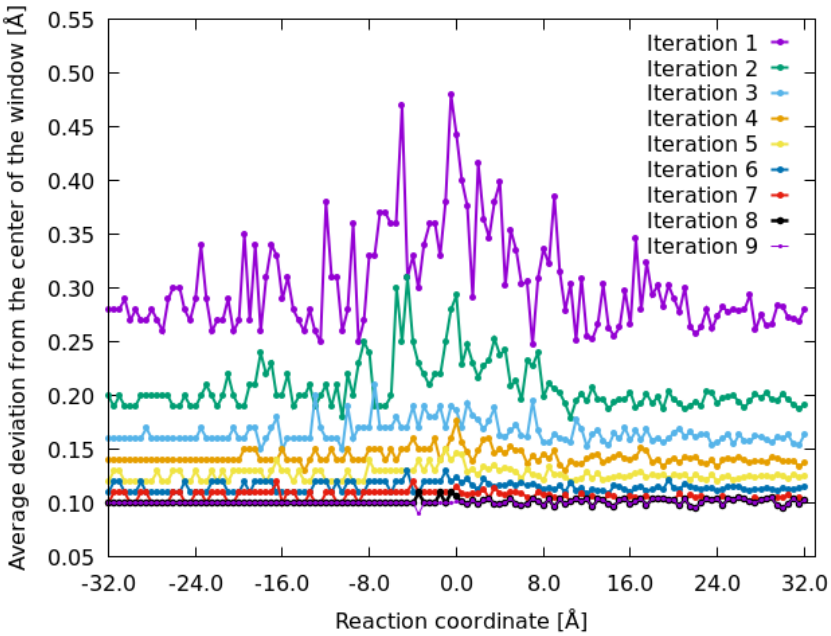

FIGURE 3.9: Average absolute deviations of one ns simulations along the different iterations performed.

Nevertheless, when performing once again the umbrella sampling for 15 ns in each window with the new value of the force constant, the resultant free energy profile represented in figure 3.10 did not improve, in partial disagreement with what was expected. The most noticeable abnormal alteration is the raising of the energy barrier to extremely high values, when this biological process is known to happen. [60] In addition, the PMF is also found to be asymmetric in almost 2 kcal· mol^{-1} when approaching to the barrier from the negative reaction coordinate values or from the positive ones. Moreover, there was not a trace of the expected minima around -15.00 Å and 15.00 Å , where the polar heads of the lipid bilayer are placed, as it occurred for the cisplatin molecule simulations performed in [22].

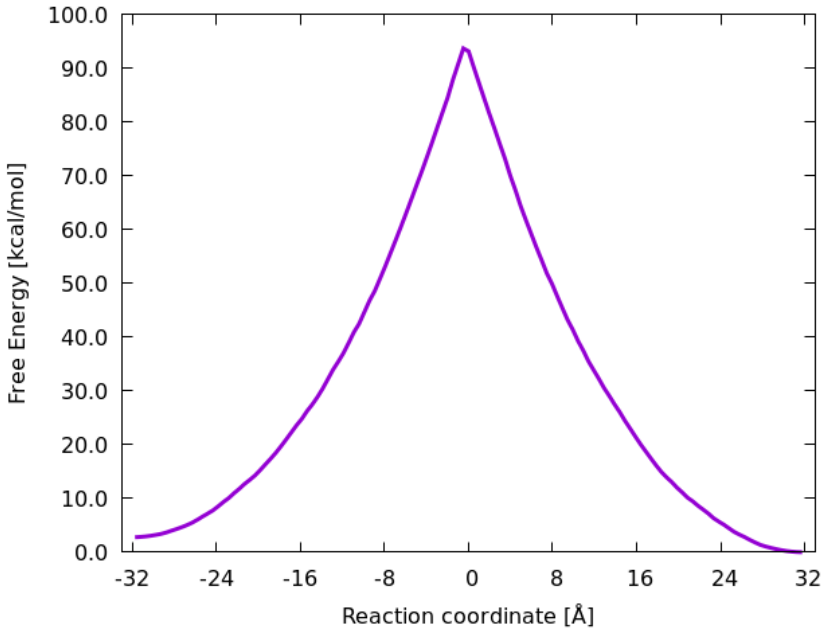

FIGURE 3.10: Free energy profile of the umbrella sampling at 303 K applying the individually calculated force constant for each window.

Analysing the possible causes of the excessively high energy values encountered, one of the most relevant aspects of the umbrella sampling technique, the overlap among windows in the vicinities, was found to be insufficient (figure 3.11). The considerably increase of the force constant in every window corrected the displacement, observing the probability distributions perfectly centred on their respective reaction coordinate value. However, this modifications on the force constants were translated into very tight probability distributions that might not have correctly sampled the entire pathway of the reaction coordinate. In fact, the overlap between neighbor windows observed in figure 3.11 is very small.

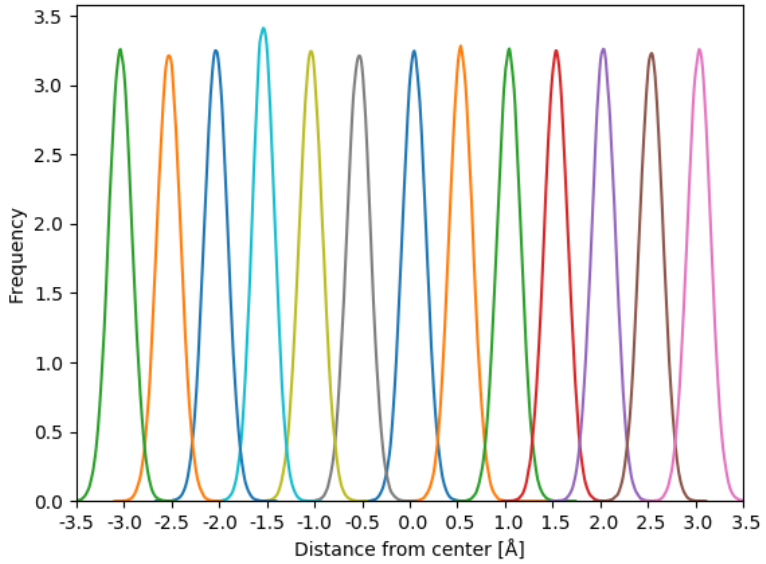

FIGURE 3.11: Histogram of the central values (from -3.00 Å to 3.00 Å) of the umbrella sampling at 303 K applying the individually calculated force constant for each window.

3.4 Selecting the number of windows

In order to evidence whether the abnormal free energy profile was due to a poor overlap among windows, the next logical step was to increase the number of windows to gain overlap among them. Whereas the previous calculation was composed of 129 windows dividing the overall pathway every 0.50 Å, the amount of windows was raised to 257 dividing the sampled space every 0.25 Å , and the window simulations were run for one nanosecond each. The reaction coordinate probability distributions for the new simulations are shown in figure 3.12.

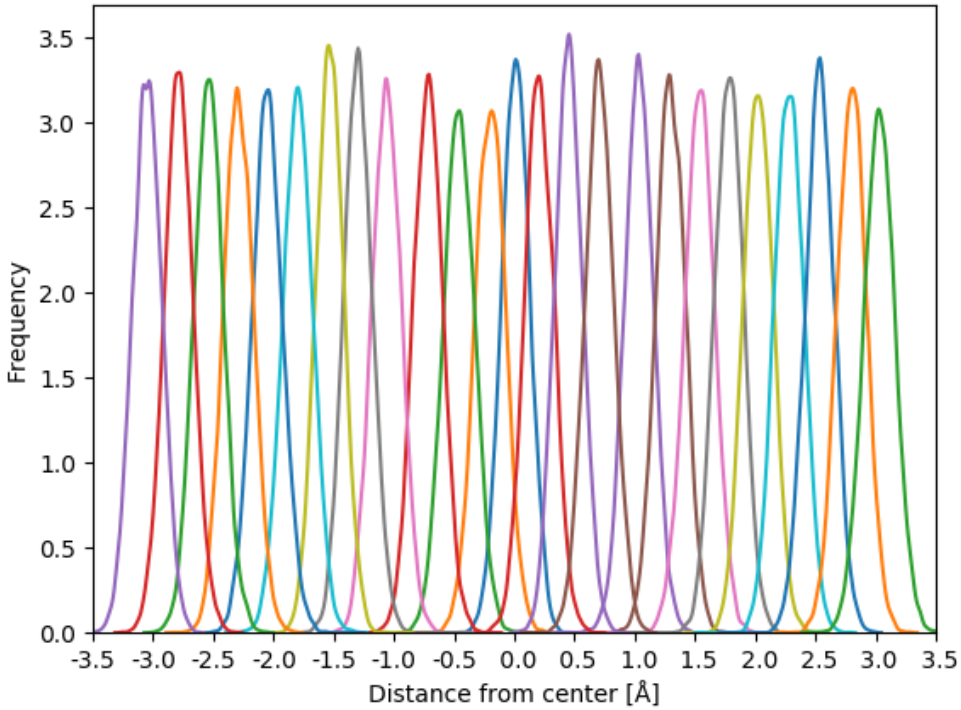

FIGURE 3.12: Histogram of the central values (from -3.00 Å to 3.00 Å) of the umbrella sampling at 303 K applying the individually calculated force constant for each window increasing by almost double the number of windows.

As it is appreciated in 3.12 the achieved overlap with a larger number of windows is much better, which seemed to be enough for the correct computation of the PMF profile. The acquired energy barrier decreased and it is in good agreement with previous simulations performed with similar Pt complexes.[61, 22] However, a very asymmetric profile was once again obtained, which is represented in green in figure 3.13, specially at around 0.00 Å. This led us to consider the fact that some deviations of certain windows could have originated those asymmetries providing that strange aspect to the profile.

Therefore, the average deviations of the windows were calculated, discovering that six of them were strongly displaced from the desired reaction coordinate value. In particular, the windows centred at -0.75, -0.50, -0.25, 0.25, 0.50, 0.75 Å were affected by very strong deviations, two of them exceeding 1.45 Å of average deviation. For that reason they were all excluded from the computation of the profile, which then presented the aspect of the purple line in figure 3.13. Although the asymmetry of the PMF profile was reduced and the two expected minima at -15 and 15 Å partially appeared, the shape of the profile did not have still the required symmetry.

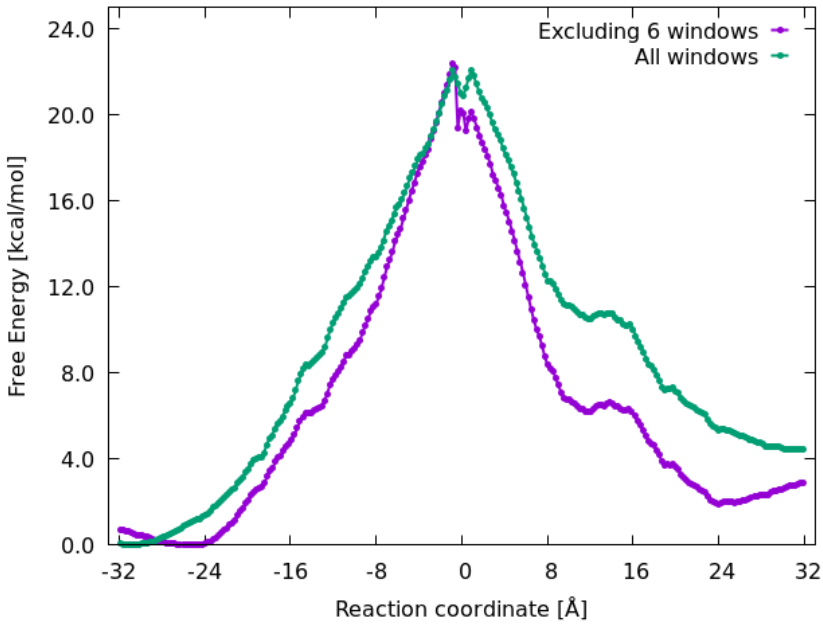

FIGURE 3.13: Free energy profile of the umbrella sampling at 303 K applying the individually calculated force constant for each window increasing by almost double the number of windows including all the windows in the green curve and excluding the six problematic ones in purple.

For further analysis of the strongly abnormal PMF profile provided by the ensemble of 0.25 Å and 0.50 Å window separation simulations, the deviations were represented for all the windows along the reaction coordinate excluding the six outlined windows previously commented (-0.75, -0.50, -0.25, 0.25, 0.50, 0.75 Å). Moreover, in order to analyse in depth the deviations, these were calculated considering its sign and also in absolute value, as done before.

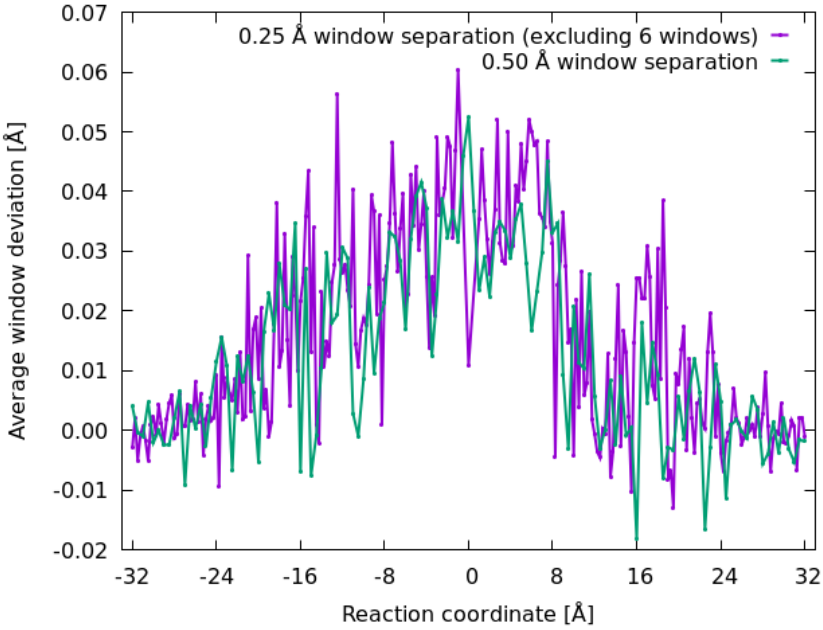

FIGURE 3.14: Average deviations of the one nanosecond simulated windows with a 0.25 Å and 0.50 Å separation between them, applying the individually calculated force constant for each window, excluding the six outlined windows in the 0.25 Åof separation (-0.75, -0.50, -0.25, 0.25, 0.50, 0.75 Å).

Although very strong errors are not observed, a bias for positive values can be seen for both ensembles of data. This may be considered as a possible reason for the profile's asymmetry. Therefore, a new procedure would need to be followed in order to reduce the deviation or to present an equivalent error in both sides of the membrane that compensates each other, providing an overall average error of zero. This could be achieved by further increasing the force constant of the artificial potential and the number of windows in a iterative manner. However, a different strategy explained in the following section was followed.

3.5 Steered molecular dynamics

At this point, the deviation of the windows was one of the most problematic issues, which was solved when the force constant values were significantly raised producing tight probability distributions, causing a poor overlap of the windows and not achieving a correct free energy profile.

A different approach was then attempted, focusing on the steered MD approximation. Specifically, pull steps were performed with different force constant values: 2.5, 5.0, 10.0, 15.0, 20.0, 50.0 and 100.0 kcal· mol^{-1} · $Å^{-2}$. The

work profiles were computed for each steered MD simulations and represented in figure 3.15. As can be seen, the thickness of the lines increases due to the dispersion of the data according to the size of the force constant applied.

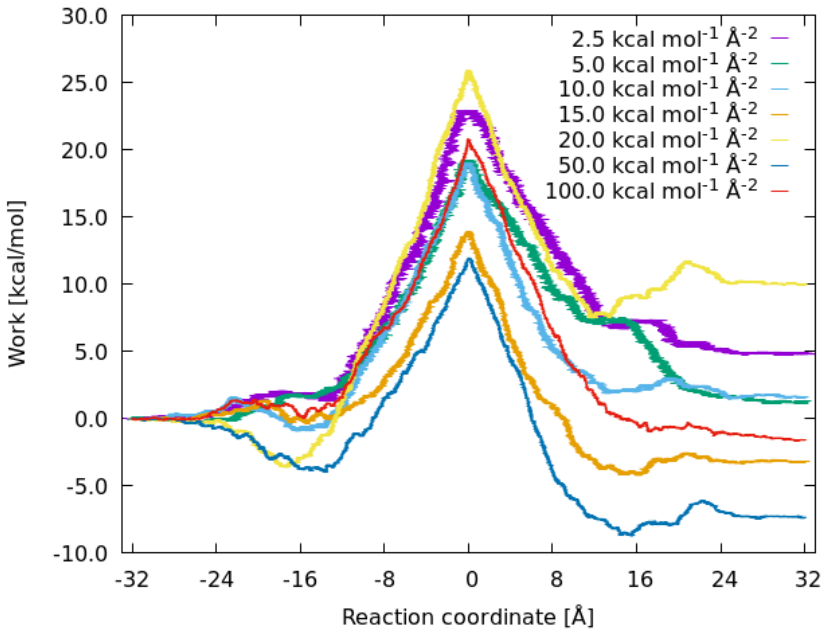

FIGURE 3.15: Work profiles applying a range of force constants from 2.5 to 100.0 kcal· mol^{-1} · $Å^{-2}$.

In order to select the most suitable force constant to carry out further simulations, the average error of the total pull was obtained and plotted in purple in figure 3.16, along with the maximum and minimum errors presented by each of the overall calculations with different force constants. As it was expected, the higher the force constant the smaller the range of deviations, as the potential to maintain the system in the corresponding reaction coordinate is better. Nevertheless, it is found that when applying a force constant of 10.0 and 15.0 kcal· mol^{-1} · $Å^{-2}$, despite their considerable range of deviations, the average error turned out to be close to zero, which is translated in equivalent errors at both sides of the center of the bilayer, a desirable property to get symmetric free-energy profiles. For the 10.0 kcal· mol^{-1} · $Å^{-2}$ force constant the average deviation accomplished was -0.005 Å while for the 15.0 kcal· mol^{-1} · $Å^{-2}$ was 0.016 Å . Therefore, the 10.0 kcal· mol^{-1} · $Å^{-2}$ force constant was the one applied in additional calculations.

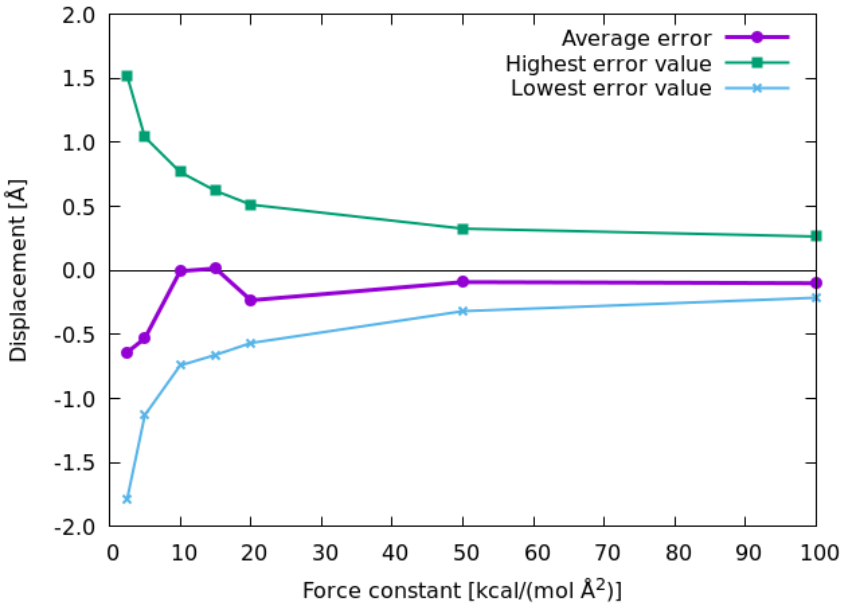

FIGURE 3.16: Average displacement along with the maximum and minimum value in Å found for the different force constants calculations.

In order to compute an accurate free energy profile from steered MD simulations, different trajectories have to be run starting with different initial conditions. Then, the work profiles obtained from the simulations have to be averaged according to Jarzynski equality (as explained in the methods section 2.19) to compute the PMF. Thus, ten different steered MD were run with a force constant of 10.0 kcal· mol^{-1} · $Å^{-2}$. The initial geometries were selected from the previous simulations and the initial velocities were generated from a Boltzmann distribution at 303 K.

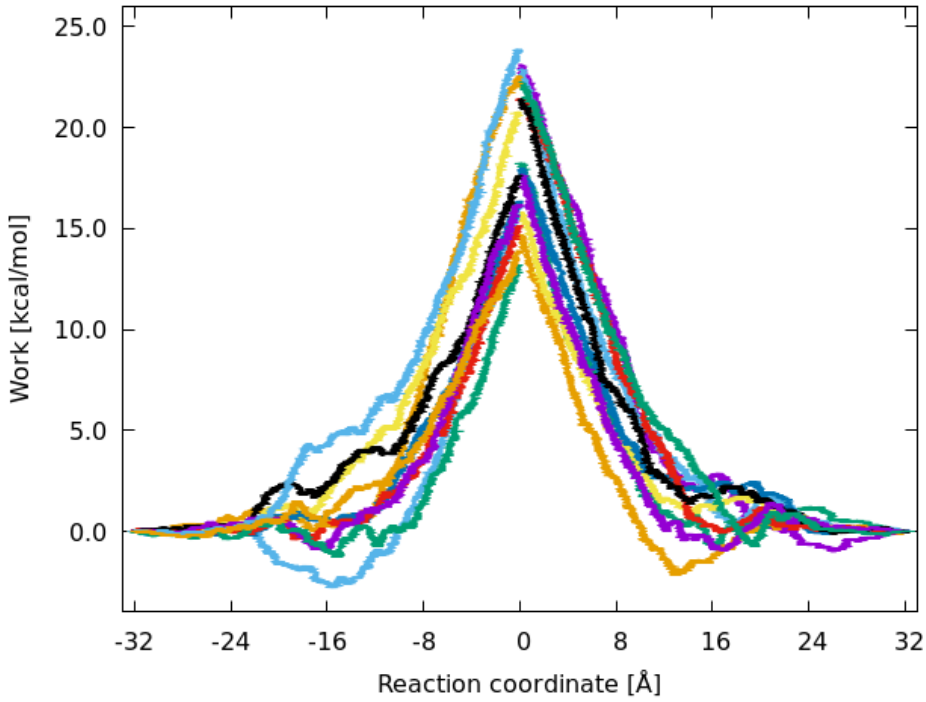

FIGURE 3.17: Work profiles of the ensemble of the ten total pulls performed with different initial geometries.

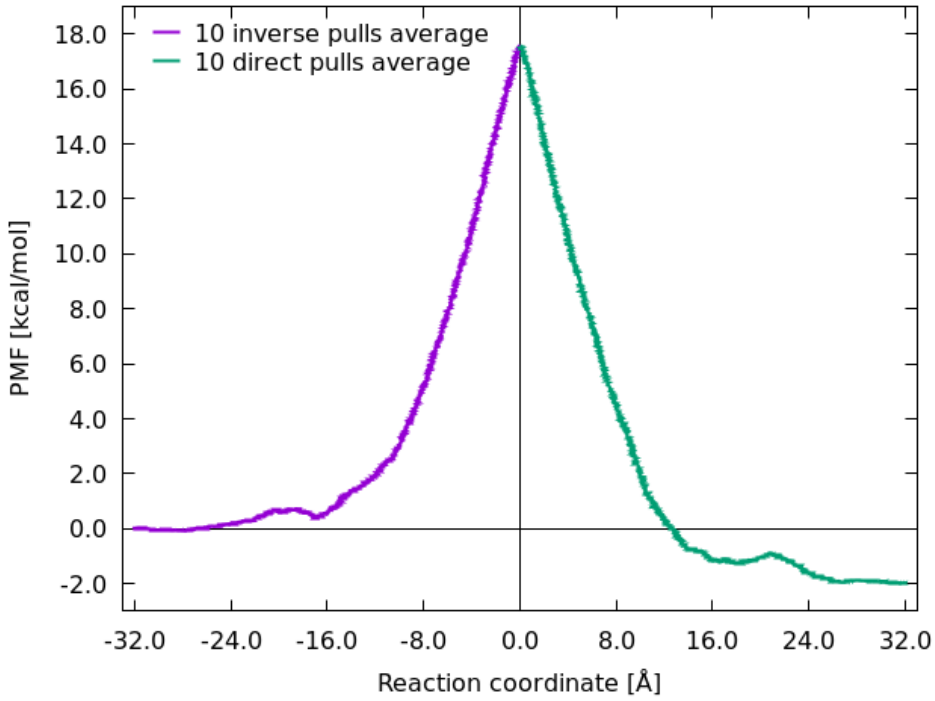

FIGURE 3.18: PMF profile of the average of work from the ten pulls.

The work profiles of the ten completed pulls are plotted in figure 3.17, and the average of all of them was computed acquiring the corresponding free energy profile (PMF) of the process which is plotted in figure 3.18. As it can be appreciated in the plot, the perfect symmetry was not achieved obtaining a difference of almost 2 kcal· mol^{-1} in the energy barrier when going from each side of the bilayer. In particular, the barrier is 17.57 kcal· mol^{-1} when approaching from negative values of the reaction coordinate, and 19.53 kcal· mol^{-1} from positive ones. Although, this error can be considered to be within the expected accuracy of the method, the obtained asymmetry is of similar magnitude as the one obtained in the initial umbrella sampling simulations, whose PMFs are plotted in figure 3.6. Therefore, it seems that this new strategy does not provide better results than the previous attempts. However, it is important to note that the PMF computed from the steered MD simulations presents the two expected minima around 16.00 Å and -16.00 Å , corresponding to the region where the polar heads of the membrane lipids are placed. These minima were found in previous studies of the permeation of the related platinum drug cisplatin.[22] In addition, the absolute maximum corresponding to the energy barrier results to be reasonable as it provides a coherent value for the activation energy of a biological process, such as the carboplatin permeation across a cell membrane, allowing to deduce that the process may occur at 303 K. Further simulations will be necessary to solve the problem of asymmetric free-energy profiles.

Chapter 4

Conclusion

One of the most widely employed platinum chemotherapeutic agents, carboplatin, undergoes a mechanism of action whose details are still not clearly defined at the atomic level. The entrance of this compound into the cells is one of the crucial steps of the overall mechanism, as it is directly related to the drug accumulation, the most important factor of the cytotoxicity of these compounds. The carboplatin uptake is claimed to be produced by diffusion across the membrane as well as by active transport, with the help of certain transport proteins. Nevertheless, not much more can be stated about such an important biological process.

The initial objective of this project was to analyse the permeation mechanism of the drug across the biological membrane in order to gain deeper insights about its driving forces. However, due to the several problems encountered when computing the free energy profile, the aim of this master thesis ended up focusing on achieving a correct PMF profile, and analyzing the factors that influence its shape. The work relied on the application of the umbrella sampling – along with the WHAM approach – and steered MD methodologies for the simulation of the diffusion process and the reconstruction of its free energy profile, resulting in considerable asymmetric profiles.

The effect of the temperature in the permeation process was evaluated leading to non-conclusive results, despite of achieving an apparent full convergence on the free energy profiles at each of the temperatures. Additionally, in order to solve the significant deviations of the reaction coordinate presented in all the windows, a systematic optimization of the ideal value of the force constant of the bias potential was developed. The ideal force constant value determined was 20.0 kcal$\cdot$ mol^{-1} $\cdot$ Å^{-2}, being considerably larger than the initial one of 2.5 kcal$\cdot$ mol^{-1} $\cdot$ Å^{-2}. The use of this optimized bias potential lead to adequate window's placements at the cost of poor overlapping probability distributions

between consecutive windows, which resulted in an extremely high energy barrier. The overlap was improved by doubling the number of windows, which lowered the activation energy but yielded to very asymmetric profile, likely caused by the short simulation time (1 ns) employed and the large reaction coordinate deviations found for some of the windows.

Finally, a different strategy was then attempted employing steered MD instead. The ideal force constant was found to be 10.0 kcal$\cdot$ mol^{-1} $\cdot$ Å^{-2}, which was applied in ten additional steered MD trajectories and in the computation of the free-energy profile according to Jarzynski's equality. This last profile presented a reasonable energy barrier and the expected minima around -16.00 Å and 16.00 Å, however, certain asymmetry was still present.

In summary, similar PFM profiles were achieved using umbrella sampling and steered MD with a reasonable activation barrier and the expected minima in the polar regions of the bilayer. Nevertheless, the desired symmetry of the profile was still not accomplished. As future work, as far as the umbrella sampling is concerned, an increment of the number of windows while increasing the force constant value could lead to more adequate PMF profiles. Regarding steered MD, a noticeable improvement could be achieved by raising the number of work profiles from the 10 computed ones in this work to, for example, 100, whose average provides the resultant free energy profile. The computation of the correct PMF profile of this biological process followed by further analyses could perhaps give us insights about the permeation mechanism of this drug.

References

[1] Saurabh Ram BihariLal Shrivastava, Prateek Saurabh Shrivastava, et al. "Working together to reduce the incidence of cancers in low-and middle-income nations: World Health Organization". In: *Clinical Cancer Investigation Journal* 9.3 (2020), pp. 61–62.

[2] Christopher Barnard. "Platinum group metal compounds in cancer chemotherapy." In: *Johns Matthey Technol Rev* 61 (2017), pp. 52–59.

[3] Timothy C Johnstone, Kogularamanan Suntharalingam, and Stephen J Lippard. "The next generation of platinum drugs: targeted Pt (II) agents, nanoparticle delivery, and Pt (IV) prodrugs". In: *Chemical reviews* 116.5 (2016), pp. 3436–3486.

[4] Michael J. Hannon. "Metal-based anticancer drugs: From a past anchored in platinum chemistry to a post-genomic future of diverse chemistry and biology". In: *Pure and Applied Chemistry* 79.12 (2007), pp. 2243–2261.

[5] Christopher H. Grant and Charlie Gourley. "Chapter 2 - Relevant Cancer Diagnoses, Commonly Used Chemotherapy Agents and Their Biochemical Mechanisms of Action". In: *Cancer Treatment and the Ovary*. Ed. by Richard A. Anderson and Norah Spears. Boston: Academic Press, 2015, pp. 21–33. ISBN: 978-0-12-801591-9.

[6] Ronald P. Miller et al. "Mechanisms of Cisplatin Nephrotoxicity". In: *Toxins* 2.11 (2010), pp. 2490–2518. ISSN: 2072-6651.

[7] Shaloam Dasari and Paul Bernard Tchounwou. "Cisplatin in cancer therapy: Molecular mechanisms of action". In: *European Journal of Pharmacology* 740 (2014), pp. 364–378. ISSN: 0014-2999.

[8] Antona J Wagstaff et al. "Carboplatin". In: *Drugs* 37.2 (1989), pp. 162–190.

[9] Mariana Castells et al. "Chapter 19 - Chemotherapy Allergy". In: *Drug Allergy Testing*. Ed. by David A. Khan and Aleena Banerji. Elsevier, 2018, pp. 197–210. ISBN: 978-0-323-48551-7.

[10] Ahlam A. Ali, Cian M. McCrudden, and Helen O. McCarthy. "Chapter 5 - Evaluation of the Impact of Nitric Oxide on Resistance to Platinum-Based Chemotherapeutics". In: *Nitric Oxide (Donor/Induced) in Chemosensitizing*. Ed. by Benjamin Bonavida. Vol. 1. Cancer Sensitizing Agents for Chemotherapy. Academic Press, 2017, pp. 71–90.

[11] Markus Galanski et al. "Synthesis, crystal structure and cytotoxicity of new oxaliplatin analogues indicating that improvement of anticancer activity is still possible". In: *European journal of medicinal chemistry* 39 (Sept. 2004), pp. 707–714.

[12] David J Stewart. "Mechanisms of resistance to cisplatin and carboplatin". In: *Critical reviews in oncology/hematology* 63.1 (2007), pp. 12–31.

[13] Stephen B. Howell et al. "Copper Transporters and the Cellular Pharmacology of the Platinum-Containing Cancer Drugs". In: *Molecular Pharmacology* 77.6 (2010), pp. 887–894.

[14] DP Gately and SB Howell. "Cellular accumulation of the anticancer agent cisplatin: a review". In: *British journal of cancer* 67.6 (1993), pp. 1171–1176.

[15] Dong Wang and Stephen J Lippard. "Cellular processing of platinum anticancer drugs". In: *Nature reviews Drug discovery* 4.4 (2005), pp. 307–320.

[16] Cara A Rabik and M Eileen Dolan. "Molecular mechanisms of resistance and toxicity associated with platinating agents". In: *Cancer treatment reviews* 33.1 (2007), pp. 9–23.

[17] Shuzhong Zhang et al. "Organic cation transporters are determinants of oxaliplatin cytotoxicity". In: *Cancer Research* 66.17 (2006), pp. 8847–8857.

[18] Aleksandra Mandic et al. "Cisplatin induces endoplasmic reticulum stress and nucleus-independent apoptotic signaling". In: *Journal of Biological Chemistry* 278.11 (2003), pp. 9100–9106.

[19] David G Breckenridge et al. "Regulation of apoptosis by endoplasmic reticulum pathways". In: *Oncogene* 22.53 (2003), pp. 8608–8618.

[20] Wojciech Szlasa et al. "Lipid composition of the cancer cell membrane". In: *Journal of Bioenergetics and Biomembranes* (2020), pp. 1–22.

[21] Xiaoyong Wang and Zijian Guo. "Targeting and delivery of platinum-based anticancer drugs". In: *Chem. Soc. Rev.* 42 (1 2013), pp. 202–224.

[22] Lorena Ruano, Gustavo Cárdenas, and Juan Jose Nogueira. "The Permeation Mechanism of Cisplatin through a Dioleoylphosphocholine Bilayer". In: *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry* (2021).

[23] Outi Salo-Ahen et al. "Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development". In: *Processes* 9 (Dec. 2020), p. 71.

[24] Ju Li. "Basic Molecular Dynamics". In: *Handbook of Materials Modeling: Methods*. Ed. by Sidney Yip. Dordrecht: Springer Netherlands, 2005, pp. 565–588.

[25] Outi Salo-Ahen et al. "Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development". In: *Processes* 9 (Dec. 2020), p. 71.

[26] Marco Rozgic. "Integration Methods,Thermostats and Barostats in Molecular Dynamics". In: (Sept. 2015).

[27] Errol Lewars. *Computational chemistry*. Ontario, Canada: Springer, 2003, p. 318.

[28] Godehard Sutmann. "Classical Molecular Dynamics". In: *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms* 10 (2002), pp. 211–254.

[29] Juan José Nogueira. *Molecular Dynamics:Classical Equations of Motion*. 2020. URL: https://youtu.be/8iHER6IP6Ds.

[30] Ralf Schneider, Amit Raj Sharma, and Abha Rai. "Introduction to molecular dynamics". In: *Computational Many-Particle Physics*. Springer, 2008, pp. 3–40.

[31] Juan José Nogueira. *Molecular Dynamics:Potential Energy and Force Fields*. 2020. URL: https://youtu.be/vpqo0q3zvJU.

[32] Hendrik Heinz et al. "Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: the INTERFACE force field". In: *Langmuir* 29.6 (2013), pp. 1754–1765.

[33] Xin Yong and Lucy T. Zhang. "Thermostats and thermostat strategies for molecular dynamics simulations of nanofluidics". In: *The Journal of Chemical Physics* 138.8 (2013), p. 084503.

[34] Giovanni Bussi, Davide Donadio, and Michele Parrinello. "Canonical sampling through velocity rescaling". In: *The Journal of Chemical Physics* 126.1 (2007), p. 014101.

[35] H. J. C. Berendsen et al. "Molecular dynamics with coupling to an external bath". In: *The Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690.

[36] Carlos Braga and Karl P. Travis. "A configurational temperature Nosé-Hoover thermostat". In: *The Journal of Chemical Physics* 123.13 (2005), p. 134101.

[37] Ronak Patel and Petety Balaji. "Effect of the Choice of the Pressure Coupling Method on the Spontaneous Aggregation of DPPC Molecules". In: *The Journal of Physical Chemistry. B* 109 (Sept. 2005), pp. 14667–74.

[38] Céline Anézo et al. "Methodological Issues in Lipid Bilayer Simulations". In: *The Journal of Physical Chemistry B* 107.35 (2003), pp. 9424–9433.

[39] Yuqing Lin et al. "Application of Berendsen barostat in dissipative particle dynamics for nonequilibrium dynamic simulation". In: *The Journal of Chemical Physics* 146.12 (2017), p. 124108.

[40] Charles L Brooks III et al. *Classical Molecular Dynamics*. 2021.

[41] Yi Isaac Yang et al. "Enhanced sampling in molecular dynamics". In: *The Journal of Chemical Physics* 151.7 (2019), p. 070902.

[42] Ognjen Perišić and Hui Lu. "On the Improvement of Free-Energy Calculation from Steered Molecular Dynamics Simulations Using Adaptive Stochastic Perturbation Protocols". In: *PLOS ONE* 9 (Sept. 2014), pp. 1–18.

[43] Sergei Izrailev et al. "Steered molecular dynamics". In: *Computational Molecular Dynamics: Challenges, Methods, Ideas*. Springer, 1999, pp. 39–65.

[44] Sanghyun Park et al. "Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality". In: *The Journal of Chemical Physics* 119.6 (2003), pp. 3559–3566.

[45] Johannes Kästner. "Umbrella sampling". In: *WIREs Computational Molecular Science* 1.6 (2011), pp. 932–942.

[46] G.M. Torrie and J.P. Valleau. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling". In: *Journal of Computational Physics* 23.2 (1977), pp. 187–199. ISSN: 0021-9991.

[47] Shankar Kumar et al. "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method". In: *Journal of Computational Chemistry* 13.8 (1992), pp. 1011–1021.

[48] Tommaso Casalini et al. "Permeation of Biopolymers Across the Cell Membrane: A Computational Comparative Study on Polylactic Acid and Polyhydroxyalkanoate". In: *Frontiers in Bioengineering and Biotechnology* 8 (2020), p. 718.

[49] Liang Chen et al. "Molecular dynamics simulations of the permeation of bisphenol A and pore formation in a lipid membrane". In: *Scientific Reports* 6.1 (2016), pp. 1–7.

[50] Sunhwan Jo et al. "CHARMM-GUI: A web-based graphical user interface for CHARMM". In: *Journal of Computational Chemistry* 29.11 (2008), pp. 1859–1865.

[51] D.A. Case, I.Y. Ben Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. *Amber18*. 2018, University of California, San Francisco.

[52] Romelia Salomon-Ferrer et al. "Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald". In: *Journal of Chemical Theory and Computation* 9.9 (2013), pp. 3878–3888.

[53] Callum J Dickson et al. "Lipid14: the amber lipid force field". In: *Journal of Chemical Theory and Computation* 10.2 (2014), pp. 865–879.

[54] William L Jorgensen et al. "Comparison of simple potential functions for simulating liquid water". In: *The Journal of Chemical Physics* 79.2 (1983), pp. 926–935.

[55] Pengfei Li, Lin Frank Song, and Kenneth M Merz Jr. "Systematic parameterization of monovalent ions employing the nonbonded model". In: *Journal of Chemical Theory and Computation* 11.4 (2015), pp. 1645–1657.

[56] Michael Crowley et al. "Adventures in improving the scaling and accuracy of a parallel molecular dynamics program". In: *The Journal of Supercomputing* 11.3 (1997), pp. 255–278.

[57] Shuichi Miyamoto and Peter A Kollman. "Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models". In: *Journal of Computational Chemistry* 13.8 (1992), pp. 952–962.

[58] Virginia Miguel et al. "Molecular view of the interaction of S-methyl methanethiosulfonate with DPPC bilayer". In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1858.1 (2016), pp. 38–46.

[59] Justin L MacCallum and D Peter Tieleman. "Computer simulation of the distribution of hexane in a lipid bilayer: spatially resolved free energy, entropy, and enthalpy profiles". In: *Journal of the American Chemical Society* 128.1 (2006), pp. 125–130.

[60] Stephen P. Binks and Miloslav Dobrota. "Kinetics and mechanism of uptake of platinum-based pharmaceuticals by the rat small intestine". In: *Biochemical Pharmacology* 40.6 (1990), pp. 1329–1336.

[61] Lukasz Nierzwicki et al. "Interaction of cisplatin and two potential antitumoral platinum (II) complexes with a model lipid membrane: a combined NMR and MD study". In: *Physical Chemistry Chemical Physics* 17.2 (2015), pp. 1458–1468.