MOLECULAR MODELLING OF CHITOSAN AND 3-HYDROXYALKANOATE POLYMERS

Supervisor Prof. Dr. Filippo Rossi and Prof. Dr. Giuseppe Perale

Co-supervisor Dr. Tommaso Casalini

Master thesis in Chemical Engineering

Amanda Rosolen Camarano Monteiro

894389

Milan

2018

CONTENTS

LIST OF FIGURES

Abstract

This thesis work evaluates the interaction of chitosan and homopolymeric poly(3-hydroxydecanoate) with a lipid bilayer in the presence of ions so to simulate the biological environment of a cell membrane. It is inserted in the European project GoBioNanoMat Project, which aims at establishing a consolidated procedure for the development and production of polymeric nanobiomaterials for drug delivery by small and medium enterprises, their suppliers and research institutes through the safe by design (SbD) approach.

Both polymeric chains were simulated individually with the lipid bilayer in a full atomistic simulation of its molecular dynamics. The work was divided into two parts: the first consisted in building the structures for Chitosan and P3HD to be simulated alone in water at 310K (body temperature) for model validation following the standard simulation procedure for molecular dynamics. With the energetically minimized structures, a joined system was built with a previously equilibrated dioleoylphosphocholine (DOPC) lipid bilayer that represents the cell membrane. The interaction and passive permeation of the bilayer by each polymer were evaluated on the second part of the work with the Potential Mean Force (PMF) method employing the Umbrella Sampling method for enhanced sampling.

For the model validation of the polymers, the root mean square displacement and radius of gyration of the chains were analyzed. It was verified that the P3HD folded when solvated in water while chitosan did not exhibit any meaningful changes in conformation. Further study of the energies was performed and the link between the hydrophobic effect and the van der Waals interactions was demonstrated.

The free energy landscape for the membrane partitioning was calculated in accordance with the Weighted Histograms Analysis Method (WHAM) with the sampling obtained with a spring-like biased potential of harmonic constant equal to 2.5 kcal/mol. The overall sampling obtained was considered good with the exception of a few high energy coordinates. Chitosan was not able to penetrate the bilayer and its PMF accused a bidding region on the surface of the bilayer. The P3HD chain displayed opposite results with an energy maximum on the surface of the bilayer.

The results obtained will be used in future work inside GoBioNanoMat project for comparison with experimental results and attainment of the parameters for the construction of the coarse grain models for the simulation of nanoparticles.

SOMMARIO

Questo lavoro di tesi valuta l'interazione di chitosano e dell'omopolimero poli (3-idrossidiecanoato) (P3HD) con un doppio strato lipidico in presenza di ioni in modo da simulare l'ambiente biologico di una membrana cellulare.

È inserito nel progetto europeo GoBioNanoMat Project, che mira a stabilire una procedura consolidata per lo sviluppo e la produzione di nanobiomateriali polimerici per il rilascio controllato di farmaci da parte di piccole e medie imprese, i loro fornitori e istituti di ricerca attraverso un approccio *Safe by Design* (SbD).

La permeazione delle singole catene attraverso il doppio strato lipidico è stata caratterizzata attraverso simulazione su atomistica mediante dinamica molecolare. Il lavoro consisteva in due parti: la prima parte si è basata sulla costruzione delle strutture per chitosano e P3HD da simulare in acqua a 310 K (temperatura corporea) per la convalida del modello.

Con le strutture energeticamente minimizzate è stato costruito un sistema unito con un doppio strato lipidico dioleoilfosfocolina (DOPC) precedentemente equilibrato che rappresenta la membrana cellulare. L'interazione e la permeazione passiva del doppio strato da parte di ciascun polimero sono state valutate sulla seconda parte del lavoro con il *Umbrella Sampling* per il *sampling* delle conformazioni, adottando la distanza tra i centri di massa del polimero e della membrana come coordinata di reazione. Il Potenziale di Forza Media (PMF), rappresentativo della energia libera del sistema, è stato in seguito calcolato in funzione della coordinata di reazione con il metodo degli istogrammi ponderati (WHAM).

Per la validazione del modello dei polimeri sono stati analizzati lo spostamento quadratico medio e il raggio di girazione delle catene. È stato verificato che la catena di P3HD si ripiegava in soluzione acquosa mentre il chitosano non mostrava cambiamenti significativi nella conformazione, rimanendo completamente esteso. È stato eseguito un ulteriore studio delle energie e il legame tra l'effetto idrofobico e le interazioni di van der Waals è stato dimostrato.

Il profilo di energia libera per la partizione della membrana è stato calcolato in base al metodo di analisi degli istogrammi ponderati adottando un potenziale armonico di bias avente costante armonica pari a 2,5 kcal/mol. Il campionamento complessivo ottenuto è stato considerato buono ad eccezione di poche coordinate ad alta energia. Il chitosano non era in grado di penetrare il doppio strato ed il profilo di energia libera evidenzia un minimo all'interfaccia tra la membrana ed il solvente. La catena di P3HD è stata in grado di penetrare

nella membrana e la conformazione in cui la catena è al centro della membrana e non più esposta al solvente rappresenta il minimo del profilo di energia libera.

I risultati ottenuti saranno utilizzati nei lavori futuri all'interno del progetto GoBioNanoMat per il confronto con i risultati sperimentali e l'ottimizzazione dei parametri per la costruzione dei modelli *coarse grain* per la simulazione di nanoparticelle.

1. INTRODUCTION

1. The problem of drug delivery and nanobiomaterials today

The biggest challenge in drug delivery is to overcome the natural barriers imposed by the body. The traditional larger-sized materials currently in use face poor solubility, poor absorption inside the body and in vivo instability problems;[1] they reduce the efficacy of the drug delivery system requiring larger doses to be administrated and pose as an additional hazard of the drug. The recent application of nano-scaled materials for use in diagnostics tools and drug delivering have shown improved efficacy of pre-existing and new drugs.[1,2] Their reduced size allows for the potential crossing of the barriers making the drugs more target specific. The effectiveness of these delivery systems can also be attributed to the reduced drug toxicity, controlled drug release and modification of the drug pharmacokinetics and biological distribution.[3]

The understanding and control of matter at the nanoscale (<1000 nanometers) is the definition given by the National Nanotechnology Institute[4] to nanotechnology. Nanotechnology has already proven itself to be a bridge between biological and physical science.[4] At this scale, the properties of the materials can differ greatly from their bulk behavior, and this unique phenomenon opens prospects for novel applications, especially in nanomedicine, Figure 1.
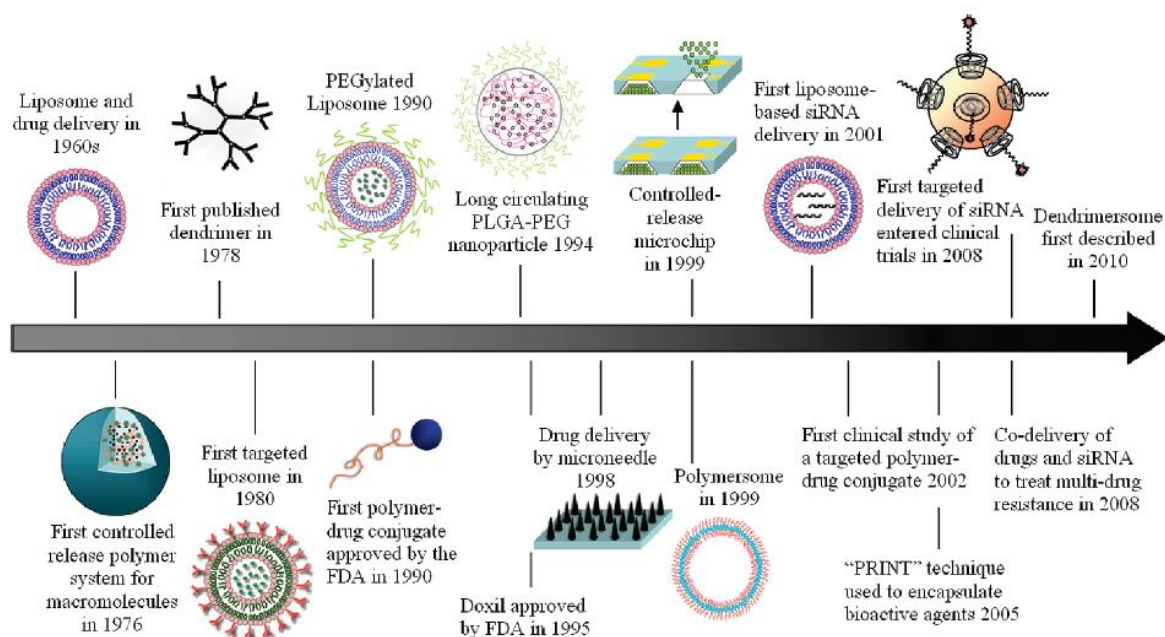
Figure 1. Application of nanotechnology for drug delivery systems throughout the years - important milestones. One of the first nanotechnology products for drug delivery systems were the lipid vesicles, the liposomes. More complex systems were described around the '80s. Later the use was polyethylene glycol (PEG), shown to increase circulation times of liposomes and polymeric nanoparticles paved the way to the approval of Doxil.[5] From that even more complex system have been developed with the goal to improve the targeting, delivering and release of the drugs. Taken from 1. Shi J. *et al.,*2010.[2]

The nanomedicine field has faced rapid growth in recent years with the search for more effective drugs especially when dealing with cancer treatments and the crossing of the blood-brain barrier.[1,3] The nanomaterials have revolutionized the drug delivery systems as they allow for the construction of a variety of systems of different materials, sizes, and properties, all in accordance with the drug's specifications, producing better target-oriented drugs, minimizing their risks and side effects. The main benefit of these nanoparticles is related to their surface properties which allow for the affixation of many proteins.[1]

Compared to traditional drug delivery system, nanobiomaterials are linked to reduced toxicity of the drugs, but at the same time, they impose an unknown risk as carriers. Natural products exhibit remarkable chemical and biological properties with macromolecular specificity and lower toxicity.[1] Drugs based on those products could possess unique advantages, such as lowered side effects with good therapeutic potential if the concerns related to biocompatibility and toxicity were to be cleared out.[1] Little is known about the toxicology of nanobioparticles and most part of the knowledge on this area is focused on their inhalation.[6]

For even more efficient usage in drug delivery, it is necessary to better understand their interaction with the human body.[3] Computational studies have helped to better picture molecular interactions of the drugs aiding on the development of next-generation drugs.[1] A multitude of substances is currently under investigation for the production of new drugs.

The primary goals for research of nanobiotechnologies in drug delivery include:[6]

- more specific drug targeting and delivery;
- reduction in toxicity while maintaining therapeutic effects;
- greater safety and biocompatibility;
- and faster development of new safe medicines.

Adjustments in nanostructures size, shape, hydrophobicity, and surface changes can further enhance the bioactivity of these nanomaterials[1] and their impact in biological and environmental systems,[7] that is why having a clear understanding of their properties is essential for the design of the carriers.

Another problem related to the production of nanodrugs and nano-based delivery system is related to the lack of regulatory guidelines that hinder the potential development and application of such materials.[1]

2. The "GoBioNanoMat" project

Figure 2. GoBioNanoMat Project logo[15]

GoBioNanoMat is a European project founded with the objective to develop a knowledge base and guidelines for the development and production of polymeric nanobiomaterials to be used in drug delivery systems using the safe-by-design (SbD) approach. The aim is to benefit industrial enterprises, especially small and medium companies, their suppliers, and research institutes working on the field.

Nowadays, due to the lack of knowledge on nanobioparticles, the use of the material in an industrial and commercial application is still limited. The goal is to provide a consolidated step-by-step development procedure to the institutes and small/medium enterprises, reducing uncertainties and commercial risks for polymeric nanobiomaterials drug delivery applications. With the knowledge base generated and joint work with regulatory agencies, based on the regulatory preparedness (RP) concept, it is expected that newly developed applications will have facilitated entry in the market.

The project consortium is composed by the departments of Particles-Biology Interactions, and Technology and Society of EMPA Swiss Laboratory of Materials Science and Technology; Life Technologies of University of Applied Science and Artes Western Switzerland (HES-SO); the School of Pharmaceutical Science of University of Geneva; the Department of Innovative Technologies of SUPSI; the Center of Neuroscience and Cell Biology from University of Coimbra; the Hightech Zentrum Aargau AG; the Institut für praxisorientierte Qualifizierung IPQ; the Faculty of Chemistry, Biology, Geography and Advances Environmental Research Laboratories of West University of Timisoara; and Acrostak AG. Together they combine expertise in the areas of nanobiomaterials science and technology, life science, pharmaceutical science, nano-safety, and lifecycle thinking.[15]

The Safe by Design principle has become a popular concept to address the risk of emerging technologies like nanotechnology.[16] With the application of SbD thinking on the innovation process, the goal is to achieve an early identification and management of potential risks for environmental health and safety, reducing uncertainties and increasing the trust so to accelerate outcomes and market access for innovative and safe nanobiomaterials. It aims at optimizing the hazard/advantage ratio during the materials life cycle, balancing costs and risk levels.

Besides the SbD thinking in which the design of the nanobiomaterial together with its unintended side effects, efficacy, and environmental risks are investigated, the regulatory framework and the product's lifecycle are considered. The regulatory framework is dealt with based on the RP concept of sharing expertise and knowledge towards guidance for registration or market approval.[17]

Nanobiomaterials represent a new and extensive field, full of applications yet to be exploited. Their use in the production of drug delivery systems is believed to represent an increase in the efficiency of the system and reduction of adverse responses. The GoBioNanoMat

project focuses on the study of three polymeric bionanomaterials regarding drug delivery applications: Polylactic Acid (PLA),[18] Chitosan and Polyhydroxyalkanoates (PHA); to consolidate and verify the implementation of the SbD approach in real industrial case studies.

On the present thesis, the interactions of a Polyhydroxyalkanoate (PHA) and of Chitosan with a lipid bilayer, used here to simulate the cell membrane, were evaluated through molecular modeling with the aim of generating more data and information to be used in the following phases of this project. In particular, the outcomes of the present work will be employed to test and tune a coarse-grained model, which allows simulating the permeation of entire nanoparticles across a cellular membrane, thus moving towards a system of practical interest.

3. Polyhydroxyalkanoates

Polyhydroxyalkanoates (PHA), due to their origin from natural sources and unique physiochemical and mechanical properties, have received special attention for their application as a drug delivery material; they are considered a potential biomaterial for the assembling of new nanometric drug carriers.

Many polymers are currently used as drug carriers due to their versatility. Polymers can be manipulated and especially produced to present preset characteristics; they have proved to be an effective carrier for targeted drugs increasing the drug efficacy.[9] Apart from having the right physical and chemicals properties, for a polymer to be used in the medical field, especially in drug delivery, it must be biocompatible and biodegradable.

The general molecular structure of PHA monomer is shown in Figure 3. The class of hydroxyalkanoates consists of 150 different monomers all characterized by a polyester backbone and a side R chain responsible for the variety of compounds and properties, physical and chemical, of the PHA biopolymers, as shown in Figure 4. They are subclassified, according to the length of the chain, in short (scl), medium (mcl) and long chain (lcl) monomers, being long-chain-lengths rarely synthesized. Small-chain-lengths have small side chains and are characterized to have a high degree of crystallinity while medium-chain-lengths PHAs have longer, and mostly hydrophobic mobile side chain which lower their molecular order.[19]

5
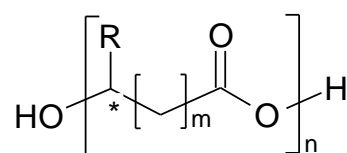
Figure 3. PHA general molecular structure. The asterisk denotes the chiral center for the monomer. The index m is related to the size of the backbone of the structure. The R group is responsible for the variety of the PHAs found, it can assume from a simple form as a proton to chains as long as 14 carbon atoms, including aromatic structures and ending functional groups. Taken from GoBioNanoMat report, 2018.[19]

R group	Carbon no.	PHA polymer
methyl	C_4	Poly(3-hydroxybutyrate)
ethyl	C_5	Poly(3-hydroxyvalerate)
propyl	C_6	Poly(3-hydroxyhexanoate)
butyl	C_7	Poly(3-hydroxyheptanoate)
pentyl	C_8	Poly(3-hydroxyoctanoate)
hexyl	C_9	Poly(3-hydroxynonanoate)
heptyl	C_{10}	Poly(3-hydroxydecanoate)
octyl	C_{11}	Poly(3-hydroxyundecanoate)
nonyl	C_{12}	Poly(3-hydroxydodecanoate)
decyl	C_{13}	Poly(3-hydroxytridecanoate)
undecyl	C_{14}	Poly(3-hydroxytetradecanoate)
dodecyl	C_{15}	Poly(3-hydroxypentadecanoate)
tridecyl	C_{16}	Poly(3-hydroxyhexadecanoate)

Figure 4. The nomenclature and carbon number of some PHA monomers as a function of the R group in place for a backbone of m=1. Adapted from Tan, G.-Y.A., *et al*, 2014.[20]

PHAs are linear hydrophobic polyesters that can be produced synthetically or naturally. The advantage of naturally producing the polymer is that it avoids the use of chemicals that could later affect the toxicology of the final product with their remaining residues. The biosynthesis of PHA is done by specific microorganisms, mainly gram positive and negative bacteria and archaea, in which energy and carbon are accumulated in their cytoplasm as granules under nutrient limiting-conditions and excess carbon.[21] There are three main metabolic pathways for the synthesis, depending on the limiting conditions and carbon source PHB is synthesized by the polymerization of (R)-3- hydroxybutyryl-CoA molecules, as summarized in Figure 5. The main limitation of the biosynthesis is the ability to control the molecular weight of the polymers generated with precision.
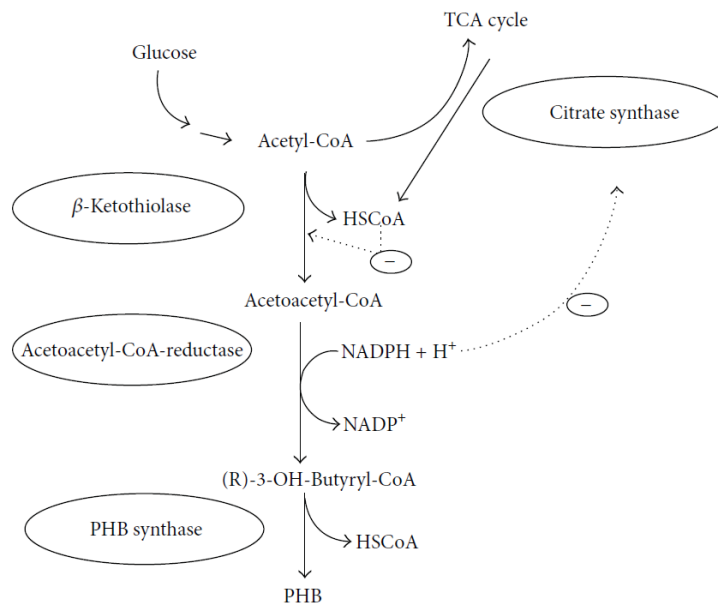
Figure 5. Metabolic Synthesis Pathway for the production of PHB from R. *eutropha.* Two acetyl-CoA molecules are coupled to form acetoacetyl-CoA, the product is then stereoselectively reduced to (R)-3-hydroxybutyryl-CoA. Finally, PHB is synthesized by polymerization of (R)-3-hydroxybutyryl-CoA molecules by the PHB synthase. Taken from Shrivastav A, Kim H, Kim Y.,2013.[21]

Polyhydroxyalkanoates are widely researched for their use as drug delivery carriers because of their biodegradability, biocompatibility, and non-toxicity. It is one of the oldest bioplastics since the discovery of polyhydroxybutyrate (PHB), which hydrolytically degrades to the monomer D-3-Hydroxybutyricacid, naturally found in the bloodstream as a product of cell metabolism produced during fatty acid oxidation in the liver,[21] leading to the belief that it would be well tolerated in the biological environment; some researches have even reported that its deprivation *in vivo* is faster[21] than the hydrolysis at body temperature *in vitro* indicating that some enzymes might play a role as catalyst. PHA can be also completely broken down by some microorganisms found in different environments. mcl-PHA and scl-PHA biodegradability are influenced by their specific characteristics such as molecular weight and crystallinity. Their stability is closely related to the degradability and can be modified and improved with the addition of additives at the cost of decreasing the biocompatibility.
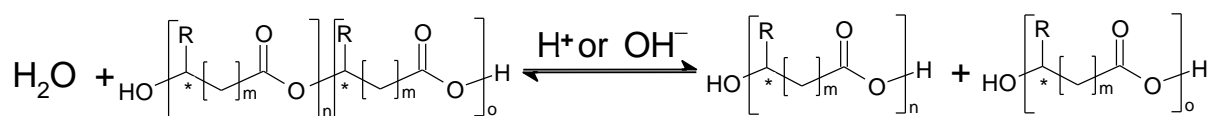
Figure 6. Degradation of PHA by hydrolysis of ester bonds of the backbone chain under acidic or basic conditions. Taken from GoBioNanoMat report, 2018.[19]

The polymer can be turned into films, porous matrices, microcapsules, microspheres, and nanoparticles. As a hydrophobic material, one of its main uses as drug delivery has been the microencapsulation of active drugs. PHA nanoparticles can be used on the constitution of the hydrophobic core of polymeric micelles, nanostructures made of amphiphilic block copolymers that form a core-shell structure in the aqueous solution. The hydrophobic core can be loaded with hydrophobic drugs, at the same time the hydrophilic shell makes the whole system soluble in water and stabilizes the core.[1] The slow surface degradation by lipases also indicates potential use for the controlled release of the drug compound although most of the medical applications of PHA still need to be evaluated *in vivo.*

4. Chitosan

Chitosan is a polysaccharide soluble in water derived from chitin, Figure 7, found in nature. It is the second most abundant polysaccharide[26] following cellulose and it has several unique properties that have drawn attention to its use in the medical field for many applications, including drug delivery, making it the most important polysaccharide.
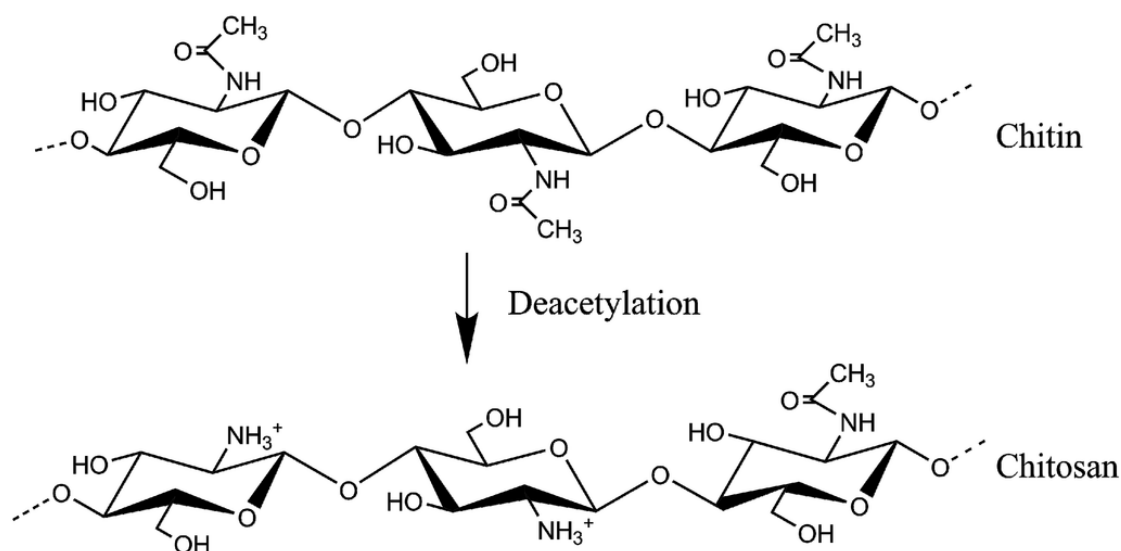
Figure 7. Chitosan polymer from chitin. Taken from Nilsen-Nygaard, J. *et al.*,2015.[27]

Its structure is similar to celluloses, differing on its primary amino groups which makes chitosan chemically active. It is a linear binary heteropolysaccharide composed of β-1,4-linked glucosamine with various degrees of N-acetylation of glucosamine residues.[28] It is

biocompatible and biodegradable with antimicrobial properties, immune stimulating activity, and capability of absorbing toxic metals. The cationic character of the amino groups is responsible for its mucoadhesive character, transfection, *in situ* gelation and permeation enhancement, and enables for the controlled delivery of drugs.[26]

Chitosan can be produced from chitin in regulated molecular weight and deacetylation degrees. Chitin is a natural, structural element found in the exoskeleton of crustaceans such as shrimps, lobsters, and crabs. Their shells grounded into powder are the raw material to produce chitosan. It can be prepared by alkaline N-deacetylation using concentrated sodium hydroxide solutions at high temperature for long periods or by N-deacetylation using enzymes under relatively mild conditions, chitin with a degree of deacetylation $\geq 75\%$ is considered chitosan. The alkaline N-deacetylation, Figure 8, is more common due to its simplicity. The process occurs in two steps. In the first step, the linear chain of N-acetyl-D-glucosamine joined together by β-1,4 linkage is extracted. The second step consists of the deacetylation in an aqueous solution of NaOH.
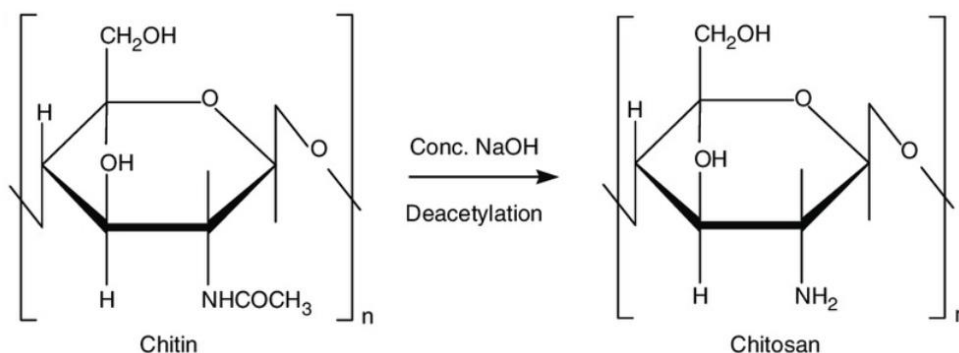
Figure 8. Alkaline N-deacetylation of chitin monomer to chitosan monomer. Taken from Dutta, J *et al.,* 2012.[29]

Chitosan with different deacetylation degree (DD) and molecular weights present different properties and functionalities such as solubility, viscosity, the reactivity of proteinaceous material coagulation, and heavy metal ion chelation. Another reason why chitosan attracts attention is due to its antimicrobial and antifungal activity. Its solubility, also influenced by a number of factors like DD, is mainly inversely affected by pH, having higher activity at lower values of pH due to its poor solubility in pH numbers higher than 6.5, Figure 9.
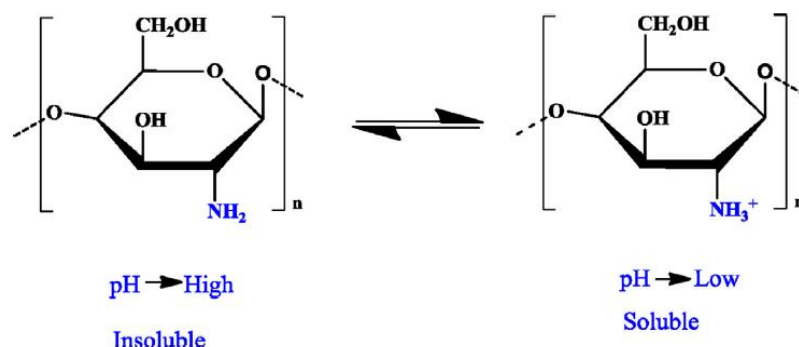
Figure 9. pH driven protonation of chitosan amino group. In basic conditions, it contains free amino groups -NH$_2$ while in acidic environments it protonates to -NH$_3^+$.[30] Taken from Ali A, Ahmed S., 2018.[26]

The biopolymer is non-toxic and soft tissue compatible. It has shown the ability to increase membrane penetration of large molecules across mucosal surface.[31,32] It is believed that its nanoparticles will only enhance such characteristics. It degrades slowly to amino sugars, harmless products completely absorbed by the human body.

Chitosan can be produced in many different forms, such as powder, films, and fibers.[30] The form of production of the nanoparticles will depend on the desired characteristics. It is used in the encapsulation or coating of various types of nanoparticles, thus producing different particles with multiple functions for their potential uses in the detection and diagnosis of different types of diseases.[1]

5. Molecular modeling

Molecular modeling studies the molecular structure and function through model building and simulation. It is a relatively young field that only recently has been capable to simulate time scales compatible with the biological processes thanks to the computational development. Nowadays it has become one of the most versatile computational technics for the simulation of biomolecules.[41] It enables the simulation of a variety of systems that could have not been simulated otherwise, either because of unfeasible conditions (extreme high temperature or pressure), safety limitations or high cost.[42] The molecular modeling allows for the simulation of systems at the microscopic level and the observation of properties and phenomena of very short scales of time and space like conformational changes and ligand binding. It is a connection between the macroscopic and the microscopic world provided by the

theory of statistical mechanics. Two different approaches can be used for the modeling and simulation of the molecules: the classical approximation and the quantum approximation.

For the classical approximation, molecular mechanics and molecular dynamics are used to calculate the properties of the system during the simulation. Molecular mechanics relies on the use of force fields (FF) to simulate and calculate all the forces acting on the molecules. Those forces are derived from the potential energy. A great number of packages, like AMBER and MM, provides specific force fields depending on the type of system and molecule one wants to simulate. Molecular dynamics deals with the calculation of the time-dependent properties and have as output for the simulations the trajectories of the molecules based on the iterative solution of Newton's second principle over time.

The empirical force fields are a potential energy function that allows for the computation of the total potential energy of the system as a function of atomic coordinates. It is described as the sum of the *bonded interaction potentials* and the *non-bonded interactions potentials*.[41] Each empirical force field is tuned with a set of parameters to model the specific type of molecules and they should be implemented accordingly. The mixing of different FF is not recommended unless stated otherwise by the force field developers.

The use of computer simulations in the description of microscopic systems has increased in recent years. It allows for an atomic-level comprehension of molecules and, thanks to the advances in high-performance computing and algorithm simplicity of molecular modeling, one is able to simulate hundreds of thousands of atoms in one routine. Even bigger systems can be simulated depending on the detail level. Full atomistic simulations, in which all atoms are represented, provide the deepest detail level. Although their accuracy is higher, so is the computational effort, therefore the simulation of larger systems and/or time scales beyond the ones accessible by full atomistic simulations (nanoseconds to microseconds) are unfeasible. For the simulation of bigger systems and to access longer time scales, an alternative is a coarse-grained model in which groups of atoms are enclosed into beads, parametrized in order to best reproduce properties like hydrophobicity or polarity of the embedded atoms.
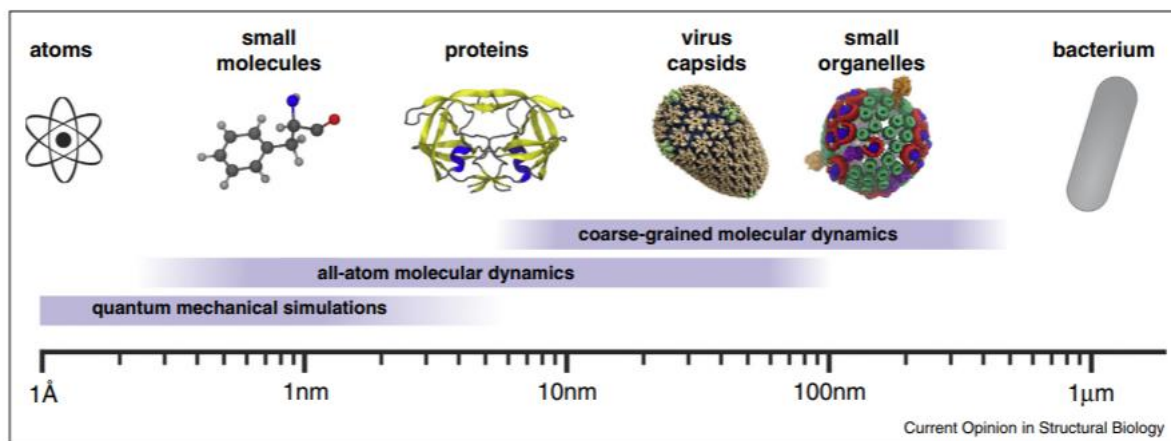
Figure 10. Different systems and the recommended approach for their modeling and simulation represented on a logarithmic scale from 1 ångström to 1 micrometer. Taken from Perilla, J.R. *et al.,*2012.[43]

The representation method for the solvent is also important when it comes to computational effort and is a key issue on the system definition. Explicit inclusion of solvent molecules leads to a comprehensive description of the system since solvation effects are accounted for in detail. The drawback is the increase of system size and, thus, a loss of computational efficiency. Many force fields have been developed to explicitly describe the water molecule as a solvent, such as the Transferable Intermolecular Potentials (TIP) family (TIP3P, TIP4P, TIP5P), the punctual charges models, SPC, and the ST2 model.[41] It is worth mentioning that often the choice of the water model is determined by the employed force field since FF are usually optimized adopting a specific water model.

For bigger systems the explicit representation of the solvent becomes unfeasible and some implicit models have been developed, none with equivalent accuracy. These models consider the solvent as a continuum medium computing its average influence on the solute. The method is specially employed in the simulation of electrostatic solvation for which the solvent is considered a dielectric medium.[41] The efficiency and precision of a simulation are proportional to the model and algorithm applied.

2. THEORETICAL BACKGROUND

1. Molecular dynamics

In molecular dynamics, molecules are described as a series of mutually interacting charged points, *i.e.*, the atoms. The main output is constituted by the atomic coordinates of the system as a function of simulation time obtained through the numerical integration of Newton's equation of motion.

For that purpose, the estimation of the forces acting on the atoms is needed, they can be derived from the potential energy of the system, Eq. (1).

$$F_i = m_i \ddot{r}_i = -\frac{\partial U}{\partial r_i} \tag{1}$$

Where $U(\vec{r})$ is the potential energy of the system; explicit function of atomic coordinates, F_i is the force acting on the ith atom, m_i its mass, and $\ddot{r}_i$ is the second derivative with respect to time of the spatial coordinates of the ith atom, r_i.

The potential energy function U(r) is also called force field, which is constituted by a suitable function form and set of parameters. There are general purpose force fields, like GAFF or OPLS, and FF specifically tailored for given categories of macromolecules, such as proteins, carbohydrates, nucleic acid, and lipids. At first glance, the accuracy of the results obtained by the simulation depends on the reliability of the chosen FF. Many force fields with different degrees of complexity are currently available for the description of a variety of systems. AMBER-based force fields were used in this work.

The potential energy accounts for non-bonded interactions and bonded interactions. The non-bonded interactions are described by the electrostatic and van der Waals forces and the bonded-interaction by the bond stretching, angle bending, and dihedral and improper torsions.[46] The functional form of AMBER-based force fields is reported in Eq. (2). The parameters used in such expressions are obtained by *ab initio* methods, fitting of experimental data or semi-empirical quantum mechanics calculations.

$$U = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\phi - \delta)]$$

$$+ \sum_{improper} V_{imp} + \sum_{LJ} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) + \sum_{elec} \frac{q_i q_j}{r_{ij}}$$

(2)

k_k: spring constants

r: bond length between two atoms

r_0: equilibrium/ reference bond length between the two atoms

θ: the actual bond angle between three atoms

θ_0: equilibrium/ reference bond angle between the three atoms

ϕ: torsional angle

δ: phase factor

V_n: height of the potential barrier

n: number of minima or maxima between 0 and 2π

V_{imp}: term accounting for the positive contribution of out-of-plane motion, usually a function of the improper angle (deviation from planarity)

ϵ_{ij}, σ_{ij}: van der Waals interaction parameter between two atoms.

r_{ij}: the distance between non bonded atoms i and j.

q_i, q_j: atomic charges for atoms i and j.

While bond stretching and angle bending are high-frequency motion and can be approximated to rigid, the bond torsions are less stiff and the right representation of this harmonic motion play an essential role on ensuring the correct degree of rigidity of the molecule and its conformation changes[46] and, therefore, on determining the local structure and relative stability of different molecular conformations. The additional term of improper torsion is needed to describe the planarity of some groups such as aromatic rings. As for the non-bonded,

(12-6) Lennard Jones potentials and Coulomb's law are used to describe the van der Waals and the electrostatic interactions, respectively.

In some force fields, the cross interaction between the components is also considered. Each atom is characterized by a partial atomic charge, which is a fraction of the electron charge and accounts for the different electronegativity. Atomic charges can be computed by means of different approaches. The restrained electrostatic potential method (RESP) is adopted in this work to be consistent with the parametrization procedure of AMBER-based force fields, as recommended by the developer.

Overall, a FF is a mathematical expression that describes the dependence of the energy of the system from its coordinates. For this reason, the calculations of the trajectory are performed iteratively as represented in Figure 11. Meanwhile, Eq.(1), represents a system of N coupled second-order non-linear differential equations that needs to be solved numerically through an appropriate integration algorithm.
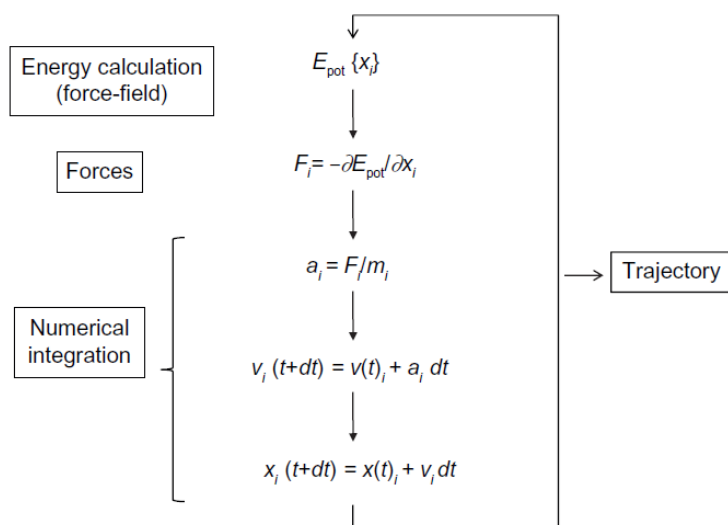
Figure 11. Molecular Dynamics iterative procedure. Taken from Gelpi, J. *et al.*,2015.[49]

- Boundary conditions

MD simulations systems consist of a small number of molecules if compared to a macroscopic bulk system. As a result, the percentage of molecules on the surface is much higher leading to a misrepresentation of the surface effects; the right selection of the boundary

conditions is then essential for obtaining the right results. In a simulation where the solvent is explicitly represented, its boundary conditions must be represented as well. To simulate the bulk characteristics, periodic boundary conditions (PBC) are used, which means that the simulation box is surrounded by an infinite number of replicas of itself, as depicted in Figure 12. This allows simulating the bulk phase of, *e.g.*, a water solution, with the disadvantage of introducing a periodicity not present in nature. The total number of atoms inside the periodic box is conserved in time, *i.e.*, if an atom leaves the box on one side, it will enter the system from the opposite side at the same position and with the same velocity.

In the simulations to follow, attention was taken so that the box size for replication would be large enough so that the molecule simulated would not be subject to any influence from its other periodic images. The use of this periodic boundary is possible from the assumption that the potential range is not too long;[42] for cases where long-range correlations play an important role, it should receive special attention.
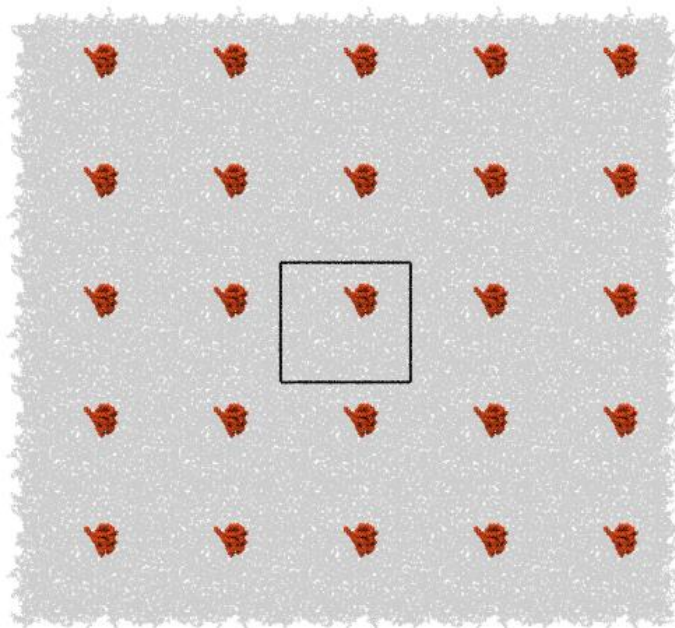
Figure 12. Periodicity of PDB boundary. – The box used for periodicity is represented in the middle and its periodic images can be visualized on the surroundings. The size of the box is so that the polymer simulated did not interact with its images.

- Initial position estimation

For obtaining a reliable trajectory, besides choosing the right force field to simulate it, it is fundamental to have a good input. The input for MD is composed of all the parameters used and the initial configuration of the system, *i.e.*, the initial coordinates, velocities and time.[48] The position of the atoms must be correctly estimated, especially because the Leonard Jones interactions are distance dependent, increasing very rapidly as intermolecular distance decreases.

The initial configuration is not necessarily a good representation of the system if it is analyzed in terms of energy configuration and spatial arrangement,[46] but it is a system that will smoothly converge to the desired conditions to be explored. And, since the MD simulation is programmed to explore a restricted region around the initial, a good starting point will most likely change the course of simulation and allow it to be performed in feasible time spans.

- Time step

The integration of the motion equations is performed through algorithms based on the finite-difference method in which the integration is performed over small time intervals, the time step. For a good representation of the actual motion, the time step must be small enough so to simulate even the highest frequency movement of the system. The intramolecular frequencies are considerably higher than the intermolecular, they call for a low time step which is responsible for a substantial increase in the computational cost. For that reason, the bond stretching and angle bending vibrations are assumed rigid. Thus, the time step can be set based on the translational and rotational movements increasing the efficiency of the simulation. For atomistic simulations, the value varies around 1 fs to 2 fs.[49]

- Solvation

As mentioned, the method used for the representation of the solvent plays a very important role in the computational efficiency of the simulation. In fact, when using an explicit representation, most of the molecules in the system are actually solvent. To be coherent with the atomistic level of detail of the molecules simulated in this work the solvation was performed explicitly and a number of water and ions molecules were added to the simulation box so to simulate the *in vivo* environment. The water coordination is usually taken from pure water simulations experiments and the ions follow the biopolymer charge distribution.[50]

2. Restrained electrostatic potential (RESP) method

The RESP method is the approach employed in Amber force fields during the parameterization process for the calculation of partial atomic charges. The concept of atomic charges is fundamental to all of chemistry and their values are essential for simulating the physical properties of condensed-phase matter.[51] The major strength of deriving charges from the electrostatic potential (ESP) is that it is able to reproduce the intermolecular interaction properties of the molecules with a simple two-body additive potential, the introduction of restrains in the magnitude of partial atomic charges to the fitting process in the form of a penalty function is a refinement to reduce problems of conformational dependency of charges and large charges.[52] The procedure is based on a multistage approach to ensure that rotationally degenerated atoms have equivalent charges.[52] In RESP, atom centered point charges are computed through fitting to reproduce the electrostatic potential around the molecules. To simulate a loaded molecule system its partial atomic charges must be computed accordingly in order to ensure the reliability of results.[53]

The first step of the procedure is to optimize the geometry of the structure, then the electrostatic potential, a combination of contributions from nuclei and electrons, is computed through quantum chemistry calculations, Eq. (3).[53]

$$\phi(r) = \phi_{nucl}(r) + \phi_{elec}(r) = \sum_{i=1}^{Nnuclei} \frac{Z_i}{|r - R_i|} - \int \frac{dr'\rho(r)}{|r' - r|} \tag{3}$$

$\phi(r)$: the electrostatic potential at the point r;

Z_i : the charge of the ith nucleus placed at the R_i point;

$\rho(r)$: electron density as a function of the coordinates.

The ESP is calculated in the solvent accessible region around the molecule at each point of a shell grid placed at 1.4, 1.6, 1.8, and 2.0 the Van der Walls radii. The least square procedure

is used to fit the charge q_j to each atomic center j in the molecule and the electrostatic potential is given by Eq. (4).

$$\phi_i^{calc} = \sum_{j=1}^{Natoms} \frac{q_j}{r_{ij}} \tag{4}$$

r_{ij}: distance from point i and charge j

The function to be minimized through the least-square procedure is defined in Eq. (5). At the minimum point, the first derivative should be zero with respect to all charges q_j, Eq. (6).

$$\chi_{esp}^2 = \sum_{i=1} (\phi_i - \phi_i^{calc})^2 \tag{5}$$

$$\frac{\partial \chi_{esp}^2}{\partial q_j} = 2 \sum_i (\phi_i - \phi_i^{calc}) \frac{\partial \phi_i^{calc}}{\partial q_j} = -2 \sum_i \frac{(\phi_i - \phi_i^{calc})}{r_{ij}} = 0 \tag{6}$$

So, if the molecule has N atoms with an overall charge Z, the Nth charge depends on the value of the previously calculated charges, Eq. (7).

$$q_N = Z - \sum_{j=1}^{N-1} q_j \tag{7}$$

In the restrained electrostatic potential method, the restraint penalty function is added to the objective function, Eq. (8) and the new least-square minimum is calculated by Eq. (9).

$$\chi^2 = \chi_{esp}^2 + \chi_{rstr}^2 \tag{8}$$

$$\frac{\partial \chi^2}{\partial q_j} = \frac{\partial \chi^2_{esp}}{\partial q_j} + \frac{\partial \chi^2_{rstr}}{\partial q_j} = 0 \tag{9}$$

The penalty function assumes hyperbolic form, Eq. (10), which, due to its non-linear derivatives in q_j, entails the iterative solution of the equations to self-consistency.

$$\chi^2_{rstr} = a \sum_j \left[(q_j^2 + b^2)^{\frac{1}{2}} - b \right] \tag{10}$$

$$\frac{\partial \chi^2_{rstr}}{\partial q_j} = aq_j (q_j^2 + b^2)^{-\frac{1}{2}} \tag{11}$$

A is the scale factor to determine the asymptotic limits of the strength of the restraint. The system of equations can be written in a matrix form, Eq. (12).

$$Aq = B \tag{12}$$

Where

$$A_{jk} = \sum_i \frac{1}{r_{ij} r_{ik}} \tag{13}$$

$$A_{jj} = \sum_i \frac{1}{r_{ij}^2} + \frac{\partial \chi^2_{rstr}}{\partial q_j} \tag{14}$$

$$B_j = \sum_i \frac{\phi_i}{r_{ij}} + q_{0,j} \frac{\partial \chi^2_{rstr}}{\partial q_j} \tag{15}$$

$q_{0,j}$ is the target charge for the restraint.

This method represents an improvement from the ESP method. The electrostatic potential method fails in describing correctly the intramolecular interactions. The application of the penalty function will mostly correct this issue with little cost to intermolecular phenomena description.

3. The simulation procedure

- Energy minimization

After the system has been correctly assembled, it is necessary to minimize its energy in order to guarantee that the structure of each molecule is credible to reality. The initial position given by the building blocks is adequately correct but might not represent the actual conformation. Therefore, the minimization process will optimize the geometry of the structure, removing unstable conformations by minimizing the potential energy of the system. This step is important to relieve strains of the guessed initial conformation and to relax possible steric clashes removing occasional bad contacts. There are many algorithms for the minimization such as the steepest descent method and the Newton-Raphson method,[41] the first has reduced computational efficiency but it is a more robust and reliable whilst the latter is more efficient but it might run into convergence problems if the initial structure is flawed.

- Heating

Once the potential energy has been minimized it is necessary to gradually heat the system to the temperature of interest. Since the system here simulated represents an *in vivo* environment the temperature was set to 310K (37ºC). The heating of the system is performed at constant volume, by assigning a pseudorandom value to the velocities of the system so that the total kinetic energy of the system matches the desired temperature, Eq. (16).[50]

$$\langle E_k \rangle = \frac{1}{2} \langle \sum_{i=1}^{3N} m_i v_i^2 \rangle = \frac{3N(k_B T)}{2} \tag{16}$$

Where 3N is the number of degrees of freedom.

To control the temperature during the simulations here described a Langevin thermostat is used. This thermostat applies random friction and random force to momenta, p_i, according to Eq. (17).[55]

$$\frac{dp_i}{dt} = \sum_{j}^{N} \left(F_{ij} |q_i - q_j| \right) - \gamma p_i + R_i(T) \tag{17}$$

- Equilibration

Even though the geometry was corrected and the temperature elevated to the desired value, the system is still far away from the thermodynamic equilibrium. This next step consists then in the initialization of the system after which the trajectories for analysis can be produced (production stage). The simulation must be long enough to guarantee the relaxation of the properties of interest through an exchange between kinetic and potential energy after the increment on kinetic energy from the heating procedure.

- Production

Once the molecules have been correctly minimized, heated, and equilibrated the trajectories for post-processing and analysis can be generated. In this step, the structural characteristics and dynamic behavior will be sampled at the desired temperature after which the behavior of the system is analyzed. The simulations must be long enough to guarantee the correct sampling of the behavior to be investigated.

4. Post processing

- RMSD

The Root mean square deviation (RMSD), Eq. (18), is a statistical measure of the difference between values of the same data set. For this work the measurement was used to describe the displacement of the oligomers from a reference structure, here set as the starting structure to remove the rigid-body motions of the MD frames.[56]

$$RMSD = \left[\frac{1}{N} \sum_{i=1}^{N} \left(r_i(t) - r_i(0) \right)^2 \right]^{\frac{1}{2}} \tag{18}$$

Where $r_i(t)$ is the coordinate of the atom i at time t and $r_i(0)$ its coordinate at time 0, the reference, being N the total number of atoms in the polymer.

The RMSD allows for the identification of the attainment of the relaxed structure.[53]

- Radius of gyration

The radius of gyration (R_g) gives a sense of size for the molecule,[57] describing the dimension of the polymer chain. It consists of the averaged squared-distance of the atoms from the center of mass of the molecule, Eq.(19).

$$R_g = \sqrt{\frac{\sum_i^n m_i |r_i - r_{cm}|^2}{\sum_i^n m_i}} \tag{19}$$

$$r_{cm} = \frac{\sum_i^n m_i r_i}{\sum_i^n m_i} \tag{20}$$

Where r_{cm} is the position of the center of mass of the molecule at that step.

The measures of the radius in time give a good description of solvation phenomena that is taking place. For hydrophobic polymeric chains, for example, the R_g can illustrate the folding of the chain as its value will reduce in time till a stable value, identified as the final folded structure.

- SASA

The solvent accessible surface area is defined as the surface characterized around a protein by a hypothetical center of the solvent sphere with the van der Waals surface of the molecule, Figure 13.[58]
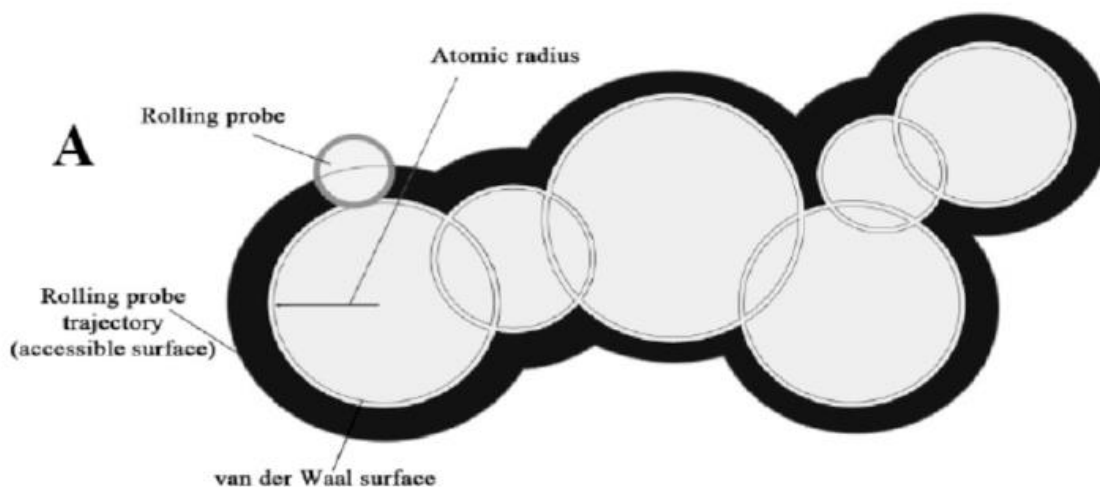
Figure 13. Illustration of the solvent accessible surface area. Adapted from Ali, S.A., *et al.,* 2014.[58]

It provides a good picture of the molecule inside the system and it can also be used for stability studies. There are many different methods for the calculation of the property, the one used in AMBER is the linear combination of pairwise overlaps (LCPO) method. It is based on the linear combinations of terms composed from pairwise overlaps of hard spheres.[58]

- Energies

The energy profile gives a good estimation of the stability of the molecule. It is a good indicator for the convergence and equilibrium of the system, it also gives new insights for the phenomena happening. For the computation of the energies related to the polymers alone in water, the MM-PBSA method was used. This approach calculates the free energy difference between two different solvated conformations of the same molecule.[59] The free energy profile of the polymer-membrane system is more complex and need special methods and procedures for its mapping.

5. Free energy calculations

The understanding of the free energy landscape is of outermost importance for the description of many microscopic phenomena happening inside the system such as protein folding, protein-ligand association, solvation, etc. It provides a link between the microscopic

behavior of a system and the most important thermodynamic equilibrium properties,[53] making it essential for the application of the materials on drug systems.

The free energy calculation has then become one of the most interesting and challenging problems on the modeling of biomolecules.[53] Standard MD simulations sample the space according to the Boltzmann distribution, therefore, events of low probability, corresponding to high-energy states, rarely occur in feasible simulation times. Consequently, it is difficult to obtain a reliable free energy landscape, especially for liquid systems due to their high number of minimum energy configurations. In addition, if two metastable states are separated by an energy barrier higher than the thermal energy, the system remains stuck in one state since this barrier would rarely be crossed, thus leading to insufficient sampling. To address the problem, many enhanced sampling methods have been developed to overcome the energy barriers and map the entire free energy landscape among which are Metadynamics (MT) and Umbrella Sampling (US).
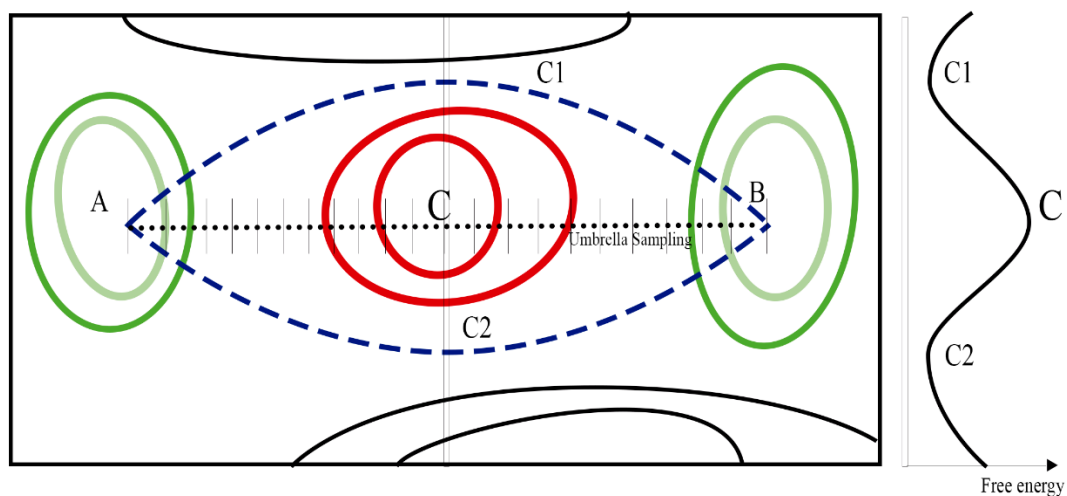
Figure 14. Free energy landscape illustration. Regions A and B represent low energy regions while C represents a high energy region not usually sampled, but of interest to the phenomena. The straight black line connecting A and B is one example of an energy landscape that would need enhanced sampling. The Umbrella Sampling method would sample this path in N windows (perpendicular black fine traces). Adapted from Yand, M., *et al.*, 2014.[60]

Statistical thermodynamics states the link between free energy and the partition function of the system, Q, Eq (21).

$$F = -\kappa_B T \ln Q \qquad (21)$$

Where κ_B is the Boltzmann constant and T is the absolute temperature. If it involves a constant number of particles, volume, and temperature its relation is to the Helmholtz free energy, if instead of volume, the pressure is kept constant, then it is related to the Gibbs free energy.[61]

The canonical partition function Q can be calculated by integrating the whole phase space as in Eq. (22).

$$Q = \int exp\left(-\frac{U}{\kappa_B T}\right) dr \qquad (22)$$

U is the potential energy independent from momentum.

Absolute free energies, Eq (21), however, can only be calculated for small and simple systems in which the analytical expression of the partition function can be obtained. For a more complex system, only the free energy differences can be computed.

$$\Delta F_{BA} = F_B - F_A = -\kappa_B T \ln\frac{Q_B}{Q_A} \qquad (23)$$

The membrane permeability and water-membrane partitioning of small particles are examples of simulations in which a great energy barrier is involved and at the same time, the free energy landscape is of importance in the analysis of the interaction. Hence, it is of interest to compute the differences free energy over one specific pathway, for that the potential mean force method was applied.

6. Potential Mean Force (PMF) method

The Potential Mean Force method allows for the description of the change in free energy along a reaction path.[65] The reaction coordinate can be a molecular degree of freedom or a collective coordinate obtained from a combination of degrees of freedom.[53] When evaluating the insertion of a drug into a membrane the method can provide good indications of how

component and membrane would interact, predicting and/or confirming the crossing. For the case here presented, the reaction path chosen is the distance that separates the center of mass of both the biopolymer and the lipid membrane along the z-axis.
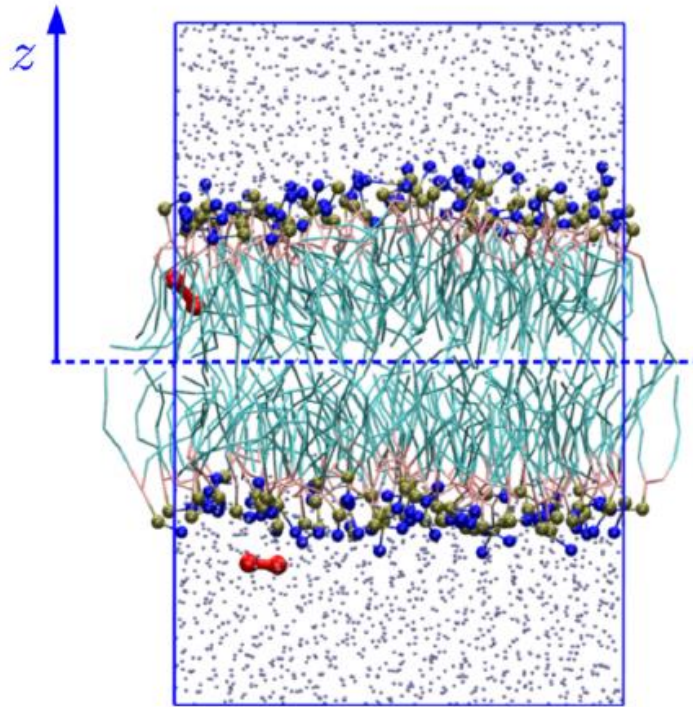
Figure 15. Illustration of the system simulated and the reaction path along the z-axis. Taken from Bochicchio, D., *et al.*,2015.[66]

Analyzing the problem in terms of reaction coordinate, q, the partition function can be calculated integrating Eq. (22) in all degrees of freedom with the exception of the reaction path itself. This value corresponds to the probability of finding the system in a small interval dq around q, a probability distribution p(q), Eq. (24).

$$p(q) = \frac{\int \delta[q_0(r) - q] \exp\left(-\frac{U(r)}{\kappa_B T}\right) dr}{\exp \int \left(-\frac{U(r)}{\kappa_B T}\right) dr} \tag{24}$$

The free energy at certain coordinate, Eq.(25), can, therefore, be obtained in terms of probability of finding the system at that state, $p(q)$, from the relation stated on Eq. (23).[53]

$$E(q) = E(q_0) - \kappa_B T ln \frac{p^u(q)}{p_0^u(q_0)} \tag{25}$$

If the energy barrier that separates two regions in space is much higher than $\kappa_B T$ it is unlikely that a standard sampling will comprise it. For that reason, the PMF method is coupled with the Umbrella Sampling Method and the Weighted Histogram Analysis Method (WHAM). The first will enhance the sample of the space while the later will derive the PMF from an iterative solution.

7. Umbrella sampling

The umbrella sampling is an enhanced sampling method that augments the transition between metastable stated through the addition of a harmonic bias potential W_i. The modified potential can be represented by Eq.(26).[66]

$$U_b(\vec{r}) = U_0(\vec{r}) + W_i(q) \tag{26}$$

With q being the coordinate of interest in the sampling, U_b the biased potential and U_0 the original potential. The subscript i refers to the ith window of the sampling.

To sample the entire region of interest N simulations, the windows, are needed. In each one of them, the system will be restrained to the position q_i with a force constant k_i, the restrained potential will shift the sample distribution along the reaction path for each window.[67] The good value for N will vary with the biased potential imposed. It is of outermost importance that the entire region is sampled so to obtain a good potential energy profile at the end. Different bias potentials can be applied to different positions along the path if needed.

$$W_i(q) = \frac{1}{2} k_i (q - q_i)^2 \tag{27}$$

Where k_i is a harmonic constant and q_i the position along q. The harmonic constant must be chosen wisely so to correctly sample all the space of interest; if it is too high the sampling will be narrowed to the chosen position if too low the sample will be sparse and derive unclear results. For different bias potentials, more than one harmonic constant is used.
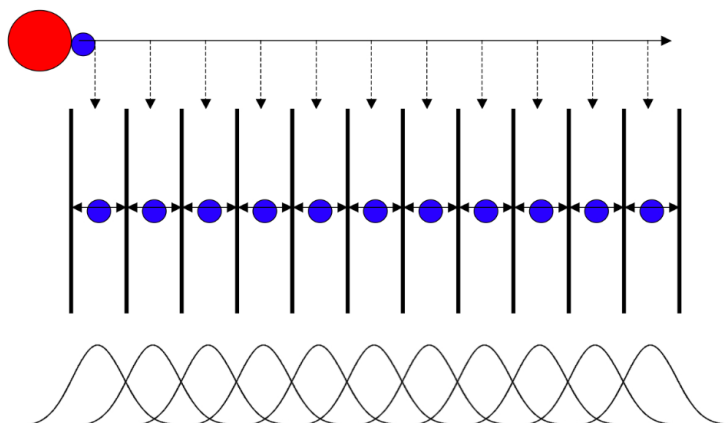
Figure 16. Reaction path on Umbrella Sampling.[68] The molecule of interest initially stranded to one position (red circle) due to energetic metastable state is restrained at a series of different coordinates along the reaction path (blue circles) for enhanced sampling.

The result of the US is a set of overlapping histograms $h_i(q)$, illustrated in Figure 16, that provides a probability distribution function $p_i^b(q)$, biased by the $w_i(q)$ previously applied.[64] The unbiased probability function is obtained from the biased with the weighted histograms analysis method (WHAM). From the unbiased probability distribution function $p_u(q)$ the free energy profile is obtained according to Eq. (25).

8. Weighted Histogram Analysis Method (WHAM)

The WHAM method is an extension of the Umbrella Sampling method for the calculation of the free energies of a system. It estimates the statistical uncertainty of the unbiased probability distribution from the set of umbrella histograms, $h_i(q)$, and subsequently computes the potential free energy.[64]

The histograms obtained with the US method represent a probability distribution from a biased potential, Eq. (26). The partition function for the biased system is described by Eq. (28).

$$Q_b = \int exp\left(-\frac{U_b}{\kappa_B T}\right) dr = \int exp\left(-\frac{U_0(\vec{r}) + W_i(q)}{\kappa_B T}\right) dr = \exp\left(-\frac{E_b}{\kappa_B T}\right) \tag{28}$$

Where E_b is the free energy of the restrained system.

$$E_b = -\kappa_B T \ln Q_b \tag{29}$$

While the partition of the unperturbed system is written on Eq. (30).

$$Q_0 = \int exp\left(-\frac{U_0(\vec{r})}{\kappa_B T}\right) dr = \exp\left(-\frac{E_0}{\kappa_B T}\right) \tag{30}$$

E_0 is the free energy of the unperturbed system. The partition function of the restrained system can be transcribed as a function of the unperturbed partition function and the perturbation potential averaged on the unperturbed state as stated on Eq. (31).[53]

$$\frac{Q_b Q_0}{Q_0} = \frac{Q_0 \int exp\left(-\frac{U_0(\vec{r})}{\kappa_B T}\right) exp\left(-\frac{W_i(q)}{\kappa_B T}\right) dr}{\int exp\left(-\frac{U_0(\vec{r})}{\kappa_B T}\right) dr} =$$

$$= Q_0 \int exp\left(-\frac{U_0(\vec{r})}{\kappa_B T}\right) \rho^u(q) dq = Q_0 \left\langle exp\left(-\frac{W_i(q)}{\kappa_B T}\right)\right\rangle_0 \tag{31}$$

From the relation, the biased, Eq. (33), and the unbiased, Eq. (32), probability distributions can be expressed in terms of the partition function.

$$p_i^u(q) = \frac{\int \delta[q_0(r) - q] \exp\left(-\frac{U_0(r)}{\kappa_B T}\right) dr}{\exp \int \left(-\frac{U_0(r)}{\kappa_B T}\right) dr} = \frac{\int \delta[q_0(r) - q] \exp\left(-\frac{U_0(r)}{\kappa_B T}\right) dr}{Q_0} \tag{32}$$

$$p_i^b(q) = \frac{\int \delta[q_0(r) - q] \exp\left(-\frac{U_0(\vec{r}) + W_i(q)}{\kappa_B T}\right) dr}{\int \exp\left(-\frac{U_0(\vec{r}) + W_i(q)}{\kappa_B T}\right) dr} = \tag{33}$$

$$= \frac{\int \delta[q_0(r) - q] \exp\left(-\dfrac{U_0(\vec{r}) + W_i(q)}{\kappa_B T}\right) dr}{Q_b}$$

Inserting the perturbed partition function, the biased distribution functions can be expressed in terms of the unbiased distribution functions, Eq (31).

$$p_i^b(q) = \frac{\int \delta[q_0(r) - q] \exp\left(-\dfrac{U_0(\vec{r}) + W_i(q)}{\kappa_B T}\right) dr}{Q_0 \left\langle \exp\left(-\dfrac{W_i(q)}{\kappa_B T}\right)\right\rangle_0}$$

$$= \left\langle \exp\left(-\frac{W_i(q)}{\kappa_B T}\right)\right\rangle_0^{-1} p_i^u(q) \exp{-\frac{W_i(q)}{\kappa_B T}}$$

(34)

The goal of the WHAM approach is to determine a PMF from the probability distribution functions so that it agrees on the regions of overlapping. Without a sufficient overlap of the histograms, the error related to this estimation increases.
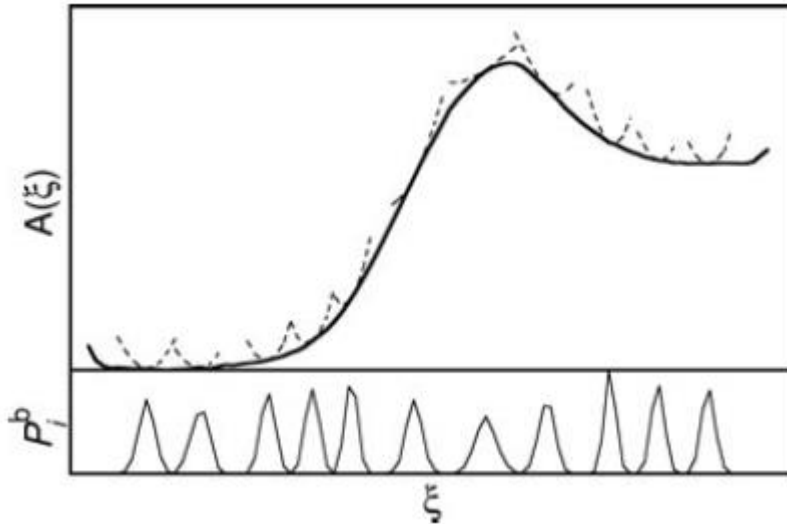
Figure 17. Illustration of a PMF obtained from a set of histograms obtained with Umbrella Sampling. The Global free energy is represented by the thick solid curve with the contribution of some of the windows in dashed lines. Below some of the windows of sampling are shown along the reaction coordinate, ξ. Taken from Kastner, J. 2011.[61]

Defining F_i as the best estimation of the free energy from the biased computation for the ith window.

$$\exp\left(-\frac{F_i}{\kappa_B T}\right) = \int exp\left(-\frac{W_i(q)}{\kappa_B T}\right)p_i^b(q)dq = \left\langle\exp\left(-\frac{W_i(q)}{\kappa_B T}\right)\right\rangle \tag{35}$$

The PMF related to the ith window is computed.

$$F(q) = F(q_0) - \kappa_B T ln\frac{p_i^u(q)}{p_0^u(q_0)} - W_i(q) + F_i \tag{36}$$

The global probability distribution $P(q)$ is computed by a weighted sum of the N windows and their unbiased partial probability distribution $p_i^u(q)$, Eq. (37).

$$P(q) = C\sum_{i=1}^{N_w}\omega_i(q)p_i^u(q) \tag{37}$$

Where $\omega_i(q)$ is the weighting factor and C is a normalization constant. The sum of the weights must be equal to one, Eq. (38), and their values must be set with the objective of minimizing the statistical error of the global probability $P(q)$, Eq. (39).

$$\sum_{i=1}^{N_w}\omega_i(q) = 1 \tag{38}$$

$$\frac{\partial\sigma^2(P(q))}{\partial\omega_i} = 0 \tag{39}$$

Where $\sigma^2(P(q))$ is the variance of the global distribution. Hence, the weights can be described as in Eq. (40).

$$\omega_i(q) = \frac{n_i \exp\left(-\dfrac{U_{bi}(q) - F_i}{\kappa_B T}\right)}{\sum_{j=1}^{N_w} n_j \exp\left(-\dfrac{U_{bj}(q) - F_j}{\kappa_B T}\right)} \tag{40}$$

n_i is defined as the number of points used on the construction of the histograms, the bins.

Introducing the previous equation to the global probability distribution function the Eq. (41) is derived.

$$P(q) = \sum_{i=1}^{N_w} \frac{n_i}{\sum_{j=1}^{N_w} n_j \exp\left(-\dfrac{W_j(q) - F_j}{\kappa_B T}\right)} p_i^b(q) \tag{41}$$

Since the global distribution function and the free energy F_i are interdependent the problem, Eq.(35) and Eq.(41), must be solved iteratively to convergence.

3. AIM OF THE WORK

The goal of this work is to produce enough data to aid in future developments for the use of poly(3-hydroxylalkanoates) and chitosan in drug delivery applications. The work aims at developing reliable models of the polymers for MD simulations of nanoparticles composed by any of the two components.

It consisted of two main parts:

1. MD simulation of the polymer alone in the water solution;
2. MD simulation of permeation of the polymer on a DOPC lipid bilayer.

Through the atomistic molecular dynamics simulation of a small chain, the initial goal is to be able to parametrize a coarse grain model that will allow for the simulation of bigger particles and bigger time intervals, thus moving to simulations of systems of practical interest.

Once the model for the polymer is validated by comparison of the phenomena observed with literature, the simulation of its interaction with the lipid bilayer is analyzed. The objective is to compute the free energy profile of the membrane partition by the oligomer.

4. METHODOLOGY

1. Building the molecules

- ▪ P3HD

The first part of the work consisted of the atomistic simulation of a PHA chain in water. The PHA simulated in this work was a small chain of h-P3HD, homopolymeric poly(3-hydroxydecanoate), an mcl-PHA that can be produced pure at high yields from fatty acids.[19] Chain ends were saturated with a methyl group on one side and a methoxy group on the other.
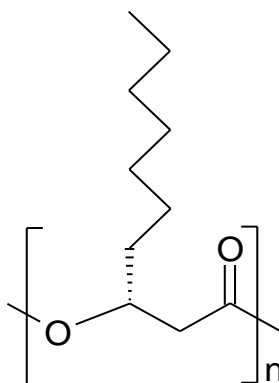
Figure 18. 3-hydroxydecanoate monomer[19]

The oligomer was simulated using the 2nd generation general AMBER force field (GAFF2). GAFF2 is a general and complete force field that covers most of the organic chemical species with all the parameters needed for composing molecules made of C, O, N, H, S, P, Cl, Br, and I.[76] The parametrization was performed entirely with HF/6-31G* RESP charges formalism and, to keep consistency, so was this method used on the computation of the partial atomic charges.

For the modeling of the polymer, an initial structure was optimized according to the density functional theory (DFT) calculation *in vacuo* at B3LPYP/6-31G (d,p) level of theory and fed as input to the computation of the electrostatic potentials *in vacuo*. On a two-stage process, the charges were computed by, first, making the overall molecule neutrally charged and then setting for equivalent atoms to have equivalent charges. Three building blocks fragments were generated, a head unit (3PI), Figure 19, a mid-chain unit (3PH), Figure 20, and a tail unit (3PF), Figure 21. The full description of each building block can be found in

Appendix A – P3HD Building Blocks. With these fragments, P3HD polymeric chains of any length can be built for an all-atom simulation.
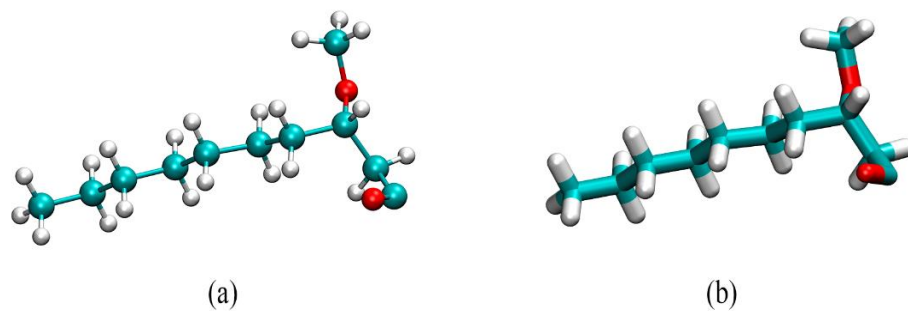
Figure 19. 3PI building unit (a) CPK representation (b) licorice representation. Carbon atoms are represented in cyan, hydrogen atoms in white and oxygen atoms in red.
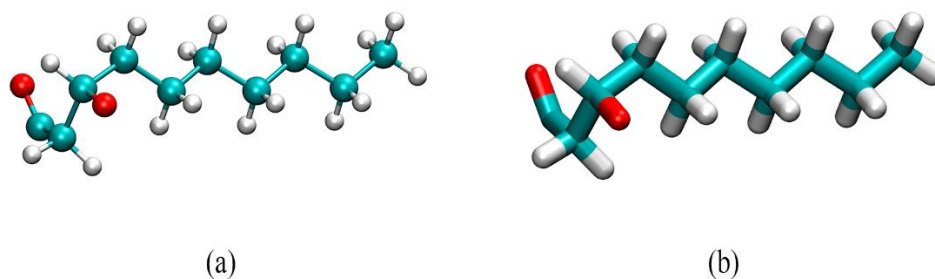
Figure 20. 3PH building unit (a) CPK representation (b) licorice representation. Carbon atoms are represented in cyan, hydrogen atoms in white and oxygen atoms in red.
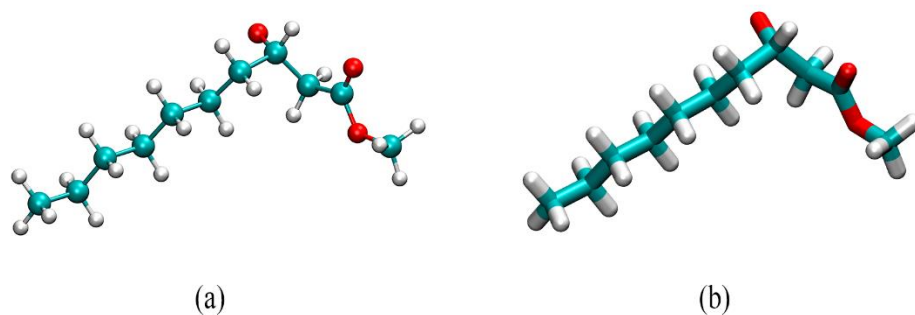
Figure 21. 3PF building unit (a) CPK representation (b) licorice representation. Carbon atoms are represented in cyan, hydrogen atoms in white and oxygen atoms in red.

The obtained building blocks were used to assemble a polymer chain containing 20 monomer units, as shown in Figure 22. The monomers were placed in order to assure the correct chirality as depicted in Figure 23.
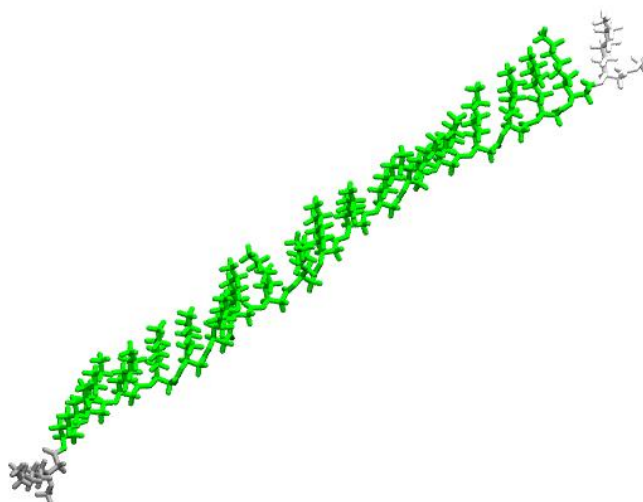
Figure 22. Assembled PHA composed of one 3PI building unit (gray), eighteen 3PH units (green), and one 3PF (white)
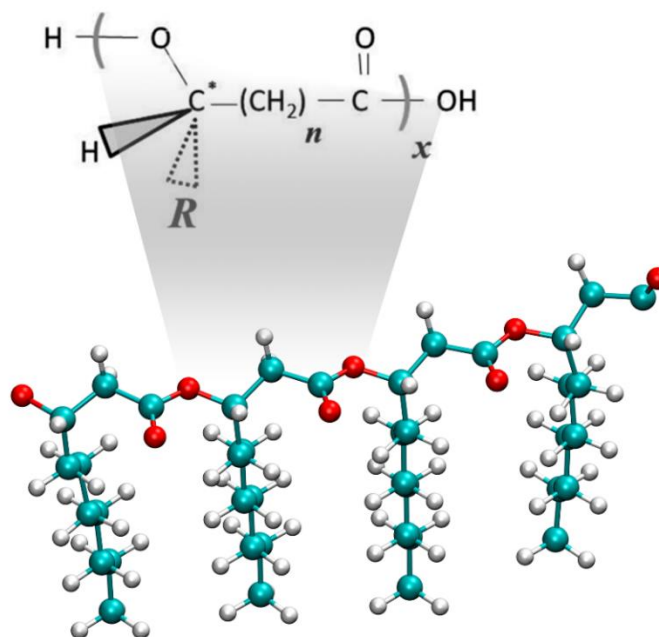
Figure 23. A segment of the h-P3HD chain. The chirality obtained is compared to the theoretical molecular structure from Koller, M.,2015.[77] Carbon atoms are represented in cyan, hydrogen atoms in white and oxygen atoms in red.

The LEaP program was used to prepare the input for the AMBER molecular dynamics simulation. The FF, PDB, and libraries for the building blocks were loaded onto the program for the creation of the topology and parameters files.

The structure was explicitly solvated in water with the pre-equilibrated box of TIP3P water, the TIP3PBOX, with a buffer of 10Å around the oligomer meaning that a solute box of 10Å was placed between the molecule and the end of the simulation box. No ions were added to assure electroneutrality since the polymer under investigation is not charged.

The molecular dynamics simulations were run through *pmemd* (Particle Mesh Ewald Molecular Dynamics), a primary engine within AMBER software, developed for improved performance with parallel and highly performant GPU capabilities.[76] Following the general protocol, an initial minimization simulation of 5000 steps was run to allow for the structure to relax and remove possible bad contacts created by solvation. The heating procedure was performed on the steepest descendent algorithm for the first 2500 cycles followed by the conjugate gradient algorithm for other 50000 steps. The non-bonded cut off distance was set to 10Å. The system was heated to 310K on 50 ps in NVT ensemble. The temperature control was

38

done with the Langevin thermostat with a collision frequency of 1 ps^{-1}. For this stage, the P3HA was kept restrained to a fixed position. The SHAKE constraint was enabled to all covalent bonds involving hydrogen and their forces were not calculated allowing the use of a 0.002 ps time step. The equilibration was run at constant pressure (1 atm) and temperature (310K), NPT ensemble, during 1 ns with periodic boundary conditions. The pressure was controlled with a Berendsen barostat. Finally, 140 ns of simulation were run under the same conditions for the production of trajectories to be used for analysis.

The RMSD and radius of gyration of the structure were obtained with *cpptraj*,[78] the main program in AMBER for processing coordinate trajectories and data files,[76] they were computed for each frame of the production using the initial structure as a reference. The solvation energy and the total energy were obtained with the MM-PBSA approach,[76] where the molecular mechanical energies are computed with the sander program from AMBER and the solvation free energy is calculated with a numerical solver for the Poisson-Boltzmann (PB) method.[76]

The next step consisted of the construction of the polymer-lipid bilayer system. The final structure of P3HD obtained was added to the simulation box containing DOPC lipid bilayer, previously equilibrated in past work of GoBioNanoMat project.[18] For that, the *AddToBox* tool of AMBER was used, then the chain was manually positioned with *vmd*[79] to assure a non-interacting distance with the bilayer. Later, the system was explicitly solvated with the addition of 36000 water molecules, also 93 molecules of K$^+$ and 93 molecules of Cl$^-$ were added to simulate the ionic concentration found *in vivo* approximately 0.15M of KCl. The system was minimized, heated and equilibrated to remove any bad connections and arrive to a minimal energy state.

With the equilibrated system, umbrella sampling simulations were performed to investigate the penetration of the chain into the bilayer. The chosen collective coordinate was defined as the distance between the center of mass of both structures (the bilayer and the PHA) along the z-axis. The biased probability distribution was computed at 40 windows set from 0Å, the center of the bilayer, Figure 25, to 40Å (spam of 1Å) using the AMBER center-of-mass (COM) umbrella restraint code.

The restraining harmonic constant, Eq. (27), was chosen based on previous work of the project.[18] Since the value showed to be too high and the histograms did not show a good overlap,

the value was decreased to 2.5 kcal/mol. Two sets of short simulations, 10 ns each, were performed at both ends of the reaction path (0Å, 01Å, 02Å and 30Å, 31Å, 32Å).
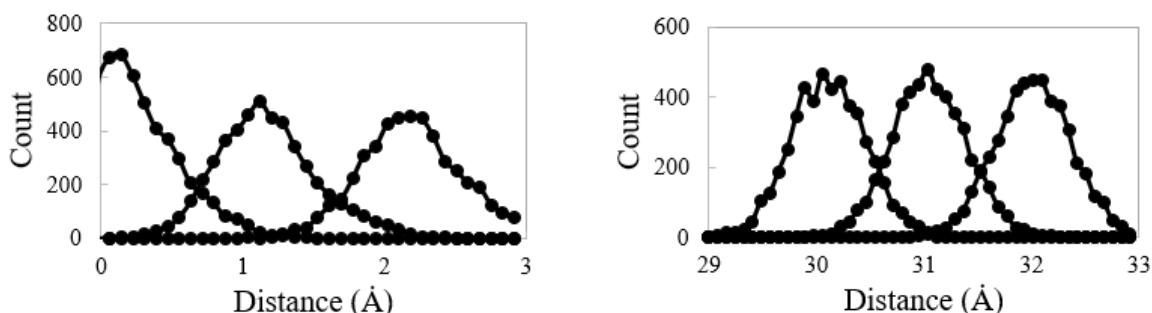
Figure 24. Restraining harmonic constant trial: histograms produced with k=2.5 kcal/mol

Since a good overlap was identified, the harmonic constant chosen was 2.5 kcal/mol. Each window was simulated for 40 ns with a time step of 0,002 ps in anisotropic pressure scaling conditions. To reduce computational costs, only the position of the atoms of the hydrophilic part was used for the computation of the center of mass of the membrane. The periodic boundary conditions were applied in all directions.
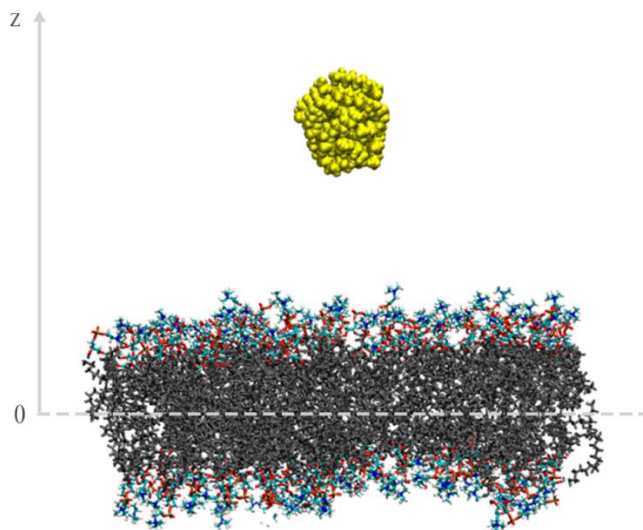
Figure 25. Reaction coordinate for the US simulations. The P3HD molecule (yellow) was brought from the equilibrium distance from the COM of the bilayer to q_i along the z-axis for each window i. The hydrophilic part of the bilayer is represented atom by atom being carbon in cyan, hydrogen atoms in white and oxygen atoms in red. The tail group is drawn in gray.

The molecular mechanical and internal energies and van der Waals and electrostatic interactions were calculated with *Sander* for each window. The PMF was built from the histograms with WHAM as described on 2. Theoretical Background with the aid of MatLab. The SASA and R_g were obtained with *cpptraj* and, finally, the free energy related to the interaction between the two parts was calculated by subtracting the free energy equivalent to the polymer alone solvated in water and the one equivalent to the membrane alone solvated in water, Eq, (42).

$$\Delta E = E_{system} - E_{P(3HD)} - E_{bilayer} \qquad (42)$$

- Chitosan

The molecular modeling of Chitosan was performed with the aid of GlyCam website,[80] a platform built by the University of Georgia with the objective of facilitating quality modeling for carbohydrates. With the carbohydrate builder tool, the structure of chitin (16 β-1,4-linked N-acetylglucosamine saturated by a methyl ester) was built. The program is implemented with a predictive tool for the 3D structure so to provide a good initial structure.[80]
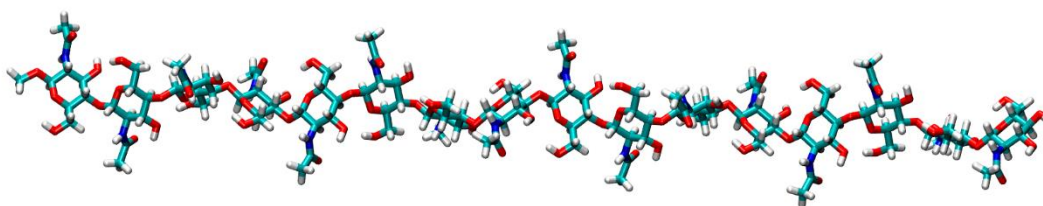
Figure 26. Chitin chain built with GlyCam WEB. Carbon atoms are represented in cyan, hydrogen atoms in white, oxygen atoms in red and nitrogen in blue.

The deacetylation to chitosan was performed manually by removing the N-acetyl groups from the structure, Figure 27. As mentioned, chitosan with different degrees of deacetylation presents different properties. In the present work, the molecule was modeled with a DD of 100%.
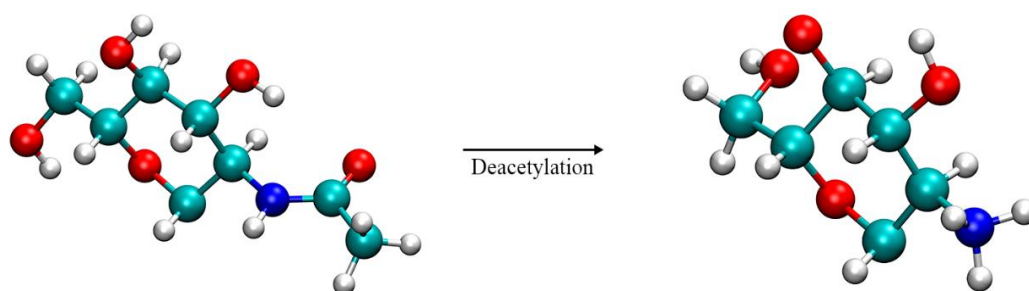
Figure 27. Deacetylation of Chitin monomer. Carbon atoms are represented in cyan, hydrogen atoms in white, oxygen atoms in red and nitrogen in blue.

The final structure can then be divided into three building blocks, Figure 28, the initial 0YP, Figure 29, the chain repeating unit 4YP, Figure 30, and the final unit, the methyl ester building block OME, Figure 31. Their detailed description of charges and positions can be found in Appendix B – Chitosan Building Blocks, with the exception of the final unit which is a standard residue of the FF.
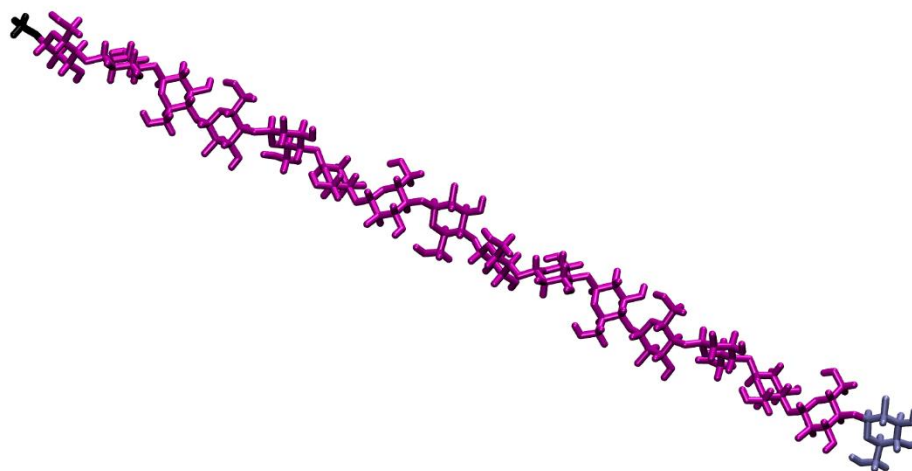
Figure 28. Chitosan structure assembled composed of one 0YP building unit (ice-blue), fifteen 4YP units (purple), and one OME (black)
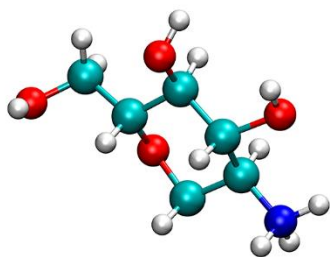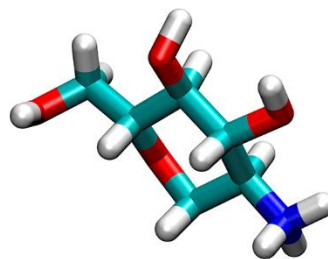
Figure 29. 0YP residue in CPK (a) and licorice (b). Carbon atoms are represented in cyan, hydrogen atoms in white, oxygen atoms in red and nitrogen in blue.
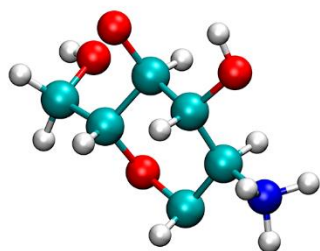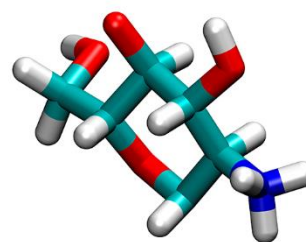
Figure 30. 4YP residue in CPK (a) and licorice (b). Carbon atoms are represented in cyan, hydrogen atoms in white, oxygen atoms in red and nitrogen in blue.

Figure 31. OME residue in CPK (a) and licorice (b). Carbon atoms are represented in cyan, hydrogen atoms in white, oxygen atoms in red and nitrogen in blue.

The MD simulation of the oligosaccharide was performed with the GLYCAM-06 force field. GLYCAM-06 torsion terms were developed by fitting to quantum mechanical data (B3LYP/6-31++G(2d,2p)//HF/6-31G(d)) for small molecules, therefore often highly symmetric molecules, asymmetric phase shifts were not required in the parameters allowing the set to be used for both α- and β- carbohydrates anomers regardless of monosaccharide ring size or conformation.[76,81]

The input for the MD simulation was prepared with tLeap. The structure was explicitly solvated in water with the pre-equilibrated box of TIP3P water, the TIP3PBOX. With a buffer of 10Å around the oligomer. To balance the positively charged molecule, 16 atoms of Cl- were added to obtain a system electrostatically neutral.

The molecular dynamics simulations were run through *pmemd*. Following the general protocol, an initial minimization simulation of 5000 steps was run to allow for the structure to relax and remove possible bad contacts created by solvation. The system was heated to 310K on 50 ps an NVT simulation. The non-bonded cut off distance was set to 10Å. The temperature control was done with the Langevin thermostat with a collision frequency of 1ps^{-1} with chitosan restrained to a fixed position. The SHAKE constraint was enabled to all bonds involving hydrogen, therefore, the time step employed was of 0.002 ps. The equilibration was run at constant pressure (1 atm) and temperature (310K), NPT ensemble, during 1 ns with periodic boundary conditions. The pressure was controlled with a Berendsen barostat. Finally, 140 ns of simulation were run under the same conditions for the production of trajectories to be used for analysis.

For the analysis of the model, the RMSD and radius of gyration of the structure were obtained with *cpptraj*[76] for each frame of the production simulations using as reference structure the initial input. The solvation energy and the total energy were obtained with the MM_PBSA approach.[76]

The final structure of chitosan was extracted by stripping water and Cl- out of the system and then placed on a system together with the previously minimized bilayer with *vmd*. After that, the *AddToBox* tool was used to solvate the new system with 36000 water molecules; 93 molecules of KCl were added and an additional 16 Cl[-] ions were included to neutralize the system. The system was minimized, heated and equilibrated according to the same procedure used for the Chitosan molecule alone.

With the equilibrated system, the umbrella sampling simulations were performed to investigate the penetration of the oligosaccharide into the bilayer. The reaction path chosen was the z-axis and the collective coordinate defined as the distance between the center of mass of both structures. The biased probability distribution was computed at the 40 windows set from 0Å, the center of the bilayer, to 40Å (spam of 1Å) using the AMBER COM umbrella restraint code. With a restraining harmonic constant of 2.5 kcal/mol, each window was simulated for 40ns with a time step of 0,002 ps in anisotropic pressure scaling conditions. To reduce computational costs, only the position of the atoms of the hydrophilic part was used for the computation of the center of mass of the membrane. The periodic boundary conditions were applied in all directions.

2. Summary of simulations

Table 1. Summary of simulations performed

System	Box Size [nm x nm x nm]	Water molecules	K^+ molecules	Cl^- molecules	Simulation Time [ns][1]
P3HD alone	99.311x84.484x80.618	21849	0	0	140
P3HD + Bilayer	96.437x88.897x217.612	36000	93	93	40*
Chitosan alone	86.306x63.105x101.682	17948	0	16	140
Chitosan + Bilayer	96.437x88.897x217.612	36000	93	109	40*

[1]Production stage; *For each window of the US

5. RESULTS

1. PHA

The MD simulation of the P3HD chain alone in water is essential for the validation of the structure and to ensure that an adequate set of force fields and parameters were applied. The initial and final conformations of the polymer are shown in Figure 32. As expected, the initially linear chain folded once placed in the polar solvent due to the hydrophobic behavior of the PHAs; such comportment serves as a good indicator for the reliability of the model.

(a)   (b)
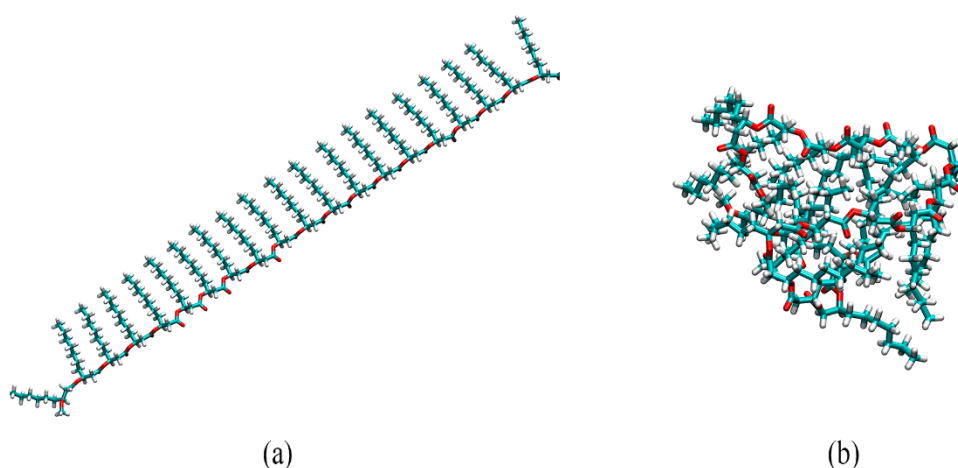
Figure 32. Input (a) and output (b) conformations of the MD simulation for P3HD solvated in water. Carbon atoms are represented in cyan, hydrogen atoms in white and oxygen atoms in red.

Because of its non-polar characteristic, the compound tends to minimize the surface area exposed to the polar solvent. The modeled P3HD exhibited the behavior, achieving a completely folded conformation after 3 ns of simulation. Snapshots of some conformations observed in time are shown in Figure 33. Further simulations were run to ensure the convergence of the oligomer both structurally and energetically.
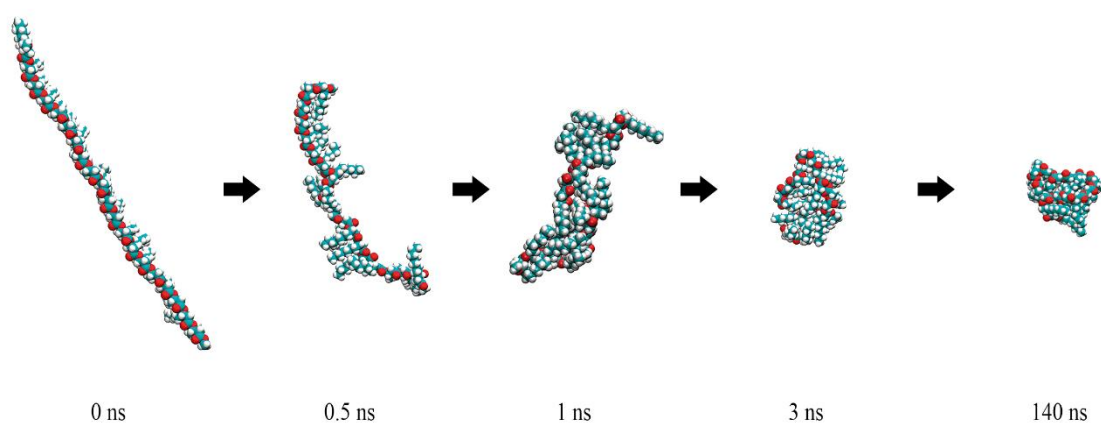
Figure 33. P3HD folding over time

The convergence of the simulation was checked by the computation of the radius of gyration and RMSD of P3HD as a function of simulation time. The R_g showed a stable value of 9.13 ± 0.22Å and its complete profile is shown in Figure 34. The RMSD, Figure 35, converged to 24.59 ± 0.752Å.
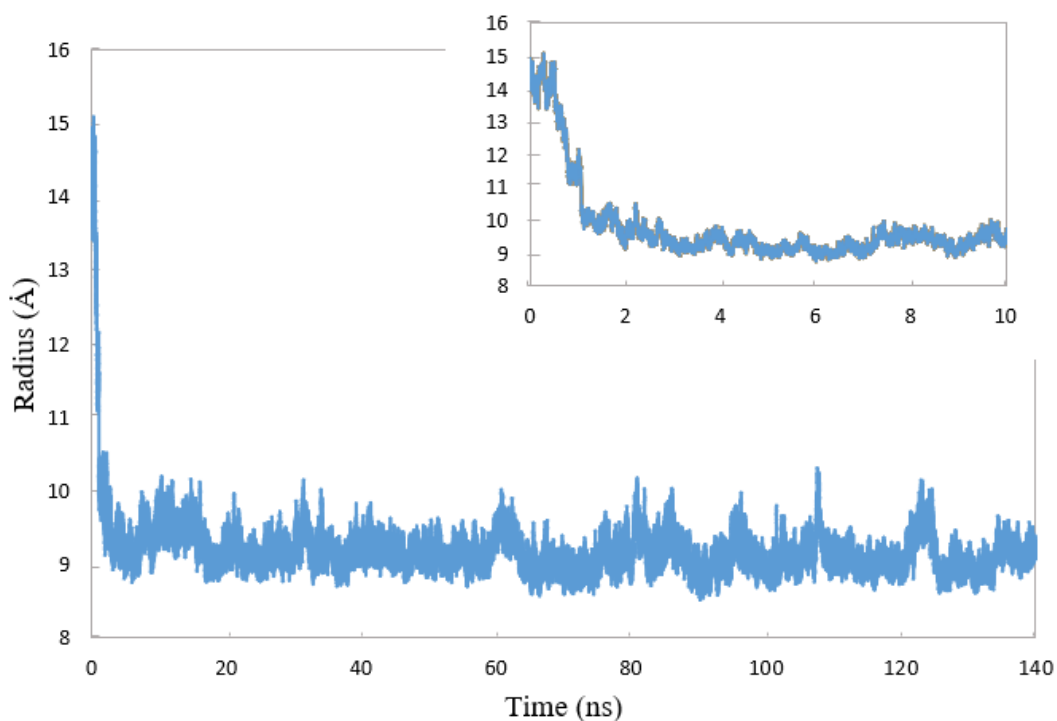
Figure 34. The radius of gyration of P3HD chain solvated in water over time
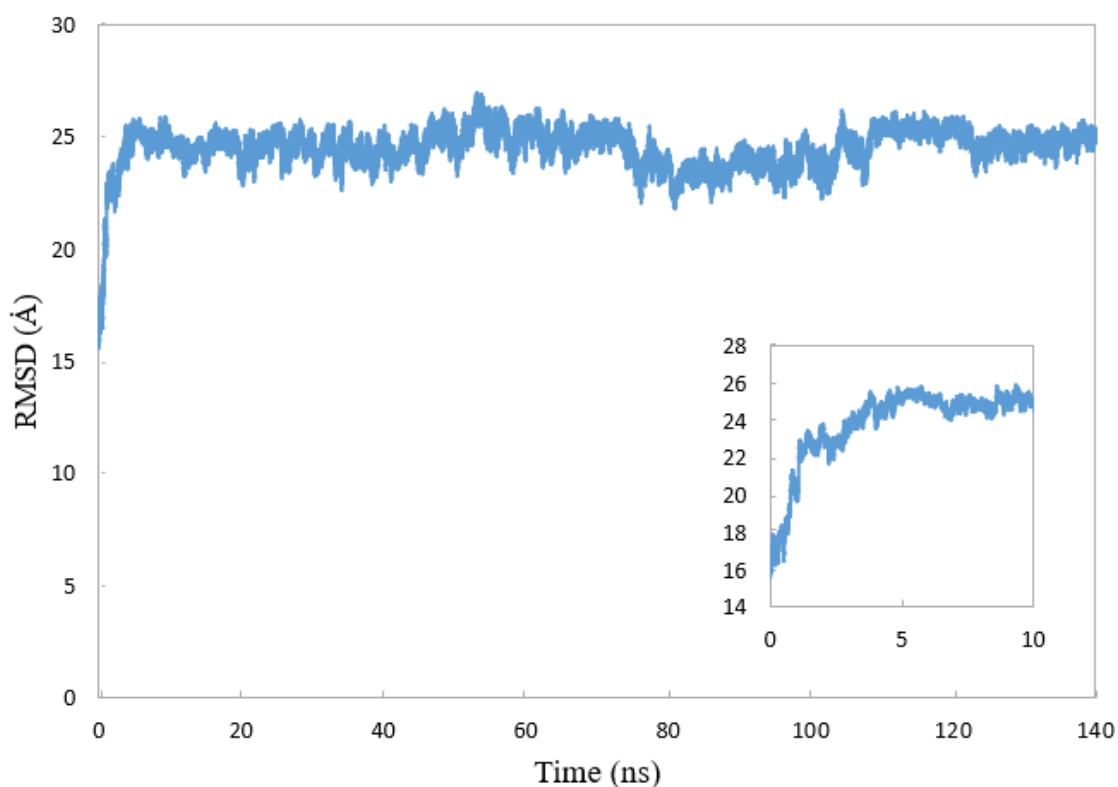
Figure 35. RMSD of P3HD chain solvated in water over time

Since only the trajectories from the production stage were considered on this analysis and the reference structure used is the initial input to be minimized, the initial value for the RMSD curve is not zero as it would be expected from Eq. (18) if all trajectories were to be used in the analysis. Nonetheless, the value increases until stabilization as the atoms move during the folding, distancing themselves from their initial position. The opposite trend is observed for the radius of gyration as the distance of the atoms from the center of mass of the polymer grows smaller with the phenomena. Thus, the initial value of 14.94Å observed at the beginning of the production decreases to equilibrium.

The total energy of the system through time is shown in Figure 36. It reaches an average value of -75.47 ± 21.73 kcal/mol. The rapid oscillations on the energy value and, hence, the high value of deviation is expected once the energies are based on the high-frequency oscillations of the atomic interactions.
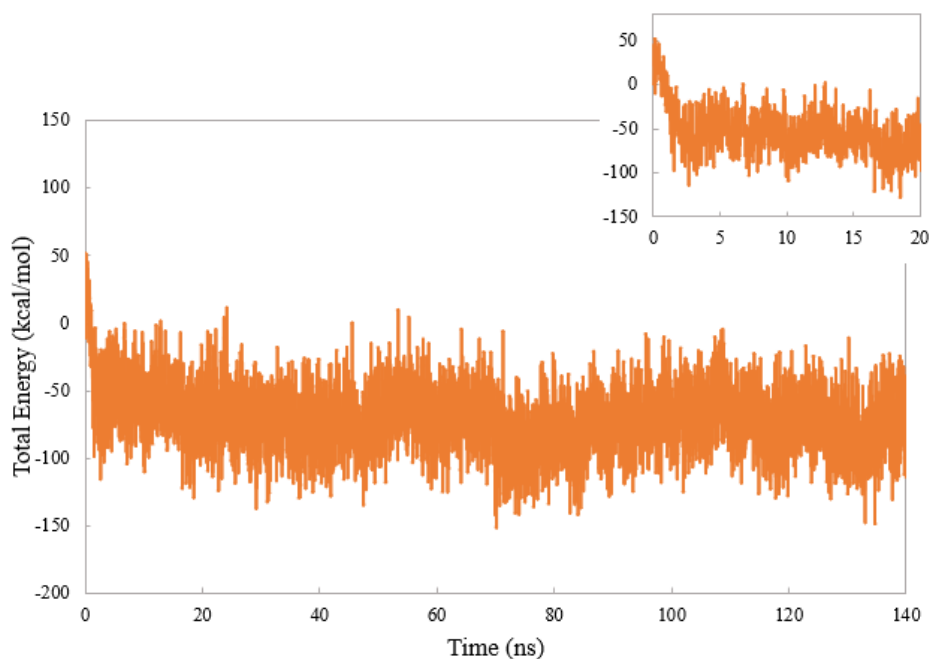
Figure 36. Total energy of the system over time

For a better picture of the energy of the system, it is interesting to focus the analysis on the van der Waals interactions over time, Figure 37.
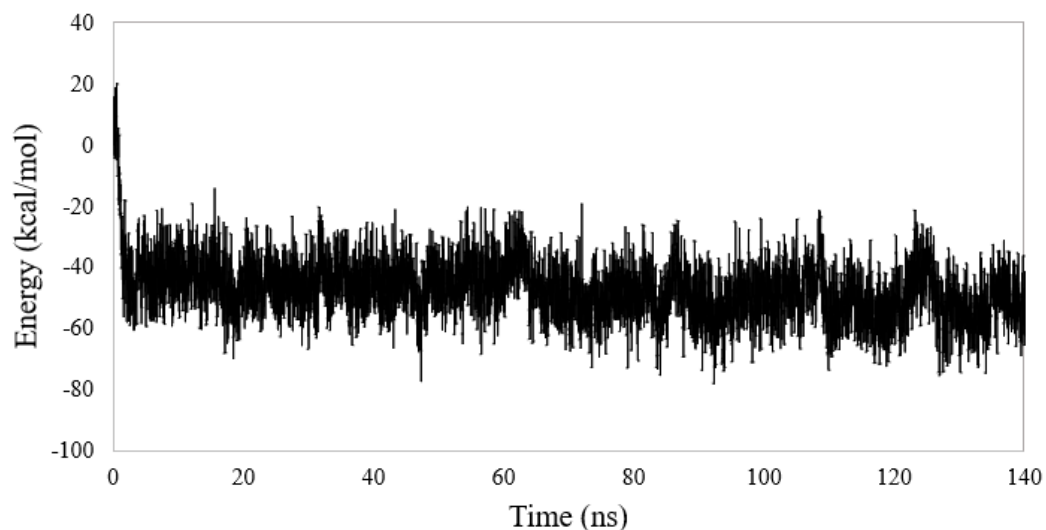
Figure 37. Van der Waals interaction energy for P3HD solvated in water system over time

The same trend of the total energy is observed for the van der Waals interaction energy as it reaches a stable value after few nanoseconds equal to -48.18 ± 8.88 kcal/mol, indicating

that the interaction plays an important role on the system dynamics. The reached plateau further confirms the attainment of a reasonably equilibrated conformation.

The van der Waals interactions are short-range weak interactions resulting from transient induced dipoles in the electron cloud surrounding the atom and one of the main driving forces of the hydrophobic effects.[82] Since it can be an attractive or repulsive force depending on the distance between atoms, the most favorable interaction occurs when the van der Waals surfaces are in contact without interpenetrating; it arrives at the absolute minimum energy configuration at the distance r_0, Figure 38. During the folding the atoms are attracted by the forces until they are distanced r_0 apart from each other, being r_0 a specific value of each element.
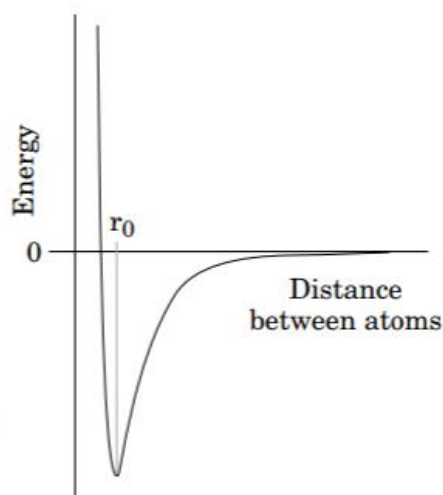
Figure 38. Graphical representation of Lennard Jones potential (12-6). The energy minimum is observed at an optimal distance r_0 between atoms. Taken from Brandt, M, 2016.[82]

Throughout, the results have shown to be consistent. The plateau for the energy has been reached at around the same time as the RMSD and radius of gyration confirming that the interaction is a major driver for the hydrophobic effect of the system simulated, as expected from literature.[83] Therefore, the simulated molecule of P3HD has initially shown to be a good model for the real biopolymer.

2. PHA-Membrane System

The final equilibrated polymer-lipid bilayer system is illustrated in Figure 39.
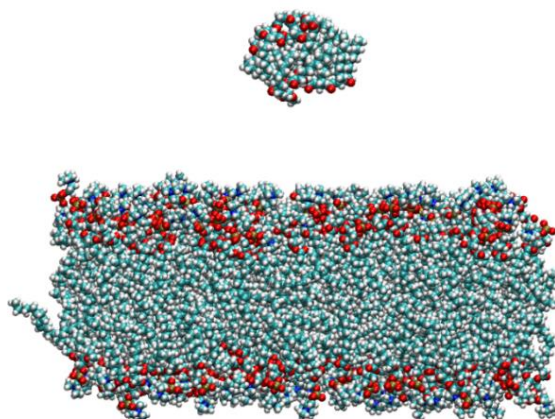
Figure 39. P3HD and Lipid Bilayer system. Solvated in water and ions (not represented for clarity).

Three snapshots of the final conformation obtained for the system with the US at distances 40Å, 20Å and 0Å is shown in Figure 40. From a visual analysis, it can be observed that the molecule was able to permeate and reach the center of the membrane.
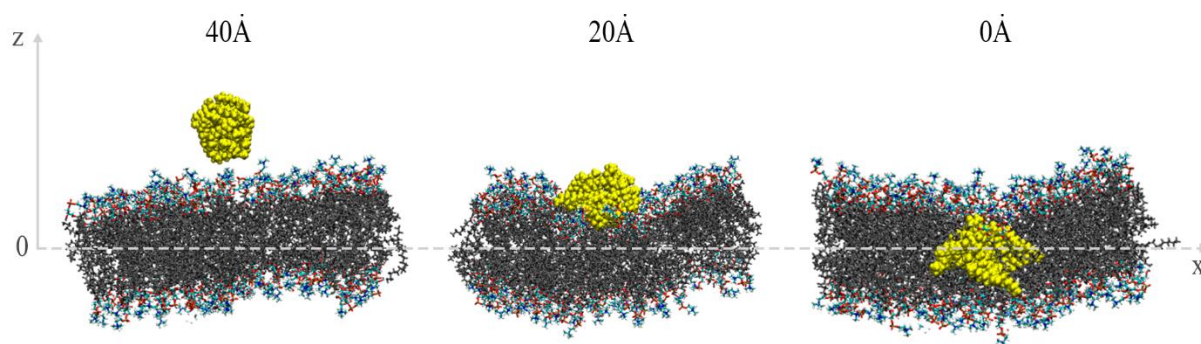
Figure 40. Permeation of lipid membrane by the P3HD, snapshots at distances of 40Å (non-interacting structures), 20Å (crossing of the hydrophilic surface) and 0Å (total permeation of the oligomer inside the membrane).

The histograms produced, Figure 41, showed overall a good overlap with the exception of three individual coordinates, 19Å, 23Å and 26Å those three distances represent the first contact with the membrane and the crossing of its surface.
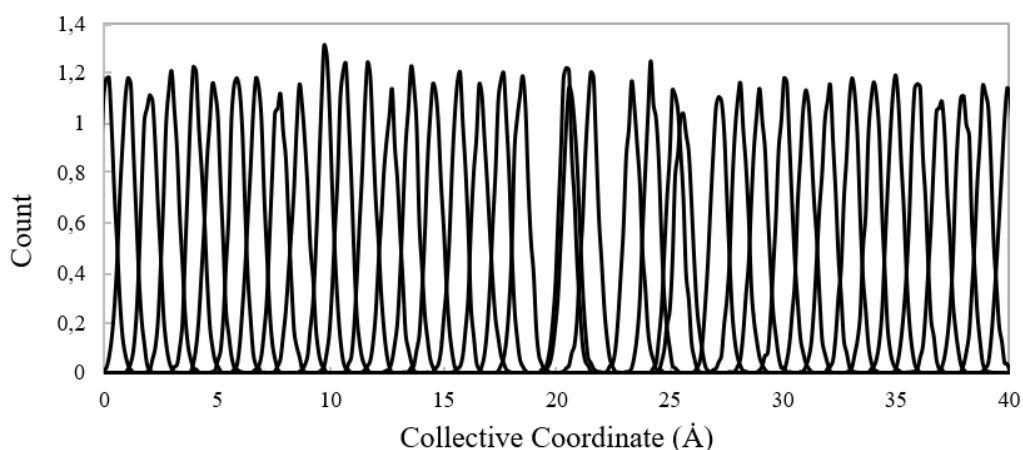
Figure 41. Histograms from the US applied to the permeation of P3HD inside the lipid bilayer

The lack of sampling on the aforementioned regions indicates that the chosen harmonic constant is not adequate in these ranges and a higher value should be used for those specific coordinates. Also, it indicates that the method used for the estimation of a k was not efficient and needs to be improved. A good alternative would have been performing trials for k also at the distance of the first contact between the membrane and the polymer. The PMF obtained is shown in Figure 42. The plot can be translated as the free energy profile over the reaction path; a low energy state is observed from 0Å to 5Å and it can be translated as a bigger affinity between the polymer and the interior of the membrane. As the distance from the center increases so does the value for the energy, reaching a maximum around 23Å. After it, the free energy decreases to a stable value, independent from distance.
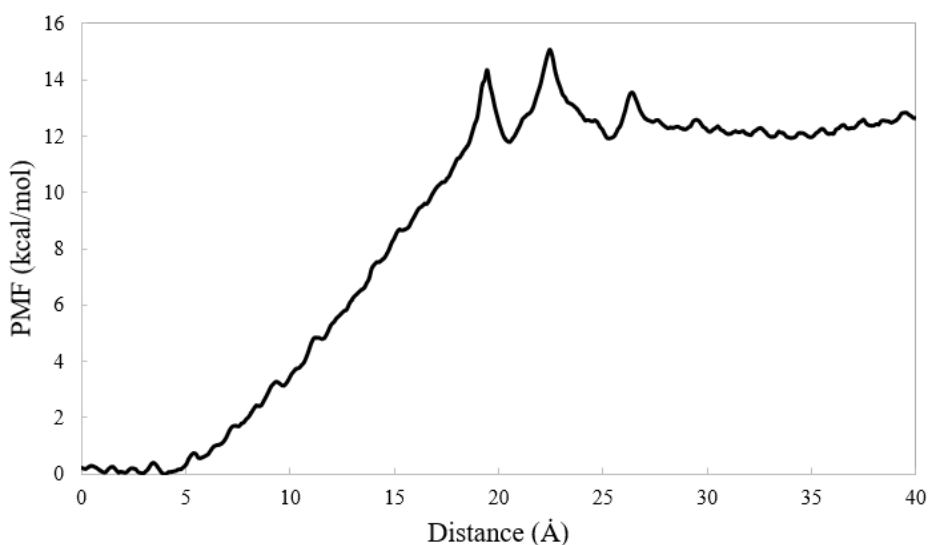
Figure 42. PMF of the membrane partitioning by a small chain of P3HD

53

WHAM works with the estimation of the free energy value from the region of overlapping (as a rule of thumb, highly sampled coordinates represent lower energy states). The three coordinates of flawed overlapping on the histograms were translated as high energy regions as better visualized in Figure 43. The smoothness of the PMF is then compromised, due to the lack of sampling. With further sampling using a higher restraining potential or longer simulations, a more precise PMF could be obtained.
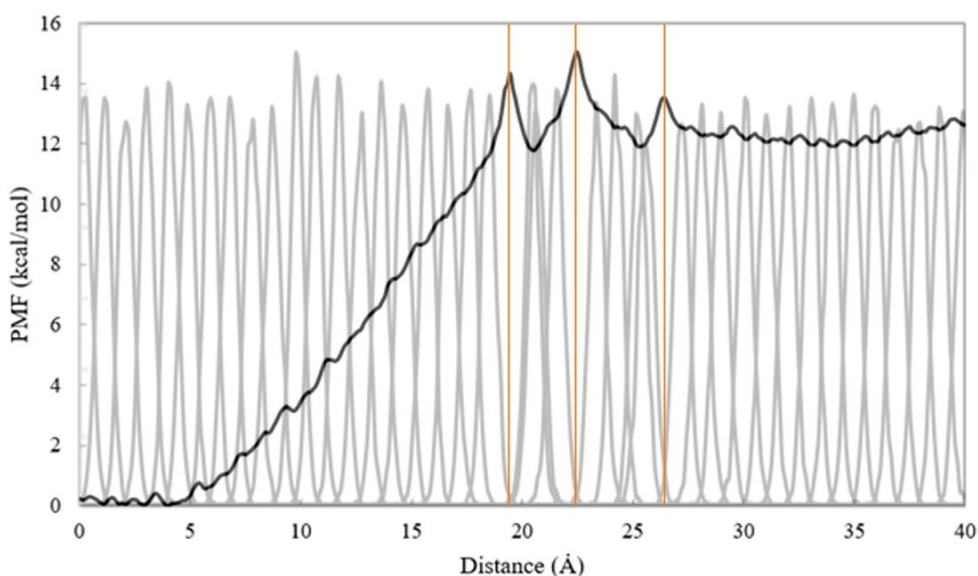
Figure 43. PMF versus sampling

Anyhow, an energy peak was expected from the nature of the membrane. The lipid membrane composed of amphiphilic phospholipids is characterized for having a highly polar region at the aqueous interface and a non-polar interior. A bigger affinity between the hydrophobic environment and molecule likewise is, therefore, expected. As the center of mass of both structures are distanced from each other and the P3HD finds itself surrounded by water the free energy increases. The energy peak around 23Å arises from the contact the oligomer with the most polarized region of the system.

At the same time, it can be observed that the radius of gyration of the molecule increases as it enters the non-polar center of the membrane and its affinity with the environment increases, Figure 44. As the short-range forces decrease closer to the center, the oligomer expands slightly.
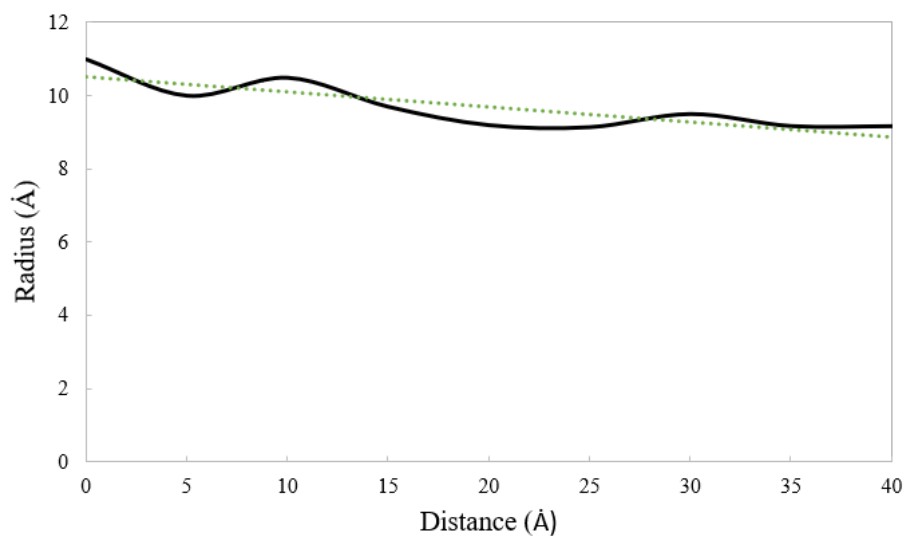
Figure 44. The radius of Gyration of P3HD computed every 5 windows. The trendline y= -0.0412x+10.523 (green) is included for better visualization.

The SASA average along the path, Figure 45, is another indicator for the membrane partitioning as the area exposed to the solvent (water) decreases with the insertion of the polymer into the bilayer.
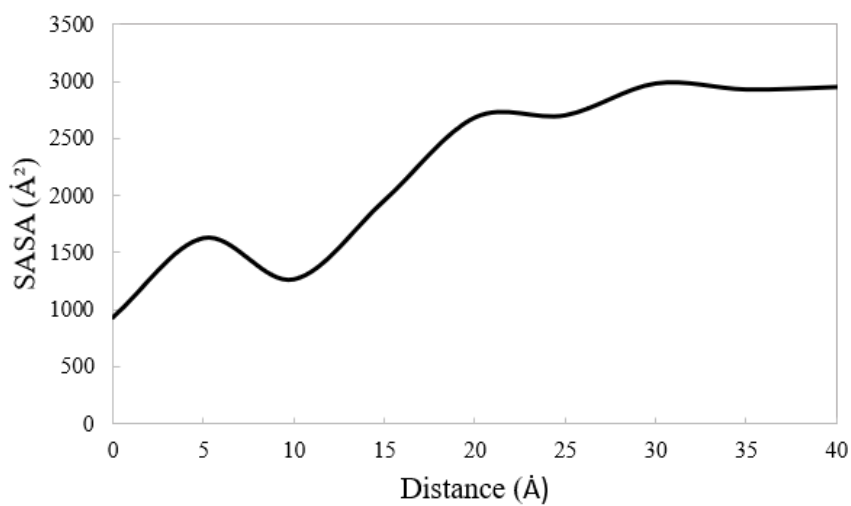
Figure 45. SASA of P3HB molecule calculated every 5 windows.

From the definition of the measurement, a value close to zero should be expected at the center of the membrane (0Å) and a crescent curve should be observed with distance. Instead, an exposed area of $930.72 \pm 139.81 Å^2$ is observed at 0Å and an oscillation is seen before the increase to a stable value. This can be explained by the fact that, while crossing the membrane,

the area accessible to the water oscillates according to the number of water molecules that also penetrate the membrane with the polymer.

3. Chitosan

The final structure after deacetylation can be seen in Figure 46.
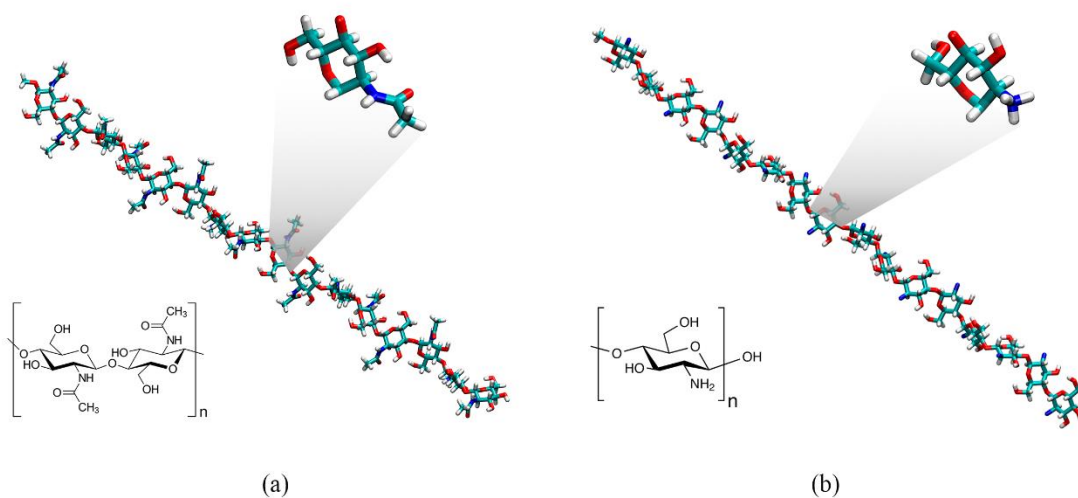
Figure 46. Deacetylation of chitin (a) to chitosan (b)

After the simulation procedure of equilibration and a production phase of 140 ns the affinity of the polymer chain with the solvent became clear as it did not present major changes in conformation, *i.e.*, did not fold. Such behavior was expected from chitosan since it is a hydrophilic biopolymer because of its charged moieties.

0 ns   140 ns
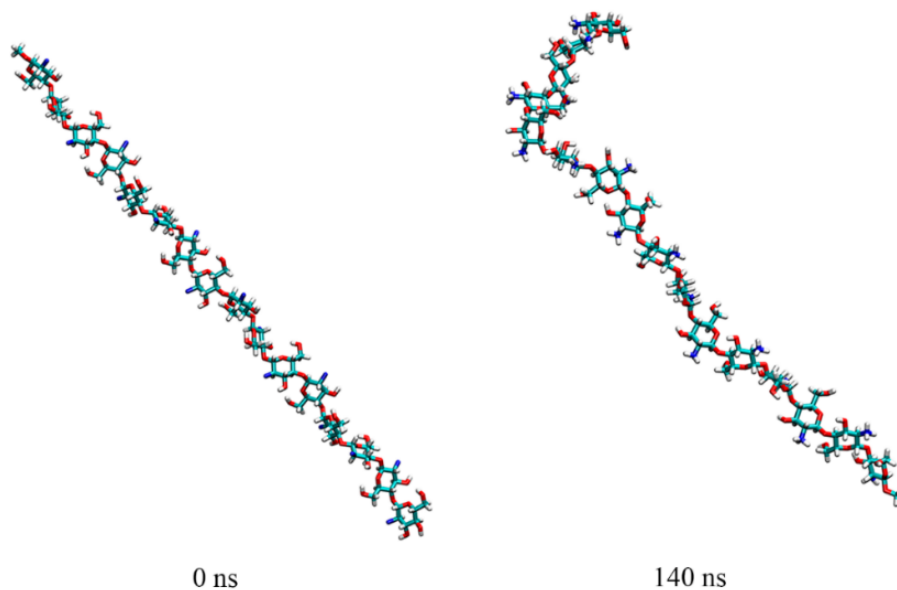
Figure 47. Chitosan chain over time - at 0 ns (left) and at 140 ns (right)

The small curve observed at the end of the chain only represents the last conformation pictured be the simulation but not necessarily the most stable among them all. During the MD simulation, the system oscillates around a stable value and, therefore, the conformation and values for one instant are not of interest, the average of the values sampled in an equilibrium state is. This becomes clearer when the RMSD for the polymer is evaluated.
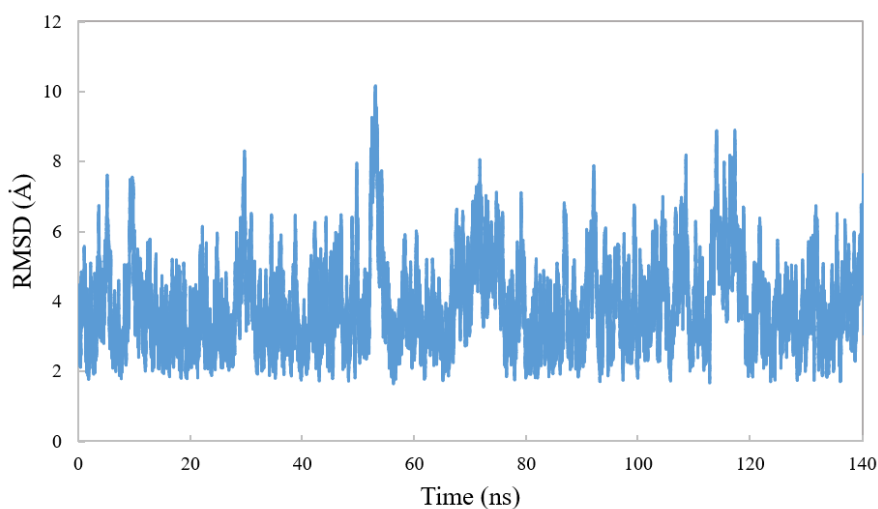
Figure 48. RMSD of chitosan solvated in water during production stage. Reference structure: initial input prior to minimization.

57

Different from the P3HD the RMSD curve for the chitosan in water did not show any significant changes over time and fluctuated around $3.87 \pm 1.25\text{Å}$. Once more, the initial value observed for the RMSD trend is not equal to zero because the conformations obtained during production are being compared to initial values prior to the minimization. The same trend is observed for the radius of gyration as is maintains a stable value of $23.91 \pm 0.61\text{Å}$. Since the chain does not fold, the distance of the atoms of the chain from its center of mass remains constant over time.
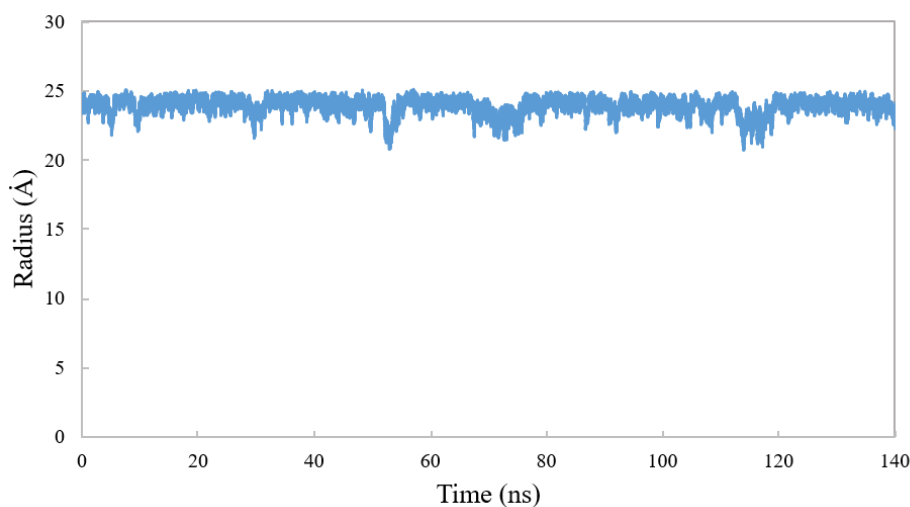
Figure 49. The radius of gyration of chitosan solvated in water during the production stage

It is also possible to notice from both graphs that the final structure pictured in Figure 47 (right) represents a lower radius of gyration and higher RMSD value than the average, 21.73Å and 7.13Å respectively.

The total energy of the system also shows a stable trend, shown in Figure 50, as expected since no solvation effect affects greatly its value.
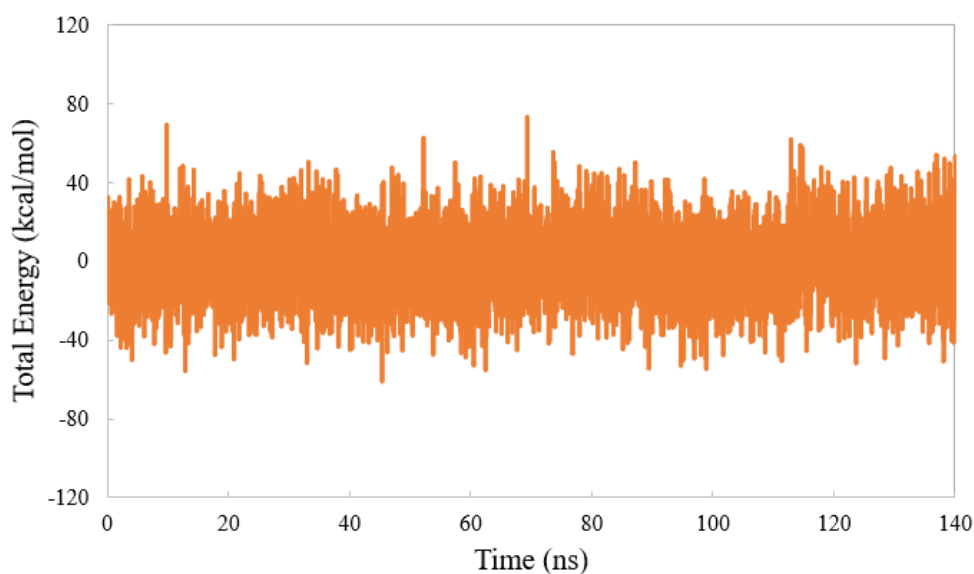
Figure 50. Total energy of the system during the production stage

In fact, if the energy related to the van der Waals interaction is analyzed, Figure 51, it also does not show any major changes.
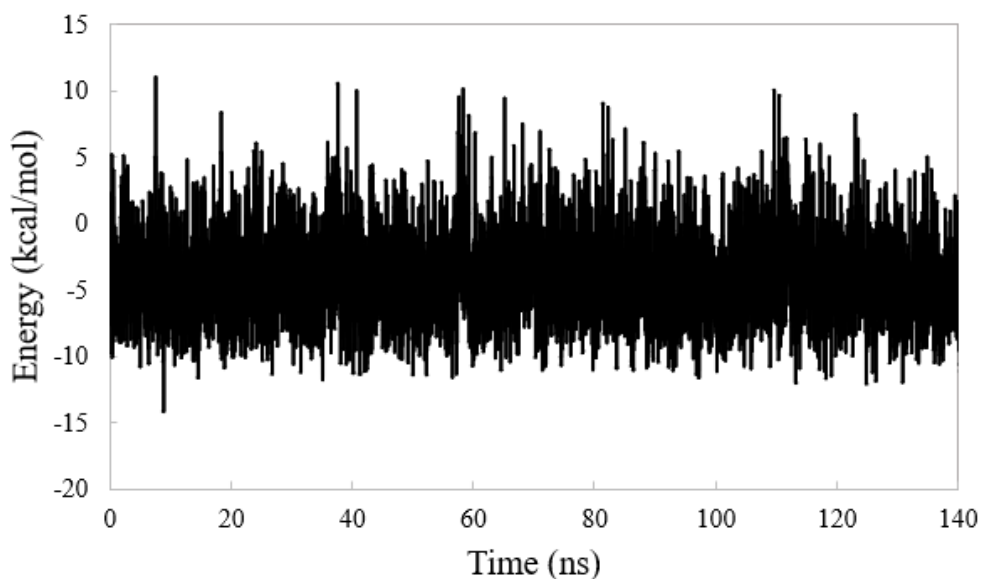
Figure 51. Van der Waals interaction energy for chitosan alone solvated in water system.

Even though the last conformation explored did not represent the average values, there was no need for further simulations since it was going to go through another equilibration cycle once inside the system with the lipid bilayer. As a matter of fact, the picture of the chitosan-membrane system can be observed in Figure 52, and a total linear chain for chitosan is observed.

59
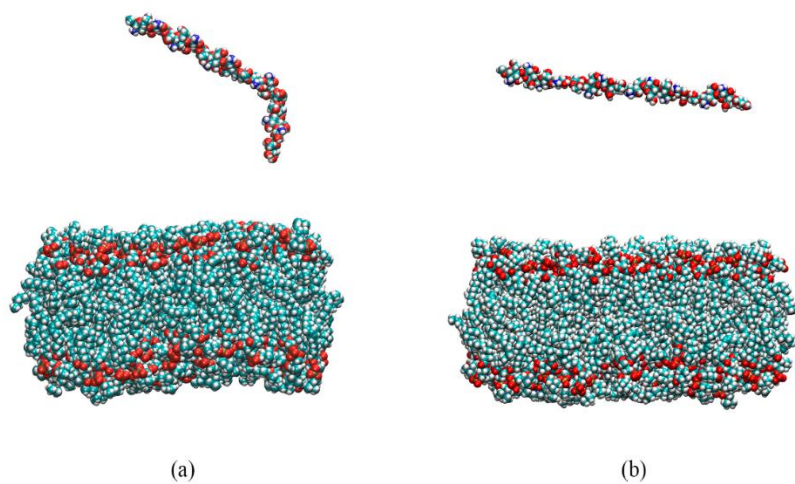
Figure 52. Chitosan + membrane system before (a) and after (b) the minimization, heating, and equilibration cycles.

4. Chitosan-membrane system

The final coordinates of three windows of the umbrella sampling of the system are shown in Figure 53.
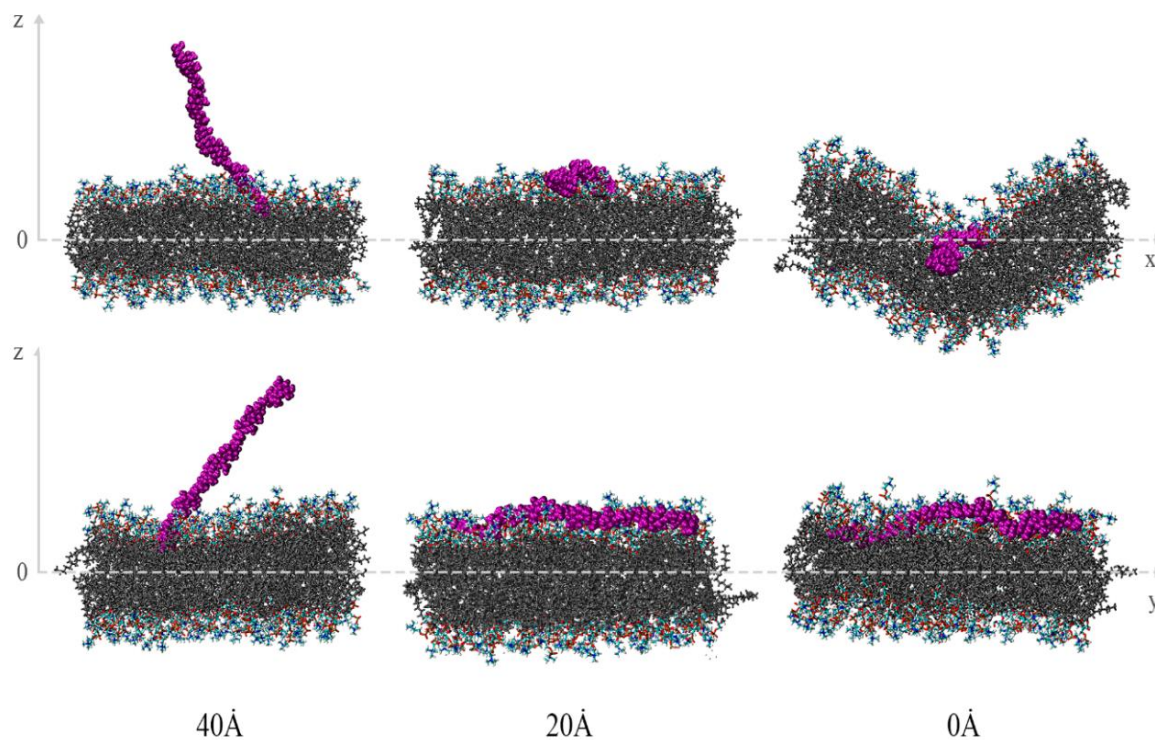
Figure 53. The US of chitosan-membrane system. Represented on x-z and y-z plane for better visualization

From the visual analysis of the trajectories, it can be said that the chitosan chain was not able to penetrate the membrane, instead, the lipid bilayer curved so to accommodate the chain at distance close to 0Å indicating that the energy required for the partition is of great magnitude and the simulations were not capable to predict such phenomenon. Also, at 40Å, even though the center of mass of both components distances 40Å from each other, it can be observed that the chain is actually in contact with the surface of the bilayer implying some favorable interaction between both parts.

The histograms obtained from all 40 windows overlap on most of the coordinates, Figure 54. With the exception of 0Å, 2Å and 3Å windows which were not sampled at all. To cover this region, a higher harmonic constant and hence a higher bias potential should've been used at them.
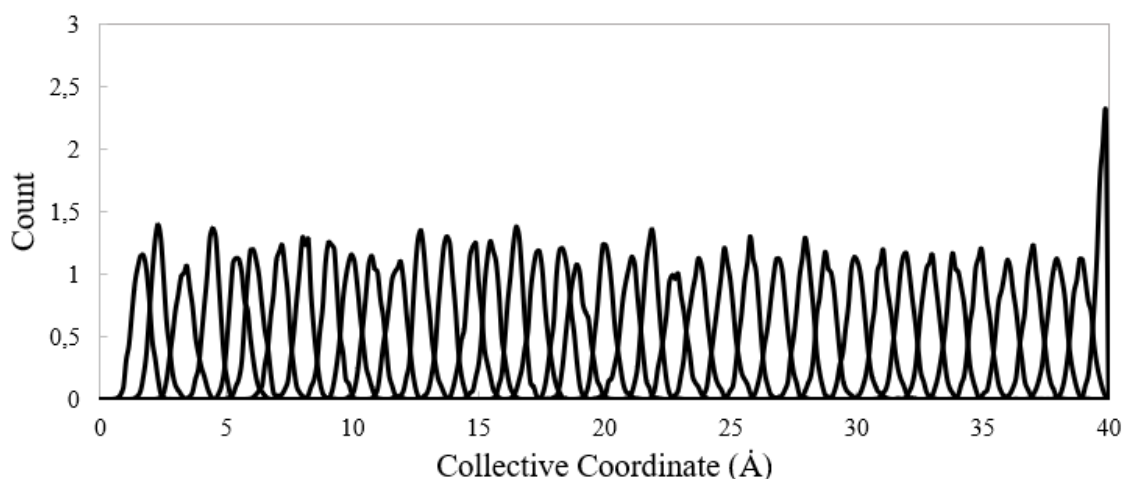
Figure 54. Histograms from the US for chitosan-membrane system

The PMF obtained with WHAM, Figure 55, shows the high energy state related to the position of the chain at the center of the membrane, infinite energy was accused at 0Å, 2Å, and 3Å.
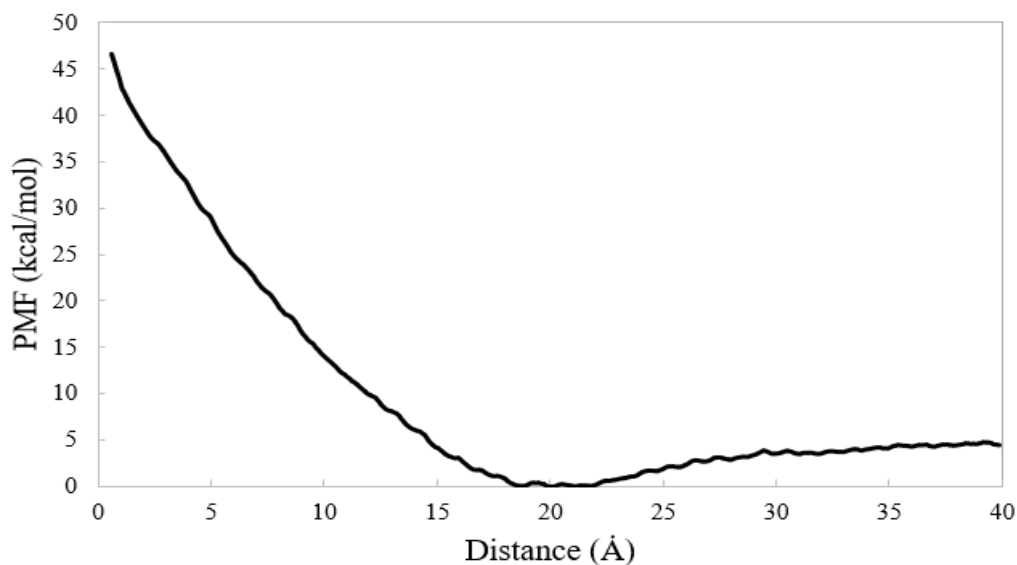
Figure 55. PMF of the membrane partitioning by a small chain of chitosan

Such profile is in accordance with what was expected since the charged polar chain will have more affinity with the surface of the membrane, polar part, than the non-polar center. The curve also shows a minimum point around 20Å accusing some energetically favored interaction between the chitosan and the hydrophilic surface of the lipid bilayer.

During the entire reaction path, the R_g of the chitosan chain, Figure 56, did not show any changes and slightly fluctuated around $23.76 \pm 0.57\text{Å}$. A foreseeable trend once the oligomer does not reach the hydrophobic session of the bilayer.
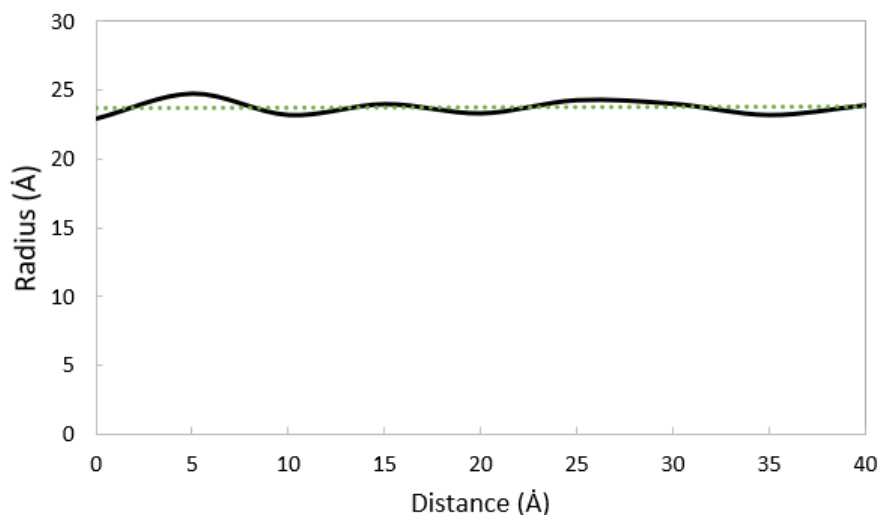
Figure 56. The radius of gyration of chitosan. The average values were calculated every 5 windows starting at 0Å. Trendline y = 0,0037x + 23,688 in green.

The solvent accessible surface area decreases with the approximation of the centers of mass of both structures, assuming negative values from 15Å, Figure 57.
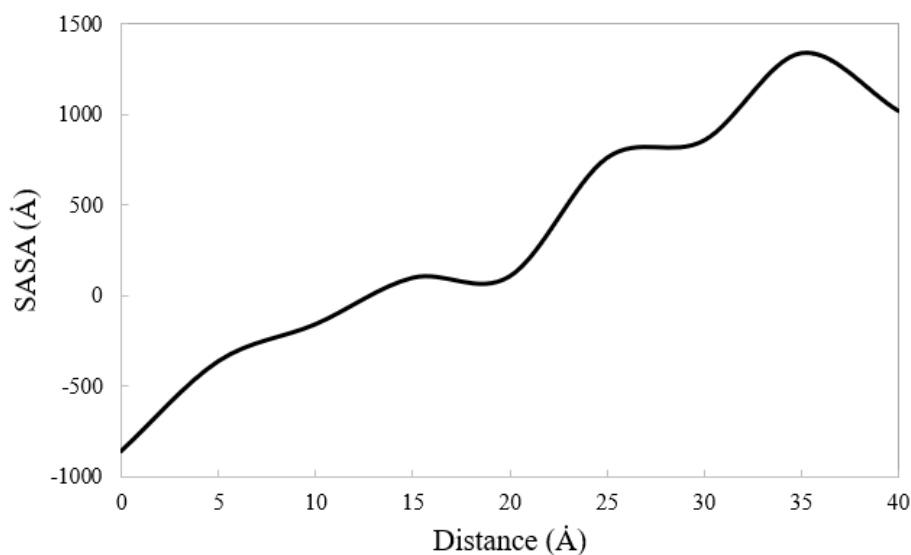
Figure 57. Solvent accessible surface area of chitosan. The average values were calculated every 5 windows starting at 0Å

A negative value of the area is accused due to the form in which the SASA is calculated. The algorithm for SASA relies on the linear combination of pairwise overlaps method. This method models each atom as a hard sphere and geometrically calculates their surfaces subtracting the how much of area is buried inside another sphere. From the mathematical formula, however, it is possible to notice that the area of one sphere buried in more than one other sphere is subtracted more than once hence underestimating the value for the area.[84]

$$A_i = P_1 S_1 + P_2 \sum_{j \varepsilon N(i)} A_{ij} + P_3 \sum_{\substack{j,k \varepsilon N(i) \\ k \varepsilon N(j) \\ k \neq j}} A_{jk} + P_4 \sum_{j \varepsilon N(i)} A_{ij} \left(\sum_{\substack{j,k \varepsilon N(i) \\ k \varepsilon N(j) \\ k \neq j}} A_{jk} \right)$$

(43)

Where A_i is the SASA for an individual atom. P_1 to P_4 are constants defined for each type of atom. For the sp3 hybridized N the first two values are such that the SASA of the entire molecule is underestimated.

5. Comparison

The DOPC lipid bilayer system can be conceptually divided into five regions. The bulk water at z > 25Å; the lipid headgroup region at aqueous interface at 25Å > z > 17Å; the dry transition region between the phosphate groups and the acyl chains 17Å > z > 10Å; the hydrophobic region z < 10Å; and very middle, the low tail density region at z = 0Å.[71] Sometimes the last two regions are described as one.[85] The z values are approximated and may vary with the model used for the bilayer.
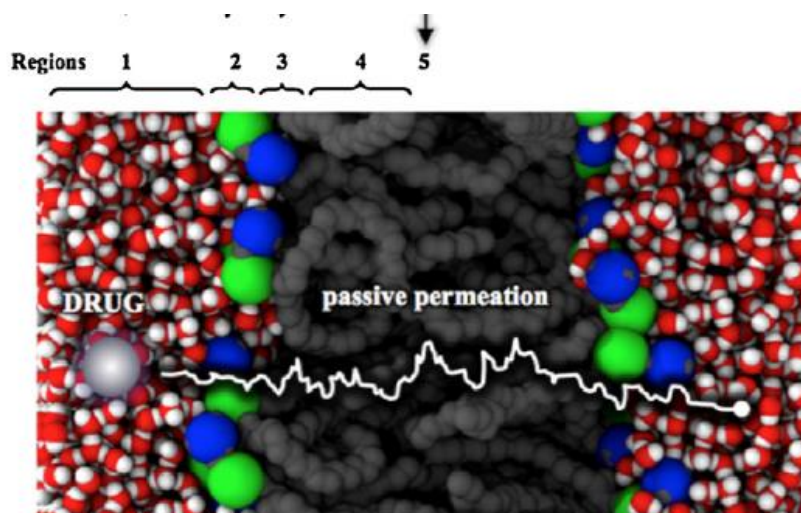
Figure 58. Regions of a DMPC lipid bilayer system. 1)Bulk water; 2)Head Group region; 3)Dry transition head group region; 4)Lipid tails; 5)Middle of the bilayer (low tail density). Adapted from Di Meo, F, *et. al.,*2016.[71]

The observed behavior of both P3HD and chitosan were in accordance with the regions. The free energy profiles of chitosan and P3HD were assumed symmetric and the energy landscape of the crossing is depicted in Figure 59.
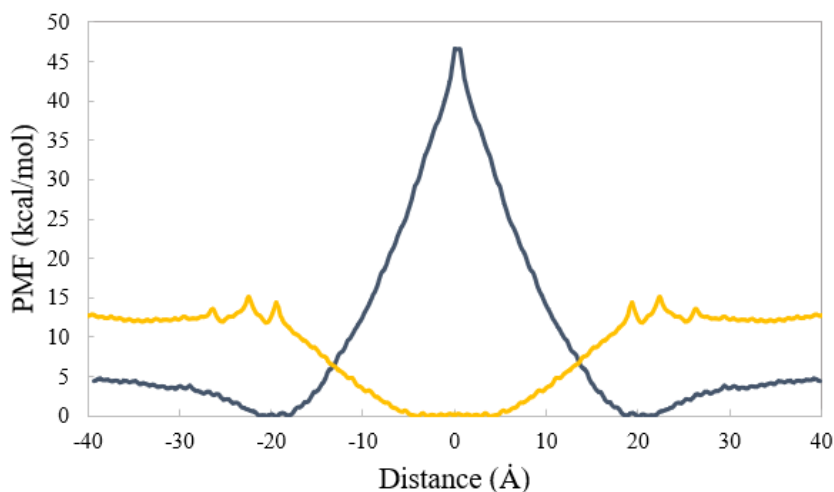
Figure 59. PMF of P3HD (yellow) and chitosan (gray) for the crossing of the membrane

Opposite behaviors were observed for the substances, which is in accordance with their opposite polarity. Chitosan is a polar, charged molecule. It has the lowest solubility in the most hydrophobic region of the membrane and because of that, a great energy barrier is involved on the crossing of the molecule. At the same time, P3HD, a hydrophobic molecule, finds its minimum energy state at the very center of the membrane. Comparing the results with other PMFs for polar and non-polar substances across a lipid bilayer found in literature similar profiles are observed indicating that the models and simulations are coherent. The PMF for water (polar) and O_2 (non-polar) are illustrated in Figure 60.
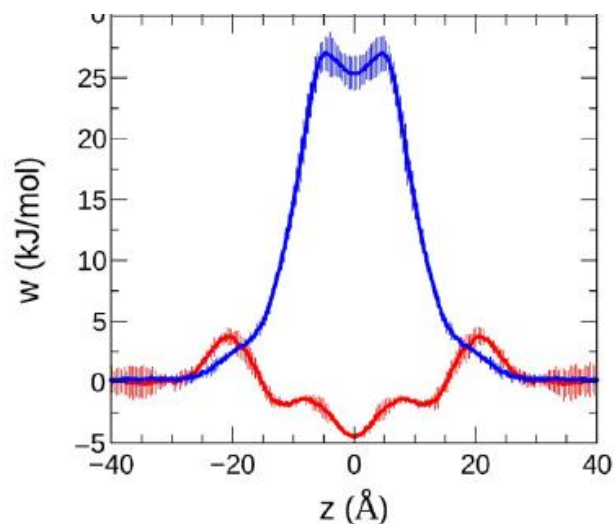
Figure 60. PMF profile of water (blue) and oxygen (red) across a DPPC bilayer. Taken from Awoonor-Williams, E., Rowley, C.N., 2016.[85]

The values for the energies cannot be directly compared due to the difference in the systems which includes but is not limited to differences in temperatures and type of bilayer. However, when comparing the profile obtained for water with the one for chitosan, it is possible to see that both involve a great energy barrier on the crossing. The non-polar molecules displayed the opposite trend with an energy minimum at the center for both O_2 and P3HD.

Analogous profiles were experimentally obtained for O_2 and chromium oxalate (polar).[86] PMF for acetamide, acetic acid, methanol, methyl acetate (only slightly polar), methylamine and water obtained with MD simulations[87] also are in accordance with chitosan profile attained in this work, while the ones for benzene and ethane resemble the one for P3HD. Moreover, the profiles are shown in Figure 61.
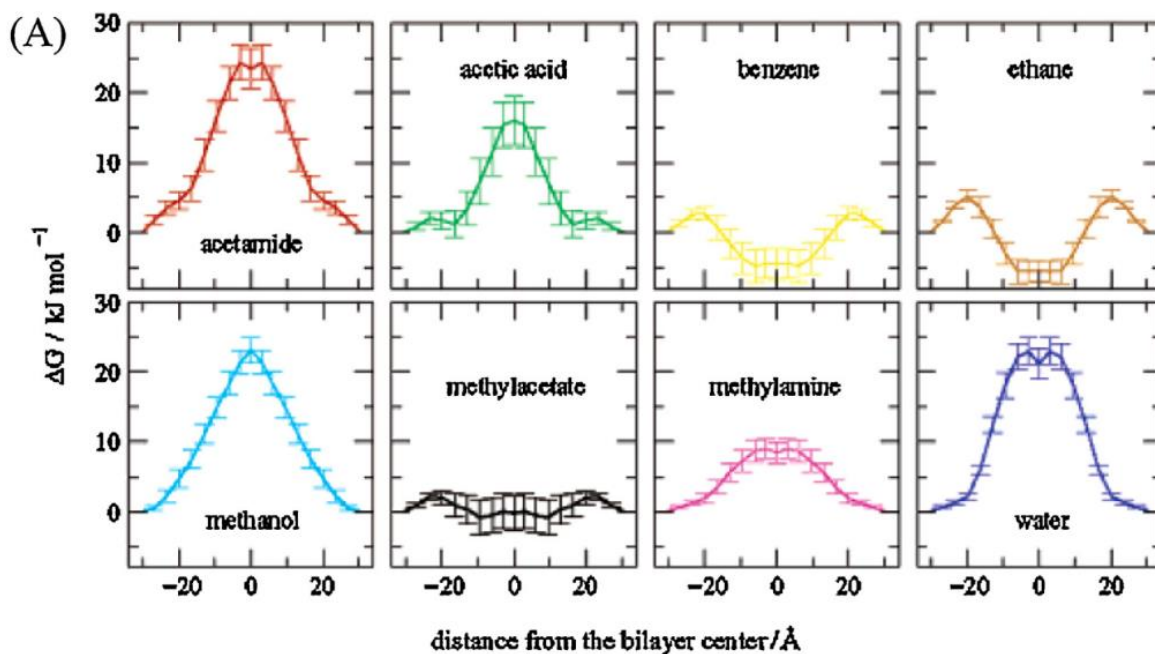
Figure 61. Free energy profiles of the crossing of a DMPC bilayer. Taken from Di Meo, F, *et. al.,*2016.[71]

Some studies characterize the influence of charges on the crossing.[88,89] The free energies for aspirin and ibuprofen simulated with GROMACS both in their neutral and anionic configurations pointed an influence of the charge state for the crossing[88] as it decreased the solubility of both drugs on the hydrophobic region of the membrane.
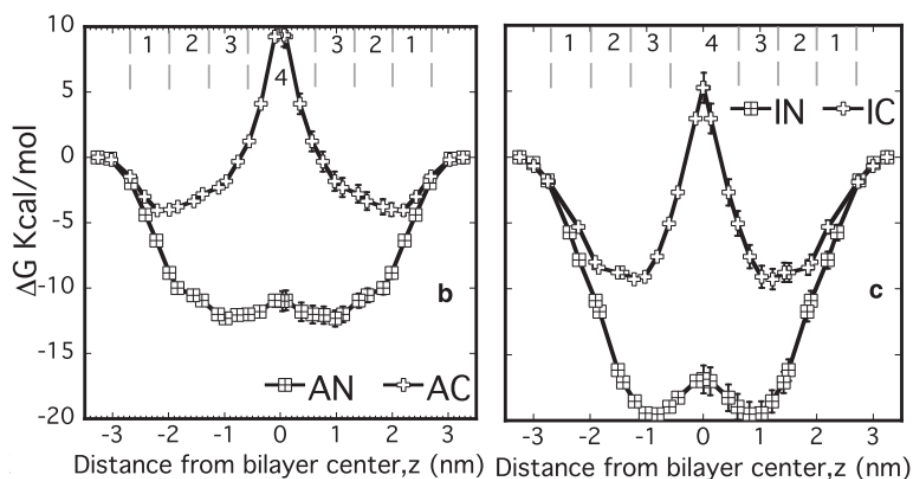
Figure 62. Free energy involved on the crossing of aspirin (left) charged (AC) and neutral (AN) and ibuprofen (right) charged (IC) and neutral (IN). Adapted from Boggara, M. *et al.,*2010.[88]

Such comparison was not performed for the chitosan since it was not the aim of the work, but it leads to the belief that the neutral structure, found in high pH solvents, could maybe have a smaller energy barrier on the crossing.

6. CONCLUSION

The objective of the study was to simulate two different polymers on an environment similar to the one found *in vivo* in drug delivery application. The molecular model for the mcl-PHA, P3HD, and chitosan chains was developed. Their interaction with a DOPC lipid bilayer was analyzed and the energy landscape for the membrane permeation was computed.

Molecular dynamics performed with AMBER tools were performed to investigate the phenomena. A P3HD chain was built with the RESP approach. When simulated in water it folded just like expected for a hydrophobic compound. The plateau reached on the RMSD, R_g and energy profiles indicated that the simulation had converged to a stable value. Opposite behavior was observed for the chitosan chain. A chitin chain was modeled with the aid of Glycam-WEB and then it was manually deacetylated to Chitosan. (DD equals to 100%). As a hydrophilic compound, it did not suffer any significant alterations on its structure and the RMSD, R_g, and energy profiles accused an equilibrated structure.

The free energy landscape of the membrane partitioning by the oligomers was computed with the PMF method through an enhanced sampling performed according to the US method, followed by the WHAM approach. The hydrophobic oligomer found its lowest energy state on the middle of the membrane at the low tail region of the bilayer, while the chitosan chain was not able to penetrate the DOPC lipid bilayer accusing a binding region at its very surface. Overall the approaches adopted proved to be adequate for the phenomena observed. Some flaws on the sampling were observed suggesting that a higher k or longer simulations should have been run in order to obtain more precise results.

The different behaviors obtained for each oligomer translate in the different possible applications of the biopolymers inside a delivery system which are being studied inside the GoBioNanoMat project. With the results obtained it will possible to tune coarse grain models for chitosan and P3HD that will enable the simulations of more practical systems, like the nanoparticles.

7. APPENDIX

A. P3HD Building Blocks

(a)   (b)
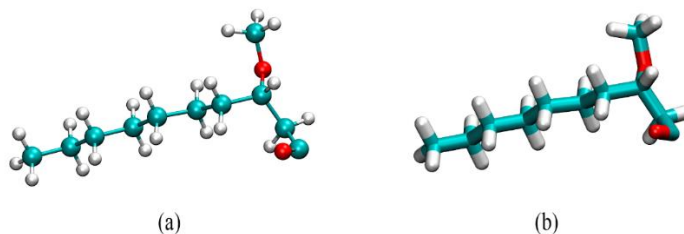
Table 2. 3PI detailing

I	Atom Name	Atom Type	X	Y	Z	Charge
1	C1	c3	-6.469	-4.954	1.559	-0.0247
2	H1	h1	-6.787	-5.292	2.549	0.064177
3	H2	h1	-6.144	-5.828	0.974	0.064177
4	H3	h1	-7.333	-4.503	1.052	0.064177
5	O1	os	-5.41	-4.039	1.759	-0.37659
6	C2	c3	-4.761	-3.611	0.558	0.195931
7	C3	c3	-3.43	-3.014	1.025	0.02251
8	H4	h1	-4.535	-4.491	-0.065	0.043429
9	C4	c3	-5.63	-2.653	-0.275	-0.04486
10	C5	c	-2.44	-2.77	-0.094	0.44417
11	H5	hc	-3.58	-2.081	1.578	-0.00326
12	H6	hc	-2.969	-3.707	1.738	-0.00326
13	H7	hc	-5.051	-2.361	-1.158	0.003417
14	H8	hc	-6.492	-3.215	-0.659	0.003417
15	C6	c3	-6.139	-1.415	0.475	0.080808
16	C7	c3	-7.103	-0.567	-0.365	0.022235
17	H9	hc	-6.635	-1.733	1.4	-0.0091
18	H10	hc	-5.294	-0.786	0.785	-0.0091
19	C8	c3	-7.628	0.674	0.367	-0.00156
20	H11	hc	-6.602	-0.257	-1.293	-0.00869

70

21	H12	hc	-7.954	-1.19	-0.678	-0.00869
22	C9	c3	-8.597	1.515	-0.472	0.023539
23	H13	hc	-8.128	0.362	1.295	-0.01066
24	H14	hc	-6.779	1.299	0.677	-0.01066
25	C10	c3	-9.126	2.756	0.258	0.051141
26	H15	hc	-8.097	1.828	-1.4	-0.0074
27	H16	hc	-9.446	0.89	-0.784	-0.0074
28	C11	c3	-10.093	3.591	-0.588	-0.07202
29	H17	hc	-9.627	2.444	1.185	-0.00831
30	H18	hc	-8.278	3.382	0.569	-0.00831
31	H19	hc	-10.45	4.468	-0.04	0.012151
32	H20	hc	-9.611	3.945	-1.506	0.012151
33	H21	hc	-10.97	3.003	-0.883	0.012151
34	O3	o	-2.592	-3.09	-1.255	-0.50503
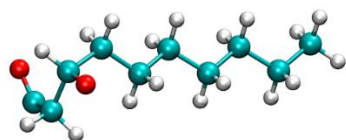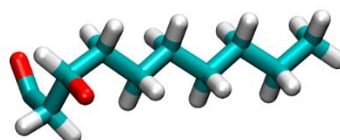
(a)   (b)

Table 3. 3PH detailing

I	Atom Name	Atom Type	X	Y	Z	Charge
1	O2	os	-1.331	-2.156	0.382	-0.47027
2	C12	c3	-0.264	-1.869	-0.571	0.033396
3	C13	c3	1.028	-2.014	0.234	-0.08048
4	H22	h1	-0.296	-2.63	-1.352	0.081515
5	C14	c3	-0.497	-0.498	-1.216	-0.0361
6	H23	hc	0.299	-0.349	-1.955	0.034207
7	H24	hc	-1.434	-0.563	-1.781	0.034207
8	C15	c3	-0.559	0.693	-0.25	0.035173

9	C16	c3	-0.903	2.009	-0.96	-0.00133
10	H25	hc	-1.304	0.493	0.53	0.006125
11	H26	hc	0.403	0.813	0.268	0.006125
12	C17	c3	-0.954	3.217	-0.018	0.003849
13	H27	hc	-0.164	2.197	-1.753	-0.0012
14	H28	hc	-1.871	1.903	-1.47	-0.0012
15	C18	c3	-1.295	4.533	-0.727	0.017203
16	H29	hc	-1.693	3.029	0.774	-0.00887
17	H30	hc	0.014	3.321	0.493	-0.00887
18	C19	c3	-1.343	5.744	0.214	0.047346
19	H31	hc	-0.556	4.72	-1.52	-0.00269
20	H32	hc	-2.264	4.432	-1.236	-0.00269
21	C20	c3	-1.678	7.055	-0.503	-0.07794
22	H33	hc	-2.083	5.558	1.004	-0.00639
23	H34	hc	-0.375	5.843	0.724	-0.00639
24	H35	hc	-1.707	7.898	0.195	0.014633
25	H36	hc	-0.934	7.287	-1.273	0.014633
26	H37	hc	-2.655	6.998	-0.997	0.014633
27	C21	c	2.272	-1.991	-0.631	0.822622
28	H38	hc	1.115	-1.245	1.008	0.054487
29	H39	hc	1.013	-2.972	0.766	0.054487
30	O5	o	2.285	-2.027	-1.844	-0.57024

(a)   (b)
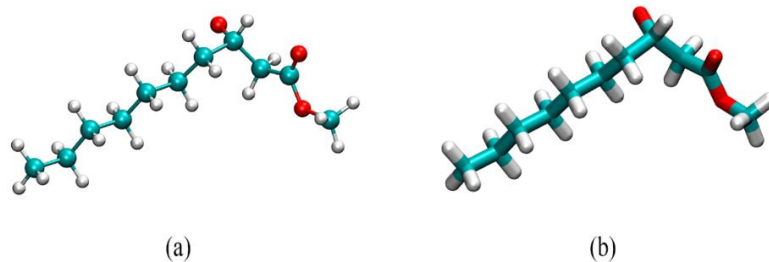
Table 4. 3PF detailing

I	Atom Name	Atom Type	X	Y	Z	Charge
1	O4	"os"	3.382	-1.953	0.14	-0.19658
2	C22	"c3"	4.672	-1.972	-0.544	0.092497
3	C23	"c3"	5.61	-2.699	0.42	-0.33728
4	H40	"h1"	4.56	-2.559	-1.458	0.126614
5	C24	"c3"	5.085	-0.547	-0.927	-0.1126
6	H41	"hc"	6.025	-0.623	-1.486	0.050422
7	H42	"hc"	4.335	-0.175	-1.635	0.050422
8	C25	"c3"	5.233	0.442	0.237	0.070565
9	C26	"c3"	5.509	1.874	-0.241	0.033821
10	H43	"hc"	4.32	0.431	0.846	-0.01672
11	H44	"hc"	6.049	0.128	0.901	-0.01672
12	C27	"c3"	5.682	2.879	0.905	0.001995
13	H45	"hc"	6.412	1.88	-0.869	-0.00651
14	H46	"hc"	4.687	2.202	-0.892	-0.00651
15	C28	"c3"	5.941	4.313	0.428	0.014559
16	H47	"hc"	4.783	2.864	1.538	-0.00775
17	H48	"hc"	6.511	2.555	1.55	-0.00775
18	C29	"c3"	6.119	5.318	1.573	0.041382
19	H49	"hc"	6.838	4.328	-0.209	-0.0029
20	H50	"hc"	5.11	4.638	-0.214	-0.0029
21	C30	"c3"	6.372	6.749	1.089	-0.07589
22	H51	"hc"	5.224	5.301	2.211	-0.00668
23	H52	"hc"	6.951	4.995	2.212	-0.00668

24	H53	"hc"	6.494	7.44	1.93	0.015491
25	H54	"hc"	7.28	6.806	0.478	0.015491
26	H55	"hc"	5.54	7.114	0.477	0.015491
27	C31	"c"	6.95	-3.037	-0.2	0.909896
28	H56	"hc"	5.773	-2.127	1.339	0.087215
29	H57	"hc"	5.144	-3.641	0.733	0.087215
30	O6	"os"	7.804	-3.504	0.735	-0.42694
31	O7	"o" 0	7.235	-2.932	-1.375	0.615711
32	C32	"c3"	9.099	-3.895	0.246	-0.04499
33	H58	"h1"	9.649	-4.245	1.119	0.092666
34	H59	"h1"	9.006	-4.693	-0.494	0.092666
35	H60	"h1"	9.609	-3.045	-0.214	0.092666

B. Chitosan Building Blocks
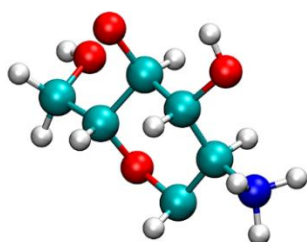
(a)
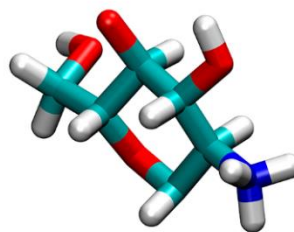
(b)

Table 5. 4YP detailing

I	Atom Name	Atom Type	X	Y	Z	Charge
1	C1	Cg	3.393597	1.433219	-1.73E-06	0.458
2	H1	H2	3.836153	0.891371	-0.85144	0
3	O5	Os	3.973148	0.998186	1.211861	-0.407
4	C5	Cg	5.40497	1.086914	1.247018	0.151
5	H5	H1	5.826396	0.450732	0.447832	0
6	C6	Cg	5.841874	0.516373	2.583014	0.33
7	H61	H1	6.928069	0.603351	2.654281	0
8	H62	H1	5.375304	1.096478	3.383289	0
9	O6	Oh	5.555191	-0.86163	2.707449	-0.674
10	H6O	Ho	4.614734	-0.98066	2.591697	0.436
11	C4	Cg	5.864983	2.527042	1.025197	0.326
12	H4	H1	5.534499	3.124203	1.887854	0
13	C3	Cg	5.270575	3.072542	-0.26184	0.123
14	H3	H1	5.705596	2.518256	-1.10785	0
15	O3	Oh	5.591417	4.453179	-0.36255	-0.641
16	H3O	Ho	6.553023	4.557945	-0.2325	0.448
17	C2	Cg	3.752599	2.889623	-0.27323	0.267
18	H2	H1	3.394918	3.508365	0.569531	0
19	N2	N3	3.238289	3.287702	-1.52991	-0.14

20	H1N	H	2.234022	3.175813	-1.5623	0.247
21	H2N	H	3.448941	4.258753	-1.71648	0.247
22	H3N	H	3.636837	2.734324	-2.27627	0.247
23	O4	Os	7.290822	2.570122	0.925058	-0.418
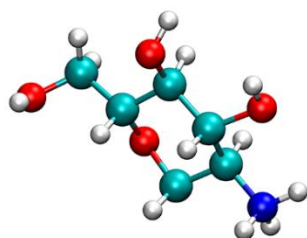
(a) (b)
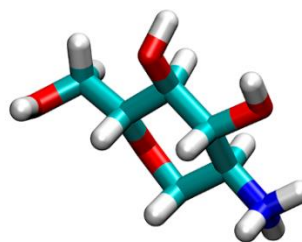
Table 6. 0YP detailing

I	Atom Name	Atom Type	X	Y	Z	Charge
1	C1	Cg	3.393597	1.433219	-1.73E-06	0.458
2	H1	H2	3.836153	0.891371	-0.85144	0
3	Os	O5	3.973148	0.998186	1.211861	-0.407
4	C5	Cg	5.40497	1.086914	1.247018	0.151
5	H5	H1	5.826396	0.450732	0.447832	0
6	C6	Cg	5.841874	0.516373	2.583014	0.33
7	H61	H1	6.928069	0.603351	2.654281	0
8	H62	H1	5.375304	1.096478	3.383289	0
9	O6	Oh	5.555191	-0.86163	2.707449	-0.674
10	H6O	Ho	4.614734	-0.98066	2.591697	0.436
11	C4	Cg	5.864983	2.527042	1.025197	0.326
12	H4	H1	5.534499	3.124203	1.887854	0
13	O4	Oh	7.290822	2.570122	0.925058	-0.663
14	H4O	Ho	7.616494	3.339364	1.394066	0.439
15	C3	Cg	5.270575	3.072542	-0.26184	0.123

16	H3	H1	5.705596	2.518256	-1.10785	0
17	O3	Oh	5.591417	4.453179	-0.36255	-0.641
18	H3O	Ho	6.553023	4.557945	-0.2325	0.448
19	C2	Cg	3.752599	2.889623	-0.27323	0.267
20	H2	H1	3.394918	3.508365	0.569531	0
21	N2	N3	3.238289	3.287702	-1.52991	-0.14
22	H1N	H	2.234022	3.175813	-1.5623	0.247
23	H2N	H	3.448941	4.258753	-1.71648	0.247
24	H3N	H	3.636837	2.734324	-2.27627	0.247

8. REFERENCES

1. Patra JK, Das G, Fraceto LF, et al. Nano based drug delivery systems: Recent developments and future prospects. *J Nanobiotechnology*. 2018;16(1):1-33. doi:10.1186/s12951-018-0392-8

2. Shi J, Votruba AR, Farokhzad OC, Langer R. Nanotechnology in drug delivery and tissue engineering: From discovery to applications. *Nano Lett*. 2010;10(9):3223-3230. doi:10.1021/nl102184c

3. Suri SS, Signh B, Fenniri H. Nanotechnology-based drug delivery systems. *J Occup Med Toxicol*. 2007;2(16):17-33. doi:10.1007/s12030-009-9028-2

4. National Nanotechnology Initiative. http://science.viniti.ru/index.php?option=com_content&task=view&id=1252&Itemid= 357. Accessed April 25, 2019.

5. Farokhzad OC, Langer R. Impact of Nanotechnology on Drug Delivery. *Am Chem Soc Nano*. 2009;3(1):16-20.

6. Jong, Wim H De; Borm PJA. Drug delivery and nanoparticles: Applications and hazards. *Int J Nanomedicine*. 2008;3(2):133-149.

7. Chen R, Riviere JE. Biological and environmental surface interactions of nanomaterials: characterization, modeling, and prediction. *WiREs Nanomedicine and Nanobiotechnology*. 2017;9(3):1-31. doi:10.1002/wnan.1440

8. Barouti G, Jaffredo CG, Guillaume SM. Advances in drug delivery systems based on synthetic poly(hydroxybutyrate) (co)polymers. *Prog Polym Sci*. 2017;73:1-31. doi:10.1016/j.progpolymsci.2017.05.002

9. Kumaresh S. Soppimath , Tejraj M. Aminabhavi * ARK, Aa, A, A WER. Biodegradable polymeric nanoparticles as drug delivery devices. *J Control Release*. 2001;70:1-20. doi:10.1007/BF00504324

10. James EN, Hanna C, Nair LS. *Nanobiomaterials for Tissue Engineering Applications*. Elsevier Inc.; 2014. doi:10.1016/B978-0-12-397157-9.00018-7

11. Pillai O, Panchagnula R. Polymers in drug delivery. *Curr Opin Chem Biol*. 2001;5:447-451.

12. Park K. Facing the truth about nanotechnology in drug delivery. *ACS Nano*. 2013;7(9):7442-7447. doi:10.1021/nn404501g

13. Hubbell JA, Chilkoti A. Nanomaterials for Drug Delivery. 2012;337(July):303-305.

14. Yang L, Zhang L, Webster TJ. Nanobiomaterials: State of the art and future trends. *Adv Eng Mater*. 2011;13(6):197-217. doi:10.1002/adem.201080140

15. GoBioNanoMat Project. 2017. http://gonanobiomat.eu/. Accessed April 25, 2019.

16. van de Poel I, Robaey Z. Safe-by-Design: from Safety to Responsibility. *Nanoethics*. 2017;11(3):297-306. doi:10.1007/s11569-017-0301-x

17. Sips A, Noorlander C, Lehmann HC, Hohener K. Safe-by-Design Concept. *NANoREG*. 2015;(c):1-32.

18. Henriques CYH. Molecular Modeling of Poly L-Latic Acid with Lipid Bilayer. 2018.

19. GoNanoBioMat Project. *SbD Polyhydroxyalkanoate (PHA)*.; 2018.

20. Tan G-YA, Chen C-L, Li L, et al. Start a Research on Biopolymer Polyhydroxyalkanoate (PHA): A Review. *Polymers (Basel)*. 2014;6:706-754.

21. Shrivastav A, Kim H, Kim Y. Advances in the Applications of Polyhydroxyalkanoate Nanoparticles for Novel Drug Delivery System. *Biomed Res Int*. 2013;2013(Mcl).

22. Sasikumar P, Ayyasamy PM, Centre D, Nadu T, Nadu T. Design and Characterization of Poly-Hydroxy Butyric Acid (PHB) Based Polymeric Nanoparticles for Controlled Release of Doxorubicin for Cancer Treatment. *Int J Curr Microbiol Appl Sci*. 2015;4(12):311-317. http://www.ijcmas.com/vol-4-12/P. Sasikumar and P.M. Ayyasamy.pdf.

23. Errico C, Bartoli C, Chiellini F, Chiellini E. Poly (hydroxyalkanoates)-Based Polymeric Nanoparticles for Drug Delivery. *J Biomed Biotechnol*. 2009;2009. doi:10.1155/2009/571702

24. Blessia TF, Sharmila DJS, Samian MR, Arsad H, Jamil NF. STRUCTURAL ANALYSIS AND MOLECULAR DYNAMICS STUDY OF PHB SYNTHASE. *Int Res J Pharm*. 2012;3(2):251-256.

25. Vilar G, Tulla-puche J, Albericio F. Polymers and Drug Delivery Systems. *Curr Drug Deliv*. 2012;9(4):1-28.

26. Ali A, Ahmed S. A review on chitosan and its nanocomposites in drug delivery. 2018;109:273-286.

27. Nilsen-Nygaard J, Strand SP, Vårum KM, Draget KI, Nordgård CT. Chitosan: Gels and Interfacial Properties. *Polymers (Basel)*. 2015;7:552-579.

28. Elgadir MA, Uddin S, Ferdosh S, et al. Impact of chitosan composites and chitosan nanoparticle composites on various drug delivery systems : A review. *J Food Drug Anal*. 2015;23:619-629.

29. Dutta J, Tripathi S, Dutta PK. *Progress in Antimicrobial Activities of Chitin, Chitosan and Its Oligosaccharides: A Systematic Study Needs for Food Applications*. Vol 18.; 2012. doi:10.1177/1082013211399195

30. Agnihotri SA, Mallikarjuna NN, Aminabhavi TM. Recent advances on chitosan-based micro- and nanoparticles in drug delivery B. 2004;100:5-28. doi:10.1016/j.jconrel.2004.08.010

31. Xu Y, Du Y. Effect of molecular structure of chitosan on protein delivery properties of chitosan nanoparticles. 2003;250:215-226.

32. Janes KA, Fresneau MP, Marazuela A, Fabra A. Chitosan nanoparticles as delivery systems for doxorubicin. 2001;73:255-267.

33. Matica A, Ostafe V. Antibacterial Properties of Chitin and Chitosans. 2017;26(1):39-54.

34. Roducts P, Matica A, Ostafe V. Biodegradability of Chitosan Based Products. 2017;26(1):75-86.

35. Tiyaboonchai W. Chitosan Nanoparticles : A Promising System for Drug Delivery. 2003;11(3):51-66.

36. Rampino A, Borgogna M, Blasi P, Bellich B, Cesàro A. Chitosan nanoparticles : Preparation , size evolution and stability. 2013;455:219-228.

37. Bugnicourt L, Ladavière C. Interests of chitosan nanoparticles ionically cross-linked with tripolyphosphate for biomedical applications. *Prog Polym Sci*. 2016;60:1-17.

38. Matica A, Menghiu G, Ostafe V. Toxicity of Chitosan Based Products. *New Front Chem*. 2017;26(1):65-74.

39. Crucho CIC, Teresa M. Polymeric nanoparticles : A study on the preparation variables and characterization methods. 2017;80:771-784.

40. Qi L, Xu Z, Jiang X, Hu C, Zou X. Preparation and antibacterial activity of chitosan nanoparticles. 2004;339:2693-2700. doi:10.1016/j.carres.2004.09.007

41. Namba AM, Silva VB da, Silva CHTP da. Dinâmica molecular: teoria e aplicações em planejamento de fármacos TT - Molecular dynamics: theory and applications in drug design. *Eclética Química*. 2008;33(4):13-24. doi:10.1590/S0100-46702008000400002

42. Kumar H, Maiti PK. Introduction to Molecular Dynamics Simulation. 2017;23:161-197. doi:10.1007/978-93-86279-50-7_6

43. Perilla JR, Goh BC, Cassidy CK, et al. Molecular dynamics simulations of large macromolecular complexes. *Curr Opin Struct Biol*. 2015;31:64-74. doi:10.1016/j.sbi.2015.03.007

44. B. R, T. S. Implicit solvent models. *Biophys Chem*. 1999;78(1-2):1-20.

45. Case DA, Cheatham TE, Darden T, et al. The Amber Biomolecular Simulation Programs. *Jounal Comput Chem*. 2005;26(16):1668-1688. doi:10.1371/journal.pone.0178059

46. González MA. Force fields and molecular dynamics simulations. *École thématique la Société Française la Neutron*. 2011;12:169-200. doi:10.1051/sfn/201112009

47. Sant'Anna CMR. Molecular modeling methods in the study and design of bioactive compounds: An introduction. *Rev Virtual Química*. 2009;1(1). doi:10.5935/1984-6835.20090007

48. Ander Malthe. Getting started with molecular dynamics modeling. 2017:28-48. https://www.uio.no/studier/emner/matnat/fys/FYS4460/v17/notes/md-2016-python.pdf.

49. Gelpi J, Hospital A, Goñi R, Orozco M. Molecular dynamics simulations: advances and applications. *Adv Appl Bioinforma Chem*. 2015:37. doi:10.2147/aabc.s70333

50. Schlick T. *Molecular Modeling and Simulation: An Interdisciplicary Guide*. Vol 21.; 2010. doi:10.1007/978-0-387-75847-3

51. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J Phys Chem*. 1993;97:10269-10280.

52. Leland B, Paul D, Krueger B, Walker R. Amber Tutorials: Setting up an Advanced System. http://ambermd.org/tutorials/advanced/tutorial1/section1.htm. Accessed June 17, 2019.

53. Casalini T. Molecular Modeling of Transport Phenomena in Hydrogels. 2012.

54. Comell WD, Cieplak P, Bayly CI, Kollman PA. Application of RESP charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free energies of Solvation. 1993;(7):9620-9631.

55. Finnerty J. Molecular dynamics meets the physical world: Thermostats and barostats. 2011.

56. Cavallari M, Ghio C, Monti S, Ferrario M, Maritan A, Carloni P. Partially folded states of HIV-1 protease: Molecular dynamics simulations and ligand binding. *J Mol Struct THEOCHEM*. 2006;769(1-3):111-121. doi:10.1016/j.theochem.2006.04.042

57. Radius of Gyration. https://www.eng.yale.edu/polymers/docs/classes/polyphys/lecture_notes/2/handout2_wse3.html. Accessed June 18, 2019.

58. Ali S, Hassan M, Islam A, Ahmad F. A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States. *Curr Protein Pept Sci*. 2014;15(5):456-476. doi:10.2174/1389203715666140327114232

59. Walker R, Steinbrecher T, McGee D, III BM, Swails J. Amber advanced tutorials: MM-PBSA. http://ambermd.org/tutorials/advanced/tutorial3/. Published 2006. Accessed June 20, 2019.

60. Yang M, Yang L, Gao Y, Hu H. Combine umbrella sampling with integrated tempering method for efficient and accurate calculation of free energy changes of complex energy surface. *J Chem Phys*. 2014;141(4).

61. Kästner J. Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci*. 2011;1(6):932-942. doi:10.1002/wcms.66

62. Chipot C, Pohorille A. *Free Energy Calculations - Theory and Applications*. Springer; 2006. http://scc.acad.bg/ncsa/articles/library/Library2016_Supercomputers-at-Work/Free_Energy/FreeEnergy_Calculation_Chemistry_and_Biology.pdf.

63. Otter WK den. Revisiting the Exact Relation between Potential of Mean Force and Free-Energy Profil. *J Chem Theory Comput*. 2013;9:3861-3865. doi:10.1021/ct300392f

64. Hub JS, De Groot BL, Van Der Spoel D. G_whams - A free Weighted Histogram Analysis implementation including robust error and autocorrelation estimates. *J Chem Theory Comput*. 2010;6(12):3713-3720. doi:10.1021/ct100494z

65. Leach AR. *Molecular Modeling Principles and Applications*. Second Edi. Pearson Education; 2001.

66. Bochicchio D, Panizon E, Ferrando R, Monticelli L, Rossi G. Calculating the free energy of transfer of small solutes into a model lipid membrane: Comparison between metadynamics and umbrella sampling. *J Chem Phys*. 2015;143(14). doi:10.1063/1.4932159

67. Shankar K, Djamal B, Swendsen RH, Kollman PA, Rosenberg JM. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The method. *J Comput Chem*. 1992;13(8):1011-1021. doi:doi:10.1002/jcc.540130812

68. Lemkul JA. Gromacs Tutorial: Umbrella Sampling. http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/umbrella/05_pull.html. Accessed June 2, 2019.

69. Okamoto Y. Generalized-ensemble algorithms: Enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J Mol Graph Model*. 2004;22(5):425-

439. doi:10.1016/j.jmgm.2003.12.009

70. Souaille M, Roux B. Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations. *Comput Phys Commun*. 2001;135(1):40-57. doi:10.1016/S0010-4655(00)00215-0

71. di Meo F, Fabre G, Berka K, et al. In silico pharmacology: Drug membrane partitioning and crossing. *Pharmacol Res*. 2016;111:471-486.

72. Perez A, Morrone JA, Simmerling C, Dill KA. Advances in free-energy-based simulations of protein folding and ligand binding. *Curr Opin Struct Biol*. 2016;36:25-31. doi:10.1016/j.sbi.2015.12.002

73. Gumbart JC. Methods for calculating Potentials of Mean Force. *NAMD Tutor*. 2017. http://www.ks.uiuc.edu/Training/Tutorials/.

74. Shinoda W. Permeability across lipid membranes ☆. *BBA - Biomembr*. 2016;1858(10):2254-2265. doi:10.1016/j.bbamem.2016.03.032

75. Hansson T, Oostenbrink Ch, Gunsteren WF van. Molecular Dynamics Simulations. *Curr Opin Struct Biol*. 2002;12:190-196. doi:10.1016/B978-0-12-409547-2.14274-X

76. 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, PAK. Amber 2018 Reference Manual. *Univ California, San Fr*. 2018.

77. Koller M. Cyanobacterial Polyhydroxyalkanoate Production: Status Quo and Quo Vadis? *Curr Biotechnol*. 2015;4(4):464-480. doi:10.2174/2211550104666150917010849

78. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput*. 2013;9(7):3084-3095.

79. Humphrey W, Dalke A, Schulten K. VMD Visual Molecular Dynamics. *J Molec Graph*. 1996;14:33-38. http://www.ks.uiuc.edu/Research/vmd/.

80. Group W. GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA. http://glycam.org/. Accessed June 20, 2019.

81. Kirschner KN, Yongye AB, Tschampel SM, et al. GLYCAM06: A generalizable Biomolecular Force Field. *J Comput Chem*. 2008;29(4):622-655. doi:10.1002/jcc.20820.GLYCAM06

82. Brandt M. Protein Folding. https://www.rose-hulman.edu/~brandt/Chem330/Protein_folding.pdf. Published 2016. Accessed June 20,

2019.

83. Absolorn DR, The I, Tan- W. The " Hydrophobic effect ": Essentially a van der Waais interaction. 1980;427:424-427.

84. Weiser J, Shenkin PS, Still CW. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J Comput Chem*. 1999;20:217-230. doi:https://doi.org/10.1002/(SICI)1096-987X(19990130)20:2<217::AID-JCC4>3.0.CO;2-A

85. Awoonor-williams E, Rowley CN. Molecular simulation of nonfacilitated membrane permeation ☆. *BBA - Biomembr*. 2016;1858(7):1672-1687. doi:10.1016/j.bbamem.2015.12.014

86. Marsh D. Polarity and permeation profiles in lipid membranes. 2001;98(14):7777-7782.

87. Orsi M, Sanderson WE, Essex JW. Permeability of Small Molecules through a Lipid Bilayer : A Multiscale Simulation Study. *J Phys Chem*. 2009;113:12019-12029.

88. Boggara MB, Krishnamoorti R. Partitioning of Nonsteroidal Antiinflammatory Drugs in Lipid Membranes : A Molecular Dynamics Simulation Study. *Biophysj*. 2010;98(4):586-595. doi:10.1016/j.bpj.2009.10.046

89. Saeedi M, Lyubartsev AP, Jalili S. Anesthetics mechanism on a DMPC lipid membrane model : Insights from molecular dynamics simulations. *Biophys Chem*. 2017;226:1-13. doi:10.1016/j.bpc.2017.03.006