UNIVERSIDADE D
COIMBRA

Susana Vieira Pinto da Cunha

Characterization of the interaction and permeation of drug-like molecules through membrane models including P-glycoprotein

Dissertação no âmbito do Mestrado em Biologia Celular e Molecular, com especialização em Neurobiologia, orientada pelo Doutor Hugo Alexandre Louro Filipe e coorientada pela Professora Doutora Ana Luísa Carvalho e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Setembro de 2019

Faculdade de Ciências e Tecnologia
da Universidade de Coimbra

Characterization of the interaction and permeation of drug-like molecules through membrane models including P-glycoprotein

Susana Vieira Pinto da Cunha

Dissertação de Mestrado em Biologia Celular e Molecular, com especialização em Neurobiologia, orientada pelo Doutor Hugo Alexandre Louro Filipe e pela Professora Doutora Ana Luísa Carvalho e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Setembro de 2019

1 2 9 0

UNIVERSIDADE D
COIMBRA

This master thesis project was performed in the Biological Chemistry Group at the Coimbra Chemistry Centre and the Computational and Systems Biology Group at the Centre for Neuroscience and Cell Biology of the University of Coimbra. The work was funded by "Fundação para a Ciência e a Tecnologia" (FCT) through the projects 007630 UID/QUI/00313/2013, PT2020-PTDC/DTP-FTO/2784/2014 and UID/NEU/04539/2019.

Agradecimentos

O trabalho que aqui apresento foi apenas possível graças à contribuição de várias pessoas que me apoiaram e ensinaram muito, enquanto aluna e pessoa, ao longo do último ano e a quem quero deixar uma palavra de agradecimento.

Antes de mais, gostaria de agradecer ao meu orientador Doutor Hugo Filipe, pelo seu apoio incansável ao longo do último ano. Agradeço imenso toda a ajuda e paciência que teve na minha introdução à Dinâmica Molecular e Biomembranas. À Professora Doutora Maria João Moreno agradeço pelas valiosas discussões e conselhos que me deu, e que enriqueceram muito este trabalho. Ao Professor Doutor Luís Loura e ao Professor Doutor Armindo Salvador agradeço os seus conselhos e sugestões ao longo deste projeto.

À Professora Doutora Ana Luísa Carvalho, enquanto orientadora interna, agradeço todo o apoio e a oportunidade de poder alargar a minha área de formação.

A todo o grupo de Química Biológica e colegas de gabinete, obrigada por me terem feito sentir tão bem-vinda e por todos os momentos de boa disposição. Deixo um especial obrigada à Patrícia Martins e à Patrícia Gonçalves.

Gostaria também de agradecer aos meus amigos, que me apoiaram e animaram muito. Em especial, à Joana Dias, pela força que me deu mesmo à distância e à Carolina Montezuma, por me ir buscar para almoçar quando começava a ficar stressada. Agradeço também à Margarida, à Joana Magalhães e à Carolina Caetano, o trio com quem sei que posso sempre contar. Carolina, que me acompanhaste neste mestrado, obrigada pela força imprescindível que me deste nestes últimos dois anos.

Por fim, agradeço imenso à minha família, que sempre me apoiou em tudo o que podiam. Ao meu pai, pela força e incentivo nesta nova aventura computacional (e por partilhares o meu exagerado entusiasmo pela linha de comandos); à minha mãe, por me ouvir e compreender quando mais preciso; ao meu irmão Pedro, por estar presente e me fazer rir mesmo quando está longe e à minha prima Rita, pelos momentos de descontração quando chego a casa.

Table of Contents

Abstract

The arrival of drugs at their target sites is influenced by their rate of permeation through physiological barriers, such as cell membranes and cell monolayers. It is therefore important for the development of new drugs to understand and characterize the various steps involved in the permeation process. In the study of membrane permeation, it is also relevant to consider the active transport of drugs by membrane proteins such as the P-glycoprotein (P-gp). The P-gp is found in various tissues, where it contributes to the efflux of drugs from the lipid bilayer to the extracellular medium. P-gp has mostly been studied in the context of multidrug resistance in cancer cells, and the permeation of drugs through the Blood Brain Barrier (BBB), that expresses high levels of this protein. In this work, coarse-grained molecular dynamics simulations with the Martini force field were used to characterize the interaction of diverse molecules with a model of a lipid bilayer containing P-gp. The model system consisted of a lipid bilayer containing P-pg, and with a diverse composition of lipids. The lipids are distributed asymmetrically between the two monolayers, mimicking the plasma membrane where P-gp is encountered. The computational simulation of this system was able to reproduce the annular phosphatidyl serine layer around P-gp that is found experimentally, as well as an enrichment in cholesterol and POPC. Potential of mean force profiles were used to study the interaction of different molecules with the lipid bilayer and/or with P-gp. The molecules addressed include an homologous series of fluorescent probes NBD-C_n, with varying chain lengths, as well as a group of phenolic phytochemicals with reported health benefits. These include capsaicin, curcumin, epigallocatechin gallate (EGCG), genistein and resveratrol. The NBD-C_n molecules were found to interact favorably with the lipid bilayer, especially the molecules with longer chains. However, the transfer of the molecules from their equilibrium positions in the membrane to the P-gp was energetically unfavorable. Moreover, estimated free energy variations for the transfer of NBD-C_n molecules from the water directly to the protein were also positive. In regard to the phytochemicals, the transfer of capsaicin and resveratrol from the membrane to the P-gp was found to be energetically unfavorable as well. On the contrary, the interaction of curcumin, genistein and EGCG with P-gp resulted in a decrease in free energy, which suggests that these molecules might be potential P-gp substrates. Estimated free energy variations for the interaction of all phytochemicals with the membrane and for the transfer from the water to the protein were negative. Lastly, the interactions of the various molecules with P-gp residues were identified. Most of these interactions were found to be common to other molecules in the literature, both substrates and inhibitors of P-gp.

Keywords: Membrane permeation; Blood-brain barrier; P-glycoprotein; Molecular Dynamics; Martini Forcefield

Resumo

A chegada de moléculas e drogas aos seus locais de destino é influenciada pela sua capacidade de permeação por barreiras biológicas tais como membranas biológicas e monocamadas celulares. Por esta razão, é crucial considerar os vários passos envolvidos na permeação no desenvolvimento de novos medicamentos. Neste contexto, tem particular relevância o transporte ativo de moléculas por proteínas membranares como a P-glicoproteína (P-gp), presente em vários tecidos, onde contribui para o efluxo de moléculas da membrana para o meio extracelular. O estudo desta proteína é especialmente importante no contexto da resistência a múltiplos fármacos em células cancerígenas e no estudo da permeação pela Barreira Hematoencefálica, que possui uma grande quantidade de transportadores. Neste estudo foram usadas simulações de dinâmica molecular com o campo de forças Martini para caracterizar a interação de diversas moléculas análogas a fármacos com um modelo membranar que inclui a P-gp. O sistema é composto por uma bicamada lipídica com uma grande diversidade de lípidos, bem como pela P-gp. Os lípidos estão distribuídos assimetricamente pelas duas monocamadas, assemelhando-se à membrana plasmática onde a P-gp se encontra. A simulação computacional reproduziu a formação de uma camada anelar de fosfatidil serina e um enriquecimento em colesterol à volta da P-gp, o que está de acordo com resultados experimentais descritos na literatura. Perfis de Potencial de Força Média foram usados para estudar a interação de moléculas com a bicamada lipídica e com a P-gp. As moléculas escolhidas incluem a série homóloga de sondas florescentes NBD-C_n, com diferentes tamanhos de cadeia, e um grupo de compostos fenólicos com benefícios para a saúde. Estes últimos incluem a capsaicina, curcumina, EGCG, genisteína e resveratrol. As moléculas NBD-C_n interagiram favoravelmente com a bicamada, especialmente as moléculas com cadeias maiores. No entanto, a transferência destas moléculas das suas posições de equilíbrio na membrana para a P-gp não foi favorável. Valores estimados para a variação de energia livre associada à transferência destas moléculas da água para a proteína são também positivos. No caso dos fitoquímicos, a transferência da capsaicina e do resveratrol da membrana para a P-gp também não foi energeticamente favorável. Pelo contrário, a interação da curcumina, genisteína e EGCG com a proteína levou a uma diminuição dos valores de energia livre, o que sugere que estes fitoquímicos poderão ser potenciais substratos da P-gp. Os valores de variação de energia livre associada à transferência dos fitoquímicos estudados da água para a membrana e da água para a proteína foram também negativos. Por último, foi feita a caracterização das interações de todas as moléculas estudadas com resíduos da P-gp. A maioria destas interações é também encontrada para outros substratos e inibidores da P-gp já descritos na literatura.

Palavras-Chave: Permeação membranar; Barreira Hematoencefálica; P-glicoproteína; Dinâmica Molecular; Campo de forças Martini

List of Abbreviations

BBB – blood-brain barrier

CG – coarse grain

CHOL – cholesterol

cLogP$_{Oct/Water}$ – calculated logarithm of the octanol/water partition coefficient

CNS – central nervous system

COM – center of mass

d – x-y distance to the COM of the TM region

DPSM – dipalmitoyl sphingomyelin

EGCG – epigallocatechin gallate

***K*$_I$** – Inhibition constant

***K*$_P$** – Partition coefficient

***K*$_M$** – Michaelis constant

LogP – logarithm of the partition coefficient

MD – molecular dynamics

MDR – multidrug resistance

NBD – 7-Nitrobenz-2-oxa-1,3-diazol-4-yl

NBD-C$_n$ – homologous series of NBD labeled fatty amines with varying alkyl chain length (n)

NB-Domain – nucleotide-binding domain

PE – phosphatidylethanolamine

P-gp – P-glycoprotein

PMF – potential of mean force

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPE – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine

POPS – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylserine

PO4 – bead for the phosphate group of the phospholipids

ROH – bead for the hydroxyl group of cholesterol

P-P$_{distance}$ – distance between the PO4 beads in opposing leaflets

PS – phosphatidylserine

RDF – radial distribution function

RMSD – root-mean-square deviation

RMSF – root-mean-square fluctuation

TM – transmembrane

TM-Domain – transmembrane domain

WHAM – weighted histogram analysis method

ΔG° – Free energy variation

1. Introduction

1.1 Molecular permeation across biological barriers

In the development of drugs, it is crucial to consider their movement through physiological barriers, namely cell monolayers and their plasma membranes. Different drug properties and characteristics of the lipid bilayer can influence permeation. [1] Even molecules that can permeate trough cellular membranes may be subjected to active transport, degradation or sequestration. [2,3] Such processes have important therapeutical consequences. For example, the increase in efflux mechanisms in cell membranes is thought to be the main cause of the development of Multidrug Resistance (MDR) in cancer cells. [4] MDR occurs when after treatment with a certain drug, the cells become resistant to other and unrelated drugs, such as different types of chemotherapeutic agents. This considerably limits the efficacy of treatments and can lead to failure of the chemotherapy [4,5]. Processes such as diffusion and transport across membranes must also be considered in the development of drugs that target the Central Nervous System (CNS). Due to the presence of the blood-brain barrier (BBB), the entry of some drugs to the brain is restricted, by a physical barrier as well as transporters and enzymes [6].

1.1.1 The blood-brain barrier

The normal functioning of the CNS depends on the homeostasis of the fluid that surrounds the neurons. To this extent, barriers between the blood and the CNS tightly control the entry of different molecules and cells, regulate the ionic composition that is required for neuronal signaling and prevent the entry of toxins to the brain parenchyma [7,8]. There are three main barriers in the CNS: the BBB composed of the endothelial cells of the blood microvessels, the blood–cerebrospinal fluid barrier in the choroid plexus and the arachnoid epithelium. However, the BBB is the most important for the exchange of molecules between the blood and CNS, due to the considerable total surface area of the microvessels [7]. The endothelial cells of the BBB are also part of the neurovascular unit, along with pericytes, astrocytes and neurons, as represented in Figure 1. This functional unit is important for the regulation of local blood supply and influences the induction and maintenance of the BBB [6,8,9].

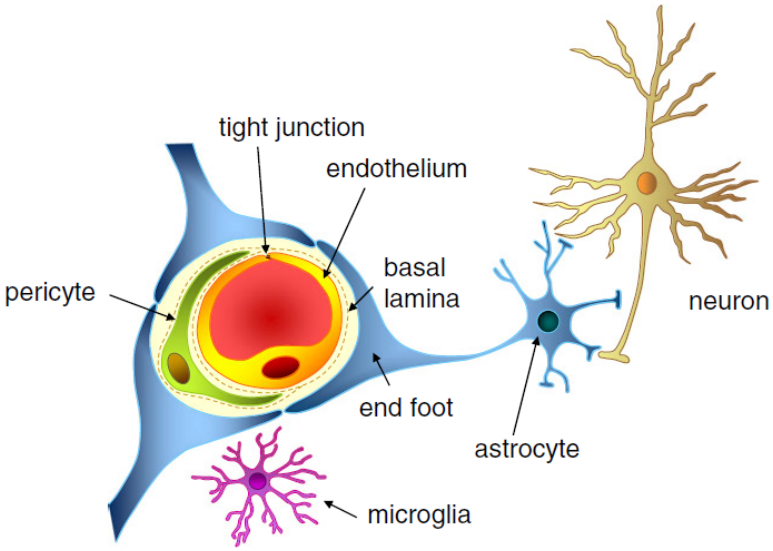

Figure 1. Representation of the neurovascular unit composed of the endothelium of the microvessels in the CNS that comprises the BBB, as well as other cellular constituents such as pericytes astrocytes and neurons that help maintain the BBB. Retrieved from [6].

The BBB is very effective as a physical barrier due to the presence of tight junctions between the endothelial cells. Tight junctions are adhesion molecules present in the membranes of adjacent cells. They greatly reduce the movement of polar molecules through the intercellular spaces (paracellular pathway), so that permeation must be done through the cells (transcellular pathway) [6,8]. Moreover, the BBB works as a transport and metabolic barrier because of the transporters and enzymes in the membrane [6]. While O_2, CO_2 and small lipophilic molecules can diffuse through the lipid bilayers of the endothelial cells, transporters are needed for the influx of ions, nutrients and other essential molecules to the brain (Figure 2). Additionally, the transport of macromolecules occurs through endocytic mechanisms such as receptor-mediated transcytosis and adsorptive-mediated transcytosis. [7,9] The transport of substances through the BBB is also conditioned by ATP-binding cassette (ABC) efflux transporters in the endothelial cell membranes, such as the P-glycoprotein (P-gp), multidrug resistance associated proteins and the breast cancer-related protein. These proteins decrease the permeation of potentially harmful lipophilic molecules that can passively diffuse through the barrier to the CNS [8,10].

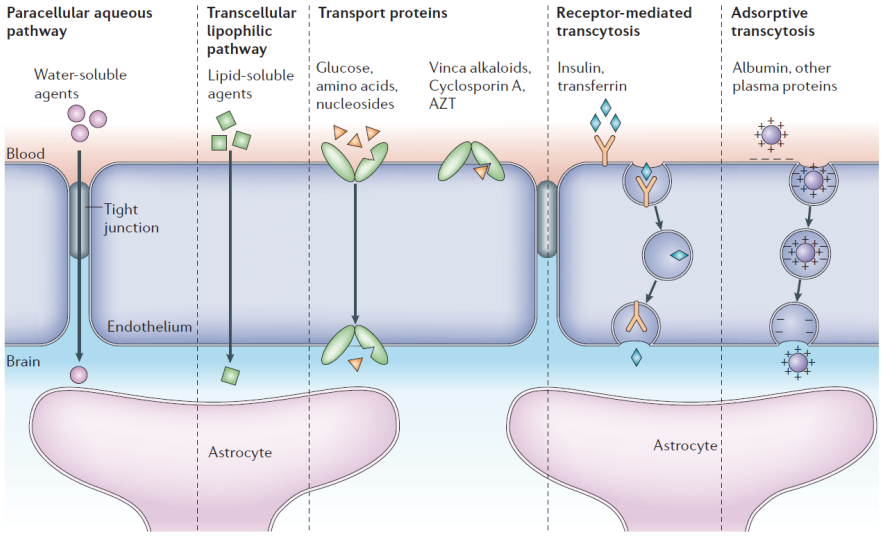

Figure 2. Transport through the blood-brain barrier. Tight junctions between the endothelial cells decrease the movement of molecules trough the paracellular pathway. Lipid-soluble molecules can diffuse through the barrier, whereas glucose and other nutrients permeate through transporters in the cell membrane. Efflux transporters such as the P-glycoprotein can also decrease the permeability of some lipophilic molecules. Proteins and other macromolecules permeate through transcytosis. Retrieved from [9].

1.1.2 Passive diffusion through biological membranes

When designing potential new drugs, it is necessary to consider their absorption, distribution, metabolism, and excretion, (ADME). The permeation through biological membranes is an essential step to take into account in this process. Given that drugs are xenobiotics, they are in general not recognized by the transporters present in the membranes. Thus, the main permeation route is by passive diffusion through the lipid bilayers of the cell membranes. To that end, efforts have been made to develop models that can predict a drug's ability to passively diffuse through lipid bilayers and cell monolayers from its physico-chemical properties, such as Lipinski's rule of 5 or more sophisticated quantitative structure–activity relationship (QSAR) models. [1,11,12]

Moreover, the permeation of a drug through a cell monolayer involves several molecular steps, as it must cross both the apical and basolateral membranes of the cell. The drug needs to partition into the outer leaflet of the apical membrane, translocate into inner leaflet of the membrane and desorb into the cytoplasm. It must then insert into the inner leaflet of the basolateral membrane, translocate into the outer leaflet and finally desorb into the target tissue. Each of these steps involves different rate constants and can be influenced both by characteristics of the membrane and by properties of the diffusing molecules. Moreover, the rate limiting-step can also vary with the properties of the drug, and needs to be identified and considered in the drug's design so that the permeation process can be optimized. [2]

Different molecule properties can determine its diffusion through lipid bilayers. Lipophilicity is considered one of the most important molecule parameters for passive diffusion, and generally leads

to an increase in permeation. In this case, descriptors such as the logarithm of the octanol-water partition coefficient (LogP) and pH-dependent logarithm of octanol-water distribution coefficient (LogD) are often used to quantify the ability of the drug to partition into the cell membrane. [1,11,13] On the other hand, hydrogen-bonding capacity, polar surface area (PSA) and molecular charge are typically associated with poor permeability [12,13]. Finally, molecular size and flexibility can also influence the molecule's ability to diffuse through membranes [1,13].

Studies *in vivo*, *in vitro* and *in silico* (computational) have provided insight into some important molecule properties that can be used to develop predictive models for permeability [14]. For example, when studying the permeation through the BBB, different permeability measures can be used. The extent of brain penetration of a molecule can be evaluated *in vivo* by calculating the logBB, which consists on the ratio between the steady-state concentrations of the drug in the brain and plasma. Alternatively, it is also possible to obtain a measure of the rate of permeation through logPS, that considers both the permeability and surface area of the vessels. *In vitro* measurements using cell culture models of the BBB have also proven useful in the study of drug permeation to the CNS. [1,14]

1.2 The P-glycoprotein

Although passive diffusion is essential for a drug to reach its target site, transporters in cell membranes can also affect the permeation of a molecule, either greatly increasing their influx or efflux [3]. Among the various efflux transporters in cell membranes, the P-glycoprotein (P-gp, ABCB1) is one of the most studied and important [13]. P-gp is an ATP-driven drug efflux pump, that belongs to the ATP-binding cassette (ABC) superfamily. Through ATP hydrolysis, P-gp is able to transport its substrates across membranes and is present in various tissues, namely in the BBB, the gastrointestinal tract, the liver, ovary and the placenta [15,16]. It is thought that P-gp may have a physiological role in protecting cells from toxic compounds by promoting their efflux directly from the membrane, and thus decreasing their entry into the cells [6]. Notwithstanding, P-gp is mostly known due to its contribution to the development of MDR in cancer cells. By promoting the efflux of different cytotoxic drugs, P-gp lowers their intracellular concentrations and can impair the response to chemotherapeutic treatments [5,17].

Moreover, P-gp is present at high levels in the luminal membrane of the endothelial cells of the BBB, that protects the CNS from deleterious molecules. Once more, although it contributes to the role of the BBB in preventing the entry of xenobiotics to the brain parenchyma, it is also an additional obstacle to drug delivery to the brain, by actively transporting them back to the blood stream [10]. This decreases the penetration of the majority of drugs such as benzodiazepines, antiemetics, some antidepressants and drugs used in the treatment of epilepsy and schizophrenia [18,19].

1.2.1 Molecular structure and transport cycle

In 2009, Aller *et al.* [20] obtained the first x-ray crystal structure of the inward-facing conformation of the mouse P-gp (Figure 3). In this conformation, P-gp has an inverted V shape, opened towards the cytoplasm. Since then, the structure has been refined [21], and the inward-faced structures of P-gp from other species have also been obtained [22,23]. Moreover, in 2018, Kim and Chen[16] reported the structure of the outward-facing conformation of human P-gp. Along with other biochemical and simulation studies [24], this has allowed for a better comprehension of the molecular structure and transport cycle of P-gp.

The P-gp is composed of two homologous halves connected by an intracellular linker, each with 6 transmembrane (TM) helices and a nucleotide binding domain (NB-Domain). In the inward conformation, the transmembrane domains (TM-Domains) limit a drug binding pocket opened to the cytoplasm, with two entry gates that allow substrate entry directly from the inner leaflet of the membrane (between TM 4 and 6 and TM 10 and 12). This pocket is composed of different binding sites and hydrophobic and aromatic residues, which allows it to interact with different types of substrates. [20,24] Pharmacological studies have suggested the existence of three main binding sites: the H site, that binds Hoechst 33342, the R site, that binds Rhodamine 123, and the P site that binds prazosin and progesterone [25,26]. Nevertheless, binding sites within P-gp may not be clearly defined and contain overlapping residues. [24,27]

Some P-gp substrates are transported without affecting the basal rate of ATP hydrolysis, while others stimulate the ATPase activity. [28] In the currently accepted mechanism for P-gp transport cycle, the protein alternates between the inward and outward facing conformation at the expense of ATP hydrolysis (Figure 4). This conformational change appears to occur even in the absence of substrates, leading to a basal rate of ATP hydrolysis [29]. In this global conformational change from the inward to the outward facing conformation of P-gp, the binding pocket and entry gates become closed, while the extracellular regions of the transmembrane domains are brought apart allowing for drug release, due to a decrease in affinity caused by changes in the drug binding sites. After release, the flexibility of the protein leads to the closure of the opening at the extracellular side, thereby preventing the entry of new substrates. Finally, ATP hydrolysis leads to the disruption of the NB-Domain dimerization and resets the inward-facing conformation of P-gp.[16,20] Recently, it was shown that substrates and inhibitors differentially modulate P-gp function. It was proposed that substrates stabilize an asymmetric post-ATP hydrolysis state, where one ATP molecule is hydrolyzed while the other remains tightly bound, which accelerates ATP hydrolysis. On the contrary, inhibitors stabilize a symmetric post-ATP hydrolysis state.[30]

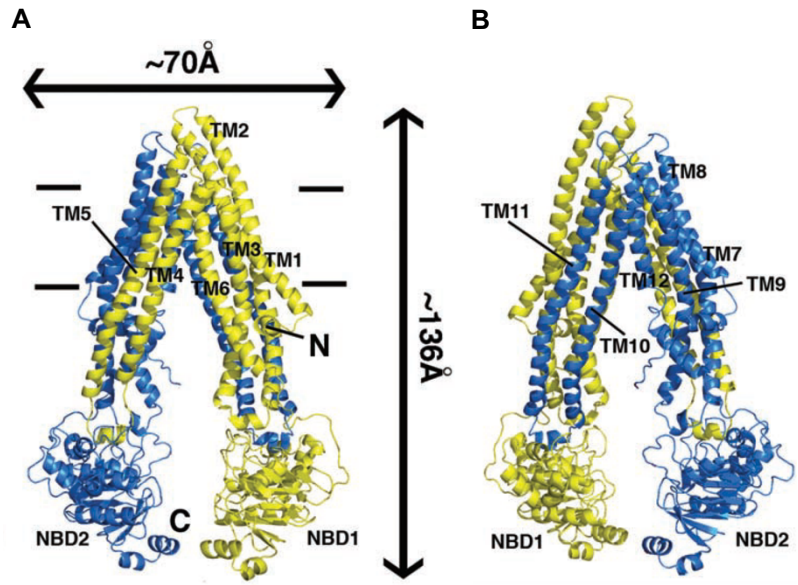

Figure 3. Crystal structure of the mouse P-glycoprotein (P-gp) in the inward conformation: A) in front and B) in back view. The P-gp is composed of two homologous halves that limit a drug-binding pocket. Each half is composed of 6 transmembrane helices and a cytosolic NB-Domain, that binds and hydrolyses ATP. The linker is not shown. Retrieved from [20].

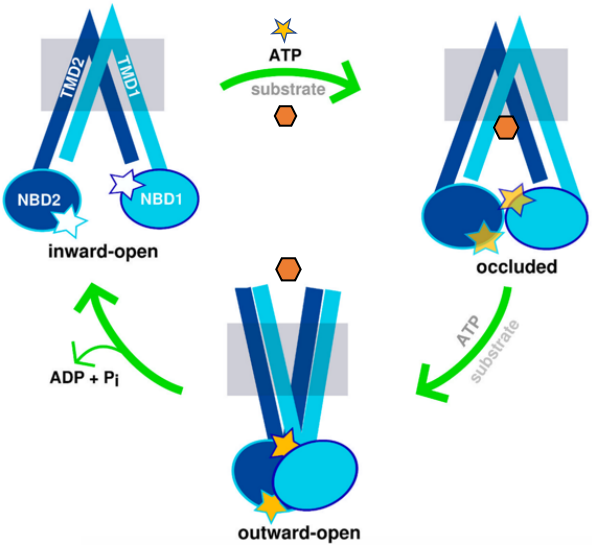

Figure 4. Transport cycle of the P-glycoprotein. In the inward-open conformation, substrates can enter the drug binding pocket through entry gates in the transmembrane domains. ATP binding induces dimerization of the NB-Domains, which leads to a global change in the protein to its outward-open conformation. The substrate is released, and ATP hydrolysis returns the protein to the initial inward-open conformation. Adapted from [31].

1.2.2 Function

Th P-gp is characterized by a broad substrate poly-specificity, as it can transport hundreds of structurally diverse molecules [24]. Most of the substrates are somewhat hydrophobic and can accumulate in the lipid bilayer. From the bilayer, they can then interact with entry gates of the P-gp on the cytoplasmic leaflet of the membrane and enter its drug binding pocket in an energetically favorable process. The possible entry of substrates directly from the cytoplasm to the binding pocket should not be disregarded. However, given the hydrophobic properties of the P-gp substrates it is widely accepted that the P-gp transports its substrates directly from the membrane rather than from the aqueous environment of the cytoplasm. [15,32,33]

Since the substrates must first partition into the membrane and translocate to the inner leaflet, the composition and fluidity of the membrane can influence the transport by P-gp. For instance, P-gp is surrounded preferentially by certain phospholipids, such as phosphatidylethanolamine (PE) and phosphatidylserine (PS) [34]. This may influence the partitioning of substrates to this membrane region, and eventually concentrate them near the P-gp, therefore increasing their binding to the protein. The lipid environment has also been shown to affect the P-gp structure and function, such as its stability and ATP binding and hydrolysis.[15] Moreover, the P-gp has also been shown to be able translocate phospholipids between the bilayer leaflets, through the same path as the transport of other drugs. This phospholipid flippase activity might therefore be responsible for a small basal activity of the protein [35,36].

Two models have been proposed for the drug efflux process. In the first, termed the “flippase” model, drugs partition into the membrane and enter the P-gp binding pocket from the cytoplasmic leaflet and are translocated or flipped to the extracellular leaflet, from which they can desorb into the extracellular medium. On the other hand, in the “hydrophobic vacuum cleaner” model, after entering the P-gp binding pocket, drugs are directly extruded to the extracellular medium. [15,32] A schematic representation of the two models is shown in Figure 5. Both of these models are currently accepted. Nevertheless, the flippase model might be more probable for more hydrophobic molecules, given the energetic cost associated to their transfer directly to the water.

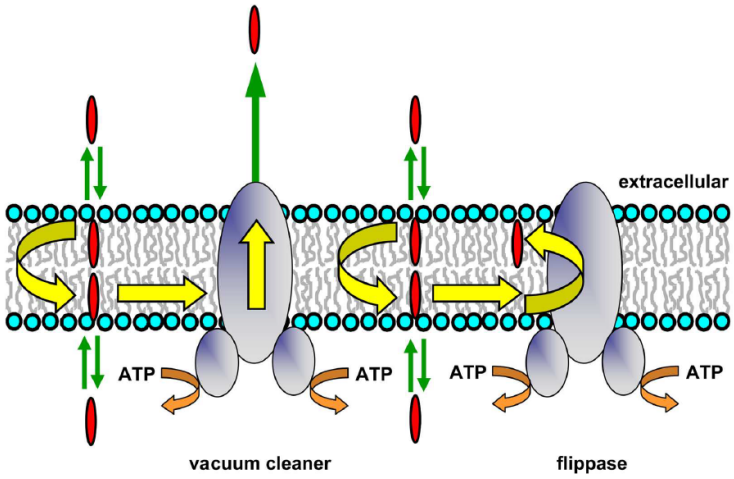

Figure 5. Drug efflux by the P-glycoprotein. In the "hydrophobic vacuum cleaner" model drugs partition into the membrane and translocate to the cytosolic leaflet, from where they are effluxed directly to the extracellular medium. In the "flippase" model, drugs are instead translocated to the extracellular leaflet, and may then desorb to the medium. Retrieved from [15].

1.3 Importance of computational studies in the understanding of biological systems

Different experimental techniques can be used to study the structure and function of biomolecules. However, since they are very dynamic and their changes in properties and conformations are relevant for their function, computational simulations can also be of great use in their study. While changes in conformations may be difficult to study, computational studies can provide detailed information on these conformations and processes. Moreover, simulations allow the study of the individual dynamics and interactions of the molecules with high spatial and temporal resolution and may help predict experimental results and provide insights into new possible experiments. [37]

The use of computational methods is of particular interest for drug design. Molecular docking studies are often used to predict the most probable conformation and orientation of a ligand in its target binding site. This is done by testing different binding conformations and determining the one with the minimum free energy. Even though docking studies are very useful in the sampling of the different binding conformations, the flexibility of the binding sites and conformational changes of the molecules are not considered. Alternatively, molecular dynamics (MD) simulations can be used to study the dynamics and temporal evolution of the interactions. [38,39]

In MD simulations, a force field is used to describe the different interactions between the particles, and their positions and velocities are updated through time according to the Newton's laws of motion [37]. Several MD simulations are done at an atomistic level by calculating the sum of forces that act on each atom at each time step. Although this gives very detailed information about the evolution of the system, it requires a lot of computational power and is not always compatible with the longer

timescales in which some cellular processes occur. Alternatively, much faster but more simplified coarse graining (CG) models can be used. With this approach, a single particle is used to represent a group of atoms, thereby reducing the degrees of freedom of the system, and number of particles to compute. [37,40]

MD can be used to simulate the dynamics of different biomolecules and cellular processes, namely conformational changes or folding of proteins, ligand binding and the interaction of drugs with membranes and transporters [37,41,42]. In fact, to study the translocation of molecules through membrane models and determine the energy barriers in the permeation process, potential of mean force (PMF) profiles can be calculated by pulling the molecule across the bilayer in MD simulations.

1.3.1 MD simulations as a tool in the study of P-gp

MD simulations have been of great use in the study of P-gp [43]. This approach has been used to show the importance of the surrounding environment in the structure of P-gp [44]. It has also been used to characterized the interactions of P-gp with its surrounding lipids in complex membrane models [45,46]. For exemple, Domicevica *et al.* [46] showed that phosphatidylserine forms an annular ring around the protein and identified possible interaction sites in the P-gp for cholesterol between TM10 and TM12. Regarding the interaction of substrates with the P-gp, this computational method has been used to identify small changes in the conformation of the protein upon substrate binding in a first step in the efflux process [47]. Moreover, PMF calculations have been used in the study of P-gp substrate entry from the water to the binding pocket and identification of possible binding sites within the binding pocket. The interactions of the substrates morphine, nicardipine were shown to be favorable and they interacted with overlapping sites in the protein [27]. Hoechst 33342, Rhodamine 123, paclitaxel, the modulator verapamil and the inhibitor tariquidar were studied in a likewise manner. All PMF profiles for these molecules had energy minima in the binding pocket and it was shown that competitive substrate binding involves overlapping locations, while non-competitive substrates likely bind to spatially separate sites, so that more than one substrate may be bound simultaneously.[48] Similarly, the entry of the substrate colchicine and inhibitor tariquidar from the bilayer core to the P-gp was shown to be favorable and possible binding sites were also identified [32].

1.3.2 The Martini Model

The Martini force field is one of the most used models in CG molecular dynamics simulations. It can be used to simulate systems with different types of molecules, such as lipids, proteins, carbohydrates [49] and DNA [50] and has been especially useful in the study of lipid membranes and their interactions with proteins. The Martini model has been used to study lipid organization in biological membranes [51], how different lipid compositions can affect membrane properties [52], and how the

interaction of different molecules with lipid bilayers can affect their properties and influence protein function [42].

The Martini model uses a 4:1 mapping, meaning that 4 heavy atoms and associated hydrogens are typically represented by a single particle or bead (Figure 6). There are 4 main types of beads considered in the model, according to the chemical nature of the corresponding atoms: charged (Q), polar (P), non-polar (N) and apolar (C). Each has different subdivisions according to their hydrogen-bonding capabilities and their degree of polarity, resulting in a total of 18 different types of beads. The strength of the interactions between the particles is then described by different parameters, according to the type of beads and type of interaction (non-bonded or bonded). Non-bonded interactions are described by a Lennard-Jones 12-6 potential, and charged groups also by a Coulombic energy function, and both are cut-off at a distance of 1.2 nm. These interactions are parameterized using a top-down approach, aiming to reproduce thermodynamic and experimental data, such as the octanol-water partition coefficient. On the other hand, bonded interactions are defined by harmonic potentials for bonds and angles, and multimodal potentials for dihedrals. The parameters for the bonded interactions are defined using structural data and to match those in atomistic simulations. This is done by mapping the molecule structure in an atomistic simulation first, and comparing the bond, angle and dihedral distribution functions to those of the CG simulation. The parameters are then adjusted for the CG simulation until the distribution matches that of the mapped atomistic simulation. [49,53]

Due to the lack of hydrogen bonds in the Martini model, the tertiary structure of proteins is not intrinsically stable. Thus, the structure is maintained by combining the Martini model with an Elastic Network model. With this method, termed ElNeDyn, a structural scaffold maintains the protein in its native conformation during the simulation. Backbone beads, that contain C_α atoms, are connected by harmonic springs characterized by a spring force constant (K_{SPRING}), that determines the rigidity of the network, and cut-off distance (R_C), that determines its extent. Springs are only defined between two beads if the distance between them is less than R_C, and beads linked by springs are excluded from non-bonded interactions given by the Lennard-Jones and Coulomb potentials. In ElNeDyn, K_{SPRING} and R_C are defined for the whole protein, but the optimal values for this parameter varies between different proteins. [49,54]

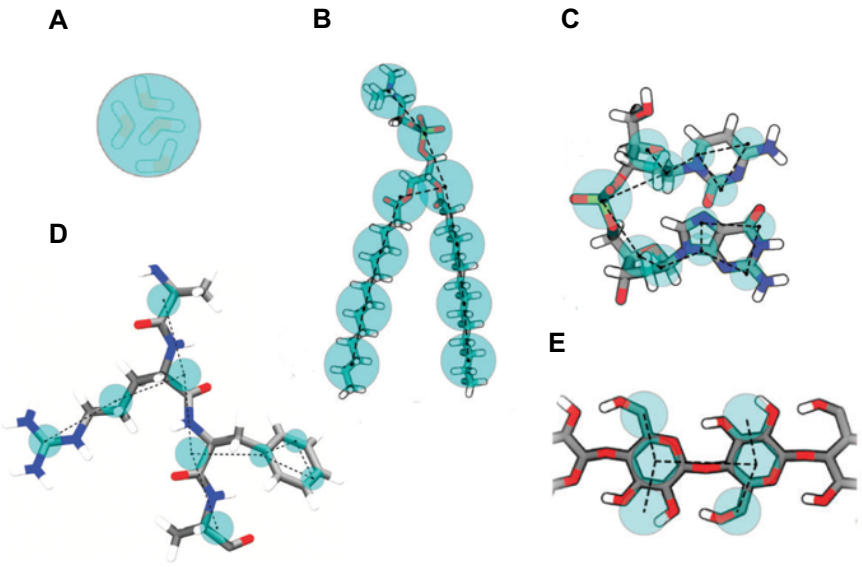

Figure 6. Examples of molecule mapping in the Martini model. A) Four water molecules are represented by a single bead, B) Lipid, C) DNA fragment, D) Peptide, and E) Polysaccharide fragment. Retrieved from [49].

1.3.3 Potential of Mean Force calculations

To study the permeation of molecules through membrane models or their interaction with P-gp and determine energy barriers in these processes, PMF profiles can be calculated by pulling the molecule across the reaction coordinate in MD simulations.

PMF profiles are used to study changes in free energy along a reaction coordinate. They are usually computed using the umbrella sampling technique, an enhanced sampling method that allows for the sampling of rare or energetically unfavorable events [55,56]. This is done by using separate umbrella simulations (referred as umbrella windows) and an harmonic potential:

$$w_i(\xi) = \frac{K_i}{2}(\xi - \xi_i^c)^2 \tag{1}$$

to restrain the simulation system at various points in the reaction coordinate, at positions ξ_i^c ($i = 1,...,N_w$) with the force constant K_i, and where N_w is the number of umbrella windows. An histogram with the biased probability distribution along the reaction coordinate is recorded for each of these umbrella windows. The PMF is then computed from the histograms using the Weighted Histogram Analysis Method (WHAM), that estimates the unbiased probability distribution with:

$$P(\xi) = \frac{\sum_{i=1}^{N_w} g_i^{-1} h_i(\xi)}{\sum_{j=1}^{N_w} n_j g_j^{-1} \exp[-\beta(w_j(\xi) - f_j)]} \tag{2}$$

and

$$\exp(-\beta f_j) = \int \mathrm{d}\xi \exp[-\beta w_j(\xi)] P(\xi) \tag{3}$$

where $\beta = 1/k_B T$, with the Boltzmann constant k_B and temperature T; n_j is the number of data points in the histogram h_j and g_j is the statistical inefficiency.

The unbiased probability distribution $P(\xi)$ is then related to the PMF, $W(\xi_0)$, by

$$W(\xi) = -\beta^{-1} \ln \left[\frac{P(\xi)}{P(\xi_0)} \right] \tag{4}$$

where ξ_0 is an arbitrary reference position for which the PMF is null, $W(\xi_0) = 0$. [56]

1.4 Aim of this study

The permeation of molecules through biological barriers is an essential biological process to consider in the design of new drugs. Although different molecular descriptors and models are already used to predict a molecule's ability to passively diffuse, there is still a need for better predictive models. Specifically, transporters such as the P-gp can determine a molecule's ability, or inability, to permeate through biological membranes and should be considered.

In this work, the interaction of a diverse group of molecules with a model of the plasma membrane including P-gp was studied using the Martini force field. First, an homologous series of fluorescent probes 7-Nitrobenz-2-oxa-1,3-diazol-4-yl (NBD)-Labeled fatty amines (NBD-C_n) shown in Figure 7A was studied. These molecules are used in the study of membrane permeation and transmembrane phospholipid distribution and their interaction with membranes is already well characterized [2,41,57,58]. Using an homologous series of molecules such as the NBD-C_n series, that varies only in alkyl chain length, allows for the characterization of the influence of this well-controlled parameter in the permeation process and interaction with P-gp.

Additionally, phenolic phytochemicals, shown in Figure 7 B-F, with reported health benefits are addressed: Epigallocatechin gallate (EGCG) (component in green tea) and resveratrol (from grapes), capsaicin (from chili peppers) and genistein (from soybeans) and curcumin (from turmeric). Several studies using these phytochemicals have been conducted to uncover their potential health benefits, namely their effects in the CNS and in cancer prevention or treatment. For example, EGCG and resveratrol have been shown to cross the BBB and have potential neuroprotective roles, which is in part attributed to their antioxidant and anti-inflammatory properties and the modulation of signaling pathways important for cell survival. [59,60] Capsaicin, genistein and curcumin have potential anticancer

effects among other benefits, due to their effect on apoptosis, cell cycle and angiogenesis [61,62]. Moreover, curcumin also has strong antioxidant and anti-inflammatory properties. Curcumin is well-established as a P-gp modulator, and experimental evidence shows that curcumin interacts with the P-gp and decreases the transport of other substrates [63]. A few studies have also reported an inhibitory effect on P-gp transport activity by EGCG [64,65] and resveratrol [66–68]. The study of these phytochemicals is compelling due to the growing interest in molecules obtained from plants that have reported health benefits, such as the ones used in this study. Studies such as this one could help shed light into the molecular mechanisms behind their function.

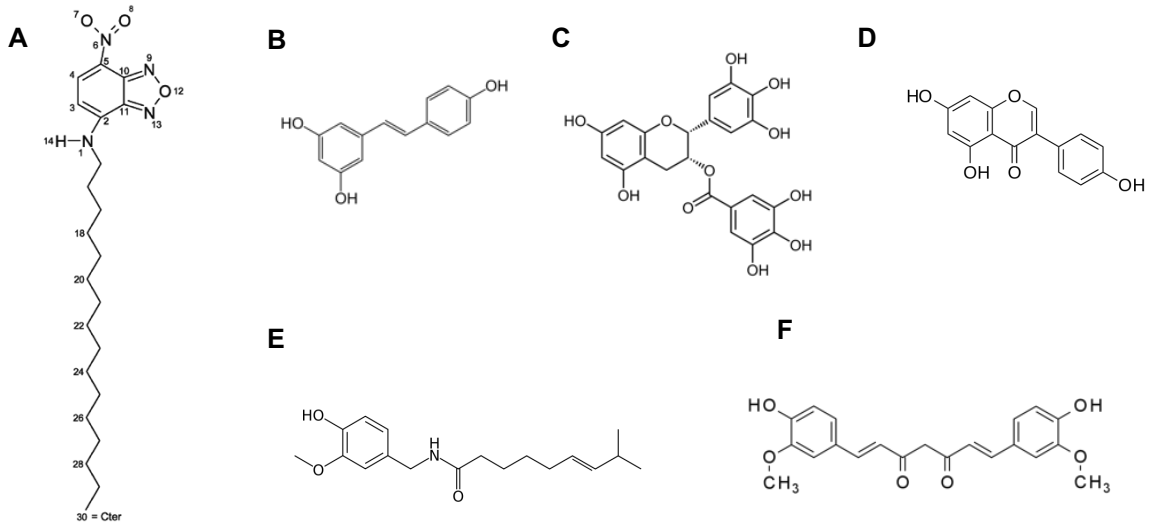

Figure 7. Chemical structure of NBD-C_{16} and the phytochemicals used in this study. A) NBD-C_{16} B) resveratrol C) EGCG D) genistein E) capsaicin F) curcumin.

The aim of this work is to characterize the individual steps of permeation of drug-like molecules through membrane models, from their interaction and partition into the membrane, together with their interaction with P-gp. The characterization of these interaction is done by PMF and free energy calculations for the different interactions, and these results are related with a possible activity of the P-gp for these molecules. Lastly, a characterization of the interactions of the molecules with diverse residues in the protein is done, as well as a comparison with residues already reported in the literature for P-gp substrates and inhibitors. MD simulations have already been used to study the interaction of other molecules with P-gp. However, in these studies the molecules were initially placed in the water [27,48] and the bilayer core [32]. In the present work, the molecules were instead placed in their equilibrium positions in the membrane, which should give a more accurate prediction of the energy variation of transfer of the molecules from the membrane to the P-gp binding pocket.

Ultimately, this work also aims to contribute to the development of a predictive permeation model started in this research group [2,41,55,58] by helping identify rate-limiting steps in the permeation

process and adding the valuable contribution of the interaction with the inward-facing P-gp conformation. An integrative characterization of both passive diffusion and interaction with an efflux transporter could contribute to the improvement of the classification of drugs according to their P-gp activity and help the development of new drugs with better permeation properties.

2. Methods

In this work, MD simulations were used to characterize the interaction of an homologous series NBD-C_n molecules and five selected phytochemicals with a membrane model of the plasma membrane including P-gp. The MD simulations and analysis were done using the GROMACS 5.1.4 simulation package [69] with the Martini CG forcefield, version 2.2 [53]. Three different systems were used in this study: one system consisting only of the complex asymmetric membrane that mimics the plasma membrane without P-gp, and two systems with P-gp, one in a complex asymmetric membrane, and one in a simpler membrane consisting of only phosphatidylcholine (POPC). The potentials of mean force (PMF) profiles were then calculated for the interaction of each molecule with the P-gp and/or membrane.

2.1 Bilayer Setup

The lipid composition of the complex asymmetric bilayers was meant to resemble the composition of a plasma membrane. To that effect, lipid concentrations were decided based on the average cell membrane composition in Ingólfsson, Melo *et al*., 2014 [51] and are detailed in Table 1. The main lipids present in plasma membranes were included: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine (POPE), dipalmitoyl sphingomyelin (DPSM), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylserine (POPS) and Cholesterol (CHOL), and distributed asymmetrically between the inner and outer layers. To study the influence of the lipid diversity of the bilayer composition on the interaction of the studied molecules in study with the P-gp, a simpler bilayer composed exclusively of POPC was also used.

To assemble the different systems, an atomistic configuration of the 4M1M P-gp, with a reconstructed linker region was converted to CG using martinize.py (see below). The protonation state of histidine residues were defined in accordance with O'Mara and Mark, 2012 [44], namely the doubly protonation of His149, His583 and His1228. The protein structural conformation was maintained using the ElNeDyn22 elastic network [54], with a force constant K_{SPRING} of 500 kJ mol^{-1} nm^{-2} and a R_C cut-off of 0.9 nm. The assembling of the surrounding membrane and solvation with water, Na^+ and Cl^- beads in a 150mM concentration was done using insane.py, creating a lipid bilayer with the selected lipids in a random distribution. Both the martinize.py and insane.py scripts were obtained from the Martini webpage (http://cgmartini.nl). A snapshot of the complex membrane system containing the P-gp is shown in Figure 8.

Table 1. Composition of the systems used in this work. Lipid percentage is showed, with absolute numbers in brackets.

	Complex asymmetric membrane with P-gp		Complex asymmetric membrane without P-gp		POPC membrane with P-gp	
Lipid	Inner Layer (%)	Outer Layer (%)	Inner Layer (%)	Outer Layer (%)	Inner Layer (%)	Outer Layer (%)
POPC	19 (84)	38 (178)	19 (91)	38 (188)	100 (257)	100 (265)
POPE	27 (124)	6 (27)	27 (135)	6 (29)	0	0
DPSM	10 (44)	20 (96)	10 (48)	20 (101)	0	0
POPS	11 (53)	0	11 (58)	0	0	0
Cholesterol	31 (138)	34 (155)	31 (150)	34 (164)	0	0
Total lipid number	443	456	482	482	257	265
Water beads	21628		27932		12901	
Ion beads	480		617		284	
Simulation box size (nm^3)	15 x 15 x 20		20 x 20 x10		13.5 x 13.5 x 20	

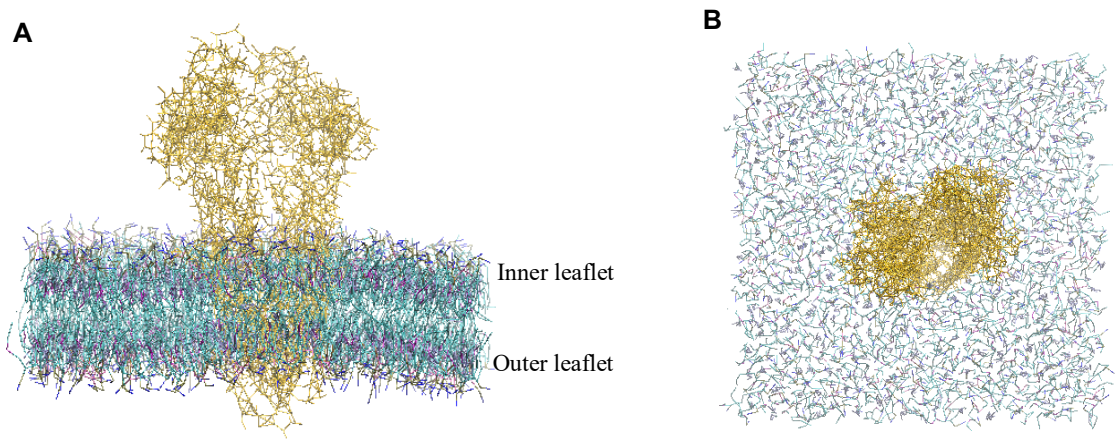

Figure 8. Snapshot of the complex membrane system containing the P-gp: A) side view and B) upper view of the NB-Domains. The P-gp is shown in yellow and inserted in the bilayer, with the NB-Domains on top. Water surrounding the membrane and P-gp is not shown.

All simulations were run using a constant number of particles, under a temperature of 298K and pressure of 1bar and using periodic boundary conditions. Temperature and pressure control were done using the V-rescale thermostat and Berendsen barostat, respectively, and with semi-isotropic pressure coupling. Coulomb interactions were calculated using the reaction field method with a cut-off of 1.1 nm and a dielectric constant of 15, and Lennard-Jones interactions were cut-off of at 1.2 nm.

For the three bilayer systems, energy minimization was then performed using the steepest descent method followed by two equilibration simulations of 1ns and 10 ns, with a 2 fs and 10 fs integration step, respectively. A production run was then carried out for 10 μs with a 20 fs integration step to allow for changes in the distribution of the lipids in the bilayer.

2.2 Parameterization of the molecules in study

A homologous series of 7-nitrobenz-2-oxa-1,3-diazol-4-yl (NBD) -labeled fatty amines (NBD-C_n) was used, with chain lengths (C_n) 4, 8, 12 and 16. The CG topology for the NBD group was obtained from Filipe *et al*., 2019 [70]. NBD-C_4, NBD-C_8 and NBD-C_{12} were obtained by removing beads from the NBD-C_{16} alkyl chain (Figure 9A).

The CG topologies of the phytochemicals capsaicin, curcumin, EGCG, genistein, and resveratrol (Figure 9B-F) were retrieved from the Martini webpage (http://cgmartini.nl).

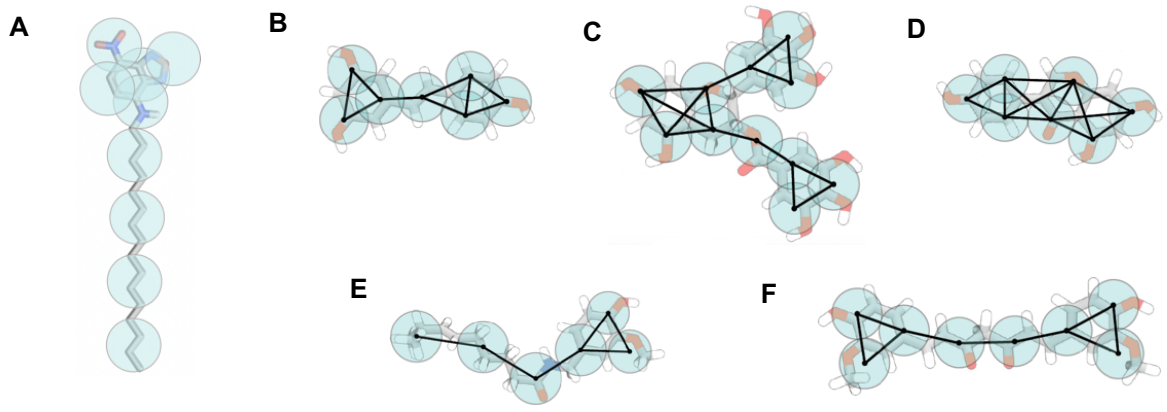

Figure 9. Coarse grain (CG) mapping of NBD-C_{16} and the phytochemicals. A) NBD-C_{16}, B) resveratrol, C) EGCG, D) genistein, E) capsaicin, and F) curcumin. The phytochemical figures were retrieved from [42].

2.3 Umbrella Sampling Simulations

2.3.1 Water - membrane system

Firstly, the permeation of the NBD-C_n series through the membrane was studied, using the complex lipid bilayer system without the P-gp. The final configuration of the 10 μs unrestrained simulation was used as the starting point for the umbrella sampling simulations. Then, for each case, two NBD-C_n molecules were added to the system, so that both leaflets of the membrane could be sampled simultaneously. One molecule was placed with the NBD group in the center of the bilayer, whereas the other was placed in the solvent, at a distance of 4 nm in z coordinate from the first molecule. The initial positions for NBD-C_{16} are showed in Figure 10 as an example. After the energy minimization, a 10 ns run with a 20 fs integration step was done with fixed z coordinate values to ensure the correct distance between the molecules. The pulling of the molecules was then done in the z axis using the pull geometry cylinder method, to decrease membrane deformations [55], and a pulling rate of 0.0002 nm ps^{-1} and a force constant of 500 kJ mol^{-1} nm^{-2}. For the umbrella sampling simulations, the windows were spaced 0.1 nm apart, using the structures obtained from the pulling simulation. A harmonic umbrella potential was applied to the NBD's center of mass (COM), with a 3000 kJ mol^{-1} nm^{-2} force constant and for 200 ns. PMF profiles were then calculated using the WHAM method implemented in the GROMACs package [56].

PMF profiles for the transfer of the phytochemicals between the water and membrane were not calculated. Alternatively, Volsurf+ [71]was also used to estimate values for the calculated logarithm of the octanol/water partition coefficient ($cLogP_{Oct/Water}$) for the NBD-C_n molecules and the phytochemicals.

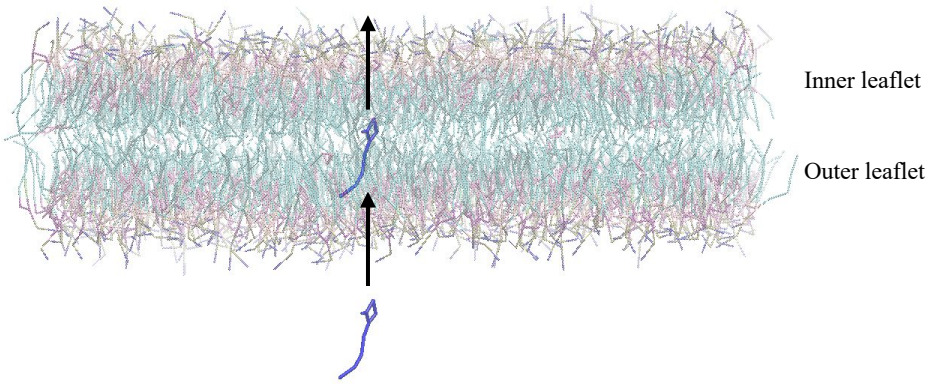

Figure 10. Initial positions and direction of pulling from the membrane to the water of NBD-C_{16} (in dark blue) in the membrane/water system. Two molecules were added in each system. One was placed in the water and another with the NBD group in the center of the bilayer, 4 nm apart. Water above and below the membrane is not shown.

2.3.2 P-gp – bilayer systems

To study the transfer of the molecules from the lipid environment to the P-gp the reaction coordinate was defined as their distance in the xy plane to the COM of the TM region of the P-gp (d=0). This region is composed exclusively of the beads in the TM helices. The different molecules were initially added to the lipid bilayer at a depth correspondent to their equilibrium positions in the inner leaflet. This was done for both the complex and POPC bilayers and at a distance of 5nm in xy plane to the COM of the TM region of the P-gp, as shown in Figure 11. The equilibrium positions of the molecules in the bilayer was obtained from Filipe *et al*., 2011 [41] for the NBD-C_n series and from Ingólfsson *et al.*, 2014 [42] for the phytochemicals. For each case, a single molecule was aligned with the P-gp gate composed of TM4 and TM6. Although the P-gp has two gates the cleft of the TM4/6 gate was larger, which allows for the entry of larger molecules and simulations have previously captured the entrance of a lipid through this gate [72].

Energy minimization was done using steepest descent, followed by a 1ns equilibration simulation with a 2 fs integration step and 10 ns simulation with a 10 fs integration step. Additionally, a 10 ns simulation, with a 20 fs integration step was run, where the x and y coordinate values of the drugs were restrained at 5nm from the protein so that the drug could adjust its z position in the bilayer. In the case of the NBD-C_n series only the NBD's COM was restrained, whereas in the case of the phytochemicals the COM of the whole molecule was used. Finally, the molecules were pulled from their positions in the bilayer at 5 nm to the COM of TM region with a pulling rate of 0.0002 nm ps^{-1} and a force constant of 500 kJ mol^{-1} nm^{-2}, using the pull geometry distance method. Ten independent pulling simulations were done for each molecule. Out of the ten simulations one was chosen based on the visualization of the trajectory of the molecule, guaranteeing that the molecule entered the P-pg through the gate and did not desorb to the water.

A total of 51 initial configurations, from the TM COM of the P-gp to 5 nm from the protein and spaced 0.1 nm in between, were then extracted from the pulling simulation. For each initial structure an independent umbrella sampling simulation was run by applying a harmonic umbrella potential to the COM of the drug, with a 3000 kJ mol^{-1} nm^{-2} force constant. Each run was done for 200 ns so that the drug could interact with the surrounding molecules. PMF profiles were then calculated using the WHAM method implemented in the GROMACs package [56].

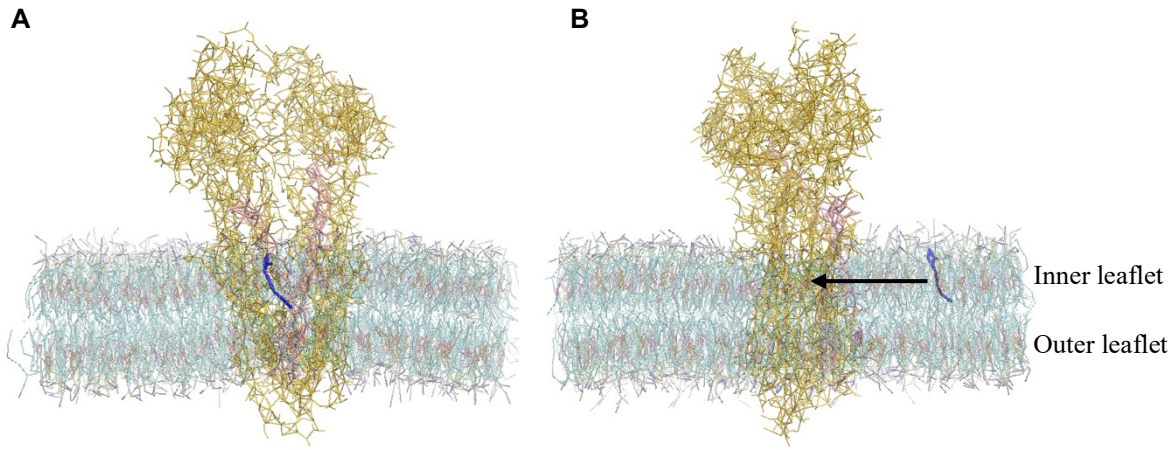

Figure 11. Initial position and direction of pulling of NBD-C_{16} from the membrane to the P-gp: A) front view and B) side view. The NBD-C_{16} molecule was placed at 5nm in xy plane from P-gp TM COM. The P-gp is shown in yellow, with the TM4 and TM6 gate in pink, and NBD-C_{16} is shown in dark blue. Water above and below the membrane is not shown.

2.4 Analysis of sampling histograms and convergence of the Free Energy Profiles

Sampling histograms with the probability distribution of the molecules in the umbrella windows along the reaction coordinate were analyzed to guarantee that there was an overlap between adjacent umbrella windows. This is required to guarantee that there is a sufficient sampling of the system and the WHAM analysis can be properly used [56]. As an example, the set of histograms for the sampling of NBD-C_8 in the membrane/P-gp and membrane/water systems are shown in Figure 12. For all molecules there was a sufficient overlap for a correct sampling of the system. In any case, the sampling in the interface between the membrane and the P-gp was noticeably more difficult for some molecules, with a smaller histogram overlap in this region. This might be due to the increase in heterogeneity of the environment near the protein.

Moreover, to guarantee that enough simulation time was used for the molecules to sample the surrounding system, the convergence of the free energy profiles was evaluated according to the method described in Filipe *et al.* [55].To do this, multiple PMF profiles were calculated for increasing simulation times, by a 10 ns interval. All profiles converged during the 200 ns. The same analysis was also done for decreasing simulation times, by discarding initial simulation times, to evaluate how much time was needed for equilibration of the system. This is important so that the sampling of the system is not dependent on the initial conditions. The final PMF profiles and subsequent analysis of the systems were calculated using only 150 ns of simulation, disregarding the first 50 ns as equilibration time. The convergence of the PMF profile for the transfer of EGCG from the membrane to the P-gp is given in Figure 13 and is a good example of the need for this analysis. It is worth noticing that while the profile converges around 50 ns (Figure 13 C and D), disregarding too large amounts of initial simulation time

results again in the divergence of the profile. This shows that the use of small simulation times, both total or by discarding too much initial time, results in altered PMF values because not enough time is being used in the analysis.

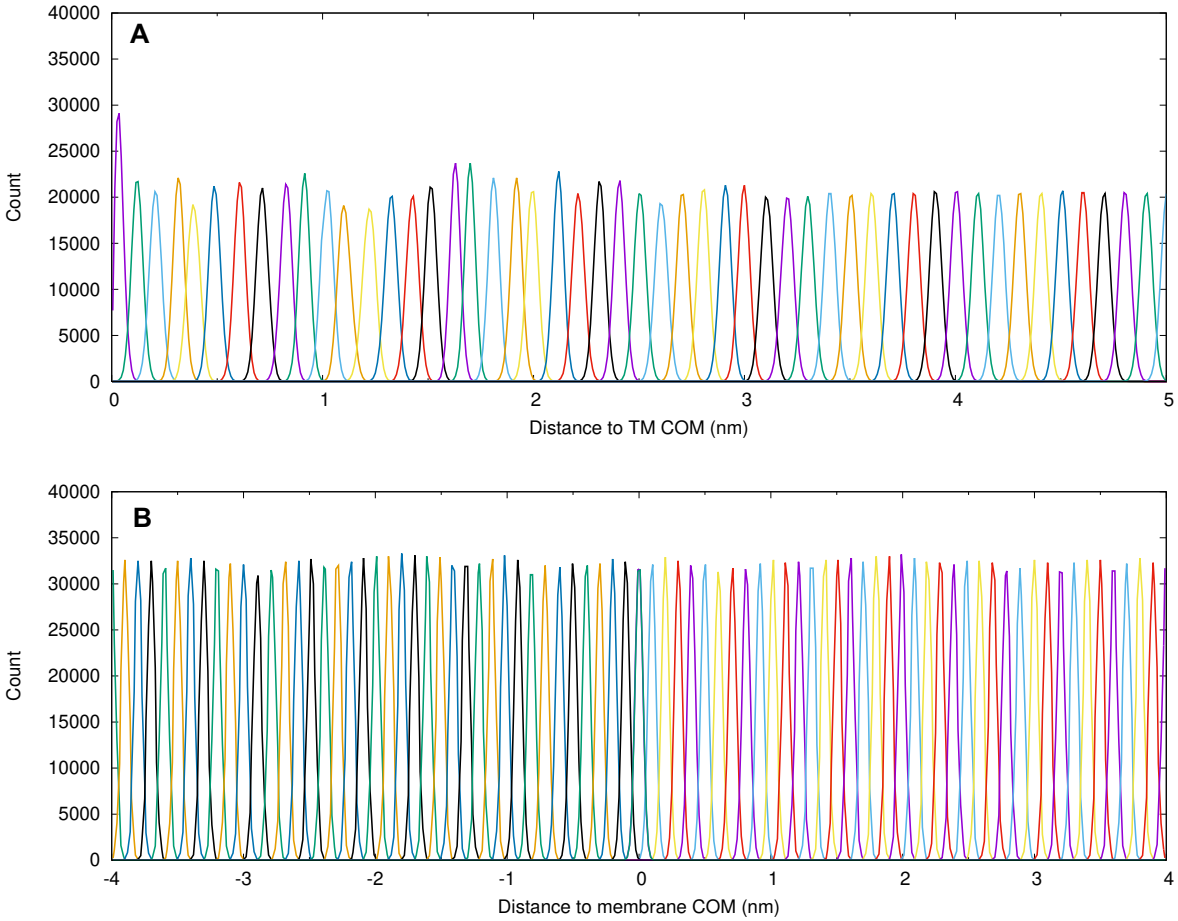

Figure 12. Examples of sampling histograms for the distribution of molecules in adjacent umbrella windows. A) sampling of NBD-C_8 in the P-gp/membrane system, B) sampling of two NBD-C_8 molecules in the water/membrane system.

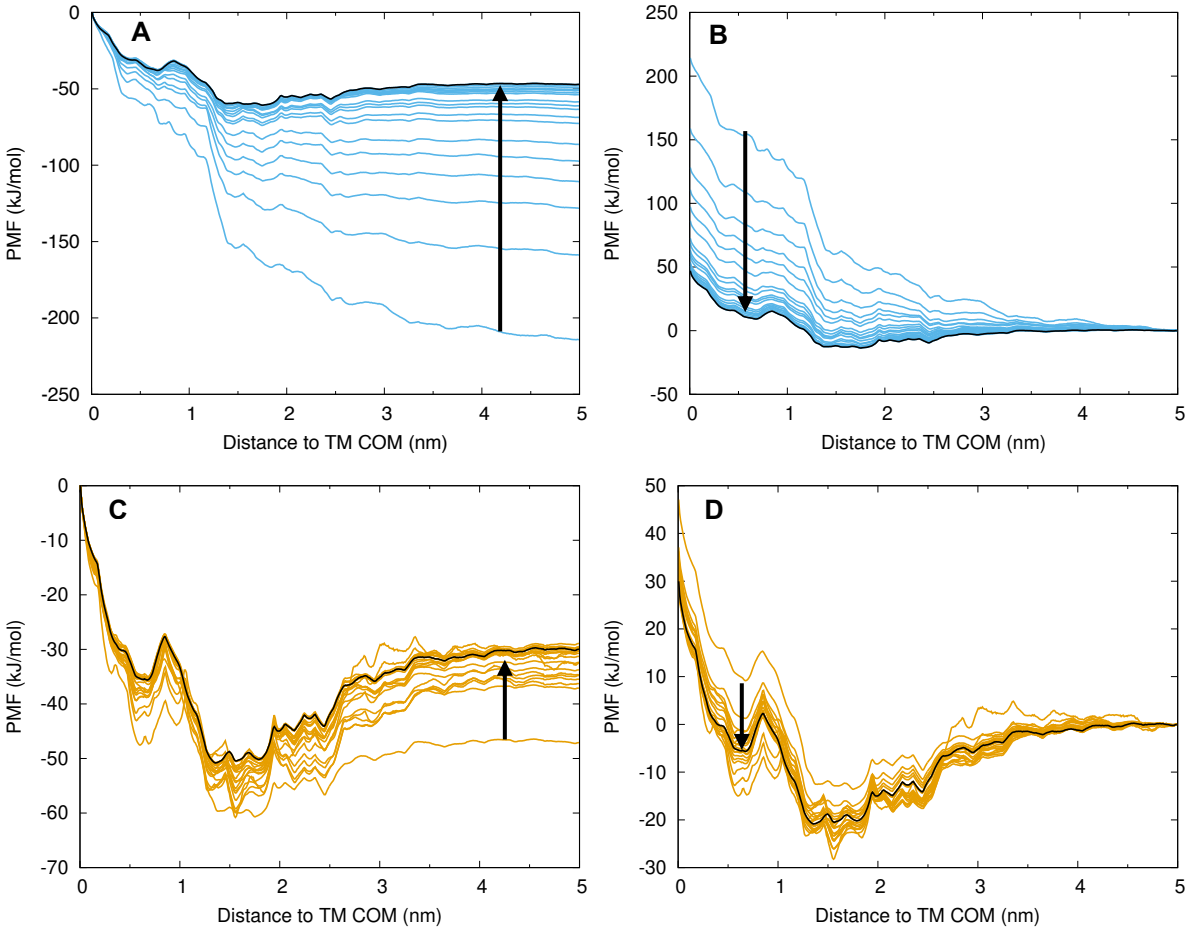

Figure 13. Convergence of the potential of mean force (PMF) profiles for the transfer of EGCG from the membrane to the P-gp, in the plasma membrane model. A) and B) show profiles for increasing simulation times by a 10 ns interval, until 200 ns (in black); C) and D) show profiles obtained by disregarding initial simulation times by a 10 ns interval, until 200 ns. The profile for 50 ns as equilibration time is shown in black. A) and C) have the energy reference position at 0 nm from the TM COM, while C) and D) have the reference at 5 nm.

2.5 Error estimation for the PMF profiles

The uncertainty associated to each PMF profile was estimated dividing the sampling data in two different sets. After disregarding the first 50 ns as equilibration time, the 150 ns sampling time was split in half, and used to calculate two additional free energy profiles for each molecule. The difference between them was then calculated and used as an error estimate.

2.6 Data Analysis and Visualization

2.6.1 RMSD, RMSF and P-gp tilt calculations

To test the effectiveness of the ElNeDyn network surrounding the P-gp, the root-mean-square deviation (RMSD) (Equation 5) of the whole protein and different domains was calculated, as well as the root-mean-square fluctuation (RMSF) for the different residues. This was done using the gmx rms and gmx rmsf GROMACs tools. The different structural motifs of the protein were defined based on the assignment of secondary structures by DSSP in the Protein Data Bank, and on the attributions in Li *et al.* [21] and Condic-Jurkic *et al*. [31]. This analysis was done with the GROMACs tools gmx rms and gmx rmsf, using the martinized P-gp starting structure as reference. Only the changes in position of the backbone beads of the molecules were considered. The RMSD is defined as,

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i \| r_i(t_1) - r_i(t_2) \|^2 \right]^{\frac{1}{2}} \tag{5}$$

with $M = \sum_{i=1}^{N} m_i$, where m_i is the mass of atom i, and $r_i(t)$ its position at time t . [73]

RMSD calculations were done for the whole protein, the linker region the NB-Domains and the TM-Domains, that include both the TM region defined for the umbrella sampling simulations as well as inter-helix regions.

The tilt of the protein was calculated using the gmx bundle GROMACs tool by calculating the angle between z axis of the membrane and a vector in the P-gp defined from the COM of the TM region to the COM of the NB-Domains.

2.6.2 Analysis of lipid distribution

The lipid distribution around P-gp may have important effects in the binding process of other molecules, such as protein substrates. Therefore, the distribution of lipids around the P-gp during the 10 μs simulation was studied by calculating the Radial Distribution Function (RDF) and by plotting 2D-density maps using GROMACs tools gmx rdf and gmx densmap. In this analysis, the distribution of the bead for the phosphate group (PO4 bead) of the phospholipids and the bead for the hydroxyl group (ROH bead) of cholesterol was considered. The RDF equation or pair correlation function $g_{AB}(r)$ between particles A and B is defined in GROMACs as:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} = \frac{1}{\langle \rho_B \rangle_{local}} \frac{1}{N_A} \sum_{i \in A}^{N_A} \sum_{j \in B}^{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2} \tag{6}$$

where $\langle \rho_B(r) \rangle$ is the B particle density, at a r radius from A, and $\langle \rho_B \rangle_{local}$ is the average B particle density within a certain radius, r_{max} which is half of the size of the simulation box. [73] This means counting all B particles positioned in $\boldsymbol{r}_j$ within a shell centered on particle A, at position $\boldsymbol{r}_i$, with a radius between r and $r+dr$. This number is divided by the volume $4\pi r^2 dr$ and averaged for all A particles. The Dirac delta function $\delta(r_{ij} - r)$, with $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$, selects $r_{ij} \sim r$. In this study, A is the COM of the TM region of the P-gp and B the surrounding phospholipid headgroup beads in the membrane.

2.6.2 Detailed analysis of the interactions with the P-gp

To study the interactions of the different molecules with the P-gp, the <z> distance of the COM of the molecules inside P-gp was calculated both in relation to the TM COM of the P-gp and to the COM of the bilayer, using the GROMACs tool gmx distance. The tool gmx select was also used to identify the P-gp residues with which the molecules interact in the d=0 umbrella window simulation and, in the case of the phytochemicals, also at the energy minimum determined from the PMF profiles. For each molecule, the P-gp residues were then split into two groups: half with the higher number of interactions (above the 50th percentile) and half with the smaller number interactions (below the 50th percentile).

The total simulation time in the present study reached more than 150 μs.

Visual Molecular Dynamics (VMD) was used to visualize the different trajectories and Gnuplot was used to plot the graphics.

3. Results and Discussion

3.1 Characterization of the lipid model system including the P-gp

3.1.1 Protein Structure

The simulation of protein systems with the Martini model requires the application of an elastic network to maintain the correct protein structure. Due to the use of the ElNeDyn network in the P-gp, no significant changes in protein structure were expected during the course of the simulations. To confirm the effectiveness of this network and test how much the protein structure changed during the 10 μs simulation, RMSD values were calculated for the backbone beads of the different domains of the P-gp, shown in Figure 14. The initially martinized structure was used as reference. For the protein inserted in the asymmetric membrane (Figure 14A), RMSD values ranged from 0.1 nm to 0.2 nm for the two NB-Domains, and 0.3 nm to 0.6 nm for the whole protein and remaining domains. These values indicate small changes in protein structure, without major alterations in the RMSD values throughout the simulation, which indicates that ElNeDyn network was effective. The values obtained are also in accordance with previous reported values for atomistic P-gp structures, namely for 4M1M and 3G5U [31]. The P-gp in the POPC membrane (Figure 14B) showed similar RMSD values, with the exception of the TM-Domain 2 region, that had smaller changes in backbone beads displacement.

RMSF values of the protein residues were also calculated to identify the regions of the protein with the larger changes in backbone beads position and are shown in Figure 15. The RMSFs were similar for the P-gps in both membranes, and the major changes were found to be in the linker region, indicating its higher flexibility despite the application of the elastic network in the protein parameterization.

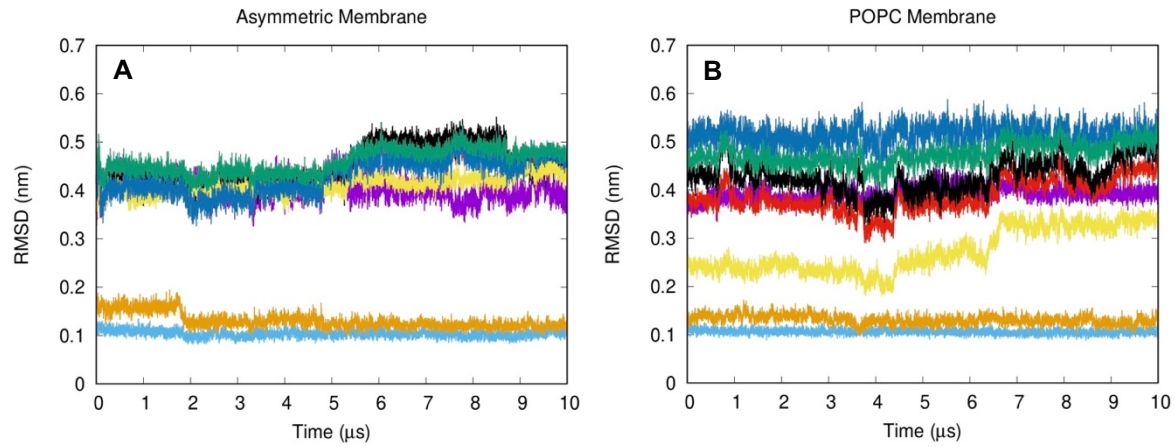

Figure 14. RMSD of the backbone beads of different P-gp domains, during the 10 μs simulation for the protein inserted: A) asymmetric membrane and B) POPC membrane. RMSD values were calculated for the whole protein (in green) and for the TM-Domains (red), the TM-Domain 1 (black), TM-Domain 2 (yellow), NB-Domains (dark blue), NB-Domain 1 (light blue), NB-Domain 2 (orange) and linker (purple).

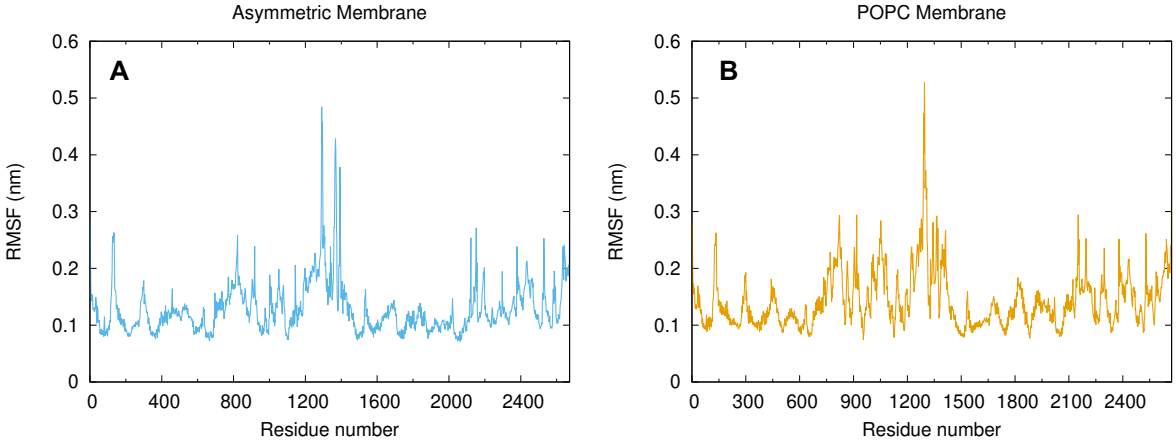

Figure 15. RMSF values for the backbone beads of the P-gp during the 10 μs simulation for the protein inserted in: A) the asymmetric membrane, and B) the POPC membrane.

While the structure of the P-gp was restricted by the elastic network, the protein was still free to change its position and orientation in the membrane. This can be influenced by the thickness of the membrane, that is dependent on the lipid composition. Therefore, the thickness of the bilayer was evaluated using the distance between the PO4 beads in opposing leaflets ($P\text{-}P_{distance}$). The $P\text{-}P_{distance}$ was on average 4.17 ± 0.02 nm for the asymmetric membrane and 3.92 ± 0.02 nm for the POPC membrane. To further characterize this model system, the tilt of the protein with reference to the z axis of the membrane was also calculated. During the simulation, the tilt of the P-gp in the asymmetric membrane, shown in Figure 16, was on average 6.34 ± 2.80°, and ranged from 0.02° to approximately 18.11°. For the P-gp in the POPC membrane, in Figure 17, the average tilt was 6.80 ± 2.95°, and ranged from 0.02° to 19.13°. The tilt of the P-gp was therefore very similar in both systems studied here, despite the differences in the thickness of the bilayer. Moreover, there is no noticeable curvature of the membrane around the protein in either case (Figure 16 and Figure 17 B and C). Together these results indicate that in both cases the more hydrophobic domain of the protein is inserted in the hydrophobic core of the membrane. Otherwise, a higher protein tilt during the simulation would be expected. These tilt values are also in agreement with those found in another simulation of the P-gp embedded in a dimiristoyl-phosphatidylcholine (DMPC) bilayer [47].

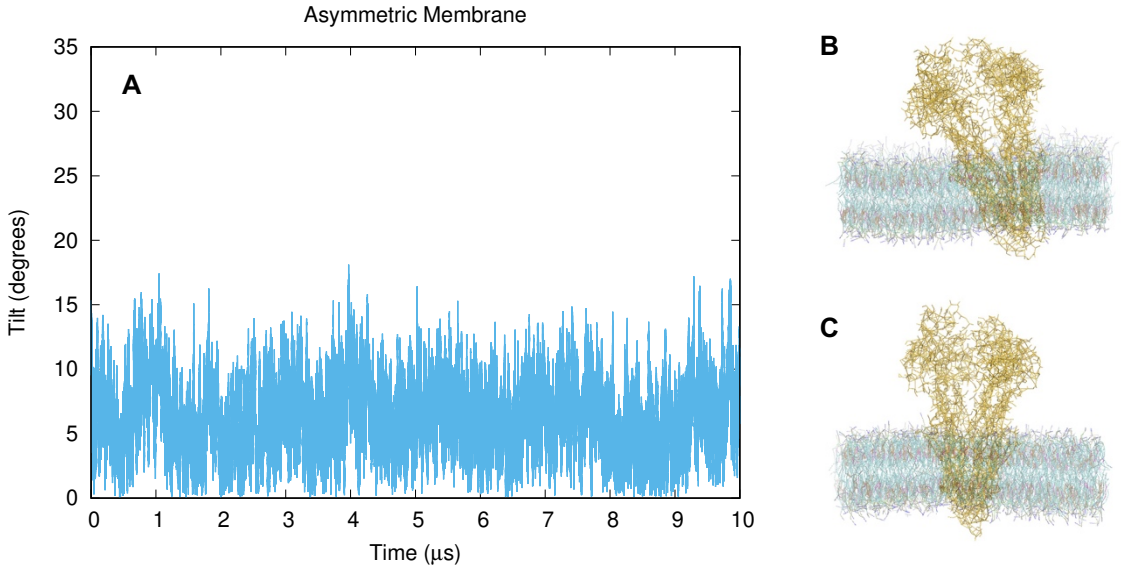

Figure 16. P-gp tilt in the asymmetric membrane: A) changes in tilt during the 10 μs simulation, with reference to the z axis of the membrane, B) snapshot of maximum axis tilt configuration (18.11°) and C) snapshot of minimum axis tilt configuration (0.02°). The P-gp is shown in yellow, and in front view. Water is not shown.

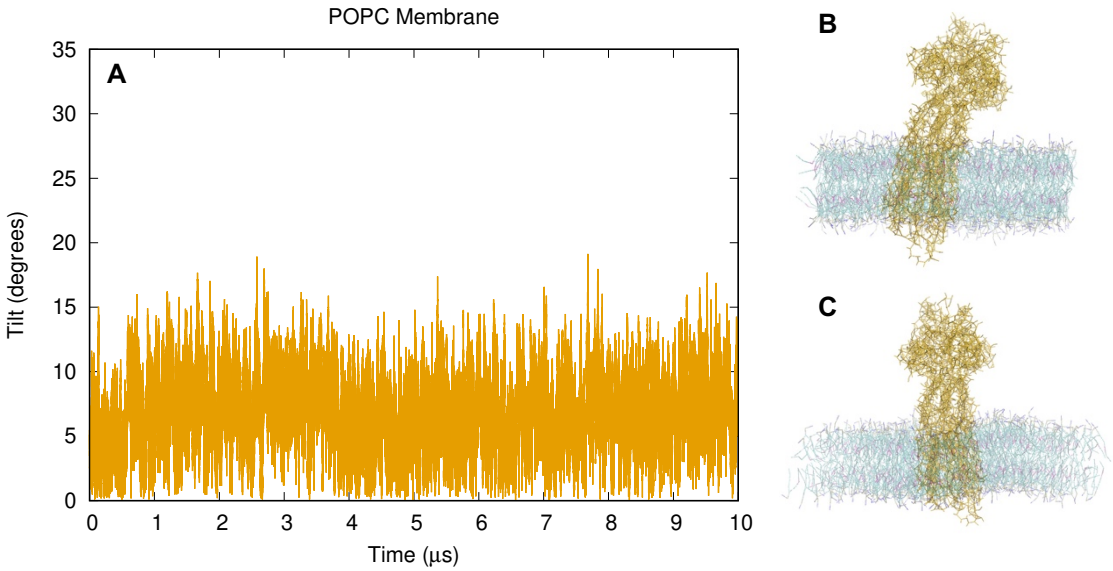

Figure 17. P-gp tilt in the POPC membrane: A) changes in tilt during the 10 μs simulation, with reference to the z axis of the membrane, B) snapshot of maximum axis tilt configuration (19.13°) and C) snapshot of minimum axis tilt configuration (0.02°). The P-gp is shown in yellow, and in side view. Water is not shown.

3.1.2 Lipid Distribution

The presence of a transmembrane protein, such as the P-gp, can have a profound effect on the lipid environment in a membrane. Namely, the distribution of lipids may be altered, depending on the interactions of their headgroups with the protein. To understand the effect of the presence of the protein in this bilayer model, lipid 2D-density maps for the x-y plane of the membrane were used to study the distribution of the lipids during the course of the 10 μs simulation. This was done firstly for all PO4 beads of the phospholipids in the POPC membrane in order to identify regions of higher density. This is shown in Figure 18, where a red region with no lipid is show in the center, corresponding to the space occupied by P-gp. The color then gradually changes as the distance from the protein increases, reflecting the increase in lipids around the protein. This gradual transition in color can be explained by the irregularities in the protein and changes in its tilt throughout the simulation. The surrounding POPC lipids show a generally uniform lipid density. Noticeably, the space occupied by the protein in the inner leaflet (Figure 18A) is larger than in the outer leaflet (Figure 18B). This is due to the V shape of the protein, that has the wider region in the inner leaflet of the membrane.

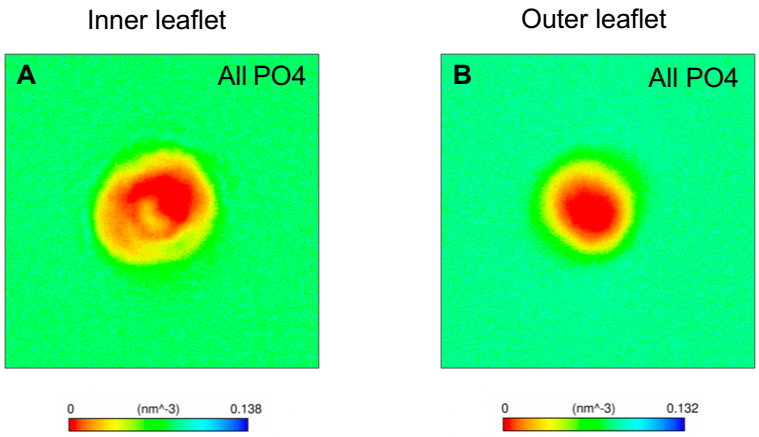

Figure 18. Lipid density maps for the POPC membrane. A) PO4 bead density in the inner leaflet, B) PO4 bead density in outer leaflet. The density values are indicated by the color scale bars. The region in red corresponds to the space occupied by P-gp.

In the complex membrane there is also a uniform distribution of phospholipids, shown in Figure 19A and G, although there are some small regions of higher density near the protein. However, in the asymmetric membrane there is a clear differential distribution of the distinct lipids. Most noticeably, POPS, a negatively charged lipid, concentrated preferentially around the P-gp in an annular ring (Figure 19F). This is further confirmed by the radial distribution function (RDF) in Figure 20, which clearly shows an enrichment in POPS from 2.5 nm to 4 nm radius from the COM of the transmembrane region of the P-gp. Additionally, a smaller enrichment in POPC and cholesterol near the protein can be identified in the RDF profile. It is to note that this enrichment, also seen in Figure 19B and E, respectively, is at a very small distance from the COM of the TM region, and on opposite sides of the protein. This suggests that these lipids interacted with the gate regions of P-gp. On the contrary, DPSM was not enriched near the protein, as is shown in its RDF profile and Figure 19D. Lastly, the distribution of POPE in the bulk region of the outer leaflet appears to be heterogeneous. However, this may not be due to the presence of higher POPE density regions, but rather result from the small amounts of this lipid in the outer leaflet. That is, there may have been an insufficient dynamic of the lipid during the course of the simulation, resulting in an heterogeneous map.

Previous studies have studied the interactions of the P-gp with its surrounding lipids. An experimental study showed that P-gp associated mainly with PE and PS [34]. Moreover, a MD study by Domicevica *et al*. using a similarly composed bilayer also reported the formation of a PS annular ring around the protein, and showed that cholesterol interacted mainly with the cavity near the TM10 and 12 gate [46], which is in agreement with the results presented here. Another MD study by Corradi *et al.* [45] showed a small enrichment in PE and PS around the protein, and a depletion in PC and DPSM. No enrichment in CHOL was found. These differences may be due to the composition of the membrane used. Particularly, the small enrichment in PS may be due to the presence of other negatively charged lipids such as phosphatidylinositol.

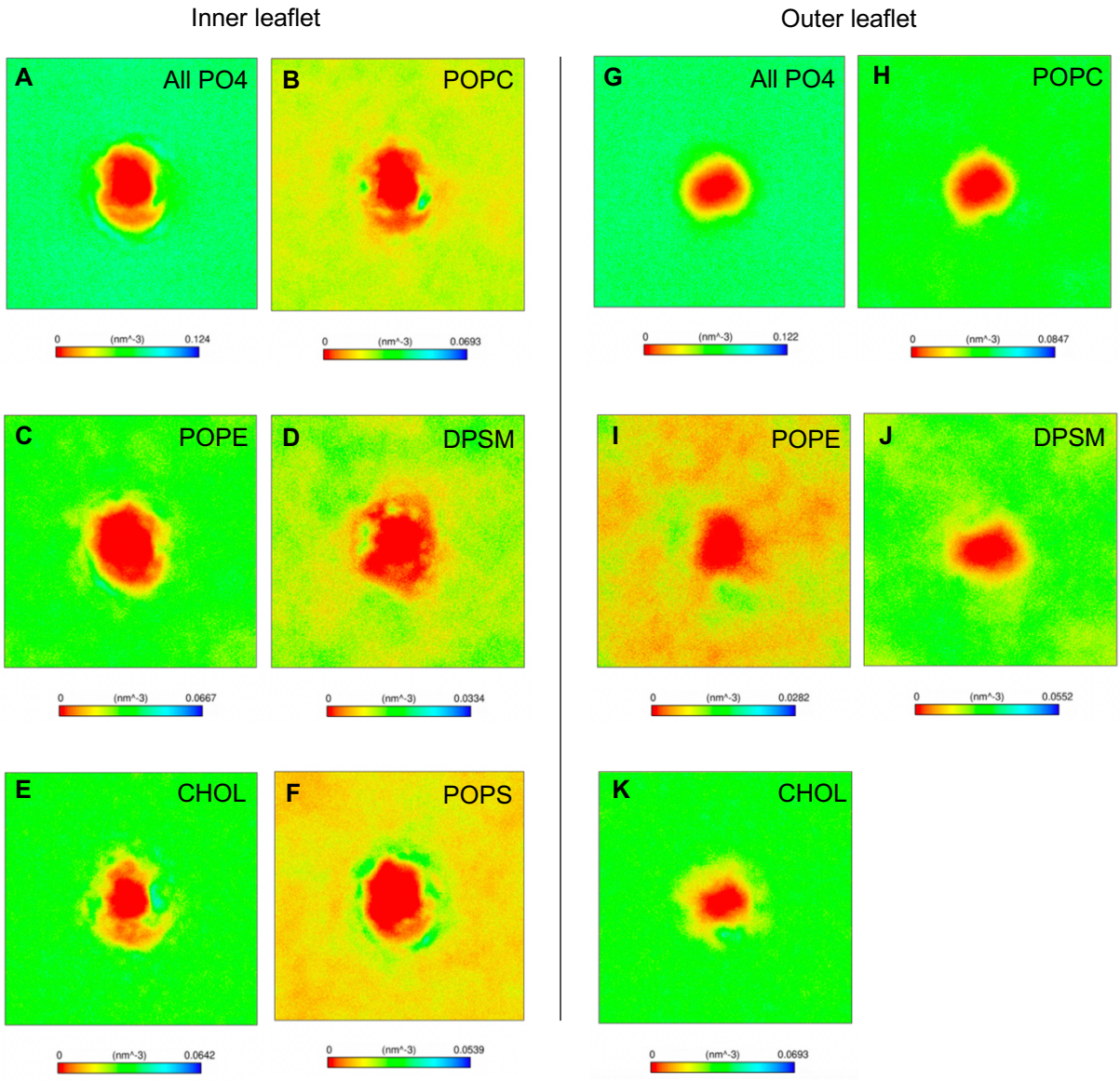

Figure 19. Lipid density maps for the system of P-gp inserted in the complex membrane. Maps on the left side correspond to the PO4 beads of phospholipids and CHOL ROH beads in the inner leaflet: A) All leaflet phospholipids, B) POPC, C) POPE, D) DPSM, E) CHOL and F) POPS. Maps on the right-hand side correspond to the PO4 beads of phospholipids and CHOL ROH beads in the outer leaflet: G) All leaflet phospholipids, H) POPC, I) POPE, J) DPSM and K) CHOL. The density values are indicated by the color scale bars. The region in red corresponds to the space occupied by P-gp.

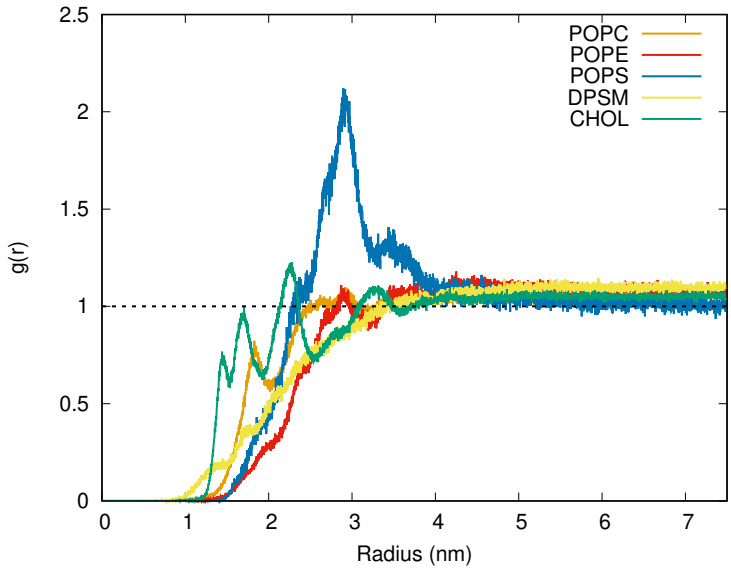

Figure 20. Radial distribution functions (RDF) for the different lipids around the P-gp in the complex membrane. The RDF function for POPC is shown in orange, POPE in red, POPS in blue, DPSM in yellow and CHOL in green. The zero radius was defined as the COM of the TM region of the P-gp.

3.2 Interaction of NBD-C_n molecules with the lipid membrane and with P-gp

Previously in this work, the models of the P-gp embedded in a lipid membrane were characterized. In this section, the interaction of an homologous series of NBD-C_n amphiphiles with both the lipid bilayer and the P-gp will be studied, in order to understand how the length of the alkyl chain influences the partition of the molecules to the membrane and their interaction with the protein.

3.2.1 Transfer of the NBD-C_n molecules between the lipid membrane and the water

The PMF profiles for the transfer of the NBD-C_n molecules from the water to the membrane were calculated using the complex bilayer system without P-gp. This was done with 2 molecules in each umbrella window simulation, so both membrane leaflets were sampled simultaneously. One molecule was placed in the membrane, with the NBD group in the center of the bilayer, and another in the water. The molecules were then pulled across the membrane, as detailed in the Methods section. The results are shown in Figure 21, where the reaction coordinate corresponds to the <z> distance of the NBD groups of the molecules to the membrane COM.

For all NBD-C_n molecules there is a decrease in the free energy with the entry of the molecules to the membrane, and all have an energy minimum when the polar NBD group is in the headgroup region of the membrane. Moreover, the positive free energy variation ($\Delta G°$) for the desorption of the molecule from the membrane to the water increases with the increase in the length of the hydrophobic chain. This has been previously reported using MD simulations with a POPC membrane [55] and is also in agreement with the experimentally measured increase of the membrane/water partition coefficient

with the chain length in amphiphiles [58]. This positive ΔG° for the desorption of the molecules to the water appears to be higher in the inner leaflet than in the outer leaflet in the case of NBD-C_{12} and NBD-C_{16}. When considering the error estimates in the case of NBD-C_{12} these values are not significantly different. However, for NBD-C_{16} there is a small difference between the two leaflets. The lower ΔG° for desorption from the outer leaflet may be due to its higher content in SM, which is typically associated with a higher membrane ordering and therefore lower membrane partition.

The obtained PMF profiles also show an energy barrier associated with the translocation of the molecule from one leaflet to the other that is similar for all NBD-C_n molecules. This suggests that the polarity of the NBD group is the main contributor to the free energy of this process. This is expected because the higher energy state in the translocation process is expected to be the solvation of the polar group in the non-polar center of the bilayer. The ΔG° between the center of the membrane and the energy minimum appears to be higher for the outer leaflet. Moreover, the location of the energy minimum appears to be slightly different for the two leaflets, being at 1.7 nm from the membrane COM for the outer leaflet, and 1.8 nm for the inner leaflet. However, when the RT characteristic energy of the molecules is considered, the values at the energy minima and their location are not statistically different between the two leaflets.

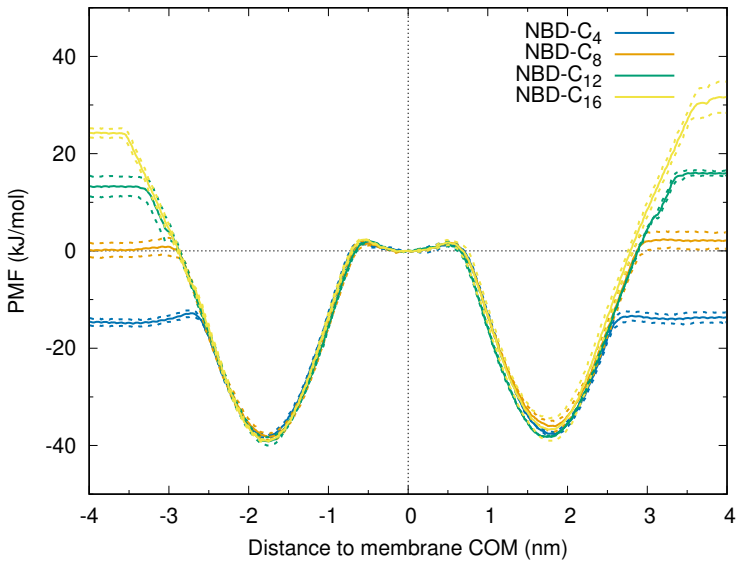

Figure 21. PMF profiles for the interaction of the NBD-C_n molecules with the asymmetric lipid membrane without P-gp. The PMF profile for NBD-C_4 is shown in blue, NBD-C_8 in orange, NBD-C_{12} in green and NBD-C_{16} in yellow. Error estimates are shown in dotted lines. The PMF energy reference was defined at the membrane COM. The negative distance values correspond to positions in the outer leaflet and positive values for the inner leaflet.

3.2.2 Transfer of the NBD-C_n molecules from the lipid membrane to the P-gp

The PMF profiles for the transfer of the NBD-C_n molecules from the membrane to the P-gp were calculated in two different systems: one with the protein embedded in the complex bilayer, and another with the protein embedded in a simpler bilayer composed only by POPC. As described in the Methods section, the PMF profiles for the interaction of the NBD-C_n series with the P-gp changed with the simulation time used for the analysis, and ultimately converged during the 200 ns simulation. Additionally, 50 ns equilibration time was also determined by discarding initial simulation times and was disregarded in subsequent analysis.

The PMF profiles for all NBD-C_n molecules in the complex membrane are approximately constant while the molecules are in the lipid environment far from the P-gp, as shown in Figure 22A and C. They then begin to increase as the molecule approaches the protein, reaching its maximum at the COM of the TM region. In the case of NBD-C_8 and NBD-C_{16} this increase starts at a greater distance from the protein, at approximately 2.3 nm and 2.4 nm respectively, while for NBD-C_4 and NBD-C_{12}, the profile remains constant until 1.9 nm and 1.6 nm (Figure 22C). For NBD-C_4 there is also a noticeable energy barrier near the gate and a free local energy minimum between 0.5 and 0.7 nm. NBD-C_8 also presents a local energy minimum at 0.3 nm, but no local energy minima inside the P-gp binding pocket are found for NBD-C_{12} and NBD-C_{16}. Moreover, the positive $\Delta G°$ for the transfer of the molecules from the membrane to the P-gp binding pocket increases with the increase in the NBD-C_n chain. This result can be explained considering the hydrophobicity of the different molecules. The homologous series of NBD-C_n molecules differ in a single parameter: the size of the highly hydrophobic alkyl chain. The longer the chain is, the higher the preference for the membrane. Therefore, longer alkyl chain molecules are more energetically stable in the membrane environment, which translates in a higher and positive $\Delta G°$ for the transfer to the P-gp binding pocket.

Similarly, for the P-gp embedded in the POPC membrane, there is also a positive $\Delta G°$ for the entry of the molecules to the protein (Figure 22B and D). However, for NBD-C_4, NBD-C_8 and NBD-C_{12} there is a decrease in the PMF as the molecule approaches the P-gp. Indeed, NBD-C_8 and NBD-C_{12} present energy minima near the gate, before the PMF begins to increase. Moreover, when setting the reference of the PMF in the P-gp binding pocket (Figure 22A and B) it can be seen that for both NBD-C_4 and especially NBD-C_8 the PMF at longer distances from the protein is higher in the POPC membrane environment than in the complex membrane. This is not the case for NBD-C_{12} and NBD-C_{16}. This suggests that the NBD group might interact favorably with the other lipids in the complex membrane. However, for NBD-C_{12} and NBD-C_{16} the positive $\Delta G°$ associated with the larger chain length might overshadow this effect.

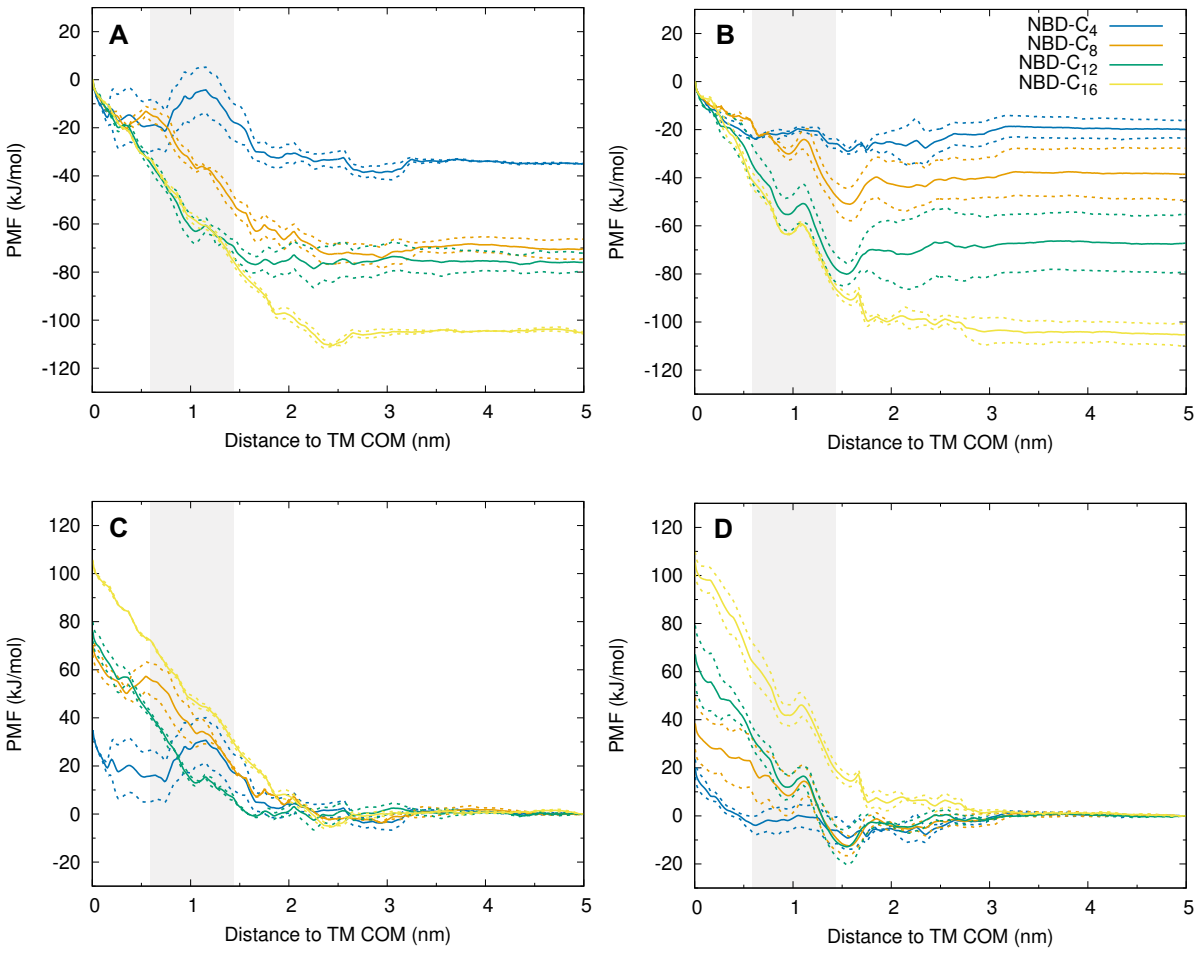

Figure 22. PMF profiles for the transfer of the NBD-C_n molecules to the P-gp, from the complex (A and C) and POPC (B and D) membranes. The PMF profile for NBD-C_4 is shown in blue, NBD-C_8 in orange, NBD-C_{12} in green and NBD-C_{16} in yellow. A) and B) have the PMF energy reference defined at 0 nm from the TM COM, while C) and D) have the PMF energy reference defined at 5 nm. The gray shading symbolizes the gate region of the P-gp. Error estimates are shown in dotted lines.

Overall, these results suggest that the interaction of the NBD-C_n series with the P-gp is unfavorable, especially for the molecules with longer chains. This result is unexpected, since the P-gp has a hydrophobic binding pocket and it is able to efflux many hydrophobic substrates. However, the tendency of the $\Delta G°$ values for the transfer of the several molecules from the membrane to the P-gp binding pocket may be explained with their different hydrophobicity. Moreover, NBD-C_4 and NBD-C_8 show a preference for the complex membrane environment. However, the fact that there is a decrease in free energy near the protein gate in the POPC membrane but not in the complex membrane suggests that the interaction of the molecule with the lipid environment near the P-gp might not be favorable. Specifically, this may be due to the interaction with cholesterol that is enriched near the gates.

3.2.3 Detailed characterization of the interaction of NBD-C_n molecules with the P-gp

The calculation of PMF profiles requires the sampling of the defined reaction coordinate. Therefore, each umbrella window simulation can be analyzed in detail to characterize the main interactions established during the sampling process. Given that there are no energy minima inside the P-gp binding pocket (except for NBD-C_4 and NBD-C_8 in the complex membrane), the interactions of the NBD-C_n series with P-gp residues were studied for the 0.0 nm umbrella window simulation, even though this was found to be the region with the maximum of free energy region in the process.

Firstly, to determine the position of the NBD-C_n molecules in the binding pocket during the 0.0 nm umbrella window simulation, the distance in <z> of the COM of the NBD groups of the molecules was calculated to the membrane COM and to the COM of the TM region of P-gp, as shown in Figure 23. The initial position of the molecules resulted from the pulling simulation, as detailed in the Methods section. All distances were higher when the membrane COM was used as reference. This is expected since some TM helices extend beyond the membrane plane, thus the COM of the TM region does not coincide with the COM of the membrane. Moreover, during this sampling simulation, there was a higher variability in the <z> distance of the NBDs to the membrane COM (Figure 23A) than to the P-gp COM (Figure 23B). This is likely due to the changes in the tilt and fluctuations in <z> position of the P-gp in the membrane.

In this umbrella window, NBD-C_4 and NBD-C_{12} were located more deeply in the binding pocket while NBD-C_8 and NBD-C_{16} were located at an upper position (closer to the NB-Domains). As shown in Figure 23B, there were a few changes in the position of the molecules at the beginning of the simulation, especially for NBD-C_{12}. There were also changes in the distance during the simulation for NBD-C_8 and to a lesser degree for NBDC_4, so the molecules might have interacted with different regions the protein. Even so, the position of the molecules remained mostly constant. This suggests that the position of the molecules was mostly dependent on their initial location. To gain insight regarding preferential location and interactions established with P-gp it would be necessary a much better sampling of the system. This could be achieved through an increase in the simulation time, and preferably through different independent simulations.

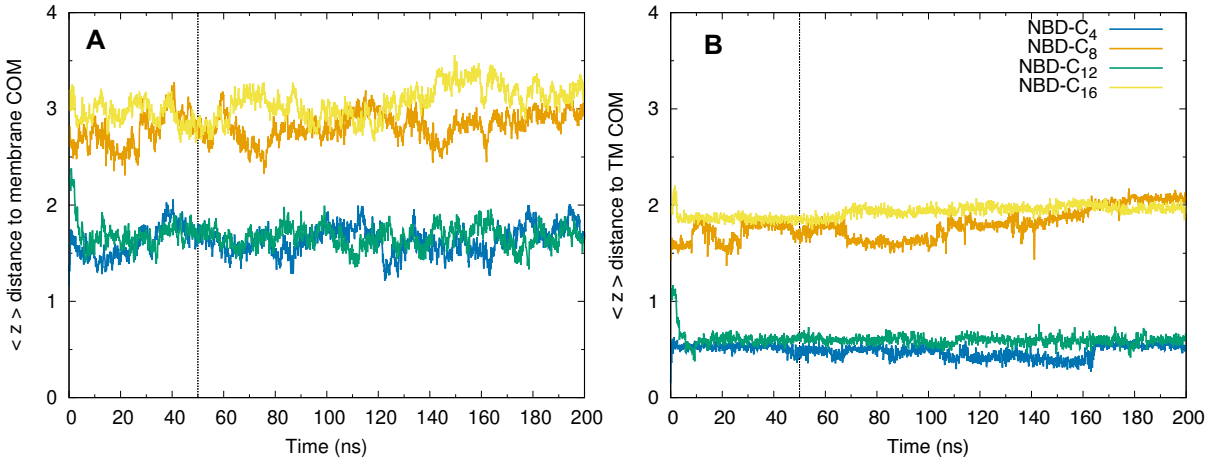

Figure 23. Distance from the COM of the NBD group of the NBD-C_n molecules to A) the membrane COM and B) to the P-gp TM COM during the 0.0 nm umbrella window simulation. The distance for NBD-C_4 is shown in blue, NBD-C_8 in orange, NBD-C_{12} in green and NBD-C_{16} in yellow. The dotted line indicates the discarded 50 ns of equilibration time in the remaining analysis.

The interactions between the NBD group of each molecule and residues of the P-gp during 0.0 nm umbrella window simulation are detailed in Figure 24 and Table 2. NBD-C_4 and NBD-C_{12} interacted with 16 and 15 residues respectively, 12 of which in common, from TM3, TM6, TM10, TM11 and TM12. Most of these P-gp residues have been previously reported to interact with other molecules, both in experimental and MD studies. Namely, with the substrates Hoechst 3342 (Met872), Verapamil (Phe938, Ser989), Nicardipine (Met188, Gln343, Ser345, Pro346), Paclitaxel (Phe938), Morphine (Gln343, Pro346, Glu349), Colchicine (Gln343, Ala991) and inhibitor tariquidar (Asn347, Ala991, Pro992). On the other hand, NBD-C_8 interacted with 17 residues and NBD-C_{16} with 11 residues, all of them common to NBD-C_8. From these residues, Lys177 has also been shown to contact with Paclitaxel, Lys996 with Tariquidar and Thr236, that interacted exclusively with NBD-C_8, has also been shown to interact with Nicardipine. [24,27,32,48] This shows that several of the identified residue interactions for the NBD groups are common to known P-gp substrates. This is especially true for NBD-C_4 and NBD-C_{12}, perhaps due to their more internal position in the binding pocket.

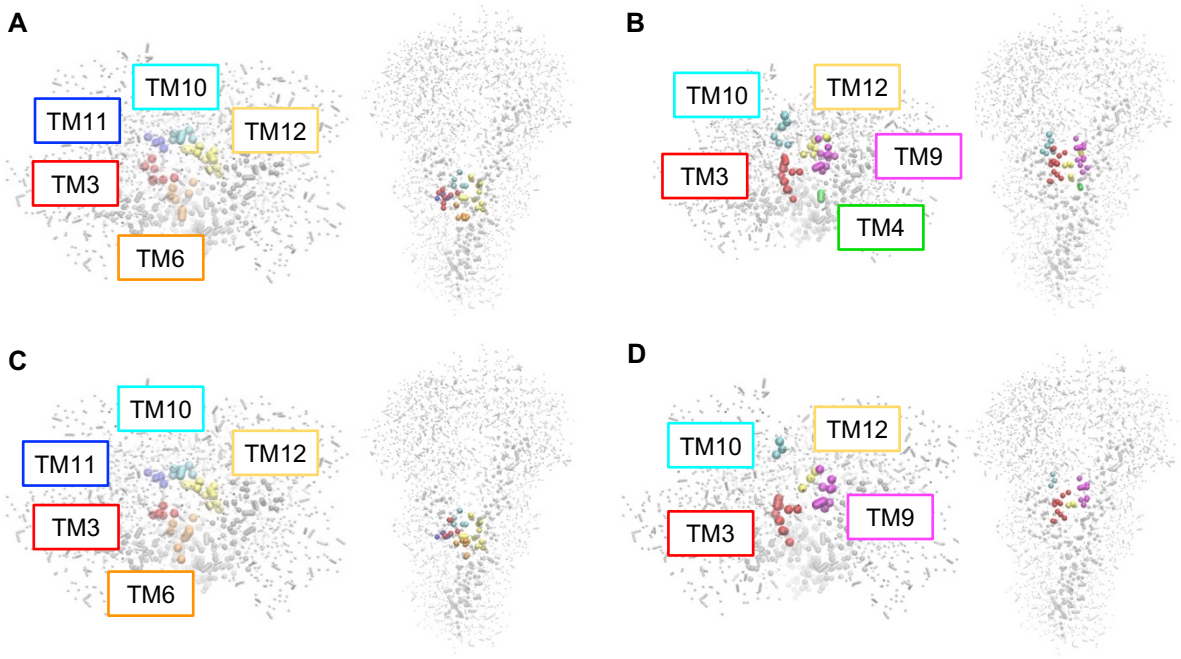

Figure 24. P-gp residues that interacted with the NBD-C_n molecules during the 0.0 nm umbrella window simulation. A) residues that interacted with NBD-C_4, B) NBD-C_8, C) NBD-C_{12} and D) NBD-C_{16}. The P-gp is shown in gray, with the TM4/6 gate in a darker shade. Residues from TM3 are highlighted in red, from TM4 in green, TM6 in orange, TM9 in pink, TM10 in light blue, TM11 in dark blue and TM12 in yellow. The protein is shown both in upper view from NB-Domains (left) and in front view (right).

Table 2. Summary of the interactions between the NBD group of the NBD-C_n molecules and P-gp residues during the 0.0 nm umbrella window simulation. The residues that made more contacts (above the 50th percentile) are highlighted in bold.

Molecule	P-gp region	Residues
NBD-C_4	TM3	Asp184, **Lys185**, Met188
	TM6	Gln343, Ala344, Ser345, **Asn347**
	TM10	**Met872**, **Leu875**, Ser876
	TM11	Phe938
	TM12	**Ser989**, **Phe990**, Ala991, Pro992, **Asp993**
NBD-C_8	TM3	Thr172, **Asp173**, Asp174, **Ser176**, **Lys177**, Glu180
	TM4	Thr236
	TM9	Thr812, Arg813, **Asn816**, Asp817, Gln820
	TM10	Lys883, Lys884
	TM12	Asp993, **Lys996**, Ser1000
NBD-C_{12}	TM3	Glu180, **Lys185**
	TM6	Gln343, Pro346, Asn347, Glu349
	TM10	Met872, **Leu875**, Ser876
	TM11	Phe938
	TM12	**Ser989**, **Phe990**, Ala991, **Pro992**, **Asp993**
NBD-C_{16}	TM3	**Asp173**, Asp174, **Ser176**, **Lys177**, Glu180
	TM9	Thr812, Arg813, **Asn816**, **Asp817**
	TM10	Lys884
	TM12	**Lys996**

3.2.4 Interplay between transfer processes of the NBD-C_n molecules and its influence on the activity of P-gp

As described in the Introduction section, P-gp's efflux activity is a very complex process [15,30,43]. In this work, only the inward-faced conformation of the P-gp was included in the performed simulations. However, some insights about the efflux process can be gained by combining all the results in an integrative approach and considering the contribution of the membrane to this process.

The affinity of a given P-gp substrate is usually evaluated experimentally by its overall effect on P-gp's activity (ATPase activity and/or substrate transport) and considering the concentration of the substrate calculated with respect to the total volume of the solution. Typical results obtained for the effect of NBD-C_4 and NBD-C_8 on P-gp's ATPase activity are shown in Figure 25A (results kindly provided by Maria João Moreno and obtained from a research collaboration with Professor Suresh

Ambudkar, Laboratory of cell biology, CCR, National Cancer Research, NIH, Bethesda). The protein was associated with crude membranes from High Five insect cells that over-expressed P-gp. Therefore, the experiment system contained P-gp, the lipid membrane and the aqueous phase. In Figure 25A it is observed that the presence of both NBD-C_4 and NBD-C_8 led to an increase in the rate of ATP hydrolysis by P-gp. At small substrate concentrations, P-gp activation was more accentuated for NBD-C_8 (Michaelis constant (K_M) = 9 μM). However, at moderately high NBD-C_8 concentrations the ATPase activity leveled off and decreased for total concentrations above 20 μM (Inhibition constant (K_I) = 50 μM). In contrast, NBD-C_4 enhanced the ATPase activity of P-gp at all concentrations tested (K_M = 40 μM), reaching higher activation values than those observed for NBD-C_8. However, to evaluate the relative affinity of both substrates to P-gp it is necessary to know their local concentration in the membrane, from where they interact with the protein. If the total substrate concentration is considered, the affinity obtained is only apparent, and it is not possible to take conclusions regarding the intrinsic affinity for P-gp. To that effect, the partition coefficient of NBD-C_8 to those membranes containing P-gp was also calculated by Maria João Moreno and collaborators, from which the concentration of NBD-C_8 that is associated with the membrane was then calculated. The affinity of NBD-C_4 was too small to be accurately measured and was calculated from the relative values obtained for both amphiphiles in POPC membranes (Cardoso *et al.* [57]). The observed ATPase activity as a function of the NBD-C_n concentration in the membrane is represented in Figure 25B. At the concentration of membrane used in these experiments, about half of the NBD-C_8 is associated with the membranes, while most of NBD-C_4 is in the aqueous phase, with only about 15% associated with the membrane. If one considers that association with P-gp takes place from the membrane (the mechanism usually assumed) then the intrinsic affinity of NBD-C_4 for P-gp is much higher than that of NBD-C_8.

The results obtained in the present work may facilitate the interpretation of these experimental results and contribute to the understanding of the factors that control the intrinsic affinity of substrates for P-gp. The PMF profile from the membrane towards P-gp (Figure 22) shows that NBD-C_4 in the membrane and associated with TM helices of the P-gp has a small difference in potential (less than 20 kJ/mol higher when bound to P-gp). On the other hand, the potential of NBD-C_8 associated with P-gp is approximately 50 kJ/mol higher than when associated with the membrane. This difference in energy is not compensated by the stabilization of NBD-C_8 in the membrane when compared to NBD-C_4 (approximately 15 kJ/mol, Figure 21). Direct comparison of experimental data with the obtained PMF profiles is very complicated and subject to several sources of uncertainty. In any case, the relation between transfer processes, from the membrane to the water and to the P-gp, will define the extent of the difference in intrinsic activity for each P-gp substrate.

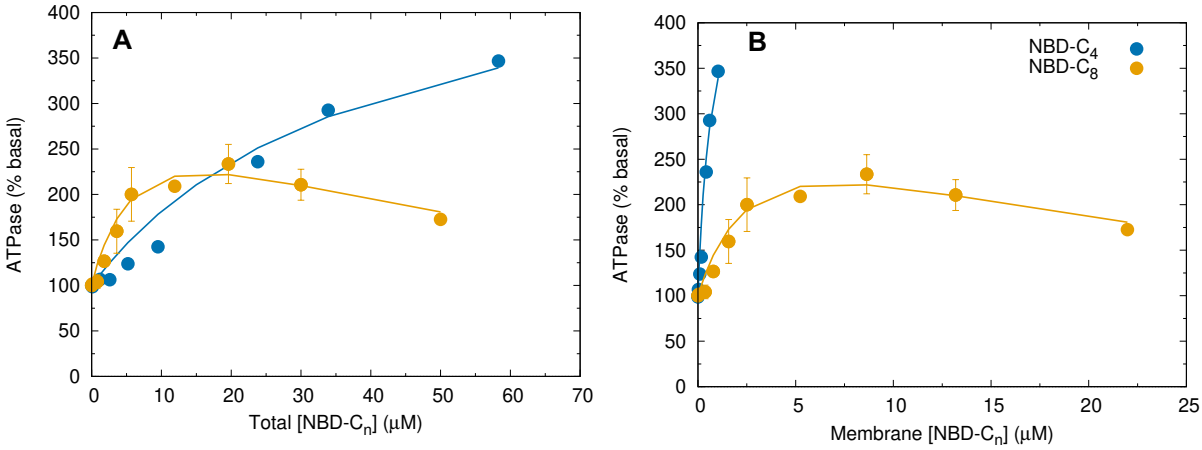

Figure 25. Experimental results for the effect of NBD-C_4 and NBD-C_8 on P-gp ATPase activity, measured by the rate of ATP hydrolysis, considering A) the total concentration of NBD-C_n and B) the concentration of NBD-C_n in the membrane. These results were kindly provided by Maria João Moreno and collaborators.

The higher affinity of NBD-C_4 than NBD-C_8 for the protein can also be evaluated by calculating the difference in energy from the water to P-gp. Having calculated the PMF profiles for the interaction of the NBD-C_n molecules both with the water/membrane and the membrane/P-gp systems, it is possible to estimate the free energy for the direct binding of the molecules in a water/P-gp transfer. Such an estimation assumes that the total sum of the free energy for the transfer of the molecules from the protein to the membrane, then from membrane to the water, and finally from the water to the protein must equal 0, as expected from a circular process. All transfer prosses are summarized in Figure 26. From this, the free energy values for the transfer of the molecules from the water to the protein can be calculated using Equation 7, where $\Delta G°_{\text{Membrane-Protein}} = -\Delta G°_{\text{Protein-Membrane}}$:

$$\Delta G°_{\text{Water-Protein}} = \Delta G°_{\text{Membrane-Protein}} - \Delta G°_{\text{Membrane-Water}} \quad (7)$$

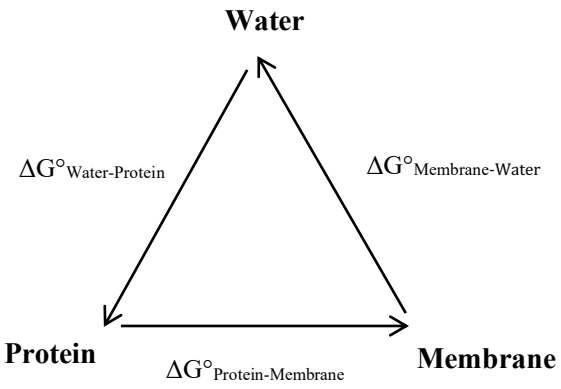

Figure 26. Transfer processes of the molecules between the water, protein and membrane and their respective free energy variations.

The free energy variations for the three processes for each NBD-C_n molecule are detailed in Table 3. Considering that not all NBD-C_n molecules have an energy minimum in P-gp, the $\Delta G°_{Membrane-Protein}$ between the membrane at 5nm and the center of the TM region (0.0 nm) were first considered. The free energy values calculated for the transfer of the molecules from the water to the protein ($\Delta G°_{Water-Protein}$) were positive, which indicates that this is an unfavorable process for all NBD-C_n molecules studied in this work. These values were larger for the molecules with longer alkyl chains, that are more hydrophobic. This is surprising considering the hydrophobicity of the P-gp binding pocket and may in part be due to the use of $\Delta G°_{Membrane-Protein}$ values for 0.0 nm, where PMF values are the highest. For this reason, the same calculations were done for 0.7 nm, where there is a free energy minimum for NBD-C_4 and PMF values are not as high. In this case, $\Delta G°_{Water-Protein}$ were smaller and negative for NBD-C_4, NBD-C_{12} and NBD-C_{16}. However, they were still positive for NBD-C_8. This may point to some technical difficulties in determining the free energy profiles for the transfer from the membrane to the P-gp, due to the size of the molecules. The P-gp may have difficulty in accommodating some large molecules within its binding pocket, especially longer chain molecules such as NBD-C_{16}. Indeed, in this study the alkyl chains of NBD-C_{12} and especially NBD-C_{16} did not fully enter the binding pocket. This could suggest that phospholipids might also not easily enter the binding pocket and be transported. However, NBD-labeled phospholipids have been shown to interact with the protein and their translocation in the membrane is facilitated by P-gp [35]. Alternatively, these results might be indicative of a limitation of the methods used in this study. While the ElNeDyn network is required to maintain the tertiary structure of the P-gp in Martini CG simulations, it might also limit flexibility and ability of the model protein to accommodate larger substrates by altering its structure. Eventually, this issue may translate in an overestimation of the ΔG^o for the transfer from the membrane to the P-gp.

Table 3. Free energies for the transfer of NBD-C_n molecules between the water, membrane and P-gp. $\Delta G°_{Water\text{-}Protein}$ values were calculated by subtracting the values for $\Delta G°_{Membrane\text{-}Water}$ and $\Delta G°_{Membrane\text{-}Protein}$, obtained from the PMF profiles in Figure 21 and Figure 22. Values are shown both for 0.0 nm from the TM region and 0.7 nm.

	0.0 nm			0.7 nm		
	$\Delta G°_{Membrane\text{-}Protein}$ (kJ/mol)	$\Delta G°_{Membrane\text{-}Water}$ (kJ/mol)	$\Delta G°_{Water\text{-}Protein}$ (kJ/mol)	$\Delta G°_{Membrane\text{-}Protein}$ (kJ/mol)	$\Delta G°_{Membrane\text{-}Water}$ (kJ/mol)	$\Delta G°_{Water\text{-}Protein}$ (kJ/mol)
NBD-C_4	34.89	24.01	10.87	13.51	24.01	-10.50
NBD-C_8	70.40	38.19	32.20	51.65	38.19	13.46
NBD-C_{12}	75.99	54.25	21.74	32.70	54.25	-21.55
NBD-C_{16}	105.25	68.35	36.90	63.12	68.35	-5.23

3.3 Interaction of the phytochemicals with P-gp

3.3.1 Transfer of the phytochemicals from the lipid membrane to P-gp

In this section the binding process of a set of phytochemical compounds is studied. In opposition to the homologous series of NBD-Cn amphiphiles, the molecular structures of this phytochemicals are very diverse. Of all the phytochemicals addressed in the present work, only curcumin has been clearly established as a modulator of P-gp. Nevertheless, EGCG and resveratrol have also emerged in the literature as possible modulators.

Here, PMF profiles for the transfer from the complex asymmetric membrane to the P-gp were calculated for all phytochemicals and organized in Figure 27 based on the presence of an energy minimum in the protein. This energy minimum suggests a favorable interaction with the P-gp. The PMF profiles for the interaction of curcumin and genistein with the P-gp are very similar (Figure 27 A and C), gradually decreasing as the molecules approach the protein. Both have an absolute energy minimum inside the P-gp's binding pocket, at approximately 0.7 nm, preceding an increase in free energy as the molecules further enter the binding pocket. For genistein, the energy minimum is of -19.8 ± 0.8 kJ/mol considering the reference of the PMF at 5 nm, and -17.2 ± 3.6 kJ/mol for curcumin (Figure 27C). Additionally, genistein has a smaller local energy minimum at 1.7 nm from the TM COM, at the membrane/protein interface. PMF values for EGCG also decrease as this molecule interacts with the P-gp. EGCG has two energy minima separated by an energy barrier, one of -20.9 ± 0.8 kJ/mol at 1.3 nm at the interface of the membrane with the gate and one of -5.7 ± 1.13 kJ/mol at 0.7 nm inside the protein. As it happens with curcumin and genistein, the maximum value of the profile is at the center of the TM

region of the P-gp. This is indicative that the interaction of the molecules is much more favorable when they are in contact with the TM helices rather than in the center of the TM region.

On the contrary, the PMF profiles for capsaicin and resveratrol (Figure 27 B and D) do not decrease as the molecules start to interact with the P-pg. In the case of capsaicin, the profile remains constant between 5 nm and 2.5 nm, followed by an increase of approximately 10 kJ/mol. Then, there is a local minimum inside the protein, at 0.6 nm, followed by a steep increase in free energy as the molecule nears the TM COM. For resveratrol, the PMF is constant from 5 nm to 3 nm, and then increases until the energy maximum at the center of the TM region. However, as resveratrol approaches and enters the P-gp, the PMF is highly variable, with 5 energy minima and small energy barriers. This might result from poor sampling, that was especially difficult for resveratrol, which would also explain the high values of the error estimate for this molecule.

Altogether, these results suggest that genistein and EGCG may be potential P-gp substrates, able to interact with the P-gp from the membrane along with curcumin. Indeed, both curcumin and EGCG have been previously reported to interact with P-gp [63–65]. On the other hand, the interaction of resveratrol and capsaicin with P-gp is not favorable, which might indicate that these molecules are not likely substrates. However, for the case of resveratrol, this conclusion based only on the membrane/P-gp transfer is contrary to experimental studies that indicate that resveratrol is capable of interacting and inhibiting P-gp function [66–68].

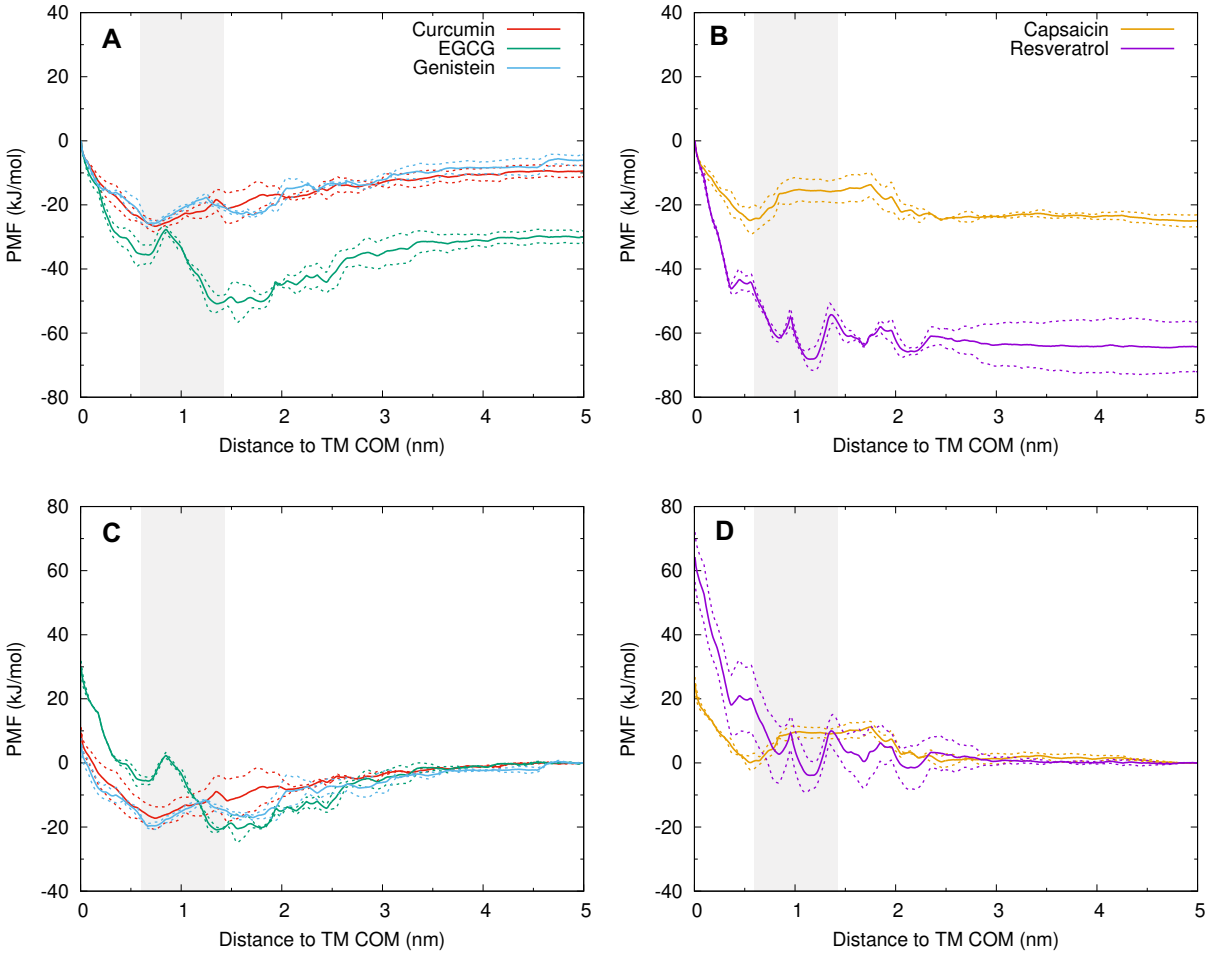

Figure 27. PMF profiles for the transfer of the phytochemicals from the complex membrane to the P-gp. A) and C) show the PMF profiles for the phytochemicals with a clear free energy minimum in the protein: curcumin (in red), EGCG (in green) and genistein (in blue). B) and D) show the PMF profiles for capsaicin (orange) and resveratrol (purple). A) and B) have the energy reference location at 0nm from the TM COM, while C) and D) have the energy reference in the membrane at 5nm from the TM COM. The gray shading symbolizes the gate region of the P-gp. Error estimates are shown in dotted lines.

3.3.2 Detailed characterization of the interaction of the phytochemicals with P-gp

Analogously to the case of the NBD-C_n homologous series, each umbrella window simulation can be analyzed in detail to characterize the main interactions established between the phytochemicals and the P-gp during the sampling process.

During the 0.0 nm umbrella window simulation, capsaicin, EGCG and genistein were located at approximately the same <z> distance from the P-gp (1.4 nm) (Figure 28B). Resveratrol interacted with more internal regions in the protein (0.6 nm) while curcumin was located further away from TM COM. Curcumin also interacted with two regions of the binding pocket, having changed its position at 85 ns, from 2.4 nm to 3nm.

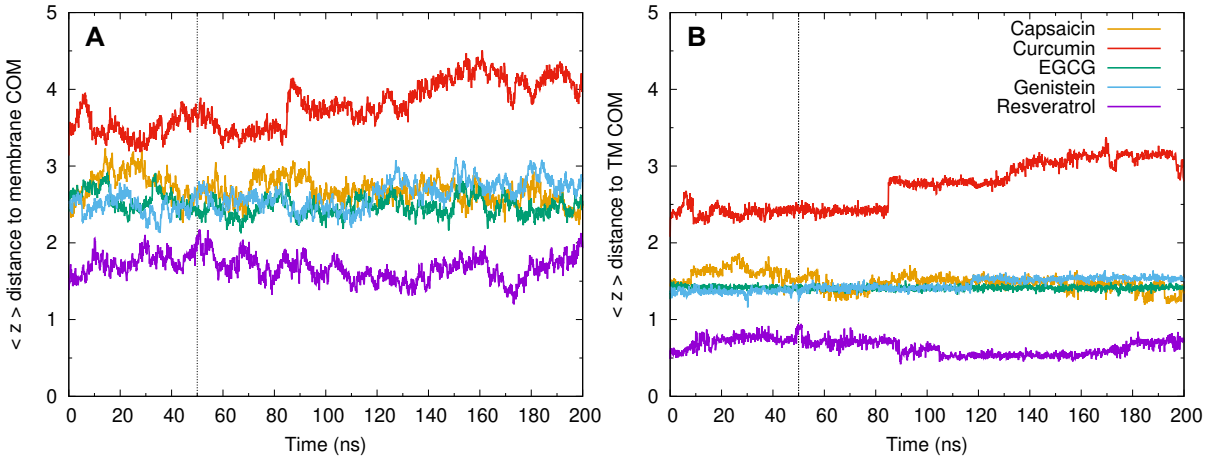

Figure 28. Distance from the COM of the phytochemicals to A) the membrane COM and B) to the P-gp TM COM, during the 0.0 nm umbrella window simulation. The distance for capsaicin is shown in orange, curcumin in red, EGCG in green, genistein in blue and resveratrol in purple. The dotted line indicates the discarded 50 ns of equilibration time in the remaining analysis.

Since there were clear energy minima identified for the phytochemicals, their interactions with protein residues were studied both at 0.0 nm, in the COM of the TM region, and at the umbrella windows corresponding to the absolute energy minima. The interactions are detailed in Figure 29 and Table 4 for the 0.0 nm umbrella windows and in Figure 30 and Table 5 for the energy minima. At 0.0 nm, curcumin interacted with 59 residues, both in the transmembrane domains of the protein, and in the linker, some of which have been shown to interact with paclitaxel (Lys177), nicardipine (Thr236, Asp237), morphine (Asn353) and tariquidar (Asn353 and Arg676). Genistein interacted with 18 residues, some that also interact with paclitaxel (Lys177), nicardipine (Ser233 and Thr236), morphine (Asn353) and tariquidar (Asn353 and Lys996). EGCG interacted with 25 residues, some that also interact with paclitaxel (Lys177), nicardipine (Ser233, Thr236, Asp237 and Pro346), morphine (Pro346, Glu349, Asn353) and tariquidar (Asn353, Pro992 and Lys996) and resveratrol with 20, some in common with paclitaxel (Lys177, Phe934 and Phe938), nicardipine (Gln343 and Pro346), morphine (Pro346), Hoechst 33342 (Met872), colchicine (Gln343, Ala991 and Ser998), verapamil (Ser989), and tariquidar (Ser988, Ala991, Pro992 and Lys996).[24,27,32,48] The fact that curcumin changed its position during the simulation might explain the increased number of interactions for this molecule. Moreover, capsaicin, EGCG and genistein, that showed similar positions in the binding pocket (Figure 28), interacted with 11 residues in common. Particularly, the majority of interactions for Genistein (16 out of 18) are also found for EGCG.

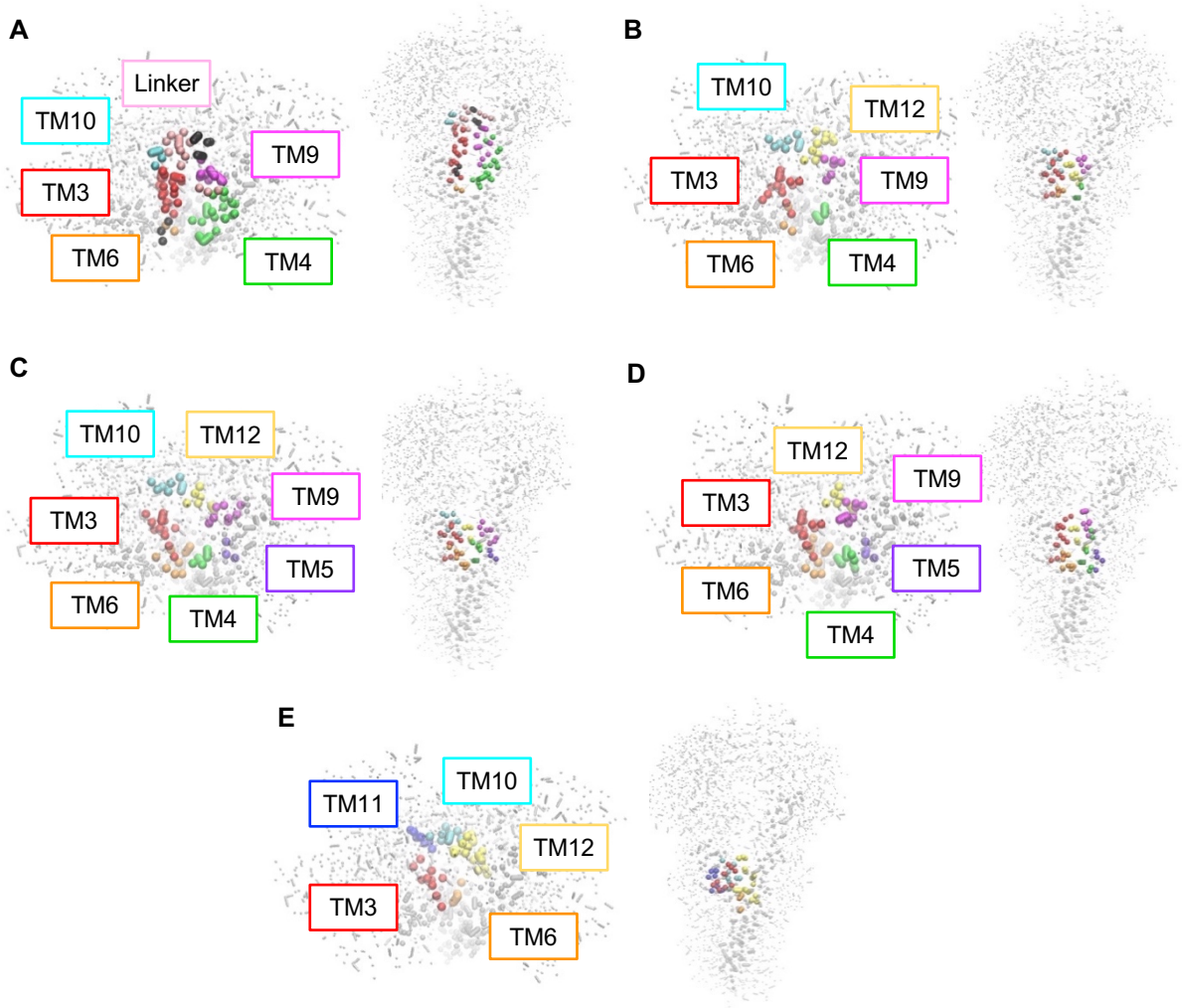

Figure 29. P-gp residues that interacted with the phytochemicals during the 0.0 nm umbrella window simulation. A) residues that interacted with curcumin, B) genistein, C) EGCG, D) capsaicin, and E) resveratrol. The P-gp is shown in gray, with the TM4/6 gate in a darker shade. Residues from TM3 are highlighted in red, from TM4 in green, TM5 in purple, TM6 in orange, TM9 in pink, TM10 in light blue, TM11 in dark blue, TM12 in yellow. The linker is highlighted in light pink and residues not attributed to these motifs in black. The protein is shown both in upper view, from the NB-Domains (left) and in front view (right).

Table 4. Summary of the interactions between the phytochemicals and P-gp residues during the 0.0 nm umbrella window. The residues that made more contacts (above the 50th percentile) are highlighted in bold.

Molecule	P-gp region	Residues
Curcumin	[a]	Lys30
	TM3	**Gly165**, **Glu166**, Asn168, **Thr169**, Arg170, Asp173, Asp174, Ser176, Lys177, Glu180
	TM4	Thr236, Asp237, Glu239, Leu240, His241, Tyr243, Ala244, Gly247, Ala248, **Glu251**
	TM6	Asn353
	Linker	Lys666, **Arg676**, Thr680, Lys681, Glu682
	[b]	Pro803, Asn805, Thr806
	TM9	Thr807, **Gly808**, Ala809, Thr811, **Thr812**, Ala815, Asn816
	TM10	Thr894, Glu898
Genistein	TM3	Asp173, **Ser176**, **Lys177**, **Glu180**, Lys185
	TM4	Ser233, Thr236
	TM6	Ala350, Asn353
	TM9	Asn816, Asp817, Gln820
	TM10	**Ser876**, Ala879, **Leu880**, **Lys883**
	TM12	Asp993, **Lys996**
EGCG	TM3	**Ser176**, **Lys177**, Asn179, **Glu180**, Lys185
	TM4	**Ser233**, Thr236, Asp237
	TM5	Lys287
	TM6	Pro346, Glu349, **Ala350**, Asn353, Ala354
	TM9	**Asn816**, Ala819, **Gln820**, Lys822
	TM10	Ser876, Ala879, **Leu880**, **Lys883**
	TM12	Pro992, Asp993, **Lys996**
Capsaicin	TM3	**Asp173**, Asp174, **Ser176**, **Lys177**, Glu180, Lys185
	TM4	Leu232, **Ser233**, **Thr236**, Leu240
	TM5	Lys287, Ala291
	TM6	**Pro346**, Asn347, Glu349, Ala350, Asn353
	TM9	Thr812, **Asn816**, Asp817, Gln820
	TM12	Pro992, Asp993, **Lys996**
Resveratrol	TM3	Lys177, Glu180, Asp184, **Lys185**
	TM6	Gln343, Pro346
	TM10	Glu871, Met872, **Leu875**, **Ser876**, Ala879
	TM11	Phe934, **Phe938**
	TM12	Ser988, **Ser989**, **Phe990**, Ala991, **Pro992**, **Asp993**, Lys996

[a] While the region was defined as part of TM-Domain 1, it is not part of any TM helix.
[b] While the region was defined as part of TM-Domain 2, it is not part of any TM helix.

For the case of the energy minimum (Figure 30 and Table 5), at 0.7 nm curcumin interacted with 30 residues, including residues that also interact with paclitaxel (Lys177), nicardipine (Lys230, Ser233, Thr236, Asp237 and Pro346), morphine (Pro346, Glu349 and Asn353) and tariquidar (Ala352, Arg355 and Ly996). Genistein interacted with 16 residues at 0.7nm, including residues that also interact with paclitaxel (Lys177), nicardipine (Asp237) and morphine (Asn353). EGCG interacted with 27 in its minimum at 1.3 nm, some in common with paclitaxel (Lys177), nicardipine (Ala229, Lys230, Ser233, Thr236, Asp237, Ser345 and Pro346), morphine (Pro346, Glu349 and Asn353) and tariquidar (Ala352 and Arg355). Capsaicin interacted with 41 residues at 0.5nm, some of which that also interact with paclitaxel (Lys177), nicardipine (Ser233, Thr236, Asp237, Met 295, Gln343, Ser345 and Pro 346), morphine (Gln343, Pro346, Glu349, Asn353), verapamil (Ser989), colchicine (Gln343, Ala991) and tariquidar (Asn347, Ala991, Pro992, Asn353 and Lys996). Resveratrol interacted with 15 residues at 1.1nm, including residues in common with nicardipine (Ala229, Lys230, Ser233, Thr236, Gln343, Ser345 and Pro 346), morphine (Pro346 and Glu349) and tariquidar (Ala352). [24,27,32,48] Notably, curcumin and genistein, that have very similar PMF profiles and energy minima at 0.7nm interacted with 11 residues in common (Lys30, Asp173, Asp174, Lys177, Asp 237, Lys238, Leu240, His241, Asn353, Ala357 and Thr812).

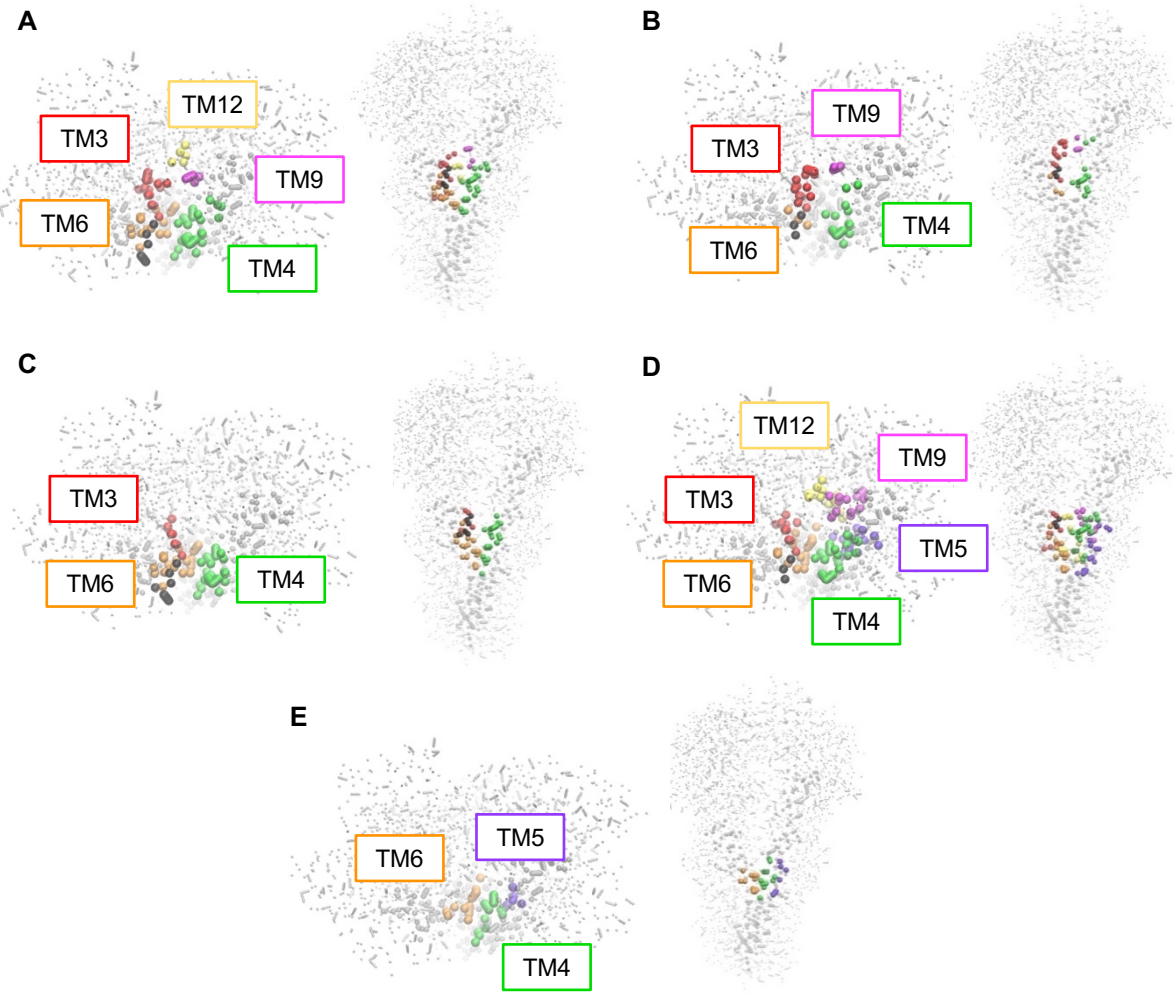

Figure 30. P-gp residues that interacted with the phytochemicals during the umbrella window simulation, at the energy minimum of each molecule. Residues that interacted with A) curcumin (0.7 nm), B) genistein (0.7 nm), C) EGCG (1.3 nm), D) capsaicin (0.5 nm), E) resveratrol (1.1 nm). The P-gp is shown in gray, with the TM4/6 gate in a darker shade. Residues from TM3 are highlighted in red, from TM4 in green, TM5 in purple, TM6 in orange, TM9 in pink, and TM12 in yellow. The linker is highlighted in light pink and residues not attributed to these motifs in black. The protein is shown both in upper view from the NB-Domains (left) and in front view (right).

Table 5. Summary of the interactions between the phytochemicals and P-gp residues during the umbrella window at the energy minima (0.7 nm for curcumin, 0.7 nm for genistein, 1.3 nm for EGCG, 0.5 nm for capsaicin,1.1nm for resveratrol). The residues that made more contacts (above the 50th percentile) are highlighted in bold.

Molecule	P-gp region	Residues
Curcumin	[a]	Lys30, Pro31
	TM3	**Asp173**, **Asp174**, Ser176, Lys177, Glu180, **Lys185**
	TM4	Lys230, **Ser233**, **Ser234**, **Thr236**, **Asp237**, Lys238, Leu240, His241, Tyr243
	TM6	Pro346, **Glu349**, Ala350, Ala352, **Asn353**, Ala354, Arg355, Gly356, Ala357
	TM9	Thr812, Asn816
	TM12	Asp993, Lys996
Genistein	[a]	**Lys30**
	TM3	**Glu166**, **Thr169**, Arg170, **Asp173**, **Asp174**, **Lys177**
	TM4	**Asp237**, Lys238, **Leu240**, **His241**, Glu251
	TM6	Asn353, Ala357
	TM9	Gly808, Thr812
EGCG	[a]	**Lys30**, Pro31, Ala32
	TM3	Asp174, Lys177, Glu180
	TM4	Gly226, Ala229, **Lys230**, **Ser233**, **Ser234**, **Thr236**, **Asp237**, Lys238, Leu240, **His241**
	TM6	Ser345, Pro346, Ile348, **Glu349**, Ala350, Ala352, **Asn353**, Ala354, Arg355, Gly356, Ala357
Capsaicin	[a]	Lys30
	TM3	Asp174, **Lys177**, Glu180, Lys185
	TM4	**Leu232**, **Ser233**, **Thr236**, Asp237, Glu239, Leu240, His241, Tyr243
	TM5	Lys281, **Lys287**, Ala288, Ala291, Asn292, Ser294, Met295
	TM6	Gln343, Ser345, **Pro346**, Asn347, Glu349, Ala350, **Asn353**, Ala354, Ala357
	TM9	Asn816, Asp817, Ala819, Gln820, Lys822, Ser827
	TM12	Ser989, Phe990, Ala991, Pro992, Asp993, Lys996
Resveratrol	TM4	Gly226, **Ala229**, Lys230, **Leu232**, **Ser233**, Thr236
	TM5	Lys287, Ala291, Ser294
	TM6	Gln343, **Ser345**, **Pro346**, Ile348, **Glu349**, Ala352

[a] While the region was defined as part of TM-Domain 1, it is not part of any TM helix.

3.3.3 Interplay between transfer processes of the phytochemicals and its influence on the activity of P-gp

As for the case of the NBD-C_n homologous series, the ΔG° for the transfer of the phytochemicals between the three main environments of the system (water/membrane/P-gp) were also estimated. However, the transfer of the phytochemicals from the water to the membrane was not simulated in this study and both $\Delta G^{\circ}_{\text{Membrane-Water}}$ and $\Delta G^{\circ}_{\text{Water-Protein}}$ needed to be calculated. For this, membrane/water partition coefficients (K_P) were estimated from calculated $\text{LogP}_{\text{Oct/Water}}$, $\text{cLogP}_{\text{Oct/water}}$, values. VolSurf+ was used to obtain the $\text{cLogP}_{\text{Oct/water}}$ values for the phytochemicals and for the NBD-C_n molecules. To validate the accuracy of this method, the $\text{cLogP}_{\text{Oct/water}}$ values for the NBD-C_n series were first compared with experimental $\text{Log}K_P$ obtained from K_P values for the interaction with POPC membranes from Cardoso *et al.* [57]. In Figure 31, a comparison between $\text{cLogP}_{\text{Oct/water}}$ calculated by Volsurf+ and $\text{Log}K_P$ values determined from experimental results is shown. $\text{cLogP}_{\text{Oct/water}}$ underestimates the $\text{Log}K_P$ for values bellow 2 and overestimates the values above 5. However, in the range between 2 and 5 there is a good agreement between the $\text{Log}K_P$ and cLogP values. Considering that the $\text{cLogP}_{\text{Oct/water}}$ of the phytochemicals are within this range, these cLogP can be considered a good estimate of $\text{Log}K_P$ for these molecules. The $\text{cLogP}_{\text{Oct/water}}$ and corresponding K_P values for both the NBD-C_n molecules and the phytochemicals are shown in Table 6, along with NBD-C_n experimental and extrapolated K_P values for POPC membranes by Cardoso *et al.* [57] and Filipe *et al.* [2], respectively.

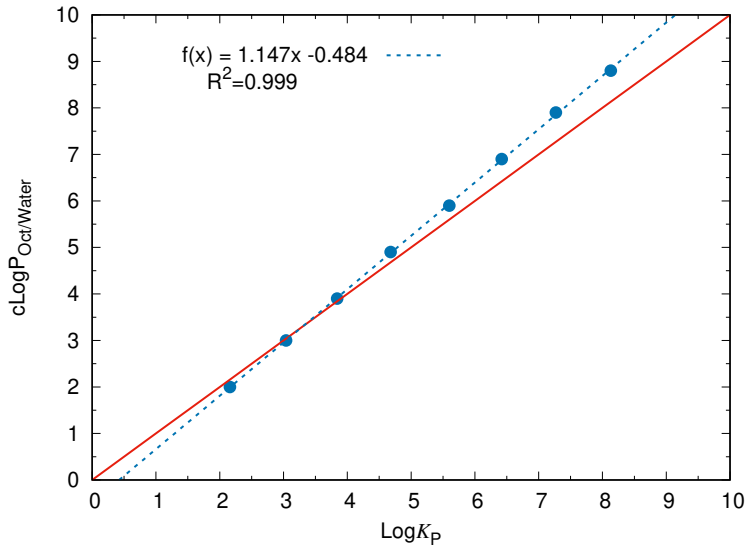

Figure 31. Comparison between the NBD-C_n series $\text{Log}K_P$ values for POPC membranes estimated from the experimental values from Cardoso *et al.* [57] and the $\text{cLogP}_{\text{Oct/Water}}$ values obtained from VolSurf+, shown in blue. The linear regression is shown in blue dotted line and the red line represents the equality $\text{Log}K_P = \text{cLogP}_{\text{Oct/Water}}$.

Table 6. $cLogP_{Oct/water}$ for the phytochemicals and NBD-C_n molecules calculated using Volsurf+, and the estimated K_P values. The experimentally obtained K_P values for the interaction of the NBD-C_n molecules with POPC membranes were retrieved from Cardoso *et* al. [57] and extrapolated values from Filipe *et al.,*2014 [2].

	$cLogP_{Oct/water}$	K_P	K_P values from Cardoso *et* al. [57] and Filipe *et al.* [2]
NBD-C_2	2.00	1.00×10^2	1.46×10^2 [a]
NBD-C_4	3.00	1.00×10^3	1.10×10^3 [b]
NBD-C_6	3.90	7.94×10^3	6.90×10^3 [b]
NBD-C_8	4.90	7.94×10^4	4.80×10^4 [b]
NBD-C_{10}	5.90	7.94×10^5	4.00×10^5 [b]
NBD-C_{12}	6.90	7.94×10^6	2.38×10^6 [a]
NBD-C_{14}	7.90	7.94×10^7	1.66×10^7 [a]
NBD-C_{16}	8.80	6.31×10^8	1.14×10^8 [a]
Curcumin	3.22	1.66×10^3	
Genistein	2.11	1.30×10^2	
EGCG	3.22	1.66×10^3	
Capsaicin	3.39	2.44×10^3	
Resveratrol	2.95	8.87×10^2	

[a] extrapolated K_P values from Filipe *et al.* [2]
[b] experimental K_P values from Cardoso *et* al. [57]

The $\Delta G°_{Membrane-Water}$ for the different phytochemicals were then estimated from the calculated K_P using:

$$\Delta G° = -RT\ \ln(K_P) \quad (8)$$

where R is the molar gas constant and T is the temperature, as described in Hinner *et al.* [74]. Finally, $\Delta G°_{Water-Protein}$ were calculated from $\Delta G°_{Membrane-Protein}$ and $\Delta G°_{Membrane-Water}$ using Equation 7. All $\Delta G°$ values are shown in Table 7, both for the 0.0 nm umbrella window and the energy minimum of each phytochemical. In the case of EGCG and resveratrol values for the absolute minimum outside the protein (1.3 nm and 1.1 nm respectively) and for the local minimum inside the protein (0.68 nm and 0.85 nm) are shown.

The free energy variation associated with the transfer of the phytochemicals from the membrane to the water is positive for all phytochemicals, which is indicative of an unfavorable process. The reverse, that is, the transfer of the phytochemicals from the water to the membrane is therefore a

favorable process, which is in agreement with data obtained experimentally for other phytochemicals[75,76].

For the process of transfer of the molecules from the water to the P-gp two scenarios were considered: i) considering the $\Delta G°_{Membrane\text{-}Protein}$ from the energy minima, ii) considering the $\Delta G°_{Membrane\text{-}Protein}$ from the 0.0 nm position, in the center of the TM region. This process is considered favorable for all molecules (negative ΔG°) when $\Delta G°_{Membrane\text{-}Protein}$ values for the energy minima were used. This suggests that even though only the entry of curcumin, EGCG and genistein to the binding pocket appears to be favorable (Figure 27), all molecules prefer the hydrophobic environment of the membrane, the P-gp binding pocket or the P-gp/membrane interface rather than the water. However, when $\Delta G°_{Membrane\text{-}Protein}$ values calculated for 0.0nm were used in the calculations, the energy associated to the transfer of the molecules from the water to the protein is negative only for curcumin and genistein. On the contrary, capsaicin, EGCG and resveratrol have positive values. Overall, all studied phytochemicals might prefer to be in contact with the TM helices inside the protein or the membrane-protein interface where they might be better positioned and hence establish more favorable interactions. Moreover, the ability of the used model of P-gp to fully accommodate these molecules inside the binding pocket can once again be questioned.

In order to understand how the polarity of the different molecules might help explain the differences in ΔG° between them for the transfer processes considered, the different ΔG° values were plotted as a function of the $cLogP_{Oct/Water}$. In Figure 32A all three ΔG° values, $\Delta G°_{Water\text{-}Membrane}$, $\Delta G°_{Membrane\text{-}Protein}$, and $\Delta G°_{Water\text{-}Protein}$, are shown for the phytochemicals and NBD-C_n molecules. Note that $\Delta G°_{Water\text{-}Membrane}$ is shown, rather than $\Delta G°_{Membrane\text{-}Water}$ as in Table 7. In the case of the NBD-C_n molecules, $\Delta G°_{Membrane\text{-}Protein}$ are higher for the more apolar molecules (higher $cLogP_{Oct/Water}$), with the exception of NBD-C_{12}. Similarly, in the case of the phytochemicals (Figure 32B), the more apolar molecules have a lower $\Delta G°_{Water\text{-}Membrane}$ and a higher $\Delta G°_{Membrane\text{-}Protein}$, with the exception of curcumin. These results suggest that within these moderately apolar phytochemicals, the least apolar interact more with the P-gp. The more apolar molecules have a stronger affinity to the membrane environment and therefore have a smaller tendency to interact with the protein. This analysis of the interplay between transfer processes can be used in the context of drug development to assessment of the propensity of a molecule to bind to the P-gp.

Table 7. Free energy variations for the transfer of the phytochemicals between the water, membrane and P-gp, both for 0.0 nm from the TM region and for the free energy minima. For EGCG and resveratrol values for the absolute minimum outside the protein and local minimum inside the protein are shown. $\Delta G°_{\text{Membrane-Water}}$ values were estimated from $cLogP_{Oct/water}$ values, and $\Delta G°_{\text{Water-Protein}}$ values were then calculated by subtracting the values for $\Delta G°_{\text{Membrane-Protein}}$ obtained from the PMF profiles in Figure 27, and $\Delta G°_{\text{Membrane-Water}}$.

	0.0 nm			minimum		
	$\Delta G°_{\text{Membrane-Protein}}$ (kJ/mol)	$\Delta G°_{\text{Membrane-Water}}$ (kJ/mol)	$\Delta G°_{\text{Water-Protein}}$ (kJ/mol)	$\Delta G°_{\text{Membrane-Protein}}$ (kJ/mol)	$\Delta G°_{\text{Membrane-Water}}$ (kJ/mol)	$\Delta G°_{\text{Water-Protein}}$ (kJ/mol)
Curcumin	9.43	18.36	-8.93	-17.23	18.36	-35.60
Genistein	6.09	12.06	-5.97	-19.67	12.06	-31.73
EGCG	29.96	18.38	11.59	-20.85 [a]	18.38 [a]	-39.23 [a]
				-5.65 [b]	18.38 [b]	-24.03 [b]
Capsaicin	24.95	19.32	5.63	0.12	19.32	-19.20
Resveratrol	64.25	16.82	47.43	-3.85 [a]	16.82 [a]	-20.67 [a]
				2.74 [b]	16.82 [b]	-14.07 [b]

[a] $\Delta G°$ calculated for the absolute minimum outside the protein (1.3 nm for EGCG and 1.1 nm for resveratrol)
[b] $\Delta G°$ calculated for the local minimum inside the protein (0.68 nm for EGCG and 0.85 nm for resveratrol)

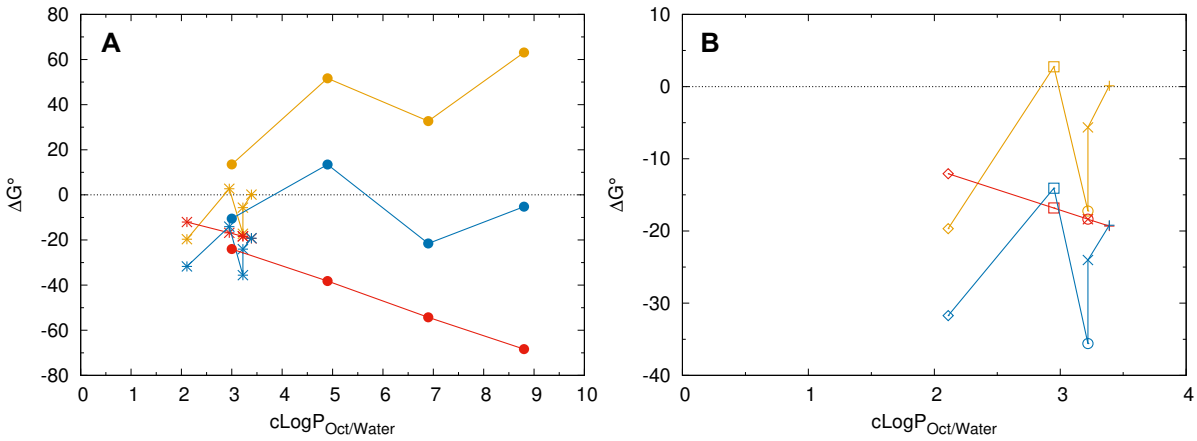

Figure 32. Variation of the $\Delta G°$ values for different transfer processes, $\Delta G°_{\text{Water-Membrane}}$ (red), $\Delta G°_{\text{Membrane-Protein}}$ (yellow), and $\Delta G°_{\text{Water-Protein}}$ (blue), as a function of the $cLogP_{Oct/Water}$ of the different molecules in study. A) $\Delta G°$ values for both the phytochemicals (✳) and the NBD-C_n molecules (●), and B) $\Delta G°$ values for the different phytochemicals: genistein (◇), resveratrol (⊡), EGCG (×), curcumin (⊙) and capsaicin (+). $\Delta G°_{\text{Membrane-Protein}}$ are for the energy minima inside the protein in the case of the phytochemicals and for 0.7 nm for the NBD-C_n molecules.

4. Conclusions

The study of the permeation of molecules through biological membranes is extremely important for the development of new drugs. This is especially true for drugs targeted to the CNS, that have to cross the BBB. Moreover, the efflux of drugs by membrane transporters such as the P-gp present in these biological barriers is another step that must be considered.

In this study, MD simulations with the Martini CG forcefield were used to study the permeation of an homologous series of NBD-C_n molecules through a membrane model of the plasma membrane. Additionally, the interaction of these molecules and a group of phytochemicals with the P-gp binding pocket, starting from their equilibrium position in the membrane, was also characterized. The membrane model used had a complex lipid composition and asymmetric lipid distribution between the two bilayer leaflets in order to resemble a plasma membrane. The model was able to reproduce the enrichment of POPS molecules around P-gp, that is found experimentally. A smaller enrichment in cholesterol and POPC near the protein was also found.

The interaction of the NBD-C_n molecules with the membrane was favorable, with an increasing affinity for the molecules with the longer alkyl chains. Moreover, the main contribution for the energy barrier for the translocation process between the two membrane leaflets must be determined by the NBD group, since it did not change for the different chain lengths of the molecules.

On the other hand, the transfer of the NBD-C_n series from the membrane to the P-gp was unfavorable, and increased with the length of the alkyl chain, both in the complex and POPC membranes. Moreover, the estimated values for the free energy of the transfer of the molecules from the water to the protein were also positive and increased with the length of the alkyl chain. Since the molecules are hydrophobic this was not expected and might be suggestive of a limitation of the used ElNeDyn network, that decreases the flexibility of the P-gp.

On the contrary, the phytochemicals genistein, curcumin and EGCG showed a favorable interaction with the P-gp when transferred from the membrane. This suggests that genistein and EGCG may be potential P-gp substrates, which agrees with experimental evidence that EGCG might interact with the protein [64,65]. The interaction of resveratrol and capsaicin with the P-gp was found to be unfavorable, which would suggest that these molecules are not likely substrates. Estimated values for the free energy of the transfer of the phytochemicals from the water to the membrane and to the protein were negative, which might result from the hydrophobicity of these molecules. Overall, the more apolar molecules, both in the case of the NBD-C_n molecules and the phytochemicals studied had a stronger affinity to the membrane environment and therefore have a smaller tendency to interact with the protein, with the exception of NBD-C_{12} and curcumin.

All NBD-C_n molecules and phytochemicals studied interacted with residues previously identified as binding sites for known P-gp substrates, namely: morphine, nicardipine, Hoechst 33342, Rhodamine 123, paclitaxel, colchicine, the modulator verapamil and the inhibitor tariquidar [27,32,48]. In the case of the NBD-C_n molecules this happened mostly for NBD-C_4 and NBD-C_{12}, that interacted with a more internal region of the binding pocket than NBD-C_8 and NBD-C_{16}. In the case of the phytochemicals, genistein, EGCG and capsaicin were found to interact within the same region in the binding pocket, whereas resveratrol interacted with a more internal region of the protein and curcumin interacted with a more external region.

Although it is possible to determine the binding sites for different molecules within the P-gp through experimental and computational methods, several of the protein residues are common to both substrates and inhibitors. Therefore, when predicting the activity of the P-gp for different molecules, it is essential to consider the interplay between the different processes involved, as discussed in this work.

In future work, PMF profiles for the transfer of the phytochemicals both from the water to the membrane and from the water directly to the protein should be computed for all molecules to confirm the free energy values estimated in this work. Furthermore, it would be important to obtain replicates of these results, using different initial positions and orientations for the NBD-C_n molecules and phytochemicals. It would also be interesting to see if the use of a more flexible protein, with a different elastic network, would result in a more favorable interaction of the NBD-C_n molecules with the P-gp.

5. References

1. Clark, D. E. In silico prediction of blood – brain barrier permeation. *Drug Discov. Today* **8**, 927–933 (2003).

2. Filipe, H. A. L., Salvador, A., Silvestre, J. M., Vaz, W. L. C. & Moreno, M. J. Beyond overton's rule: Quantitative modeling of passive permeation through tight cell monolayers. *Mol. Pharm.* **11**, 3696–3706 (2014).

3. Giacomini, K. M. *et al.* Membrane transporters in drug development. *Nat. Rev. Drug Discov.* **9**, 215–236 (2010).

4. Li, W. *et al.* Overcoming ABC transporter-mediated multidrug resistance: Molecular mechanisms and novel therapeutic drug strategies. *Drug Resist. Updat.* **27**, 14–29 (2016).

5. Gottesman, M. M., Fojo, T. & Bates, S. E. Multidrug Resistance in Cancer: Role of Atp-Dependent Transporters. *Nat. Rev. Cancer* **2**, 48–58 (2002).

6. Abbott, N. J. Blood-brain barrier structure and function and the challenges for CNS drug delivery. *J. Inherit. Metab. Dis.* **36**, 437–449 (2013).

7. Abbott, N. J., Patabendige, A. A. K., Dolman, D. E. M., Yusof, S. R. & Begley, D. J. Structure and function of the blood-brain barrier. *Neurobiol. Dis.* **37**, 13–25 (2010).

8. Daneman, R. *et al.* The Blood–Brain Barrier. *Cold Spring Harb. Perspect. Biol.* **7**, 1–23 (2015).

9. Abbott, N. J., Rönnbäck, L. & Hansson, E. Astrocyte-endothelial interactions at the blood-brain barrier. *Nat. Rev. Neurosci.* **7**, 41–53 (2006).

10. Miller, D. S. Regulation of P-glycoprotein and other ABC drug transporters at the blood-brain barrier. *Trends Pharmacol. Sci.* **31**, 246–254 (2010).

11. Artursson, P., Palm, K. & Luthman, K. Caco-2 monolayers in experimental and theoretical predictions of drug transport. *Adv. Drug Deliv. Rev.* **64**, 280–289 (2012).

12. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **64**, 4–17 (2012).

13. Rankovic, Z. Perspective CNS Drug Design : Balancing Physicochemical Properties for Optimal Brain Exposure CNS Drug Design : Balancing Physicochemical Properties for Optimal Brain

Exposure. *J. Med. Chem.* **58**, 2584–2608 (2014).

14. Abbott, N. J. Prediction of blood-brain barrier permeation in drug discovery from in vivo, in vitro and in silico models. *Drug Discov. Today Technol.* **1**, 407–416 (2004).

15. Sharom, F. J. Complex Interplay between the P-Glycoprotein Multidrug Efflux Pump and the Membrane: Its Role in Modulating Protein Function. *Front. Oncol.* **4**, 1–19 (2014).

16. Kim, Y. & Chen, J. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science* **359**, 915–919 (2018).

17. Szakács, G., Paterson, J. K., Ludwig, J. A., Booth-Genthe, C. & Gottesman, M. M. Targeting multidrug resistance in cancer. *Nat. Rev. Drug Discov.* **5**, 219–234 (2006).

18. Schinkel, A. H., Wagenaar, E., Mol, C. A. A. M. & Van Deemter, L. P-glycoprotein in the blood-brain barrier of mice influences the brain penetration and pharmacological activity of many drugs. *J. Clin. Invest.* **97**, 2517–2524 (1996).

19. O'Brien, F. E., Clarke, G., Dinan, T. G., Cryan, J. F. & Griffin, B. T. Human P-glycoprotein differentially affects antidepressant drug transport: Relevance to blood-brain barrier permeability. *Int. J. Neuropsychopharmacol.* **16**, 2259–2272 (2013).

20. Aller, S. G. *et al.* Structures of P-glycoproteins reveals a molecular basis for poly-specific drug binding. *Science* **323**, 1718–1722 (2009).

21. Li, J., Jaimes, K. F. & Aller, S. G. Refined structures of mouse P-glycoprotein. *Protein Sci.* **23**, 34–46 (2014).

22. Jin, M. S., Oldham, M. L., Zhang, Q. & Chen, J. Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* **490**, 566–569 (2012).

23. Kodan, A. *et al.* Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc. Natl. Acad. Sci.* **111**, 4049–4054 (2014)

24. Chufan, E. E., Sim, H. M. & Ambudkar, S. V. Molecular Basis of the Polyspecificity of P-Glycoprotein (ABCB1): recent biochemical and structural studies. *Adv. Cancer Res.* **125**, 71–96 (2015).

25. Shapiro, A. B. & Ling, V. Positively Cooperative Sites for Drug Transport by P-Glycoprotein with Distinct Drug Specificities. *Eur. J. Biochem.* **250**, 130–137 (1997).

26. Shapiro, A. B., Fox, K., Lam, P. & Ling, V. Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. *Eur. J. Biochem.* **259**, 841–850 (2001).

27. Subramanian, N., Condic-Jurkic, K., Mark, A. E. & O'Mara, M. L. Identification of Possible Binding Sites for Morphine and Nicardipine on the Multidrug Transporter P-Glycoprotein Using Umbrella Sampling Techniques. *J. Chem. Inf. Model.* **55**, 1202–1217 (2015).

28. Sauna, Z. E. & Ambudkar, S. V. About a switch: how P-glycoprotein (ABCB1) harnesses the energy of ATP binding and hydrolysis to do mechanical work. *Mol. Cancer Ther.* **6**, 13–23 (2007).

29. Lugo, M. R. & Sharom, F. J. Kinetic validation of the models for P-glycoprotein ATP hydrolysis and vanadate-induced trapping. Proposal for additional steps. *PLoS One* **9**, (2014).

30. Dastvan, R., Mishra, S., Peskova, Y. B., Nakamoto, R. K. & Mchaourab, H. S. Mechanisms of allosteric modulation of P-Glycoprotein by substrates and inhibitors. *Science* **364**, 689–692 (2019).

31. Condic-Jurkic, K., Subramanian, N., Mark, A. E. & O'Mara, M. L. The reliability of molecular dynamics simulations of the multidrug transporter P-glycoprotein in a membrane environment. *PLoS One* **13**, 1–24 (2018).

32. Ferreira, R. J., Ferreira, M. J. U. & Dos Santos, D. J. V. A. Do Drugs Have Access to the P-Glycoprotein Drug-Binding Pocket through Gates? *J. Chem. Theory Comput.* **11**, 4525–4529 (2015).

33. Sharom, F. J. The P-glycoprotein efflux pump: How does it transport drugs? *J. Membr. Biol.* **160**, 161–175 (1997).

34. Sharom, F. J., Yu, X., Chu, J. W. K. & Doige, C. A. Characterization of the ATPase activity of P-glycoprotein from multidrug-resistant Chinese hamster ovary cells. *Biochem. J.* **308**, 381–390 (2015).

35. Romsicki, Y. & Sharom, F. J. Phospholipid flippase activity of the reconstituted P-glycoprotein multidrug transporter. *Biochemistry* **40**, 6937–6947 (2001).

36. Pohl, A., Lage, H., Müller, P., Pomorski, T. & Herrmann, A. Transport of phosphatidylserine via MDR1 (multidrug resistance 1)P-glycoprotein in a human gastric carcinoma cell line. *Biochem. J.* **365**, 259–268 (2002).

37. Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. & Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **41**, 429–452 (2012).

38. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–949 (2004).

39. Ferreira, L. G., Dos Santos, R. N., Oliva, G. & Andricopulo, A. D. Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384–13421 (2015).

40. Ingólfsson, H. I. *et al.* The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225–248 (2014).

41. Filipe, H. A. L., Moreno, M. J. & Loura, L. M. S. Interaction of 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-Labeled Fatty Amines with 1-Palmitoyl, 2-Oleoyl- sn -glycero-3-phosphocholine Bilayers: A Molecular Dynamics Study. *J. Phys. Chem. B* **115**, 10109–10119 (2011).

42. Ingólfsson, H. I. *et al.* Phytochemicals perturb membranes and promiscuously alter protein function. *ACS Chem. Biol.* **9**, 1788–1798 (2014).

43. Subramanian, N., Condic-Jurkic, K. & O'Mara, M. L. Structural and dynamic perspectives on the promiscuous transport activity of P-glycoprotein. *Neurochem. Int.* **98**, 146–152 (2016).

44. O'Mara, M. L. & Mark, A. E. The Effect of Environment on the Structure of a Membrane Protein: P-Glycoprotein under Physiological Conditions. *J Chem Theory Comput* **8**, 3964–3976 (2012).

45. Corradi, V. *et al.* Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent. Sci.* **4**, 709–717 (2018).

46. Domicevica, L., Koldsø, H. & Biggin, P. C. Multiscale molecular dynamics simulations of lipid interactions with P-glycoprotein in a complex membrane. *J. Mol. Graph. Model.* **80**, 147–156 (2018).

47. Ferreira, R. J., Ferreira, M. J. U. & Dos Santos, D. J. V. A. Insights on P-glycoproteins efflux mechanism obtained by molecular dynamics simulations. *J. Chem. Theory Comput.* **8**, 1853–1864 (2012).

48. Subramanian, N., Schumann-Gillett, A., Mark, A. E. & O'Mara, M. L. Probing the Pharmacological Binding Sites of P-Glycoprotein Using Umbrella Sampling Simulations. *J. Chem. Inf. Model.* **59**, 2287–2298 (2019).

49. Marrink, S. J. & Tieleman, D. P. Perspective on the martini model. *Chem. Soc. Rev.* **42**, 6801–6822 (2013).

50. Uusitalo, J. J., Ingólfsson, H. I., Akhshi, P., Tieleman, D. P. & Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **11**, 3932–3945 (2015).

51. Ingólfsson, H. I. *et al.* Lipid organization of the plasma membrane. *J. Am. Chem. Soc.* **136**, 14554–14559 (2014).

52. Ingólfsson, H. I. *et al.* Computational Lipidomics of the Neuronal Plasma Membrane. *Biophys. J.* **113**, 2271–2280 (2017).

53. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & Vries, A. H. De. The MARTINI Force Field : Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).

54. Periole, X., Cavalli, M., Marrink, S.-J. & Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* **5**, 1–7 (2009).

55. Filipe, H. A. L., Moreno, M. J., Róg, T., Vattulainen, I. & Loura, L. M. S. How to tackle the issues in free energy simulations of long amphiphiles interacting with lipid membranes: Convergence and local membrane deformations. *J. Phys. Chem. B* **118**, 3572–3581 (2014).

56. Hub, J. S., de Groot, B. L. & van der Spoel, D. g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).

57. Cardoso, R. M. S. *et al.* Chain length effect on the binding of amphiphiles to serum albumin and to POPC bilayers. *J. Phys. Chem. B* **114**, 16337–16346 (2010).

58. Cardoso, R. M. S. *et al.* Chain-length dependence of insertion, desorption, and translocation of a homologous series of 7-nitrobenz-2-oxa-1,3-diazol-4-yl-labeled aliphatic amines in membranes. *J. Phys. Chem. B* **115**, 10098–10108 (2011).

59. Singh, N. A., Mandal, A. K. A. & Khan, Z. A. Potential neuroprotective properties of epigallocatechin-3-gallate (EGCG). *Nutr. J.* **15**, 1–17 (2016).

60. Bastianetto, S., Ménard, C. & Quirion, R. Neuroprotective action of resveratrol. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1852**, 1195–1201 (2015).

61. Fattori, V., Hohmann, M. S. N., Rossaneis, A. C., Pinho-Ribeiro, F. A. & Verri, W. A. Capsaicin: Current understanding of its mechanisms and therapy of pain and other pre-clinical and clinical uses. *Molecules* **21**, 1–33 (2016).

62. Spagnuolo, C. *et al.* Genistein and Cancer: Current Status, Challenges, and Future Directions. *Adv. Nutr.* **6**, 408–419 (2015).

63. Lopes-Rodrigues, V., Sousa, E. & Vasconcelos, M. H. Curcumin as a modulator of P-glycoprotein in cancer: Challenges and perspectives. *Pharmaceuticals* **9**, 1–11 (2016).

64. Knop, J. *et al.* Inhibitory effects of green tea and (-)-epigallocatechin gallate on transport by

OATP1B1, OATP1B3, OCT1, OCT2, MATE1, MATE2-K and P-glycoprotein. *PLoS One* **10**, 1–13 (2015).

65. Li, H., Krstin, S. & Wink, M. Modulation of multidrug resistant in cancer cells by EGCG, tannic acid and curcumin. *Phytomedicine* **50**, 213–222 (2018).

66. Khaleel, S. A., Al-Abd, A. M., Ali, A. A. & Abdel-Naim, A. B. Didox and resveratrol sensitize colorectal cancer cells to doxorubicin via activating apoptosis and ameliorating P-glycoprotein activity. *Sci. Rep.* **6**, 1–11 (2016).

67. Bedada, S. K., Yellu, N. R. & Neerati, P. Effect of resveratrol on the pharmacokinetics of fexofenadine in rats: Involvement of P-glycoprotein inhibition. *Pharmacol. Reports* **68**, 338–343 (2016).

68. Nabekura, T., Kamiyama, S. & Kitagawa, S. Effects of dietary chemopreventive phytochemicals on P-glycoprotein function. *Biochem. Biophys. Res. Commun.* **327**, 866–870 (2005).

69. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).

70. Filipe, H. A. L., Pokorná, Š., Hof, M., Amaro, M. & Loura, L. M. S. Orientation of nitro-group governs the fluorescence lifetime of nitrobenzoxadiazole (NBD)-labeled lipids in lipid bilayers. *Phys. Chem. Chem. Phys.* **21**, 1682–1688 (2019).

71. Cruciani, G., Crivori, P., Carrupt, P. A. & Testa, B. Molecular fields in quantitative structure-permeation relationships: The VolSurf approach. *J. Mol. Struct. THEOCHEM* **503**, 17–30 (2000).

72. Wen, P. C., Verhalen, B., Wilkens, S., Mchaourab, H. S. & Tajkhorshid, E. On the origin of large flexibility of P-glycoprotein in the inward-facing state. *J. Biol. Chem.* **288**, 19211–19220 (2013).

73. M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the G. development team. GROMACS User Manual version 5.1.4. (2016).

74. Hinner, M. J., Marrink, S. J. & De Vries, A. H. Location, tilt, and binding: A molecular dynamics study of voltage-sensitive dyes in biomembranes. *J. Phys. Chem. B* **113**, 15807–15819 (2009).

75. De Granada-Flor, A., Sousa, C., Filipe, H. A. L., Santos, M. S. C. S. & De Almeida, R. F. M. Quercetin dual interaction at the membrane level. *Chem. Commun.* **55**, 1750–1753 (2019).

76. Filipe, H. A. L. *et al.* Differential targeting of membrane lipid domains by caffeic acid and its ester derivatives. *Free Radic. Biol. Med.* **115**, 232–245 (2018).