In silico exploration of membrane transport events : prediction of passive permeation coefficients and binding sites of human MRP4

Mehdi Benmameri

Ph.D. Thesis

Université de Limoges

University of Limoges

ED 615 - Sciences Biologiques et Santé (SBS)

Inserm UMR1248 – Pharmacology and Transplantation

Submitted for the degrees of
Doctor of the University of Limoges
Discipline / Specialty: Computational Chemistry

Presented and defensed by
Mehdi Benmameri

On December 8, 2022

In silico exploration of membrane transport events: prediction of passive permeation coefficients and binding sites of human MRP4

Thesis directed by: Gabin Fabre and Patrick Trouillas

JURY:

Reporters
Dr. Maria João Moreno, Biological Chemistry Group, University of Coimbra, Portugal
Dr. Karel Berka, Physical chemistry, Palacký University, Olomouc

Examiners

Dr. Gabin Fabre, INSERM UMR1248, University of Limoges

Pr. Patrick Traouillas, INSERM UMR1248, University of Limoges
Pr. Claire Rossi. University of Technology of Compiègne
Pr. Nicolas Picard, Pr, INSERM UMR1248, University of Limoges

À mon grand-père

Acknowledgements

First of all, I thank god for everything he gave it me!

My thesis took place within the INSERM UMR1248 Pharmacology and Transplantation. I would like to thank its director, Pr. Pierre Marquet for welcoming me in the unit.

My biggest thanks and gratitude go to my supervisors Gabin Fabre and Patrick Trouillas. Gabin, it was really a big honor for me to work with you. Your experience and your guidance allowed me to explore and understand different things. It was a pleasure to travel with you, to explore different cities of Europe and to try different food cultures. Thanks you for your support over the past years, it was very important for me. Patrick, I learned a lot from you. Your expertise, your comments, your vision and your ability to manage projects are just incredible.

I thank Dr. Maria João Moreno and Dr. Karel Berka for accepting to evaluate my thesis as reporters of this thesis work.

I would like to express my thanks to the University of Limoges for the financial support. I also thank Xavier Montagutelli for the informatics support and for fixing all the problem that I faced during my calculations.

It is surely time to thank those who have surrounded me during these three years. I would like to thank Florent Di-Meo for the interesting discussions, advices and support throughout this thesis. I would also like to thank Benjamin Chantemargue for his support and guidance, especially for the membrane protein part. I would also thank all members of the our lab: Emmanuelle, Hélène, Nicolas, Jean-Seb, Roland, Anne, Nouzha, Antoine, Quentin, François, Hassan, Eliès, Marving, Rudy, Manon, Roy, Clarisse, Aurélien and all the others.

Special thanks to Emma again for the daily help, for listening and for the mental support. I thank also Karen for managing the travels and the administrative work. I thank also Maxime Jouaud for his encouragements and his mental support over the past years.

I thank my office mates, Agota and Angelika, and Veronica (Arcobaleno) for sharing science, for the interesting discussions and for the funny moments. I wish you good luck for your future projects!

To my family and friends

I want also to thank those who are dear to me. Thank you for your support, you were always here for me. I appreciate this. During the three past years, you had an important role in supporting me and organizing chilling plans so that I could change the atmosphere and take a break. This was important for the progress of my thesis. Thank you for the interesting discussions and for the funny moments.

My next thanks are addressed to my lovely sister Dounia and her kids Adam and Youcef, and my brother Zinou. I wish you the best in your life. So proud of you!

To my parents

I also have to thank my parent. Thank you for always being here for me, for listening, for your permanent support and for the unconditional love. Thank you for everything!

Also, I thank all my family members from my grandparents to the youngest fruit.

Rights

Table of Contents

List of Figures

List of Tables

Introduction

The search for novel drug candidates requires the optimization of various structural and physicochemical properties, before the drug is accepted on the pharmaceutical market (Figure 1). The structural optimization of new leads relies on the improvement of their interactions with cell components and membranes. Understanding membrane-drug interaction is of utmost importance to increase bioavailability and transport, to decrease toxicity, to avoid possible drug-drug interactions (DDI) that may lead to cytotoxicity, and eventually to adjust doses. To be approved, the “ultimate” drug candidate should be safe (from the toxic point view) and effective at low dose.

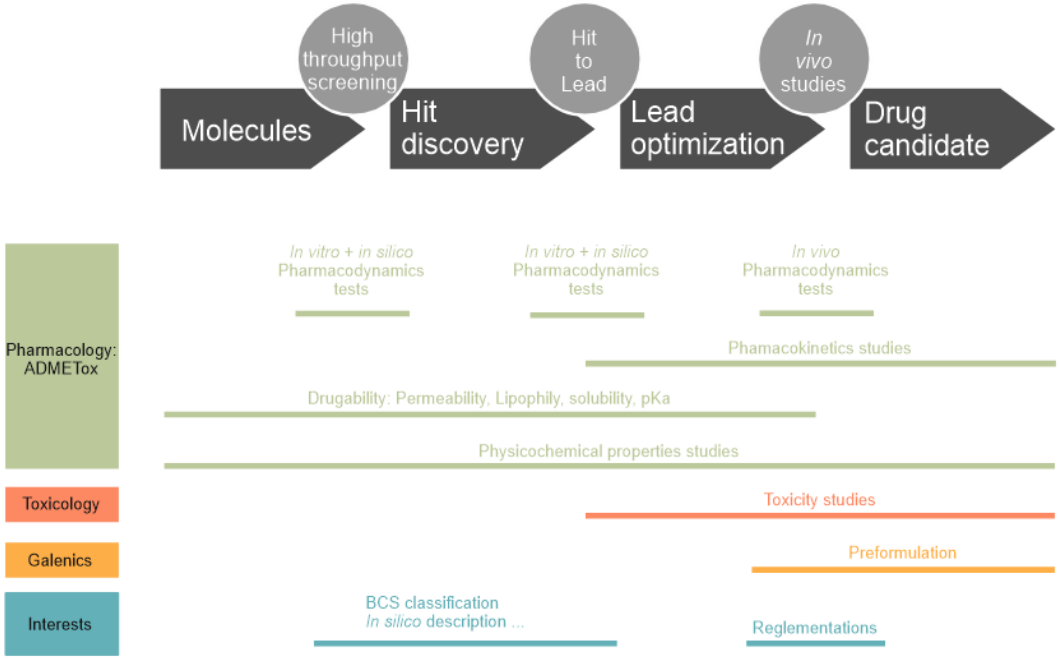


Figure 1: Different stages of drug development before clinical studies on humans

Eukaryotic cells are complex organisms that contain nucleus - the genetic reservoir -, different organelles such as mitochondria and endoplasmic reticulum, different intracellular macromolecules (i.e., enzymes), and endogenous compounds (i.e., cytokines, hormones …). They are delimited by membranes, which serve as barriers against harmful and undesired particles. Cell membranes are mainly composed of phospholipids, embedded proteins and carbohydrates. The cell membrane composition is complex and it changes from one tissue to another, i.e., the ratio of the different lipid types and the expression level of functional membrane proteins vary.[1,2] Cell membranes are highly inhomogeneous and asymmetrical, as being very complex molecular assemblies. They mainly contain lipids of many types, sugars, and embedded proteins.

Pharmacology and toxicology study the fate of foreign molecules (xenobiotics) in animals and humans. Xenobiotics can be drugs, plant-derived compounds (e.g., vitamins, polyphenols), chemicals issued from industry, food additives, or pollutants. To reach their sites of action, the vast majority of xenobiotics have to cross membranes, for instance to reach the blood from the digestive track or from the skin, or to penetrate cells. Therefore, xenobiotic-membrane

interactions are key in pharmacology and toxicology. The disposition of a xenobiotic or its metabolites in each organ cell depends on its interaction with cell membranes. Membrane crossing events can be mediated by membrane transporters or occur by spontaneous permeation through lipid bilayers, or both. However, simple lipoidal permeation is the predominant mechanism for most small and amphiphilic drug-like molecules.[3]

The interaction between xenobiotics and lipid bilayers is therefore very important in pharmacology, toxicology, and drug design in general. Partition coefficients of xenobiotics in membranes may affect the extent of the pharmacological effects, but also the toxicity caused mostly by drug accumulation inside the organ cells[4], and interactions with other molecules such as drug-drug interactions. On the other hand, passive permeation is directly related to pharmacokinetics (PK). It is characterized by the permeability coefficient, for which an accurate prediction would be highly valuable in pharmaceutic, cosmetic and food industries.

Historically, membrane passive permeation was originally described by the solubility diffusion model (SDM).[5,6] It originates from Overton's rule that states that the permeability coefficient is correlated to the partition coefficient between oil and water.[7] Several experimental techniques are used to tackle the passive permeation process, such as liposome-based[8–10] or cell-based[11,12] assays. It has been shown that *in vitro* membrane models are more permeable to amphiphilic molecules than hydrophilic compounds and ions.[13–15] A balance between hydrophilicity and lipophilicity is however required. To date, no simple model can accurately and consistently predict the coefficients of permeation through biological membranes, due to the structural complexity of biological membranes as well the complexity to decipher xenobiotic-membrane interactions.[16]

Although the Overton's and Meyer's theory[7,17] is useful to assess partition in homogeneous environments, the lipid bilayer environment is highly heterogeneous and this theory quickly reaches its limitations. Extensive studies were based on the SDM to develop the inhomogeneous solubility diffusion model (ISDM) theory.[18–22] Based on the ISDM theorem, many tools were implemented to understand membrane passive permeation.

MD simulations are efficient to describe intermolecular interactions between a xenobiotic and its target (i.e., protein, membrane …) at the atomic scale. In the present work, we study the two types of xenobiotic transport (passive permeation and transport through a membrane exporter). Concerning the former type, we built an efficient tool (MemCross) to assess membrane permeation based on ISDM and using biased MD simulations. This tool was tested on a large series of drug-like molecules and the predictions were compared to experimental data. Concerning the latter type of transport, we have explored the modes of binding of a large number of known substrates and modulators of human ABCC4/MRP4, thorough a docking study.

This thesis is divided into seven chapters. Chapter I is an overview of lipid bilayer membranes, their composition and their relationship with diseases. The different lipid properties that are useful to understand lipid membranes behaviour are listed at the end of this chapter.

Chapter II tackles the basic definitions related to the different transport processes across membranes, in particular passive permeation and active transport through membrane transporters. The different families of membrane proteins involved in the active transport of endogenous substrates and xenobiotics are overviewed, and their involvement in pharmacology is discussed. Focus is given on the specific case of the MRP4 transporter, which

is involved in drug efflux in the kidney and the liver, in relation with cytotoxicity in particular met in a specific clinical situation, namely graft rejection.

Chapter III tells the historical evolution of the permeability assessment, from the Overton's rule to the development of the inhomogeneous solubility diffusion model (ISDM), which is derived from the Smoluchowski equation. Furthermore, different *in vitro* membrane models and experimental tools for the quantification of drug-membrane interaction and assessing drugs permeability are listed.

Chapter IV highlights different theoretical methods used to assess passive permeation of drugs. It starts with the description of the underlying basic theory of molecular dynamics (MD) simulations. The chapter continues with a detailed description of the different enhanced methods, used in MD simulations, to model the potential of mean force (PMF) picturing the translocation of molecules across the lipid bilayer membrane. Another focus is given to the most used techniques employed to predict the diffusivity profiles across the membrane.

Chapter V describes a study about the first use of the Accelerated Weight Histogram method to compute the PMF profile of 13 molecules with different sizes and scaffolds. Fractional diffusivity profiles were calculated from the same AWH trajectories. Hence, permeation coefficients were calculated. Based on a comparison between the AWH method and the alchemical calculations, the AWH method rises clear advantages upon other similar methods. The AWH methodology used is named MemCross.

Chapter VI shows an extension of the usage of MemCross to predict permeation coefficients on a larger series of molecules, exhibiting different physical-chemical properties. This study aims at providing the validation of the robustness of MemCross.

Chapter VII investigates the different binding sites of the human MRP4 using a thorough molecular docking of hundreds of known hMRP4 modulators and substrates.

Chapter I. Membranes in molecular pharmacology

I.1. ADME

PK is the science that studies the fate of xenobiotics in the body. PK consists of four main steps: Absorption, Distribution, Metabolism and Excretion (ADME, see Figure 2). PK should be distinguished from pharmacodynamics (PD), which describes the effect of xenobiotics on the organism, especially based upon its interaction at its biological targets, namely the sites of action such as enzymes or receptors. The ADME process reflects the different transfers of xenobiotics between different fluid compartments. It is mainly related to xenobiotic-membrane interactions, which is often difficult to study, making prediction of ADME a difficult task.[23,24]

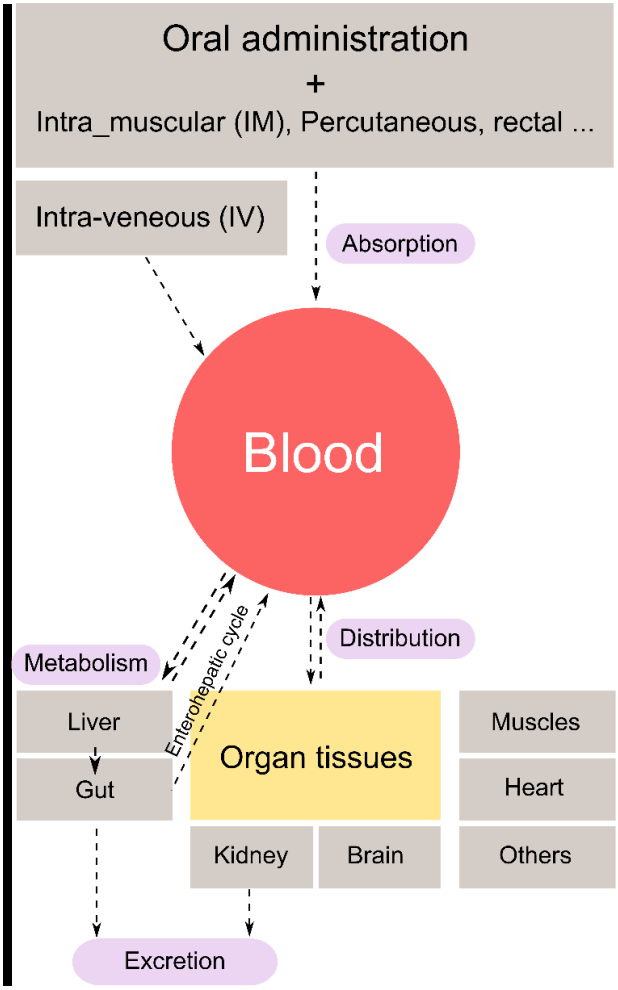


Figure 2: Interconnection of the absorption, distribution, metabolism, and excretion (ADME) steps from xenobiotics administration to excretion.

I.2. Absorption

Absorption is the passage of the xenobiotic from its site of administration into the plasmatic compartment (extracellular fluid). It determines the quantity of drugs to be administered to reach a desired concentration in the blood. This is directly related to the bioavailability of drugs, which is the portion of xenobiotics that reaches the systemic circulation before entering organ

cells. Bioavailability is highly dependent on the xenobiotic chemical structure, but also on inter-individual variability of the organisms. The most common administration mode is the oral intake (see Figure 2), for which the epithelial gastrointestinal mucosa constitutes the first barrier that xenobiotics encounter. It consists of a thin layer of epithelial cells that xenobiotics must cross to reach the systemic blood circulation.[25] Membrane crossing, in relation to ADME mechanisms, relies mainly on two processes[3]: (i) passive permeation through the lipid bilayer (see Chapter III), or (ii) carrier-mediated transport (solute carriers "SLCs", ATP-binding cassette proteins 'ABCs") (see Chapter IV).

Passive absorption depends on several physicochemical factors such as lipophilicity, solubility and acid-base constants (pK_a).[23] These parameters have a fundamental role in the high-throughput screening (HTS) where they are mostly considered as criteria to classify bad and good hit molecules. For instance, drugs with low solubility in lipids are generally poorly absorbed passively.

Overton and Meyer were the first who demonstrated that membrane crossing relies on a series of physicochemical properties of compounds (see Section III.2 for more details).[7,17] Furthermore, the rule of Lipinski (rule of five)[26] came to refine the relationship in which a series of cut-offs were set to select possible biologically active and safe molecules (see Section III.3 for more details).

I.3. Distribution

Once in the systemic circulation reached, xenobiotics spread widely *via* the bloodstream and they diffuse through a series of interconnected compartments. Thus, blood flow provides a rapid distribution independent of the chemical nature of xenobiotics. Conversely, the bioavailability of a xenobiotic in a given organ strongly depends on its physicochemical properties (e.g., molecular weight, solubility) mainly due to the fact that it should traverse cell barriers. There are three types of barriers: (i) the epithelial barrier (e.g., gastrointestinal epithelium, renal tubule); (ii) the vascular endothelium and (iii) the blood brain barrier (BBB). Their composition differs from each other and from one tissue to another[27], which may affect xenobiotics permeability. The BBB strongly limits permeation due to tight junctions between the epithelial cells, preventing highly hydrophilic molecules to cross. The distribution of xenobiotics is ruled by both passive permeation and active transport through membrane transporters. The distribution between the different fluid compartments depends on various physicochemical parameters.[4,25] Once in the cell, xenobiotics may bind to its target receptor and exert its therapeutic effect through various mechanisms (e.g., signaling, antagonism, agonism).

I.4. Metabolism and excretion

To limit their accumulation and potential toxicity, xenobiotics are gradually eliminated by the organism. However, before being excreted, most xenobiotics must be metabolized to increase their solubility. Xenobiotics metabolism consists of enzymatically-driven drug biotransformation leading to metabolites. The predominant organ where metabolism occurs is the liver through its hepatobiliary system.

Metabolic biotransformations consist of two phases named phase 1 and 2 reactions, whereby Cytochrome P450 (CYP) is the major enzymatic cell machinery system involved in metabolism of xenobiotics. Phase 1 consists of hydrolysis, oxidation or reduction reactions. This usually leads to more reactive chemical products. Phase 2 are conjugation reactions, which increases

the polarity to increase secretion in the urine and the bile. As an example, biliary acids are synthesized in the liver from cholesterol through a cascade of enzymatic reactions, mainly ensured by CYP27A1 and CYP7A1. They are then conjugated with glycine to enhance their solubilization in the bile and their delivery to the intestine.[28] A high polarity is required for an efficient elimination, otherwise, xenobiotics may be reabsorbed through the renal tubule.

Some of inactive metabolites released in the bile may be recycled again through the enterohepatic cycle (Figure 3). Once in the intestine, they are usually hydrolyzed into free drugs/chemicals and they are then reabsorbed from the intestine to start again a new cycle. This can lengthen the journey of some drugs or their metabolites in the body.

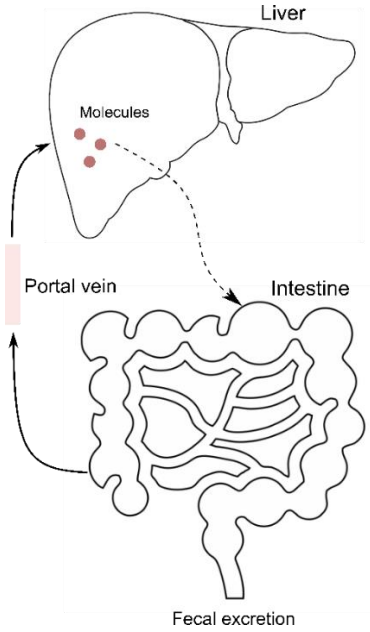


Figure 3: Enterohepatic cycle. First, bile acids are released in the intestine through the bile tract. Then, they can restart a new cycle and reach the liver through the portal vein, or they can be eliminated immediately with feces.

I.5. Lipid bilayer membranes

Each step of ADME requires xenobiotics to cross lipid bilayer membranes, either by passive permeation (detailed in Chapter III and Chapter IV), by facilitated diffusion, or by active transport (detailed in II.1.2). Understanding the complexity of lipid bilayer membranes is crucial for pharmacology. Lipid bilayer membranes (Figure 4) are key components that maintain cell architecture and function. They define cell boundaries and intracellular organelle barriers. Cell membranes ensure many essential physiological functions, wherein interactions between the membrane and cellular components (e.g., enzymes, cytokines) maintain the cellular homeostasis.[29–32] Alterations of the bilayer structure are usually associated with various diseases, [30,33,34] in particular lipid bilayer structural properties are known to influence the function of transmembrane proteins.[35–37] Membranes are highly dynamic and complex structures which make the environment along the bilayer leaflets highly heterogeneous. Apart from lipids, biological membranes contain many other components among which a wide range of proteins.[2,38] Membrane-related proteins ensure several functions such as cellular uptake or efflux of drugs or ions, biotransformation, chemical reactions and signal transduction.[39] The

conformations of the proteins within the lipid bilayer are influenced by lipid composition, lipid phase and lipid properties.[40,41]

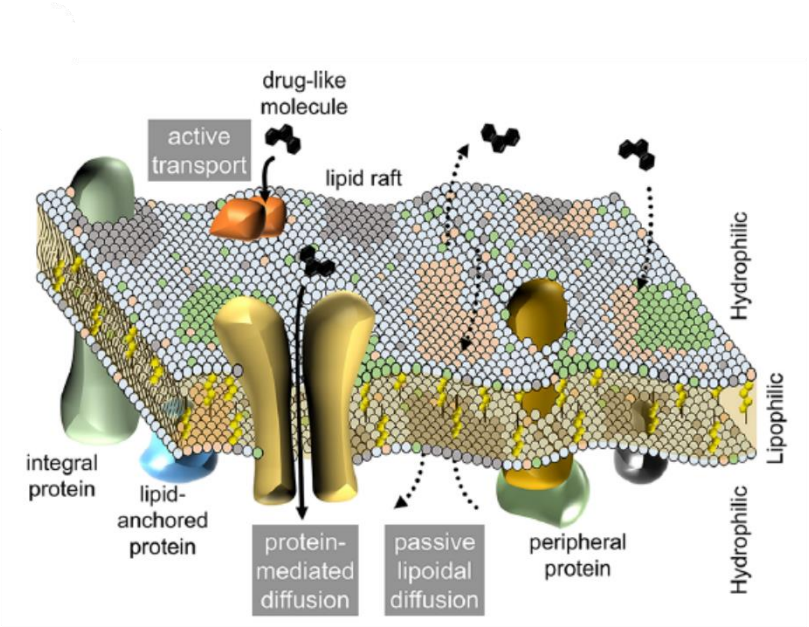


Figure 4: Fluid mosaic representation of a cell membrane.†

The lipid composition of biological lipid bilayers is diverse, and may vary from the inner to the outer leaflet, in between organelles, cell types, organs, individuals, and species.[43–45] Tackling this diversity and rationalizing its effects is still an ongoing challenge for computational, biophysical and biological research. There exist three major types of membrane lipids, namely glycerophospholipids (GPL), sphingolipids, sterols and fatty acids (see Figure 5).[31,42]

† Gathered from [42], Copyright 2022, American Chemical Society.

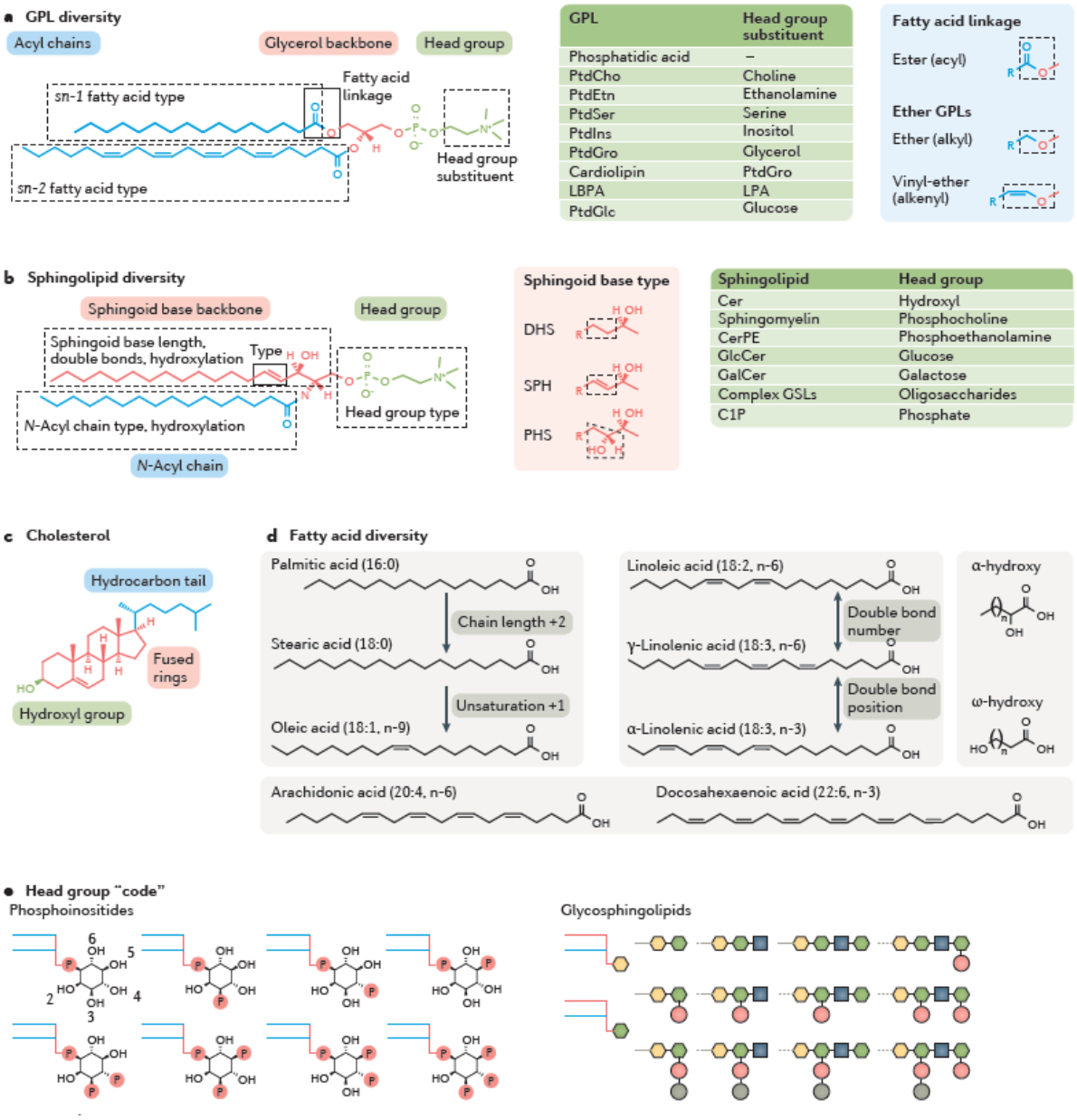


Figure 5: Chemical diversity of membrane lipids in mammals.**

I.5.1. Glycero-phospholipids

Glycero-phospholipids (GPL) are the most abundant lipids in lipid bilayer membranes. They are composed of a polar head group and two hydrophobic acyl tails. The head group and the acyl chains are linked by a glycerol moiety. The linkage between glycerol and the fatty acid chain can be either an ester, an ether or a vinyl-ether group (Figure 5a). The polar head groups may differ from the moiety linked to the phosphate group. For instance, various substituents can be adapted (Figure 5a). The hydrophobic lipid chains consist of long acyl chains that may

** Gathered from [32]

differ from their length, number, and position of the unsaturations and hydroxylations (Figure 5d). Phosphatidylcholine (PC), a typical eukaryotic phospholipid, is the most abundant component in mammal cell membranes.[46] Overall, the chemical nature of GPL head groups and lipid tails directly influences the lipid bilayer biophysical properties.

I.5.2. Sphingolipids

As for GPLs, sphingolipids consist of a head group and two hydrophobic tails. The difference between the two lipid types lies in one of the tails that consist of a sphingoid base. It may vary in length, number and position of unsaturations, and hydroxylation. The head group defines the name of the sphingolipid (Figure 5b). Sphingolipids are present in most cell types, but are found at higher concentrations, as for sphingomyelin and ceramides in neuronal tissues and the skin, respectively. Sphingolipids play a major role in modifying the lipid bilayer structural properties, but they are also involved in various cell signaling pathways.[47]

I.5.3. Sterols

The major sterol in vertebrates is cholesterol (Figure 5c).[48] In mammalian membranes, cholesterol can be found up to the same molar concentration as GPLs.[48] An increase in concentration of cholesterol favors the formation of the liquid ordered state, that is required for many biological processes to occur in membranes.[48] Other sterols are found in plants (stigmasterol, sitosterol) and fungi (ergosterol), however, they are not involved in pharmacological processes.

I.6. Structural properties of lipid bilayers

I.6.1. Membrane thickness

The membrane thickness is one the major properties of lipid bilayer membranes, it highly depends on the membrane composition. For instance, the length of acyl chains, the order, the number of unsaturations and the presence of cholesterol may strongly affect membrane thickness.[49–51] The membrane thickness can be obtained both experimentally and theoretically from MD simulations.[52–54] There exist three main definitions of membrane thickness: (i) the overall bilayer thickness (D_B, see Figure 6) calculated as the ratio of the volume per lipid V_L and the area per lipid A_L, also known as Luzzati thickness, (ii) the distance between the peaks of electron density in the head groups, or head-to-head distance (D_{HH}), and (iii) the hydrophobic thickness calculated by using the length of the acyl chains ($2D_c$, see Figure 6).

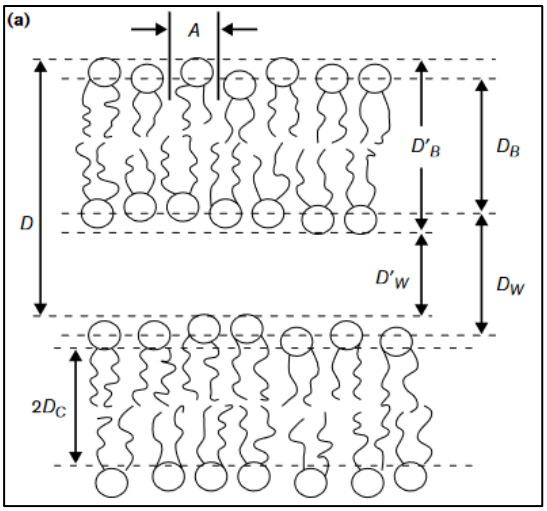


Figure 6: Schematic illustration of the different kinds of thickness in membranes.[†††]

I.6.2. Area per lipid

The area per lipid A_L is another important structural property of lipid bilayer membranes. It partially reflects lipid packing, and so lipid ordering at the surface. A_L can be affected by lipid composition but also by the presence xenobiotics in membrane, especially when lying at the membrane surface.[50,55] From MD simulations, A_L can simply be calculated by dividing the xy area of the membrane (S_{mem}), so the simulation box, by the number of lipids in one leaflet (N_l):

$$A_L = \frac{S_{mem}}{N_l} = \frac{L \times W}{N_l} \quad (1)$$

Where L and W are the X-length and the Y-width of the lipid bilayer, respectively.

I.6.3. Order parameter

The lipid bilayer order parameter, S_{CH}, corresponds to the fluctuations of the carbon−hydrogen bond in the hydrocarbon lipid tails with respect to the normal of the membrane surface.[56] These fluctuations are averaged over all lipids all along the MD simulation, using the following formula:[56]

$$S_{CH} = \frac{\langle 3\cos^2\theta - 1 \rangle}{2} \quad (2)$$

where θ is the angle between the C−H bond vector and the bilayer normal (z-axis in MD simulations). S_{CH} (see an example in Figure 7) documents on lipid ordering, and thus lipid phase. This parameter is strongly affected by lipid composition. Experimentally, the order parameter is determined on deuterated lipids, where the hydrogen atoms of the lipid tails are replaced by deuterium. Therefore, it is often referred as S_{CD}.

[†††] Gathered from [49].

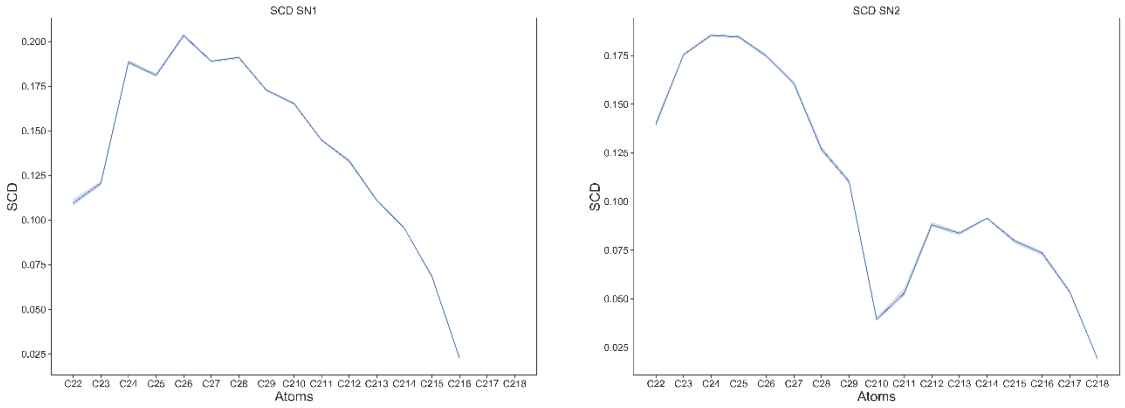


Figure 7: Order parameter of the first acyl chain (SN1, left panel) and the second acyl chain (SN2, right panel) for a 1,2-palmitoyl-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer from MD simulations with the Slipids[57] force field.

I.6.4. Lipid bilayer Phases

Depending on lipid compositions, the presence of proteins, environmental parameters including temperature, lipid bilayers can adopt different main phases[50] (see Figure 8 and Table 1), mainly (i) the liquid phase (also called liquid disordered, L_d or L_α), which describes lipid bilayers of high fluidity. Such phases are mostly composed of unsaturated lipids. They are characterized by low thickness, high A_L and low order; (ii) the liquid ordered phase (L_o), which characterizes lipid bilayers with relatively high order. It is mostly characterized by lipid mixtures containing cholesterol; and (iii) the crystalline phase, also known as the gel phase (S_O or L_β), which is characterized by increased thickness, low A_L and high order.

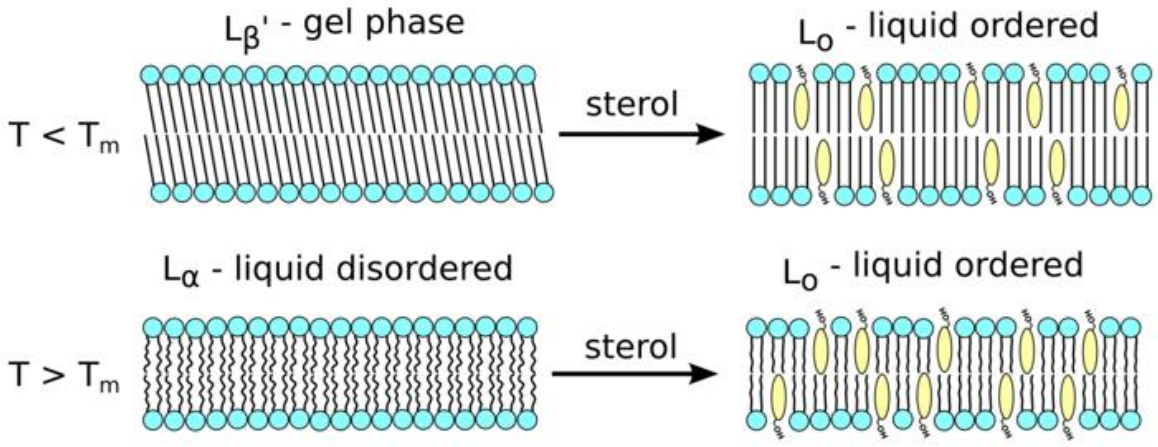


Figure 8: Comparison of the different membrane phases (liquid disordered L_d, liquid ordered L_o and gel phase Lβ).‡‡‡

Table 1 : Lipid bilayer phases and their effect on the main lipid membrane properties.

Lipid phase	**Liquid disordered (L_d)**	**Liquid ordered (L_o)**	**Gel phase (S_o)**
Thickness	Low	Medium	High
Area per lipid	High	Medium	Low
Order parameter	Low	Medium	High
Favored by increase concentrations of…	Unsaturated phospholipids	Cholesterol	Saturated phospholipids +/- Cholesterol

I.7. Membrane lipids and disease

The chemical composition of biological lipid bilayer membranes is actively regulated by the expression of lipid-metabolizing enzymes (e.g., ceramide synthases, acetyl-CoA carboxylase 1, serine palmitoyltransferase 1). Moreover, distinct lipid ratios exist between the lipid bilayer leaflets, due to flippases (i.e., enzymes that transport some lipids from one leaflet to another). These physiological variations extend to different tissues and can affect the general dynamic behavior and interactions of membrane lipids with proteins.[31,45,59] Hence, changes in lipid composition, biosynthesis, metabolism and organization can lead to pathological consequences.

One of the important factors modifying lipid composition and hence function, is a genetic one, that can occur in some diseases. A simple mutation of an enzyme related to lipid metabolism

‡‡‡ Gathered from [58].

can lead to severe outcomes.[30,60,61] For instance, it was reported that presynaptic membrane lipid disorders in the brain are associated with various neurological and psychiatric illnesses. Neurotransmission events is sensitive to lipid composition and is largely dependent on membrane protein-lipid collaborations.[62]

Another example is lipid peroxidation, a radical chain reaction that may occur in case of oxidative stress.[63] This stress arises from an imbalance between oxidants (e.g., reactive oxygen species) and antioxidants (e.g., vitamins E, C, polyphenols). The products of this chain reactions are malondialdehyde and conjugated diene, which are mutagenic and cytotoxic.[64] For example, lipid peroxidation was shown to have serious effects in various diseases such as Alzheimer and liver diseases.[34,65,66]

Chapter II. Membrane crossing

Membranes have a major role in maintaining the homeostasis in living organisms by regulating exchanges of ions and molecules (e.g., endogenous substrates, nutrients, toxic agents, drugs and metabolites) between the outer and the inner compartments of cells and organelles. This balance is maintained thanks to passive diffusion and sophisticated transport systems across membranes .[67,68]

II.1. Transport types through cell membranes

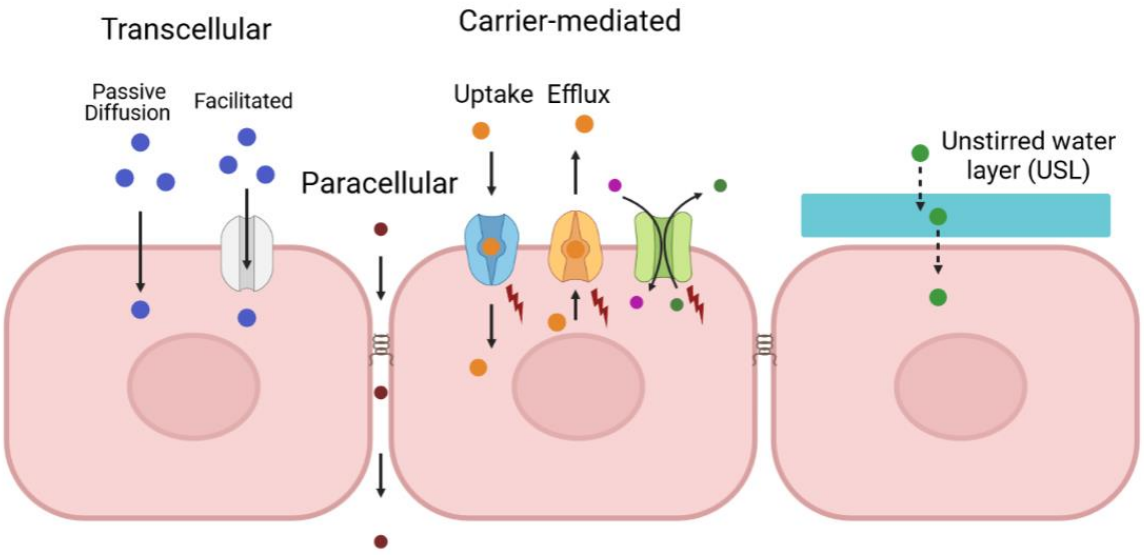


Figure 9: Schematic representation of the different types of transport across biological membranes (e.g., the intestinal epithelium)

II.1.1. Passive diffusion across the membrane

Many small and lipophilic molecules can cross lipid bilayers by passive diffusion. This process is driven by the gradient of concentrations between the aqueous solutions of both sides of the membrane (see Figure 9).[3] Conversely, the passive diffusion of large hydrophilic molecules is usually very limited, and they rather tend to cross the membrane through membrane transporters. However, these rules of thumb cannot be generalized, as many exceptions exist. The main particularity of passive diffusion is that it is not a saturable process.

The permeability describes the rate at which a given compound crosses a given membrane. The total permeability consists of a combination of different competitive transport pathways. Passive permeability can be defined as the quantity of a particle diffusing across a membrane per cross-section area in a unit time driven. This is driven by a concentration gradient (see sectionIII.1 for more details).[69]. The rate of passive permeability also depends on pressure, temperature, and concentrations in both membrane sides.

Passive permeability often depends on several intrinsic properties of the molecules including its size, chemical structure, or solubility. It also depends on the membrane composition which

differs from one tissue to another.[70–72] It can be partially understood by the Fick's first law of diffusion (see section III.1). Evaluating passive membrane permeation is required during new drug development and optimization processes to enhance their PK, in particular their absorption and delivery to reach their targets. Passive permeation is measured by the permeation coefficient P.

II.1.2. Carrier-mediated transport

Membrane proteins ensure diverse functions including signal transduction, trafficking and recognition.[73,74] They cover approximately 20% of protein sequences in the human genome.[73,75] They are confined in the membrane bilayer in different ways, namely, they can entirely span the lipid bilayer (i.e., Integral or transmembrane helical proteins) or partially (i.e., peripheral). Integral membrane proteins are assembled in lipid membranes by means of translocation and insertion through translocons.[76]

A wide range of xenobiotics and endogenous chemicals including those with low passive membrane permeability (e.g., highly hydrophilic compounds) can be transported through such proteins (carrier-mediated transport) such as ion channels (K^+, Na^+, Cl^- and Ca^{2+} channels) or importers or exporters (see Figure 9).[73,77,78] Membrane carriers regulate the entry and exit of molecules into or from cells. They have an important role in pharmacology as they are directly related to PK (ADME) of xenobiotics. They can be classified as channels or active transporters: (i) ATP-binding cassettes and (ii) Solute Carriers. Unlike passive permeation across the membrane, carrier-mediated transport is saturable and concentration dependent.[69]

Besides, membrane lipids can largely affect the structure of these membrane proteins as well as their dynamics, thus their activities. Moreover, the distribution of the various membrane proteins varies in between cell types and individuals.[62,79]

II.1.2.1. Ion channels

Channels are membrane proteins which highly facilitate diffusion of ions across the membrane. They provide a hydrophilic path that can drive an electric current carried by ion conduction. The major properties of ion channels are gating and ion selectivity. Channel-mediated transport does not directly require energy from ATP hydrolysis during the transport cycle, wherein the direction of ion crossing is along the concentration gradient. An example of ion channel is the voltage-gated calcium channels (VGCCs). VGCCs are ion channels essentially permeable to Ca^{2+} ions. They are mainly present in the membrane of excitable tissues (e.g., neurons, muscles. etc.).[80,81]

II.1.2.2. ATP-binding cassettes (ABC proteins)

ABC proteins use ATP hydrolysis as a source of energy to drive substrate transport across biological membranes. A prototypical ABC transporter consists of two different domains: (i) the two transmembrane domains (TMDs), containing each 6 α-helices and mostly embedded inside the hydrophobic core of the membrane; and (ii) the two nucleotide-binding domains (NBDs) that contain consensus/conserved domains (Walkers A and B and signature domain C). In the human organism, ABC proteins are efflux transporters (driving molecules from the inner cellular environment to the outer).

In eukaryotes, ABC genes present high sequence identity of amino acids. In mammalians, 7 distinct ABC transporter subfamilies were identified, namely: ABCA, ABCB, ABCC, ABCD,

ABCE, ABCF and ABCG. Of main interest are three members of ABC transporters: (i) ABCB1 (P-glycoprotein or P-gp) member of multidrug-resistance proteins (MDRs); (ii) ABCG2 or breast cancer-resistance protein (BCRP); and (iii) ABCC subfamily containing 9 multidrug-resistance-associated proteins or (MRP 1 to 9). [82]

II.1.2.3. Solute carriers

The solute carriers (SLCs) are mainly influx transporters, hence strongly contributing to the uptake of small molecules into cells. This uptake process is not directly driven by ATP hydrolysis.[83,84] Transporters of the SLC superfamily include multidrug and toxin extrusion proteins (MATEs), organic anion transporters (OATs), organic cation transporters (OCTs), organic anion-transporting polypeptides (OATPs), organic cation/carnitine transporters (OCTNs).

II.1.2.4. Distribution of membrane transporters in tissues

ABC transporters play a substantial role in maintaining homeostasis as they are expressed in most of epithelial cell membranes. Their expression and localization in a given tissue barrier vary between ABC and SLC members. They regulate the transport of xenobiotics (drugs and toxins) and endogenous substrates between fluid compartments. Various members of ABC and SLC transporter families have a substantial role in PK, namely ABCG2, P-gp, MRP4, OAT1 (SLC22A6), OAT3 (SLC22A8), OCT2 (SLC22A2), OATP1B1 (SLCO1B1) and OATP1B3 (SLCO1B3). (see Figure 10)

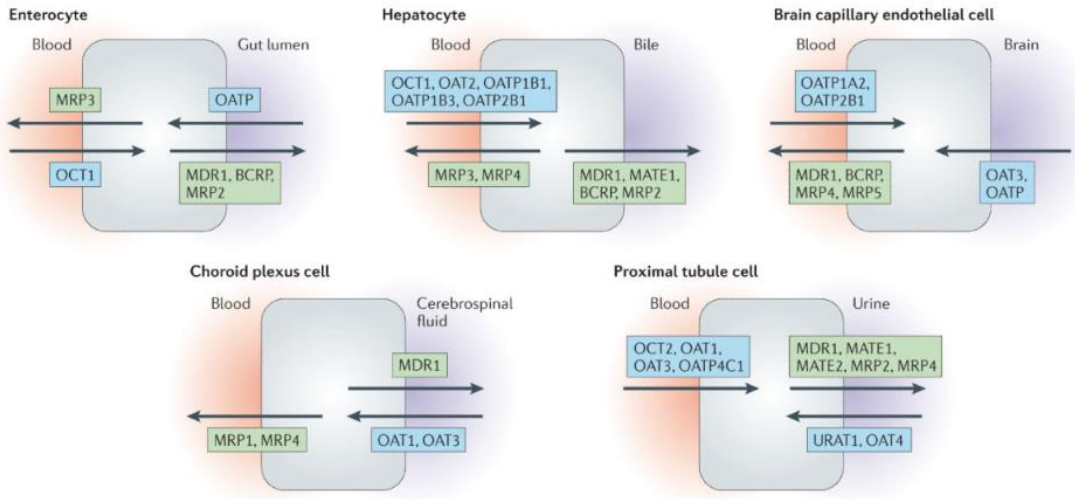


Figure 10:Membrane transporters implicated in the intake (SLCs: blue) and efflux (ABC proteins: green) of xenobiotics. †††††

††††† Gathered from [85].

Understanding the role and function of membrane transporters, in both healthy and pathological conditions, is important to develop new lead drugs or adjust drug doses. Membrane transporters are tightly related to various diseases such as cancer.

Genetic polymorphisms of membrane transporters can affect their expression level, their localization and their conformations, thus possibly altering their biological function. In addition, drug-drug interactions (DDIs), usually associated to concomitant uptake of several xenobiotics, can result from and in a dysfunction of specific membrane transporters.

II.1.3. Unstirred layers (USL)

Stagnant water layers are adjacent to the membrane. Their thickness is estimated in between 170 and 800 µm[86–89]. These unstirred layers (USL) can slow down the diffusion of molecules causing additional resistance to their permeation (see Figure 9). USL complicate the prediction of passive permeability based solely on the permeability through the lipid bilayer. The permeability through USL, P_{USL}, is related to the diffusion coefficient, D, and the thickness of the USL, δ:[90]

II.1.4. Paracellular transport

Paracellular transport consists in transport between cells, as opposed to trans-cellular transport (Figure 9). The absorption of some polar molecules and ions (< 200 Da) relies on the paracellular pathway (e.g., exogenous sugars, Na+, Cl-), whereas most drugs permeate through the transcellular pathway.[69,91–93] It is an important transport pathway in capillaries, except at the BBB. The concentration gradient drives para-cellular transport. Tight junctions – membrane protein complexes between adjacent cells – limit the para-cellular transport. In the gastrointestinal epithelium, tight junction dysfunctions are related to inflammation and metabolic disorders.[94]

II.1.5. Endocytosis

Some biologically active large molecules (e.g., cobalamin, immunoglobulins) require endocytosis to be absorbed.[95,96] This process consists in the passage of the macromolecule into the cellular environment without direct contact with the lipid tails of the lipid bilayer. The major mechanisms of endocytosis are pinocytosis, phagocytosis and receptor-mediated endocytosis. Many toxins and infectious agents (viruses) utilize this mechanism to enter into the body cells. Moreover, several targeted therapies have adopted this mechanism to increase the bioavailability of oral drugs.

II.2. Membrane transporters and organ transplantation

Organ failure or injury such as acute liver failure and kidney dysfunction are multifactorial life-threatening diseases. Drug-induced liver injury (e.g., acetaminophen-induced hepatotoxicity) or viral infections (i.e., hepatitis A, B, and E) are the main causes of liver failure. Severe subsequent complications may result in encephalopathy and other associated dysfunctions. Liver failure may be associated, in 50% of cases, with substantial renal dysfunction that may increase mortality. In particular cases, with severe complications (e.g., intracranial hypertension), transplantation can be lifesaving.[97] However, once transplanted, a patient may experience graft rejection in short, medium, or long term. The causes for graft rejection are

multifactorial, and its prediction is still challenging. The histocompatibility system, tissue typing and DDI play a substantial role in graft survival, that can result in graft rejection.[98–100]

In the context of liver and kidney transplantation, immunosuppressant drugs are broadly used to decrease the immune response toward grafts. Tacrolimus, cyclosporine A and mycophenolate mofetil (MPA) are the most commonly prescribed drugs.[101–103] Usually, immunosuppressant drugs are taken concomitantly with other xenobiotics (i.e., drugs such as antivirals, and nutrients). The accumulation of these drugs inside the graft cells may cause toxicity, leading to graft rejection. This accumulation is normally prevented by transmembrane exporters (e.g., MRP2, MRP4 and P-gp) [85,104–106]. Any dysfunction in these transporters may thus lead to dramatic drug accumulation and severe toxicity. This is particularly true with DDI that may lead to accumulation of drugs inside hepatocytes and renal cells. This accumulation is mainly due to the lack of their transport outside cells, through efflux transporters, to reach blood and urine. Drug accumulation inside liver and kidney cells results in cytotoxicity. Through a series of complex processes, involving several immune mediators such as T-cells and cytokines, this leads to graft rejection. [98,107–109]

In the context of liver and kidney transplantation, the MRP subfamily is of utmost importance. MRP4 is exclusively an efflux transporter. It is highly expressed in SNC, liver, kidney and blood cells. As a multidrug resistance protein, it mediates the elimination of various conjugated drugs though the proximal renal tubule. Also, it insures their efflux into blood compartment through the basal membrane of hepatocytes. Thus, any dysfunction of hMRP4, mainly caused by DDIs, is directly related to drug accumulation, then cytotoxicity and graft failure.[85,105,110] In Chapter VII, we will studied the potential binding sites of the available hMRP4's substrates and modulators using an advanced molecular docking approach.

Chapter III. Passive permeation of drug molecules

Plasma membranes are the center for all exchanges between the extra- and the intracellular compartments. They constitute an interface where a wide diversity of interactions between xenobiotics, biomolecules and membrane components can occur. As mentioned in Chapter II, endogenous molecules and xenobiotics can be transported in and out of the cell through membrane proteins. However, any molecule can theoretically cross the membrane by passive permeation, albeit on a very wide range of timescales. Passive permeation is non-saturable and depends on both the membrane and the permeant properties. It often plays an important role in the membrane crossing of amphiphilic and hydrophobic molecules. For a given molecule, the ability to cross biological membranes by passive permeation is evaluated by the permeation coefficient P. It is a key property during the drug discovery process, as it helps selecting potential lead molecules, and better understanding their PK.

III.1. Fick's first law

During early passive diffusion studies, membranes had been considered as a homogeneous one-slab layer through which the particle diffuses from one compartment to another following the gradient of concentration (Figure 11).[20,111]

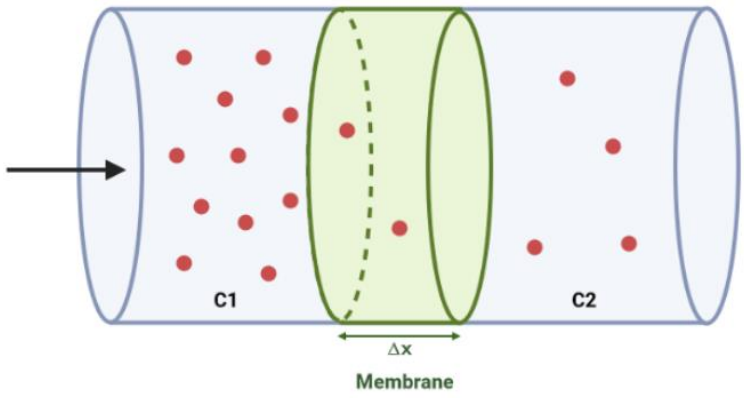


Figure 11: The one-slab membrane model (green slab) with thickness Δx separating two compartments 1 and 2 with their corresponding concentrations c1 and c2 of the particle (red circles), respectively.

The transmembrane flux J is the quantity of particles, Q, diffusing across the one slab membrane, from compartment 1 to compartment 2, per area A, during time t:

$$J = \frac{1}{A} \cdot \frac{\partial Q}{\partial t} \quad (1)$$

J is given by the Fick's first law, which is valid under the assumption of a steady state, i.e. when the concentration gradient remains constant. The one-dimensional flux J can be written

as a function the diffusivity of a particle diffusing in the membrane, D_m, and the change in concentration c_m as a function of the particle's position x along the membrane ($\frac{\partial c_m}{\partial x}$):

$$J = -D_m \frac{\partial c_m}{\partial x} \quad (2)$$

The diffusion across the homogeneous membrane layer requires partition of the particle between the aqueous medium and the membrane (Figure 11). At equilibrium, the ratio between the concentration at the lipid bilayer leaflet and the aqueous phase is the partition coefficient K given by:

$$K = \frac{c_m}{c_{aq}} \quad (3)$$

Moreover, at the steady state, the flux J is constant, and it is proportional to the concentration difference Δc of the particle between the two aqueous compartments.

$$J = P \times \Delta c \quad (4)$$

Where the proportionality factor, P, is the permeation coefficient, often expressed in $cm.s^{-1}$, and it represents the rate at which the particle can cross the membrane.

III.2. Overton's rule

The Overton's rule was proposed by Ernest Overton in the beginning of 19th century through a series of publications. The first starting point of the fundamental knowledge of cell membranes was discovered by investigating the osmotic properties (free water diffusion through cell membranes) of both plant and animal cells. Overton demonstrated that cell membranes were permeable to various undissociable organic solutes (e.g., chloroform, ether, and acetone), at different rates over time. Therefore, he could estimate the permeability of a given solute. These solutes were chosen because they did not show plasmolytic effect, contrary to impermeable solutes such as amino acids or sugars.[112–114]

This study was extended to establish a relationship between the permeability and the chemical structure of molecules. A range of solvents was used to experimentally model the partition between water and the lipid environment. The findings upon these series of experiments revealed that solutes having solubility in lipid environment (e.g., ether, lipid oils) greater than in water are likely to passively diffuse through the membrane. In addition, Overton raised evidences about selective permeability of neutral permeants with respect to their charged forms.[112,115] This highlights the essential role of the lipid/water partition coefficient on the permeability (P).[42] These findings are known today as the "Overton's Rule".[20,112,114]. It was shown that the choice of the optimal oil solvent depends on the solute. For instance, octanol is adapted when studying highly hydrated amino acids and peptides, as it forms inverse micelles trapping water molecules at the interior.[116,117] Conversely, hexadecane is more suitable for small nonelectrolytes molecules.[118]

III.3. Lipinski's rule of five

Given the wide diversity of drug-like molecules, the relationship between permeability and physicochemical properties of drug-like molecules has been studied to select potential and orally active drug candidates with increased bioavailability. Molecular properties such as

molecular weight, lipophilicity, number of H-bond donors and acceptors were shown to be important factors to predict a good bioavailability of drugs, hence, a good absorption or high permeability. A balance between nonpolar and polar properties is required for oral drugs.[119,120] In 1997 Christopher Lipinski postulated a rule of five, often coined RO5. It relies on the evaluation of four parameters to estimate a favorable drug likeliness, i.e. the probability that a drug absorbed orally is bioavailable. The four parameters are : (i) less than 10 H-bond acceptors (nitrogen and oxygen atoms), (ii) 5 H-bond donors (N-H and O-H bonds), (iii) a molecular weight lower than 500 $g.mol^{-1}$, and (iv) a $logK_{o/w}$ greater than 5.[121]

Lipinski's rule of five is widely used in high throughput screening (HTS) techniques during the drug discovery process. HTS are used to yield potential hits that exhibit good bioavailability and high therapeutic efficiency.

III.4. Biopharmaceutics classification system

Biopharmaceutics classification system (BCS) is based on the Fick's first law. It correlates *in vitro* drug solubility and intestinal permeability with its *in vivo* bioavailability.

Solubility can be defined as the ability of a molecule to dissolve in a fluid environment. To cross a membrane barrier, the molecule should exhibit the correct balance in solubility between water and the lipid medium. In aqueous medium, solubility can be defined as the sum of the concentrations of all dissolved species in the medium. In the case of weak acids HA, solubility is:[122]

$$S = [HA] + [A^-] = [HA] \times (10^{pH-pKa} + 1) \quad (5)$$

Moreover, lipid solubility is related to lipophilicity of the molecule of interest, which can be estimated (i) by measuring the partition coefficient between water and lipid (oil) or octanol ($K_{o/w}$), or (ii) theoretically by computing the free energy of partitioning where

$$K = \frac{c_1}{c_2} = e^{-\frac{\Delta G_{part}}{RT}} \quad (6)$$

Where c_1 and c_2 are the concentrations in membrane and aqueous phase, respectively, and ΔG_{part} is the free energy of partition. Lipid solubility is affected by lipophilicity, H-bonding, electrostatic bonding, ionization state of the particle and temperature.[123]

The permeability can be determined experimentally either by (i) PK studies on human subjects, (ii) intestinal perfusion on animals, (iii) *in vitro* models using intestinal epithelium, or (iv) monolayers of epithelial cells (e.g. Caco-2 cells, see section III.7).[124,125] The Biopharmaceutical classification system defines 4 classes (see Figure 12):

- Class 1: High solubility - High permeability drugs: drugs that satisfies class 1 are readily absorbed. The bioavailability of drugs of this class is usually low due the rapid metabolism.
- Class 2: Low solubility - High permeability drugs: drugs have lower absorption rate than Class 1. Their dissolution is the rate-limiting step.
- Class 3: High solubility - Low permeability drugs: permeability is the limiting step. The rate of absorption may be highly variable due the variable gastrointestinal transit.

- Class 4: Low solubility - Low permeability drugs: both drug dissolution and permeability are low, therefore bioavailability is poor.

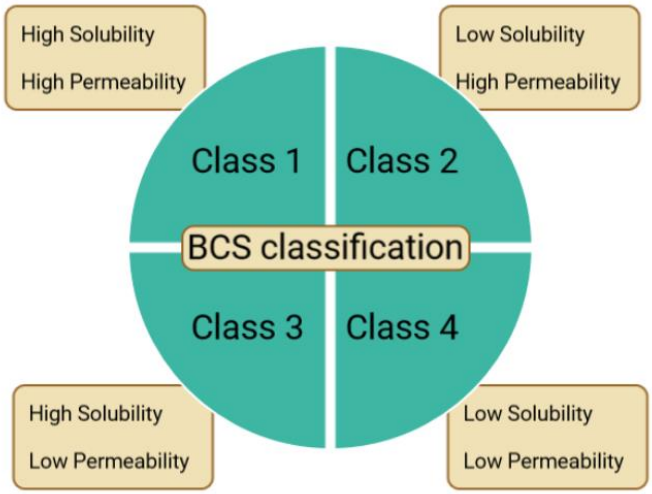


Figure 12: Schematic representation of the Biopharmaceutics Classification System (BCS)

III.5. Solubility-Diffusion Models

III.5.1. Homogeneous solubility diffusion model

Based on the Overton's rule, the homogeneous solubility diffusion model was often used to describe the permeation of small molecules through polymer membranes.[126] Here, the membrane is treated as a one-slab homogeneous layer (Figure 13). The diffusivity and the driving forces are assumed to be uniform. The Fick's second law describes the evolution of the concentration inside the one-slab membrane over time as follows:

$$\frac{\partial c_m}{\partial t} = D_m . \frac{\partial^2 c}{\partial x^2} \quad (7)$$

Where c is the concentration in the membrane, x the position, and D_m is the diffusivity of the molecule inside the membrane. At equilibrium, $\frac{\partial c_m}{\partial t} = 0$ and the flux J is constant.

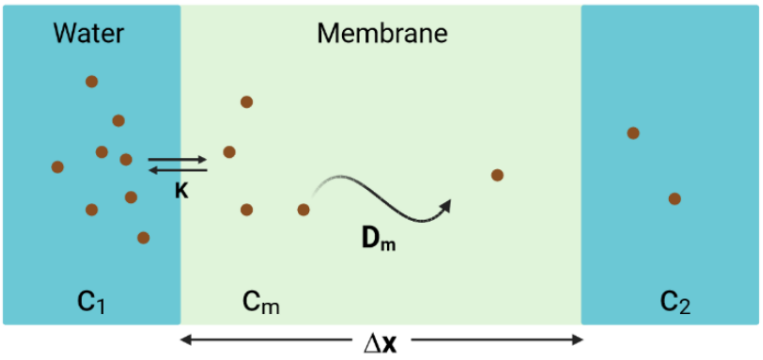


Figure 13: The Homogeneous Solubility Diffusion Model (HSDM). The one-slab membrane medium is separated by two aqueous phases with two different concentrations c1 and c2 of the diffusing particle (brown filled circles) in both sides. K is the partition coefficient between membrane and water and D_m is the diffusion coefficient of the particle inside the membrane.

At steady state, the concentration profile inside the membrane, is assumed to be linear and c_1 and c_2 are maintained constant. The flux J is constant and it is proportional to the difference in concentrations between the two aqueous compartments (Δc).

$$J = \frac{K.D_m}{\Delta x}.\Delta c_m \quad (8)$$

Where K is the partition coefficient between the aqueous phase and the membrane interior and Δx is the membrane thickness.

Combining equation (4) and equation (19) yields the expression of the permeation coefficient

$$P = \frac{K.D_m}{\Delta x} \quad (9)$$

However, there are structural differences between polymer membranes and biological lipid bilayer membranes, which makes the homogeneous model fails at describing the latter membranes. The phospholipids that compose lipid bilayers consist of a polar head groups and non-polar tails. Therefore, the polarity of lipid bilayers is very inhomogeneous and the diffusivity inside lipid bilayers is neither homogeneous.[111]

III.5.2. Inhomogeneous solubility diffusion model

The main concept of the inhomogeneous solubility diffusion model (ISDM) is that the diffusivity, the partitioning (related to the free energy landscape across the membrane) and thus the resistance to cross membrane are not uniform, in other words, they are depth-dependent. [18,127–129] This model describes the membrane as an inhomogeneous fluid medium where molecule dynamics is driven by a Brownian motion under the influence of a driving force F. Consequently, the diffusing molecule undergoes random walk along the membrane thickness, i.e., along the z-axis, but also the x-y plane.

The one-dimensional translocation rate can be straightforwardly discretized to the z-axis alone, as it represents an inhomogeneity of chemical composition, and thus a variability in diffusivity and free energy landscape as a function of the position of the molecule. [18]

The general formalism can be described by the one-dimensional Fokker-Planck equation derived into the Smoluchowski equation. It describes the motion of a diffusive particle inside the membrane under the influence of a force F(z,t):[18,130]

$$\frac{\partial \mathrm{c(z,t)}}{\partial \mathrm{t}} = \frac{\partial}{\partial \mathrm{z}} \mathrm{D(z)} \left[\frac{\partial}{\partial \mathrm{z}} - \beta \mathrm{F(z,t)} \right] \mathrm{c(z,t)} \tag{10}$$

Where $\frac{\partial \mathrm{c(z,t)}}{\partial \mathrm{t}}$ is the rate of change of the concentration at position z versus time, $\mathrm{c(z,t)}$ is the concentration of the molecule at a given depth z at time t, $\mathrm{D(z)}$ is the position-dependent diffusivity, $\mathrm{F(z)}$ is the driving force, equals to the derivative of the free energy $\mathrm{w(z)}$ (also called potential of the mean force, PMF) at position z : $\mathrm{F(z)} = -\frac{\partial \mathrm{w(z)}}{\partial \mathrm{z}}$. Here $\beta = \frac{1}{\mathrm{k}_b \mathrm{T}}$ where k_b is the Boltzmann's constant and temperature T.

The Smoluchowski equation (20) can thus be derived as a function of both the diffusivity and the free energy profile as follows:

$$\frac{\partial}{\partial \mathrm{t}} \mathrm{c(z,t)} = \frac{\partial}{\partial \mathrm{z}} \mathrm{D(z)} \left[\frac{\partial}{\partial \mathrm{z}} \mathrm{c(z,t)} + \mathrm{c(z,t)} \beta \frac{\partial}{\partial \mathrm{z}} \mathrm{w(z)} \right] \tag{11}$$

Multiplying the right side by $\frac{\mathrm{e}^{\beta \mathrm{w(z)}}}{\mathrm{e}^{\beta \mathrm{w(z)}}}$ yields:

$$\frac{\partial}{\partial \mathrm{t}} \mathrm{c(z,t)} = \frac{\partial}{\partial \mathrm{z}} \mathrm{D(z)} \frac{1}{\mathrm{e}^{\beta \mathrm{w(z)}}} \left[\mathrm{e}^{\beta \mathrm{w(z)}} \frac{\partial}{\partial \mathrm{z}} \mathrm{c(z,t)} + \mathrm{c(z,t)} \mathrm{e}^{\beta \mathrm{w(z)}} \frac{\partial}{\partial \mathrm{z}} \beta \mathrm{w(z)} \right] \tag{12}$$

Manipulation inside the square brackets of equation (12), setting $u = \mathrm{c(z,t)}$ and $\mathrm{v} = \mathrm{e}^{\beta \mathrm{w(z)}}$, and reminding that $u'v + uv' = (uv)'$, yields:

$$\frac{\partial}{\partial \mathrm{t}} \mathrm{c(z,t)} = \frac{\partial}{\partial \mathrm{z}} \left[\mathrm{D(z)} \mathrm{e}^{-\beta \mathrm{w(z)}} \frac{\partial}{\partial \mathrm{z}} \mathrm{e}^{\beta \mathrm{w(z)}} \mathrm{c(z,t)} \right] \tag{13}$$

The primitive of the left term of this equation is a flux, $\mathrm{J(z)}$, i.e., a particle quantity going through a surface per unit time, thus this equation can be re-written as:[18]

$$\mathrm{J(z)} = -\mathrm{D(z)} \mathrm{e}^{-\beta \mathrm{w(z)}} \frac{\partial}{\partial \mathrm{z}} \left[\mathrm{e}^{\beta \mathrm{w(z)}} \mathrm{c(z,t)} \right] \tag{14}$$

The integration of equation (24) over the membrane thickness yields:

$$\int_{-\frac{\mathrm{L}}{2}}^{+\frac{\mathrm{L}}{2}} \frac{\mathrm{J(z)}}{\mathrm{D(z)} \mathrm{e}^{-\beta \mathrm{w(z)}}} \mathrm{dz} = -\int_{-\frac{\mathrm{L}}{2}}^{+\frac{\mathrm{L}}{2}} \frac{\partial}{\partial \mathrm{z}} \mathrm{e}^{\beta \mathrm{w(z)}} \mathrm{c(z,t)} \, \mathrm{dz} \tag{15}$$

At steady state, the flux is constant and independent of z. This yields:

$$\mathrm{J} \int_{-\frac{\mathrm{L}}{2}}^{+\frac{\mathrm{L}}{2}} \frac{\mathrm{e}^{\beta \mathrm{w(z)}}}{\mathrm{D(z)}} \mathrm{dz} = -\left[\mathrm{e}^{\beta \mathrm{w(z)}} c(z,t) \right]_{-\frac{L}{2}}^{\frac{L}{2}} \tag{16}$$

Which leads to:

$$\mathrm{J}\int_{-\frac{\mathrm{L}}{2}}^{+\frac{\mathrm{L}}{2}}\frac{\mathrm{e}^{\beta \mathrm{w}(\mathrm{z})}}{\mathrm{D}(\mathrm{z})}\mathrm{dz} = \mathrm{e}^{\beta \mathrm{w}\left(-\frac{L}{2}\right)}c\left(-\frac{L}{2}\right) - \mathrm{e}^{\beta \mathrm{w}\left(\frac{L}{2}\right)}c\left(\frac{L}{2}\right) \quad (17)$$

At both membrane boundaries (i.e., water medium), the potential of the mean force w is equal to 0 ($\mathrm{w}\left(\frac{\mathrm{L}}{2}\right) = \mathrm{w}\left(-\frac{\mathrm{L}}{2}\right) = 0$), hence, the exponential terms equal 1. Integrating between the two sides of the membrane yields:

$$\mathrm{J}\int_{-\frac{\mathrm{L}}{2}}^{+\frac{\mathrm{L}}{2}}\frac{\mathrm{e}^{\beta \mathrm{w}(\mathrm{z})}}{\mathrm{D}(\mathrm{z})}\mathrm{dz} = \left[\mathrm{c}\left(-\frac{\mathrm{L}}{2}\right) - \mathrm{c}\left(\frac{\mathrm{L}}{2}\right)\right] = \Delta \mathrm{c} \quad (18)$$

The right side of equation (28) defines the difference in concentration between the two sides of the bilayer in the positive direction of the flux ($\Delta \mathrm{c} = c\left(-\frac{L}{2}\right) - c\left(\frac{L}{2}\right)$). Combining equation (4) and equation (28) (see section III.1):

$$P_{erm} = \frac{\mathrm{J}}{\Delta \mathrm{c}} \quad (19)$$

Hence, the resistance to permeation R, i.e., the inverse of P_{erm}, is given by:

$$R = \frac{1}{P_{erm}} = \int_{-\frac{L}{2}}^{+\frac{L}{2}}\frac{e^{\beta w(z)}}{D(z)}\,dz \quad (20)$$

Equation (20) is the central equation in the ISDM formalism, and it allows computing permeation coefficients from both the underlying Gibbs free energy and diffusivity profiles.

Four major steps were suggested to describe a molecule crossing a lipid bilayer:[131,132]

(i) For ionizable molecules, a protonation step at the water-polar head group interface is usually favorable. This step is characterized by free energy of deprotonation $\Delta G_{deprotonation}$ driving the chemical reaction (equation (23)).

(ii) Partitioning from the membrane surface in contact with the bulk water into the lipid bilayer core. This entry step is defined by the partition coefficient (K_{part}) that is based on the free energy of partitioning (ΔG_{part}), which is the difference between the free energy minima in the membrane and in water.

(iii) Flip-flop, from one to another leaflet, crossing the hydrophobic membrane center. This is characterized by a free energy difference ($\Delta G_{flip\text{-}flop}$), which is the difference between the minimum and the maximum of free energy inside the membrane.

(iv) Membrane exit to the other aqueous medium, where the molecule can be re-solvated in water. The exit of the molecule through the polar head groups requires a free energy (ΔG_{exit}), which is equal to $-\Delta G_{part}$ in the case of symmetric lipid bilayers.

Depending on the physicochemical characteristics of the molecule, any of these steps can be the rate-limiting step to membrane permeation.

III.6. Ionization state and passive permeation

It has long been suggested that only the undissociated and uncharged form of a molecule could cross passively the lipid bilayer membrane, which is known as the pH-partition hypothesis.[133]

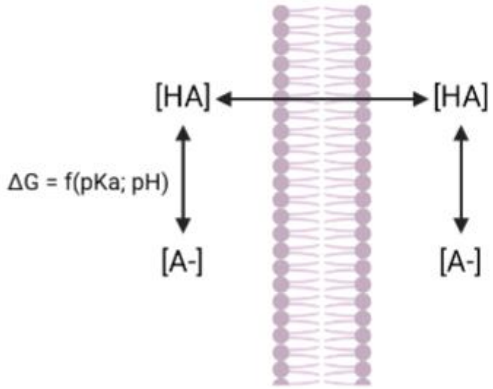


Figure 14: Diagram showing the ionization and the transfer process of weak acids across the membrane

However, the molecule-membrane interaction is highly complex to capture within a single rule and it depends on various physicochemical properties related to both the molecule and the membrane. Several physicochemical properties of the molecule were shown to affect its interactions with the membrane and thus its passive diffusion. Among others, the following parameters are key: the molecular weight (MW), lipophilicity (often expressed as the octanol/water partition coefficient, written as $logP_{o/w}$ or $logK_{o/w}$), the number of H-bond donor (HBD) and acceptor (HBA) chemical groups, the number of rotatable bonds, the presence of aromatic rings, the ionization state (acid/base properties; dissociation constant pK_a) and the polar surface area (PSA). [119,120,134–136] In some specific PK issues like absorption of xenobiotics through the gastrointestinal tract, additional parameters are at stake, such as the accessible intestinal surface area and the gastrointestinal transit.[137] Most of these xenobiotics are amphiphilic molecules (i.e., weak acids or bases). Their absorption depends on their solubility in the lipid medium (related to their lipophilicity or partition coefficient, $logK_{o/w}$), but also on their dissociation constant (pK_a) and the local pH.[123]

The charged state of an ionizable molecule is one of the limiting factors for its transport through biological membranes during the ADME steps. As an example of weak acid (HA), the ionization state at a given pH can be determined from knowing the pK_a of the molecule, as derived from the Henderson–Hasselbalch equation:

$$pK_a = pH - log\frac{[A^-]}{[HA]} \quad (21)$$

On the other hand, the free energy of deprotonation HA ⇌ A^- is:

$$\Delta G_{deprotonation} = -\beta^{-1} ln \frac{[A^-]}{[HA]} = -2.3\, \beta^{-1} log \frac{[A^-]}{[HA]} \quad (22)$$

Merging equation (21) and equation (22) give rise to the equation of the deprotonation free energy in water as a function of pK_a and pH (see Figure 14):

$$\Delta G_{deprotonation} = 2.3\, \beta^{-1}\, (pK_a - pH) \quad (23)$$

III.7. Experimental studies on passive transport across lipid bilayer membranes

The *in vivo,* the assessment of permeability of drugs can be an expensive and challenging task. The permeability depends on several factors including intrinsic physicochemical properties of the molecule but also environmental properties among which the membrane composition of the studied tissues is of key importance (e.g., presence of specific membrane transporters in the intestinal epithelium cell or the blood−brain barrier endothelial cells).[138–141]. To predict the bioavailability of orally administered drugs, the BCS has been adopted to select and classify potential biopharmaceutical drugs. According to BCS, solubility and permeability are crucial parameters that have to be evaluated to monitor the drug PK and if needed to adjust doses.[125,142] Additionally, octanol/water partition coefficients are widely used to describe passive permeation. This parameter remains controversial due to the observed differences between octanol/water and membrane/water partition coefficients as the hydrophobic media in both systems are highly different at the atomistic level.[143–145] To assess permeability with more complex and more realistic models, and to avoid unnecessary experiments on animal or human subjects, various *in vitro* experimental setups based on natural and artificial membranes have been developed to evaluate passive membrane permeation of xenobiotics (Figure 9).[23,146] In this section, we will first outline the most common *in vitro* experimental setups, and then the most recent detection systems to quantify permeation coefficients.

III.7.1. *In vitro* membrane models

Several biomimetic membrane models coupled with advanced techniques were developed to quantitatively assess the ability of drug-like molecules to passively traverse biological membranes (Figure 15). The advantage of using various membrane models is to allow rationalizing the influence of membrane properties on the permeation of molecules.

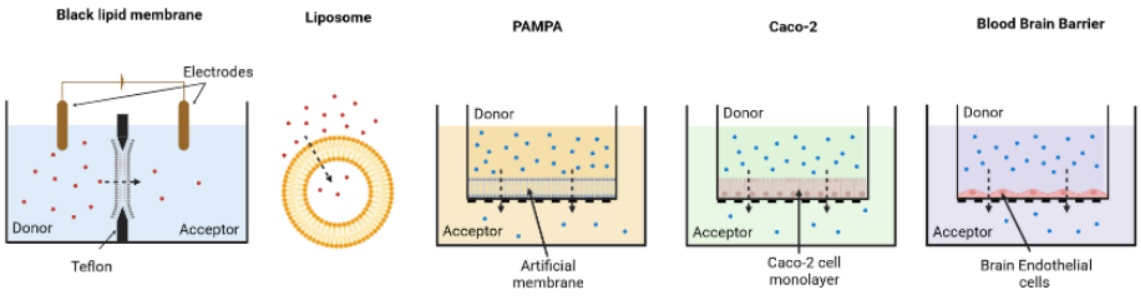


Figure 15: Schematic representation of different experimental membrane models namely Black lipid membranes (BLM), Liposomes, PAMPA, Caco-2 and BBB.

III.7.1.1. Cell-based permeability assays

Cell-based experimental assays are used to evaluate the permeability and the interactions with membrane transporters of drugs.[147]

III.7.1.1.1. Caco-2/MDCK cell-based assay

Cultured colorectal adenocarcinoma-derived cells (Caco-2) (see Figure 15) and Madin-Darby canine kidney epithelial cell lines (MDCK) have been used to evaluate drug absorption through the gastrointestinal tract.[11,12,148–152] These cells model the epithelial barrier, which xenobiotics may cross from the luminal medium into the blood (systemic circulation). They often contain membrane transporters such as H+/peptide co-transporter (PEPT1), Glucose transporter 2 (GLUT2), organic cation transporter (OCTs), Na^+/glucose transporter 1 (SGLT1), organic anion-transporting polypeptide (OATPs), breast cancer resistance protein transporter (BCRP) and multidrug resistance protein 1 (MDR1, known also as P-gp).[153,154] The presence of active transport in this system may result in a biased estimation of oral absorption due to the over- or under-expression of active carriers in the cultured cells.[155]

These assays can thus describe transcellular, paracellular and carrier mediated transport of xenobiotics. However, they require cell culture, which can be expensive and time consuming. Therefore, cell-based systems are not suited for high throughput investigations. In addition, USL can affect the measured (effective) permeability coefficients. Their contribution can be limited – but not entirely eliminated – by vigorous stirring process in order to thin out these stagnant aqueous layers.[51, 70, 79]

III.7.1.1.2. Blood Brain Barrier cell-based assay

The blood brain barrier (BBB) consists of highly specialized endothelial cells (astrocytes, pericytes and vascular endothelial cells) that regulate the interplay, molecular trafficking and communication between the brain and the periphery.[157] The blood and the brain parenchyma endothelial cells of the central nervous system contain specific membrane transporters that regulate the uptake and/or efflux of both endogenous substrates and xenobiotics. This includes organic anion transporter 3 (OAT3), organic anion-transporting polypeptide (OATP1A2), mono-carboxylate transporter 1 (MCT1), breast-cancer-resistance protein (BCRP/ABCG2), P-glycoprotein (P-gp; MDR1), and multidrug-resistance-associated protein 4 (MRP4).[158,159] Cell-based assays consist of primary cell lines or cultured cells arranged as a thin monolayer of BBB endothelial cells, attached with tight junctions (Figure 15).[160–164]

Different types of transport are represented in BBB cell-based assays, namely transcellular, carrier-mediated transport, transport across unstirred layers, and paracellular, although the latter is severely limited due to tight junctions between cells.[165–167] *In vitro*, ions transport through the cell layer is very low due to the tight junctions that minimize the leakage of hydrophilic molecules between the endothelial cells. This produces a trans-endothelial electrical resistance (TEER) that can be monitored to evaluate the integrity of the BBB. [168,169]

BBB cell-based assays suffer from controlling the expression of membrane transporters. This issue constitutes the main drawback of using these models to evaluate transport through a specific membrane protein. In addition, cells are not exposed to the blood flow. Therefore, they lack the influence of shear stress on the endothelial cells present in the *in vivo* conditions.

III.7.1.2. Parallel Artificial Membrane Permeability Assay

The parallel artificial membrane permeability assay (PAMPA) is an artificial membrane model that consists of a porous hydrophobic filter matrix impregnated with a lipid membrane separating two aqueous buffer compartments, the donor well and the acceptor well. The test compound is added into the donor well at a known concentration. Then, it migrates into the acceptor well through the artificial membrane (see Figure 15). [42,142,170,171] PAMPA systems are mainly related to passive diffusion, and are used to predict oral absorption and bioavailability of drugs in the gastrointestinal tract.[172] The main advantages of PAMPA are its low cost and time efficiency compared to cell-based models. It can tolerate pH variations and therefore can be used as a high throughput technique to predict permeability. Most PAMPA models do not include active transport, which is the preferred absorption route of most drug.[14,173,174]

Various compositions of the artificial membrane have been used in PAMPA systems. For instance, n-dodecane solution of DOPC (DOPC-PAMPA), dilute dodecane solution of egg lecithin-based lipid compositions (double sink PAMPA or PAMPA-DS), biomimetic lipid mixture (BM-PAMPA), porcine brain tissue-based (BBB-PAMPA), and skin PAMPA system to assess transdermal entry/entrance.[172,173,175–179]

PAMPA-DS is the most common PAMPA system as it mimics the physiological conditions of the gastrointestinal tract. The first sink is the pH gradients between the donor and the acceptor wells. The pH of the donor well can be adjusted between 3 and 10 depending on the compound and the absorption medium.[14] In the acceptor well, the pH is usually set at 7.4, reflecting the physiological value in the blood. For instance, to mimic the absorption of acidic xenobiotics in the stomach, pH is set at 5, so that the acid molecule enters the gastric epithelium under its neutral form and exit into the blood circulation under its ionized form, then it can be trapped by blood transporters to be distributed in the targeted cells. The second sink is the presence of micellar surfactant agents that play a role of chemical scavengers in the acceptor compartment. They trap lipophilic compounds in the acceptor well and hinder their accumulation inside the hydrophobic membrane core, favoring an unidirectional permeation.

Several factors can alter the correlation of passive permeation coefficients measured by PAMPA to *in vivo*, *in vitro* or *in silico* experiments. For instance, the presence of unstirred water layers (USL), the presence or absence of sinks in the acceptor solution, the formation of aqueous pores, transmembrane pH conditions, or the retention of highly lipophilic compounds.[14,180,181]

III.7.1.3. Black lipid membranes

Black lipid membranes (BLM) were first introduced by Mueller *et al.* in 1962.[182] They are simplistic artificial membrane models composed of a thin film of phospholipids isolated from natural sources such as egg lecithin. Different phospholipids can be utilized including phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylinositol (PI) or sphingomyelin (SM). Essentially, a thin single bilayer film is formed between two sheets of Teflon or polyethylene submerged in aqueous solution. A small lateral

tension is applied to hold the bilayer (Figure 15). BLM has been considered as simplistic model to assess passive permeation of xenobiotics as well as to study the physicochemical characteristics of complex lipid membranes. Upon projection of a light source on the central region of the bilayer to control its quality, it turns out as black spots all over the film, hence the name of black lipid membranes. Moreover, BLM are sensitive to modifications in the electrical properties and allow the measurement of electrical characteristics of membranes that are built in this system.[183,184]

The main drawback of BLM is the instability of the freestanding phospholipid bilayer. For instance, a slight evaporation in one well may cause large fluctuations in lateral membrane pressure, which in turn can disrupt the bilayer.

III.7.1.4. Liposomes

Liposomes are biomimetic membranes widely used to model transport and permeability of xenobiotics thanks to their easiness of use. They are either made of a single phospholipid type or of a mixture of lipids. To prepare liposomes, lipids are solubilized in an organic solvent, which is then evaporated. Lipids deposit on the surface forming a film, which is then hydrated with an aqueous solution to form a liposome suspension.[9]

The diameters of liposomes are in the ranges 20-100 nm (small unilamellar vesicles, SUV), 100 nm - 1 µm (large unilamellar vesicles, LUV) and 1-100 µm (giant unilamellar vesicles, GUV).[42,185–189] When using certain lipid mixtures, care should be taken in the choice of the organic solvent. In addition, oxidation reactions can occur when using unsaturated phospholipids (e.g., PS, PE) and cholesterol in aerobic conditions, and/or high illumination.

The potential presence of several layers in a single liposome can be an issue in predicting permeation through single lipid bilayer membranes.[184]

III.7.2. Experimental methods to assess membrane permeability

The permeation coefficient ($logP_{erm}$) can be experimentally determined using several experimental techniques and various *in vitro* membrane models. In this section, the recent detection techniques to quantify $logP_{erm}$ will be described.

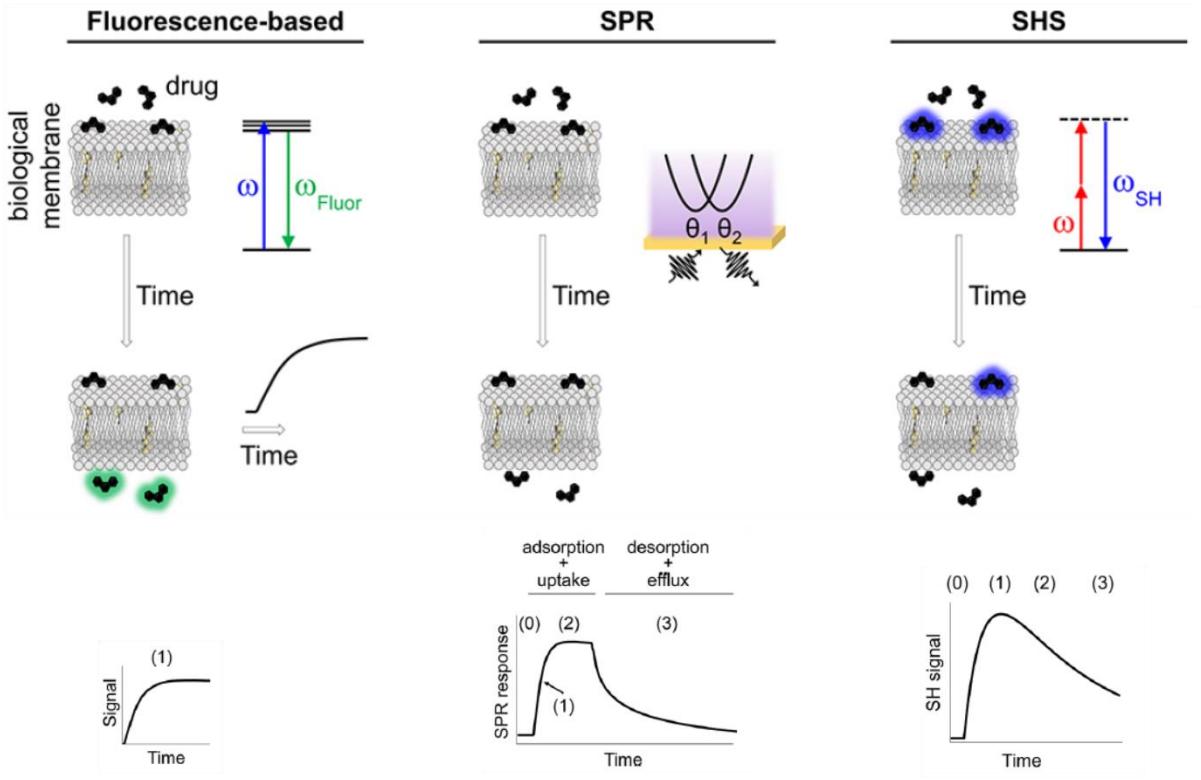


Figure 16: Recent experimental techniques to quantify permeation coefficients.††

III.7.2.1. Fluorescence-based techniques

Based on spectroscopy and imaging techniques, fluorescence-based permeability assays have been widely used to study passive transport of molecules across membrane model systems.[190] Many drug-like molecules contain chromophoric aromatic moieties detectable by UV-visible spectrophotometry. However, these molecules must exhibit high solubility to be detectable.[190–192] Otherwise, a fluorescent dye-host complex can be used to indirectly monitor transport kinetics, such as in the fluorescent artificial receptor membrane assay (FARMA).[193–195] (Figure 16 and Figure 17)

†† Adapted from [42]. Copyright 2022, American Chemical Society.

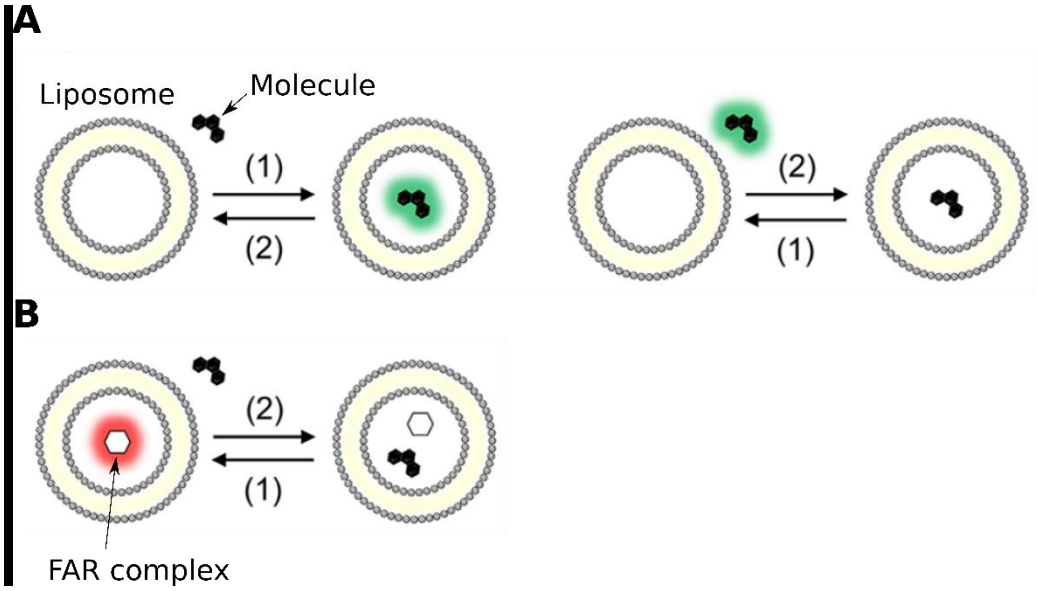


Figure 17: The fluorescence-based techniques. (A) Based on a fluorescent molecule, (B) Based on fluorescent host-dye complex: the fluorescent artificial receptor membrane assay (FARMA). When the molecule of interest (black) binds the fluorescent artificial receptor FAR (red), its fluorescence is turned off, resulting in a decrease of signal intensity.[§§]

As an example, the fluorescent artificial receptor (FAR) consists of a fluorescent complex of a dye (dicationic D1–D3) and a host (macrocycle cucurbit[8] uril; CB8). Upon the competitive binding of the translocating molecule of interest with FAR, the dye is displaced from the host. Thus, the fluorescence intensity decreases. The change in signal intensity can be quantitatively measured in real time using UV-visible spectroscopy, even for non-fluorescent drugs.[196,197] However, the tested compounds should possess high binding affinities to the host. FARMA are used to track the transport of large set of compounds with different chemical scaffolds in liposomes.[193,198] One of the advantages of this assay is that it allows high throughput screening using microplates.

The fluorescence intensity is proportional to the concentration of the molecule c(t) in the inner compartment of the liposome:

$$\mathrm{I(t)} = \mathrm{I_0} + \beta \times c(t) \tag{24}$$

Where I_0 is the initial offset, β is the scaling factor. Correspondingly, the concentration c(t) is related to the rate constant k_p of crossing the vesicle membrane and thus the permeation coefficient P through a phenomenological model, [42,197] where

$$P = \frac{k_p r}{3} \tag{25}$$

Where r is the inner radius of the liposome.

[§§] Adapted from [42]. Copyright 2022, American Chemical Society.

III.7.2.2. Surface plasmon resonance

Surface plasmon resonance (SPR) biosensing is a real time analytical technique that was developed to assess surface-related drug- and protein-membrane interactions, as well as membrane proteins and ligand interactions during the drug discovery processes.[199–202] Recently, several studies on drug−membrane interactions at the surface of lipid bilayers and transport kinetics were carried out using SPR technique. [199,203–205]

SPR is a label-free, sensitive and sufficient technique that allows quantitative kinetics determination and screening assays on both cell-based and artificial planar lipid bilayer systems without the need of fluorescence or radiolabelling. For permeation studies, liposomes are mainly used to assess the transport of molecules through the lipid bilayer. Experimentally, liposomes are immobilized on a SPR sensor surface using DNA strands, attached onto the surface for one strand and to a cholesterol molecule for the other. [206–208]

For SPR technique, a polarized light is emitted from an optical source in a prism, then reflected on biosensor (membrane-solution surface) to a detection unit. The internal reflected light produces an electromagnetic field wave extending a short distance into the solution (evanescent-wave phenomenon) changing the plasmon resonance. (Figure 16)

Upon binding of the molecule of interest at the surface of the immobilized membranes, the refractive index of the environment close to the surface is altered. The amount of the bound molecules, and their affinities to the membrane and association/dissociation kinetics can be measured in real time.[200]

Tracking the SPR response as a function of time can highlight several phenomena: (i) adsorption and transport of the molecule of interest characterized by an increase of the SPR response, (ii) desorption/efflux or dissociation, transfer out of the liposome to regenerate the first free liposome.[42] Passive permeation coefficients can be determined by fitting to the time-resolved SPR response evolution. The linear dependency of the passive permeation rate on the concentration gradient is represented by the observed rate constant k_p.[199] As for the FARMA assay, the permeation coefficient can be determined using equation 34.

SPR has many advantages for permeation studies: it is sensitive, label-free, cost-efficient, can be automatized, and the liposomes can be reused. However it requires large molecular compounds to detect the mass change at the sensor-membrane surface.[209]

III.7.2.3. Second-Harmonic Light Scattering

Second-Harmonic Light Scattering (SHS) is a label-free optical bioassay that is used to investigate surface-specific and surface-related interactions between biomolecular systems in real-time, in particular interactions at the lipid bilayer interfaces.[210–212] It includes binding of proteins at a membrane surface, tracking the kinetics of permeating molecules across the lipid bilayer membrane of liposomes and investigating specific orientations of an adsorbed molecule to the lipid membrane.[213–216] SHS relies on the use of resonant targets to enhance the second harmonic response,[217,218] which is related to the number of molecules bound at the membrane surface. For passive transport studies, time-evolution of the adsorption of the molecules of interest on both sides of the lipid bilayer membrane with opposite orientations yields a decrease in SH response.[219–222] This decrease is proportional to the transport rate of the molecules across the membrane.[223] As for other methods, the transport rate is characterized by rate constants that can be directly converted into permeation coefficients.

SHS provides a distinct analytical advantage compared to linear spectroscopic methods (e.g., fluorescence and UV-visible spectroscopy) as it is highly sensitive and capable to detect low concentrations of the biomolecules of interest (i.e., proteins, nanoparticles, small molecules). SHS also has the advantage over SPR and fluorescent techniques to detect low molecular weight species. However, the SH response is only produced in non-centrosymmetric media such as collagen.[224]

III.7.3. Comparison between experimental models

Comparing passive permeation coefficients between *in vivo* and *in vitro* experiments requires a normalization regarding the differences in both measurement conditions (e.g., paracellular, efflux, and carrier mediated permeabilities, unstirred layers contribution (USL), and accessible surface area). In addition, pH conditions play an important role in the transport of xenobiotics and should be adjusted regarding the type of experiment.[78,86,87,225–230] Avdeef's compilation sets a correction of permeation coefficient for USL, carrier-mediated transport contribution, pH-dependent ionization and non-transcellular effects.[14,137,165,231] Moreover, the same corrections should be carried out to correlate to theoretical predictions of passive permeation coefficients. The total (effective) permeability P_{tot} is the sum of inverse different *n* sequential permeabilities (P_{USL}, P_{trans} …):

$$\frac{1}{P_{tot}} = \sum_{n} \frac{1}{P_n} \tag{26}$$

In the intestine, the Surface Accessibility Area (ASA) is the available epithelial surface that is accessible to the molecule in order to be absorbed. In *in vitro* membrane models (e.g., cell-based), ASA is much smaller than *in vivo*. This may play an important role in reproducing the *in vivo* membrane permeability, especially for hydrophilic molecules whose absorption is supposed to be slow so that they are more exposed to the whole surface of the adjacent cell membranes.[232]

Chapter IV. Theoretical determination of passive permeation coefficients

Based on the ISDM theory (equation (20) section III.5.2), the determination of the passive permeation coefficient (P_{erm}) requires the modelling of two different ingredients/parameters that vary along the membrane normal (z-coordinate), namely, the potential of mean force (PMF, w(z)) and the diffusivity (D(z)).

Given the inhomogeneity of lipid bilayer membranes, the dynamics of the molecule diffusion inside the membrane depends on its interactions with the surrounding lipid components. When a molecule reaches a membrane, it first interacts with the polar head groups. In this region, the molecule can be blocked outside, or insertion can be initiated, depending on its molecular properties. Upon interaction of the molecule with the polar head groups, this region can be perturbed. The subsequent reorganization may either trap the molecule in this region or allow easier diffusion of the molecule to be inserted deeper in the membrane, then to reach the lipid hydrophobic core. Here, the molecule can be trapped at a given depth of the bilayer leaflet due to favorable interactions, or it can diffuse a lot and can possibly jump from one to another leaflet (flip-flop). The transmembrane crossing is usually the rate-limiting step due the energetic penalty encountered by the molecule at the membrane center. The height of the crossing barriers is quantified by the potential of mean force (PMF).[20,233] Most drugs are amphiphilic compounds, their partition in the membrane is usually favorable, with possibly some reorientations in particular in the lipid head group region.

MD simulations can be used to model the PMF and the diffusivity profiles across the membrane.[70,233–235] To get a correct PMF profile along the membrane translocation path, a sufficient sampling of the configurational space along the normal to the membrane surface is required.

IV.1. Introduction to molecular dynamics simulations

Molecular dynamics simulations (MD) simulations are a powerful tool that can be used to describe molecular events, e.g., molecular processes which allow xenobiotics to cross lipid bilayer membranes.[20,233] Based on numerically and iteratively solvation of classical physics equations, MD simulations allow monitoring time-dependent evolution of a molecular system, in particular atom positions, velocities and potential energies.[236] Usually, the algorithms used to proceed MD simulations are applied within specified ensembles of physical constants, mainly the canonical ensembles NVT or NPT, considering constant the number of particles in the system (N), the temperature (T) and either the volume (V) the pressure (P), respectively.[236]

Based on the ISDM theory, permeation coefficients can be calculated by means of the Smoluchowski differential equation (20) in section III.5.2. This is suitable by modeling both (i) the free energy surface w(z) (PMF) picturing the molecule's translocation along the bilayer normal, and (ii) the position-dependent diffusivity D(z) that describe the local motion of the molecule.[237]

IV.1.1. Equations of motion

The dynamics of molecular systems is described by different independent variables, for which the most convenient one are the spatial positions q and momenta p of all atoms. All possible sets/combinations of coordinates (q, p) define the phase space with 6N dimensions, with N being the number of atoms.[238]

For each of the atoms i in the system, the time-evolution of the coordinate q and momenta p are described by the Hamilton equations:[239]

$$\begin{cases} \frac{dp_i}{dt} = -\frac{dV}{dq_i} & (1) \\ \frac{dq_i}{dt} = \frac{dT_i}{dp_i} & (2) \end{cases}$$

Where V is the potential energy function of the system and T_i is the kinetic energy for atom i.

In practice, Newton's equations are used instead. Therefore, during MD simulations, the solution of these equations can be solved numerically by considering the velocity v and the coordinate q of the atoms that moves within a short time step Δt. At time t=0, initial velocities are guessed and the initial coordinates (e.g., obtained from experiments) allow to initiate the iterative solvation of the equations. This is achieved step by step, considering that the coordinates at time t plus an interval of time Δt, depend on the coordinates at t, q(t), as follows:[239,240]

$$q(t + \Delta t) = q(t) + \Delta t.\, v(t) - \frac{\Delta t^2}{2m}\frac{dV(t)}{dq} \quad (3)$$

Similarly, the new velocities at $t+\Delta t$ are given by:

$$v(t + \Delta t) = v(t) - \frac{\Delta t}{2m}\left(\frac{dV(t + \Delta t)}{dq} + \frac{dV(t)}{dq}\right) \quad (4)$$

From equations (3) and (4), we notice that the resolution of the equations of motion strongly depends on the gradient of the potential energy. During MD simulations, the latter can be computed using a force field.

IV.1.2. Force fields

The force field describes all inter-atomic interactions of a molecular system. It relates the potential energy with the internal coordinates of a system (bonds, angles, dihedrals…). These interactions are divided into bonded and non-bonded interactions.[20,241] The former correspond to all covalently linked atoms and consider the potential energy of bonds, angles and dihedrals. The latter are non-covalent interactions, mainly Coulomb and van der Waals. These interactions depend strongly on the nature of the interacting atoms (C, N, O, …), their electronic properties (hybridization, partial charges…), bond types, angle types, or torsion types, which all define the force field parameters (see Figure 18).

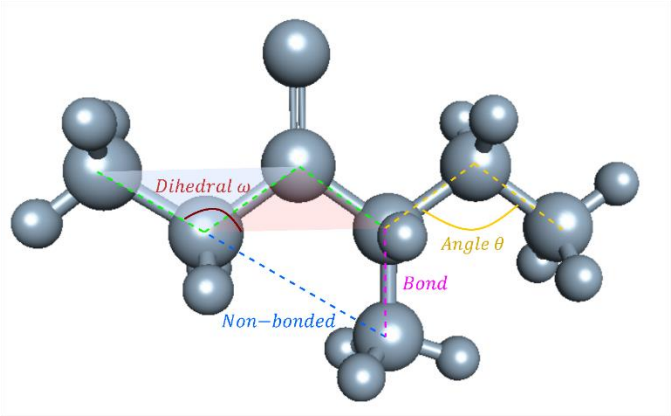


Figure 18: Representation of the force filed bonded and non-bonded interactions

Here, we will only consider additive and non-polarizable force fields, for which atomic partial charges do not change during the MD simulations. Polarizable force fields with varying charge distributions along the simulation are not discussed.

IV.1.2.1. Bonded interactions

Bond potential (V_{bond}):

In classical MD simulations, the potential energy of the bond is described by a harmonic potential:

$$V_{bond} = \frac{1}{2} k_b (r - r_0)^2 \quad (5)$$

Where k_b is the bond force constant, r_0 is the equilibrium distance and r is the distance between the two corresponding atoms.

Angle potential (V_{angle}):

The potential energy of angles is also defined by a harmonic potential:

$$V_{angle} = \frac{1}{2} k_b (\theta - \theta_0)^2 \quad (6)$$

Where k_a is the angle force constant, θ is the angle defined by the atoms.

Torsion potential ($V_{dihedral}$):

The potential energy of the torsion around a bond, is defined by the dihedral angle of four adjacent atoms (1→4) and is described by a Fourier potential as follows:

$$V_{dihedral} = k_t (1 - \cos(n\omega - \gamma))^2 \quad (7)$$

Where k_t, n and γ are parameters defining the shape of the potential, ω is the dihedral angle between the two planes, both defined by three bonded atoms (i.e., atoms 1,2,3 and 2,3,4).

IV.1.2.2. Non-bonded interactions:

Coulomb Interactions (V_{coul}):

The electrostatic potential between two atoms i and j is:

$$V_{coul} = \frac{q_i q_j}{4\pi\varepsilon r_{ij}} \tag{8}$$

Where q_i and q_j are the charges of the atoms i and j, respectively.

van der Waals Interactions (V_{vdw}):

These interactions are modeled by the Lennard-Jones potential that combines an attractive and a repulsion contribution:

$$V_{LJ} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{9}$$

Where ε_{ij} is the depth of the V_{LJ} potential between atoms i and j, σ_{ij} is the separation between the atoms when $V_{LJ} = 0$, and r_{ij} is the distance between atoms i and j. V_{LJ} is supposed to be different for each pair of atoms.

When adding all the energetic terms which picture most of inter-atomic interactions, the general formula for the potential energy of the system is given by:

$$V_{tot} = \sum_{i=1}^{N_b}\frac{1}{2}k_b(r-r_0)^2 + \sum_{i=1}^{N_a}\frac{1}{2}k_b(\theta-\theta_0)^2 + \sum_{i=1}^{N_d}k_t(1-\cos(n\omega-\gamma))^2 + \sum_{ij}\frac{q_i q_j}{4\pi\varepsilon r_{ij}} + \sum_{ij}4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{10}$$

IV.2. Biased MD methods to obtain the free energy

With the current computational facilities, conventional MD simulations cannot easily capture molecular motions for systems too large (number of atoms greater than a million) or too long molecular processes (longer than tens of µs).[242] As both parameters are inversely proportional to the simulation speed (in ns/day), tradeoffs are possible, for instance very large systems but for short simulations. However, some events occur in time scales beyond the reach of current unbiased MD simulations. In particular, the translocation of molecules across lipid bilayer membranes is driven by free energy barriers that hinder the straightforward exploration of the configurational space, especially inside the central hydrophobic membrane core.[20,236] For instance, for small molecules, the potential of mean force underlying their translocation through a lipid bilayer membrane would require averaging over many crossings in a large number of

long simulations, hence necessitating huge computational effort[20]. This type of non-biased simulations can only sample states for relatively small molecular systems and ns to µs time scale. The probability distribution of a molecule to be in different states can be readily calculated leading the estimation of the PMF profile. However, this is not affordable for most amphiphilic drug-like molecules where the different possible states are separated by very high energy barriers, which are not visited by using conventional MD simulations, leading to a non-ergodic behavior.

Alternatively, enhanced sampling can be afforded by biased MD simulations, which have proven their time-effectiveness and robustness to extend the range of sampling and to assess the free energy landscape of complex molecular systems (i.e., allowing suitable translocation between the different states). Various methods have been developed to model the free energy surfaces of different molecular systems. For instance, umbrella sampling (this was used by Marrink and Berendsen in their original study)[243,244], Adaptive Biased Force method, Metadynamics[245,246], Alchemical free energy calculations[70] or Accelerated Weight Histogram (AWH) method[247]. In the following sections, we will cover basic concepts of these methods and their applicability in the context of predicting PMF profiles of membrane permeability.

If the energy barrier to cross between two states is larger than the thermal free energy (k_bT), the system will likely be trapped in given regions close to one state and cannot populate the other states (quasi-nonergodicity) (see Figure 19). Usually, molecular systems contain many degrees of freedom that cannot be all sampled during conventional MD simulations. Most enhanced sampling methods account for one to three degrees of freedom, called reaction coordinates. For passive permeation processes, the reaction coordinates can straightforwardly be related to the z-axis, normal to the lipid bilayer surface.

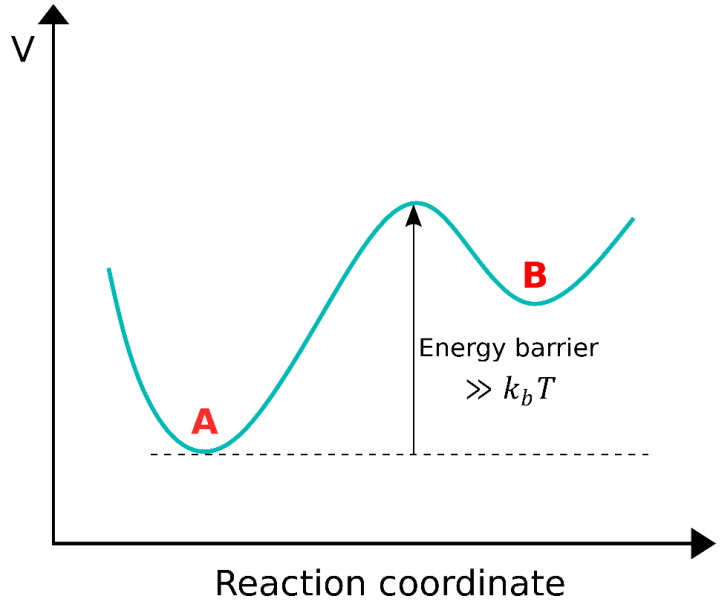


Figure 19: Schematic representation showing the high-energy barrier in the free energy surface that hinders the transitions from state A to state B.

IV.2.1. Introduction to statistical thermodynamics

At thermodynamic equilibrium, the free energy F of the system for a thermodynamic state equals to: [248]

$$F = -\frac{1}{\beta}\ln Q \tag{11}$$

Where Q is the partition function describing the number of states accessible at a given temperature. Q is related to the Hamiltonian H as follows:

$$Q = \frac{1}{N!\,h^{3N}}\iint \mathrm{e}^{-\beta H(x,p)}dxdp \tag{12}$$

Where N is the number of particles, h the Planck's constant, β is the inverse temperature and it equals to $1/(k_bT)$, x are the coordinates and p the momenta.

The Hamiltonian describes the total energy of the thermodynamic system:

$$H(x,p) = \sum_{i=1}^{N} K(p) + V(x) \tag{13}$$

Where K(p) is the kinetic energy which depends on the momentum p, and V(x) is the position-dependent potential energy. In classical MD simulations, V(x) is described by a force field. At thermodynamic equilibrium, only the configurational term of the partition function is considered since the temperature and particle's masses do not change. Therefore, the kinetic contribution to the Hamiltonian is considered constant. Moreover, in the isothermal-isobaric NPT ensemble, the low change in average volume at p = 1atm does not affect the free energy between states.[249]

The configurational partition function Z can be used instead of Q. In the NPT ensemble, the free energy is described by the Gibbs free energy G, which approximates the Helmholtz free energy A in the phase space Γ:[236,248]

$$G \approx A = -\frac{1}{\beta}\ln Z = -\frac{1}{\beta} ln \int_{\Gamma} \mathrm{e}^{-\beta U(x)} \tag{14}$$

The normalized configurational distribution at state k is described by the Boltzmann distribution as follows:

$$P_k = \frac{1}{Z}\mathrm{e}^{-\beta U_k} = \frac{\mathrm{e}^{-\beta U_k}}{\int_k \mathrm{e}^{-\beta U_k}} = e^{F_k - \beta U_k} \tag{15}$$

Where:

$$F_k = \frac{1}{\beta}\int_k \mathrm{e}^{-\beta U_k} \tag{16}$$

For complex molecular systems, an analytical expression for Z cannot be derived. Only free energy differences between macrostates or between systems can be computed. For instance, the free energy difference between two states A and B is given by:

$$\Delta G_{A,B} = G_B - G_A = -\frac{1}{\beta}\ln Z_B + \frac{1}{\beta}\ln Z_A = -\frac{1}{\beta}\ln\frac{Z_B}{Z_A} \quad (17)$$

In MD simulations, we can compute the free energy from the potential energy if we have a sufficient overlap between two phase spaces. Conversely, when the two targeted states are distant from each other, intermediate states can be simulated and equation (17) can be applied between neighboring states.

IV.2.2. Alchemical calculations

Alchemical free energy calculations are efficient and practical approaches to calculate free energies of transfer between two end-states (e.g., bound and unbound states) by decoupling the interactions (electrostatics and van der Waals) between the molecule and its environment (i.e., aqueous medium, membrane or protein binding site), using bridging potential energy functions (see Figure 20).[250] These functions are represented by non-physical (alchemical) intermediates connecting the end-states using a coupling parameter λ.

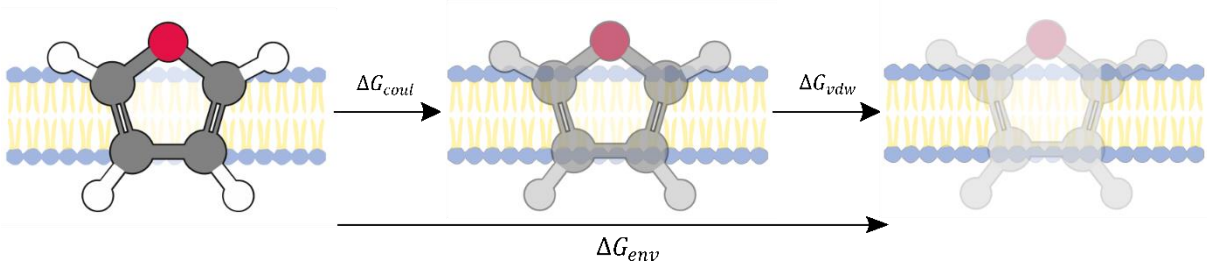


Figure 20: Illustration of the interaction decoupling during alchemical free energy calculations between a molecule and a membrane at a given depth z. First the electrostatic interactions are removed gradually, then the vdW interactions.

Recently, alchemical free energy calculations have been used to model the PMF profiles of small molecules in lipid bilayer membranes.[70] The PMF profiles can then be used to compute passive permeation coefficient.

Based on equation (17), the free energy of transition between the fully uncoupled (λ_B) and the fully coupled (λ_A) states is:[249,250]

$$\Delta G_{env} = -\frac{1}{\beta}\ln\frac{Z(\vec{\lambda}_B)}{Z(\vec{\lambda}_A)}) = -\frac{1}{\beta}\ln\frac{\int_\Gamma e^{\left(u\left(\vec{q};\vec{\lambda}_B\right)\right)d\vec{q}}}{\int_\Gamma e^{\left(u\left(\vec{q};\vec{\lambda}_A\right)\right)d\vec{q}}} \quad (18)$$

Where $u(\vec{q};\vec{\lambda})$ is the reduced potential which can be simplified, in a thermodynamic equilibrium, as follows:

$$u(\vec{q};\vec{\lambda}) = \beta U(\vec{q};\vec{\lambda}) \quad (19)$$

Here, U is the potential energy function of the coordinate q at state λ_k. When the two end-states are distant from each other, the free energy cannot be directly estimated from computing the partition function of the two end-states, as the overlap between these two states is poor or absent. In practice, the decoupling of a molecule from its environment (e.g., membrane or water) relies on introducing intermediate coupling parameters λ connecting the two end-states.

In alchemical calculations, the potential is scaled alchemically, namely applying a linear transformation, in the simplest case, so that the net potential energy, $U(\lambda, \vec{q})$, can be computed as follows:

$$U(\lambda, \vec{q}) = (1-\lambda)U_0(\vec{q}) + \lambda U_1(\vec{q}) \quad (20)$$

Where U_0 and U_1 are the potentials of the two end states.

Generally, for each λ, a separate MD simulation is performed. Therefore, the free energy between pairs of neighboring λ states can be computed by:

$$\Delta G_{env} = \sum_{i=0}^{i-1} \Delta F(\vec{\lambda}_i, \vec{\lambda}_{i+1}) = \sum_{i=0}^{i-1} -\frac{1}{\beta} \ln \frac{Z(\vec{\lambda}_{i+1})}{Z(\vec{\lambda}_i)} \quad (21)$$

Where $\Delta F(\vec{\lambda}_i, \vec{\lambda}_{i+1})$ is the free energy between two successive λ states. Moreover, the absolute free energy of environment (ΔG_{env}) can be computed using various estimators such as Thermodynamic Integration (TI), Bennett Acceptance Ratio (BAR) or the Multistate Bennett Acceptance Ratio (MBAR).[250,251]

IV.2.2.1. Thermodynamic Integration

The Thermodynamic Integration (TI) is used to calculate the free energy difference between two separate states A and B of λ = 0 and λ = 1, respectively. The free energy difference ∆f can be obtained by integrating the ensemble average, over all λ states between A and B, of the derivative of the potential energy function with respect to λ:

$$\Delta f = \int_0^1 \frac{df}{d\vec{\lambda}} = \int_0^1 \left\langle \frac{du(\vec{\lambda}, \vec{q})}{d\vec{\lambda}} \right\rangle_{\vec{\lambda}} d\vec{\lambda} \quad (22)$$

In practice, the simplest numerical integration scheme is the trapezoid rule where ∆f can be expressed as follows:

$$\Delta f_{ij} = \frac{\Delta\lambda_{ij}}{2} \left[\left\langle \frac{\partial u_\lambda}{\partial \lambda} \right\rangle_{\lambda_i} + \left\langle \frac{\partial u_\lambda}{\partial \lambda} \right\rangle_{\lambda_j} \right] \quad (23)$$

In the context of membrane permeation, the PMF profile can be modeled using alchemical calculations.[70] It has been suggested that in order to obtain the overall PMF along the membrane thickness, the molecule of interest can put at specified depths of the membrane, and then the absolute free energy difference is calculated using alchemical transformation calculations. [70] These depths are specified by the user and should reflect the different chemical interfaces of the membrane system (i.e., aqueous medium, polar head groups, glycerol

moieties, hydrophobic core of the acyl chains and the membrane center). The overall PMF profile can be obtained by interpolating in between these points. The main issue of this method is the bias of the numerical quadrature that increase the level of statistical noise, i.e., may lead to large uncertainties during the propagation of the error through λ.

IV.2.2.2. Bennett Acceptance Ratio (BAR) method

The Bennett Acceptance Ratio (BAR) method is used to significantly improve the estimate of the free energy difference between neighboring states. The free energy difference between i and j neighboring states can be written as follows:

$$\Delta f_{ij} = \beta \Delta G_{ij} = -\ln\langle e^{-\Delta u_{ij}} \rangle_i \tag{24}$$

From equation (24), Δf is critically dependent on the convergence of the ensemble average. In other words, Δu_{ij} should converge and the distributions of the simulated molecule in both states should overlap. BAR method is used to solve this issue by considering the potential free energy differences: (i) from the state i to the state j and (ii) from the state j to the state i. The reduced free energy difference equals:

$$\Delta f_{ij} = \ln \frac{\left\langle \frac{1}{1 + e^{-\Delta u_{ij} + \Delta f_{ij}}} \right\rangle_j}{\left\langle \frac{1}{1 + e^{-\Delta u_{ij} - \Delta f_{ij}}} \right\rangle_i} + \Delta f_{ij} - ln \frac{N_j}{N_i} \tag{25}$$

Where N_i and N_j are the number of the samples performed in states i and j, respectively. Equation (25) relies on solving the following formula:

$$\sum_{i=1}^{N_i} \frac{1}{1 + e^{-\Delta u_{ij} + \Delta f_{ij}}} = \sum_{j=1}^{N_j} \frac{1}{1 + e^{-\Delta u_{ij} - \Delta f_{ij}}} \tag{26}$$

Given the samples from both states i and j, this equation can be solved in a self-consistence manner and then Δf_{ij} can be predicted. The overall free energy difference between end-states is the sum of Δf_{ij} over the intermediate λ-states.

IV.2.2.3. Multistate Bennett acceptance ratio method

In the same spirit as BAR, the Multistate Bennett acceptance ratio (MBAR) method is used to estimate the free energy differences between all states instead of two neighboring states for BAR method. The general formula for the free energy f_i at state i is:

$$f_i = -\ln \sum_{n=1}^{N} \frac{e^{-u_i(\vec{q}_n)}}{\sum_{k=1}^{k} N_k e^{f_k - u_k(\vec{q}_n)}} \tag{27}$$

Where N is the total number of samples over all states, N_k is the number of samples at each state k. MBAR has the lowest variance of the free energy[252] compared to the the TI method.

IV.2.3. Umbrella sampling

The Umbrella Sampling (US) method allows computing the free energy landscape underlying the translocation of a molecule along the reaction coordinate ξ, projected on the z-axis perpendicular to the membrane surface.[253] ξ is discretized into several overlapped windows i (Figure 21).

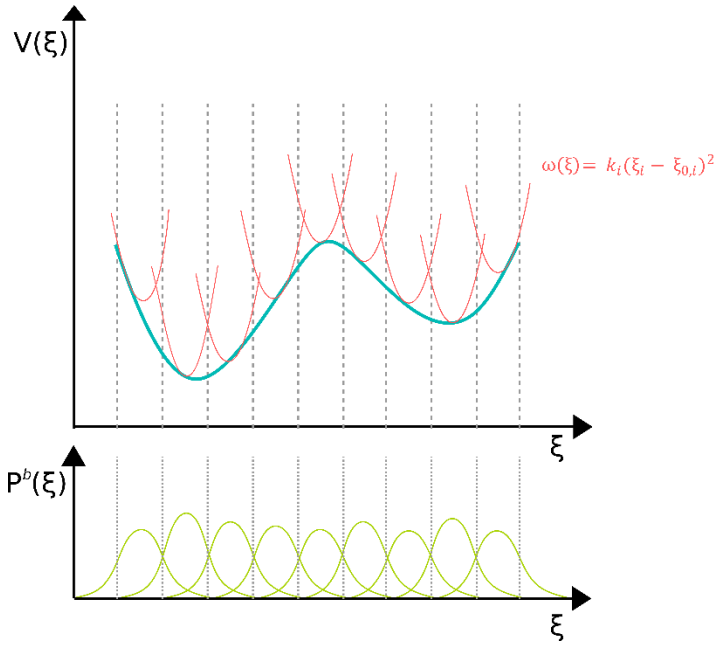


Figure 21: Illustration of the Umbrella sampling protocol. The reaction coordinate ξ is split into several windows. At each window, MD simulations are run using harmonic potential (umbrella) as a bias (red). The bottom panel shows the computed biased probabilities. They can be used to yield the unbiased free energy profile (blue) using WHAM method.

At each window i of ξ, the MD simulation runs are performed using a harmonic potential ω(ξ) that is applied on the restrained molecule of interest. The biased potential V^b of the system is:

$$V^b = V^u + \omega(\xi) \quad (28)$$

Where V^u is the unbiased potential and the artificial potential is:

$$\omega(\xi) = k_i(\xi_i - \xi_{0,i})^2 \quad (29)$$

Where $\xi_{0,i}$ and ξ_i are the window center and the interval positions of window i, respectively, and k_i the force constant. Subsequently, the probability distributions $P^b(\xi)$ at each of the overlapped windows i are obtained (i.e., along ξ).

In the canonical NVT ensemble, the Helmholtz free energy can be calculated from the probability distribution:[253]

$$A^u(\xi) = -\frac{1}{\beta} \ln P^u(\xi) \quad (30)$$

In the NPT ensemble, the Gibbs free energy can be computed as follows:

$$\mathrm{G}^{u}(\xi) = -\frac{1}{\beta}\ln \mathrm{P}^{u}(\xi) \tag{31}$$

To compute the unbiased free energy $\mathrm{G}^{u}(\xi)$, the unbiased probability $\mathrm{P}^{u}(\xi)$ has to be estimated. This requires to relate $\mathrm{P}^{u}(\xi)$ and $\mathrm{P}^{b}(\xi)$. Therefore, the equation is solved iteratively until convergence. The Weighted Histogram Analysis Method (WHAM)[249,254] can be used to combine the distributions and yield the free energy profile (i.e., the PMF) along ξ.

$$\mathrm{G}^{u}(\xi) = f\left(\mathrm{P}^{b}(\xi)\right) \tag{32}$$

The choice of the k_i parameter is a rate-limiting step during the US process. Setting this parameter to a very small value will drive the system far from the free energy barrier and the total biased potential will be close to the unbiased one. In other words, we are falling in the conventional MD simulations limits ($\mathrm{V^b} \approx \mathrm{V^u}$). Conversely, very high force constant will produce narrow probability distributions, which will lead to an inaccurate free energy profile. Commonly, a convenient k_i value, which can reproduce overlapped probability distributions, is required.

The rate of convergence in each window is the major disadvantage of US as it is comparable to classical MD, and may lack of sampling for the molecule of interest. The global simulation time is increased compared to other biased methods such as alchemical calculations. US can combined with other enhanced sampling techniques to overcome these issues (e.g., replica exchange).[255,256]

IV.2.4. Well-tempered Metadynamics

Metadynamics is a widely used biasing method to predict the free energy surface and overcome sampling limits.[257,258] In metadynamics, a history-dependent bias potential U_b (ξ, t) is applied along a predefined collective variable ξ. U_b (ξ, t) consists of a sum of Gaussian kernels that are periodically applied along ξ and time t; it is updated at each of n iterations (see Figure 22) and the sum is given by:

$$\mathrm{U}^{b}(\xi, \mathrm{t}) = \sum_{k=1}^{n} e^{-\frac{\beta}{\gamma-1}\mathrm{U}^{b}_{k-1}(\xi_k)} G(\xi, \xi_k) = \sum_{k=1}^{n} h\, e^{-\frac{\beta}{\gamma-1}\mathrm{U}^{b}_{k-1}(\xi_k)} e^{-\frac{1}{2}\sum_{i=1}^{d}\frac{\left(\xi_i-\xi_{k,i}\right)^2}{\sigma_i^2}} \tag{33}$$

Where γ is the bias factor, h is the Gaussians height and G(ξ, ξ_k) is the Gaussian function:

$$G(\xi, \xi_k) = h\, e^{-\frac{1}{2}\sum_{i=1}^{d}\frac{\left(\xi_i-\xi_{k,i}\right)^2}{\sigma_i^2}} \tag{34}$$

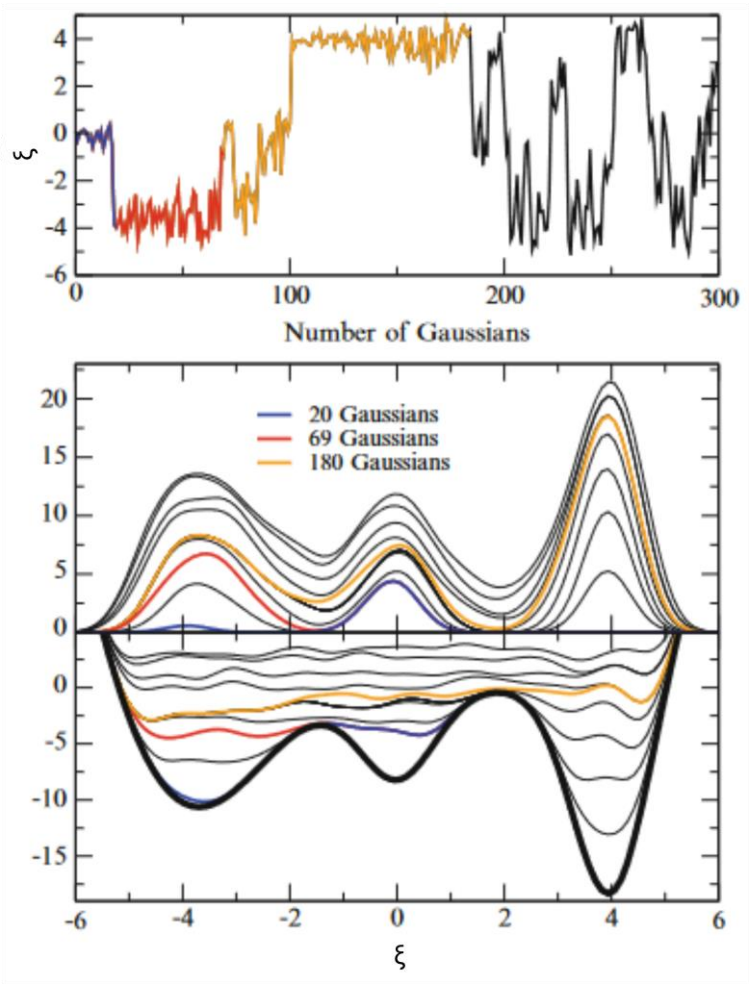


Figure 22: Illustration of the Metadynamics procedure. The upper panel shows the evolution of the collective variable ξ over the simulation trajectory. The middle panel shows results obtained with different numbers of Gaussians (evolution of bias potential G). The system escapes to the second minimum after filling the first minimum using Gaussian functions (blue line); red line reflects when also the second minimum is filled and the orange line reflects when the whole profile is filled. In the lower panel, the evolution of the free energy surface with added bias potential with different number of Gaussians, and the final energy surface (thick black line) at convergence.[§§§§]

Well-tempered metadynamics (WT-MetaD) is a form of metadynamics where the more a region is sampled, the lower the height of the Gaussian hills (h) becomes. Upon the increase of the simulation time, the bias potential exhibits asymptotic regime and becomes much smaller. This accelerates the convergence of the free energy profiles.[246,260]

$$\mathrm{U}^{b}(\xi, \mathrm{t} \to \infty) = -\left(1 - \frac{1}{\gamma}\right) A(\xi) + c(t) \tag{35}$$

[§§§§] Adapted from [259]

Following this equation, the free energy surface A(ξ) can be obtained from the bias potential using several reweighting methods.[261–265] WT-MetaD and US have been used to assess the PMF profile of permeating molecules such as water and arginine by discretizing the reaction coordinate to the distance between the centers of mass (COM) of the molecule and the membrane, along z-axis (the membrane normal).[266] The accurate choice of input parameters constitutes the main limitation of WT-metaD as it affects the rate of convergence of the free energy profiles.

IV.2.5. Adaptive biasing force method

The adaptive biasing force (ABF) method[267,268] is based on thermodynamic integration formalism[236,269] where the average force is calculated to yield the PMF. ABF is used to assess the free energy surface in molecular systems exhibiting high free energy barriers. In ABF, the free energy along the one-dimensional reaction coordinate (ξ) is seen as the potential, V, that results from the average force F where $F(\xi) = -\frac{\partial V(\xi)}{\partial \xi}$. During the simulation, the gradient of the free energy surface A(z) is related to the average force applied on ξ as follows:[2]

$$\frac{\partial A}{\partial \xi} = -\langle F_{\xi}(x)\rangle_{\xi=z} \tag{36}$$

In ABF, the gradient of the free energy (i.e., the average force) is calculated after a sufficient sampling time, and is then used to apply a time-dependent force $F_t^{bias}(z)$ on z that opposes the estimated average force. Therefore, the resulting force cancels out, leading to a local flat free energy. By this way, the exploration of the reaction coordinate is enhanced by bypassing the free energy barriers. While the simulation is under progress, mainly at long simulation time, the negative gradient of the free energy converges and equals the applied biasing force. Consequently, the effective free energy surface can be estimated by integrating all the average forces over the transition coordinate ξ(x).

$$\mathrm{A}(\xi) = \int -\langle \mathrm{F}_{\xi}(\mathrm{x})\rangle_{\xi(\mathrm{x})=\mathrm{z}}\, \mathrm{d}\xi(\mathrm{x}) \tag{37}$$

In practice, the reaction coordinate is usually split into several overlapping strata, or windows, ($\xi(x) \in [z_{min}, z_{max}]$) in order to improve the efficiency of the free energy estimate (i.e., exploring more efficiently the reaction coordinate).[269–271] The ABF method was used to estimate the free energy profiles in several biological phenomena.[271–275] Moreover, it has been used to model the PMF profile underlying the translocation of small molecules across lipid bilayers[70,270,276].

IV.2.6. Accelerated Weight Histogram method

The accelerated weight histogram (AWH) method is an efficient extended-ensemble method that can be used to predict the free energy landscape.[247,277,278] In AWH, a history-dependent bias potential is applied adaptively to flatten the free energy barriers along the chosen the reaction coordinate ξ. To proceed in the free energy calculations, the AWH algorithm acts on a reference the reaction coordinate ξ is discretized into grid points, λ, which constitutes a dynamic variable.

During MD simulations, the bias g(λ) is assigned as a weighting factor. The states are sampled according to a joint distribution P(x, λ), where

$$P(x,\lambda) = \frac{1}{\mathrm{Z}} e^{g(\lambda)-\mathrm{E}(\mathrm{x},\lambda)} \tag{38}$$

Where x is the configuration of the system (x ∈ X). This yields the marginal distribution P(λ)

$$P(\lambda) = \int P(x,\lambda)\, dx = \frac{1}{\mathrm{Z}} e^{g(\lambda)-\mathrm{f}(\lambda)} \tag{39}$$

Where f(λ) is the dimensionless free energy (f(λ) = βF(λ)). The weighs $e^{g(\lambda)}$ are adjusted during the simulation so that P(λ) approaches a predefined target distribution $\rho(\lambda)$, often chosen to be flat. The bias function is thus equal to:

$$g(\lambda) = f(\lambda) + \ln \rho(\lambda) \tag{40}$$

In the case of uniform (flat) target distribution, the bias function g(λ) is tuned to be equal to the desired free energy f(λ) along the system parameter λ though several updates. Here, neither the free energy nor the bias is known. Therefore, it requires to proceed adaptively until the convergence is reached (i.e., the marginal probability approaches the target distribution).[247,279]

In practice, the transition of λ follows a conditional probability given for the current configuration x

$$\omega(\lambda|x) = P(\lambda|x) = \frac{e^{g(\lambda)-E(x,\lambda)}}{\sum_{\lambda'} e^{g(\lambda')-E(x,\lambda')}} \tag{41}$$

which can computed using a Gibbs sampler[280].

After collecting n_λ samples, the free energy is updated ($f \leftarrow f + \Delta f$) based on the weights ω(λ|x) that consists of the samples of the AWH method, where:

$$\Delta f(\lambda) = -\ln\left(1 + \frac{n_\lambda\, \bar{\omega}(\lambda)}{N\rho(\lambda)}\right) + c \tag{42}$$

Here c is a constant, N is the total number of the samples, n_λ is the number of λ samples after each update, (λ) is the target distribution and $\bar{\omega}(\lambda)$ corresponds to $n_\lambda \bar{\omega}(\lambda) = \sum_{i=1}^{n_\lambda} \omega^i(\lambda)$. The effective number of samples is then updated by n_λ: $N \leftarrow N + n_\lambda$. It is worth noting that the update size Δf is proportional to 1/N.

Finally, the bias, g(λ), is updated following equation (40) and is equal to the new free energy (f(λ)) in the case of uniform (flat) target distribution. The bias update evolves inversely proportional to the number of samples N. However, at the very beginning of the simulation, the hyperbolic decay of update size Δf would cause the PMF to depend largely on the initial value of the reaction coordinate ξ. Therefore, instead of a rapid hyperbolic decay, the simulation is divided in two stages: (i) during the initial stage, the update size Δf is kept constant until the whole reaction coordinate has been explored; by this way, the same (coarse) update size is applied along the whole reaction coordinate. Then, each time the whole reaction coordinate is explored (i.e. a covering has occurred), the update size is divided by 3 (Figure 23). (ii) the final stage is simply to let the number of samples evolve linearly over time, thus the update size decreases as 1/t. The final stage allows the fine-tuning of the bias.

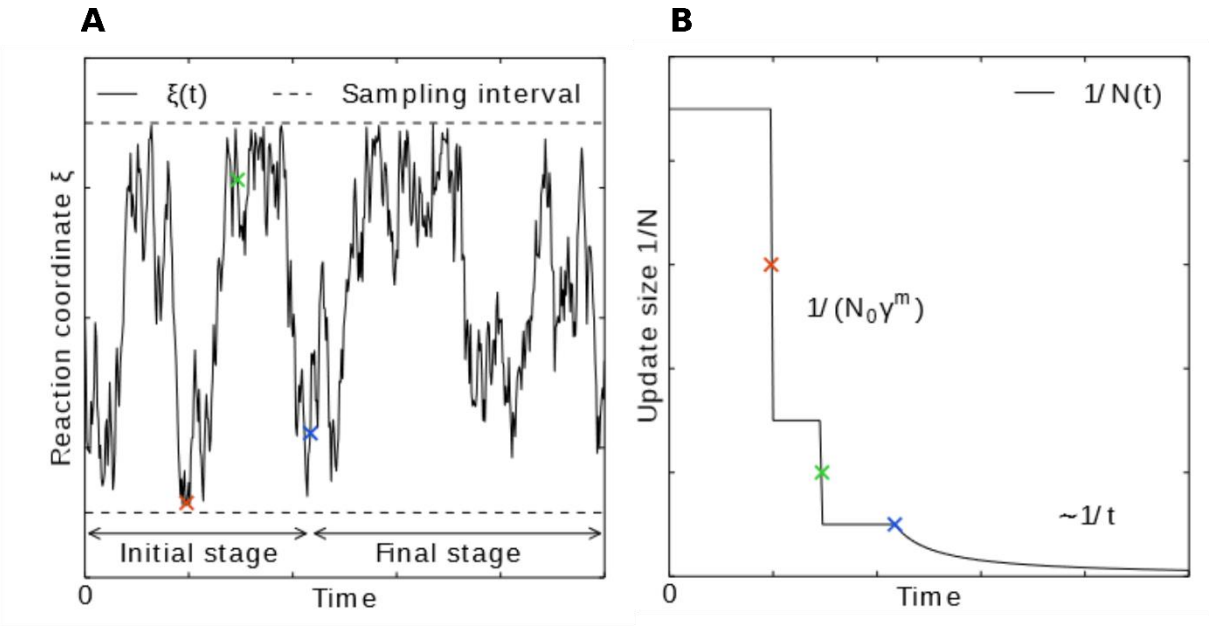


Figure 23: (A) Evolution of the reaction coordinate over time (t) for a Brownian particle in a double-well potential, where the particle covers the reaction coordinate several times in the initial stage before entering the final stage. (B) The times of covering are shown as ×-markers of different colors. At these times the free energy update size $\sim 1/N$ at different times of covering (shown in panel A), where N is decreased by scaling N by a factor of $\gamma = 3$. [*****]

The AWH algorithm works internally along the discretized coordinate λ. However, the free energy of interest is the PMF along ξ (Φ(ξ)). It can be estimated by coupling the system to harmonic potentials Q(ξ, λ) (on-the-fly reweighting procedure) that relates the reaction coordinate ξ to λ:

$$e^{-f(\lambda)} = \int e^{-Q(\xi,\lambda)} e^{-\Phi(\xi)} d\xi \tag{43}$$

Where the umbrella potential $Q(\xi,\lambda) = -\frac{k}{\beta}(\xi-\lambda)^2$, k is the force constant. For large k, F(λ) ≈ Φ(ξ).

Before running the simulation, the initial update size N_0 has to be set by the user and it is linked to the user-defined diffusion coefficient D and the initial error of the free energy ε_0:

$$\frac{1}{N_0} \sim D\varepsilon_0^2 \tag{44}$$

The performance of the algorithm is largely insensitive to these values, as the update size exponentially converges at the beginning of the simulation. The diffusion coefficient can be roughly guessed from a previously calculated diffusion coefficient of the molecule along the reaction coordinate. If the input parameters D and ε_0 are difficult to estimate, they can also be determined by a simple trial and error procedure. Indeed, if D and ε_0 were set too low, the

[*****] Gathered from [279]

system might be pulled apart, and if they were set too high, the update size would be too small, therefore this yields slow convergence.

In the context of passive membrane permeation, we have used the AWH method, for the first time on all-atom membrane systems, to estimate the PMF profile along the POPC lipid bilayer membrane for a set of different molecules (i.e, small and drug-like xenobiotics) (Chapter 5 and 6). The reaction coordinate (ξ) was specified to be the distance between the centers of mass of the molecule and the membrane. The AWH method shows its simplicity to setup and time efficiency compared to alchemical calculations or other adaptive biasing methods (e.g., US, ABF, metadynamics).

Several advantages can be derived from the AWH method: (i) simple setup as it requires less input parameters than other adaptive biasing techniques, (ii) it allows exploration of multiple pathways of the system and thus being quite independent on the initial configurations and (iii) the Gibbs sampler allows large transitions in parameter space λ.

IV.3. Diffusivity

Along with the PMF, the diffusivity is an important parameter in the prediction of passive permeation coefficients. Using conventional MD simulations, the dynamics of diffusion of a molecule inside the membrane can be limited when it is trapped in regions with low free energy. There exist several methods to determine the diffusivity coefficients from the trajectories of equilibrium and non-equilibrium MD simulations.

IV.3.1. Mean-Square Displacement (MSD) method

MD simulations allow to compute one-dimensional diffusion coefficients D(z) (i.e., along z-axis, the reaction coordinate) based on calculating the mean square displacements (MSD(z)) of the molecule as a function of several time steps. Here, the position-dependent diffusion coefficient can be computed using the Einstein formula:[281]

$$\mathrm{D(z)} = \frac{1}{2}\lim_{t\to\infty}\frac{d}{dt}\mathrm{MSD(z)} = \frac{1}{2}\lim_{t\to\infty}\frac{d}{dt}\langle\big(z(t)-z(0)\big)^2\rangle \qquad (45)$$

From equation (45), we notice that the accurate estimate of the MSD requires sufficiently long simulations. In other words, the diffusive process occurs when averaging the MSDs over long time-steps. This is often a limiting step in conventional MD simulations, which pictures local MSDs as the exploration of the reaction coordinate is limited in the PMF profile. A convenient way to overcome this issue is to compute local diffusion constants over short times (1-5ps) where the local PMF (w(z)) is supposed to be relatively constant.[126] In the absence of a mean force applied on the molecule during its translocation through the lipid bilayer, i.e., flat free energy landscape, the MSD can be computed using longer lag-times.[270]

IV.3.2. Autocorrelation methods

For restrained systems using a harmonic potential (e.g., a given window of the reaction coordinate)[282,283], the degrees of freedom of the molecule are limited and thus its motion can be described by the generalized Langevin equation for a harmonic oscillator.[284] Therefore, the position-dependent diffusivity can be calculated using the velocity autocorrelation function as introduced by Roux and co-workers.[285,286] Moreover, Hummer[237] reformulated Roux and co-

workers' formula and derived the position autocorrelation function to calculate the position-dependent diffusion coefficient. For a harmonically restrained molecule at a certain position i (e.g., a given window of the reaction coordinate)[282,283], the diffusion coefficient can be computed using the following equation:[237,283]

$$\mathrm{D}(\xi_\mathrm{i}) = \frac{\langle \delta\xi^2 \rangle_\mathrm{i}}{\tau_\mathrm{i}} \tag{46}$$

Where ξ is the reaction coordinate, τ is the time correlation function of the reaction coordinate ξ which equals:

$$\tau_\mathrm{i} = \frac{\int_0^\infty dt\ \langle \delta\xi(\mathrm{t})\delta\xi(0) \rangle_\mathrm{i}}{\langle \delta\xi^2 \rangle_\mathrm{i}} \tag{47}$$

With $\delta\xi = \xi(\mathrm{t}) - \langle\xi\rangle$.

IV.3.3. Force correlation method

This method was used to predict diffusion coefficients over free energy barriers.[126] Position-dependent diffusion (D(z)) and friction coefficient (f(z)) are related through the Einstein's equation:[20,126,287]

$$\mathrm{D}(\mathrm{z}) = \frac{\mathrm{k_b T}}{\mathrm{f}(\mathrm{z})} \tag{48}$$

On the other hand, based on the fluctuation-dissipation theorem, the local time-dependent friction constant f(z,t) can be calculated from the autocorrelation function of the random forces that are acting on the molecule:

$$\mathrm{f}(\mathrm{z},\mathrm{t}) = \frac{\langle \Delta F(z,t)\ \Delta F(z,0) \rangle}{\mathrm{k_b T}} \tag{49}$$

Where $\Delta F(z,t) = F(z,t) - \langle F(z,t) \rangle$.

From equation (49), the position-dependent friction coefficient can be calculated by integrating over time the local time friction coefficients:

$$f(z) = \int_0^\infty f(z,t)\, dt = \frac{1}{k_b T} \int_0^\infty \langle \Delta F(z,t)\ \Delta F(z,0) \rangle\, dt \tag{50}$$

Combining equation (48) and equation (50), the diffusion coefficient is given as a function of the force autocorrelation function:

$$\mathrm{D}(\mathrm{z}) = \frac{(\mathrm{k_b T})^2}{\int_0^\infty \langle \Delta F(z,t)\ \Delta F(z,0) \rangle\, dt} \tag{51}$$

IV.3.4. Bayesian method

IV.3.4.1. Classical diffusivity from Smoluchowski diffusion equation

For most drug-like molecules that are diffusing across lipid bilayer membranes, the energetic variations in the PMF (w(z)) are higher than k_bT in some regions of the membrane, in particular at the hydrophobic membrane center. Therefore, estimating diffusivity using Einstein equation (45) is not an appropriate approximation, as an unacceptable bias is introduced.[126] The Bayesian inference method can be used to determine, self-consistently, the position-dependent diffusivity D(z) but also the free energy w(z) from the trajectory data of MD simulations.[237,288] Long MD runs or short independent replica runs can be used to assess these parameters. Furthermore, Comer and co-workers[276,289] developed a methodology to compute the position-dependent diffusivity in biased MD simulations.

The motion of the molecule can be modeled using the Smoluchowski diffusion equation (equation (11) section III.5.2), which describes the probability of the molecule to move from position z_1 at time t to position z_2 at time t+Δt. For biased simulations, the diffusivity can be estimated on a flat energy landscape along the one-dimensional reaction coordinate. In this case, the molecule can diffuse freely without the interference of a deterministic force F(z) (i.e., the effective free energy surface is flattened, $w_{eff}(z) = 0$, as in the case of ABF method).[270]

The trajectories obtained from MD simulations can be discretized into a series of transitions over a time step Δt. The likelihood P[Z(t)|D(z)], which is the probability of the observed trajectory Z(t) given the initially guessed D(z), can be computed as the product of the probabilities of each observed transitions of the molecule from $Z(t_j)$ to $(Z(t_j,+\Delta t)$ over time interval Δt.[270]

$$P[Z(t)|D(z)] = \prod_j p\left[Z\left(t_j + \Delta t\right)|Z\left(t_j\right), D(z)\right] \tag{52}$$

As suggested by Chipot and Comer,[270] the probability of transition can be calculated numerically by solving equation (11), see section III.5.2 using the Crank-Nicolson method.[290]

Using Bayes' theorem, the desired posterior probability $P[D(z)|Z(t)]$ of the unknown position-dependent diffusion coefficient given the observed trajectory can be obtained from the likelihood $P[Z(t)|D(z)]$:

$$P[D(z)|Z(t)] = p_{prior}[D(z)] \times P[Z(t)|D(z)] \tag{53}$$

Where $p_{prior}[D(z)]$ is the prior distribution of the diffusion coefficient. The prior can be set uniform; furthermore, it can be defined as the product of priors assuming scale invariance and smoothness of the diffusivity. The position dependent diffusivity between neighboring intervals of the reaction coordinate is usually smooth. This prior knowledge can be imposed to obtain a smooth diffusivity by setting the prior ($p_{prior}[D(z)]$) as a weighting function that applies a harmonic penalty on the deviations of D(z) ($|D(z_i) - D(z_{i+1})|$):[237,270]

$$p_{smooth}\,[D(z)] = \prod_i e^{-\frac{[D(z_i) - D(z_{i+1})]^2}{2h^2\gamma^2}} \tag{54}$$

And

$$p_{scale}\,[D(z)] = \prod_i \frac{1}{D(z_i)} \tag{55}$$

Here, the smaller γ, the smoother the diffusivity profile. Then, the posterior probability is maximized using a Monte Carlo-based procedure through several iterations.[289,291]

Using trajectory data from independent MD runs or from the same trajectories of biased MD simulations (e.g., ABF[70,270], AWH[292]), the Bayesian inference scheme can be performed to estimate the diffusivity profiles using the open-source DiffusionFusion tool.[270,289,293]

However, it has been shown that for membrane permeation, the estimated diffusivity profiles for small alcohols largely depend on the lag time Δt used in the Bayesian inference, in particular in the membrane core (Figure 24A).[270] In this case, the longest feasible lag-time (e.g., Δt = 64 ps) can be used to compute the diffusivity profile that will later be used to obtain the permeation coefficient.

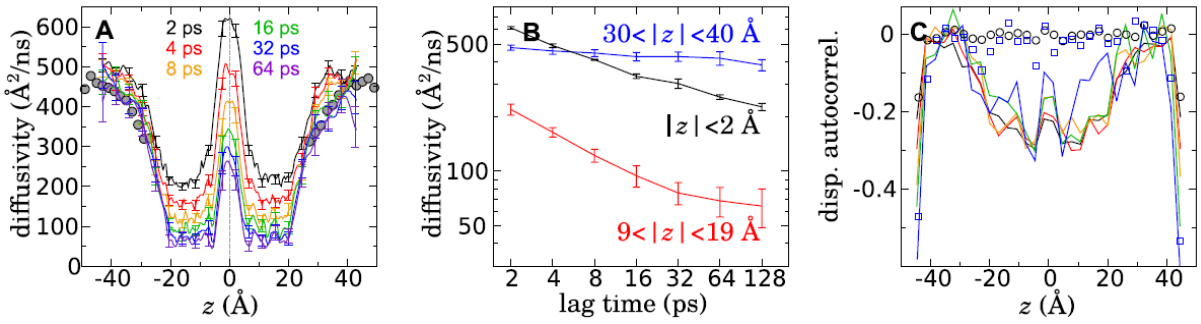


Figure 24: (A) example of position-dependent diffusivity (D(z)) profile as a function of lag-time (Δt) , using Bayesian analysis, of methanol crossing a POPC lipid bilayer membrane. The grey filled circles reflect the diffusivity as calculated using MSD method from a simulation containing 5% of methanol in aqueous solution. (B) Variation of the mean diffusivity as a function of lag-time within two different regions of the membrane (|z| < 2 Å and 9 < |z| < 19 Å from the membrane center) and one region in the aqueous medium (30 < |z| < 40 Å from the membrane center). (C) Normalized correlations of consecutive displacements of methanol over z in a flat energy landscape for different lag-times. The symbols reflect the same correlation in an ideal Brownian dynamics trajectory of equal length using the position-dependent diffusivity from A with Δt = 32 ps.†††††

IV.3.4.2. Fractional diffusivity: subdiffusion

The diffusive motion for many biological systems was known to exhibit a subdiffusive behavior, where the mean square displacement as a function of lag time Δt, obeys a power law, with a factor α, with $0 < \alpha < 1$; ($\langle x^2 \rangle \sim t^{\alpha}$).[294–297] The lateral diffusion of membrane phospholipids was also shown to obey a subdiffusive motion on time-scales spanning between 100 fs and 10 ns.[298,299] In the context of passive permeation of molecules, Chipot and Comer[270] reported the evidence of a subdiffusive behavior for small molecules that are crossing lipid bilayers. Accordingly, they developed a convenient framework to assess subdiffusion in membrane

††††† Gathered from [270]

permeation of small molecules along the one-dimensional reaction coordinate (i.e., the membrane normal) using MD simulations.[270]

In water, methanol molecules exhibited a classical diffusion described by the linear fit (r=0.99) of the MSD along z ($\langle \Delta Z^2 \rangle \sim t$), but also tested on x and y planes, over a range of lag-time values (Figure 25A). The spatial probability distribution of the molecule exhibits approximately a Gaussian form, a hallmark of a classical diffusion.

The diffusivity calculated using the MSD method was in agreement with the results obtained using Bayesian scheme for the range |z| > 30 Å (grey filled circles in Figure 24A).

Conversely, in the hydrophobic membrane medium (i.e., 9 < z < 19 Å) and near the center of the membrane, the MSD is nonlinear and it follows a power law, written as $\langle Z^2 \rangle \sim t^{\alpha}$ where the exponent α = 0.65 and 0.76, respectively. An example of the fit between MSD on the x-axis and lag-time is depicted in Figure 25A-B. This is a strong indicator of an anomalous diffusion compared to diffusion in water. In addition, the spatial probability distribution is not Gaussian anymore (i.e., cusp-like shape, in Figure 2 C), reflecting that the molecule spends more time at the initial position before moving out, which is a typical behavior in a subdiffusive regime.[300]

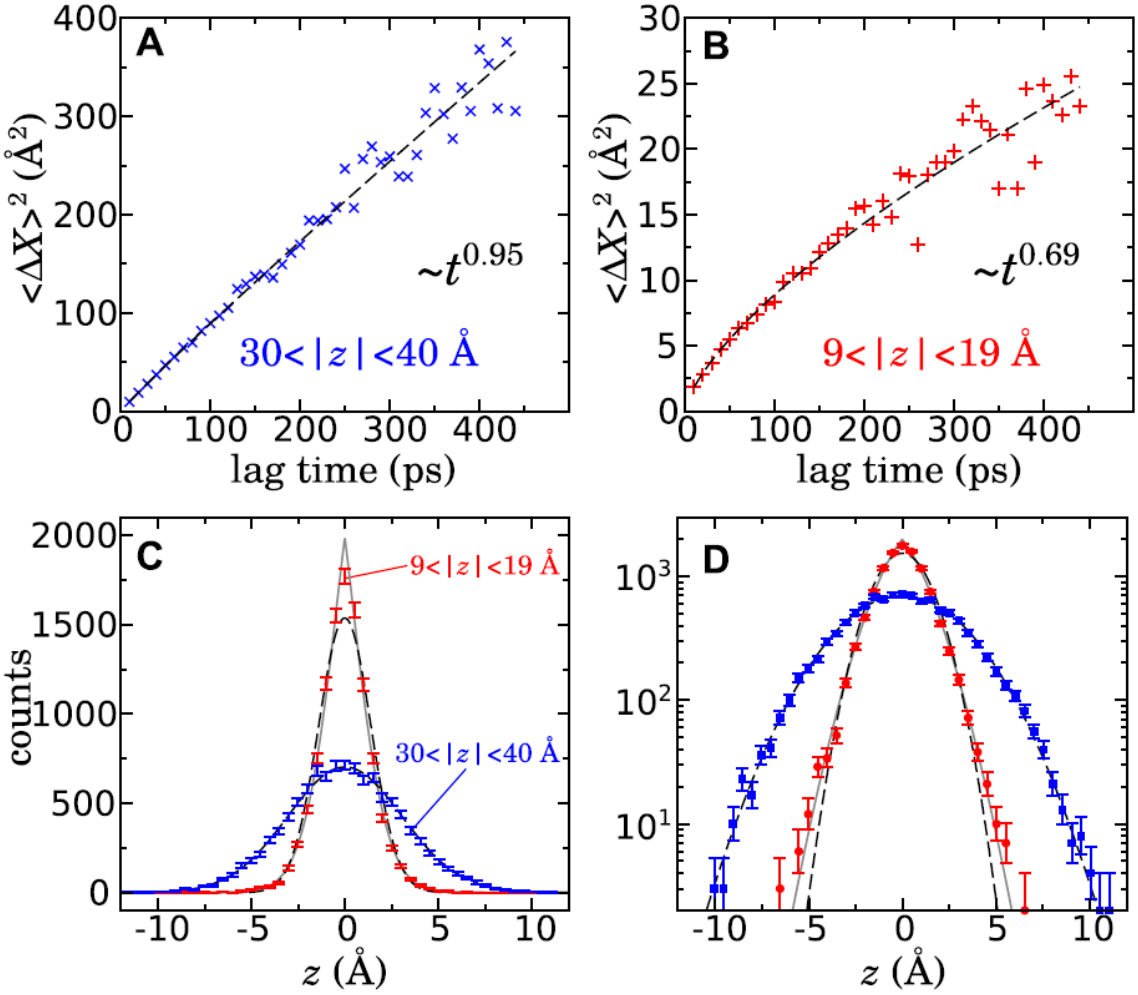


Figure 25: Example of variation of the MSD along the x-axis of methanol as a function of lag-time within two different regions of z; in (A) the aqueous medium and (B) the membrane core, with power-

law fits to the data (dashed curves). (C) The probability of methanol displacements for Δt = 10 ps and (D) the same histogram on a logarithmic scale. ******

Another evidence of subdiffusion across lipid bilayer membranes is shown by calculating the mean correlation between consecutive displacements over $2\Delta t$ intervals. This highlighted a negative correlation within the membrane suggesting that the displacements are anti-correlated, reflecting a subdiffusive behavior as opposed to a Markovian behavior in the aqueous phase (Figure 24C). The origin of this subdiffusive process can be assigned to the formation of transient empty volumes within the membrane where the molecule may spend more time before a next jump occurs.[270,300]

Similar to classical diffusivity, the Bayesian scheme can be applied to a MD trajectory Z(t) to estimate the position-dependent fractional diffusivity of fractional order α, $K_\alpha(z)$ (instead of D(z)). Using the Bayesian scheme, the position-dependent fractional order α(z) and fractional diffusivity $K_\alpha(z)$ can be optimized simultaneously. The probability of the observed Z(t) is given by:

$$P[Z(t)|K_\alpha(z),\alpha(z)] = \prod_j P\left[Z\left(t_j+\Delta t\right)|Z\left(t_j\right),K_\alpha(z),\alpha(z)\right] \quad (56)$$

Which is the product of all the observed transitions of the molecule from $Z(t_j)$ to $(Z(t_j,+\Delta t)$ over time interval Δt. Each of the observed transitions $P\left[Z\left(t_j+\Delta t\right)|Z\left(t_j\right),K_\alpha(z),\alpha(z)\right]$ can be determined using the Crank−Nicolson finite difference method.[301] This is used to numerically solve the fractional Smoluchowski equation with the fractional order α given by the following equation:[270]

$${}_t\mathrm{D}_*^{\alpha(z)}P(z,t) = \frac{\partial^\alpha P(z,t)}{\partial t^\alpha} = \frac{\partial}{\partial z}\left[K_\alpha(z)\frac{\partial}{\partial z} - \beta K_\alpha(z)F(z,t)\right]P(z,t) \quad (57)$$

Then, the posterior probability can be maximized using a Monte Carlo procedure

$$P[K_\alpha(z),\alpha(z)|Z(t)] = p_{prior}[K_\alpha(z)] \times P[Z(t)|K_\alpha(z),\alpha(z)] \quad (58)$$

The same rule can be followed to smooth out the resulting fractional diffusivity profiles. Chipot and Comer[270] showed that the resulting α(z) values from the Bayesian scheme were in agreement with the observations cited above, namely in the aqueous medium, α(z) approaches 1 suggesting a classical diffusion, while α(z) was around 0.7 within the membrane. This reflects regions with negative correlation (Figure 24C). Moreover, using the fractional model for several lag-times ranging from 2 to 32 ps, yielded fractional diffusivity profiles with similar values that are independent from the lag-time value (Figure 26B).

****** Gathered from [270]

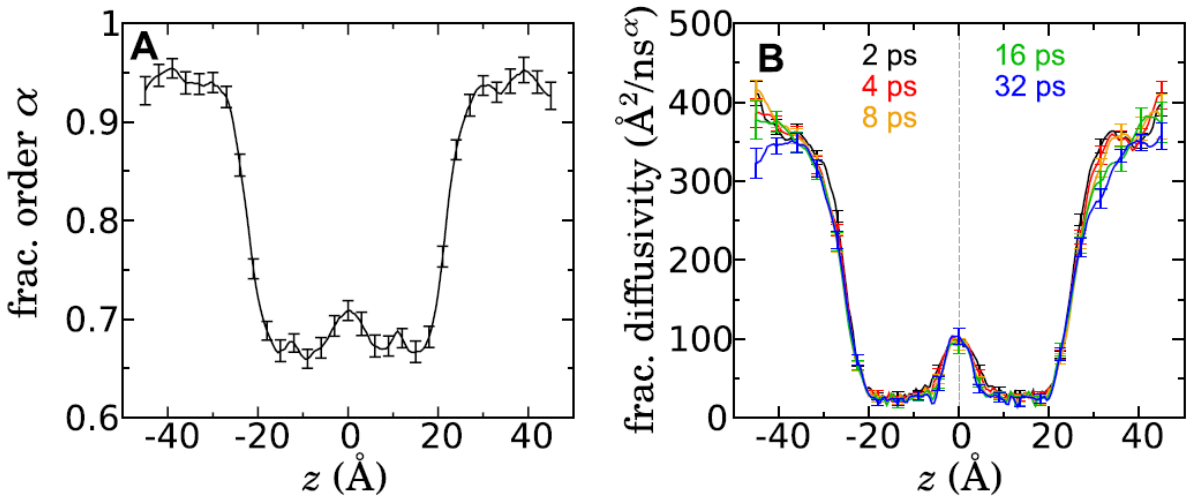


Figure 26: (A) Example of estimated fractional order α values over z, using the fractional Smoluchowski model and Bayesian analysis. (B) Estimated position-dependent fractional diffusivity for different lag-times using the Bayesian scheme. *******

IV.4. Permeability from methods using implicit solvent models

Implicit solvent models have been used to predict thermodynamic and physiological properties with a time-effectiveness, as the total number of atoms in the simulation is greatly reduced.[302–304] One of the important assumption of these methods is that the surrounding solvent (i.e., water, organic solvents) as well as the lipid bilayer[304,305] can replaced by a continuum characterized by its dielectric constants ε. In this formalism, a solute (molecule of interest) is defined as a molecular volume (vdW volume) (Figure 27) and the corresponding solvent accessible surface (SAS), on which the charge density σ is applied.

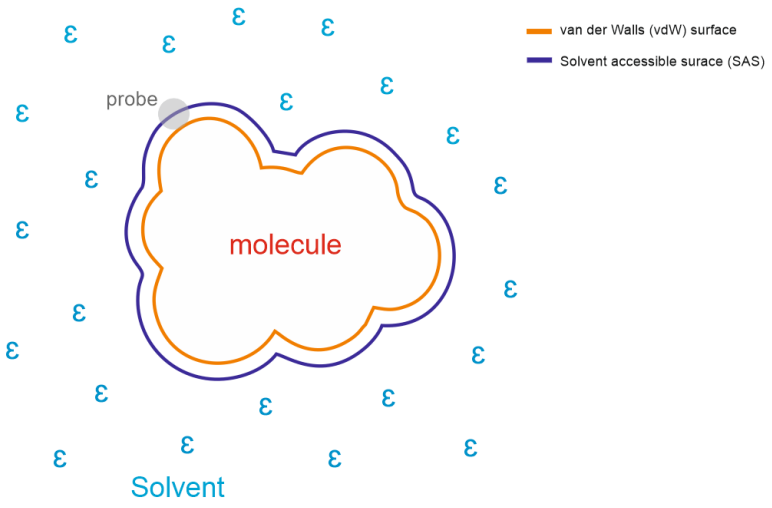


Figure 27: Illustration of the molecular vdW surface (orange) that is calculated from vdW radii, and the solvent accessible surface (blue) defined by the center of the probe (grey circle; it reflects an ideal

******* Gathered from [298]

solvent molecule). The solvent molecules are not represented explicitly, but implicitly by the dielectric constant of the solvent.

In these models, all information required about the conformation and the charge distribution of the molecule can be gathered from quantum calculations (e.g., density functional theory calculations).[306] In COSMO (conductor-like screening model)[307], SAS is divided in tesserae, for which the surface charge q* can be calculated and assigned. For solvents with finite ε values, the charge q is approximated to;

$$q = f(\varepsilon)q \tag{59}$$

Where f is a lowering factor equal to

$$f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon + 0.5} \tag{60}$$

Consequently, by knowing q* and q, the electrostatic interaction between the molecule and the solvent can be determined.

COSMO-RS (COSMO for Real Solvents)[308–310] is a COSMO-based approach to predict chemical potentials μ in solvents. COSMO-RS combines the dielectric continuum solvation model and statistical thermodynamics.[306] It considers the polarization surface charge density σ from DFT/COSMO[307]. Indeed, the molecular system can be considered as a charged surface interacting with a continuum solvent. Such interactions can picture both electrostatic and hydrogen-bonding interactions.[306] The polarization charge density σ on the molecular surface can be parametrized to calculate the polarization free energy:

$$E_{es}(\sigma, \sigma') = \frac{\alpha'(\sigma+\sigma')^2}{2} \tag{61}$$

And

$$E_{hbond}(\sigma, \sigma') = c_{hbond} \min\{0, \sigma\sigma' + \sigma^2{}_{hbond}\} \tag{62}$$

where σ and σ' are the counteracting surface densities and a', c_{hbond} and σ_{hbond} are adjustable parameters.[306] Then, a statistical thermodynamic procedure can be used. First the σ potential μ(σ) is equal to:

$$\mu_S(\sigma) = -\frac{RT}{a_{eff}} \ln\left[\int p_S(\sigma')\, e^{\frac{a_{eff}}{RT}[\mu_S(\sigma') - E(\sigma,\sigma')]}\, d\sigma'\right] \tag{63}$$

Where a_{eff} is an average molecular contact area, $p_S(\sigma)$ is the distribution function of the ensemble S (i.e., the system may consist of a mixture of compounds X) and $E(\sigma,\sigma')$ is the sum of electrostatic and hydrogen-bonding contribution energies (equations 61 and 62). Next, $\mu_S(\sigma)$ can be solved iteratively as it appears in both sides of equation (63). Finally, the chemical potential of compound X (μ^X) can be calculated by integrating the σ potential $\mu_S(\sigma)$ over the molecular surface of each compound X:

$$\mu_S^X = \int p^X(\sigma)\, \mu_S(\sigma) d\sigma + \mu_{combS}^X \tag{64}$$

Here, $p^X(\sigma)$ is a distribution function of σ and μ_{combS} is a combinatorial contribution that considers shape effects and size of the solute and the solvent.[306,311] The resulting chemical potential allows the assessment for other thermodynamic properties such as solvation free energy and partition coefficients.

In the context of passive membrane permeation, a convenient method (COSMOperm)[312] that is based on the COSMO-RS model, was developed to predict passive membrane permeability in an lipid bilayer membranes for neutral and ionized molecules. COSMOperm relies on calculating the free energy profiles ΔG(z) along the z-axis of the lipid bilayer membrane (i.e., the membrane normal), using another COSMO extension method (COSMOmic)[304], as well as the diffusivity profiles D(z) using COSMO-RS parameters. COSMOmic is specific to inhomogeneous systems such as lipid bilayer membranes. The virtual membrane is seen as a layered liquid system with different composition per layer (i.e., different σ-profiles). To predict an accurate layer-dependent free energy and a membrane-water partition coefficient, all orientations and conformations of the molecule of interest are sampled.

Similar to MD-simulations, COSMOperm is based on the ISDM theory to compute passive membrane coefficient P. Moreover, COSMOperm prediction of passive membrane permeation coefficients showed a good correlation with experiments for a large set of neutral and ionized molecules.[312]

Another recent implicit solvent-based method (PerMM) was developed to predict the passive membrane permeation coefficients. The PeMM method is based on the PPM model (positioning of proteins in membranes)[313,314] to predict free energy of transfer from water to the lipid bilayer of membrane proteins. This was developed to optimize their rotational and translational positions.[314] In the PPM theory, the hydrocarbon core of the membrane is considered as a planar slab with interfacial polarity profiles. The transfer free energy of a molecule from water to positions z, along the membrane bilayer normal, ΔGtransfer(z), is the sum of short-range terms, long-range terms and deionization energy of ionizable groups.[22,315] The short-range terms, related to the solvent-accessible area (ASA) depict vdW and hydrogen-bonding interactions, as well as hydrophobic effects. The long-range terms depict dipole-dipole interactions, and polarization terms constitute the electrostatic contribution of the dipole moment μj. The transfer free energy (ΔGtransfer(z)) is given by the following formula:

$$\Delta G_{transfer}(d,\varphi,\tau) = \sum_{i=1}^{N} \sigma_i^{wat\rightarrow memb}(z_i)\, ASA_i + \sum_{j=1}^{M} \eta_j^{wat-memb}(z_j)\mu_j + \sum_{k=1}^{L} \min\left\{\Delta E_k^{ion},\, \Delta E_k^{neut}\right\} \quad (65)$$

Where d, j, and τ are spatial positions of the molecule with respect to the lipid bilayer, $\sigma^{wat\rightarrow memb}$ is the atomic solvation parameter that describes the surface transfer energy (per Å^2) of atom i from water to the point z_i, ASA is the solvent-accessible area for atom i, N is the number of atoms of the molecule, $\eta^{wat\rightarrow memb}(z_j)$ is the dipolar solvation parameter that describes the penalty energy of transferring the dipole moment of 1D from water to the point z_j, μ_j is the dipole moment of group j, M is the number of group dipoles, E^{neut} and E^{ioniz} are transfer energies of the ionizable group k in the neutral and ionized states, respectively. The solvation parameters σ_i depends on the position-dependent polarity parameters of the lipid bilayer and can be estimated according to the universal solvation model.[22,315,316] Indeed, the partition coefficient K(z) equals:

$$K(z) = e^{-\beta \Delta G_{transfer}(z)} \quad (66)$$

In the PerMM method, the diffusivity coefficient D is related to ASA as follows:

$$D_i = \frac{kD_0(\eta)}{ASA_i^n} \tag{67}$$

Where D_0 is a constant for a particular membrane type with micro-viscosity n and the subscript n is ~2/3. Correspondingly, the intrinsic permeation coefficient equals:

$$logP_{\sum i} = -\log\left(ASA_i^n \int_{-\frac{L}{2}}^{\frac{L}{2}} \frac{dz}{K_i(z)}\right) \tag{68}$$

Using the PerMM method, the calculated permeation coefficients of many molecules, in a DOPC bilayer, showed a good correlation with experiments, in particular with black lipid membrane (BLM) models. The correlations with PAMPA-DS and cell-based membrane models, Caco-2 and Blood Brain Barrier (BBB), exhibited lower correlation. This might be explained by the difference in membrane compositions and contribution of other transport mechanisms.[22]

Other implicit models were used to estimate solute partitioning into lipid bilayer considering either isotropic or anisotropic properties.[313,317] Other physics-based methods were also developed to predict passive membrane permeation.[318–322] Although these implicit solvent models are useful for simulations of bimolecular systems and in the prediction of membrane permeation properties, with a remarkable gain of computational time, they still miss some detailed atomic descriptions that are critical to fully capture the underlying molecular processes, e.g., arrangement of water molecules around the permeating molecule or specific interactions or impact of the molecule within the lipid bilayer system.

Chapter V. MemCross: Accelerated Weight Histogram method to assess membrane permeability

Mehdi Benmameri,[†] Benjamin Chantemargue, [‡] Antoine Humeau, [†] Patrick Trouillas, [†¶] and Gabin Fabre*[†]

[†]INSERM, UMR 1248, F-87000 Limoges, France;

[‡]INSILIBIO, 1 avenue Ester Technopôle, 87069 Limoges Cedex, France;

[¶]CATRIN RCPTM, 779 00 Olomouc - Holice, Czech Republic

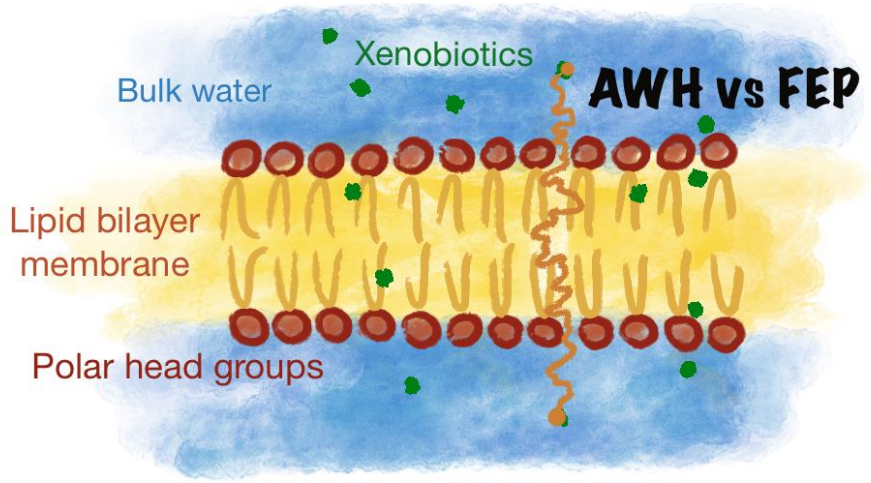
doi : 10.26434/chemrxiv-2022-rk82h

Foreword

In this study, we describe our developed methodology (MemCross) to compute passive permeation coefficients through a POPC lipid bilayer. Accordingly, we used the AWH method to model the PMF and Diffusivity profiles. Subdiffusion and ionization state of ionizable molecules were taken into account. This article was submitted in chemrxiv.org (doi : 10.26434/chemrxiv-2022-rk82h) and will be submitted in a journal in the near future.

V.1. ABSTRACT

Passive permeation events across biological membranes are determining steps in the pharmacokinetics of xenobiotics. To reach an accurate and rapid prediction of membrane permeation coefficients of drugs is a complex challenge, which can efficiently support drug discovery. Such predictions are indeed highly valuable as they may guide the selection of potential leads with optimum bioavailabilities prior to synthesis. Theoretical models exist to predict these coefficients. Many of them are based on molecular dynamics (MD) simulations, which allow calculation of permeation coefficients through the evaluation of both the potential of mean force (PMF) and the diffusivity profiles. However, these simulations still require intensive computational efforts, and novel methodologies should be developed and benchmarked. Free energy perturbation (FEP) method was recently shown to estimate PMF with a significantly reduced computational cost compared to the adaptive biasing force method. This benchmarking was achieved with small molecules, namely short-chain alcohols. Here, we show that to estimate the PMF of bulkier, drug-like xenobiotics, conformational sampling is a critical issue. To reach a sufficient sampling with FEP calculations requires a relatively long time-scale, which can lower the benefits related to the computational gain. In the present work, the Accelerated Weight Histogram (AWH) method was employed for the first time in all-atom membrane models. The AWH-based protocol, named MemCross, appears affordable to estimate PMF profiles of a series of drug-like xenobiotics, compared to other enhanced sampling methods. The continuous exploration of the crossing pathway by MemCross also allows modelling subdiffusion by computing fractional diffusivity profiles. The method is also versatile as its input parameters are largely insensitive to the molecule properties. It also ensures a detailed description of the molecule orientations along the permeation pathway, picturing all intermolecular interactions at an atomic resolution. Here, MemCross was applied on a series of 12 xenobiotics, including four weak acids, and a coherent structure-activity relationship was established.



Keywords: Permeation, AWH, FEP, Membrane, Protonation state, PMF, Fractional diffusivity.

V.2. INTRODUCTION

Most xenobiotics (*i.e.*, drugs, vitamins, toxics, nutrients, which can be natural, bio-transformed or synthetic compounds) are crossing many biological membranes before reaching their targets. The crossing events can occur either through membrane proteins/channels or by passive permeation through the lipid bilayer, both processes being possibly in competition. The latter process is a spontaneous thermodynamic process, which is called "passive" as it does not require any additional energy but the intermolecular interactions between the xenobiotics and its environment. Virtually, all xenobiotics can cross passively, however the kinetics of the process is key and can take a few microseconds up to centuries. To rationalize this membrane permeability process and to estimate its timescale is of particular interest in the development of new drugs or improvement of existing treatments, as it is directly related to ADME/Tox.[323,324,18] A fast and accurate prediction of permeation coefficients (P_{erm}) is a major issue in pharmaceutical industries. One of the major reasons for clinical trial failure is the lack of efficacy, part of which can be attributed to poor bioavailability.[325] A variety of experimental techniques have been developed to estimate P_{erm} values.[23,42] The most common techniques rely on cell-based lipid membranes (Caco-2, MDCK),[326] planar bilayers,[327] PAMPA[328] or liposomes,[195] by creating a concentration gradient between the inner and the outer compartments of the biomimetic or biological membranes, and measuring the xenobiotic transfers. However, such experimental assessments do not easily allow large screening on large drug databases, and they cannot be used to predict permeability of drugs not yet synthetized (*e.g.*, prediction of the impact of chemical modifications from a collection of hit compounds).[42]

As an alternative to experimental techniques, many computational approaches have been developed to predict P_{erm} values.[329–332] Among them, physics-based methods have been developed to rapidly estimate the partition and the permeation of molecules.[115,333,22] Although these methods allow screening on large databases, they cannot capture the atomic resolution of the permeation process to fully describe interplays between xenobiotics and membrane constituents. In this respect, molecular modeling methods have also been used to follow the passive permeation process, which can be seen as a 3-step mechanism, namely, membrane insertion, flip-flop from one to another leaflet, and resolvation into the second water compartment. Because the timescale of permeation events is often beyond the reach of free molecular dynamics (MD) simulations, there is a need of using biased MD methods. Among others, Umbrella sampling[243], Adaptive Biasing Force method (ABF)[269,334], Free Energy Perturbation (FEP)[70], or Metadynamics[335,258] have been tested to quantify drug P_{erm} values. They provide an atomic-scale description of the positioning and orientation of molecules along the permeation path. The main issue of these methods is that they still require highly demanding computational resources, and the setup of some of them often requires delicate parametrizations to reach accuracy. As it is a highly statistically-driven process, Kinetic Markov State models can also be used to assess P_{erm} values.[336–338] Although they represent the most precise *ab initio* technique, their high computational cost still prohibits large scale screening.

Taking the chemical variability and inhomogeneity of lipid bilayer structure into account, a quantitative theory based on inhomogeneous solubility-diffusion model (ISDM) and consistent with MD simulations was developed to assess passive permeation[20,128,339]. However, the ISDM formalism does not account for the subdiffusion of xenobiotics through the lipid bilayer membrane. Chipot and Comer, demonstrated that subdiffusion exists for small alcohol molecules,[270] and that it is necessary to take it into account for a more accurate prediction of P_{erm}.[340] Contrary to the classical diffusion, subdiffusion means that the mean squared

displacements of a molecule inside the lipid bilayer is not proportional to time, but it obeys a power-law fit, namely $\langle x^2 \rangle \sim t^{\alpha}$ with $0 < \alpha < 1$. The factor α is a key variable to estimate the motion of the molecules along the permeation path. Although the methodology developed in [270] provided an accurate estimation on prototypical compounds, the need of performing multi-μs long ABF simulations hampers its use as a high-throughput screening tool. More recently, the same authors proposed a faster method to obtain the free energy profile w(z) of short chain alcohols across membranes, using FEP calculations, thus reducing the computational cost by one order of magnitude.[70]

Here, we computed P_{erm} values of a set of 12 xenobiotics (quercetin, resveratrol, curcumin, rutin, viniferin, caffeic acid, coumarin, testosterone, ceefourin1/2, MK571, methotrexate) bulkier than the compounds used in [70] (Figure S1). These xenobiotics exhibit a wide variety of chemical scaffolds, and are of pharmacological relevance, as most of them serve as modulators or substrates of crucial membrane proteins such as human ABC membrane proteins.[341–345] Ethanol was added as reference compound to compare to existing data from the literature. The potential of the mean force (PMF) profiles through a prototypical lipid bilayer membrane was obtained for a subset of four compounds (ethanol, quercetin, resveratrol and curcumin) to choose the most appropriate method between Free energy perturbation (FEP)[346] and Accelerated Weight Histogram (AWH) methods[247,278,347,348]. To the best of our knowledge, this study proposes the first use of the AWH method on xenobiotic-membrane interaction simulations with an all-atom model. AWH was recently successfully applied to the determination of the permeabilities of small compounds with a coarse-grain model.[349] Then fractional diffusivity profiles (*i.e.*, taking subdiffusion into account) were computed by using the AWH trajectories. Finally, using this AWH-based protocol coined MemCross, P_{erm} values were calculated for the whole series of compounds providing structure-property comparisons.

V.3. METHODS

V.3.1. Theoretical background

ISDM takes the normal to the membrane surface (named z-axis) as the reaction coordinate along which the molecule undergoes a random walk. At steady state, the P_{erm} value is the inverse of integral of the resistance R(z) to permeation events along the z-axis: [18,270]

$$P_{erm}^{-1} = \int_{-L/2}^{L/2} R\,(z)\,dz = \int_{-L/2}^{L/2} \frac{e^{\beta w(z)}}{D(z)}\,dz \qquad (1)$$

Where $\beta = 1/k_bT$, k_b is the Boltzmann constant and T the temperature.

As seen in **equation (1)**, P_{erm} depends on both the potential of the mean force *w*(z) picturing its translocation from an initial (aqueous) to a final (aqueous) compartment through the membrane, and the diffusivity *D*(z) along the z-axis. Uncoupled methodological effort can thus be given to calculate the two profiles separately, and then to merge them again to obtain P_{erm} values.

V.3.2. Computational setup

The lipid bilayer system, containing 128 molecules of 1-palmitoyl-2-oleyl–sn-glycero-3-phosphocholine (POPC), was built using CHARMM-GUI membrane builder.[350] To ensure reliable molecular organization and membrane properties, the system was equilibrated for 500 ns. The dimensions of the simulation box were 66 Å × 66 Å × 90 Å. A pure POPC membrane model was chosen here as a reasonable model of mammal cell membranes, which contain a majority of phospholipids, POPC being one of the major representative lipids.[70]

All xenobiotics (Figure S1) were parametrized using GAFF2 (second generation generalized amber force field).[351] Particular attention was given to all dihedral angles, but only the dihedral between the aromatic rings of quercetin, rutin and ceefourin 2 needed to be reparametrized as described in the literature to ensure a correct π-conjugation of the systems.[352] Different conformers were then generated using the Confab software[353] with 100 kcal/mol as the energy cutoff, and with a limit of 10 conformers. Afterwards, geometry optimization was carried out on all conformers with the Gaussian09 software[354] using B3LYP/cc-pVDZ in an implicit diethylether solvent. The atomic RESP partial charges were derived from Gaussian calculations using R.E.D.III software and Duan *et al.* method.[355]

All MD simulations were performed using Gromacs2020.3[241] and the Slipids 2016 force field[53] for POPC molecules, GAFF2 for xenobiotics and TIP3P for water molecules. Temperature and pressure were maintained at 310 K and 1 atm, respectively. Bonds involving hydrogen atoms were constrained by applying the LINCS algorithm[356]. The Particle-Mesh-Ewald method was used to evaluate the long range electrostatic forces[357]. The simulation time step was set at 2 fs.

V.3.3. Alchemical free energy perturbation (FEP)

Alchemical FEP calculations were performed to determine the one-dimensional PMF profile along the reaction coordinate of the permeation process for ethanol, quercetin, resveratrol and curcumin. To estimate the complete PMF profile, the free energy of binding was calculated at six different depths of one layer of the membrane, namely at z = 0, 4, 13, 22, 30 and 40 Å, 0 being the center of the lipid bilayer (see Figure S2**)**.[70] In order to maximize sampling, the orientation θ of the molecule was set and monitored during equilibration simulations. θ was set as the angle between the z-axis and the longest vector of each molecule, as shown in Figure S3A. Three different initial θ-values (0°, 90° and 180°) of each molecule were subjected to equilibration at each depth, except in water, *i.e.*, at 30 and 40 Å, as in this region sampling reorientations of the molecules is very fast. Inspection of the orientation of quercetin over time revealed that when convergence of θ was reached (Figure S3B), the standard deviation of the θ was lower than 20°. Therefore, for all molecules, an equilibration simulation was considered converged if the standard deviation of θ was less than 20° during the last 10 ns of equilibration (Figure S3B and Table S1). The equilibration simulations lasted from 10 to 100 ns, with an average of 35 ns per window. The geometries resulting from the equilibration process were taken as initial coordinates to run the FEP calculations.

Once the orientations were converged at a given depth, FEP calculations were performed. During these alchemical calculations, the molecule is progressively decoupled from its environment (molecule vanishing) using the coupling parameter λ. Here we performed 16 simulation windows, corresponding to 16 λ-values (2 at the edges of the process, *i.e.*, both the

molecule fully present in its environment and the molecule fully vanished, respectively and 14 intermediate windows). A restraint was applied to the distance between the molecule and the center of mass of the membrane using a harmonic potential with a force constant of 50000 $kJ.mol^{-1}.nm^{-2}$. The restraint was applied along the z-axis letting the molecule freely explore the *xy* plane at a given depth. The simulation time per λ window was about 10 ns for ethanol, as it was sufficient for the three replicas to converge. For bulkier molecules, this time was increased to 20 ns for quercetin 25 ns for resveratrol and curcumin. As discussed in the Results section, this increased simulation time was not sufficient to obtain convergence between the 3 replicas for resveratrol and curcumin. As the computational effort was too high compared to the AWH method, the FEP calculations were not extended further.

The decoupling was done separately for both electrostatic and van der Waals interactions. The decoupling of electrostatic interactions was performed using replica exchange over five λ-values ranging from 0 to 1 with $\Delta\lambda = 0.25$. These simulations were carried out on CPUs as Gromacs 2020 does not allow the perturbation of charges on GPUs. The decoupling of van der Waals interactions were done using 11 λ-values ranging from 0 to 1 with $\Delta\lambda = 0.1$ on GPUs. The free energy of solvation of the molecule in a given environment was estimated using alchemlyb.pymbar library[358–360] and the Multistate Bennett Acceptance Ratio (MBAR)[361] method.

The adequate spacing between λ-values was ensured by calculating the overlap matrix for both electrostatics and van der Waals interaction decoupling and for all z-windows (Figure S4). The matrix was estimated using multistate Bennett acceptance ratio estimator (MBAR),[361] mapped over λ-states. Briefly, the overlap matrix carries the information about the overlap of the phase space probability distributions between the neighboring states where the probability of overlap fraction should exceed 3%.[358]

To obtain the complete PMF profile along the z-axis, the six free energy values, obtained at the six different depths, were interpolated using a cubic Hermite interpolator. The corresponding profile was symmetrized to build the full PMF profile, picturing the energy cost of molecule translocation from one to another water compartment through the membrane, *i.e.,* from z = -40 to 40 Å.

V.3.4. Accelerated Weight Histogram method (AWH)

The AWH method was also performed to compute the PMF along the z-axis. The one-dimensional reaction coordinate was chosen as the distance between the center of mass of both the xenobiotics and the membrane, projected on the z-axis. Similar to other adaptive biasing methods,[243,258,269,334] an history-dependent biasing potential was added to the system in an adaptive manner to flatten the free energy barriers along the reference reaction coordinate and to reach the chosen target distribution. The main advantage of the AWH method lies in its fast convergence, exponential during the initial stage, and asymptotic afterwards.[247] This allows rapid determination of PMF profiles with a very few input parameters, that are only very weakly dependent on the xenobiotics.[348]

Each xenobiotic was initially placed in the water bulk, at z ~ 40 Å. Here, the targeted distribution was chosen to be flat. The initial diffusion constant D_0 was set at 5.10^{-3} $nm^2.ps^{-1}$ and the error estimate (ε_0) was limited to 3 $kJ.mol^{-1}$. The force constant for the umbrella harmonic potential was chosen to be large enough (128000 $kJ.mol^{-1}.nm^{-2}$). The AWH simulations were run for a minimum of 1 µs. The time required to escape the initial stage was monitored, and for some

molecules, the simulations were prolonged up to 4 µs to ensure the convergence of the PMF profiles for the 3 replicas.

V.3.5. Position-dependent diffusivity

Subdiffusion was shown to occur for small molecules in membranes.[270] In brief, analyzing the mean square displacement along z- and x-axes of methanol in a molecular POPC membrane system using increased lag times exhibited a linear dependence on time in water. However, a power law dependence (t^{α} for $0 < \alpha < 1$) was highlighted inside the membrane.[270] A fractional-order Smoluchowski equation was developed by Chipot and Comer, in which the diffusion coefficient D(z) is replaced by the position-dependent fractional diffusivity profiles $K_{\alpha}(z)$. [270] In order to estimate the diffusivity within this more accurate theoretical framework, we estimated $K_{\alpha}(z)$ from the second half of the AWH trajectories, using a Bayesian inference scheme implemented in the open-source DiffusionFusion tool.[270,289] The z-dependent fractional order α-values were gathered from the α-values used for ethanol in a POPC bilayer.[270] Several lag times (Δt), ranging from 2 ps to 16 ps, were investigated as suggested in reference [270] to confirm the time-independency of the fractional diffusivity. Diffusions at each z-step, every 0.5 Å from 0 to 40 Å, were initialized with a value of 100 $nm^2.ps^{-1}$. The probability of the observed trajectory is computed as the product of the probabilities of the observed movements from $z(t_j)$ to $z(t_j + \Delta t)$, in other word the probabilities that the molecule at $z(t_j + \Delta t)$ under the condition that it was at $z(t_j)$:

$$P[z(t)|K_{\alpha}(z)] = \prod_j p\left[z(t_j + \Delta t)|z(t_j), K_{\alpha}(z)\right] \qquad (2)$$

The posterior probability is maximized using Monte Carlo procedure with 2000 cycles:

$$P[K_{\alpha}(z)|Z(t)] = p_{prior}[K_{\alpha}(z)] \times P[Z(t)|K_{\alpha}(z)] \qquad (3)$$

Smoothed fractional diffusivity can be obtained by multiplying the likelihood $P[Z(t)|K\alpha(z)]$ by a weighting function using a harmonic potential of grid space *h* and of parameter γ.[237] The smoothness of the diffusivity is given by:

$$p_{smooth}\,[K_{\alpha}(z)] = \prod_i e^{-[K_{\alpha}(z_i) - K_{\alpha}(z_{i+1})]^2 / 2h^2\gamma^2} \qquad (4)$$

V.3.6. Protonation state assessment

The protonation state and the total charge are major factors ruling lipid bilayer crossing and only neutral molecule are expected to cross membrane at physiological pH (pH-partition hypothesis). However and against this hypothesis, acidic compounds were seen to partition in membranes at physiological pH under their anionic forms.[192] Here, we considered two different charge states for the four acid molecules, namely caffeic acid, methotrexate, MK571 and ceefourin 2. The simulations were performed with *(i)* the mono-anionic form that is most probable ionized state in water at pH = 7.4, and *(ii)* the neutral form, as it is usually the most likely to diffuse through the membrane interior.[133] For such compounds, acid-base exchanges can occur at the polar head-group surface of the lipid bilayer or deeper,[133] thus impacting on the PMF profiles. To take this acid-base property into account, the free energy and fractional diffusivity profiles were calculated for both ionization forms and then they were combined to provide the most realistic profile [362]. To do so, the free energy profile of the protonated form was shifted by the free energy difference between those two forms ($\Delta G_{deprotonation}$) in water (at 40 Å from the membrane center) with respect to the free energy profiles of the ionized form.

To set the reference in water, the profile of the deprotonated form was shifted to 0 in water. $\Delta G_{deprotonation}$ was derived from the Henderson Hasselbalch equation:

$$\Delta G_{deprotonation} = 2.3\, k_b T\, (pKa - pH) \qquad (5)$$

Where k_b and T are the Boltzmann constant and temperature, respectively.

The combined profile was then obtained as the minimum energy of the two profiles all along the z-axis Figure S5A). The crossover point between the two profiles corresponded to the most probable depth ($z_{transition}$) for acid-base exchange and transition between the two forms. Knowing $z_{transition}$, a combined diffusivity profile was obtained by keeping the diffusivity value corresponding to the chosen protonation state (Figure S5B). The permeation coefficients of these acid molecules can be calculated using **equation (1)** and these combined free energy and diffusivity profiles.

V.3.7. Error estimation of the permeation coefficient

A Monte Carlo procedure was used to propagate the error estimation of the parameters w(z) and $K_\alpha(z)$ from the 3 replicas on the estimation of P_{erm} coefficient obtained using **equation (1).** This procedure generated 1000 values of $logP_{erm}$ from 1000 values of w(z) and $K_\alpha(z)$, leading to the mean permeation coefficient ($logP^{calc}_{erm}$) and its corresponding standard deviation ($\pm\, \sigma_{logP}$).

w(z) was generated from a normal distribution of parameters $\mu_{w(z)}$, $\sigma_{w(z)}$. Instead of using directly the estimates of $\mu_{w(z)}$ and $\sigma_{w(z)}$ from the 3 replicas, which could underestimate the real uncertainty of the parameters, $\mu_{w(z)}$ and $\sigma_{w(z)}$ are randomly generated for each simulation.

$\mu_{w(z)}$ is generated in a normal distribution of mean $\bar{x}$ (the PMF mean over 3 replicas) and standard deviation $\sigma_{\bar{x}} = \frac{s_{w(z)}}{\sqrt{n}}$, where $s_{w(z)}$ is the standard deviation of the PMF over 3 replicas, $\sigma_{\bar{x}}$ is the standard deviation of the mean and n number of replicas (i.e., n=3).

Similarly, $\sigma_{w(z)}$ was generated in a normal distribution with mean $\sigma_{\bar{x}}$ and standard deviation $SD(\sigma_{\bar{x}})$ given by the following equation;

$$SD(\sigma_{\bar{x}}) = \sigma_{\bar{x}} \cdot \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \cdot \sqrt{\frac{n-1}{2} - \left(\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)^2} \qquad (6)$$

Where Γ is the gamma function, and n the number of replicas. This equation arises from the distribution of the standard error of a sample in a normal distribution. Likewise, the same procedure was used to estimate $K_\alpha(z)$.

The MemCross protocol regroups the use of the AWH method, fractional diffusivities, the protonation state assessment, and the error estimation procedure.

V.4. RESULTS AND DISCUSSION

V.4.1. Decision making about the methodology to compute P_{erm}

V.4.1.1. Alchemical free energy perturbation (FEP)

The membrane crossing was decomposed into two steps corresponding to two different free energy terms read on the PMF (Figure 28A). The difference in free energy between the water bulk and the PMF minimum was named $\Delta G_{partition}$, as it corresponds to partitioning, *i.e.*, insertion of xenobiotics into the lipid bilayer. The opposite of $\Delta G_{partition}$ corresponds to the energy required for the molecule to exit from the lipid bilayer into water (resolvation). The second term ($\Delta G_{flip\text{-}flop}$) was the difference in free energy between the local PMF maximum near the lipid bilayer center, and the PMF minimum. This term corresponds to the energy required to cross the center of the membrane, in other words, the energy required for the molecule to flip-flop from one to another leaflet.

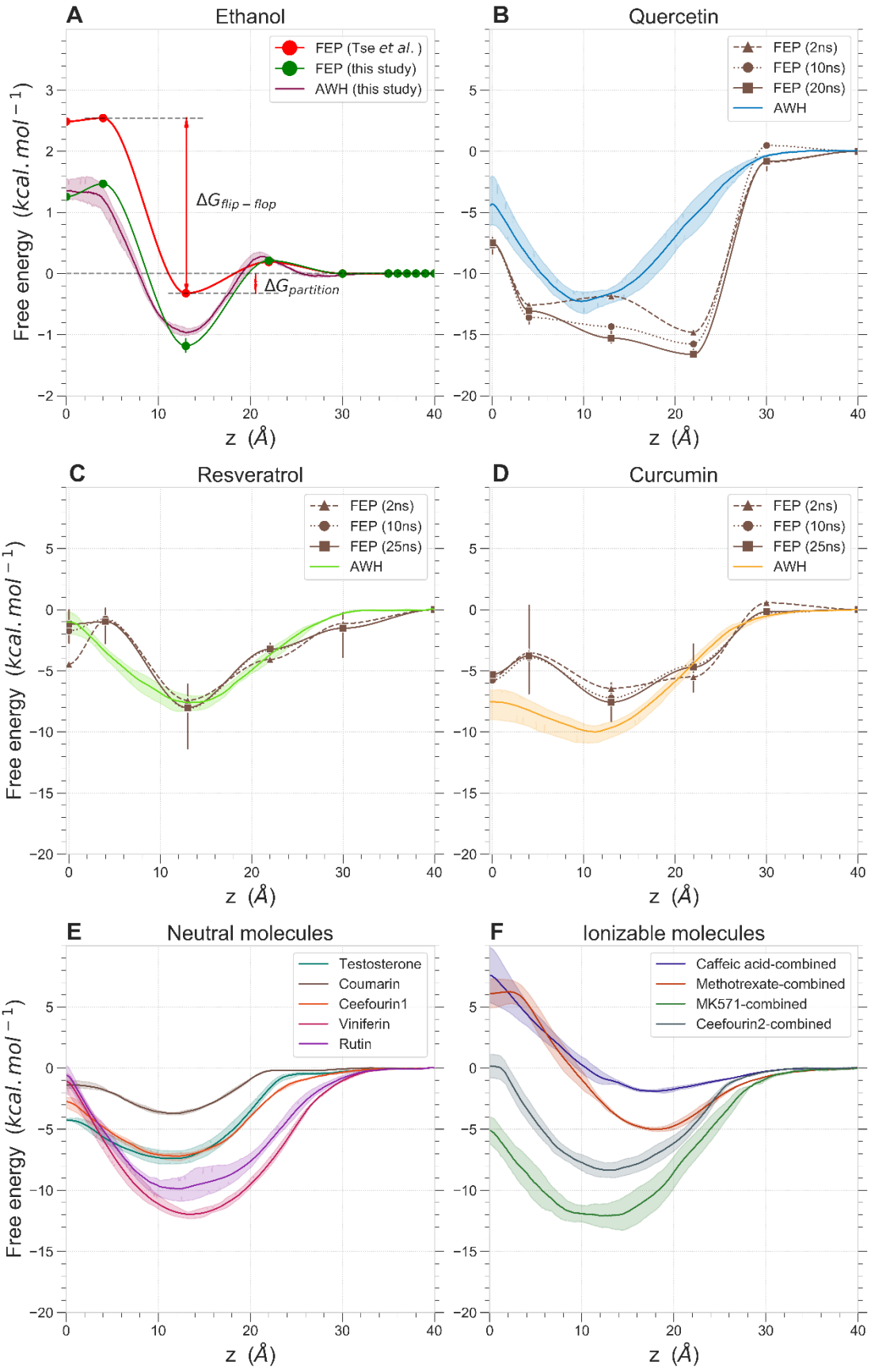


Figure 28. Calculated PMF profiles picturing the translocation of molecules from the water bulk (z=40 Å) to the center of the membrane (z=0 Å): A) for ethanol using alchemical FEP (red: gathered from [70], purple: the current study) and the AWH method (blue); B-D)

comparison between PMF profiles of quercetin, resveratrol and curcumin using FEP for different time scales per λ-value (2, 10 and 20-25 ns) and the AWH method; E) for a set of neutral molecules using the AWH method; F) combined PMF profiles for caffeic acid, methotrexate, MK571 and ceefourin 2. For the FEP method (Panels B to D), error bars are shown for t_{sim} = 25 ns.

The PMF profile of ethanol obtained here was similar to the one predicted by Tse *et al.*[70] (

A). Both profiles exhibited a small local maximum at around z = 22 Å (0.18 kcal.mol^{-1}). This is explained by the unfavorable local rearrangements of the lipid polar head groups in the presence of ethanol. The $\Delta G_{partition}$ value of ethanol (-1.04 kcal.mol^{-1}) calculated here was slightly lower than the one reported in the literature (-0.32 kcal.mol^{-1}, Figure S6A)[70]. The $\Delta G_{flip-flop}$ value was 2.35 kcal.mol^{-1}, very close to 2.81 kcal.mol^{-1} as obtained in [70]. Such slight differences in the free energy of binding below the polar head groups likely arises from the two different force fields used in both studies. The ΔG value at 13 Å and at the center of the membrane, obtained with the CHARMM force field by Tse *et al.*[70] are 0.7 and 1.0 kcal/mol higher, respectively, than those obtained here with Slipids and GAFF. This fully agrees with the benchmarking study performed with these forcefields for small molecules and membrane interactions.[363] Whatever the methodology used, the driving forces appeared adequately pictured for such small molecules, and FEP can provide reliable estimates of the ΔG profiles.

The PMF profiles were calculated with the same method for three test compounds with much larger molecular size than ethanol, namely quercetin, resveratrol, and curcumin (Figure 28B-D). To optimize the sampling time for these large compounds during the FEP calculations, the orientation of the molecules in the lipid bilayer, pictured by θ (see Methods section), was followed during equilibration simulations. The relaxation time of θ was in the order of tens of ns. This is comparable with the time of each FEP simulation (10-25 ns). Therefore, despite equilibration simulations, an exhaustive sampling of θ (hence of conformational relaxation of the molecule) is unlikely with FEP, for which the molecule is fully coupled to its environment.

For quercetin, the 3 replicas converged, as the free energy profile at 20 ns exhibited low standard deviations at each depth (Figure 28B). However, the free energy minimum was at z = 22 Å, *i.e.*, unexpectedly outside the membrane. This disagrees with a previous work[364], where the most favorable position of quercetin was just below the lipid head groups, due to its amphiphilic character, showing that 22 Å as a minimum is not realistic. For resveratrol and curcumin, FEP calculations yielded large standard deviations of several kcal/mol between replicas, even after 25 ns of simulation per λ, especially in the lipid tail region of the bilayer (Figure 28C-D).

Overall, to implement FEP calculations for molecules of such large size compared to ethanol, the calculation time was increased from 2 ns to 25 ns per λ window. This amounts to 6-7.5 µs of global simulation time per xenobiotic, which was even insufficient for a complete sampling. It must be stressed that alteration of the lipid bilayer properties in the presence of relatively large molecules are likely to occur,[365] which requires longer sampling to capture the entire process. This is intricate with the fact that the bias applied during these simulations could locally change the membrane properties leading to a misarrangement of the molecule, here in the polar head groups. Hidden barriers can arise, particularly in the latter region and may impact on the sampling in this region.[366] Without a sufficient sampling of conformational reorganization, a correct estimate of the free energies is unlikely. Based on these three prototypical compounds, *absolute* alchemical FEP simulations of such large molecules suffers

from a lack of sampling of intermolecular interactions, which could be solved by a drastic lengthening of the simulation time. Nevertheless, FEP is very efficient at estimating *relative* binding free energies. That is, for molecules A and B differing by a few atoms, the free energy of molecule B can be quickly calculated upon progressively adding or removing these few atoms to molecule A. In database screening, these relative FEP-derived free energy values could be useful for families of structurally related compounds, if the absolute free energy profile is known for at least one of them.

V.4.1.2. AWH method

AWH is a recent method that was used to accurately predict the PMF profile of a DNA base pair splitting, followed along a hydrogen-bond distance between the two residues as a reaction coordinate.[348] In the current study, we used the AWH method as an alternative to estimate the PMF profile picturing the permeation process.

To validate the effect of the input parameters on the sampling of the initial stage, we first computed the PMF profiles for ethanol using different values of key parameters of this method: D_0 and ε_0. D_0 can be roughly set within a few orders of magnitude of the expected diffusivity.[241] Only a weak effect on the PMF profile of ethanol when D_0 was set between 5.10^2 and 5.10^{-4} $nm^2.ps^{-1}$ (Figure S7). A value of 5.10^{-3} $nm^2.ps^{-1}$ is close to the measured diffusivity in water[340] and allowed the time required to exit the initial stage of the AWH algorithm to be shorter than the whole simulation time (Figure S7). Similarly, changing ε_0 from 3 to 5 $kJ.mol^{-1}$ did not affect the PMF. Therefore, for the sake of simplicity and to standardize the MD inputs, the following two values (D_0 = 5.10^{-3} $nm^2.ps^{-1}$ and ε_0 = 3 $kJ.mol^{-1}$) were used for ethanol and later all xenobiotics.

The profiles of ethanol, and the three test compounds, namely quercetin, resveratrol, and curcumin exhibited low standard deviations (Figure 28A-C). As for the alchemical FEP, using the same force field, the AWH method showed a very good agreement in predicting the PMF profile of ethanol (Figure 28A). However, for the larger quercetin, curcumin and resveratrol, AWH exhibited many advantages (Figure 28B-D), especially the lower standard deviations between replicas. The AWH-derived minimum free energy of quercetin, which was located inside the lipid bilayer below the polar head groups, also agreed with the most stable position previously found just below the lipid head groups (10 to 15 Å from membrane center).[56, 58] Additionally, the computational time required to obtain converged estimate of permeability coefficients was relatively short, at around 600 ns per replica (Figure S8).

To rationalize convergence issues, it must be stressed that although the overall simulation time per xenobiotic was longer in FEP than in AWH, individual simulations were 40-50 times shorter with FEP (20-25 ns vs 1000 ns). Therefore, the sampling of the orientation of the xenobiotics was at stake. This was confirmed by following the normalized distributions of the θ-angle (between the z-axis and the longest vector) of quercetin, resveratrol and curcumin (Figure S3A) with both FEP and AWH calculations over the three replicas (Figure 29). For the four molecules (*i.e.,* ethanol, quercetin, resveratrol and curcumin), AWH exhibited similar distributions with the 3 replicas, *i.e.*, similar distribution irrespective from the initial orientation (see Figure 29). This confirms the ability of this methodology to sample the orientation inside the membrane. Although orientation was well sampled for ethanol (highly fluctuating θ-angle as seen in Figure S9) with FEP, large differences were observed in between the different replicas and for each window for quercetin, resveratrol and curcumin. For the latter molecules and at each depth of the FEP calculations, the distribution of θ orientation often remained

concentrated around the initial orientation, especially inside the lipid bilayer (Figure 29). Orientation sampling thus appears to be one of the major factors rationalizing the uncertainties in the absolute FEP-based PMF profiles.

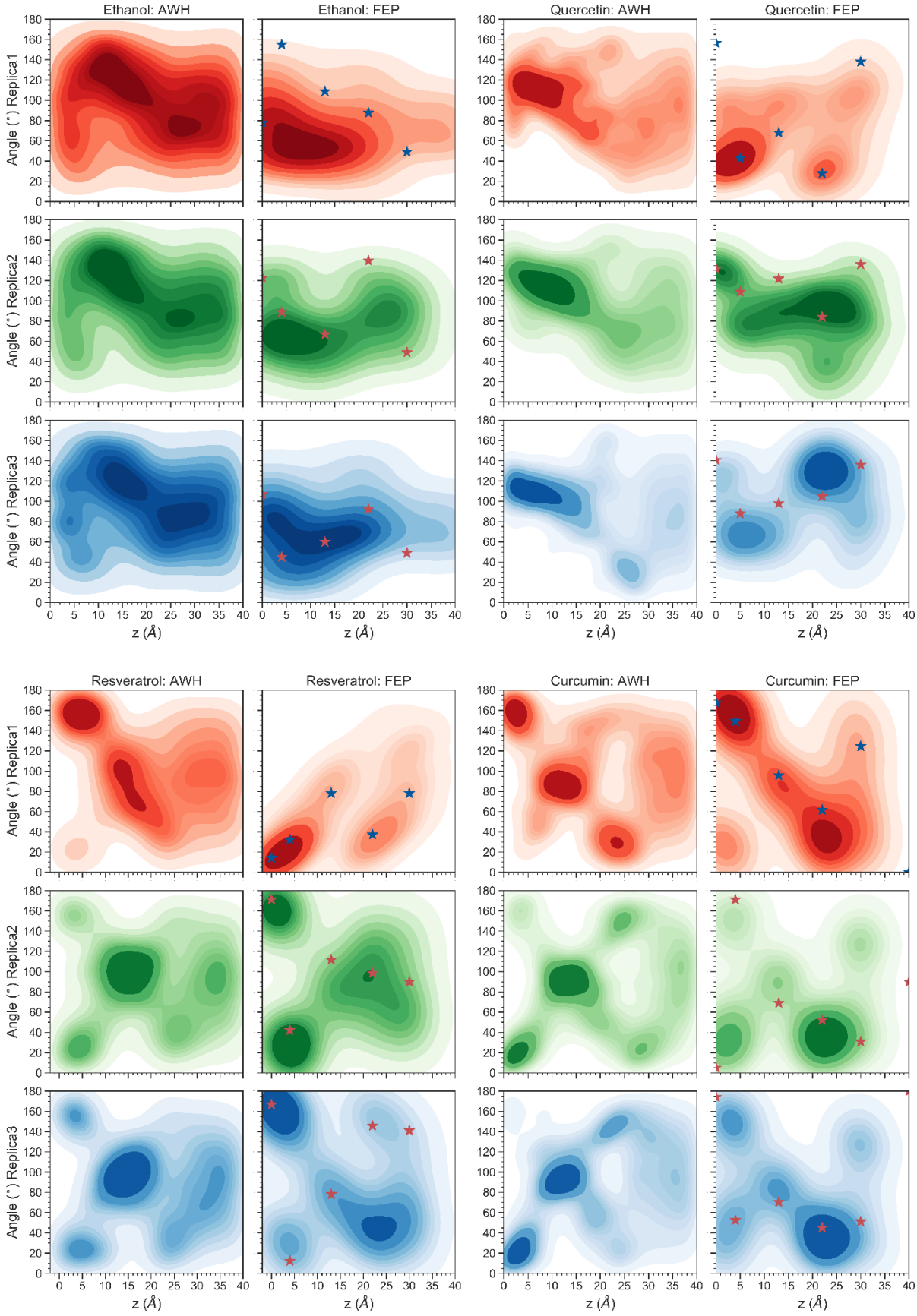


Figure 29. θ-angle distributions for ethanol, quercetin, resveratrol and curcumin through one layer of the membrane obtained with both FEP and AWH methods, and over three replicas. The star makers reflect the initial orientations of the molecules (each corresponding to a

replica), gathered at the end the equilibration simulations before running FEP. For FEP, θ was measured only for the first λ, in which the molecule is fully interacting with its environment.

V.4.1.3. Validation of the α-values for the fractional diffusivity calculations

Position-dependent fractional diffusivity profiles were computed using lag times ranging from 2 to 16 ps for quercetin and ethanol. Quercetin is taken as a prototypical case for drug-like xenobiotics, as it is amphiphilic, aromatic, and has a molecular weight in the range of the great majority of drugs.[26] Using the same power-law coefficient α as in ref [270], the fractional diffusivity of ethanol was roughly independent of all lag times, in agreement with Chipot and Comer[270] (Figure S10). The independence of lag times was also observed for quercetin, using the same α-values (Figure 30A). Therefore, α-values can be applicable to larger molecules.

Moreover, as the software Diffusion-Fusion was validated on trajectories obtained with the ABF method, here we reproduced the fractional diffusivity profile of ethanol using the AWH method. Figure 30B shows the smoothed fractional diffusivity (K_α) profiles for ethanol using a lag time of 4 ps. The diffusivity profile of ethanol ($K_\alpha \approx 363$ Å^2.ns$^{-\alpha}$ in water and $K_\alpha \approx 50$ Å^2.ns$^{-\alpha}$ at the membrane center) was very similar to that reported in the literature using Adaptive Biasing Force (ABF) method.[70,340] This suggests that the positions obtained from AWH simulations where the molecule diffuses on a flat PMF along the z-axis can indeed be used to compute the fractional diffusivity profiles.

V.4.2. Comparison between molecules

The use of AWH and the MemCross protocol was extended to all other xenobiotics of the series of twelve exhibiting a variety of chemical structures, namely quercetin, resveratrol, curcumin, coumarin, viniferin, rutin, caffeic acid, testosterone**,** methotrexate, MK571 and ceefourin 1 and 2.

The PMF profiles of neutral molecules obtained with the AWH method (Figure 28E) exhibited a variety of partition and membrane-center crossing free energies ($\Delta G_{partition}$ and $\Delta G_{flip\text{-}flop}$, see also Figure S6B). AWH-based minima were coherent for all molecules as well as relative energies. To give only a few examples, it is interesting to note that curcumin exhibited a low $|\Delta G_{flip\text{-}flop}|$ value of 2.5 ± 0.43 kcal.mol^{-1}*,* suggesting that curcumin can cross the center of the lipid bilayer at a low cost, which agrees with the literature showing its capability to flip-flop in several lipid bilayer systems.[368] Also, the difference in angle distribution between replica 1 and the other replicas of curcumin near the center of membrane or near the polar head groups, *i.e.,* 20-22 Å, can be explained by the symmetry of the curcumin molecule, which can adopt two orientations of the C-C vector: ~20° or the complementary angle ~160° with respect to z-axis. The orientation of quercetin ranges from 90° to 120° with respect to the z-axis, while resveratrol and curcumin had similar orientations, relatively parallel to z-axis, *i.e.,* ~20°/160°, near the center of the membrane, and parallel to the membrane surface below the polar head groups, *i.e.,* angle of the longest vector of the molecule and the z-axis ~90°. Such orientations have already been described for these polyphenols in lipid bilayer.[369,370]

Taking all compounds into account, interesting correlations were highlighted between free energy differences and molecule descriptors such as the molecular weight (MW), the number

of H bond donors (HDon) and acceptors (HAc) (Figure 30D). A strong negative correlation was highlighted between $\Delta G_{partition}$ and the molecular weight (R = -0.796) as well as with the average diffusivity in water (R = -0.868). On the other hand, translocation free energy ($\Delta G_{flip\text{-}flop}$) was positively correlated to the molecular weight (R = 0.713). Thus, for the molecules studied here the larger the molecule, the stronger its partition, and the harder its translocation through the bilayer center. This is in agreement with Lipinski's rule-of-five[26], however the limited dataset in this study prevents from generalization of these correlations.

V.4.2.1. Protonation state assessment

The combined free energy profiles of the four acids are shown in

F. More detailed profiles (neutral and ionized) are shown in Figure S5. Table 2 compares the depth of transition from ionized to the neutral form ($z_{transition}$) to the depth at which the molecule exhibits an energy minimum (z_{min}). For all four acids, $z_{transition}$ values were deeper in the lipid bilayer than the z_{min} of both forms. This means that the xenobiotics enter inside the membrane in their ionized form, and may then reprotonate to cross the lipid bilayer center, as already described[362]. It must be stressed that $z_{transition}$ was very different for all molecules, *i.e.,* 3-4 Å deeper than z_{min} for caffeic acid and methotrexate, and almost at the membrane center for MK571 and ceefourin 2. This indicates that the latter two molecules may permeate the membrane center equally in their ionized or neutral forms. At $z_{transition}$, all forms still exhibited several contacts with water molecules (Figure S11).

At neutral and alkaline pH, caffeic acid action on membranes was shown to be confined at the membrane surface in the lipid head groups, suggesting that at this pH, the carboxylic group is under its ionized form and points toward the positive choline moiety of the polar head group. Conversely, at acidic pH, only neutral form is present and therefore can permeate the membrane.[371] This agrees with the PMF profiles of both forms of caffeic acid (Figure S5A), as the barrier at the center of the membrane is lower for the neutral than for the charged form.

V.4.2.2. Fractional diffusivity calculations

Figure 30B shows the smoothed fractional diffusivity (K_α) profiles for the twelve xenobiotics. As expected, all twelve compounds diffuse much slower than ethanol due to bulkier chemical structures. As expected, the diffusivity coefficients drastically decrease inside the membrane.

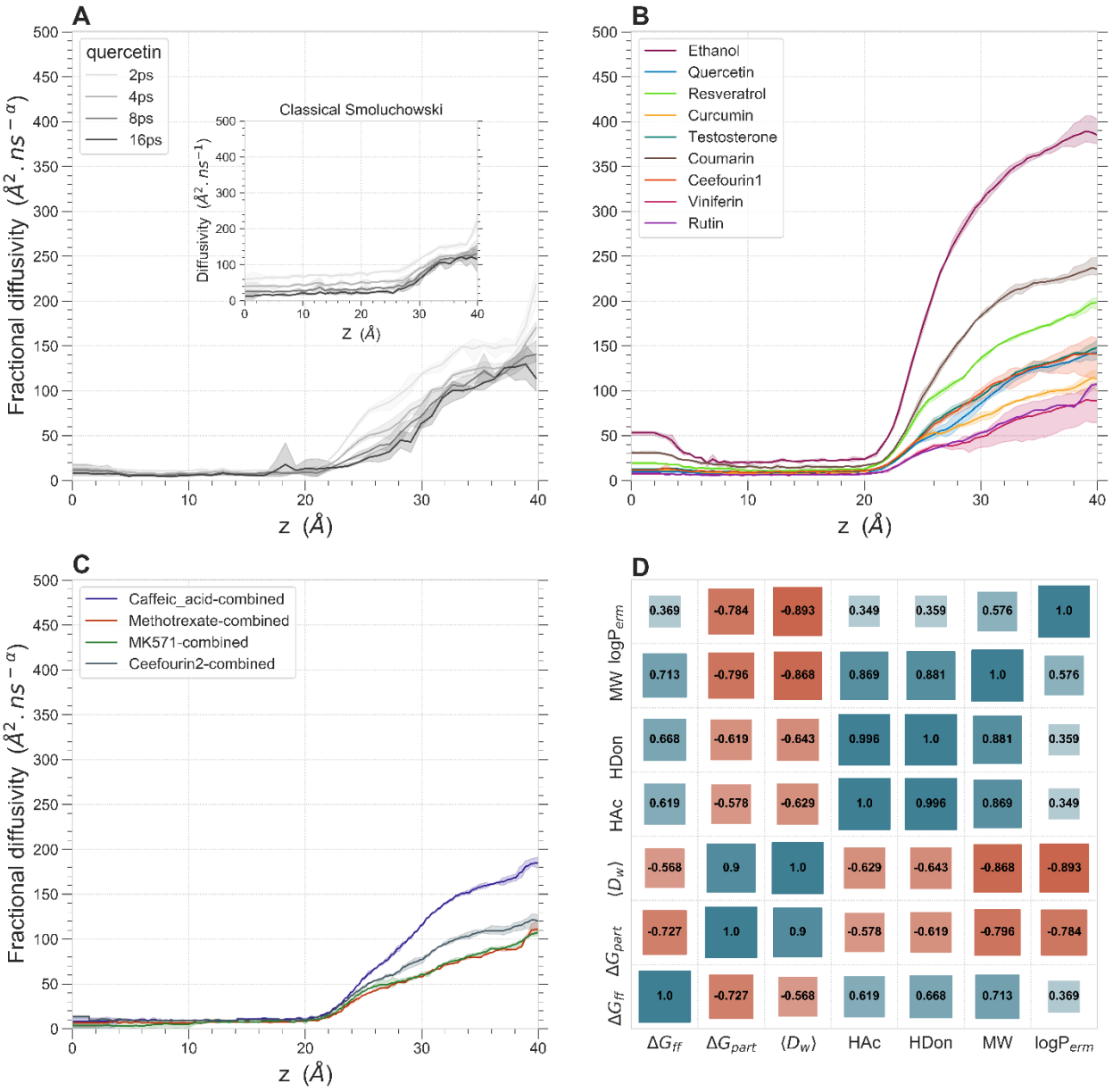


Figure 30. A) Fractional diffusivity (K_α) profiles of quercetin for different lag-times Δt compared to the diffusivity profiles calculated based on the classical Smoluchowski equation. B) Smoothed Fractional diffusivity profiles for Δt = 4 ps of all molecules. C) Combined Fractional diffusivity profiles for caffeic acid, methotrexate, MK571 and ceefourin 2. Standard errors were calculated over three replicas. D) Heatmap matrix showing the correlation coefficients R between free energy differences ($\Delta G_{partition}$ and $\Delta G_{flip\text{-}flop}$), three molecular descriptors (HAc: number of hydrogen-bond acceptors, HDon: number of hydrogen-bond donors and MW: molecular weight), averaged diffusivity in water of neutral molecules ($\langle D_w \rangle$) and $logP_{erm}$.

The fractional diffusivity of coumarin was $K_\alpha \approx 219$ Å^2.ns$^{-\alpha}$ in water. As for ethanol, but to a lesser extent due to larger structures (MW = 46.07 and 146.14 g.mol^{-1} for ethanol and coumarin, respectively), coumarin exhibited a slight increase in diffusivity at the membrane center, due to the presence of the free volumes. Resveratrol diffuses faster in the aqueous phase with a diffusivity coefficient $K_\alpha \approx 169$ Å^2.ns$^{-\alpha}$, while rutin exhibited the lowest diffusivity in water with $K_\alpha \approx 77$ Å^2.ns$^{-\alpha}$. Being glycosylated, rutin preferentially binds to the polar heads

through hydrogen bonding formed with its sugar moiety, thus being less prone to diffuse than non-glycosylated and smaller molecules. All molecules but coumarin, highlighted a decrease in diffusivity at the membrane center. The correlation matrix in Figure 30D shows a negative correlation between the diffusivity coefficient of a molecule in water and its molecular weight, confirming that the larger the molecule, the lower its diffusivity.

Diffusivity profiles were computed for both molecule forms of the four acid molecules (Figure S5B). There were no significant differences between the diffusivity profiles of the two forms suggesting that the diffusivity of a molecule is not related to its ionization state (Figure 30C).

Table 2. z-coordinates of the energy minima (z_{min}) and transition points ($\langle z_{transition} \rangle$) of acid molecules used in this study.

Molecule	**z_{min} (Å)**			**$\langle z_{transition} \rangle$ (Å)**
	Neutral	Ionized	Combined	
caffeic acid	14	18	18	14
methotrexate	15	18	18	11
MK571	12	13	13	0
ceefourin2	10	13	13	1

V.4.2.3. Permeability assessment

Interestingly, most xenobiotics exhibited faster permeation ability compared to ethanol, see Table 3. The following hierarchy for permeation capacity was observed with all xenobiotics: MK571 > quercetin > curcumin > ceefourin 1 > resveratrol > viniferin > rutin > coumarin > ceefourin2 > ethanol >> methotrexate > caffeic acid (see Table 3). Based on **equation (1),** P_{erm}^{calc} is exponentially related to the PMF especially at the membrane interior. The higher the barrier, the lower the $\log P_{erm}^{calc}$ value. Besides, diffusivity values are comparable for all xenobiotics, and thus play a minor role. Hence, the free energy barrier at the center of the membrane is the key factor in $\log P_{erm}$. The lower values of rutin and viniferin compared to the other polyphenols is likely due to their higher molecular mass, and to the presence of a polar sugar moiety for rutin, which strongly anchors the molecule to the polar head group region. Although their hydrophobic moieties partition deeper in the membrane (low $\Delta G_{partition}$, see Figure S6B), they exhibit high free energy barriers to cross the bilayer center (high $\Delta G_{flip\text{-}flop}$, see Figure S6B). Regarding the ionizable molecules, MK571 exhibited the highest permeation coefficient, in agreement with its low free energy value at the membrane center, followed by ceefourin 2. Conversely, caffeic acid and methotrexate exhibited much lower $\log P_{erm}^{calc}$ values (-4.85 ± 0.32 and -4.09 ± 0.17, respectively), almost precluding passive permeation for these two compounds. This is rationalized by their high PMF barrier at the membrane center (Figure 28F).

Table 3. Computed permeation coefficient ($\log P_{erm}^{calc}$) with MemCross compared to experimental permeation coefficient ($\log P_{erm}^{exp}$), calculated partition coefficient ($\log P_{art}^{calc}$) and octanol-water partition coefficient ($\log P_{o/w}$) for the set of molecules used in this study.

Molecules	$\mathbf{\log P_{erm}^{calc}}$ (**MemCross**)	$\mathbf{\log P_{erm}^{exp}}$	$\mathbf{\log P_{art}^{calc}}$	$\mathbf{\log P_{o/w}}$
ethanol	0.47 ± 0.0	-4.42[372], -3.28[22]	0.69	-0.31[373], -0.16[374]
quercetin	1.73 ± 0.03	-4.77[22], -4.4[22], -3.2[22], -3.38[22], -4.03[22]	8.68	1.82[375]
resveratrol	1.64 ± 0.02	-4.38[22], -4.4[22], -4.82[22]	5.39	3.1[376], 1.53[377]
curcumin	1.70 ± 0.0		7.04	~ 3[378]
testosterone	1.59 ± 0.0	-2.83[22], -3.99[22], -3.58[22], -3.10[22]	5.35	3.32[373]
coumarin	1.40 ± 0.0		2.63	1.39[373]
ceefourin1	1.68 ± 0.0		5.09	
viniferin	1.60 ± 0.03		8.50	1.58[377]
rutin	1.49 ± 0.03		7.04	-0.64[375]
ceefourin2	1.23 ± 0.07		6.01	
caffeic acid	-4.35 ± 0.24		1.38	1.15[379]
MK571	1.74 ± 0.01		8.65	
methotrexate	-4.33 ± 0.16	-7.04[22], -9[22], -5.57[22]	3.54	-1.85[373]

Convergence of $\log P_{erm}^{calc}$ was assessed over the simulation time with a time step of 100 ns. As shown in Figure S8, 500-700 ns per replica were suitable to get converged $\log P_{erm}$ values for all neutral molecules except large and polar ones. For rutin, viniferin, methotrexate and caffeic acid, 1 to 3 µs were required to converge. Regarding rutin, it exhibited the slowest convergence among all the molecules studied here. Therefore, the MemCross may have limitations for large xenobiotics containing many polar moieties. For large xenobiotics, biasing only one collective variable (the projection of their center-of-mass on the z-axis) may not be sufficient to achieve fast convergence. Additional collective variables such as the xenobiotic orientation or solvation may help reduce sampling time. Alternatively, collective-variable independent biased methods such as GaWTM-eABF[380] can be used.

V.4.2.4. Comparison with experimental values

Experimental permeation coefficients are only available for a subset of the xenobiotics studied here (ethanol, quercetin, resveratrol, testosterone, methotrexate, Table 3), that does not allow a proper benchmarking of the method. Methotrexate was however correctly identified as the xenobiotic with the lowest $\log P_{erm}$ value. This is likely due to its ionizable acid moieties that confer more resistivity to permeation across the lipid bilayer. Calculated $\log P_{erm}^{calc}$ values with MemCross were systematically higher than the experimental ones by 3 to 6 log units for these molecules. Such discrepancy is akin to what was described in the literature for other theoretical methods.[22,340,381] Several reasons may contribute to these large differences. The free energy

estimation is the most critical to calculate P_{erm}, due to the exponential function. Force field inaccuracies could lead to errors in the free energy profiles. Although the Slipids force field in combination with GAFF for small molecules was shown to be accurate to reproduce experimental $\Delta G_{partition}$ for small molecules in lipid bilayers,[363] that may not be the case for $\Delta G_{flip\text{-}flop}$ which drives $\log P_{erm}^{calc}$. Although the correlation between $\log P_{art}$ and $\log P_{o/w}$ in this study is weak (Table 3), the largest differences can be attributed to the distinction between octanol and POPC. Indeed, polar molecules such as MK571 or rutin exhibit a strong affinity for POPC and not for octanol. This can be attributed to the increased polarity of phospholipid head group, absent in octanol. Another source of error correlating with the overestimation of the permeation is the overestimated diffusivity of the TIP3P model.[382] Discrepancies may also arise from the difference between intrinsic and apparent permeabilities. Apparent permeability is typically measured experimentally, and involves the slow diffusion in a broader aqueous medium, including unstirred water layers.[111] Intrinsic permeability only considers the lipid bilayer. It is evaluated by simulations, or it can be obtained through corrections applied to experimental apparent permeabilities.[326,383,384] Additionally, the variety of experimental setups and membrane composition further hinders the comparison with simulations, discrepancies within $\log P_{erm}^{exp}$ values reported here are often up to 1.5 log units (Table 3). Finally, MemCross can also account for a part of the discrepancies within experimental values. Other theoretical methods, such as ABF, were able to reproduce experimental results for a few drug-like molecules with a better precision (i.e., $\log P_{erm}^{calc}$ within one log units).[385] However, ABF required multi-µs simulations to achieve convergence. Therefore, MemCross appears as a compromise among other theoretical methods. We showed that it can provide converged estimates of $\log P_{erm}^{calc}$ within a few hundreds of ns for drug-like molecules, with full atomistic details. Although absolute predictions of experimental values are poor, relative predictions are promising. Regarding the relatively low computational cost, this may allow navigating and extending databases.

V.5. CONCLUSION

AWH is a promising method to assess all ingredients (PMF and diffusivity) required to calculate P_{erm}. One of the major advantages of AWH is that the individual simulations allow an efficient sampling of the movements of the xenobiotics inside the membrane. This leads to convergence of the PMF profiles along the translocation path. Fractional diffusivity profiles can be directly calculated from the AWH trajectories with the MemCross protocol. The computational cost to obtain a converged estimate of P_{erm} with MemCross for drug-like compounds was often lower than other enhanced sampling methods, such as ABF or umbrella sampling, which require multi-µs simulations for each replica[385–387]. AWH has the advantages of both an exponential convergence, and of being almost insensitive to the input parameters D_0 and ε_0. Overall, all these advantages make MemCross a very promising method to evaluate permeation coefficients of series of structurally unrelated compounds. Although the agreement with experimental permeability values was limited and the small number of xenobiotics prevented from a comprehensive exploration of all drawbacks and limitations, MemCross appeared as a very promising method to study and evaluate permeation coefficients on large databases, hence considering a wide diversity of organic molecules, which is particularly fruitful in the pharmacological context.

Chapter VI. MemCross, a robust computational tool to evaluate permeability coefficients of drug-like molecules

Mehdi Benmameri [a], Benjamin Chantemargue[b], Patrick Trouillas[a,c], Gabin Fabre[a]

[a]INSERM, UMR 1248, F-87000 Limoges, France;

[b]INSILIBIO, 1 avenue Ester Technopôle, 87069 Limoges Cedex, France;

[c]CATRIN RCPTM, 779 00 Olomouc - Holice, Czech Republic

Foreword

In this study, we benchmark the MemCross methodology to compute passive permeation coefficients and compare with different sets of experimental measurements. Convergence of the permeation coefficient and the impact of diffusivity calculations in the logPerm assessment were also discussed. This study will be submitted in the near future.

VI.1. INTRODUCTION

One of the most important causes of drug attrition in the pharmaceutical industry is a poor bioavilability. Simple structure-activity relationships based on common molecular descriptors do not correlate with preclinical studies outcomes.[388] Therefore, a better rationalization of a drug bioavailability is very desirable. One of the critical feature of bioavailability is membrane crossing, which is often the rate-limiting step in pharmacokinetics. The crossing of biological membranes often goes through a complex process combining active transport through membrane proteins, para-cellular crossing, and spontaneous passive permeation across lipid bilayers.[3,73,81,82,93] In the case of lipophilic or amphiphilic compounds, the latter process is dominant, especially for drug influx.[18,42,70,115,122,127,318,339] Therefore, the prediction of permeation coefficients is key to rationalize and predict the pharmacokinetics of xenobiotics.

There exist several *in vitro* methods that have extensively been used to predict drug absorption in the human body. The most common are based on cultured colorectal adenocarcinoma-derived, Caco-2, Madin-Darby canine kidney epithelial (MDCK) and blood brain barrier (BBB) cell lines.[11,42,137,147,148,165,167] These cell-based models contain membrane transporters,[155,158,159] which complexifies the evaluation of passive permeation, as it is difficult to discriminate passive permeation events from translocation through membrane proteins. Furthermore, with these *in vitro* models, the time required for cell culture precludes from high-throughput screening. Artificial lipid bilayer models as used in the parallel artificial membrane permeability assay (PAMPA) or liposomes and black lipid membranes (BLM) have raised an interesting alternative.[9,10,14,173,183] These biomimetic models offer many advantages against their cell-based counterparts, mainly because active transport is absent from these models. The community has thoroughly produced a large amount of experimental data, however with significant variability in the protocols, and this lack of standards hinders from easy comparison between data.[14,137,165]

In an effort to speed up screening processes under the specification of comparison between compounds, many theoretical models have been developed to predict membrane permeation coefficients. Historically, membrane permeation has been assessed by evaluating the partition

coefficient of organic molecules between aqueous phase and organics solvents like octanol ($logP_{octanol/water}$).[389–391] Several quantitative structure−activity relationship (QSAR) models were developed to predict $logP_{octanol/water}$ values,[392–394] with very low computational resources. However, biological membranes are more complex than simple organic solvents such as octanol. This is a major drawback to allow a comprehensive description of passive permeation based solely on $logP_{octanol/water}$, in turn precluding accurate predictions of permeation coefficients. Lipinski's rule of five introduced other factors to improve this prediction, however many compounds are found to be exceptions to that rule.[115] Another theoretical model, the solubility diffusion model (SDM), was developed to determine permeation coefficients of molecules diffusing through membranes.[6,111] At steady state, it relates the permeation coefficient to the flux of molecules crossing the membrane and to the concentration gradient. The homogeneous solubility diffusion model (HSDM) has shown good correlations in predicting passive diffusion through polymer membranes.[20,126] The theoretical framework of HSDM considers the permeation coefficient related to the diffusivity in the homogeneous membrane, the partition coefficient between both phases (aqueous solvent and membrane interior) and the membrane thickness. Yet, this model has failed at describing the anisotropic nature of lipid bilayers. Later on, the inhomogeneous solubility diffusion model (ISDM) was developed to evaluate the translocation of particles across inhomogeneous lipid bilayer membranes, under the influence of a deterministic force that is acting on the particle along the permeation path. This was formulated through the Smoluchowski diffusion equation.[18,128,129] The ISDM theory is a promising formalism to predict permeation coefficients that can be calculated by introducing a resistance, R, to crossing events given by:

$$\mathrm{R} = \frac{1}{\mathrm{P_{erm}}} = \int_{-\frac{\mathrm{L}}{2}}^{\frac{\mathrm{L}}{2}} \frac{\mathrm{e}^{\beta \mathrm{w(z)}}}{\mathrm{D(z)}} \mathrm{dz} \qquad (1)$$

where w(z) is the free energy profile along the permeation path, $\mathrm{D(z)}$ is the diffusivity at a given depth, z, and $\beta = \frac{1}{kT}$ with k the Boltzmann constant and T the temperature. z is the coordinate along the permeation path and it is integrated from one to the other side of the membrane.

The computational methodologies to evaluate the major two ingredients of equation (1) (i.e., t w(z) and $\mathrm{D(z)}$) should be developed for a chemical library large enough to sample the wide chemical diversity of compounds, while keeping an affordable computational cost. Many studies were performed on series made of only a few compounds, which has prevented from generalization and validation of the methodologies for large molecule screening applicable to drug design.[70,395–397] To test the robustness of a computational method requires comparison to experimental data. However, this is the apparent permeability which is usually measured, whereas theoretical methods provide an intrinsic permeability. By intrinsic permeability we mean that focus is only given on the neutral form of the molecule crossing passively the lipid bilayer, without taking into account the large unsitirred water layers, transport by membrane proteins, or paracellular permeability.[14] In other words, the theoretical methods usually only consider the free energy underlying the translocation of the neutral form of the molecule and diffusivity profiles along the membrane normal (i.e., z-axis) that is delimited by membrane boundaries (from –L/2 to L/2).[22,398]

Many physics-based models have shown promising results in predicting passive membrane permeation coefficients. The advantage of such methods is their low computational cost and thus their relatively easy usage in screening drug databases. However, the inherent physical

formalisms do not allow to capture the process at the atomic scale. The passive permeation prediction based on molecular dynamics (MD) simulations has been a promising solution to address this understanding together with predicting permeation coefficients. Although the computational cost of observing membrane-crossing events within conventional MD simulations is prohibitive, biased MD simulations allow the determination of permeation coefficients in a reasonable computational time. The main benefit of using MD simulations is the explicit description of all interatomic interactions. Nevertheless, the accuracy of the permeation coefficient predictions is still a methodological challenge. The use of several biased MD methods have been reported, such as Umbrella sampling, Adaptive Biasing Force or metadynamics. [70,269,270,395–397] We have developed a protocol named MemCross based on the AWH algorithm that has shown efficacy at predicting permeation coefficients of molecules through a POPC lipid bilayer by assessing both free energy and diffusivity profiles from a single simulation.[292] MemCross was tested on 12 different molecules and it showed promising results obtained at a lower computational cost compare to other biased methods. Here, we assess the accuracy of MemCross by computing the permeation coefficients of more than 350 compounds, for which experimental permeation coefficients are available. The impact of diffusivity on the permeation coefficient values are also discussed. The most adapted simulation times were also highlighted according to the molecular sizes. Finally, we compared our results with experimental measurements obtained with biomimetic membrane models (liposomes and BLM, double-sink PAMPA) and cell-based membrane models (BBB and Caco2/MDCK).

VI.2. METHODS

VI.2.1. MD simulations

MD simulations were performed using the Gromacs 2021 software.[241] The Slipids 2020 force field[57] was used for the lipids and the TIP3P model for water molecules. The pdb files for the permeants (i.e., undissociated and ionized forms, neutrals, acids, bases, zwitterions) used in this study were gathered from Lomize *et al*.[22,399] (Table S2). They were parametrized using GAFF2 (second generation generalized amber force field)[351]. First, several conformers were generated using Confab software[353] within a 100 kcal/mol energy cutoff, then geometry optimization was carried out using the Gaussian09 software[354] with the B3LYP hybrid functional and the cc-pVDZ basis set in an implicit diethylether solvent. Atomic RESP partial charges were derived from Gaussian calculations using the R.E.D.III software and Duan *et al.* method[355]. Among permeants, monoatomic ions and large molecules (i.e., molecular weight (MW) greater than 800) were omitted. Periodic boundary conditions were applied in all three dimensions. The MD time step was set to 2 fs. The cutoffs for electrostatic and van der Waals interactions were set to 14 Å. Long-range electrostatic interactions were calculated using the Particle-Mesh-Ewald (PME) method. The V-rescale thermostat was used to maintain the temperature at 310 K and the Parrinello-Rahman barostat was used to maintain the pressure at 1.013 bar.

A 1-palmitoyl-2-oleyl–sn-glycero-3-phosphocholine (POPC) lipid bilayer containing 128 molecules was used. The pure POPC membrane was chosen as a model of lipid bilayers used experimentally.[70] The x-y box dimensions were 66 Å × 66 Å. The box dimension along the z axis was set according to the permeant size, as follows:

$$z_{box} = 2 \times \left(z_{init} + z_{permeant} + z_{margin}\right) \quad (2)$$

Where z_{init} was set to 38 Å, which corresponds to the initial positioning of the molecule in the water bulk with respect to the membrane center (z_{center} = 0 Å), $z_{permeant}$ was the length on the permeant longest axis, and z_{margin} represents an extra thickness of solvent set to 6 Å. The smallest value for z_{box} was 90 Å.

VI.2.2. Potential of the mean force

We used the same protocol to predict the potential of the mean force (PMF) and diffusivity profiles as described in our previous work.[292] In brief, the accelerated weight histogram (AWH) is an enhanced sampling method that can be used to predict free energy landscapes of biomolecular systems.[247] As we showed in our previous work, AWH can compute the PMF along the translocation pathway (i.e., membrane normal that is the z-axis). The molecule was initially placed in the water compartment at z ~ 40 Å. The distance between the center of masses of both the molecule of interest and the membrane projected on the z-axis constitutes the one-dimensional reaction coordinate. This coordinate is biased within the AWH simulation. The target distribution was set to be uniform. This allows the free energy along the chosen reaction coordinate to be flattened during the simulations by applying a history-dependent bias potential. The error on the free energy was limited to 3 kJ.mol^{-1} and the initial diffusion constant was set to 5.10^{-3} nm^2.ps^{-1}. The force constant for the harmonic potential was chosen to be large enough (128000 kJ.mol^{-1}.nm^{-2}). For each molecule, three replicas with identical inputs were simulated. The simulation time was 1 µs per replica for all membrane-molecule systems, regardless the size of the molecule.

VI.2.3. Position-dependent diffusivity

The subdiffusive behavior of molecules within the POPC lipid bilayer membrane was taken into account.[70,270,292] As reported in our previous work[292], the position-dependent fractional diffusivity profiles $K_\alpha(z)$ were estimated from the AWH trajectories using a Bayesian scheme provided in the DiffusionFusion software.[270,289] The likelihood of the observed trajectory is computed as the product of the probabilities of the observed movements from Z(t_j) to Z(t_j + Δt):

$$\mathrm{P}[\mathrm{Z}(\mathrm{t})|\mathrm{K}_\alpha(\mathrm{z})] = \prod_{\mathrm{j}} \mathrm{p}\left[\mathrm{Z}\left(\mathrm{t}_\mathrm{j} + \Delta\mathrm{t}\right)|\mathrm{Z}\left(\mathrm{t}_\mathrm{j}\right), \mathrm{K}_\alpha(\mathrm{z})\right] \quad (3)$$

In practice, the observed movements are calculated by numerically solving the fractional Smoluchowski diffusion equation using Crank-Nicolson finite difference method.[301]

Then, the posterior probability is maximized using a Monte Carlo procedure:

$$\mathrm{P}[\mathrm{K}_\alpha(\mathrm{z})|\mathrm{Z}(\mathrm{t})] = \mathrm{p}_{\mathrm{prior}}[\mathrm{K}_\alpha(\mathrm{z})] \times \mathrm{P}[\mathrm{Z}(\mathrm{t})|\mathrm{K}_\alpha(\mathrm{z})] \quad (4)$$

The $p_{prior}[K_\alpha(z)]$ probability was set uniform with a value of K_α = 100 Å^2.ns^{-1}. The z-dependent fractional order α values were ranging from 0.7 at the membrane center to 1 in the water bulk, as previously described.[270,292]

Classical diffusivity profiles were also evaluated using the same Bayesian algorithm, but with the classical Smoluchowski diffusion equation. A lag time of 4 ps was used in the diffusivity calculations. For convergence assessment (Figure 32), a lag time of 2 ps was used.

VI.2.4. Ionization state assessment

The full permeation from one side to the other side of the membrane of ionizable molecules usually requires protonation and deprotonation events. Here, we only considered either the neutral form or the ionized form at a time, to mimic different acid-base conditions.[133,292] The PMF and the fractional diffusivity profiles were calculated for both forms. In water, the free energy difference driving the deprotonation reaction ($\Delta G_{deprotonation}$) is derived from the Henderson Hasselbalch equation and is equal to:

$$\Delta G_{deprotonation} = 2.3\ k_b T\ (pKa - pH) \quad (5)$$

Where k_b and T are the Boltzmann constant and the temperature, respectively.

As the ionized form is predominant in water, the free energy of the ionized form in water was set to 0. Then, the PMF profile of the neutral form was shifted by $\Delta G_{deprotonation}$. A combined PMF profile was defined as the minimum free energy of both PMFs along the z-axis. The z-coordinate corresponding to the transition point between the PMFs of both forms was used to combine the diffusivity profiles in the same manner. These combined profiles were used to calculate the permeation coefficient of ionizable molecules.

VI.2.5. Determination of permeation coefficients

The permeation coefficient P_{erm} was evaluated based on ISDM (equation 1). To reproduce the experimental conditions, the integration interval was set to -22 Å to +22 Å (i.e., membrane thickness).

VI.2.6. Estimation of the error on permeation coefficients

The propagation of the error of the PMF (w(z)) and of the diffusivity $K_\alpha(z)$ on the permeation coefficient (equation 1) was evaluated using a Monte Carlo procedure with 1000 cycles from the three replicas, as previously reported.[292] Using this procedure, one thousand $logP_{erm}$-values were generated for each molecule. This yielded a more accurate estimate of an averaged value of $\log P_{erm}^{MemCross}$ and its corresponding standard deviation.

VI.2.7. Linear fit to experimental permeation coefficients

The experimental permeation coefficients ($\log P_{erm}^{exp}$) values of the *in vitro* different membrane types (i.e., liposomes, BLM, PAMPA-DS, BBB, Caco2/MDCK) were collected from the PerMM database.[399] These experimental values come from different sources, which makes difficult the correlation between theoretical and experimental $\log P_{erm}^{exp}$. First, a linear regression model fit between the calculated $\log P_{erm}^{MemCross}$ and the experimental $\log P_{erm}^{exp}$ values was performed with each experimental dataset. Then, linear corrections from the regressions were applied to the theoretical data in the form:

$$\log P_{erm}^{f} = a \times \log P_{erm}^{MemCross} + b \quad (6)$$

Where $\log P_{erm}^{f}$ is the refitted $\log P_{erm}^{MemCross}$ and a and b are the slope and the intercept, respectively.

For each correlation fit, the quality of the linear regression fit was evaluated by the coefficient of determination R^2. Outliers were defined as compounds that deviated more than twice the standard deviation from the fit line.

VI.3. RESULTS AND DISCUSSION

VI.3.1. Potential of mean force (PMF) and diffusivity profiles of selected examples

Here, we applied the AWH-based protocol that have been previously reported[292] to model the free energy profiles across a POPC lipid bilayer membrane. This method was capable to predict the PMF as well as the fractional diffusivity profiles of a set of drug-like molecules, and thus the permeation coefficients from equation (1). It also highlighted its accuracy to sample the orientation of the permeants all along the bilayer membrane normal (i.e., z-axis). Moreover, the permeation of ionizable compounds can be assessed by considering both the neutral and the ionized (at a given pH) forms, then integrate over the combined free energy and factional diffusivity profiles (equation 5).In order to validate the accuracy of our tool, the prediction of permeation coefficients for a larger set of molecules (i.e., more than 417 compounds, see Table S2) was established and compared with experimental values.

The PMF and diffusivity profiles of 10 representative molecules with different molecular structures are plotted in Figure 31A.

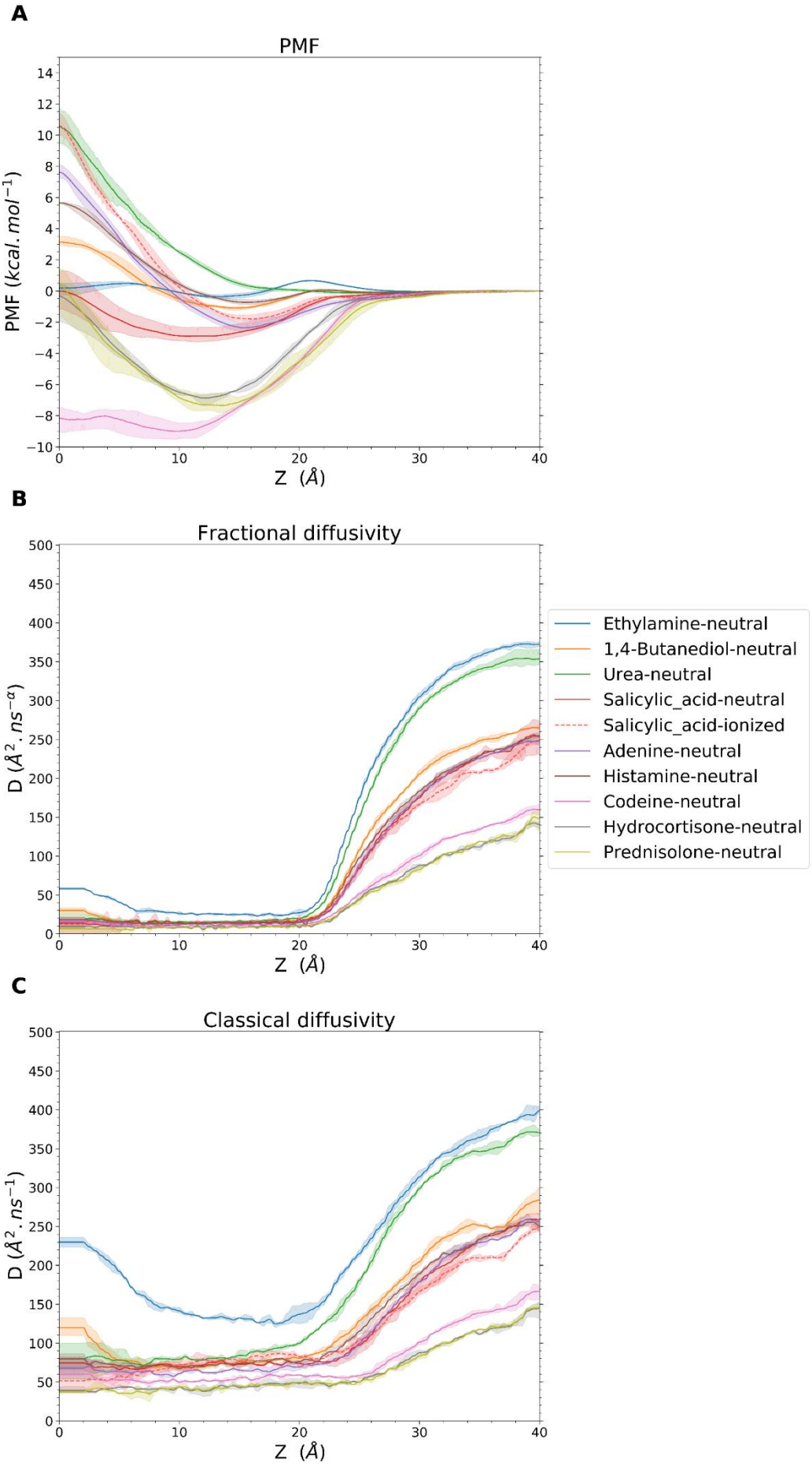


Figure 31: (A) Potential of mean force (PMF), (B) Fractional diffusivity profiles and (C) Classical diffusivity profiles for 9 molecules with different molecular sizes. The two forms (neutral and ionized) of salicylic acid are shown with red dotted lines and red straight lines, respectively.

Codeine is a prototypical example of hydrophobic molecules. It exhibited the lowest free energy barrier at the membrane center, and thus its accumulation in the core of the membrane. The relatively flat PMF in the lipid tail region of the membrane (i.e., between the center and 11 Å) suggested frequent flip-flop events from one to another membrane leaflet.

Both prednisolone and hydrocortisone are taken as prototypical amphiphilic molecules. They exhibited a favorable partition below the lipid polar head groups, where the polar moieties interact with the polar head groups, and the lipophilic moieties with the lipid tails. Conversely to codeine, an increase in the free energy was observed at the center of the membrane for both compounds, due to unfavorable interactions of polar moieties with the highly apolar environment in this region. Ethylamine is a different example of amphiphilic molecule, exhibiting a relatively iso-energetic PMF with a slight increase in the surrounding of the polar head group region.

The 1,4-butanediol, histamine, adenine, and urea compounds are typical neutral hydrophilic molecules. As expected, they exhibited a high free energy barrier at the center of the membrane, consistent with a disruption of the bilayer organization caused by the presence of polar moieties in this region.

Salicylic acid is a prototypical polar and ionizable drug. Due to the non-favorable interactions of its ionized carboxylic group with the surrounding hydrophobic environment, its ionized (charged) form exhibited a much higher energy penalty at the membrane center than the neutral form. Salicylic acid is expected to approach the membrane in its ionized form, then to re-protonate below the polar head groups at 0.65 nm, when the number of surrounding water molecules decreases. This would ease crossing of the membrane center in its neutral form.

The standard deviations between the 3 replicas of the PMF profiles of Figure 31A were lower than 1 kcal/mol. The reproducibility of the results indicates that the simulation time of 1 µs per replica allowed a sufficient conformational sampling for most of the molecules and their direct environments along the permeation path.

The fractional and classical diffusivities were calculated using the respective Smoluchowski equations (Figure 31B-C). In the water phase, both fractional and classical diffusivities exhibited similar values. This is expected because the two Smoluchowski equations are identical in water, with a fractional order of 1. Within the membrane core, the fractional diffusivity profiles were relatively flat and lower than 50 $Å^2.ns^{-\alpha}$. Conversely, classical diffusivity profiles (Figure 31C) exhibited larger variations within the membrane, according to the size of molecule. Interestingly, both fractional and classical diffusivity profiles exhibited an increased diffusivity at the membrane center for the small compounds, namely ethylamine and 1,4-butanediol, albeit the effect is smaller with fractional diffusivity. This agrees with the literature about short-chain alcohols.[70] Moreover, both forms of salicylic acid exhibited similar diffusivity profiles, indicating the non-dependence of the diffusivity on the charge of the molecule.

VI.3.2. $logP_{erm}^{MemCross}$ convergence

The convergence of the calculated permeation coefficient values, referred here as $\log P_{erm}^{MemCross}$, was evaluated for all the molecules used in this study. It was assessed by following the average of differences between $\log P_{erm}^{MemCross}$ along the simulation (every 100 ns) and $\log P_{erm}^{MemCross}$ obtained at 1 µs (Figure 32). The $\log P_{erm}^{MemCross}$ values were considered converged when the difference was around *half log unit*. By subdividing the permeants according to their MW, the results shown in Figure 32 confirmed that the smaller the molecule, the faster the convergence.

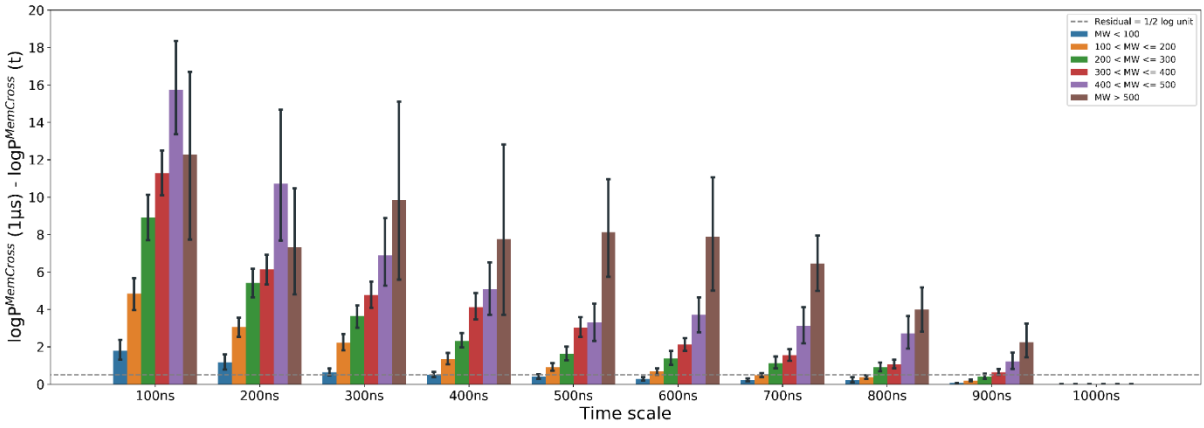


Figure 32: Convergence of $logP_{erm}^{MemCross}$ along the simulation for as a function of MW. The convergence is followed by the difference between $logP_{erm}^{MemCross}$ at time t (along the simulation, x-axis) and the $logP_{erm}^{MemCross}$ at time t = 1µs.

After 100 ns MD simulation, the differences were high for all the molecules, along with high standard deviations between permeants. This can be explained by the coarse bias updates during the initial stage of the AWH procedure, and it indicates that longer simulations are necessary for a finer estimate of $\log P_{erm}^{MemCross}$. For very small molecules (MW lower than 100 g.mol^{-1}), $\log P_{erm}^{MemCross}$ convergence was reached within only 300 ns. To reach convergence for molecules with MW lower than 400 g.mol^{-1}, the required simulation time was in the range 800-900 ns. Molecules with a MW larger than 500 g.mol^{-1} may require more than 1 µs of simulation time to reach convergence. However, their drug-likeliness is likely lower according to Lipinski's rule of five.[26]

VI.3.3. Diffusivity prediction

To the best of our knowledge, here we have computed the largest collection of diffusivity profiles of permeants in membrane from all-atom MD simulations. Based on this dataset, we have investigated whether the fractional diffusivity can be predicted from a few molecular descriptors, namely the logarithm of the molecular weight (logMW), the logarithm of radius of gyration, the accessible surface area (ASA), the number of rotatable bonds, and the inverse of the smallest cross-sectional area (Inv-CSA). Figure 33 shows the correlation matrix between each pair of the parameters.

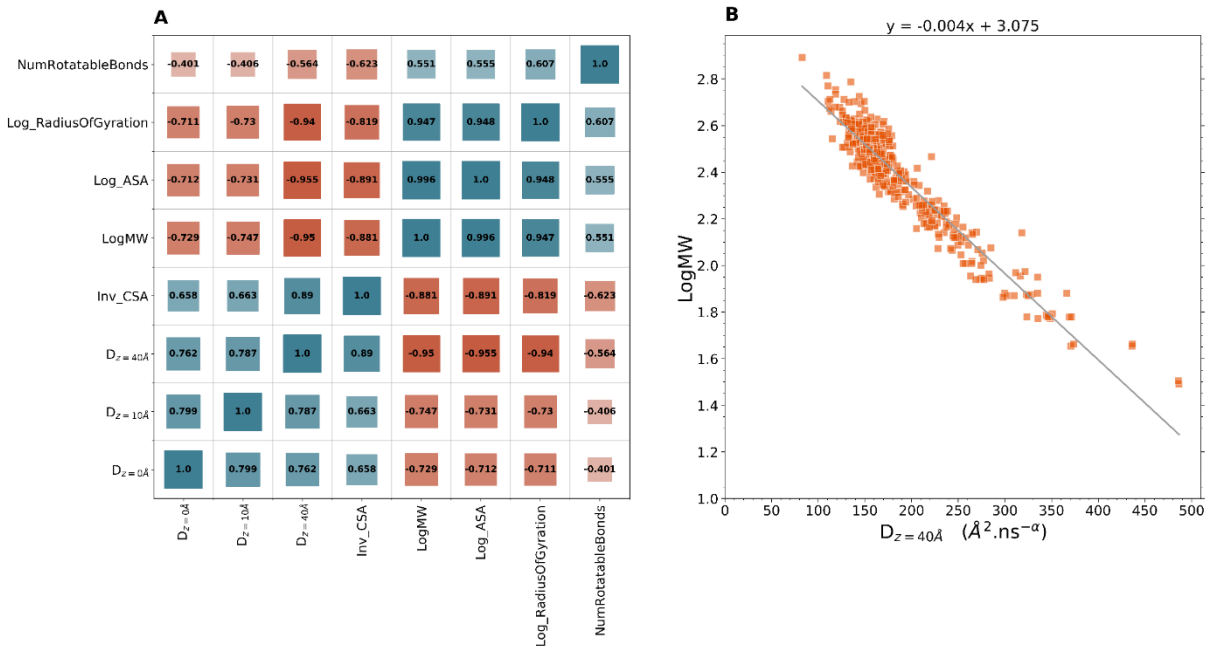


Figure 33: (A) Correlation matrix between diffusivity values (in water, within the membrane and at the membrane center) and different molecular descriptors.

Strong negative correlation was highlighted between the diffusivity in water ($D_{z=40A}$) and logMW (Figure 33A and B) with a correlation coefficient R = -0.95. Less, but still important negative correlation was found between logMW and the diffusivity in the region of the lipid chains ($D_{z=10A}$, R = 0.75) and at the membrane center ($D_{z=0A}$, R = 0.73). Both the logarithm of the radius of gyration and the logarithm of the accessible surface area (logASA) were highly correlated to the MW and thus they also exhibited high correlations with the diffusivity values (Figure 33). On the other hand, the inverse of cross-sectional area was strongly correlated to $D_{z=40A}$ with R = 0.89, but with less correlation with the diffusivities in the membrane including the membrane center. The number of rotatable bonds exhibited relatively weak correlation with all diffusivity values. Given the abovementioned results, the diffusivity in water can be predicted from MW with high confidence. Conversely, diffusivity within the membrane appears to be multifactorial.

The impact of the diffusivity on the calculated permeation coefficients was investigated by computing $\log P_{erm}^{MemCross}$ using both classical and fractional diffusivity profiles (Figure S12A-B). The only significant difference was a decrease of the intercept of the linear regression equation (-3.30 and -3.93 log units for fractional and classical models, respectively). Consequently, by comparing $\log P_{erm}^{MemCross}$ values using both the classical and the fractional diffusivities (Figure S12C), a very good correlation (R^2 = 0.997) was highlighted. This arises from equation 1, where the exponential term of the PMF drives the value of P over the linear proportionality between P and diffusivity. Although the fractional model provides a more accurate description of the dynamics within the lipid bilayer, the classical diffusivity model can be used without losing accuracy by adding a corrective term of 0.69 log units to $\log P_{erm}^{MemCross}$ (Figure S12C). Furthermore, $\log P_{erm}^{MemCross}$ calculated using a constant value for diffusivity (10 $Å^2.ns^{-1}$) yielded an excellent correlation to $\log P_{erm}^{MemCross}$ obtained with fractional diffusivity (y = 0.98x – 0.77, R^2 = 0.994, Figure S12D). Therefore, carefully defining a constant value of the diffusivity may

appear almost as accurate as using the fractional diffusivity, since $\log P_{erm}^{MemCross}$ is (exponentially) more dependent on the PMF than on the diffusivity.

VI.3.4. Benchmarking against experimental permeability coefficients

To assess the ability of this AWH-based protocol at predicting experimental results, we compared the $\log P_{erm}^{MemCross}$ values obtained here with the $\log P_{erm}^{exp}$ values issued from 4 different datasets corresponding to different experimental methods split as follows, 114, 254, 168 and 146 obtained from the phosphatidylcholine-based (PC-based) method, the PAMPA-DS system, the BBB cells and the Caco2/MDCK protocol, respectively.

VI.3.4.1. PC-based methods

An exclusion criterion was defined as follows, a $\log P_{erm}^{MemCross}$ standard deviation σ greater than half the root mean square error of the original dataset. According to this criterion, two compounds, phloretin and loperamide, were excluded from the dataset, i.e., $\log P_{erm}^{MemCross}$ values were poorly reproduced. Taking these two exclusions into consideration, a linear regression model was found:

$$logP_{erm}^{f} = 0.98\, logP_{erm}^{th} - 3.3 \qquad (7)$$

With a coefficient of determination of $R^2 = 0.69$, and RMSE = 3.1 log units. The correlation for the whole set of molecules highlighted all sugars as outliers namely D-glucose, D-fructose and sucrose (see Figure S12A). This suggests that the Gaff2 force-field may not be well suited for carbohydrates. To better describe the correlation between $\log P_{erm}^{MemCross}$ and $\log P_{erm}^{exp}$, a linear regression step was performed after excluding these three sugar molecules. This yielded the following linear regression,

$$\log P_{\mathrm{erm}}^{\mathrm{f}} = 0.94 \log P_{erm}^{MemCross} - 3.23 \qquad (8)$$

With a slight increase in the correlation coefficient to $R^2 = 0.71$, and a decrease in the RMSE to 2.62 log units (Figure 34A).

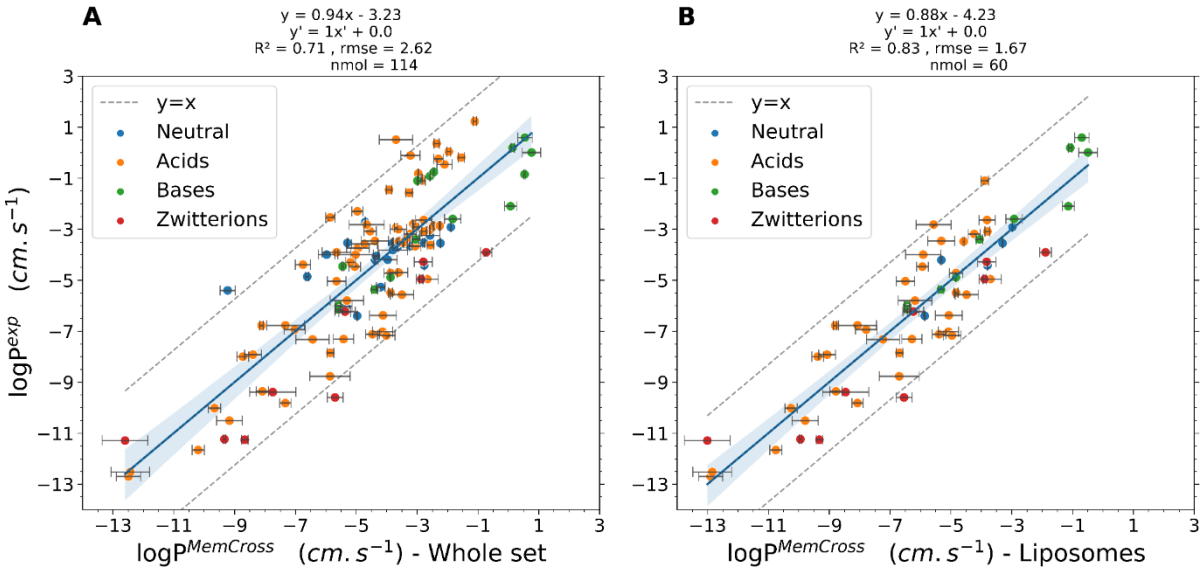


Figure 34: Fitted linear regression models between $logP_{erm}^{MemCross}$ and $logP_{erm}^{exp}$ for (A) the whole set of liposomes/BLM membrane measurements and (B) the set of liposomes measurements for different molecule types: neutrals, acids, bases and zwitterions.

Given the variability of the different experimental conditions, we then compared $\log P_{erm}^{MemCross}$ values to the experimental ones for measurements using only liposomes (Figure 34B). This provided the following regression:

$$\log P_{erm}^{f} = 0.88 \log P_{erm}^{MemCross} - 4.23 \quad (9)$$

With a good correlation coefficient R²=0.83 and low deviation error from the mean (RMSE =1.67 log units) (Figure 34A).

Interestingly, both data sets exhibited negative intercept values (b = -3.23 and -4.23 log units for the whole set and for liposomes only, respectively). This means that $\log P_{erm}^{MemCross}$ values were overestimated compared to $\log P_{erm}^{exp}$ (i.e., underestimation of the free energy barriers). Several factors can rationalize this discrepancy including differences in membrane composition temperature, inherent force-field or protocol-based errors, or impact of diffusivity at different scales [22]. Another plausible hypothesis to the underestimation of the free energy barrier within the membrane is the underestimation of the partial charges of the permeants.

VI.3.4.2. PAMPA-DS

The MemCross permeation coefficients ($\log P_{erm}^{MemCross}$) were evaluated for compounds tested with the more complex PAMPA-DS membrane systems. Here, we considered only intrinsic permeability values, as reported by Avdeef after corrections accounting for permeability through the unstirred layers (USL) and all non-passive transport.[14,137,165]

The linear regression obtained was:

$$\log P_{erm}^{f} = 0.77 \log P_{erm}^{MemCross} - 4.68 \quad (10)$$

With R² = 0.66 and RMSE = 2.39 log units.

The correlation was slightly weaker than for more simplistic membrane models (liposomes and BLM), which can partly be rationalized by the imperfect correlation between both experimental methods (RMSE = 1.75 log units).[22]

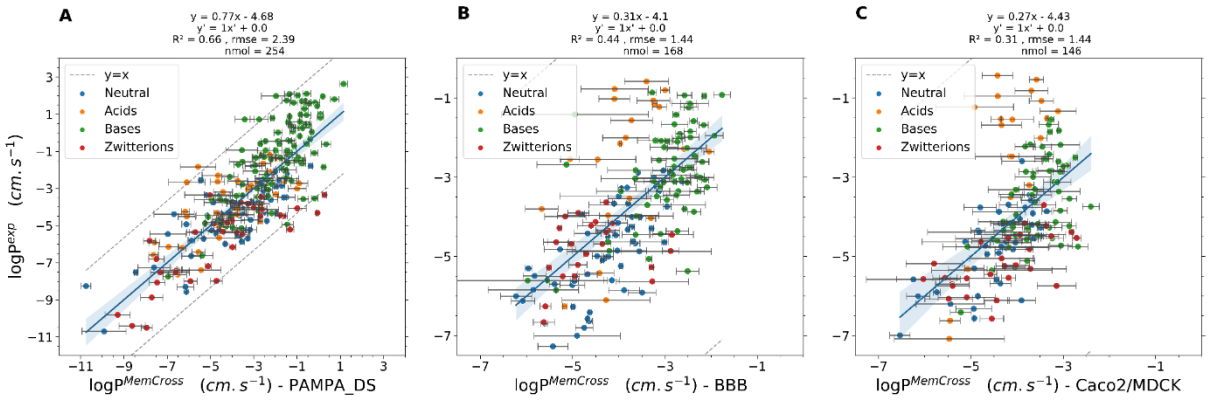


Figure 35: Fitted linear regression models between $logP_{erm}^{MemCross}$ and $logP_{erm}^{exp}$ for (A) the set of PAMPA-DS and (B) Blood Brain Barrier, and (C) Caco2/MDCK measurements for different molecule types: neutrals, acids, bases and zwitterions.

VI.3.4.3. Cell-based assays

The linear regression for BBB was:

$$\log P_{\mathrm{erm}}^{\mathrm{f}} = 0.31 \log P_{erm}^{MemCross} - 4.1 \qquad (11)$$

With R^2 = 0.44 and RMSE = 1.44 log units.

For Caco2/MDCK, the linear regression was:

$$\log P_{\mathrm{erm}}^{\mathrm{f}} = 0.27 \log P_{erm}^{MemCross} - 4.43 \qquad (12)$$

With R^2 = 0.31 and RMSE = 1.44 log units.

The correlations of $\log P_{erm}^{MemCross}$ with BBB and Caco2/MDCK (Figure 35B-C) yielded low correlation coefficients and low linear regression slopes. Since MemCross can achieve relatively accurate predictions for simple systems such as liposomes, the most probable explanation of the poorer performance on cell-based assays lies in the higher complexity of the experimental models. Indeed, cells may have different membrane compositions, paracellular transport, and most importantly active transport. This is in agreement with the imperfect correlations found between experimental methods.[22,137] For instance, the experimental values of permeation coefficients of a few acids in BBB and Caco2/MDCK measurements were overestimated compared to the predicted $\log P_{erm}^{MemCross}$ values (Figure 35B-C), whereas the acids were not outliers in simpler systems. This could be due to the existence of active membrane transport specific for these acids in cells. Interestingly, the RMSE values (1.44 log units) were lower than for PAMPA-DS and even liposomes/BLM. Overall, this means that MemCross can be used to obtain coarse estimates of the permeation coefficients on cell-based models for non-acidic molecules.

VI.3.5. Source of the error between theory and experiments

To rationalize the sources of error between $\log P_{erm}^{MemCross}$ and $\log P_{erm}^{exp}$, we evaluated correlated the residual errors between both with respect to different descriptors. Only the dataset of liposomes and BLM measurements was used, as they yielded the best correlation to theoretical predictions and the errors observed when comparing to the cell-based methods were already discussed above. The descriptors used were the mean PMF barrier at the membrane center, the mean standard deviation of the whole PMF profile from three replicas (on the whole profile and at the membrane center), the numbers of aliphatic hydroxyl and carboxylic acid moieties, the number of OH and NH moieties, the number of rotatable bonds, and the MW (Figure 36).

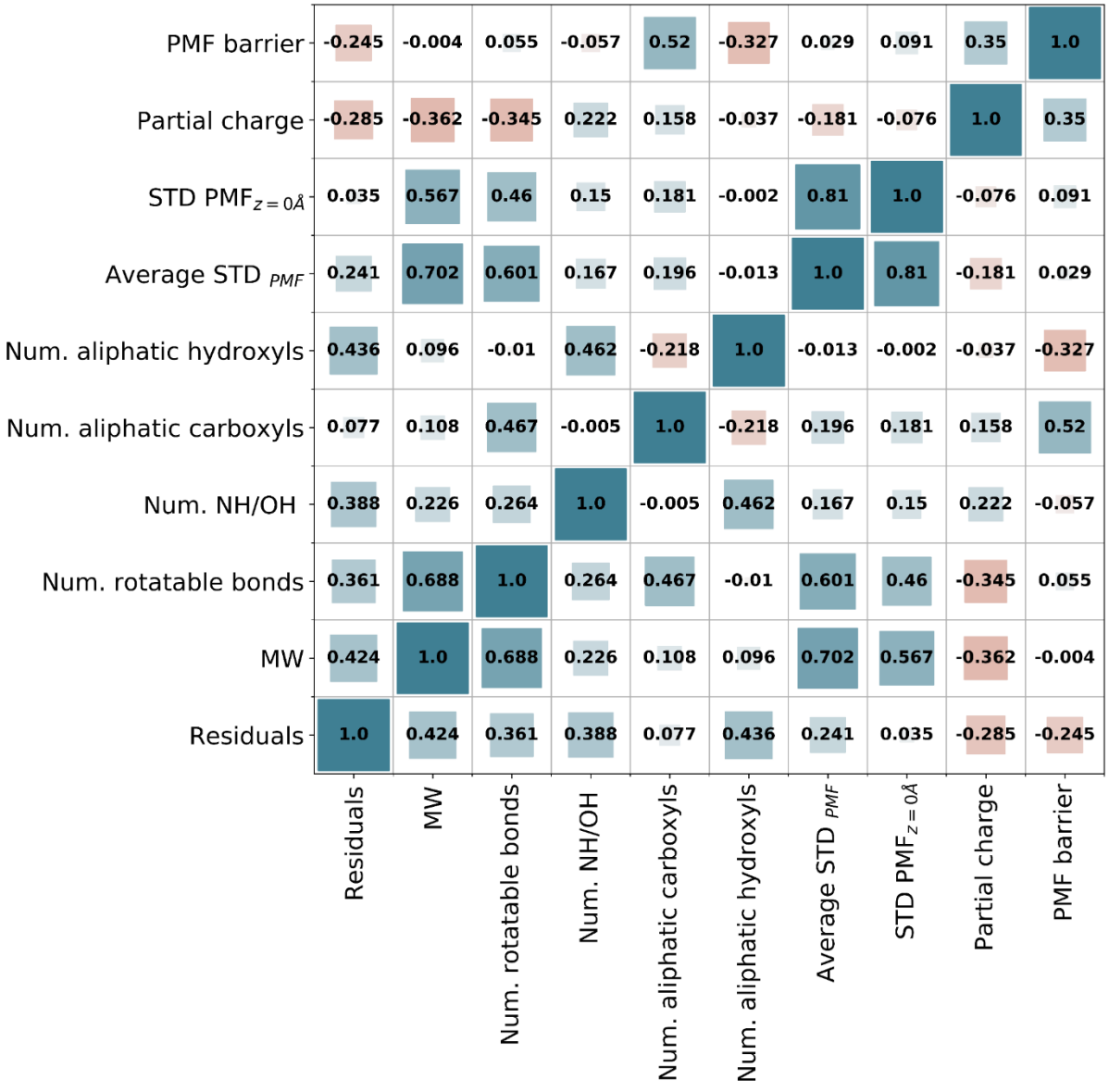


Figure 36: Correlation matrix of the $\log P_{erm}$ residual and different molecular descriptors and PMF-linked parameters.

Weak positive correlations between MW, number of rotatable bonds, number of aliphatic hydroxyl groups and number of polar NH and OH groups were noticed (R = 0.42, 0.36, 0.43 and 0.39, respectively). However, no single descriptor was identified to fully rationalize the residual error, confirming that the error is multifactorial. Multivariate linear combination models were probed using all possible combinations between the parameters (except the ASA parameter as it is correlated to MW), see Figure 36. The quality of each of these models was assessed by computing the Akaike information criterion (AIC), the lower the AIC, the better the model. The whole Liposomes/BLM dataset was randomly split into the training set (80% of molecules) and the test set (20% of molecules). The multivariate model (on the training set) corresponds to the lowest AIC value and a combination of the following 6 parameters: MW, number of NH/OH groupments, number of aliphatic carboxylic acids, number of aliphatic hydroxyl groups, standard deviation of PMF at z = 0, partial charge and the height of the PMF barrier:

$$\begin{aligned}\mathrm{Residual}_{predicted} = {} & 0.004\ \mathrm{MW} + 0.22\ \mathrm{NHOHCount} + 0.95\ \mathrm{fr}_{\mathrm{AI_{COO}}} + 0.37\ \mathrm{fr}_{\mathrm{AI_{OH}}} \\ & -1.04\ \mathrm{STD}_{\mathrm{PMF}_{z=0\text{Å}}}, -4.03\ \mathrm{Partial}_{\mathrm{charge}} - 0.05\ \mathrm{PMF}_{\mathrm{barrier}} + 2.94 \qquad (13)\end{aligned}$$

With AIC = 270.9 and R^2 = 0.47.

Applying equation (13) to the test set resulted in a very poor prediction correlation between the predicted and actual residual errors (R^2 = 0.001, Figure S13). In other words, the best multivariate model was not able to predict the residual error. This suggests that the error is very difficult to predict, and that it depends on other parameters, which might be numerous.

VI.3.6. Comparison to another theoretical model

In order to improve the benchmarking of MemCross, the $\log P_{erm}^{MemCross}$ values were compared betwenn MemCross and a recently developped physics-based method (PerMM) by Lomize *et al*[22]. Here, for the correlations to each experimental dataset, we investigated whether common outlier molecules were identified. As previously described, a molecule was considered as an outlier if it was outside the linear regression line by more than twice the residual standard deviation. Outliers common to both theoretical methods were excluded from the benchmark, as an issue with the experimental measurment is plausible. Outliers found in only one method were kept, as in this case, the discrepancy is more likely attributed to the theoretical model. For PAMPA-DS data set, ganciclovir was highlighted as a common outlier, hence it was ommited. Finally, for the BBB and Caco2/MDCK data sets, ascorbic acid, neotrofin, ganciclovir and terfenadine were identified as common outliers.

Correlations between $\log P_{erm}^{exp}$ and $\log P_{erm}$ from MemCross and from PerMM on the liposomes/BLM data set are shown in Figure S14A-B. When considering the whole data set, the correlation coefficients were 0.71, 0.79, and RMSE 2.62 and 1.90 for MemCross and PerMM, respectively. Therefore, the performance of PerMM was slighly higher. Conversely, regarding the liposome-only data subset (Figure S14C-D), MemCross correlation was slighlty better (R^2 = 0.83, RMSE = 1.67) compared to PerMM (R^2 = 0.80, RMSE = 1.93).

Regarding the PAMPA-DS dataset, MemCross exhibited a slighly better correlation (R^2 = 0.66, RMSE = 2.38) compared to PerMM (R^2 = 0.59, RMSE = 2.83, Figure S15A-B). The slope of regression was lower for PerMM (0.59) than for MemCross (0.79). A plausible explanation comes from the smaller integration interval used in PerMM for PAMPA-DS (between -15 and 15 Å). Reducing the integration interval results in a higher estimated $\log P_{erm}$ for lipophilic

molecules (i.e., with a low free energy barrier at the membrane core), which can flatten the linear regression.

As for cell-based assays, both MemCross and PerMM exhibited low correlation coefficients (0.47 and 0.57, respectively for BBB, 0.29 and 0.34, respectively for Caco2/MDCK, see Figure S15C-F). RMSE and slopes were also very similar between both methods. This confirms the difficulty for theoretical methods to catch the entirety of the permeation process occuring in cell membranes.

Overall, comparable correlation coefficients and RMSE values were obtained with both methods, MemCross being slightly more accurate for liposomes and for PAMPA-DS. It is clear that the more complex the membrane, the weaker the correlation. Therefore, the *in silico* models should match as close as possible to the experimental conditions (e.g., lipid composition, temperature, ion concentrations, pH), so that to improve the accuracy of the predicted $\log P_{erm}^{MemCross}$. It should also be noted that the computational cost of PerMM is much lower than for MemCross. However, MemCross offers a full atomistic description of the molecule-membrane interaction, as well as complete diffusivity profiles, and its computational time could also be lowered with multiple walkers and replica-exchange AWH for instance.

VI.4. CONCLUSION

This work is a benchmark study to validate the robustness of a novel computational methodology, named MemCross, to estimate membrane permeation coefficients. Both PMF and diffusivity profiles along the permeation path through the membrane were computed for more than 400 xenobiotics, mostly drugs. To the best of our knowledge, this is the largest study performed with drug-like database, while performing all-atom MD simulations, totaling more than 1.2 millisecond simulation time. MemCross accurately reproduced experimental data obtained with simple biomimetic models, such as liposomes and BLM. This success was attributed to the similarities between the *in silico* POPC lipid bilayer model used in this study and these PC-based *in vitro* membrane models. It is worth mentioning that due to variability in experimental protocols, the membrane permeability coefficients are experimentally relevant within an error range of one order of magnitude.[400] MemCross also exhibited a good correlation with PAMPA-DS experiments, although some discrepancies could be rationalized by the difference in membrane composition and the contribution of aqueous boundary layers. The comparison to cell-based assays appeared more limited, which is likely due to the presence of other transport types such as active membrane and para-cellular transports that are competing with the passive diffusion process. Such discrepancies have also been reported with other theoretical models.[22] Overall, the correlations obtained with MemCross were comparable to PerMM, a physics-based model developed to predict permeation coefficients.

The simulation time per replica required to get a converged estimate of the permeation coefficient was 300 and 900 ns for molecules with a molecular weight lower than 100 $g.mol^{-1}$ and in the range 100-400 $g.mol^{-1}$, respectively. Apart from sugars, the PMF, the diffusivity and the permeation coefficients of all chemical scaffolds can be readily modelled by MemCross. Although the fractional diffusivity model provides the best correlations with experimental data, however we highlighted a weak dependence of the permeation coefficient on fractional, classical, or even constant diffusivity profiles.

To summarize, MemCross is an affordable and useful tool to predict permeation coefficients. It provides a full atomistic description of molecule-membrane interactions, which is very desirable in the pharmacological context targeted by pharmaceutical industry.

Chapter VII. Thorough evaluation of binding sites in the human ABCC4/MRP4 Alphafold2 structure

Mehdi Benmameri, [†] Benjamin Chantemargue, [‡] Patrick Trouillas, [†¶] and Gabin Fabre*[†]

[†]INSERM, UMR 1248, F-87000 Limoges, France;

[‡]INSILIBIO, 1 avenue Ester Technopôle, 87069 Limoges Cedex, France;

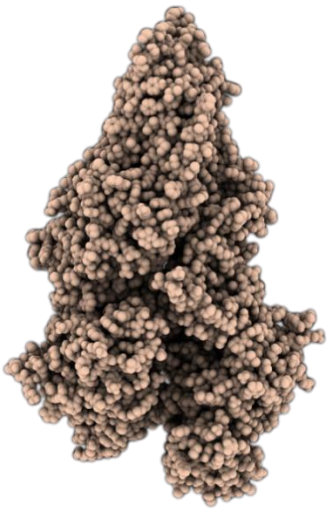
[¶]CATRIN RCPTM, 779 00 Olomouc - Holice, Czech Republic

Foreword

In this study, we investigate the potential binding sites in the human MRP4 protein of a range of MRP4 substrates and modulators. The AlphaFold2 protein structure was equilibrated during an MD simulation step, then a thorough docking procedure was established on the whole protein. Here, binding sites, molecules entry and key residues that are involved in the binding are discussed. This study will be submitted in the near future.

VII.1. ABSTRACT:

The human Multidrug Resistance Protein 4 (*h*MRP4) is a membrane protein, of clinical importance, responsible for the active efflux of large number of xenobiotics from inside the liver and kidney cells. The uptake of concomitant drugs and/or nutrients may favor drug-drug or drug-nutrient interactions that can lead to several outcomes such as cytotoxicity caused by drug accumulation inside the liver and kidney cells. No experimental structure was reported for *h*MRP4 so far. Here, we investigate the binding sites in the *h*MRP4 structure, issued from AlphaFold2, using thorough molecular docking of 129 known substrates and inhibitors. We highlighted five major regions, the chamber at the central part of the protein being the most favorable binding site for all ligands. Key residues for ligand binding in the different regions were listed. Connections between the regions through iso-energetic paths were studied, highlighting gating routes to the main chamber.



Keywords: Multi-Drug Resistance Protein 4 (MRP4), Membrane Transporters, Molecular Docking, MD simulations, Binding sites.

VII.2. INTRODUCTION

The ATP-binding cassette (ABC) transporters constitute one of the largest superfamily of membrane proteins in all living organisms.[401] They are involved in the translocation of a wide range of substrates including endogenous compounds, xenobiotics, metabolites, ions or macromolecules.[402,403] In particular, they are known to efflux drugs across biological membranes in the human body. They play a substantial role in maintaining cellular homeostasis by favoring outer-to-inner cell exchanges in various organ tissues. Due to their diversity and low specificity, they play a key role in the pharmacokinetics of xenobiotics.

In eukaryotes, ABC transporters typically ensures the unidirectional substrate efflux across the membrane.[404–407] They are composed of four distinct domains, namely two cytoplasmic nucleotide binding domains (NBDs) and two helical transmembrane domains (TMDs).[408] Each of the NBDs is responsible for the binding of one ATP molecule. The ATP hydrolysis provides the energy required for conformational changes during the transport cycle. The NBDs of the ABC membrane transporters are characterized by seven highly conserved motifs, named A-loop, Walker A, Q-loop, the signature motif or C-loop (LSGGQ), Walker B, D-loop and H-loop.[409,410] The TMDs are composed of twelve transmembrane helices (TMHs) and represent the most variable sequences between ABC transporters. They confer the specificity of substrate binding to most of these proteins.[411]

In eukaryotes, apart few exceptions,[412–415] ABC proteins are all efflux transporters.[416] Dysfunctions in the efflux process ensured by these proteins are often associated with pathological outcomes. First, these dysfunctions can be attributed to mutations in the ABC coding genes, which are associated with various diseases (*e.g.,* cystic fibrosis[417], Tangier disease[418], obstetric cholestasis[419], and Scott syndrome[420]). A second type of dysfunction may come from the overexpression of these transporters. To follow these dysfunctions has gained a growing interest in the treatment of cancer, as they can be responsible for multidrug resistance (MDR), in turn leading to treatment failure.[421,422] This makes these transporters are known as multidrug resistance proteins (MRPs).[423–425] Among ABC transporters, breast cancer resistance protein (BCRP), P-glycoprotein (P-gp, MDR1) and Multi-Drug Resistance Proteins (MRPs, ABCC family) play an important role in the disposition and pharmacokinetics of a wide range of drugs (*e.g.,* antibiotics, immunosuppressants) thanks to their ubiquity/versatility (*i.e.,* in the liver, the kidney, the intestine and the brain).[426,427] Drug-drug interactions (DDI) can also be associated with dysfunctions of these transporters, leading to the accumulation of xenobiotics inside cells, hence causing cytotoxicity. DDI may occur when various xenobiotics compete in the binding sites of the ABC transporters.

The human Multidrug Resistance Protein 4 (*h*MRP4) is a member of the ABCC subfamily. It exhibits a versatile localization in different organs, in particular in the liver and the kidneys.[428–430] In basolateral membranes of the hepatic cells, *h*MRP4 is involved in the export of substrates from the liver cells to the portal blood. In the kidneys, it is responsible for their excretion from the renal cells into urine across the apical membrane of renal proximal tubule cells.[428,431–433] The *h*MRP4 exports diverse organic (sometimes aromatic) anions such as endogenous cell signaling molecules (*e.g.,* cAMP), but also immunosuppressants and antiviral drugs (*e.g.,* methotrexate, tenofovir) and glucuronide conjugates (*e.g.,* mycophenolate glucuronide).[428,434–436] In *h*MRP4, as for other eukaryotic MRPs, the two NBD are not symmetrical: NBD2 contains the canonical site (NBS2) which is the main ATP hydrolysis site, while NBD1 contains the degenerate site (NBS1) and strongly binds to the ATP altering its hydrolysis.[437] Several DDI were reported for *h*MRP4, for example between methotrexate, one of the most prescribed

immunosuppressant drugs to prevent kidney graft rejection, and anti-inflammatory drugs (NSAIDs), which slow down the export of the former[438,439]. This leads to methotrexate accumulation in cells and subsequent cytotoxicity. Therefore, DDI represent a plausible cause of graft rejection.[341,440–442]

Understanding the structural patterns and binding modes of *h*MRP4 ligands (substrates and modulators) is of utmost clinical importance to decipher its transport cycle and rationalize DDI. *In vitro* studies showed a series of key residues involved in *h*MRP4 function, namely PHE368, PHE369, GLU374, ARG375, GLU378, TRP995 and ARG998.[435,443,444] However, to the best of our knowledge, there is no experimentally resolved structure for hMRP4 yet. Various *in silico* *h*MRP4 3D models were proposed based on homology modeling approaches. The main drawback of these models is the low sequence identity between *h*MRP4 and other MRPs of the same subfamily, leading to possible structural errors in key domains. Recently, AlphaFold[445,446] provided a new structure that outperforms the quality of the previous homology models. Here, we use the *h*MRP4 model from AlphaFold to perform a thorough molecular docking then to establish a comprehensive list of the binding sites, based on the mode of binding of 129 known substrates and modulators of hMRP4.

VII.3. METHODS

VII.3.1. *h*MRP4 structure model

The inward facing (IF) conformation of the wild-type human MRP4 structure was obtained from the AlphaFold 2 protein structure database.[445,446] The model confidence assigned by Alphafold is “confident” or “very high” for a substantial part of both TMD and NBD domains (Figure S16). This corresponds to 1172 residues. However, the confidence is lower for the flexible loops, in particular for the experimentally unresolved linker (L1), i.e., the GLY624-PHE698 sequence in between the first nucleotide-binding domain (NBD1) and the second transmembrane domain (TMD2). The linker lies inside the chamber of the protein, which is a solvent-accessible area, close to the inner polar head groups. The dihedral angles of the protein backbone model were checked using Procheck[447] Ramachandran diagram (Figure S17).[448,449] About 92.4% of residues were in most favored regions of the Ramachandran plot, confirming a good quality model (> 90%). Additionally, 99.5% of residues were in the most favored or allowed regions (Figure S17). The ERRAT server estimated the overall quality factor to 96.8%, equivalent to a good high-resolution structure (> 95%).[449,450] The protonation state of the HIS residues at pH=7.4 was verified using ProteinPrepare server.[451,452] This yielded the following protonation states: δ-nitrogen protonation (HID) was assigned to HIS: 35, 55, 152, 153, 158, 172, 411, 466, 583, 592, 710, 798, 831, 930, 1037, 1111, 1141, 1225, 1233, 1299, 1302 and 1305, and ε-nitrogen protonation (HIE) was assigned to HIS 213, 472, 572, 903, 934 and 1060. The orientation of the protein inside the lipid bilayer core was predicted using PPM3[453]. Then, it was embedded into a POPC:CHOL (3:1) membrane using CHARMM-GUI membrane builder[350]. This membrane model was thus made of 315 POPC molecules and 105 cholesterol molecules. The whole system (protein + membrane) was solvated with water. Then, Na^+ and Cl^- ions were added to neutralize the system with a physiological concentration of 0.154 M. Two ATP molecules, each associated with one Mg^{2+} ion each, were added in their corresponding nucleotide binding sites. To do so, the NBD structures of hMRP4 and bovine MRP1[454] (PDB ID 6UY0) were aligned.[437]

VII.3.2. Setup of molecular dynamics simulations

MD simulations were performed using the Amber20 software[455]. The Lipid17 force field[456] was used for the lipids, Amber FF14SB[457] for the protein residues and TIP3P[458] for water molecules. ATP force field parameters were gathered from the Amber Parameter database[459]. The modified DNA.OL15[460,461] force field was used to the ATP molecules . During the thermalization step, the distance between different atoms of the ATP, Mg^{2+} and the ATP-binding site were restrained using a harmonic potential as suggested in Wen *et al* [462] and Tóth *et al* [415]. Periodic boundary conditions were applied in the x, y and z directions.

Electrostatic and van der Waals (vdW) interaction cutoffs were set to 10 Å. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald (PME)[463] method. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm. The time step was set to 2 fs.

The system was thermalized in two steps. First, from 0 to 100 K in the NVT ensemble for water and ions. Then, up to 310 K in the NPT ensemble for the whole system. Finally, in order to allow relaxation of the protein's side chains, the whole molecular system was equilibrated in 3 replicas for 1 µs each. The temperature was maintained at 310 K during the production runs using Langevin thermostat.[464] Berendsen barostat[465] was also used to maintain the pressure at 1 atm.

VII.3.3. Selection of reliable protein conformations for docking

Representative protein structures were extracted from the MD simulations based on the free energy landscape with the InfleCS clustering method[466] . The InfleCS method is based on a Gaussian mixture free energy estimator (GMM). Free energy-based clusters were defined using the following parameters: NBD distance, NBD twist, cavity volume of the central protein chamber and the root mean square displacements (RMSD) of each ARG residue of the protein chamber in contact with water. ARG residues were selected for this analysis because they exhibited more flexibility than other aminoacids. The density of states was estimated with an 80×80 grid and 5 iterations. Up to 16 gaussian components were allowed for Gaussian mixture model.

First, a one-dimensional clustering was carried out for each variable. Three variables exhibited at least two energy minima: NBD twist, NBD distance, and the RMSD or ARG998 (Figure S18). ARG998 was selected because it was identified as a key residue of MRP4,[443] and because its RMSD exhibited two separated clusters, contrary to other ARG residues in the protein chamber (Figure S18). These variables were selected for further 2D clustering, yielding 25 clusters (Figure S19). For the sake of limiting the calculation time, 13 clusters were selected (with probability greater than or equal to 0.03), which provided 13 conformations for further docking, thus mimicking a partially dynamic docking. For each of them, the conformation corresponding to the cluster center was selected as a representative protein structure for docking calculations.

VII.3.4. Molecular docking

A thorough docking was performed on the 13 protein structures using Autodock Vina 1.2 software.[467] 26 overlapping cubic grid boxes, each 29Å wide, were needed to cover the whole protein (Table S3)., to cover the whole protein. The overlap between neighboring boxes was 10 Å. For each ligand and for each grid, 10 independent replicas, and the exhaustiveness

parameter set to 40, ensured a thorough exploration of the conformational space. Up to 20 poses were saved for each docking calculation, with a maximum energy range of 3 kcal.mol^{-1} with respect to the best pose. This extensive procedure generated more than 8 million poses.

In order to normalize the number of poses in each docked area of the protein structure, among poses that were completely inside the common area between overlapped grids, only the poses of one of the grids were kept. Following this procedure, about 2 million poses were omitted.

Additionally, an ATP molecule was also docked to both Nucleotide Binding Sites (NBS1 and NBS2) using two grids of 29Å wide (Table S3) and 20 replicas each.

VII.3.5. Protein structures

The PDBQT files of the *h*MRP4 structures were generated using AutoDock tools. The selected structures were considered rigid during the docking runs. Although rigid docking does not allow the sampling of rotameric states during the docking procedure, the use of 13 structures from different clusters allowed mimicking a dynamic docking while drastically reducing the computational cost.

VII.3.6. Ligands

*h*MRP4 substrates and modulators were gathered from the CHEMBL database.[468,469] Only the ligands with available standard values (Km, IC_{50}, Activity and Inhibition) were selected. Ligands with Km and IC_{50} values upper than 100 µM were then omitted, because these ligands are likely to be neither substrates nor inhibitors of MRP4. ceefourin 1 and 2, which are specific inhibitors or MRP4,[345] were added to the dataset. Eventually, 129 ligands remained after this selection process. Ligands were divided into three groups, namely substrates (11), inhibitors (112) and substrates/inhibitors (6) (see Table S4). The PDBQT files of ligands were generated using Open babel software[470] from the their corresponding smiles. The protonation states of ligands were set with respect to pH = 7.4. 3D conformations were generated and minimized with the GAFF force field using Open babel, and then saved as PDBQT files.[471]

VII.3.7. MOLE channels analysis

Gating channels in a representative structure of *h*MRP4 from MD simulations were analyzed using MoleOnline.[472] the accessible radius for the channel entry was set to 5 Å.

VII.4. RESULTS AND DISCUSSION

VII.4.1. Relaxation of the *h*MRP4 structure

The evolution of several structural parameters was followed during MD simulations. The backbone RMSD (Figure S20E) increases up to 5 Å during the first 300 ns (relaxation of the protein in the lipid bilayer). Then it stabilizes, showing the convergence towards a stable protein structure in its new environment. Prior to adding ATP-Mg molecules, the inward-facing domains of the protein tend to close,as the distance between the two NBDs decreased. . Although the time scale afforded by conventional MD simulations hinders from observing the complete transport cycle, this trend confirms that the presence of ATP molecules favors the closure towards an inward-facing closed structure, as already suggested for the alternating-

access transport cycle.[412,473] The linker (L1) exhibited a higher flexibility than the rest of the protein backbone (Figure S20). However, at the end of the equilibration stage, L1 stabilizes and resides far from residues of interest.

VII.4.2. Molecular docking highlights five binding regions

The docking procedure was applied for the 129 *h*MRP4 ligands. For each ligand, only poses having an energy of binding 1 kcal.mol^{-1} greater than the most favorable one were kept for further analyses. The selection of binding regions was based on a visual inspection of the density of centers of mass (COMs) of the molecules. Among the twelve identified binding regions, five (Regions 1 - 5) exhibited a particularly high density of ligand COMs (Figure 37A and Table S5). Most of them were located in the central region of the protein. Other seven minor regions (Regions 6-12) were observed (Figure 37B and Table S5).

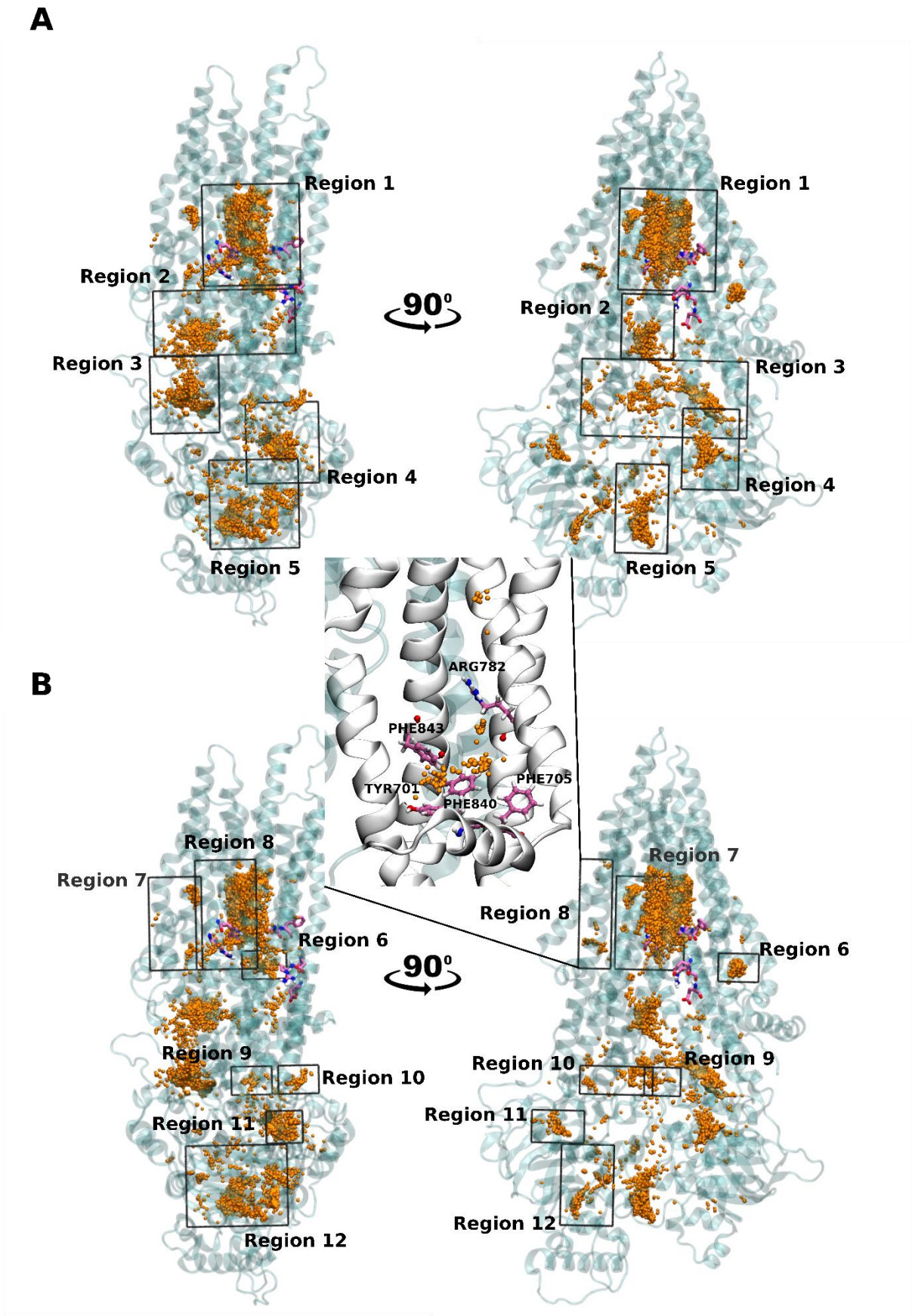


Figure 37: (A) Most important regions defined after projecting the center of masses (COM) of all ligands exhibiting a binding energy not greater than 1 kcal/mol compare to the best pose. Regions 1 -

5 exhibited high density of COMs. (B) Peripheral regions of ligand COMs (Regions 6 - 12). Region 8 contains only inhibitors and the residues involved in the binding are depicted in the top of the panel B.

Region 1 is the center of the protein chamber, which constitutes the main binding site for all ligand types. This region lies inside the inner cavity of the transmembrane core of the protein and it contains key residues which have been reported to rationalize the MRP4 protein function.[435,443,444] It exhibited the highest density of poses compared to other regions (Figure 37). The solvent-accessible part of this region 1 contains 13 % of positively charged, less than 9 % anionic, 27 % of polar and 22 % of aromatic residues, and 28 % residues with hydrophobic side chain (Figure 38 and Table S6). The relatively high proportion of cationic and aromatic residues correlates with the aromatic and anionic substrates transported by MRP4.

Region 2 is between region 1 and a part of the linker L1 (from PHE654 to SER668). It is located at the entry of the protein chamber at the top of the polar head groups in the lipid bilayer membrane. Compared to region 1, it contains a lower proportion of aromatic residues (11 %), and a higher proportion of cationic (28 %) and anionic (21 %) residues (Figure 38 and Table S6). This suggests that typical aromatic ligands may have a better affinity for region 1 than region 2, while charged ligands may bind rather in region 2. Region 2 also shares three known key residues with region 1, namely GLU374, ARG375 and GLU378.

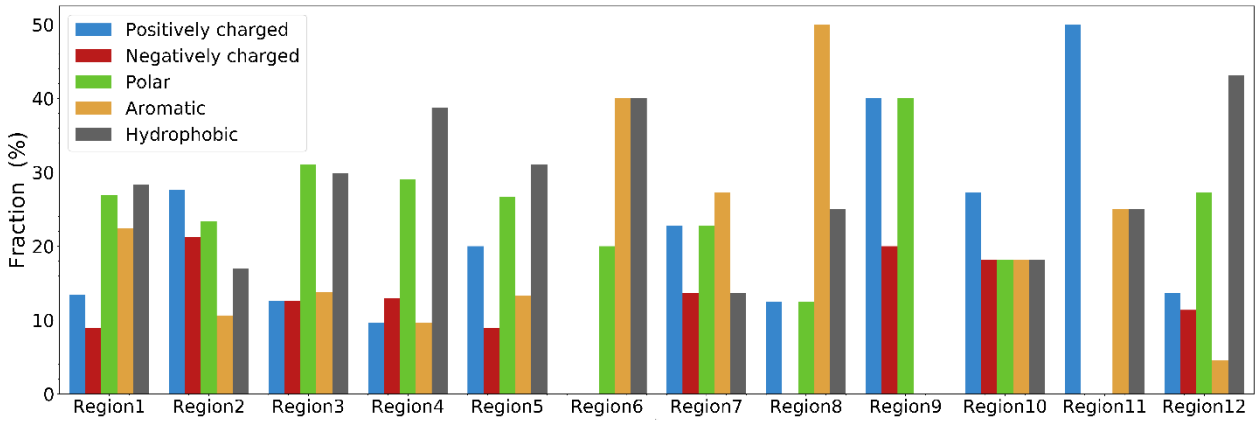


Figure 38: Fraction of positively charged (ARG, LYS), negatively charged (GLU, ASP), polar (SER, THR, ASN, GLN, CYS), aromatic (TRP, PHE, TYR, HID, HIE) and hydrophobic residues in each of the identified regions.

Region 3 lies in an outer solvent-exposed area in between TMD1 and NBD1. It contains mostly polar residues. The analysis of access channels by MOLE highlighted an entry channel (brown surface in Figure 39), which connected the path observed between Regions 1, 2 and 3. (Figure 39 and Figure 37A). Interestingly, nearly half of the ligands in this study (60 ligands: 8 substrates, 48 inhibitors and 4 substrates/inhibitors) were found to bind to all these three regions, with poses within 1 kcal.mol^{-1} of each other (Table S7). Therefore, interconnection between these three regions may consist of an iso-energetic path through which ligands can access the center of the protein chamber. Interestingly, a flexible part of linker L1 is located between regions 2 and 3. It may play a regulatory role in this potential entry channel.

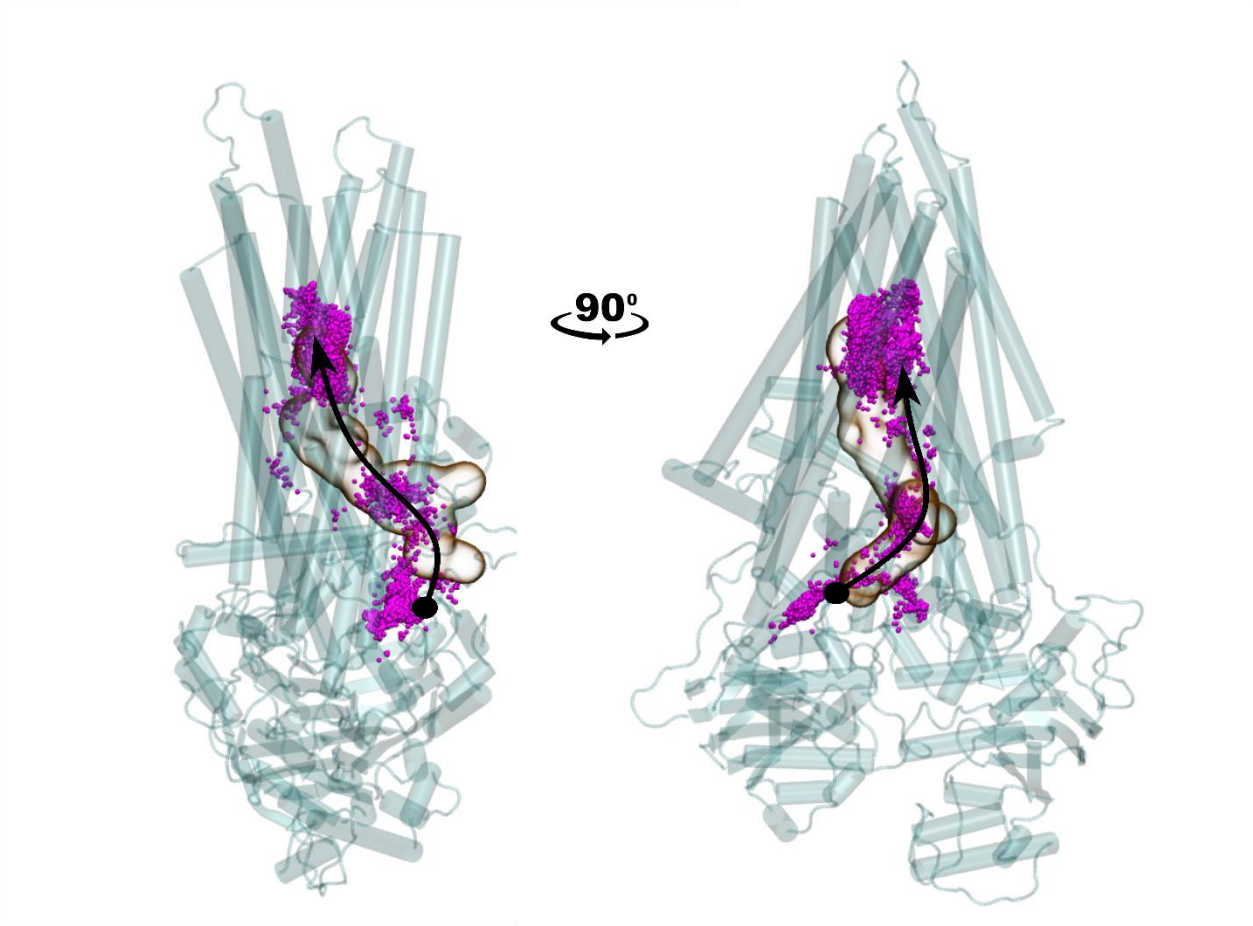


Figure 39: Pore channel predicted using MOLE (brown transparent surface) and ligand COMs from Region 3 (black circle) to Region 2 (the elbow of the arrow) to Region 1 (head of the arrow) projected on a representative hMRP4 structure.

Region 4 is located at the degenerate ATP binding site 1 (NBS1, see Figure 37A). Most of the substrates binding this region possess a purine-like moiety, such as mercaptopurine (CHEMBL1425), adefovir (CHEMBL484) and tenofovir (CHEMBL483). Likewise, inhibitors such as thioguanine (CHEMBL727), indoprofen (CHEMBL15870), nalidixic acid (CHEMBL5) also bind this region. The aromatic moiety of the bound ligands is mainly interacting with A-loop of NBS1.

Region 5 is located in between both NBDs. The proportion of cationic residues is higher than anionic ones (Figure 38). The inhibitors which are binding here could hinder the ABC closing, thus slowing down or even blocking the transport cycle.

The seven minor regions (Regions 6-12) are smaller and exhibit a lower density of ligands (Figure 37B). Region 6 lies in the vicinity of the TMD1 between TMH1, 2, 3, and 6 at the outer surface of the protein chamber. This region is isolated, as channel analysis revealed no interconnection with other regions. It lies near the lipid polar head groups. Region 6 highlighted no charged residues (Figure 39).

Region 7 and 8 are in the outer surface of the chamber, in contact with the lipid bilayer part. Interestingly, region 8 contains many aromatic residues such as TYR701, PHE705, PHE840 and PHE843 and the positively charged ARG782 which is a residue exclusive to MRP4 and

MRP5 proteins[435] (Figure 37B and Figure 39). Accordingly, region 8 binds amphiphilic molecules with both a hydrophobic (often aromatic) moiety, and a polar moiety (carboxyl, nitro or sulfonyl). Interestingly, only inhibitors bind in region 8, namely ibuprofen, ketoprofen, flurbiprofen, acitretin, undecylenic acid, entacapone and etoricoxib.

Region 10 is in the entry from the solvated open part while region 9 is more in contact with the L1 and TMH3. Region 9 does not contain any solvent-accessible aromatic residue (Figure 38). Region 10 lies between the top of NBD2, THM4 and TMH5. Regions 11 and 12 are both in the NBD2 but far from NBS2.

VII.4.3. Identifying new key residues

To identify key residues in the binding regions, the contacts between ligands and aminoacids were counted. A contact was defined as a distance lower than 5 Å between any C, N or O atoms of the ligand and the aminoacids. Since *h*MRP4 is an efflux transporter for aromatic and negatively charged substrates,[428] only positively charged and aromatic residues were considered here. Concerning the ligands, only poses having an energy of binding 1 kcal.mol^{-1} lower than the best ranked pose were taken into account. The contact counts are reported in Figure 40. Contacts considering all non hydrophobic residues are reported in Figure S21.

Key residues already identified in the literature as necessary for *h*MRP4 function (namely: TRP995, ARG998 and PHE368)[443,444] were found in contact of most ligands (Figure 40). They are located in Region 1, i.e., in the main part of the chamber.

Interestingly, we highlighted many other residues (Figure 40). Among them, ARG362, PHE324, TYR728 and LYS106 interacted strongly with known substrates, inhibitors, and substrate-inhibitors. These residues of region 1 may constitute potential targets for mutagenesis studies to further investigate their role in functioning and transport cycle of *h*MRP4.

Another interesting residue common to all ligand classes, TRP419, is located at the degenerate ATP binding site of NBS1 (Region 4), where it binds the aromatic adenosine moiety of the ATP. TRP419 is actually the A-loop, a residue known in ABC transporters to bind the aromatic moieties of purine-like substrates. However, it has been reported that, for ABC exporters, the aromatic TRP residue, in NBS1, confers strong binding to the adenosine moiety compared to the canonical site where it is replaced by the canonical TYR residue.[437] This suggests that the binding of any ligand to the degenerate site requires better binding affinity than the ATP molecule. An example of mercaptopurine bound to NBS1 and interacting with TRP419 by π-π stacking, was comparable to the binding of the ATP purine moiety (Figure S22). The docking score for mercaptopurine was -5.3 kcal.mol^{-1} in NBS1 and it was -8.4 kcal.mol^{-1} for ATP. This significant difference indicates a stronger binding of ATP compare to mercaptopurine.

TYR985 exhibited numerous contacts only with substrates, while LYS1116 was found close to only inhibitors. This may suggest the potential role of TYR985 in the substrate export mechanism. It is worth noting that the aromatic residue from Region 6 (TYR88) were found to interact only with inhibitors, which may suggest a probable allosteric site to hinder the conformational change and, hence, the transport of substrates through *h*MRP4 protein.

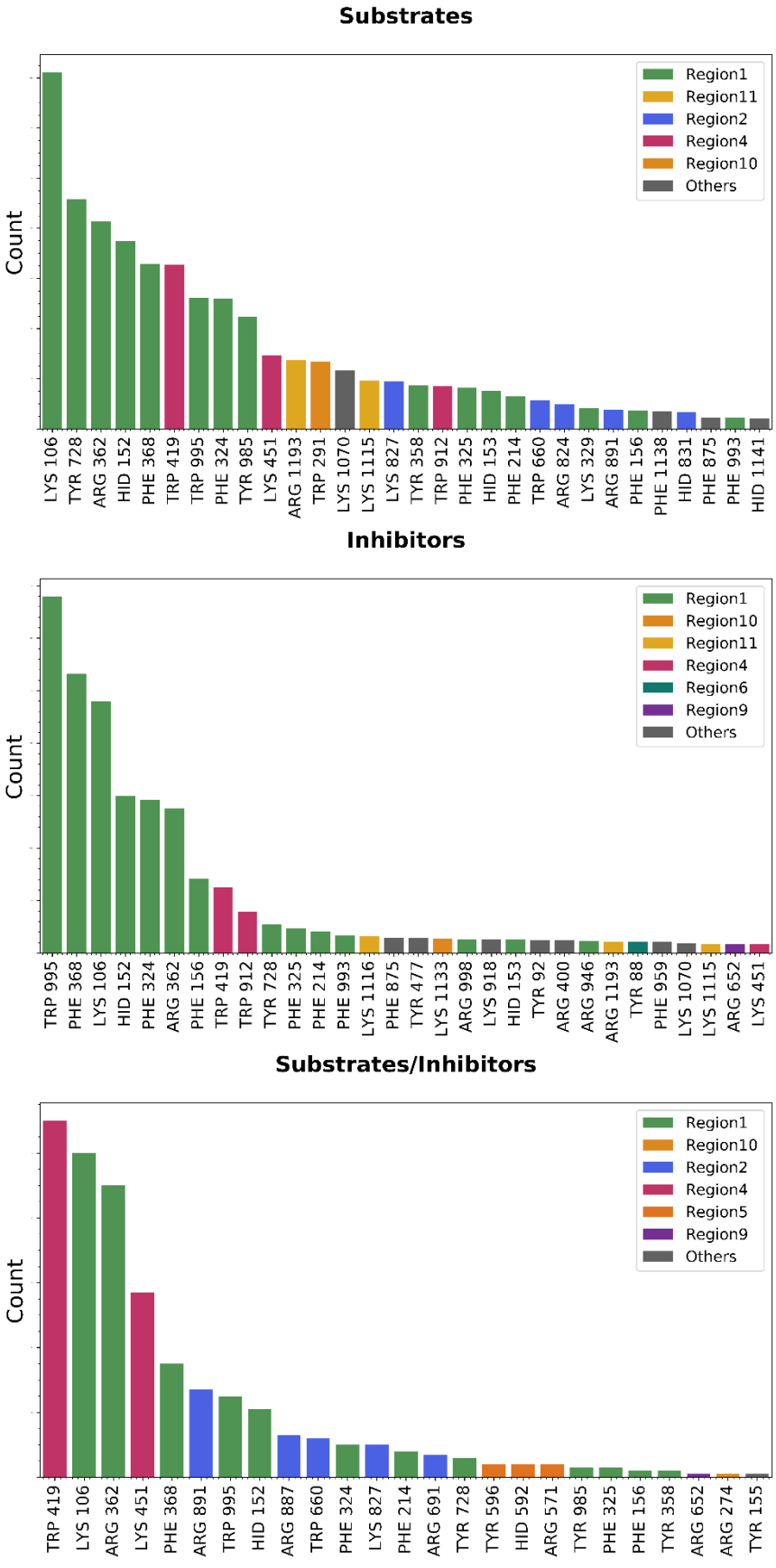


Figure 40: Count of the first 30 cationic and aromatic aminoacids in contact with the ligands (substrates, inhibitors and substrates/inhibitors) and their corresponding regions. Counts are normalized in each ligand class by the maximum number of counts among residues.

VII.4.4. Differences in binding regions for substrates and inhibitors

Comparing the binding positions of both inhibitors and substrates over the whole protein highlighted that the inhibitors can occupy all substrate binding regions (Figure S23). The contrary is not true, as a few small binding regions are specific to inhibitors (e.g. Region 8).

In order to assess whether some binding regions could be associated primarily with ligand binding for transport or inhibition, the binding of substrates, inhibitors, and substrates-inhibitors was analyzed separately (Figure 41). Region 1 exhibited the highest binding fraction for all ligand types (i.e., substrates, inhibitors and substrates/inhibitors). This correlates well with the fact that region 1 corresponds to the main binding pocket. This corresponded to 35% of substrates that are bound in Region 1, while it was the most favored region for 54% of inhibitors and 33% of substrate-inhibitors. Since the majority of inhibitors bind in this region, it is likely that the inhibitors may act via competitive inhibition.

Moreover, estradiol 17β-D-glucuronide, a known MRP4 substrate, as well as several ligands (e.g., clofazimine, sildenafil, and estradiol-disulfate) were highlighted to bind only in region 1 (see Table S8).

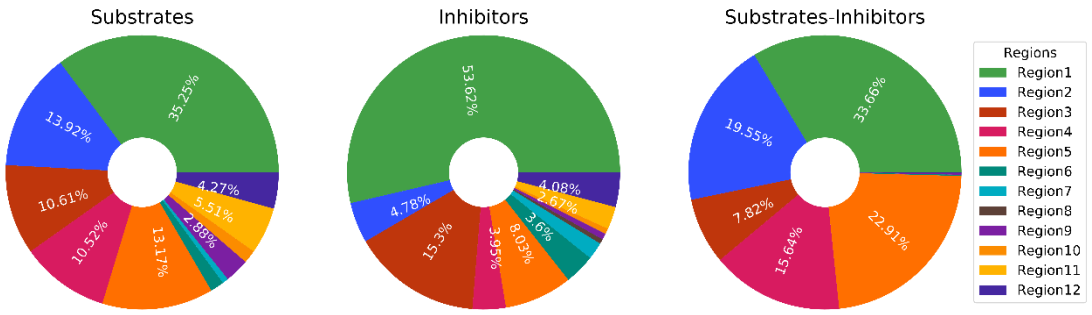


Figure 41: Fractions of the bound ligands (substrates, inhibitors and substrates/inhibitors) at every defined region. Here, as in the all the results in this work, only poses within 1 kcal.mol-1 from the best ligand pose were considered.

Regions 2 and 3 account for 24% of all substrate binding. This supports the hypothesis that substrates bind equivalently to several adjacent zones from Region 3, following the iso-energetic path (Figure 39) to the central protein chamber.

Region 2 binds a very small fraction of inhibitors (5% of bound poses). Inhibitors are mostly concentrated in Region 1 and to lesser extent in Region 3 (15%). In Region 4 (ATP binding site), only 5% of inhibitors were found, including thioguanine, indoprofen and nalidixic acid. Substrates-inhibitors exhibited a distribution in binding regions similar to substrates, with a higher proportion in region 5 (23%). One prototypical example of substrate-inhibitor binding in Region 4 is cyclic adenosine monophosphate (cAMP), which is a well-known endogenous

substrate of MRP4 and shares most of its structure with ATP. However, the small number of substrates-inhibitors in this study (i.e., 6) prevents generalization.

Region 6 exhibited higher binding to all types of ligands than the other minor binding regions (Regions 5- 12) that accounted for 12% of binding of substrates and inhibitors, and almost no substrates-inhibitors.

VII.5. CONCLUSION

Here, we performed a thorough docking study for 129 substrates, inhibitors and substrate/inhibitors of *h*MRP4 on the overall structure of the protein. Five major binding regions were highlighted, where Region 1 is the protein chamber in which the key residues previously reported in the literature were present. It is the most favorable binding region for all ligands. An equivalent implication of Regions 2 and 3 was observed, which suggested an iso-energetic gating path leading the central chamber of the protein. This was supported by MOLE channel analysis, which highlighted an overlap of the resulting channels from Region 3 to Region 1. Other small regions were also identified and account for less ligand binding. Several regions were common for substrates and inhibitors, in particular Region 1 and 3, while Region 8 was exclusively specific to inhibitors, which may suggest an allosteric binding site for inhibitors with its own key residues.

Moreover, the analysis of residues in contact with each type of ligands was established. We identified key residues namely TRP995, PHE368 and ARG998 (correlating with the literature) as well as other potential residues, namely ARG362, PHE324, TYR728 and LYS106 to be in tight interaction with all types of ligands. Interestingly, TYR985 was found to be specific to substrates. These identified residues may guide further targeted mutagenesis studies to investigate their role in substrate binding and translocation, and in the inhibition process.

Conclusion

Transport of xenobiotics inside the body involves their passage through lipid bilayer membranes. This mostly consists of two major membrane transport processes: (i) passive permeation along the concentration gradient across the membrane, and (ii) active transport requiring energy via membrane proteins. Passive permeation is common to most hydrophobic and amphiphilic molecules. Understanding the ability of xenobiotics to spontaneously cross the membrane is important in the drug discovery process. Passive permeation is commonly measured by the logarithm of the passive permeation coefficient ($logP_{erm}$). It can be measured experimentally or calculated by theoretical methods. Based on the ISDM theory, computational methods must determine two main ingredients across the lipid bilayer normal: the potential of mean force (free energy), and the diffusivity profiles.

In Chapter V and Chapter VI of this thesis, we used the AWH method for the first time on all-atom POPC membrane system to model both PMF and diffusivity profiles, creating a method named Memcross. In Chapter V, 13 different molecules (mostly substrates for ABC transporters) were subjects for determining their $logP_{erm}$. First, we compared the performance of FEP and Memcross methods to model the PMF profiles of a small alcohol molecule (i.e., ethanol) and larger polyphenols (quercetin, resveratrol and curcumin). Memcross has the advantage over FEP for being more accurate to assess the free energy profile and the orientation of the three molecules along the lipid bilayer, and faster in terms of simulation time. Moreover, fractional diffusivity profiles were computed from AWH trajectories, which allowed to take the subdiffusive behavior of permeants in the membrane core into account. We extended the use of Memcross to another set of 9 molecules including 4 acids, for which the protonation state across the permeation path was assessed.

In Chapter VI, to evaluate the performance of Memcross for $logP_{erm}$ predictions, we compared $logP_{erm}$ values of a large number of structurally different compounds with four different experimental measurements (Liposomes/BLM, PAMPA-DS, BBB and Caco2/MDCK). This includes neutral molecules, acids, bases and zwitterions. The correlations yielded a very good correlation with simplistic PC-based liposomes (R^2=0.83). Adding measurements from mixed datasets (liposomes and BLM) slightly decreased the correlation (R^2=0.71). We identified that Memcross might not be suited for carbohydrates. Furthermore, a relatively good agreement with PAMPA-DS experimental $logP_{exp}$ was highlighted (R^2=0.66). Weaker correlations were identified with datasets from cell-based membrane model (i.e., BBB and Caco2/MDCK) mostly due to the difference in membrane composition, (e.g., cholesterol and other non-PC-based lipids), the presence of non-suspected membrane transporters, USLs, and transport across tight junctions. The residual error between calculated and experimental $logP_{erm}$ was shown to be multifactorial. In addition, we evaluated the required simulation time for each type of molecules based on their molecular weight as it was shown to be strongly related to the convergence of $logP_{erm}$ during the AWH simulations. The diffusivity profile calculated either by using the fractional, classical or set as a constant value does not have a substantial role in determining $logP_{erm}$ for most hydrophilic and amphiphilic molecules.

These projects raise future perspectives to optimize the correlation of the calculated $logP_{erm}$ with PC-based membrane models. First, a benchmark of $logP_{erm}$ values over different methods for predicting the partial charges of the ligand atoms should be considered. Regarding carbohydrates, proper force filed such as GLYCAM can be used instead of GAFF2. For ions, standard dedicated forcefields can be used. Second, a benchmark over different forcefields

such as CHARMM36, Lipid21 and Slipids2021 could be of interest. Third, it would be interesting to evaluate the effect of different membrane compositions using mixtures of lipids including sphingolipids and cholesterol on Memcross performance. Fourth, biasing additional collective variables such as the orientation of the molecules could improve sampling and reduce the computational effort. Finally, and simulations using the whole translocation pathway (i.e., from one water medium to the other water side), particularly in the case of asymmetric membranes, could be adopted in the AWH algorithm to improve the free energy surface prediction. Furthermore, thanks to the full atomic resolution of our model and its flexibility, the permeation of molecules in the presence of other compounds (e.g., membrane proteins, peptides, or other small molecules) in the lipid bilayer medium can be investigated. Correspondingly, the impact of these accompanying biomolecules may affect the permeation could improve the correlation to more complex membrane systems (BBB and Caco2).

In Chapter VII, we tackled the active transport event through the human MRP4 membrane protein. First, the best model of hMRP4 structure was gathered from AlphaFold2, an artificial intelligence-based tool, which provides models with an accuracy rivaling the best experimental methods. The protein was embedded into a POPC and cholesterol lipid bilayer and relaxed using MD simulations. Second, a thorough molecular docking was established on the whole protein structure to identify potential binding sites of known hMRP4 substrates and inhibitors. Mainly, 6 binding regions were highlighted to bind all ligands. The main chamber of the protein (Region 1) was identified to be the most occupied region for all ligands. The interconnection between these regions, in particular from region 3 to 1 suggested an iso-energetic path that could allow these ligands to enter the main chamber. The ATP-binding site 1 was also identified to bind substrates/inhibitors in the A-loop domain. Other small regions were identified to bind specific ligands (Region 8 for inhibitors). Finally, the analysis of binding residues highlighted previously reported key residues (ARG368 and PHE995). New potential key residues were identified from this study. They constitute potential targets to investigate the substrates translocations or in the inhibitory mechanism. Based on these results, a further step would be to study the drug-drug interaction mechanism. One way to do this could be to select a known couples of substrates and inhibitors and to study the impact of the inhibitors on the dynamics of the protein structure using previously developed tools in our lab (e.g., clustering, allosteric network analysis, machine learning-based studies of the transport cycle).

Bibliography

(1) David J., T. Drug−Membrane Interactions. Analysis, Drug Distribution, Modeling. Methods and Principles in Medicinal Chemistry, Volume 15 By J. K. Seydel and M. Wiese. Wiley-VCH, Weinheim, Germany. 2002. Xix + 349 Pp. 17.5 × 24.5 Cm. ISBN 3 527 3042 7. $128.00. *J. Med. Chem.* **2003**, *46* (2), 318–318. https://doi.org/10.1021/jm0204688.

(2) Casares, D.; Escribá, P. V.; Rosselló, C. A. Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *IJMS* **2019**, *20* (9), 2167. https://doi.org/10.3390/ijms20092167.

(3) Smith, D.; Artursson, P.; Avdeef, A.; Di, L.; Ecker, G. F.; Faller, B.; Houston, J. B.; Kansy, M.; Kerns, E. H.; Krämer, S. D.; Lennernäs, H.; van de Waterbeemd, H.; Sugano, K.; Testa, B. Passive Lipoidal Diffusion and Carrier-Mediated Cell Uptake Are Both Important Mechanisms of Membrane Permeation in Drug Disposition. *Mol. Pharmaceutics* **2014**, *11* (6), 1727–1738. https://doi.org/10.1021/mp400713v.

(4) Järvinen, E. Human Efflux Transporters in Drug Disposition: In Vitro Transport of Glucuronide Metabolites. 74.

(5) Mansy, S. S. Membrane Transport in Primitive Cells. *Cold Spring Harbor Perspectives in Biology* **2010**, *2* (8), a002188–a002188. https://doi.org/10.1101/cshperspect.a002188.

(6) Parisio, G.; Stocchero, M.; Ferrarini, A. Passive Membrane Permeability: Beyond the Standard Solubility-Diffusion Model. *J Chem Theory Comput* **2013**, *9* (12), 5236–5246. https://doi.org/10.1021/ct400690t.

(7) Overton, C. E. *Studien Über Die Narkose: Zugleich Ein Beitrag Zur Allgemeinen Pharmakologie*; 1901.

(8) Nasr, G.; Greige-Gerges, H.; Elaissari, A.; Khreich, N. Liposomal Membrane Permeability Assessment by Fluorescence Techniques: Main Permeabilizing Agents, Applications and Challenges. *Int J Pharm* **2020**, *580*, 119198. https://doi.org/10.1016/j.ijpharm.2020.119198.

(9) Bangham, A. D. Surrogate Cells or Trojan Horses. The Discovery of Liposomes. *Bioessays* **1995**, *17* (12), 1081–1088. https://doi.org/10.1002/bies.950171213.

(10) Clary, L.; Verderone, G.; Santaella, C.; Vierling, P. Membrane Permeability and Stability of Liposomes Made from Highly Fluorinated Double-Chain Phosphocholines Derived from Diaminopropanol, Serine or Ethanolamine. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1997**, *1328* (1), 55–64. https://doi.org/10.1016/S0005-2736(97)00076-X.

(11) Artursson, P.; Karlsson, J. Correlation between Oral Drug Absorption in Humans and Apparent Drug Permeability Coefficients in Human Intestinal Epithelial (Caco-2) Cells. *Biochemical and Biophysical Research Communications* **1991**, *175* (3), 880–885. https://doi.org/10.1016/0006-291X(91)91647-U.

(12) Collett, A.; Sims, E.; Walker, D.; He, Y.; Ayrton, J.; Rowland, M.; Warhurst, G. Comparison of HT29-18-C1 and Caco-2 Cell Lines as Models for Studying Intestinal Paracellular Drug Absorption. *Pharmaceutical Research* **1996**, *13* (2), 216–221. https://doi.org/10.1023/A:1016082829111.

(13) Kotova, E. A.; Kuzevanov, A. V.; Pashkovskaya, A. A.; Antonenko, Y. N. Selective Permeabilization of Lipid Membranes by Photodynamic Action via Formation of Hydrophobic Defects or Pre-Pores. *Biochimica et Biophysica Acta (BBA) -*

Biomembranes **2011**, *1808* (9), 2252–2257. https://doi.org/10.1016/j.bbamem.2011.05.018.

(14) Avdeef, A. *Avdeef A Permeability—PAMPA In Absorption and Drug Development*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(15) Cooper, G. M. Cell Membranes. *The Cell: A Molecular Approach. 2nd edition* **2000**.

(16) Conrado, D. J.; Duvvuri, S.; Geerts, H.; Burton, J.; Biesdorf, C.; Ahamadi, M.; Macha, S.; Hather, G.; Francisco Morales, J.; Podichetty, J.; Nicholas, T.; Stephenson, D.; Trame, M.; Romero, K.; Corrigan, B.; Group, the D. D. T. in the A. D. C. (DDT-A. W. Challenges in Alzheimer's Disease Drug Discovery and Development: The Role of Modeling, Simulation, and Open Data. *Clinical Pharmacology & Therapeutics* **2020**, *107* (4), 796–805. https://doi.org/10.1002/cpt.1782.

(17) Meyer, A. H. Zur Theorie Der Alkoholnarkose. *Archiv für Experimentelle Pathologie und Pharmakologie* **1899**. https://doi.org/10.1007/BF01834479.

(18) Awoonor-Williams, E.; Rowley, C. N. Molecular Simulation of Nonfacilitated Membrane Permeation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2016**, *1858* (7), 1672–1687. https://doi.org/10.1016/j.bbamem.2015.12.014.

(19) Gauthier, A.; Joós, B. Stretching Effects on the Permeability of Water Molecules across a Lipid Bilayer. *The Journal of Chemical Physics* **2007**, *127* (10), 105104. https://doi.org/10.1063/1.2764079.

(20) Venable, R. M.; Krämer, A.; Pastor, R. W. Molecular Dynamics Simulations of Membrane Permeability. *Chem. Rev.* **2019**, *119* (9), 5954–5997. https://doi.org/10.1021/acs.chemrev.8b00486.

(21) Shinoda, W. Permeability across Lipid Membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2016**, *1858* (10), 2254–2265. https://doi.org/10.1016/j.bbamem.2016.03.032.

(22) Lomize, A. L.; Pogozheva, I. D. Physics-Based Method for Modeling Passive Membrane Permeability and Translocation Pathways of Bioactive Molecules. *J. Chem. Inf. Model.* **2019**, *59* (7), 3198–3213. https://doi.org/10.1021/acs.jcim.9b00224.

(23) Avdeef, A. *Absorption and Drug Development: Solubility, Permeability, and Charge State*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(24) Lipinski, C. A. Drug-like Properties and the Causes of Poor Solubility and Poor Permeability. *Journal of Pharmacological and Toxicological Methods* **2000**, *44* (1), 235–249. https://doi.org/10.1016/S1056-8719(00)00107-6.

(25) James Ritter; Flower, R.; Henderson, G.; Rang, H. *Rang and Dale's Pharmacology*, 7th ed.; 1950.

(26) 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 Settings1PII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in Advanced Drug Delivery Reviews 23 (1997) 3–25.1. *Advanced Drug Delivery Reviews* **2001**, *46* (1), 3–26. https://doi.org/10.1016/S0169-409X(00)00129-0.

(27) Guidotti, G. The Composition of Biological Membranes. *Arch Intern Med* **1972**, *129* (2), 194. https://doi.org/10.1001/archinte.1972.00320020038003.

(28) Russell, D. W. The Enzymes, Regulation, and Genetics of Bile Acid Synthesis. *Annu. Rev. Biochem.* **2003**, *72* (1), 137–174. https://doi.org/10.1146/annurev.biochem.72.121801.161712.

(29) Kuivenhoven, J. A.; Hegele, R. A. Mining the Genome for Lipid Genes. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2014**, *1842* (10), 1993–2009. https://doi.org/10.1016/j.bbadis.2014.04.028.

(30) Lamari, F.; Mochel, F.; Sedel, F.; Saudubray, J. M. Disorders of Phospholipids, Sphingolipids and Fatty Acids Biosynthesis: Toward a New Category of Inherited Metabolic Diseases. *J Inherit Metab Dis* **2013**, *36* (3), 411–425. https://doi.org/10.1007/s10545-012-9509-7.

(31) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane Lipids: Where They Are and How They Behave. *Nat Rev Mol Cell Biol* **2008**, *9* (2), 112–124. https://doi.org/10.1038/nrm2330.

(32) Harayama, T.; Riezman, H. Understanding the Diversity of Membrane Lipid Composition. *Nat Rev Mol Cell Biol* **2018**, *19* (5), 281–296. https://doi.org/10.1038/nrm.2017.138.

(33) Halliwell, B. Antioxidants in Human Health and Disease. *Annu. Rev. Nutr.* **1996**, *16* (1), 33–50. https://doi.org/10.1146/annurev.nu.16.070196.000341.

(34) Aruoma, O. I. Free Radicals, Oxidative Stress, and Antioxidants in Human Health and Disease. *J Amer Oil Chem Soc* **1998**, *75* (2), 199–212. https://doi.org/10.1007/s11746-998-0032-9.

(35) Bogdanov, M.; Mileykovskaya, E.; Dowhan, W. Lipids in the Assembly of Membrane Proteins and Organization of Protein Supercomplexes: Implications for Lipid-Linked Disorders. In *Lipids in Health and Disease*; Quinn, P. J., Wang, X., Eds.; Harris, J. R., Quinn, P. J., Series Eds.; Subcellular Biochemistry; Springer Netherlands: Dordrecht, 2008; Vol. 49, pp 197–239. https://doi.org/10.1007/978-1-4020-8831-5_8.

(36) Corradi, V.; Mendez-Villuendas, E.; Ingólfsson, H. I.; Gu, R.-X.; Siuda, I.; Melo, M. N.; Moussatova, A.; DeGagné, L. J.; Sejdiu, B. I.; Singh, G.; Wassenaar, T. A.; Delgado Magnero, K.; Marrink, S. J.; Tieleman, D. P. Lipid–Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent. Sci.* **2018**, *4* (6), 709–717. https://doi.org/10.1021/acscentsci.8b00143.

(37) Palsdottir, H.; Hunte, C. Lipids in Membrane Protein Structures. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004**, *1666* (1–2), 2–18. https://doi.org/10.1016/j.bbamem.2004.06.012.

(38) Singer, S. J.; Nicolson, G. L. The Fluid Mosaic Model of the Structure of Cell Membranes: Cell Membranes Are Viewed as Two-Dimensional Solutions of Oriented Globular Proteins and Lipids. *Science* **1972**, *175* (4023), 720–731. https://doi.org/10.1126/science.175.4023.720.

(39) Escribá, P. V. Membrane-Lipid Therapy: A Historical Perspective of Membrane-Targeted Therapies — From Lipid Bilayer Structure to the Pathophysiological Regulation of Cells. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2017**, *1859* (9), 1493–1506. https://doi.org/10.1016/j.bbamem.2017.05.017.

(40) White, S. H.; Ladokhin, A. S.; Jayasinghe, S.; Hristova, K. How Membranes Shape Protein Structure. *Journal of Biological Chemistry* **2001**, *276* (35), 32395–32398. https://doi.org/10.1074/jbc.R100008200.

(41) *Membrane Proteins Diffuse as Dynamic Complexes with Lipids | Journal of the American Chemical Society*. https://pubs.acs.org/doi/10.1021/ja101481b (accessed 2022-08-14).

(42) Sharifian Gh., M. Recent Experimental Developments in Studying Passive Membrane Transport of Drug Molecules. *Mol. Pharmaceutics* **2021**, *18* (6), 2122–2141. https://doi.org/10.1021/acs.molpharmaceut.1c00009.

(43) Sohlenkamp, C.; Geiger, O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiology Reviews* **2016**, *40* (1), 133–159. https://doi.org/10.1093/femsre/fuv008.

(44) Yamashita, A.; Hayashi, Y.; Nemoto-Sasaki, Y.; Ito, M.; Oka, S.; Tanikawa, T.; Waku, K.; Sugiura, T. Acyltransferases and Transacylases That Determine the Fatty Acid Composition of Glycerolipids and the Metabolism of Bioactive Lipid Mediators in Mammalian Cells and Model Organisms. *Progress in Lipid Research* **2014**, *53*, 18–81. https://doi.org/10.1016/j.plipres.2013.10.001.

(45) Grösch, S.; Schiffmann, S.; Geisslinger, G. Chain Length-Specific Properties of Ceramides. *Progress in Lipid Research* **2012**, *51* (1), 50–62. https://doi.org/10.1016/j.plipres.2011.11.001.

(46) Uphoff, A.; Hermansson, M.; Haimi, P.; Somerharju, P. Analysis of Complex Lipidomes. In *Medical Applications of Mass Spectrometry*; Elsevier, 2008; pp 223–249. https://doi.org/10.1016/B978-044451980-1.50013-6.

(47) Hannun, Y. A.; Obeid, L. M. Sphingolipids and Their Metabolism in Physiology and Disease. *Nat Rev Mol Cell Biol* **2018**, *19* (3), 175–191. https://doi.org/10.1038/nrm.2017.107.

(48) Dufourc, E. J. Sterols and Membrane Dynamics. *J Chem Biol* **2008**, *1* (1–4), 63–77. https://doi.org/10.1007/s12154-008-0010-6.

(49) Nagle, J. F.; Tristram-Nagle, S. Lipid Bilayer Structure. *Curr Opin Struct Biol* **2000**, *10* (4), 474–480.

(50) Kučerka, N.; Nieh, M.-P.; Katsaras, J. Fluid Phase Lipid Areas and Bilayer Thicknesses of Commonly Used Phosphatidylcholines as a Function of Temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2011**, *1808* (11), 2761–2771. https://doi.org/10.1016/j.bbamem.2011.07.022.

(51) Kučerka, N.; Perlmutter, J. D.; Pan, J.; Tristram-Nagle, S.; Katsaras, J.; Sachs, J. N. The Effect of Cholesterol on Short- and Long-Chain Monounsaturated Lipid Bilayers as Determined by Molecular Dynamics Simulations and X-Ray Scattering. *Biophys J* **2008**, *95* (6), 2792–2805. https://doi.org/10.1529/biophysj.107.122465.

(52) Katsuta, H.; Sawada, Y.; Sokabe, M. Biophysical Mechanisms of Membrane-Thickness-Dependent MscL Gating: An All-Atom Molecular Dynamics Study. *Langmuir* **2019**, *35* (23), 7432–7442. https://doi.org/10.1021/acs.langmuir.8b02074.

(53) Jämbeck, J. P. M.; Lyubartsev, A. P. Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *J. Phys. Chem. B* **2012**, *116* (10), 3164–3179. https://doi.org/10.1021/jp212503e.

(54) Klauda, J. B.; Kučerka, N.; Brooks, B. R.; Pastor, R. W.; Nagle, J. F. Simulation-Based Methods for Interpreting X-Ray Data from Lipid Bilayers. *Biophys J* **2006**, *90* (8), 2796–2807. https://doi.org/10.1529/biophysj.105.075697.

(55) Poojari, C.; Wilkosz, N.; Lira, R. B.; Dimova, R.; Jurkiewicz, P.; Petka, R.; Kepczynski, M.; Róg, T. Behavior of the DPH Fluorescence Probe in Membranes Perturbed by Drugs. *Chemistry and Physics of Lipids* **2019**, *223*, 104784. https://doi.org/10.1016/j.chemphyslip.2019.104784.

(56) Piggot, T. J.; Allison, J. R.; Sessions, R. B.; Essex, J. W. On the Calculation of Acyl Chain Order Parameters from Lipid Simulations. *J. Chem. Theory Comput.* **2017**, *13* (11), 5683–5696. https://doi.org/10.1021/acs.jctc.7b00643.

(57) Grote, F.; Lyubartsev, A. P. Optimization of Slipids Force Field Parameters Describing Headgroups of Phospholipids. *J. Phys. Chem. B* **2020**, *124* (40), 8784–8793. https://doi.org/10.1021/acs.jpcb.0c06386.

(58) *3.1: Membrane Phase Transitions*. Physics LibreTexts. https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_241_-_Membrane_Biology/03%3A_Membrane_Phases_and_Morphologies/3.01%3A_Membrane_Phase_Transitions (accessed 2022-08-14).

(59) Sezgin, E.; Levental, I.; Mayor, S.; Eggeling, C. The Mystery of Membrane Organization: Composition, Regulation and Roles of Lipid Rafts. *Nat Rev Mol Cell Biol* **2017**, *18* (6), 361–374. https://doi.org/10.1038/nrm.2017.16.

(60) Nakahara, K.; Ohkuni, A.; Kitamura, T.; Abe, K.; Naganuma, T.; Ohno, Y.; Zoeller, R. A.; Kihara, A. The Sjögren-Larsson Syndrome Gene Encodes a Hexadecenal Dehydrogenase of the Sphingosine 1-Phosphate Degradation Pathway. *Molecular Cell* **2012**, *46* (4), 461–471. https://doi.org/10.1016/j.molcel.2012.04.033.

(61) Braverman, N. E.; Raymond, G. V.; Rizzo, W. B.; Moser, A. B.; Wilkinson, M. E.; Stone, E. M.; Steinberg, S. J.; Wangler, M. F.; Rush, E. T.; Hacia, J. G.; Bose, M. Peroxisome Biogenesis Disorders in the Zellweger Spectrum: An Overview of Current Diagnosis, Clinical Manifestations, and Treatment Guidelines. *Molecular Genetics and Metabolism* **2016**, *117* (3), 313–321. https://doi.org/10.1016/j.ymgme.2015.12.009.

(62) Lauwers, E.; Goodchild, R.; Verstreken, P. Membrane Lipids in Presynaptic Function and Disease. *Neuron* **2016**, *90* (1), 11–25. https://doi.org/10.1016/j.neuron.2016.02.033.

(63) Pizzino, G.; Irrera, N.; Cucinotta, M.; Pallio, G.; Mannino, F.; Arcoraci, V.; Squadrito, F.; Altavilla, D.; Bitto, A. Oxidative Stress: Harms and Benefits for Human Health. *Oxid Med Cell Longev* **2017**, *2017*, 8416763. https://doi.org/10.1155/2017/8416763.

(64) Niki, E. Lipid Peroxidation: Physiological Levels and Dual Biological Effects. *Free Radical Biology and Medicine* **2009**, *47* (5), 469–484. https://doi.org/10.1016/j.freeradbiomed.2009.05.032.

(65) Ostrea, E. M.; Cepeda, E. E.; Fleury, C. A.; Balun, J. E. Red Cell Membrane Lipid Peroxidation and Hemolysis Secondary to Phototherapy. *Acta Paediatr Scand* **1985**, *74* (3), 378–381. https://doi.org/10.1111/j.1651-2227.1985.tb10987.x.

(66) Frei, B. *Natural Antioxidants in Human Health and Disease. Academic Press.*, ed.; Academic Press; 2012.

(67) Walrant, A.; Cardon, S.; Burlina, F.; Sagan, S. Membrane Crossing and Membranotropic Activity of Cell-Penetrating Peptides: *Dangerous Liaisons*? *Acc. Chem. Res.* **2017**, *50* (12), 2968–2975. https://doi.org/10.1021/acs.accounts.7b00455.

(68) Zheng Koh, D. H.; Saheki, Y. Regulation of Plasma Membrane Sterol Homeostasis by Nonvesicular Lipid Transport. *Contact* **2021**, *4*, 25152564211042452. https://doi.org/10.1177/25152564211042451.

(69) You, G. *Drug Transporters: Molecular Characterization and Role in Drug Disposition*; Wiley: Hoboken, NJ, 2014.

(70) Tse, C. H.; Comer, J.; Sang Chu, S. K.; Wang, Y.; Chipot, C. Affordable Membrane Permeability Calculations: Permeation of Short-Chain Alcohols through Pure-Lipid Bilayers and a Mammalian Cell Membrane. *J. Chem. Theory Comput.* **2019**, *15* (5), 2913–2924. https://doi.org/10.1021/acs.jctc.9b00022.

(71) Frallicciardi, J.; Melcr, J.; Siginou, P.; Marrink, S. J.; Poolman, B. Membrane Thickness, Lipid Phase and Sterol Type Are Determining Factors in the Permeability of Membranes to Small Solutes. *Nat Commun* **2022**, *13* (1), 1605. https://doi.org/10.1038/s41467-022-29272-x.

(72) Guler, S. D.; Ghosh, D. D.; Pan, J.; Mathai, J. C.; Zeidel, M. L.; Nagle, J. F.; Tristram-Nagle, S. Effects of Ether vs. Ester Linkage on Lipid Bilayer Structure and Water Permeability. *Chem Phys Lipids* **2009**, *160* (1), 33–44. https://doi.org/10.1016/j.chemphyslip.2009.04.003.

(73) Gromiha, M. M.; Ou, Y.-Y. Bioinformatics Approaches for Functional Annotation of Membrane Proteins. *Briefings in Bioinformatics* **2014**, *15* (2), 155–168. https://doi.org/10.1093/bib/bbt015.

(74) Almeida, J. G.; Preto, A. J.; Koukos, P. I.; Bonvin, A. M. J. J.; Moreira, I. S. Membrane Proteins Structures: A Review on Computational Modeling Tools. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2017**, *1859* (10), 2021–2039. https://doi.org/10.1016/j.bbamem.2017.07.008.

(75) Piccoli, S.; Suku, E.; Garonzi, M.; Giorgetti, A. Genome-Wide Membrane Protein Structure Prediction. *Curr Genomics* **2013**, *14* (5), 324–329. https://doi.org/10.2174/13892029113149990009.

(76) Johnson, A. E.; van Waes, M. A. The Translocon: A Dynamic Gateway at the ER Membrane. *Annu. Rev. Cell Dev. Biol.* **1999**, *15* (1), 799–842. https://doi.org/10.1146/annurev.cellbio.15.1.799.

(77) Peyronnet, R.; Tran, D.; Girault, T.; Frachisse, J.-M. Mechanosensitive Channels: Feeling Tension in a World under Pressure. *Front. Plant Sci.* **2014**, *5*. https://doi.org/10.3389/fpls.2014.00558.

(78) Sugano, K.; Kansy, M.; Artursson, P.; Avdeef, A.; Bendels, S.; Di, L.; Ecker, G. F.; Faller, B.; Fischer, H.; Gerebtzoff, G.; Lennernaes, H.; Senner, F. Coexistence of Passive and Carrier-Mediated Processes in Drug Transport. *Nat Rev Drug Discov* **2010**, *9* (8), 597–614. https://doi.org/10.1038/nrd3187.

(79) Bondar, A.-N.; del Val, C.; Freites, J. A.; Tobias, D. J.; White, S. H. Dynamics of SecY Translocons with Translocation-Defective Mutations. *Structure* **2010**, *18* (7), 847–857. https://doi.org/10.1016/j.str.2010.04.010.

(80) Catterall, W. A.; Perez-Reyes, E.; Snutch, T. P.; Striessnig, J. International Union of Pharmacology. XLVIII. Nomenclature and Structure-Function Relationships of Voltage-Gated Calcium Channels. *Pharmacol Rev* **2005**, *57* (4), 411–425. https://doi.org/10.1124/pr.57.4.5.

(81) Yamakage, M.; Namiki, A. Calcium Channels--Basic Aspects of Their Structure, Function and Gene Encoding; Anesthetic Action on the Channels--a Review. *Can J Anaesth* **2002**, *49* (2), 151–164. https://doi.org/10.1007/BF03020488.

(82) Dean, M.; Hamon, Y.; Chimini, G. The Human ATP-Binding Cassette (ABC) Transporter Superfamily. *J Lipid Res* **2001**, *42* (7), 1007–1017.

(83) Ahn, S.-Y.; Nigam, S. K. Toward a Systems Level Understanding of Organic Anion and Other Multispecific Drug Transporters: A Remote Sensing and Signaling Hypothesis. *Mol Pharmacol* **2009**, *76* (3), 481–490. https://doi.org/10.1124/mol.109.056564.

(84) DeGorter, M. K.; Xia, C. Q.; Yang, J. J.; Kim, R. B. Drug Transporters in Drug Efficacy and Toxicity. *Annu Rev Pharmacol Toxicol* **2012**, *52*, 249–273. https://doi.org/10.1146/annurev-pharmtox-010611-134529.

(85) Nigam, S. K. What Do Drug Transporters Really Do? *Nature Reviews Drug Discovery* **2015**, *14* (1), 29–44. https://doi.org/10.1038/nrd4461.

(86) Thomson, A. B. R.; Dietschy, J. M. Derivation of the Equations That Describe the Effects of Unstirred Water Layers on the Kinetic Parameters of Active Transport Processes in the Intestine. *Journal of Theoretical Biology* **1977**, *64* (2), 277–294. https://doi.org/10.1016/0022-5193(77)90357-5.

(87) Hgerle, M. L.; Winne, D. Drug Absorption by the Rat Jejunum Perfused in Situ: Dissociation from the PH-Partition Theory and Role of Microclimate-PH and Unstirred Layer. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1983**, *322* (4), 249–255. https://doi.org/10.1007/BF00508339.

(88) Chiou, W. L. Effect of 'Unstirred' Water Layer in the Intestine on the Rate and Extent of Absorption after Oral Administration. *Biopharm. Drug Dispos.* **1994**, *15* (8), 709–717. https://doi.org/10.1002/bdd.2510150808.

(89) Fagerholm, U.; Lennernäs, H. Experimental Estimation of the Effective Unstirred Water Layer Thickness in the Human Jejunum, and Its Importance in Oral Drug Absorption. *European Journal of Pharmaceutical Sciences* **1995**, *3* (5), 247–253. https://doi.org/10.1016/0928-0987(95)00027-B.

(90) Pohl, P.; Saparov, S. M.; Antonenko, Y. N. The Size of the Unstirred Layer as a Function of the Solute Diffusion Coefficient. *Biophysical Journal* **1998**, *75* (3), 1403–1409. https://doi.org/10.1016/S0006-3495(98)74058-5.

(91) Madara, J. L.; Pappenheimer, J. R. Structural Basis for Physiological Regulation of Paracellular Pathways in Intestinal Epithelia. *J. Membrain Biol.* **1987**, *100* (1), 149–164. https://doi.org/10.1007/BF02209147.

(92) Jöns, T.; Wittschieber, D.; Beyer, A.; Meier, C.; Brune, A.; Thomzig, A.; Ahnert-Hilger, G.; Veh, R. W. K+-ATP-Channel-Related Protein Complexes: Potential Transducers in the Regulation of Epithelial Tight Junction Permeability. *Journal of Cell Science* **2006**, *119* (15), 3087–3097. https://doi.org/10.1242/jcs.03041.

(93) Pappenheimer, J. R.; Reiss, K. Z. Contribution of Solvent Drag through Intercellular Junctions to Absorption of Nutrients by the Small Intestine of the Rat. *J. Membrain Biol.* **1987**, *100* (1), 123–136. https://doi.org/10.1007/BF02209145.

(94) Lee, B.; Moon, K. M.; Kim, C. Y. Tight Junction in the Intestinal Epithelium: Its Association with Diseases and Regulation by Phytochemicals. *J Immunol Res* **2018**, *2018*, 2645465. https://doi.org/10.1155/2018/2645465.

(95) Mostov, K. E. Transepithelial Transport of Immunoglobulins. *Annu. Rev. Immunol.* **1994**, *12* (1), 63–84. https://doi.org/10.1146/annurev.iy.12.040194.000431.

(96) Dan, N.; Cutler, D. F. Transcytosis and Processing of Intrinsic Factor-Cobalamin in Caco-2 Cells. *J Biol Chem* **1994**, *269* (29), 18849–18855.

(97) Bernal, W.; Wendon, J. Acute Liver Failure. *N Engl J Med* **2013**, *369* (26), 2525–2534. https://doi.org/10.1056/NEJMra1208937.

(98) Christopher Zusi, F.; Pitts, W. J.; Burke, J. R. IKK-2/NF-KB-DEPENDENT TRANSCRIPTION. In *Target Validation in Drug Discovery*; Elsevier, 2007; pp 199–221. https://doi.org/10.1016/B978-012369393-8/50012-2.

(99) Dallman, M. J. Graft Rejection. In *Encyclopedia of Immunology*; Elsevier, 1998; pp 1011–1015. https://doi.org/10.1006/rwei.1999.0261.

(100) Weston, B. C. Tissue Typing and Matching. In *Corneal Surgery*; Elsevier, 2009; pp 301–304. https://doi.org/10.1016/B978-0-323-04835-4.50040-9.

(101) Colombo, D.; Ammirati, E. Cyclosporine in Transplantation - a History of Converging Timelines. *J Biol Regul Homeost Agents* **2011**, *25* (4), 493–504.

(102) Ong, P. W.; Kee, T.; Ho, Q. Y. Impact of Tacrolimus versus Cyclosporine on One-Year Renal Transplant Outcomes: A Single-Centre Retrospective Cohort Study. *Proceedings of Singapore Healthcare* **2020**, *29* (4), 217–222. https://doi.org/10.1177/2010105820957370.

(103) Okour, M.; Jacobson, P. A.; Ahmed, M. A.; Israni, A. K.; Brundage, R. C. Mycophenolic Acid and Its Metabolites in Kidney Transplant Recipients: A Semimechanistic Enterohepatic Circulation Model to Improve Estimating Exposure. *J Clin Pharmacol* **2018**, *58* (5), 628–639. https://doi.org/10.1002/jcph.1064.

(104) Miners, J.; Yang, X.; Knights, K.; Zhang, L. The Role of the Kidney in Drug Elimination: Transport, Metabolism, and the Impact of Kidney Disease on Drug Clearance. *Clin. Pharmacol. Ther.* **2017**, *102* (3), 436–449. https://doi.org/10.1002/cpt.757.

(105) Berthier, J.; Benmameri, M.; Sauvage, F.-L.; Fabre, G.; Chantemargue, B.; Arnion, H.; Marquet, P.; Trouillas, P.; Picard, N.; Saint-Marcoux, F. MRP4 Is Responsible for the Efflux Transport of Mycophenolic Acid β- D Glucuronide (MPAG) from Hepatocytes to Blood. *Xenobiotica* **2020**, 1–10. https://doi.org/10.1080/00498254.2020.1813352.

(106) George, B.; You, D.; Joy, M. S.; Aleksunes, L. M. Xenobiotic Transporters and Kidney Injury. *Advanced Drug Delivery Reviews* **2017**, *116*, 73–91. https://doi.org/10.1016/j.addr.2017.01.005.

(107) Aouad, H.; Faucher, Q.; Sauvage, F.-L.; Pinault, E.; Barrot, C.-C.; Arnion, H.; Marquet, P.; Essig, M. *A Multi-Omics Study to Investigate Tacrolimus Nephrotoxicity Mechanisms*; preprint; Pharmacology and Toxicology, 2021. https://doi.org/10.1101/2021.07.29.454229.

(108) Alexander, S. I.; Younes, S. B.; Zurakowski, D.; Mirza, N. M.; Dubey, D. P.; Drew, M. P.; Harmon, W. E.; Yunis, E. J. CELL-MEDIATED CYTOTOXICITY: A PREDICTOR OF CHRONIC REJECTION IN PEDIATRIC HLA HAPLOIDENTICAL RENAL TRANSPLANTS1: *Transplantation* **1997**, *63* (12), 1756–1761. https://doi.org/10.1097/00007890-199706270-00009.

(109) Burat, B.; Faucher, Q.; Čechová, P.; Arnion, H.; Di Meo, F.; Sauvage, F.; Marquet, P.; Essig, M. Cyclosporine A Inhibits MRTF-SRF Signaling through Na^+/K^+ ATPase Inhibition and Actin Remodeling. *FASEB BioAdvances* **2019**, *1* (9), 561–578. https://doi.org/10.1096/fba.2019-00027.

(110) Giacomini, K. M.; Huang, S.-M.; Tweedie, D. J.; Benet, L. Z.; Brouwer, K. L. R.; Chu, X.; Dahlin, A.; Evers, R.; Fischer, V.; Hillgren, K. M.; Hoffmaster, K. A.; Ishikawa, T.; Keppler, D.; Kim, R. B.; Lee, C. A.; Niemi, M.; Polli, J. W.; Sugiyama, Y.; Swaan, P. W.; Ware, J. A.; Wright, S. H.; Wah Yee, S.; Zamek-Gliszczynski, M. J.; Zhang, L.; The International Transporter Consortium. Membrane Transporters in Drug Development. *Nature Reviews Drug Discovery* **2010**, *9* (3), 215–236. https://doi.org/10.1038/nrd3028.

(111) Hannesschlaeger, C.; Horner, A.; Pohl, P. Intrinsic Membrane Permeability to Small Molecules. *Chem. Rev.* **2019**, *119* (9), 5922–5953. https://doi.org/10.1021/acs.chemrev.8b00560.

(112) Kleinzeller, A. Ernest Overton's Contribution to the Cell Membrane Concept: A Centennial Appreciation. *Physiology* **1997**, *12* (1), 49–53. https://doi.org/10.1152/physiologyonline.1997.12.1.49.

(113) Kleinzeller, A. Chapter 1 Charles Ernest Overton's Concept of a Cell Membrane. In *Current Topics in Membranes*; Elsevier, 1999; Vol. 48, pp 1–22. https://doi.org/10.1016/S0070-2161(08)61039-4.

(114) Kleinzeller, A. *Exploring the Cell Membrane: Conceptual Developments.*; Elsevier Science: Amsterdam, 2014.

(115) 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. Pharmaceutics* **2014**, *11* (10), 3696–3706. https://doi.org/10.1021/mp500437e.

(116) Roux, B. Lonely Arginine Seeks Friendly Environment. *Journal of General Physiology* **2007**, *130* (2), 233–236. https://doi.org/10.1085/jgp.200709819.

(117) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. Distribution of Amino Acids in a Lipid Bilayer from Computer Simulations. *Biophysical Journal* **2008**, *94* (9), 3393–3404. https://doi.org/10.1529/biophysj.107.112805.

(118) Walter, A.; Gutknecht, J. Permeability of Small Nonelectrolytes through Lipid Bilayer Membranes. *J. Membrain Biol.* **1986**, *90* (3), 207–217. https://doi.org/10.1007/BF01870127.

(119) Leeson, P. D.; Davis, A. M. Time-Related Differences in the Physical Property Profiles of Oral Drugs. *J. Med. Chem.* **2004**, *47* (25), 6338–6348. https://doi.org/10.1021/jm049717d.

(120) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45* (12), 2615–2623. https://doi.org/10.1021/jm020017n.

(121) 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 1PII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in Advanced Drug Delivery Reviews 23 (1997) 3–25. 1. *Advanced Drug Delivery Reviews* **2001**, *46* (1–3), 3–26. https://doi.org/10.1016/S0169-409X(00)00129-0.

(122) Avdeef, A. Solubility of Sparingly-Soluble Ionizable Drugs. *Advanced Drug Delivery Reviews* **2007**, *59* (7), 568–590. https://doi.org/10.1016/j.addr.2007.05.008.

(123) Avdeef, A. *Absorption and Drug Development: Solubility, Permeability, and Charge State*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(124) Wu, C.-Y.; Benet, L. Z. Predicting Drug Disposition via Application of BCS: Transport/Absorption/ Elimination Interplay and Development of a Biopharmaceutics Drug Disposition Classification System. *Pharm Res* **2005**, *22* (1), 11–23. https://doi.org/10.1007/s11095-004-9004-4.

(125) Amidon, G. L.; Lennernäs, H.; Shah, V. P.; Crison, J. R. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. *Pharmaceutical Research* **1995**, *12* (3), 413–420. https://doi.org/10.1023/A:1016212804288.

(126) Marrink, S.-J.; Berendsen, H. J. C. Simulation of Water Transport through a Lipid Membrane. *J. Phys. Chem.* **1994**, *98* (15), 4155–4168. https://doi.org/10.1021/j100066a040.

(127) Avdeef, A. Physicochemical Profiling (Solubility, Permeability and Charge State). *CTMC* **2001**, *1* (4), 277–351. https://doi.org/10.2174/1568026013395100.

(128) Diamond, J. M.; Katz, Y. Interpretation of Nonelectrolyte Partition Coefficients between Dimyristoyl Lecithin and Water. *J. Membrain Biol.* **1974**, *17* (1), 121–154. https://doi.org/10.1007/BF01870176.

(129) Ghysels, A.; Venable, R. M.; Pastor, R. W.; Hummer, G. Position-Dependent Diffusion Tensors in Anisotropic Media from Simulation: Oxygen Transport in and through Membranes. *J. Chem. Theory Comput.* **2017**, *13* (6), 2962–2976. https://doi.org/10.1021/acs.jctc.7b00039.

(130) Peters, B. Continuous Stochastic Variables. In *Reaction Rate Theory and Rare Events Simulations*; Elsevier, 2017; pp 403–433. https://doi.org/10.1016/B978-0-44-456349-1.00015-5.

(131) KrÃ¤mer, S.; Lombardi, D.; Primorac, A.; Thomae, A.; Wunderli-Allenspach, H. Lipid-Bilayer Permeation of Drug-Like Compounds. *C&B* **2009**, *6* (11), 1900–1916. https://doi.org/10.1002/cbdv.200900122.

(132) Kamp, F.; Hamilton, J. A. How Fatty Acids of Different Chain Length Enter and Leave Cells by Free Diffusion. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **2006**, *75* (3), 149–159. https://doi.org/10.1016/j.plefa.2006.05.003.

(133) Shore, P. A.; Brodie, B. B.; Hogben, C. A. M. THE GASTRIC SECRETION OF DRUGS: A PH PARTITION HYPOTHESIS. *J Pharmacol Exp Ther* **1957**, *119* (3), 361.

(134) Alex, A.; Millan, D. S.; Perez, M.; Wakenhut, F.; Whitlock, G. A. Intramolecular Hydrogen Bonding to Improve Membrane Permeability and Absorption in beyond Rule of Five Chemical Space. *Med. Chem. Commun.* **2011**, *2* (7), 669. https://doi.org/10.1039/c1md00093d.

(135) Pardridge, W. M. Transport of Small Molecules through the Blood-Brain Barrier: Biology and Methodology. *Adv Drug Deliv Rev* **1995**, *15* (1–3), 5–36.

(136) Manallack, D. T.; Prankerd, R. J.; Yuriev, E.; Oprea, T. I.; Chalmers, D. K. The Significance of Acid/Base Properties in Drug Discovery. *Chem. Soc. Rev.* **2013**, *42* (2), 485–496. https://doi.org/10.1039/C2CS35348B.

(137) Avdeef, A. *Avdeef A Permeability: Caco-2/MDCK In Absorption and Drug Development*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(138) Lennernäs, H. Intestinal Permeability and Its Relevance for Absorption and Elimination. *Xenobiotica* **2007**, *37* (10–11), 1015–1051. https://doi.org/10.1080/00498250701704819.

(139) Chadwick, V. S.; Phillips, S. F.; Hofmann, A. F. Measurements of Intestinal Permeability Using Low Molecular Weight Polyethylene Glycols (PEG 400). *Gastroenterology* **1977**, *73* (2), 247–251. https://doi.org/10.1016/S0016-5085(19)32197-3.

(140) Vidon, N.; Evard, D.; Godbillon, J.; Rongier, M.; Duval, M.; Schoeller, J.; Bernier, J.; Hirtz, J. Investigation of Drug Absorption from the Gastrointestinal Tract of Man. II. Metoprolol in the Jejunum and Ileum. *British Journal of Clinical Pharmacology* **1985**, *19* (S2), 107S-112S. https://doi.org/10.1111/j.1365-2125.1985.tb02750.x.

(141) Knutson, T.; Fridblom, P.; Ahlström, H.; Magnusson, A.; Tannergren, C.; Lennernäs, H. Increased Understanding of Intestinal Drug Permeability Determined by the LOC-I-GUT Approach Using Multislice Computed Tomography. *Mol. Pharmaceutics* **2009**, *6* (1), 2–10. https://doi.org/10.1021/mp800145r.

(142) M. Reis, J.; Sinko, B.; H.R. Serra, C. Parallel Artificial Membrane Permeability Assay (PAMPA) - Is It Better than Caco-2 for Human Passive Permeability Prediction? *MRMC* **2010**, *10* (11), 1071–1076. https://doi.org/10.2174/1389557511009011071.

(143) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41* (7), 1007–1010. https://doi.org/10.1021/jm970530e.

(144) Herbette, L. G.; Chester, D. W.; Rhodes, D. G. Structural Analysis of Drug Molecules in Biological Membranes. *Biophys J* **1986**, *49* (1), 91–94. https://doi.org/10.1016/S0006-3495(86)83605-0.

(145) Mason, R. P.; Rhodes, D. G.; Herbette, L. G. Reevaluating Equilibrium and Kinetic Binding Parameters for Lipophilic Drugs Based on a Structural Model for Drug Interactions with Biological Membranes. *J. Med. Chem.* **1991**, *34* (3), 869–877. https://doi.org/10.1021/jm00107a001.

(146) Krämer, S. D. Quantitative Aspects of Drug Permeation across in Vitro and in Vivo Barriers. *European Journal of Pharmaceutical Sciences* **2016**, *87*, 30–46. https://doi.org/10.1016/j.ejps.2015.10.013.

(147) Volpe, D. A. Drug-Permeability and Transporter Assays in Caco-2 and MDCK Cell Lines. *Future Medicinal Chemistry* **2011**, *3* (16), 2063–2077. https://doi.org/10.4155/fmc.11.149.

(148) Adson, A.; Burton, P. S.; Raub, T. J.; Barsuhn, C. L.; Audus, K. L.; Ho, N. F. H. Passive Diffusion of Weak Organic Electrolytes across Caco-2 Cell Monolayers: Uncoupling the Contributions of Hydrodynamic, Transcellular, and Paracellular Barriers. *Journal of Pharmaceutical Sciences* **1995**, *84* (10), 1197–1204. https://doi.org/10.1002/jps.2600841011.

(149) Stewart, B. H.; Chan, O. H.; Lu, R. H.; Reyner, E. L.; Schmid, H. L.; Hamilton, H. W.; Steinbaugh, B. A.; Taylor, M. D. Comparison of Intestinal Permeabilities Determined in Multiple in Vitro and in Situ Models: Relationship to Absorption in Humans. *Pharmaceutical Research* **1995**, *12* (5), 693–699. https://doi.org/10.1023/A:1016207525186.

(150) Simmons, N. L. Cultured Monolayers of MDCK Cells: A Novel Model System for the Study of Epithelial Development and Function. *General Pharmacology: The Vascular System* **1982**, *13* (4), 287–291. https://doi.org/10.1016/0306-3623(82)90047-7.

(151) Cho, M. J.; Thompson, D. P.; Cramer, C. T.; Vidmar, T. J.; Scieszka, J. F. The Madin Darby Canine Kidney (MDCK) Epithelial Cell Monolayer as a Model Cellular Transport Barrier. *Pharmaceutical Research* **1989**, *06* (1), 71–77. https://doi.org/10.1023/A:1015807904558.

(152) Sun, H.; Chow, E. C.; Liu, S.; Du, Y.; Pang, K. S. The Caco-2 Cell Monolayer: Usefulness and Limitations. *Expert Opinion on Drug Metabolism & Toxicology* **2008**, *4* (4), 395–411. https://doi.org/10.1517/17425255.4.4.395.

(153) Tsuji, A.; Tamai, I. Carrier-Mediated Intestinal Transport of Drugs. *Pharmaceutical Research* **1996**, *13* (7), 963–977. https://doi.org/10.1023/A:1016086003070.

(154) Pangeni, R.; Kang, S.; Jha, S. K.; Subedi, L.; Park, J. W. Intestinal Membrane Transporter-Mediated Approaches to Improve Oral Drug Delivery. *J. Pharm. Investig.* **2021**, *51* (2), 137–158. https://doi.org/10.1007/s40005-021-00515-1.

(155) Chong, S.; Dando, S. A.; Soucek, K. M.; Morrison, R. A. In Vitro Permeability through Caco-2 Cells Is Not Quantitatively Predictive of in Vivo Absorption for Peptide-like Drugs Absorbed via the Dipeptide Transporter System. *Pharmaceutical Research* **1996**, *13* (1), 120–123. https://doi.org/10.1023/A:1016045820933.

(156) Lennernas, H. Modeling Gastrointestinal Drug Absorption Requires More In Vivo Biopharmaceutical Data: Experience from In Vivo Dissolution and Permeability Studies in Humans. *CDM* **2007**, *8* (7), 645–657. https://doi.org/10.2174/138920007782109823.

(157) Erickson, M. A.; Wilson, M. L.; Banks, W. A. In Vitro Modeling of Blood–Brain Barrier and Interface Functions in Neuroimmune Communication. *Fluids Barriers CNS* **2020**, *17* (1), 26. https://doi.org/10.1186/s12987-020-00187-3.

(158) Urquhart, B. L.; Kim, R. B. Blood−brain Barrier Transporters and Response to CNS-Active Drugs. *Eur J Clin Pharmacol* **2009**, *65* (11), 1063–1070. https://doi.org/10.1007/s00228-009-0714-8.

(159) 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. *Neurobiology of Disease* **2010**, *37* (1), 13–25. https://doi.org/10.1016/j.nbd.2009.07.030.

(160) Pardridge, W. M. Holy Grails and in Vitro Blood–Brain Barrier Models. *Drug Discovery Today* **2004**, *9* (6), 258. https://doi.org/10.1016/S1359-6446(03)03003-4.

(161) Hitchcock, S. A. Blood–Brain Barrier Permeability Considerations for CNS-Targeted Compound Library Design. *Current Opinion in Chemical Biology* **2008**, *12* (3), 318–323. https://doi.org/10.1016/j.cbpa.2008.03.019.

(162) Cecchelli, R.; Berezowski, V.; Lundquist, S.; Culot, M.; Renftel, M.; Dehouck, M.-P.; Fenart, L. Modelling of the Blood–Brain Barrier in Drug Discovery and Development. *Nat Rev Drug Discov* **2007**, *6* (8), 650–661. https://doi.org/10.1038/nrd2368.

(163) Terasaki, T.; Ohtsuki, S.; Hori, S.; Takanaga, H.; Nakashima, E.; Hosoya, K. New Approaches to in Vitro Models of Blood–Brain Barrier Drug Transport. *Drug Discovery Today* **2003**, *8* (20), 944–954. https://doi.org/10.1016/S1359-6446(03)02858-7.

(164) Bickel, U. How to Measure Drug Transport across the Blood-Brain Barrier. *Neurotherapeutics* **2005**, *2* (1), 15–26. https://doi.org/10.1602/neurorx.2.1.15.

(165) Avdeef, A. *Avdeef A Permeability: Blood−Brain Barrier In Absorption and Drug Development*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(166) Pardridge, W. M.; Triguero, D.; Yang, J.; Cancilla, P. A. Comparison of in Vitro and in Vivo Models of Drug Transcytosis through the Blood-Brain Barrier. *J Pharmacol Exp Ther* **1990**, *253* (2), 884–891.

(167) Booth, R.; Kim, H. Characterization of a Microfluidic in Vitro Model of the Blood-Brain Barrier (MBBB). *Lab Chip* **2012**, *12* (10), 1784. https://doi.org/10.1039/c2lc40094d.

(168) Williams-Medina, A.; Deblock, M.; Janigro, D. In Vitro Models of the Blood–Brain Barrier: Tools in Translational Medicine. *Front. Med. Technol.* **2021**, *2*, 623950. https://doi.org/10.3389/fmedt.2020.623950.

(169) Linz, G.; Djeljadini, S.; Steinbeck, L.; Köse, G.; Kiessling, F.; Wessling, M. Cell Barrier Characterization in Transwell Inserts by Electrical Impedance Spectroscopy. *Biosensors and Bioelectronics* **2020**, *165*, 112345. https://doi.org/10.1016/j.bios.2020.112345.

(170) Kansy, M.; Avdeef, A.; Fischer, H. Advances in Screening for Membrane Permeability: High-Resolution PAMPA for Medicinal Chemists. *Drug Discovery Today: Technologies* **2004**, *1* (4), 349–355. https://doi.org/10.1016/j.ddtec.2004.11.013.

(171) Sun, H.; Nguyen, K.; Kerns, E.; Yan, Z.; Yu, K. R.; Shah, P.; Jadhav, A.; Xu, X. Highly Predictive and Interpretable Models for PAMPA Permeability. *Bioorganic & Medicinal Chemistry* **2017**, *25* (3), 1266–1276. https://doi.org/10.1016/j.bmc.2016.12.049.

(172) Sugano, K.; Nabuchi, Y.; Machida, M.; Aso, Y. Prediction of Human Intestinal Permeability Using Artificial Membrane Permeability. *Int J Pharm* **2003**, *257* (1–2), 245–251. https://doi.org/10.1016/s0378-5173(03)00161-3.

(173) Bermejo, M.; Avdeef, A.; Ruiz, A.; Nalda, R.; Ruell, J. A.; Tsinman, O.; González, I.; Fernández, C.; Sánchez, G.; Garrigues, T. M.; Merino, V. PAMPA—a Drug Absorption in Vitro Model. *European Journal of Pharmaceutical Sciences* **2004**, *21* (4), 429–441. https://doi.org/10.1016/j.ejps.2003.10.009.

(174) Zhu, C.; Jiang, L.; Chen, T.-M.; Hwang, K.-K. A Comparative Study of Artificial Membrane Permeability Assay for High Throughput Profiling of Drug Absorption Potential. *European Journal of Medicinal Chemistry* **2002**, *37* (5), 399–407. https://doi.org/10.1016/S0223-5234(02)01360-0.

(175) Ottaviani, G.; Martel, S.; Carrupt, P.-A. Parallel Artificial Membrane Permeability Assay: A New Membrane for the Fast Prediction of Passive Human Skin Permeability. *J Med Chem* **2006**, *49* (13), 3948–3954. https://doi.org/10.1021/jm060230+.

(176) Avdeef, A.; Strafford, M.; Block, E.; Balogh, M. P.; Chambliss, W.; Khan, I. Drug Absorption in Vitro Model: Filter-Immobilized Artificial Membranes. 2. Studies of the Permeability Properties of Lactones in Piper Methysticum Forst. *Eur J Pharm Sci* **2001**, *14* (4), 271–280. https://doi.org/10.1016/s0928-0987(01)00191-9.

(177) Sugano, K.; Nabuchi, Y.; Machida, M.; Asoh, Y. Permeation Characteristics of a Hydrophilic Basic Compound across a Bio-Mimetic Artificial Membrane. *Int J Pharm* **2004**, *275* (1–2), 271–278. https://doi.org/10.1016/j.ijpharm.2004.02.010.

(178) Dagenais, C.; Avdeef, A.; Tsinman, O.; Dudley, A.; Beliveau, R. P-Glycoprotein Deficient Mouse in Situ Blood-Brain Barrier Permeability and Its Prediction Using an in Combo PAMPA Model. *Eur J Pharm Sci* **2009**, *38* (2), 121–137. https://doi.org/10.1016/j.ejps.2009.06.009.

(179) Mensch, J.; Melis, A.; Mackie, C.; Verreck, G.; Brewster, M. E.; Augustijns, P. Evaluation of Various PAMPA Models to Identify the Most Discriminating Method for the Prediction of BBB Permeability. *Eur J Pharm Biopharm* **2010**, *74* (3), 495–502. https://doi.org/10.1016/j.ejpb.2010.01.003.

(180) Bendels, S.; Tsinman, O.; Wagner, B.; Lipp, D.; Parrilla, I.; Kansy, M.; Avdeef, A. PAMPA–Excipient Classification Gradient Map. *Pharm Res* **2006**, *23* (11), 2525–2535. https://doi.org/10.1007/s11095-006-9137-8.

(181) Ismael J. Hidalgo. Assessing the Absorption of New Pharmaceuticals. *CTMC* **2001**, *1* (5), 385–401. https://doi.org/10.2174/1568026013395010.

(182) Mueller, P.; Rudin, D. O.; Ti Tien, H.; Wescott, W. C. Reconstitution of Cell Membrane Structure in Vitro and Its Transformation into an Excitable System. *Nature* **1962**, *194* (4832), 979–980. https://doi.org/10.1038/194979a0.

(183) Tsemperouli, M.; Amstad, E.; Sakai, N.; Matile, S.; Sugihara, K. Black Lipid Membranes: Challenges in Simultaneous Quantitative Characterization by Electrophysiology and Fluorescence Microscopy. *Langmuir* **2019**, *35* (26), 8748–8757. https://doi.org/10.1021/acs.langmuir.9b00673.

(184) Yeagle, P. L. Laboratory Membrane Systems. In *The Membranes of Cells*; Elsevier, 2016; pp 95–114. https://doi.org/10.1016/B978-0-12-800047-2.00006-1.

(185) Weinberger, A.; Tsai, F.-C.; Koenderink, G. H.; Schmidt, T. F.; Itri, R.; Meier, W.; Schmatko, T.; Schröder, A.; Marques, C. Gel-Assisted Formation of Giant Unilamellar Vesicles. *Biophysical Journal* **2013**, *105* (1), 154–164. https://doi.org/10.1016/j.bpj.2013.05.024.

(186) Nishimura, K.; Matsuura, T.; Sunami, T.; Fujii, S.; Nishimura, K.; Suzuki, H.; Yomo, T. Identification of Giant Unilamellar Vesicles with Permeability to Small Charged Molecules. *RSC Adv.* **2014**, *4* (66), 35224. https://doi.org/10.1039/C4RA05332J.

(187) Nishimura, K.; Suzuki, H.; Toyota, T.; Yomo, T. Size Control of Giant Unilamellar Vesicles Prepared from Inverted Emulsion Droplets. *Journal of Colloid and Interface Science* **2012**, *376* (1), 119–125. https://doi.org/10.1016/j.jcis.2012.02.029.

(188) Heider, E. C.; Barhoum, M.; Edwards, K.; Gericke, K.-H.; Harris, J. M. Structural Characterization of Individual Vesicles Using Fluorescence Microscopy. *Anal. Chem.* **2011**, *83* (12), 4909–4915. https://doi.org/10.1021/ac200632h.

(189) Bagatolli, L. A.; Needham, D. Quantitative Optical Microscopy and Micromanipulation Studies on the Lipid Bilayer Membranes of Giant Unilamellar Vesicles. *Chemistry and Physics of Lipids* **2014**, *181*, 99–120. https://doi.org/10.1016/j.chemphyslip.2014.02.009.

(190) Marks, J. R.; Placone, J.; Hristova, K.; Wimley, W. C. Spontaneous Membrane-Translocating Peptides by Orthogonal High-Throughput Screening. *J. Am. Chem. Soc.* **2011**, *133* (23), 8995–9004. https://doi.org/10.1021/ja2017416.

(191) Wohnsland, F.; Faller, B. High-Throughput Permeability PH Profile and High-Throughput Alkane/Water Log *P* with Artificial Membranes. *J. Med. Chem.* **2001**, *44* (6), 923–930. https://doi.org/10.1021/jm001020e.

(192) Thomae, A. V.; Wunderli-Allenspach, H.; Krämer, S. D. Permeation of Aromatic Carboxylic Acids across Lipid Bilayers: The PH-Partition Hypothesis Revisited. *Biophysical Journal* **2005**, *89* (3), 1802–1811. https://doi.org/10.1529/biophysj.105.060871.

(193) Biedermann, F.; Ghale, G.; Hennig, A.; Nau, W. M. Fluorescent Artificial Receptor-Based Membrane Assay (FARMA) for Spatiotemporally Resolved Monitoring of Biomembrane Permeability. *Commun Biol* **2020**, *3* (1), 383. https://doi.org/10.1038/s42003-020-1108-9.

(194) Sezer, D.; Oruç, T. Protonation Kinetics Compromise Liposomal Fluorescence Assay of Membrane Permeation. *J. Phys. Chem. B* **2017**, *121* (20), 5218–5227. https://doi.org/10.1021/acs.jpcb.7b01881.

(195) Eyer, K.; Paech, F.; Schuler, F.; Kuhn, P.; Kissner, R.; Belli, S.; Dittrich, P. S.; Krämer, S. D. A Liposomal Fluorescence Assay to Study Permeation Kinetics of Drug-like Weak Bases across the Lipid Bilayer. *Journal of Controlled Release* **2014**, *173*, 102–109. https://doi.org/10.1016/j.jconrel.2013.10.037.

(196) Wu, X.; Busschaert, N.; Wells, N. J.; Jiang, Y.-B.; Gale, P. A. Dynamic Covalent Transport of Amino Acids across Lipid Bilayers. *J. Am. Chem. Soc.* **2015**, *137* (4), 1476–1484. https://doi.org/10.1021/ja510063n.

(197) Nilam, M.; Collin, S.; Karmacharya, S.; Hennig, A.; Nau, W. M. Membrane Permeability and Its Activation Energies in Dependence on Analyte, Lipid, and Phase Type Obtained by the Fluorescent Artificial Receptor Membrane Assay. *ACS Sens.* **2021**, *6* (1), 175–182. https://doi.org/10.1021/acssensors.0c02064.

(198) Sinn, S.; Biedermann, F. Chemical Sensors Based on Cucurbit[*n*]Uril Macrocycles. *Isr. J. Chem.* **2018**, *58* (3–4), 357–412. https://doi.org/10.1002/ijch.201700118.

(199) Brändén, M.; Tabaei, S. R.; Fischer, G.; Neutze, R.; Höök, F. Refractive-Index-Based Screening of Membrane-Protein-Mediated Transfer across Biological Membranes. *Biophysical Journal* **2010**, *99* (1), 124–133. https://doi.org/10.1016/j.bpj.2010.03.059.

(200) Cooper, M. A. Advances in Membrane Receptor Screening and Analysis. *J. Mol. Recognit.* **2004**, *17* (4), 286–315. https://doi.org/10.1002/jmr.675.

(201) Cooper, M. A. Optical Biosensors in Drug Discovery. *Nat Rev Drug Discov* **2002**, *1* (7), 515–528. https://doi.org/10.1038/nrd838.

(202) Daghestani, H. N.; Day, B. W. Theory and Applications of Surface Plasmon Resonance, Resonant Mirror, Resonant Waveguide Grating, and Dual Polarization Interferometry Biosensors. *Sensors* **2010**, *10* (11), 9630–9646. https://doi.org/10.3390/s101109630.

(203) Brändén, M.; Dahlin, S.; Höök, F. Label-Free Measurements of Molecular Transport across Liposome Membranes Using Evanescent-Wave Sensing. *ChemPhysChem* **2008**, *9* (17), 2480–2485. https://doi.org/10.1002/cphc.200800614.

(204) Baird, C. L.; Courtenay, E. S.; Myszka, D. G. Surface Plasmon Resonance Characterization of Drug/Liposome Interactions. *Analytical Biochemistry* **2002**, *310* (1), 93–99. https://doi.org/10.1016/S0003-2697(02)00278-6.

(205) Danelian, E.; Karlén, A.; Karlsson, R.; Winiwarter, S.; Hansson, A.; Löfås, S.; Lennernäs, H.; Hämäläinen, M. D. SPR Biosensor Studies of the Direct Interaction

between 27 Drugs and a Liposome Surface: Correlation with Fraction Absorbed in Humans. *J. Med. Chem.* **2000**, *43* (11), 2083–2086. https://doi.org/10.1021/jm991156g.

(206) Pfeiffer, I.; Höök, F. Bivalent Cholesterol-Based Coupling of Oligonucletides to Lipid Membrane Assemblies. *J. Am. Chem. Soc.* **2004**, *126* (33), 10224–10225. https://doi.org/10.1021/ja048514b.

(207) Stamou, D.; Duschl, C.; Delamarche, E.; Vogel, H. Self-Assembled Microarrays of Attoliter Molecular Vessels. *Angew. Chem. Int. Ed.* **2003**, *42* (45), 5580–5583. https://doi.org/10.1002/anie.200351866.

(208) Christensen, S. M.; Stamou, D. Surface-Based Lipid Vesicle Reactor Systems: Fabrication and Applications. *Soft Matter* **2007**, *3* (7), 828. https://doi.org/10.1039/b702849k.

(209) Tran, R. J.; Sly, K. L.; Conboy, J. C. Applications of Surface Second Harmonic Generation in Biological Sensing. *Annual Rev. Anal. Chem.* **2017**, *10* (1), 387–414. https://doi.org/10.1146/annurev-anchem-071015-041453.

(210) Roke, S.; Gonella, G. Nonlinear Light Scattering and Spectroscopy of Particles and Droplets in Liquids. *Annu. Rev. Phys. Chem.* **2012**, *63* (1), 353–378. https://doi.org/10.1146/annurev-physchem-032511-143748.

(211) Roke, S. Nonlinear Spectroscopy of Bio-Interfaces. *International Journal of Materials Research* **2011**, *102* (7), 906–912. https://doi.org/10.3139/146.110535.

(212) Gomopoulos, N.; Lütgebaucks, C.; Sun, Q.; Macias-Romero, C.; Roke, S. Label-Free Second Harmonic and Hyper Rayleigh Scattering with High Efficiency. *Opt. Express* **2013**, *21* (1), 815. https://doi.org/10.1364/OE.21.000815.

(213) Salafsky, J. S.; Eisenthal, K. B. Second Harmonic Spectroscopy: Detection and Orientation of Molecules at a Biomembrane Interface. *Chemical Physics Letters* **2000**, *319* (5–6), 435–439. https://doi.org/10.1016/S0009-2614(00)00116-0.

(214) Liu, Y.; Yan, E. C. Y.; Zhao, X.; Eisenthal, K. B. Surface Potential of Charged Liposomes Determined by Second Harmonic Generation. *Langmuir* **2001**, *17* (7), 2063–2066. https://doi.org/10.1021/la0011634.

(215) Liu, J.; Subir, M.; Nguyen, K.; Eisenthal, K. B. Second Harmonic Studies of Ions Crossing Liposome Membranes in Real Time. *J. Phys. Chem. B* **2008**, *112* (48), 15263–15266. https://doi.org/10.1021/jp806690z.

(216) Zeng, J.; Eckenrode, H. M.; Dounce, S. M.; Dai, H.-L. Time-Resolved Molecular Transport across Living Cell Membranes. *Biophysical Journal* **2013**, *104* (1), 139–145. https://doi.org/10.1016/j.bpj.2012.11.3814.

(217) Stokes, G. Y.; Conboy, J. C. Measuring Selective Estrogen Receptor Modulator (SERM)–Membrane Interactions with Second Harmonic Generation. *J. Am. Chem. Soc.* **2014**, *136* (4), 1409–1417. https://doi.org/10.1021/ja409250y.

(218) Nguyen, T. T.; Conboy, J. C. High-Throughput Screening of Drug–Lipid Membrane Interactions via Counter-Propagating Second Harmonic Generation Imaging. *Anal. Chem.* **2011**, *83* (15), 5979–5988. https://doi.org/10.1021/ac2009614.

(219) Moreaux, L.; Sandre, O.; Mertz, J. Membrane Imaging by Second-Harmonic Generation Microscopy. *J. Opt. Soc. Am. B* **2000**, *17* (10), 1685. https://doi.org/10.1364/JOSAB.17.001685.

(220) Hermann, J. P.; Ducuing, J. Absolute Measurement of Two-Photon Cross Sections. *Phys. Rev. A* **1972**, *5* (6), 2557–2568. https://doi.org/10.1103/PhysRevA.5.2557.

(221) Heinz, T. F.; Chen, C. K.; Ricard, D.; Shen, Y. R. Spectroscopy of Molecular Monolayers by Resonant Second-Harmonic Generation. *Phys. Rev. Lett.* **1982**, *48* (7), 478–481. https://doi.org/10.1103/PhysRevLett.48.478.

(222) Shen, Y. R. Surface Properties Probed by Second-Harmonic and Sum-Frequency Generation. *Nature* **1989**, *337* (6207), 519–525. https://doi.org/10.1038/337519a0.

(223) Eisenthal, K. B. Second Harmonic Spectroscopy of Aqueous Nano- and Microparticle Interfaces. *Chem. Rev.* **2006**, *106* (4), 1462–1477. https://doi.org/10.1021/cr0403685.

(224) Fuentes-Corona, C. G.; Licea-Rodriguez, J.; Younger, R.; Rangel-Rojo, R.; Potma, E. O.; Rocha-Mendoza, I. Second Harmonic Generation Signal from Type I Collagen Fibers Grown in Vitro. *Biomed. Opt. Express* **2019**, *10* (12), 6449. https://doi.org/10.1364/BOE.10.006449.

(225) *Transport Processes in Pharmaceutical Systems: Biological Transport Phenomena in the Gastrointestinal Tract.*, 0 ed.; Amidon, G. L., Lee, P. I., Topp, E. M., Eds.; CRC Press, 1999. https://doi.org/10.1201/9780203909478.

(226) Wilson, J. P. Surface Area of the Small Intestine in Man. *Gut* **1967**, *8* (6), 618–621. https://doi.org/10.1136/gut.8.6.618.

(227) Travis, S.; Menzies, I. Intestinal Permeability: Functional Assessment and Significance. *Clinical Science* **1992**, *82* (5), 471–488. https://doi.org/10.1042/cs0820471.

(228) Oliver, R. E.; Jones, A. F.; Rowland, M. What Surface of the Intestinal Epithelium Is Effectively Available to Permeating Drugs? *Journal of Pharmaceutical Sciences* **1998**, *87* (5), 634–639. https://doi.org/10.1021/js9701216.

(229) Ano, R.; Kimura, Y.; Shima, M.; Matsuno, R.; Ueno, T.; Akamatsu, M. Relationships between Structure and High-Throughput Screening Permeability of Peptide Derivatives and Related Compounds with Artificial Membranes: Application to Prediction of Caco-2 Cell Permeability. *Bioorganic & Medicinal Chemistry* **2004**, *12* (1), 257–264. https://doi.org/10.1016/j.bmc.2003.10.002.

(230) Yamashita, S.; Furubayashi, T.; Kataoka, M.; Sakane, T.; Sezaki, H.; Tokuda, H. Optimized Conditions for Prediction of Intestinal Drug Permeability Using Caco-2 Cells. *European Journal of Pharmaceutical Sciences* **2000**, *10* (3), 195–204. https://doi.org/10.1016/S0928-0987(00)00076-2.

(231) Avdeef, A. *Avdeef A PKa Determination In Absorption and Drug Development*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

(232) Yamashita, S.; Tanaka, Y.; Endoh, Y.; Taki, Y.; Sakane, T.; Nadai, T.; Sezaki, H. Analysis of Drug Permeation across Caco-2 Monolayer: Implication for Predicting in Vivo Drug Absorption. *Pharmaceutical Research* **1997**, *14* (4), 486–491. https://doi.org/10.1023/A:1012103700981.

(233) Di Meo, F.; Fabre, G.; Berka, K.; Ossman, T.; Chantemargue, B.; Paloncýová, M.; Marquet, P.; Otyepka, M.; Trouillas, P. In Silico Pharmacology: Drug Membrane Partitioning and Crossing. *Pharmacological Research* **2016**, *111*, 471–486. https://doi.org/10.1016/j.phrs.2016.06.030.

(234) Manrique, R.; Wu, W.; Chang, J.-S. A Molecular Dynamics Study on the CO2 Permeability of Microalgae Lipid Membrane. *J Appl Phycol* **2020**, *32* (1), 291–297. https://doi.org/10.1007/s10811-019-01940-4.

(235) Cao, Z.; Zhang, X.; Wang, C.; Liu, L.; Zhao, L.; Wang, J.; Zhou, Y. Different Effects of Cholesterol on Membrane Permeation of Arginine and Tryptophan Revealed by Bias-Exchange Metadynamics Simulations. *J. Chem. Phys.* **2019**, *150* (8), 084106. https://doi.org/10.1063/1.5082351.

(236) Hénin, J.; Lelièvre, T.; Shirts, M. R.; Valsson, O.; Delemotte, L. Enhanced Sampling Methods for Molecular Dynamics Simulations. arXiv February 8, 2022.

(237) Hummer, G. Position-Dependent Diffusion Coefficients and Free Energies from Bayesian Analysis of Equilibrium and Replica Molecular Dynamics Simulations. *New J. Phys.* **2005**, *7*, 34–34. https://doi.org/10.1088/1367-2630/7/1/034.

(238) Li, J. BASIC MOLECULAR DYNAMICS. 24.

(239) Schneider, R.; Sharma, A. R.; Rai, A. Introduction to Molecular Dynamics. In *Computational Many-Particle Physics*; Fehske, H., Schneider, R., Weiße, A., Eds.; Lecture Notes in Physics; Springer: Berlin, Heidelberg, 2008; pp 3–40. https://doi.org/10.1007/978-3-540-74686-7_1.

(240) Schlaich, A.; Kesselheim, S.; Rühle, F. 1: Equations of Motion and Integrators. 8.

(241) Lindahl; Abraham; Hess; Spoel, V. D. GROMACS 2020.3 Manual. **2020**. https://doi.org/10.5281/ZENODO.3923644.

(242) Cardenas, A. E.; Elber, R. Modeling Kinetics and Equilibrium of Membranes with Fields: Milestoning Analysis and Implication to Permeation. *J Chem Phys* **2014**, *141* (5), 054101. https://doi.org/10.1063/1.4891305.

(243) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *Journal of Computational Physics* **1977**, *23* (2), 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.

(244) Roux, B. The Calculation of the Potential of Mean Force Using Computer Simulations. *Computer Physics Communications* **1995**, *91* (1), 275–282. https://doi.org/10.1016/0010-4655(95)00053-I.

(245) Sun, R.; Dama, J. F.; Tan, J. S.; Rose, J. P.; Voth, G. A. Transition-Tempered Metadynamics Is a Promising Tool for Studying the Permeation of Drug-like Molecules through Membranes. *J. Chem. Theory Comput.* **2016**, *12* (10), 5157–5169. https://doi.org/10.1021/acs.jctc.6b00206.

(246) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100* (2), 020603. https://doi.org/10.1103/PhysRevLett.100.020603.

(247) Lindahl, V.; Lidmar, J.; Hess, B. Accelerated Weight Histogram Method for Exploring Free Energy Landscapes. *The Journal of Chemical Physics* **2014**, *141* (4), 044110. https://doi.org/10.1063/1.4890371.

(248) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. Basic Ingredients of Free Energy Calculations: A Review. *J. Comput. Chem.* **2009**, NA-NA. https://doi.org/10.1002/jcc.21450.

(249) Aldeghi, M.; Bluck, J. P.; Biggin, P. C. Absolute Alchemical Free Energy Calculations for Ligand Binding: A Beginner's Guide. In *Computational Drug Discovery and Design*; Gore, M., Jagtap, U. B., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; Vol. 1762, pp 199–232. https://doi.org/10.1007/978-1-4939-7756-7_11.

(250) Mey, A. S. J. S.; Allen, B. K.; Bruce Macdonald, H. E.; Chodera, J. D.; Hahn, D. F.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A.; Scheen, J.; Shirts, M. R.; Tresadern, G.; Xu, H. Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *LiveCoMS* **2020**, *2* (1). https://doi.org/10.33011/livecoms.2.1.18378.

(251) Paliwal, H.; Shirts, M. R. A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J. Chem. Theory Comput.* **2011**, *7* (12), 4115–4134. https://doi.org/10.1021/ct2003995.

(252) Tan, Z. On a Likelihood Approach for Monte Carlo Integration. *Journal of the American Statistical Association* **2004**, *99* (468), 1027–1036. https://doi.org/10.1198/016214504000001664.

(253) Kästner, J. Umbrella Sampling. *WIREs Computational Molecular Science* **2011**, *1* (6), 932–942. https://doi.org/10.1002/wcms.66.

(254) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *Journal of Computational Chemistry* **1992**, *13* (8), 1011–1021. https://doi.org/10.1002/jcc.540130812.

(255) Oshima, H.; Re, S.; Sugita, Y. Replica-Exchange Umbrella Sampling Combined with Gaussian Accelerated Molecular Dynamics for Free-Energy Calculation of Biomolecules. *J. Chem. Theory Comput.* **2019**, *15* (10), 5199–5208. https://doi.org/10.1021/acs.jctc.9b00761.

(256) Olson, M. A.; Lee, M. S. Application of Replica Exchange Umbrella Sampling to Protein Structure Refinement of Nontemplate Models. *J Comput Chem* **2013**, *34* (20), 1785–1793. https://doi.org/10.1002/jcc.23325.

(257) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annual Review of Physical Chemistry* **2016**, *67* (1), 159–184. https://doi.org/10.1146/annurev-physchem-040215-112229.

(258) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proceedings of the National Academy of Sciences* **2002**, *99* (20), 12562–12566. https://doi.org/10.1073/pnas.202427399.

(259) Bussi, G.; Laio, A.; Tiwary, P. Metadynamics: A Unified Framework for Accelerating Rare Events and Sampling Thermodynamics and Kinetics. In *Handbook of Materials Modeling: Methods: Theory and Modeling*; Andreoni, W., Yip, S., Eds.; Springer International Publishing: Cham, 2020; pp 565–595. https://doi.org/10.1007/978-3-319-44677-6_49.

(260) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-Tempered Metadynamics Converges Asymptotically. *Phys. Rev. Lett.* **2014**, *112* (24), 240602. https://doi.org/10.1103/PhysRevLett.112.240602.

(261) Tiwary, P.; Parrinello, M. A Time-Independent Free Energy Estimator for Metadynamics. *J. Phys. Chem. B* **2015**, *119* (3), 736–742. https://doi.org/10.1021/jp504920s.

(262) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics. *Journal of Computational Chemistry* **2009**, *30* (11), 1615–1621. https://doi.org/10.1002/jcc.21305.

(263) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with Adaptive Gaussians. *J. Chem. Theory Comput.* **2012**, *8* (7), 2247–2254. https://doi.org/10.1021/ct3002464.

(264) Schäfer, T. M.; Settanni, G. Data Reweighting in Metadynamics Simulations. *J. Chem. Theory Comput.* **2020**, *16* (4), 2042–2052. https://doi.org/10.1021/acs.jctc.9b00867.

(265) Marinova, V.; Salvalaglio, M. Time-Independent Free Energies from Metadynamics via Mean Force Integration. *J. Chem. Phys.* **2019**, *151* (16), 164115. https://doi.org/10.1063/1.5123498.

(266) *Free Energy Calculations of Membrane Permeation: Challenges Due to Strong Headgroup–Solute Interactions | Journal of Chemical Theory and Computation*. https://pubs.acs.org/doi/10.1021/acs.jctc.7b01159 (accessed 2022-08-11).

(267) Darve, E.; Pohorille, A. Calculating Free Energies Using Average Force. *The Journal of Chemical Physics* **2001**, *115* (20), 9169–9183. https://doi.org/10.1063/1.1410978.

(268) Darve, E.; Wilson, M. A.; Pohorille, A. Calculating Free Energies Using a Scaled-Force Molecular Dynamics Algorithm. *Molecular Simulation* **2002**, *28* (1–2), 113–144. https://doi.org/10.1080/08927020211975.

(269) Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. The Adaptive Biasing Force Method: Everything You Always Wanted To Know but Were Afraid To Ask. *J. Phys. Chem. B* **2015**, *119* (3), 1129–1151. https://doi.org/10.1021/jp506633n.

(270) Chipot, C.; Comer, J. Subdiffusion in Membrane Permeation of Small Molecules. *Sci Rep* **2016**, *6* (1), 35913. https://doi.org/10.1038/srep35913.

(271) Hénin, J.; Tajkhorshid, E.; Schulten, K.; Chipot, C. Diffusion of Glycerol through Escherichia Coli Aquaglyceroporin GlpF. *Biophysical Journal* **2008**, *94* (3), 832–839. https://doi.org/10.1529/biophysj.107.115105.

(272) Hénin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. Exploring Multidimensional Free Energy Landscapes Using Time-Dependent Biases on Collective Variables. *J. Chem. Theory Comput.* **2010**, *6* (1), 35–47. https://doi.org/10.1021/ct9004432.

(273) Weroński, P.; Jiang, Y.; Rasmussen, S. Molecular Dynamics Study of Small PNA Molecules in Lipid-Water System. *Biophysical Journal* **2007**, *92* (9), 3081–3091. https://doi.org/10.1529/biophysj.106.097352.

(274) Gorfe, A. A.; Babakhani, A.; McCammon, J. A. Membrane Insertion Free Energy Profile of H-Ras Membrane Anchor. *Angew Chem Int Ed Engl* **2007**, *46* (43), 8234–8237. https://doi.org/10.1002/anie.200702379.

(275) Vaitheeswaran, S.; Thirumalai, D. Interactions between Amino Acid Side Chains in Cylindrical Hydrophobic Nanopores with Applications to Peptide Stability. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (46), 17636–17641. https://doi.org/10.1073/pnas.0803990105.

(276) Comer, J.; Schulten, K.; Chipot, C. Calculation of Lipid-Bilayer Permeabilities Using an Average Force. *J. Chem. Theory Comput.* **2014**, *10* (2), 554–564. https://doi.org/10.1021/ct400925s.

(277) Lidmar, J. Improving the Efficiency of Extended Ensemble Simulations: The Accelerated Weight Histogram Method. *Phys. Rev. E* **2012**, *85* (5), 056708. https://doi.org/10.1103/PhysRevE.85.056708.

(278) Lindahl, V.; Gourdon, P.; Andersson, M.; Hess, B. Permeability and Ammonia Selectivity in Aquaporin TIP2;1: Linking Structure to Function. *Sci Rep* **2018**, *8* (1), 2995. https://doi.org/10.1038/s41598-018-21357-2.

(279) Bauer, P.; Hess, B.; Lindahl, E. GROMACS 2022.2 Manual. **2022**. https://doi.org/10.5281/zenodo.6637572.

(280) Chodera, J. D.; Shirts, M. R. Replica Exchange and Expanded Ensemble Simulations as Gibbs Sampling: Simple Improvements for Enhanced Mixing. *J. Chem. Phys.* **2011**, *135* (19), 194110. https://doi.org/10.1063/1.3660669.

(281) Maginn, E. J.; Messerly, R. A.; Carlson, D. J.; Roe, D. R.; Elliot, J. R. Best Practices for Computing Transport Properties 1. Self-Diffusivity and Viscosity from Equilibrium Molecular Dynamics [Article v1.0]. *Living Journal of Computational Molecular Science* **2019**, *1* (1), 6324–6324. https://doi.org/10.33011/livecoms.1.1.6324.

(282) Carpenter, T. S.; Kirshner, D. A.; Lau, E. Y.; Wong, S. E.; Nilmeier, J. P.; Lightstone, F. C. A Method to Predict Blood-Brain Barrier Permeability of Drug-Like Compounds Using Molecular Dynamics Simulations. *Biophysical Journal* **2014**, *107* (3), 630–641. https://doi.org/10.1016/j.bpj.2014.06.024.

(283) Bonhenry, D.; Tarek, M.; Dehez, F. Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane. *J. Chem. Theory Comput.* **2013**, *9* (12), 5675–5684. https://doi.org/10.1021/ct400576e.

(284) Gaalswyk, K.; Awoonor-Williams, E.; Rowley, C. N. Generalized Langevin Methods for Calculating Transmembrane Diffusivity. *J. Chem. Theory Comput.* **2016**, *12* (11), 5609–5619. https://doi.org/10.1021/acs.jctc.6b00747.

(285) Woolf, T. B.; Roux, B. Molecular Dynamics Simulation of the Gramicidin Channel in a Phospholipid Bilayer. *Proc Natl Acad Sci U S A* **1994**, *91* (24), 11631–11635.

(286) Woolf, T. B.; Roux, B. Conformational Flexibility of O-Phosphorylcholine and o-Phosphorylethanolamine: A Molecular Dynamics Study of Solvation Effects. *J. Am. Chem. Soc.* **1994**, *116* (13), 5916–5926. https://doi.org/10.1021/ja00092a048.

(287) Roux, B.; Karplus, M. Ion Transport in a Gramicidin-like Channel: Dynamics and Mobility. *J. Phys. Chem.* **1991**, *95* (12), 4856–4868. https://doi.org/10.1021/j100165a049.

(288) Sriraman, S.; Kevrekidis, I. G.; Hummer, G. Coarse Master Equation from Bayesian Analysis of Replica Molecular Dynamics Simulations. *J. Phys. Chem. B* **2005**, *109* (14), 6479–6484. https://doi.org/10.1021/jp046448u.

(289) Comer, J.; Chipot, C.; González-Nilo, F. D. Calculating Position-Dependent Diffusivity in Biased Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (2), 876–882. https://doi.org/10.1021/ct300867e.

(290) Mathews, J. H.; Fink, K. D. *Numerical Methods Using MATLAB*, 4th ed.; Prentice-Hall of India: New Delhi, 2005.

(291) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **1953**, *21* (6), 1087–1092. https://doi.org/10.1063/1.1699114.

(292) Benmameri, M.; Chantemargue, B.; Humeau, A.; Trouillas, P.; Fabre, G. Accelerated Weight Histogram Method to Assess Membrane Permeability. **2022**. https://doi.org/10.26434/chemrxiv-2022-rk82h.

(293) Comer, J. DiffusionFusion, 2022.

(294) *Time Fractional Diffusion: A Discrete Random Walk Approach | SpringerLink*. https://link.springer.com/article/10.1023/A:1016547232119 (accessed 2022-08-25).

(295) Weiss, M.; Elsner, M.; Kartberg, F.; Nilsson, T. Anomalous Subdiffusion Is a Measure for Cytoplasmic Crowding in Living Cells. *Biophys J* **2004**, *87* (5), 3518–3524. https://doi.org/10.1529/biophysj.104.044263.

(296) Kou, S. C.; Xie, X. S. Generalized Langevin Equation with Fractional Gaussian Noise: Subdiffusion within a Single Protein Molecule. *Phys. Rev. Lett.* **2004**, *93* (18), 180603. https://doi.org/10.1103/PhysRevLett.93.180603.

(297) *Phys. Rev. E 70, 051915 (2004) - Fractional diffusion modeling of ion channel gating*. https://journals.aps.org/pre/abstract/10.1103/PhysRevE.70.051915 (accessed 2022-08-25).

(298) Flenner, E.; Das, J.; Rheinstädter, M. C.; Kosztin, I. Subdiffusion and Lateral Diffusion Coefficient of Lipid Atoms and Molecules in Phospholipid Bilayers. *Phys Rev E Stat Nonlin Soft Matter Phys* **2009**, *79* (1 Pt 1), 011907. https://doi.org/10.1103/PhysRevE.79.011907.

(299) Cardenas, A. E.; Elber, R. Markovian and Non-Markovian Modeling of Membrane Dynamics with Milestoning. *J. Phys. Chem. B* **2016**, *120* (33), 8208–8216. https://doi.org/10.1021/acs.jpcb.6b01890.

(300) Metzler, R.; Klafter, J. The Random Walk's Guide to Anomalous Diffusion: A Fractional Dynamics Approach. *Physics Reports* **2000**, *339* (1), 1–77. https://doi.org/10.1016/S0370-1573(00)00070-3.

(301) Sweilam, N. Crank-Nicolson Finite Difference Method for Solving Time-Fractional Diffusion Equation. *Journal of Fractional Calculus and Applications* **2012**, *2*, 1–9.

(302) Diedenhofen, M.; Eckert, F.; Klamt, A. Prediction of Infinite Dilution Activity Coefficients of Organic Compounds in Ionic Liquids Using COSMO-RS. *J. Chem. Eng. Data* **2003**, *48* (3), 475–479. https://doi.org/10.1021/je025626e.

(303) Klamt, A.; Eckert, F.; Hornig, M.; Beck, M. E.; Bürger, T. Prediction of Aqueous Solubility of Drugs and Pesticides with COSMO-RS. *J Comput Chem* **2002**, *23* (2), 275–281. https://doi.org/10.1002/jcc.1168.

(304) Klamt, A.; Huniar, U.; Spycher, S.; Keldenich, J. COSMOmic: A Mechanistic Approach to the Calculation of Membrane−Water Partition Coefficients and Internal Distributions within Membranes and Micelles. *J. Phys. Chem. B* **2008**, *112* (38), 12148–12157. https://doi.org/10.1021/jp801736k.

(305) Jakobtorweihen, S.; Zuniga, A. C.; Ingram, T.; Gerlach, T.; Keil, F. J.; Smirnova, I. Predicting Solute Partitioning in Lipid Bilayers: Free Energies and Partition Coefficients from Molecular Dynamics Simulations and COSMOmic. *J. Chem. Phys.* **2014**, *141* (4), 045102. https://doi.org/10.1063/1.4890877.

(306) Hornig, M.; Klamt, A. COSMOfrag: A Novel Tool for High-Throughput ADME Property Prediction and Similarity Screening Based on Quantum Chemistry. *J Chem Inf Model* **2005**, *45* (5), 1169–1177. https://doi.org/10.1021/ci0501948.

(307) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, No. 5, 799–805. https://doi.org/10.1039/P29930000799.

(308) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99* (7), 2224–2235. https://doi.org/10.1021/j100007a062.

(309) Marsh, K. N. COSMO-RS from Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design. By A. Klamt. Elsevier: Amsterdam, The Netherlands, 2005. 246 Pp. $US 165. ISBN 0-444-51994-7. *J. Chem. Eng. Data* **2006**, *51* (4), 1480–1480. https://doi.org/10.1021/je0602317.

(310) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102* (26), 5074–5085. https://doi.org/10.1021/jp980017s.

(311) Klamt, A.; Eckert, F. COSMO-RS: A Novel and Efficient Method for the a Priori Prediction of Thermophysical Data of Liquids. *Fluid Phase Equilibria* **2000**, *172* (1), 43–72. https://doi.org/10.1016/S0378-3812(00)00357-5.

(312) Schwöbel, J. A. H.; Ebert, A.; Bittermann, K.; Huniar, U.; Goss, K.-U.; Klamt, A. COSMO *Perm:* Mechanistic Prediction of Passive Membrane Permeability for Neutral Compounds and Ions and Its PH Dependence. *J. Phys. Chem. B* **2020**, *124* (16), 3343–3354. https://doi.org/10.1021/acs.jpcb.9b11728.

(313) Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. Anisotropic Solvent Model of the Lipid Bilayer. 2. Energetics of Insertion of Small Molecules, Peptides, and Proteins in Membranes. *J Chem Inf Model* **2011**, *51* (4), 930–946. https://doi.org/10.1021/ci200020k.

(314) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. A.; Mosberg, H. I. Positioning of Proteins in Membranes: A Computational Approach. *Protein Sci* **2006**, *15* (6), 1318–1333. https://doi.org/10.1110/ps.062126106.

(315) Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. Anisotropic Solvent Model of the Lipid Bilayer. 1. Parameterization of Long-Range Electrostatics and First Solvation Shell Effects. *J. Chem. Inf. Model.* **2011**, *51* (4), 918–929. https://doi.org/10.1021/ci2000192.

(316) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396. https://doi.org/10.1021/jp810292n.

(317) Parisio, G.; Ferrarini, A. Solute Partitioning into Lipid Bilayers: An Implicit Model for Nonuniform and Ordered Environment. *J. Chem. Theory Comput.* **2010**, *6* (8), 2267–2280. https://doi.org/10.1021/ct100210u.

(318) Leung, S. S. F.; Mijalkovic, J.; Borrelli, K.; Jacobson, M. P. Testing Physical Models of Passive Membrane Permeation. *J Chem Inf Model* **2012**, *52* (6), 1621–1636. https://doi.org/10.1021/ci200583t.

(319) Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J Am Chem Soc* **2006**, *128* (43), 14073–14080. https://doi.org/10.1021/ja063076p.

(320) Swift, R. V.; Amaro, R. E. Modeling the Pharmacodynamics of Passive Membrane Permeability. *J Comput Aided Mol Des* **2011**, *25* (11), 1007–1017. https://doi.org/10.1007/s10822-011-9480-7.

(321) Swift, R. V.; Amaro, R. E. Back to the Future: Can Physical Models of Passive Membrane Permeability Help Reduce Drug Candidate Attrition and Move Us beyond QSPR? *Chem Biol Drug Des* **2013**, *81* (1), 61–71. https://doi.org/10.1111/cbdd.12074.

(322) Brocke, S. A.; Degen, A.; MacKerell, A. D.; Dutagaci, B.; Feig, M. Prediction of Membrane Permeation of Drug Molecules by Combining an Implicit Membrane Model with Machine Learning. *J. Chem. Inf. Model.* **2019**, *59* (3), 1147–1162. https://doi.org/10.1021/acs.jcim.8b00648.

(323) Batchelor, H.; Gershkovich, P. Introduction to Biopharmaceutics Measures. In *Biopharmaceutics*; Batchelor, H., Ed.; Wiley, 2022; pp 31–38. https://doi.org/10.1002/9781119678366.ch3.

(324) Klocke, R. A. Velocity of CO2 Exchange in Blood. 13.

(325) Arrowsmith, J.; Miller, P. Phase II and Phase III Attrition Rates 2011–2012. *Nature Reviews Drug Discovery* **2013**, *12* (8), 569–569. https://doi.org/10.1038/nrd4090.

(326) Avdeef, A. Permeability: Caco-2/MDCK. In *Absorption and Drug Development*; John Wiley & Sons, Ltd, 2012; pp 499–574. https://doi.org/10.1002/9781118286067.ch8.

(327) *Planar Lipid Bilayers*; Elsevier, 1993. https://doi.org/10.1016/C2009-0-03331-5.

(328) Fujikawa, M.; Nakao, K.; Shimizu, R.; Akamatsu, M. QSAR Study on Permeability of Hydrophobic Compounds with Artificial Membranes. *Bioorganic & Medicinal Chemistry* **2007**, *15* (11), 3756–3767. https://doi.org/10.1016/j.bmc.2007.03.040.

(329) Bloch, I. A Theory of Membrane Permeability: I. *Bulletin of Mathematical Biophysics* **1944**, *6* (3), 85–92. https://doi.org/10.1007/BF02478238.

(330) Finkelstein, A. Water and Nonelectrolyte Permeability of Lipid Bilayer Membranes. *Journal of General Physiology* **1976**, *68* (2), 127–135. https://doi.org/10.1085/jgp.68.2.127.

(331) MacCallum, J. L.; Tieleman, D. P. Chapter 8 Interactions between Small Molecules and Lipid Bilayers. In *Current Topics in Membranes*; Elsevier, 2008; Vol. 60, pp 227–256. https://doi.org/10.1016/S1063-5823(08)00008-2.

(332) Orsi, M.; Sanderson, W. E.; Essex, J. W. Permeability of Small Molecules through a Lipid Bilayer: A Multiscale Simulation Study. *J. Phys. Chem. B* **2009**, *113* (35), 12019–12029. https://doi.org/10.1021/jp903248s.

(333) Schwöbel, J. A. H.; Klamt, A. Mechanistic Skin Penetration Model by the COSMOperm Method: Routes of Permeation, Vehicle Effects and Skin Variations in the Healthy and Compromised Skin. *Computational Toxicology* **2019**, *11*, 50–64. https://doi.org/10.1016/j.comtox.2019.02.004.

(334) Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. *The Journal of Chemical Physics* **2008**, *128* (14), 144120. https://doi.org/10.1063/1.2829861.

(335) Cao, Z.; Bian, Y.; Hu, G.; Zhao, L.; Kong, Z.; Yang, Y.; Wang, J.; Zhou, Y. Bias-Exchange Metadynamics Simulation of Membrane Permeation of 20 Amino Acids. *IJMS* **2018**, *19* (3), 885. https://doi.org/10.3390/ijms19030885.

(336) Noé, F.; Rosta, E. Markov Models of Molecular Kinetics. *J. Chem. Phys.* **2019**, *151* (19), 190401. https://doi.org/10.1063/1.5134029.

(337) Rosta, E.; Hummer, G. Free Energies from Dynamic Weighted Histogram Analysis Using Unbiased Markov State Model. *J. Chem. Theory Comput.* **2015**, *11* (1), 276–285. https://doi.org/10.1021/ct500719p.

(338) Wu, H.; Paul, F.; Wehmeyer, C.; Noé, F. Multiensemble Markov Models of Molecular Thermodynamics and Kinetics. *Proc Natl Acad Sci USA* **2016**, *113* (23), E3221–E3230. https://doi.org/10.1073/pnas.1525092113.

(339) Marrink, S. J.; Berendsen, H. J. C. Permeation Process of Small Molecules across Lipid Membranes Studied by Molecular Dynamics Simulations. *J. Phys. Chem.* **1996**, *100* (41), 16729–16738. https://doi.org/10.1021/jp952956f.

(340) Comer, J.; Schulten, K.; Chipot, C. Permeability of a Fluid Lipid Bilayer to Short-Chain Alcohols from First Principles. *J. Chem. Theory Comput.* **2017**, *13* (6), 2523–2532. https://doi.org/10.1021/acs.jctc.7b00264.

(341) Ivanyuk, A.; Livio, F.; Biollaz, J.; Buclin, T. Renal Drug Transporters and Drug Interactions. *Clin Pharmacokinet* **2017**, *56* (8), 825–892. https://doi.org/10.1007/s40262-017-0506-8.

(342) Bobrowska-Hägerstrand, M.; Lillås, M.; Mrówczyñska, L.; Wróbel, A.; Shirataki, Y.; Motohashi, N.; Hägerstrand, H. Resveratrol Oligomers Are Potent MRP1 Transport Inhibitors. *Anticancer Res* **2006**, *26* (3A), 2081–2084.

(343) Planas, J. M.; Alfaras, I.; Colom, H.; Juan, M. E. The Bioavailability and Distribution of Trans-Resveratrol Are Constrained by ABC Transporters. *Archives of Biochemistry and Biophysics* **2012**, *527* (2), 67–73. https://doi.org/10.1016/j.abb.2012.06.004.

(344) Chen, Y.; Yuan, X.; Xiao, Z.; Jin, H.; Zhang, L.; Liu, Z. Discovery of Novel Multidrug Resistance Protein 4 (MRP4) Inhibitors as Active Agents Reducing Resistance to Anticancer Drug 6-Mercaptopurine (6-MP) by Structure and Ligand-Based Virtual Screening. *PLoS ONE* **2018**, *13* (10), e0205175. https://doi.org/10.1371/journal.pone.0205175.

(345) Cheung, L.; Flemming, C. L.; Watt, F.; Masada, N.; Yu, D. M. T.; Huynh, T.; Conseil, G.; Tivnan, A.; Polinsky, A.; Gudkov, A. V.; Munoz, M. A.; Vishvanath, A.; Cooper, D. M. F.; Henderson, M. J.; Cole, S. P. C.; Fletcher, J. I.; Haber, M.; Norris, M. D. High-Throughput Screening Identifies Ceefourin 1 and Ceefourin 2 as Highly Selective Inhibitors of Multidrug Resistance Protein 4 (MRP4). *Biochemical Pharmacology* **2014**, *91* (1), 97–108. https://doi.org/10.1016/j.bcp.2014.05.023.

(346) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics* **1954**, *22* (8), 1420–1426. https://doi.org/10.1063/1.1740409.

(347) Lindahl, V.; Lidmar, J.; Hess, B. Riemann Metric Approach to Optimal Sampling of Multidimensional Free-Energy Landscapes. *Phys. Rev. E* **2018**, *98* (2), 023312. https://doi.org/10.1103/PhysRevE.98.023312.

(348) Lindahl, V.; Villa, A.; Hess, B. Sequence Dependency of Canonical Base Pair Opening in the DNA Double Helix. *PLoS Comput Biol* **2017**, *13* (4), e1005463. https://doi.org/10.1371/journal.pcbi.1005463.

(349) Frallicciardi, J.; Melcr, J.; Siginou, P.; Marrink, S. J.; Poolman, B. Membrane Thickness, Lipid Phase and Sterol Type Are Determining Factors in the Permeability of Membranes to Small Solutes. *Nat Commun* **2022**, *13* (1), 1605. https://doi.org/10.1038/s41467-022-29272-x.

(350) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865. https://doi.org/10.1002/jcc.20945.

(351) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(352) Small Molecules Dihedral Parametrization.

(353) O'Boyle, N. M.; Vandermeersch, T.; Flynn, C. J.; Maguire, A. R.; Hutchison, G. R. Confab - Systematic Generation of Diverse Low-Energy Conformers. *J Cheminform* **2011**, *3* (1), 8. https://doi.org/10.1186/1758-2946-3-8.

(354) Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016. Gaussian09.

(355) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24* (16), 1999–2012. https://doi.org/10.1002/jcc.10349.

(356) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *Journal of Computational Chemistry* **1998**, *18* (12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

(357) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *The Journal of Chemical Physics* **1995**, *103* (19), 8577–8593. https://doi.org/10.1063/1.470117.

(358) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. *J Comput Aided Mol Des* **2015**, *29* (5), 397–411. https://doi.org/10.1007/s10822-015-9840-9.

(359) Pymbar.

(360) Alchemlyb.

(361) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *The Journal of Chemical Physics* **2008**, *129* (12), 124105. https://doi.org/10.1063/1.2978177.

(362) Yue, Z.; Li, C.; Voth, G. A.; Swanson, J. M. J. Dynamic Protonation Dramatically Affects the Membrane Permeability of Drug-like Molecules. *J. Am. Chem. Soc.* **2019**, *141* (34), 13421–13433. https://doi.org/10.1021/jacs.9b04387.

(363) Paloncýová, M.; Fabre, G.; DeVane, R. H.; Trouillas, P.; Berka, K.; Otyepka, M. Benchmarking of Force Fields for Molecule–Membrane Interactions. *J. Chem. Theory Comput.* **2014**, *10* (9), 4143–4151. https://doi.org/10.1021/ct500419b.

(364) Košinová, P.; Berka, K.; Wykes, M.; Otyepka, M.; Trouillas, P. Positioning of Antioxidant Quercetin and Its Metabolites in Lipid Bilayer Membranes: Implication for Their Lipid-Peroxidation Inhibition. *J. Phys. Chem. B* **2012**, *116* (4), 1309–1318. https://doi.org/10.1021/jp208731g.

(365) Hossain, S. I.; Saha, S. C.; Deplazes, E. Phenolic Compounds Alter the Ion Permeability of Phospholipid Bilayers *via* Specific Lipid Interactions. *Phys. Chem. Chem. Phys.* **2021**, *23* (39), 22352–22366. https://doi.org/10.1039/D1CP03250J.

(366) Neale, C.; Pomès, R. Sampling Errors in Free Energy Simulations of Small Molecules in Lipid Bilayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2016**, *1858* (10), 2539–2548. https://doi.org/10.1016/j.bbamem.2016.03.006.

(367) Fabre, G.; Bayach, I.; Berka, K.; Paloncýová, M.; Starok, M.; Rossi, C.; Duroux, J.-L.; Otyepka, M.; Trouillas, P. Synergism of Antioxidant Action of Vitamins E, C and Quercetin Is Related to Formation of Molecular Associations in Biomembranes. *Chem. Commun.* **2015**, *51* (36), 7713–7716. https://doi.org/10.1039/C5CC00636H.

(368) Kopeć, W.; Telenius, J.; Khandelia, H. Molecular Dynamics Simulations of the Interactions of Medicinal Plant Extracts and Drugs with Lipid Bilayer Membranes. *FEBS J* **2013**, *280* (12), 2785–2805. https://doi.org/10.1111/febs.12286.

(369) Lyu, Y.; Xiang, N.; Mondal, J.; Zhu, X.; Narsimhan, G. Characterization of Interactions between Curcumin and Different Types of Lipid Bilayers by Molecular Dynamics Simulation. *J. Phys. Chem. B* **2018**, *122* (8), 2341–2354. https://doi.org/10.1021/acs.jpcb.7b10566.

(370) Fabre, G. Molecular Interaction of Natural Compounds with Lipid Bilayer Membranes: Towards a Better Understanding of Their Biological and Pharmaceutical Actions. **2015**, 162.

(371) Naumowicz, M.; Kusaczuk, M.; Zając, M.; Jabłońska-Trypuć, A.; Mikłosz, A.; Gál, M.; Worobiczuk, M.; Kotyńska, J. The Influence of the PH on the Incorporation of Caffeic Acid into Biomimetic Membranes and Cancer Cells. *Sci Rep* **2022**, *12* (1), 3692. https://doi.org/10.1038/s41598-022-07700-8.

(372) Ly, H. V.; Longo, M. L. The Influence of Short-Chain Alcohols on Interfacial Tension, Mechanical Properties, Area/Molecule, and Permeability of Fluid Lipid Bilayers.

Biophysical Journal **2004**, *87* (2), 1013–1033. https://doi.org/10.1529/biophysj.103.034280.

(373) *Exploring QSAR*; Hansch, C., Leo, A., Hoekman, D. H., Eds.; ACS professional reference book; American Chemical Society: Washington, DC, 1995.

(374) Moldoveanu, S.; David, V. Solvent Extraction. In *Modern Sample Preparation for Chromatography*; Elsevier, 2015; pp 131–189. https://doi.org/10.1016/B978-0-444-54319-6.00006-2.

(375) Rothwell, J. A.; Day, A. J.; Morgan, M. R. A. Experimental Determination of Octanol−Water Partition Coefficients of Quercetin and Related Flavonoids. *J. Agric. Food Chem.* **2005**, *53* (11), 4355–4360. https://doi.org/10.1021/jf0483669.

(376) Herbig, M. E.; Evers, D.-H. Correlation of Hydrotropic Solubilization by Urea with LogD of Drug Molecules and Utilization of This Effect for Topical Formulations. *European Journal of Pharmaceutics and Biopharmaceutics* **2013**, *85* (1), 158–160. https://doi.org/10.1016/j.ejpb.2013.06.022.

(377) Privat, C.; Telo, J. P.; Bernardes-Genisson, V.; Vieira, A.; Souchard, J.-P.; Nepveu, F. Antioxidant Properties of *Trans* -ε-Viniferin As Compared to Stilbene Derivatives in Aqueous and Nonaqueous Media. *J. Agric. Food Chem.* **2002**, *50* (5), 1213–1217. https://doi.org/10.1021/jf010676t.

(378) Priyadarsini, K. The Chemistry of Curcumin: From Extraction to Therapeutic Agent. *Molecules* **2014**, *19* (12), 20091–20112. https://doi.org/10.3390/molecules191220091.

(379) Sangster, J.; Box, P. O. A Databank of Evaluated Octanol-Water Partition Coefficients. 3.

(380) Chen, H.; Fu, H.; Chipot, C.; Shao, X.; Cai, W. Overcoming Free-Energy Barriers with a Seamless Combination of a Biasing Force and a Collective Variable-Independent Boost Potential. *Journal of Chemical Theory and Computation* **2021**. https://doi.org/10.1021/acs.jctc.1c00103.

(381) Ghaemi, Z.; Minozzi, M.; Carloni, P.; Laio, A. A Novel Approach to the Investigation of Passive Molecular Permeation through Lipid Bilayers from Atomistic Simulations. *J. Phys. Chem. B* **2012**, *116* (29), 8714–8721. https://doi.org/10.1021/jp301083h.

(382) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960. https://doi.org/10.1021/jp003020w.

(383) Avdeef, A. Permeability—PAMPA. In *Absorption and Drug Development*; John Wiley & Sons, Ltd, 2012; pp 319–498. https://doi.org/10.1002/9781118286067.ch7.

(384) Avdeef, A. Permeability: Blood–Brain Barrier. In *Absorption and Drug Development*; John Wiley & Sons, Ltd, 2012; pp 575–680. https://doi.org/10.1002/9781118286067.ch9.

(385) Tse, C. H.; Comer, J.; Wang, Y.; Chipot, C. Link between Membrane Composition and Permeability to Drugs. *J. Chem. Theory Comput.* **2018**, *14* (6), 2895–2909. https://doi.org/10.1021/acs.jctc.8b00272.

(386) Ribeiro, R. P.; Coimbra, J. T. S.; Ramos, M. J.; Fernandes, P. A. Diffusion of the Small, Very Polar, Drug Piracetam through a Lipid Bilayer: An MD Simulation Study. *Theor Chem Acc* **2017**, *136* (4), 46. https://doi.org/10.1007/s00214-017-2073-3.

(387) Nitschke, N.; Atkovska, K.; Hub, J. S. Accelerating Potential of Mean Force Calculations for Lipid Membrane Permeation: System Size, Reaction Coordinate, Solute-Solute Distance, and Cutoffs. *J. Chem. Phys.* **2016**, *145* (12), 125101. https://doi.org/10.1063/1.4963192.

(388) Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. An Analysis of the Attrition of Drug Candidates from Four Major Pharmaceutical Companies. *Nat Rev Drug Discov* **2015**, *14* (7), 475–486. https://doi.org/10.1038/nrd4609.

(389) Lipnick, R. L. Narcosis, Electrophile and Proelectrophile Toxicity Mechanisms: Application of SAR and QSAR. *Environ Toxicol Chem* **1989**, *8* (1), 1–12. https://doi.org/10.1002/etc.5620080101.

(390) Hansch, C. The Advent and Evolution of QSAR at Pomona College. *J Comput Aided Mol Des* **2011**, *25* (6), 495–507. https://doi.org/10.1007/s10822-011-9444-y.

(391) Sangster, J. *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*; Wiley series in solution chemistry; Wiley: Chichester, 1997.

(392) Guziałowska-Tic, J. The Use of QSAR Methods for Determination of N-Octanol/Water Partition Coefficient Using the Example of Hydroxyester HE-1. *E3S Web Conf.* **2017**, *19*, 02034. https://doi.org/10.1051/e3sconf/20171902034.

(393) Leahy, D. E.; Taylor, P. J.; Wait, A. R. Model Solvent Systems for QSAR Part I. Propylene Glycol Dipelargonate (PGDP). A New Standard Solvent for Use in Partition Coefficient Determination. *Quantitative Structure-Activity Relationships* **1989**, *8* (1), 17–31. https://doi.org/10.1002/qsar.19890080104.

(394) Van De Waterbeemd, H.; Van Bakel, P.; Jansen, A. Transport in Quantitative Structure-Activity Relationships VI: Relationship Between Transport Rate Constants and Partition Coefficients. *Journal of Pharmaceutical Sciences* **1981**, *70* (9), 1081–1082. https://doi.org/10.1002/jps.2600700930.

(395) Nitschke, N.; Atkovska, K.; Hub, J. S. Accelerating Potential of Mean Force Calculations for Lipid Membrane Permeation: System Size, Reaction Coordinate, Solute-Solute Distance, and Cutoffs. *The Journal of Chemical Physics* **2016**, *145* (12), 125101. https://doi.org/10.1063/1.4963192.

(396) Comer, J.; Schulten, K.; Chipot, C. Permeability of a Fluid Lipid Bilayer to Short-Chain Alcohols from First Principles. *J. Chem. Theory Comput.* **2017**, *13* (6), 2523–2532. https://doi.org/10.1021/acs.jctc.7b00264.

(397) Pokhrel, N.; Maibaum, L. Free Energy Calculations of Membrane Permeation: Challenges Due to Strong Headgroup-Solute Interactions. *J. Chem. Theory Comput.* **2018**, *14* (3), 1762–1771. https://doi.org/10.1021/acs.jctc.7b01159.

(398) Lohman, S. W. *Definitions of Selected Ground-Water Terms, Revisions and Conceptual Refinements*; 1988; U.S. Government Printing Office, 1972; pp 1–21. https://doi.org/10.3133/wsp1988.

(399) *PerMM: A Web Tool and Database for Analysis of Passive Membrane Permeability and Translocation Pathways of Bioactive Molecules - PubMed*. https://pubmed.ncbi.nlm.nih.gov/31259547/ (accessed 2022-09-18).

(400) Sugita, M.; Sugiyama, S.; Fujie, T.; Yoshikawa, Y.; Yanagisawa, K.; Ohue, M.; Akiyama, Y. Large-Scale Membrane Permeability Prediction of Cyclic Peptides Crossing a Lipid Bilayer Based on Enhanced Sampling Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2021**, *61* (7), 3681–3695. https://doi.org/10.1021/acs.jcim.1c00380.

(401) Linton, K. J.; Holland, I. B. *The ABC Transporters of Human Physiology and Disease: Genetics and Biochemistry of ATP Binding Cassette Transporters*; WORLD SCIENTIFIC, 2011. https://doi.org/10.1142/7371.

(402) Hyde, S. C.; Emsley, P.; Hartshorn, M. J.; Mimmack, M. M.; Gileadi, U.; Pearce, S. R.; Gallagher, M. P.; Gill, D. R.; Hubbard, R. E.; Higgins, C. F. Structural Model of ATP-

Binding Proteins Associated with Cystic Fibrosis, Multidrug Resistance and Bacterial Transport. *Nature* **1990**, *346* (6282), 362–365. https://doi.org/10.1038/346362a0.

(403) Rees, D. C.; Johnson, E.; Lewinson, O. ABC Transporters: The Power to Change. *Nat Rev Mol Cell Biol* **2009**, *10* (3), 218–227. https://doi.org/10.1038/nrm2646.

(404) Saurin, W.; Hofnung, M.; Dassa, E. Getting in or out: Early Segregation between Importers and Exporters in the Evolution of ATP-Binding Cassette (ABC) Transporters. *J Mol Evol* **1999**, *48* (1), 22–41. https://doi.org/10.1007/pl00006442.

(405) Hediger, M. A.; Clémençon, B.; Burrier, R. E.; Bruford, E. A. The ABCs of Membrane Transporters in Health and Disease (SLC Series): Introduction. *Mol Aspects Med* **2013**, *34* (2–3), 95–107. https://doi.org/10.1016/j.mam.2012.12.009.

(406) Vasiliou, V.; Vasiliou, K.; Nebert, D. W. Human ATP-Binding Cassette (ABC) Transporter Family. *Hum Genomics* **2008**, *3* (3), 281. https://doi.org/10.1186/1479-7364-3-3-281.

(407) Mourez, M.; Hofnung, M.; Dassa, E. Subunit Interactions in ABC Transporters: A Conserved Sequence in Hydrophobic Membrane Proteins of Periplasmic Permeases Defines an Important Site of Interaction with the ATPase Subunits. *EMBO J* **1997**, *16* (11), 3066–3077. https://doi.org/10.1093/emboj/16.11.3066.

(408) Linton, K. J.; Zolnerciks, J. K.; Schmitt, L. General Introduction, Structure and Likely Mechanism of Action of ABC Transport Proteins. In *The ABC Transporters of Human Physiology and Disease*; WORLD SCIENTIFIC, 2011; pp 1–27. https://doi.org/10.1142/9789814280075_0001.

(409) Davidson, A. L.; Dassa, E.; Orelle, C.; Chen, J. Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol Mol Biol Rev* **2008**, *72* (2), 317–364. https://doi.org/10.1128/MMBR.00031-07.

(410) Schneider, E.; Hunke, S. ATP-Binding-Cassette (ABC) Transport Systems: Functional and Structural Aspects of the ATP-Hydrolyzing Subunits/Domains. *FEMS Microbiol Rev* **1998**, *22* (1), 1–20. https://doi.org/10.1111/j.1574-6976.1998.tb00358.x.

(411) Higgins, C. F. ABC Transporters: Physiology, Structure and Mechanism--an Overview. *Res Microbiol* **2001**, *152* (3–4), 205–210. https://doi.org/10.1016/s0923-2508(01)01193-7.

(412) Thomas, C.; Tampé, R. Structural and Mechanistic Principles of ABC Transporters. *Annu Rev Biochem* **2020**, *89*, 605–636. https://doi.org/10.1146/annurev-biochem-011520-105201.

(413) Kroll, T.; Prescher, M.; Smits, S. H. J.; Schmitt, L. Structure and Function of Hepatobiliary ATP Binding Cassette Transporters. *Chem Rev* **2020**. https://doi.org/10.1021/acs.chemrev.0c00659.

(414) Scortecci, J. F.; Molday, L. L.; Curtis, S. B.; Garces, F. A.; Panwar, P.; Van Petegem, F.; Molday, R. S. Cryo-EM Structures of the ABCA4 Importer Reveal Mechanisms Underlying Substrate Binding and Stargardt Disease. *Nat Commun* **2021**, *12* (1), 5902. https://doi.org/10.1038/s41467-021-26161-7.

(415) Tóth, Á.; Janaszkiewicz, A.; Crespi, V.; Di Meo, F. *On the Interplay between Lipids and Asymmetric Dynamics of an NBS Degenerate ABC Transporter*; preprint; Biophysics, 2022. https://doi.org/10.1101/2022.05.16.492073.

(416) El-Awady, R.; Saleh, E.; Hashim, A.; Soliman, N.; Dallah, A.; Elrasheed, A.; Elakraa, G. The Role of Eukaryotic and Prokaryotic ABC Transporter Family in Failure of Chemotherapy. *Frontiers in Pharmacology* **2017**, *7*.

(417) Riordan, J. R.; Rommens, J. M.; Kerem, B.; Alon, N.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.; Plavsic, N.; Chou, J. L. Identification of the Cystic Fibrosis Gene:

Cloning and Characterization of Complementary DNA. *Science* **1989**, *245* (4922), 1066–1073. https://doi.org/10.1126/science.2475911.

(418) Remaley, A. T.; Rust, S.; Rosier, M.; Knapper, C.; Naudin, L.; Broccardo, C.; Peterson, K. M.; Koch, C.; Arnould, I.; Prades, C.; Duverger, N.; Funke, H.; Assman, G.; Dinger, M.; Dean, M.; Chimini, G.; Santamarina-Fojo, S.; Fredrickson, D. S.; Denefle, P.; Brewer, H. B. Human ATP-Binding Cassette Transporter 1 (ABC1): Genomic Organization and Identification of the Genetic Defect in the Original Tangier Disease Kindred. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (22), 12685–12690. https://doi.org/10.1073/pnas.96.22.12685.

(419) Strautnieks, S. S.; Bull, L. N.; Knisely, A. S.; Kocoshis, S. A.; Dahl, N.; Arnell, H.; Sokal, E.; Dahan, K.; Childs, S.; Ling, V.; Tanner, M. S.; Kagalwalla, A. F.; Németh, A.; Pawlowska, J.; Baker, A.; Mieli-Vergani, G.; Freimer, N. B.; Gardiner, R. M.; Thompson, R. J. A Gene Encoding a Liver-Specific ABC Transporter Is Mutated in Progressive Familial Intrahepatic Cholestasis. *Nat Genet* **1998**, *20* (3), 233–238. https://doi.org/10.1038/3034.

(420) Albrecht, C.; McVey, J. H.; Elliott, J. I.; Sardini, A.; Kasza, I.; Mumford, A. D.; Naoumova, R. P.; Tuddenham, E. G. D.; Szabo, K.; Higgins, C. F. A Novel Missense Mutation in ABCA1 Results in Altered Protein Trafficking and Reduced Phosphatidylserine Translocation in a Patient with Scott Syndrome. *Blood* **2005**, *106* (2), 542–549. https://doi.org/10.1182/blood-2004-05-2056.

(421) Sodani, K.; Patel, A.; Kathawala, R. J.; Chen, Z.-S. Multidrug Resistance Associated Proteins in Multidrug Resistance. *Chin J Cancer* **2012**, *31* (2), 58–72. https://doi.org/10.5732/cjc.011.10329.

(422) Roepe, P. D. What Is the Precise Role of Human MDR 1 Protein in Chemotherapeutic Drug Resistance. *Current Pharmaceutical Design 6* (3), 241–260.

(423) Schinkel, A. H.; Mayer, U.; Wagenaar, E.; Mol, C. A.; van Deemter, L.; Smit, J. J.; van der Valk, M. A.; Voordouw, A. C.; Spits, H.; van Tellingen, O.; Zijlmans, J. M.; Fibbe, W. E.; Borst, P. Normal Viability and Altered Pharmacokinetics in Mice Lacking Mdr1-Type (Drug-Transporting) P-Glycoproteins. *Proc Natl Acad Sci U S A* **1997**, *94* (8), 4028–4033. https://doi.org/10.1073/pnas.94.8.4028.

(424) Cole, S.; Bhardwaj, G.; Gerlach, J.; Mackie, J.; Grant, C.; Almquist, K.; Stewart, A.; Kurz, E.; Duncan, A.; Deeley, R. Overexpression of a Transporter Gene in a Multidrug-Resistant Human Lung Cancer Cell Line. *Science* **1992**, *258* (5088), 1650–1654. https://doi.org/10.1126/science.1360704.

(425) Zhang, Y.-K.; Wang, Y.-J.; Gupta, P.; Chen, Z.-S. Multidrug Resistance Proteins (MRPs) and Cancer Therapy. *AAPS J* **2015**, *17* (4), 802–812. https://doi.org/10.1208/s12248-015-9757-1.

(426) Schinkel, A. H.; Jonker, J. W. Mammalian Drug Efflux Transporters of the ATP Binding Cassette (ABC) Family: An Overview. *Adv Drug Deliv Rev* **2003**, *55* (1), 3–29. https://doi.org/10.1016/s0169-409x(02)00169-2.

(427) Slot, A. J.; Molinski, S. V.; Cole, S. P. C. Mammalian Multidrug-Resistance Proteins (MRPs). *Essays Biochem* **2011**, *50* (1), 179–207. https://doi.org/10.1042/bse0500179.

(428) van Aubel, R. A. M. H.; Smeets, P. H. E.; Peters, J. G. P.; Bindels, R. J. M.; Russel, F. G. M. The MRP4/ABCC4 Gene Encodes a Novel Apical Organic Anion Transporter in Human Kidney Proximal Tubules: Putative Efflux Pump for Urinary CAMP and CGMP. *J Am Soc Nephrol* **2002**, *13* (3), 595–603. https://doi.org/10.1681/ASN.V133595.

(429) Berthier, J.; Arnion, H.; Saint-Marcoux, F.; Picard, N. Multidrug Resistance-Associated Protein 4 in Pharmacology: Overview of Its Contribution to Pharmacokinetics,

Pharmacodynamics and Pharmacogenetics. *Life Sciences* **2019**, *231*, 116540. https://doi.org/10.1016/j.lfs.2019.06.015.

(430) Sampath, J.; Adachi, M.; Hatse, S.; Naesens, L.; Balzarini, J.; Flatley, R.; Matherly, L.; Schuetz, J. Role of MRP4 and MRP5 in Biology and Chemotherapy. *AAPS J* **2002**, *4* (3), 22–30. https://doi.org/10.1208/ps040314.

(431) Ci, L.; Kusuhara, H.; Adachi, M.; Schuetz, J. D.; Takeuchi, K.; Sugiyama, Y. Involvement of MRP4 (*ABCC4*) in the Luminal Efflux of Ceftizoxime and Cefazolin in the Kidney. *Mol Pharmacol* **2007**, *71* (6), 1591–1597. https://doi.org/10.1124/mol.106.031823.

(432) Hasegawa, M.; Kusuhara, H.; Adachi, M.; Schuetz, J. D.; Takeuchi, K.; Sugiyama, Y. Multidrug Resistance–Associated Protein 4 Is Involved in the Urinary Excretion of Hydrochlorothiazide and Furosemide. *JASN* **2007**, *18* (1), 37–45. https://doi.org/10.1681/ASN.2005090966.

(433) Imaoka, T.; Kusuhara, H.; Adachi, M.; Schuetz, J. D.; Takeuchi, K.; Sugiyama, Y. Functional Involvement of Multidrug Resistance-Associated Protein 4 (MRP4/ABCC4) in the Renal Elimination of the Antiviral Drugs Adefovir and Tenofovir. *Mol Pharmacol* **2007**, *71* (2), 619–627. https://doi.org/10.1124/mol.106.028233.

(434) Berthier, J.; Benmameri, M.; Sauvage, F.-L.; Fabre, G.; Chantemargue, B.; Arnion, H.; Marquet, P.; Trouillas, P.; Picard, N.; Saint-Marcoux, F. MRP4 Is Responsible for the Efflux Transport of Mycophenolic Acid β-d Glucuronide (MPAG) from Hepatocytes to Blood. *Xenobiotica* **2021**, *51* (1), 105–114. https://doi.org/10.1080/00498254.2020.1813352.

(435) Russel, F.; Koenderink, J.; Masereeuw, R. Multidrug Resistance Protein 4 (MRP4/ABCC4): A Versatile Efflux Transporter for Drugs and Signalling Molecules. *Trends in Pharmacological Sciences* **2008**, *29* (4), 200–207. https://doi.org/10.1016/j.tips.2008.01.006.

(436) Ritter, C. A.; Jedlitschky, G.; Meyer zu Schwabedissen, H.; Grube, M.; Köck, K.; Kroemer, H. K. Cellular Export of Drugs and Signaling Molecules by the ATP-Binding Cassette Transporters MRP4 (ABCC4) and MRP5 (ABCC5). *Drug Metab Rev* **2005**, *37* (1), 253–278. https://doi.org/10.1081/dmr-200047984.

(437) Stockner, T.; Gradisch, R.; Schmitt, L. The Role of the Degenerate Nucleotide Binding Site in Type I ABC Exporters. *FEBS Letters* **2020**, *594* (23), 3815–3838. https://doi.org/10.1002/1873-3468.13997.

(438) El-Sheikh, A. A. K.; van den Heuvel, J. J. M. W.; Koenderink, J. B.; Russel, F. G. M. Interaction of Nonsteroidal Anti-Inflammatory Drugs with Multidrug Resistance Protein (MRP) 2/ABCC2- and MRP4/ABCC4-Mediated Methotrexate Transport. *J Pharmacol Exp Ther* **2007**, *320* (1), 229–235. https://doi.org/10.1124/jpet.106.110379.

(439) El-Sheikh, A. A. K.; Greupink, R.; Wortelboer, H. M.; van den Heuvel, J. J. M. W.; Schreurs, M.; Koenderink, J. B.; Masereeuw, R.; Russel, F. G. M. Interaction of Immunosuppressive Drugs with Human Organic Anion Transporter (OAT) 1 and OAT3, and Multidrug Resistance-Associated Protein (MRP) 2 and MRP4. *Translational Research* **2013**, *162* (6), 398–409. https://doi.org/10.1016/j.trsl.2013.08.003.

(440) Calcagno, A.; Kim, I.-W.; Wu, C.-P.; Shukla, S.; Ambudkar, S. ABC Drug Transporters as Molecular Targets for the Prevention of Multidrug Resistance and Drug-Drug Interactions. *CDD* **2007**, *4* (4), 324–333. https://doi.org/10.2174/156720107782151241.

(441) Glavinas, H.; Krajcsi, P.; Cserepes, J.; Sarkadi, B. The Role of ABC Transporters in Drug Resistance, Metabolism and Toxicity. *CDD* **2004**, *1* (1), 27–42. https://doi.org/10.2174/1567201043480036.

(442) Yin, J.; Wang, J. Renal Drug Transporters and Their Significance in Drug–Drug Interactions. *Acta Pharmaceutica Sinica B* **2016**, *6* (5), 363–373. https://doi.org/10.1016/j.apsb.2016.07.013.

(443) El-Sheikh, A. A. K.; van den Heuvel, J. J. M. W.; Krieger, E.; Russel, F. G. M.; Koenderink, J. B. Functional Role of Arginine 375 in Transmembrane Helix 6 of Multidrug Resistance Protein 4 (MRP4/ABCC4). *Mol Pharmacol* **2008**, *74* (4), 964–971. https://doi.org/10.1124/mol.107.043661.

(444) Wittgen, H. G. M.; van den Heuvel, J. J. M. W.; Krieger, E.; Schaftenaar, G.; Russel, F. G. M.; Koenderink, J. B. Phenylalanine 368 of Multidrug Resistance-Associated Protein 4 (MRP4/ABCC4) Plays a Crucial Role in Substrate-Specific Transport Activity. *Biochemical Pharmacology* **2012**, *84* (3), 366–373. https://doi.org/10.1016/j.bcp.2012.04.012.

(445) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Research* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.

(446) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(447) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J Appl Cryst* **1993**, *26* (2), 283–291. https://doi.org/10.1107/S0021889892009944.

(448) Pontius, J.; Richelle, J.; Wodak, S. J. Deviations from Standard Atomic Volumes as a Quality Measure for Protein Crystal Structures. *J Mol Biol* **1996**, *264* (1), 121–136. https://doi.org/10.1006/jmbi.1996.0628.

(449) *SAVESv6.0 - Structure Validation Server*. https://saves.mbi.ucla.edu/ (accessed 2022-07-27).

(450) Colovos, C.; Yeates, T. O. Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. *Protein Sci* **1993**, *2* (9), 1511–1519. https://doi.org/10.1002/pro.5560020916.

(451) *ProteinPrepare: a protein preparation app for MD simulations [WEB APP]*. https://www.playmolecule.com/proteinPrepare/ (accessed 2022-07-27).

(452) Martínez-Rosell, G.; Giorgino, T.; De Fabritiis, G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2017**, *57* (7), 1511–1516. https://doi.org/10.1021/acs.jcim.7b00190.

(453) Lomize, A. L.; Todd, S. C.; Pogozheva, I. D. Spatial Arrangement of Proteins in Planar and Curved Membranes by PPM 3.0. *Protein Sci* **2022**, *31* (1), 209–220. https://doi.org/10.1002/pro.4219.

(454) Wang, L.; Johnson, Z. L.; Wasserman, M. R.; Levring, J.; Chen, J.; Liu, S. Characterization of the Kinetic Cycle of an ABC Transporter by Single-Molecule and Cryo-EM Analyses. *eLife* **2020**, *9*, e56451. https://doi.org/10.7554/eLife.56451.

(455) D.A. Case; K. Belfon; I.Y. Ben-Shalom; S.R. Brozell; D.S. Cerutti; T.E. Cheatham; III; V.W.D. Cruzeiro; T.A. Darden; R.E. Duke; G. Giambasu; M.K. Gilson; H. Gohlke; A.W. Goetz,R Harris; S. Izadi; S.A. Iz- mailov; K. Kasavajhala; A. Kovalenko; R. Krasny; T. Kurtzman; T.S. Lee; S. LeGrand; P. Li; C. Lin; J. Liu; T. Luchko; R. Luo; V. Man; K.M. Merz; Y. Miao; O. Mikhailovskii; G. Monard; H. Nguyen; A. Onufriev; F. Pan; S. Pantano; R. Qi; D.R. Roe; A. Roitberg; C. Sagui; S. Schott-Verdugo; J. Shen; C.L. Simmerling; N.R. Skrynnikov; J. Smith; J. Swails; R.C. Walker; J. Wang; L. Wilson; R.M. Wolf; X. Wu; Y. Xiong; Y. Xue; D.M. York; P.A. Kollman. AMBER 2020.

(456) Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10* (2), 865–879. https://doi.org/10.1021/ct4010307.

(457) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(458) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(459) *AMBER parameter database (Bryce Group: Computational Biophysics and Drug Design - University of Manchester)*. http://amber.manchester.ac.uk/ (accessed 2022-07-27).

(460) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12* (8), 4114–4127. https://doi.org/10.1021/acs.jctc.6b00186.

(461) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of Polyphosphate Parameters for Use with the AMBER Force Field. *J Comput Chem* **2003**, *24* (9), 1016–1025. https://doi.org/10.1002/jcc.10262.

(462) 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* **2013**, *288* (26), 19211–19220. https://doi.org/10.1074/jbc.M113.450114.

(463) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.

(464) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates OfN-Acetylalanyl-N?-Methylamide. *Biopolymers* **1992**, *32* (5), 523–535. https://doi.org/10.1002/bip.360320508.

(465) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *The Journal of Chemical Physics* **1984**, *81* (8), 3684–3690. https://doi.org/10.1063/1.448118.

(466) Westerlund, A. M.; Delemotte, L. InfleCS: Clustering Free Energy Landscapes with Gaussian Mixtures. *J. Chem. Theory Comput.* **2019**, *15* (12), 6752–6759. https://doi.org/10.1021/acs.jctc.9b00454.

(467) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, NA-NA. https://doi.org/10.1002/jcc.21334.

(468) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res* **2012**, *40* (Database issue), D1100-1107. https://doi.org/10.1093/nar/gkr777.

(469) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: Towards Direct Deposition of Bioassay Data. *Nucleic Acids Res* **2019**, *47* (D1), D930–D940. https://doi.org/10.1093/nar/gky1075.

(470) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *Journal of Cheminformatics* **2011**, *3* (1), 33. https://doi.org/10.1186/1758-2946-3-33.

(471) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(472) Pravda, L.; Sehnal, D.; Svobodová Vařeková, R.; Navrátilová, V.; Toušek, D.; Berka, K.; Otyepka, M.; Koča, J. ChannelsDB: Database of Biomacromolecular Tunnels and Pores. *Nucleic Acids Research* **2018**, *46* (D1), D399–D405. https://doi.org/10.1093/nar/gkx868.

(473) Hofmann, S.; Januliene, D.; Mehdipour, A. R.; Thomas, C.; Stefan, E.; Brüchert, S.; Kuhn, B. T.; Geertsma, E. R.; Hummer, G.; Tampé, R.; Moeller, A. Conformation Space of a Heterodimeric ABC Exporter under Turnover Conditions. *Nature* **2019**, *571* (7766), 580–583. https://doi.org/10.1038/s41586-019-1391-0.

Appendix 1. Supplementary information for Chapter VI: "MemCross, a robust computational tool to evaluate permeability coefficients of drug-like molecules"

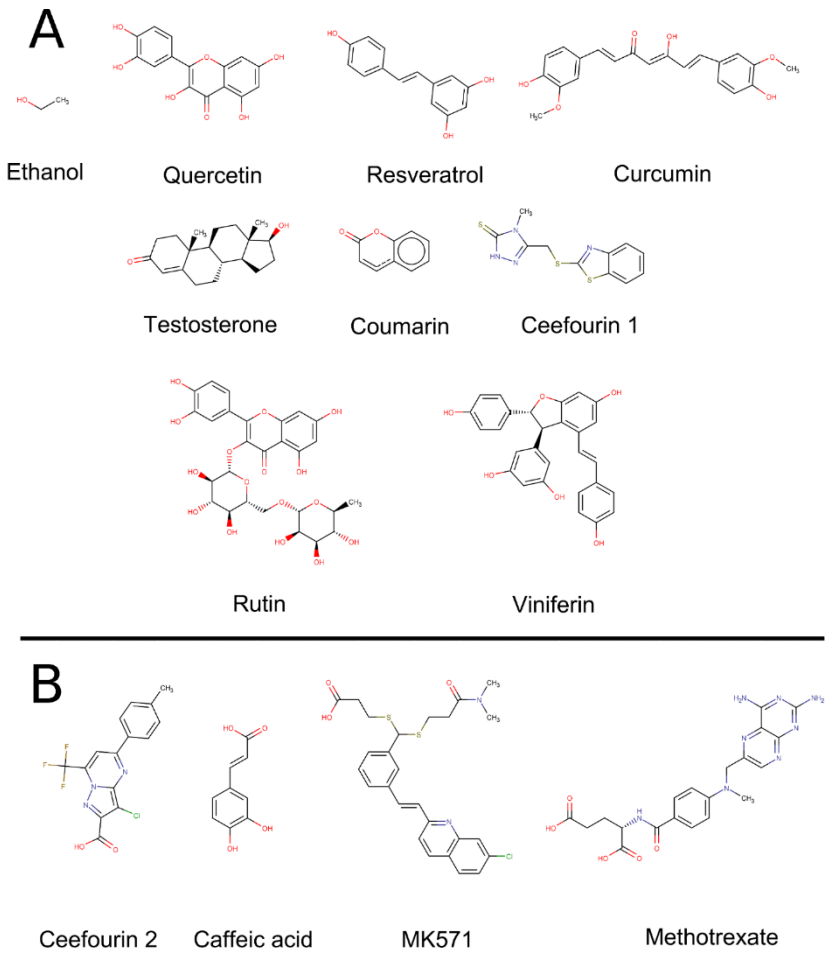


Figure S1: Chemical structures of the set of molecules used in the current study. (A): neutral and (B): acids.

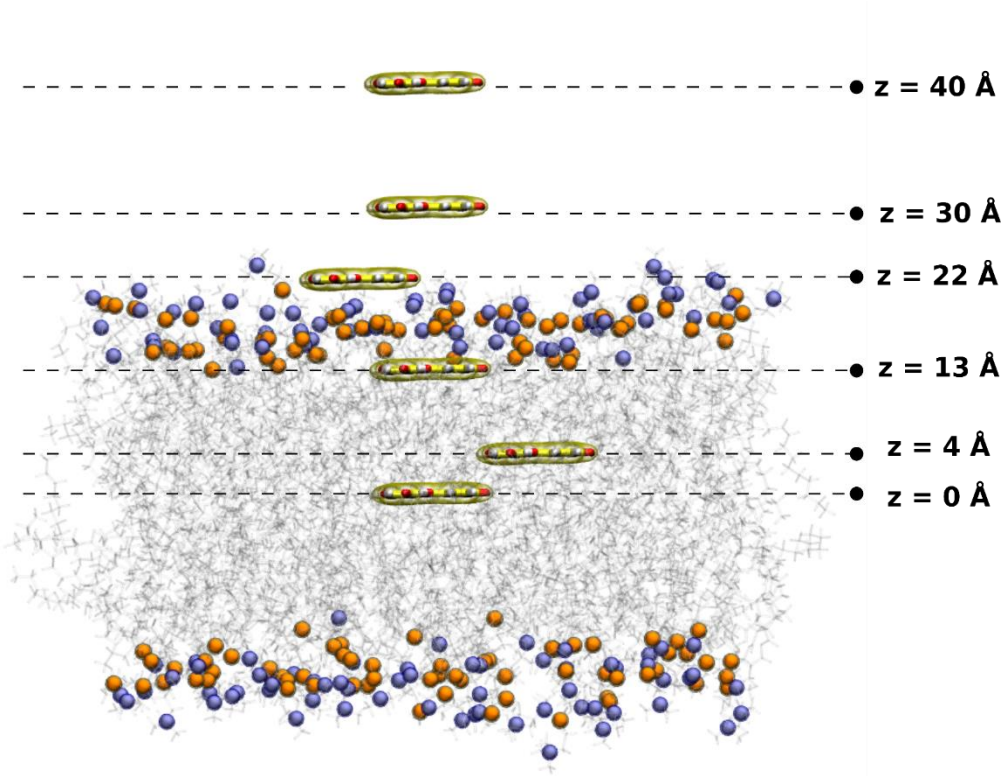


Figure S2: A representative snapshot showing the different depths of quercetin used in FEP calculations. z = 0 Å is the center of membrane and z = 40 Å is water.

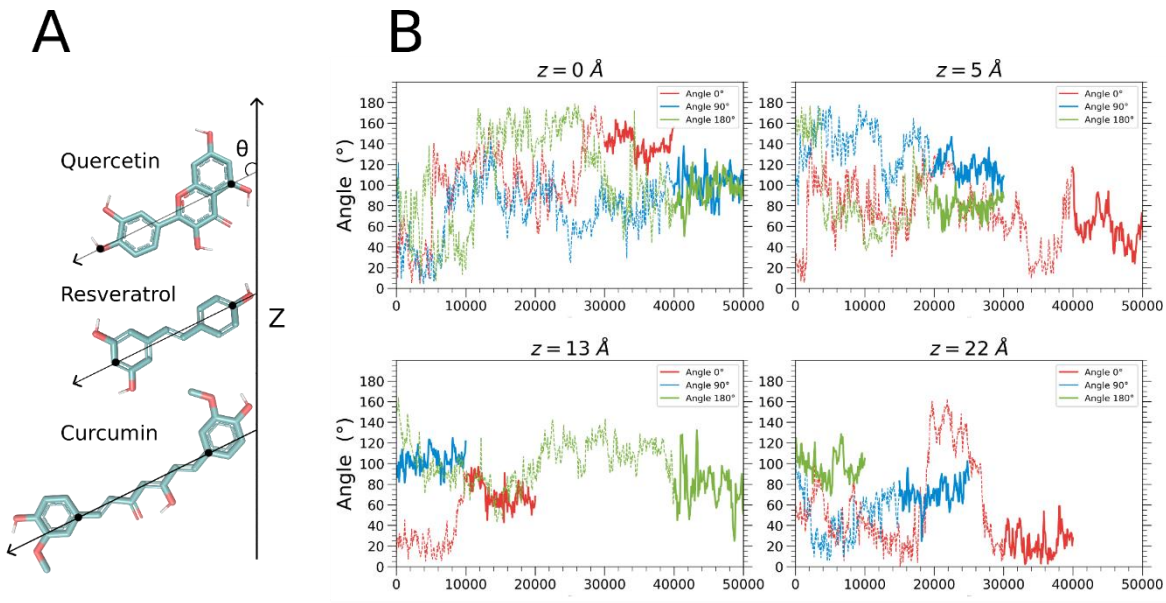


Figure S3: A) Molecular vectors used to calculated angle distributions and angle convergence assessment for quercetin, resveratrol and curcumin. B) Example of angle convergence of quercetin during equilibration runs before starting FEP calculations. The different panels represent different depths (windows) where the molecules were put with three different orientations over three runs.

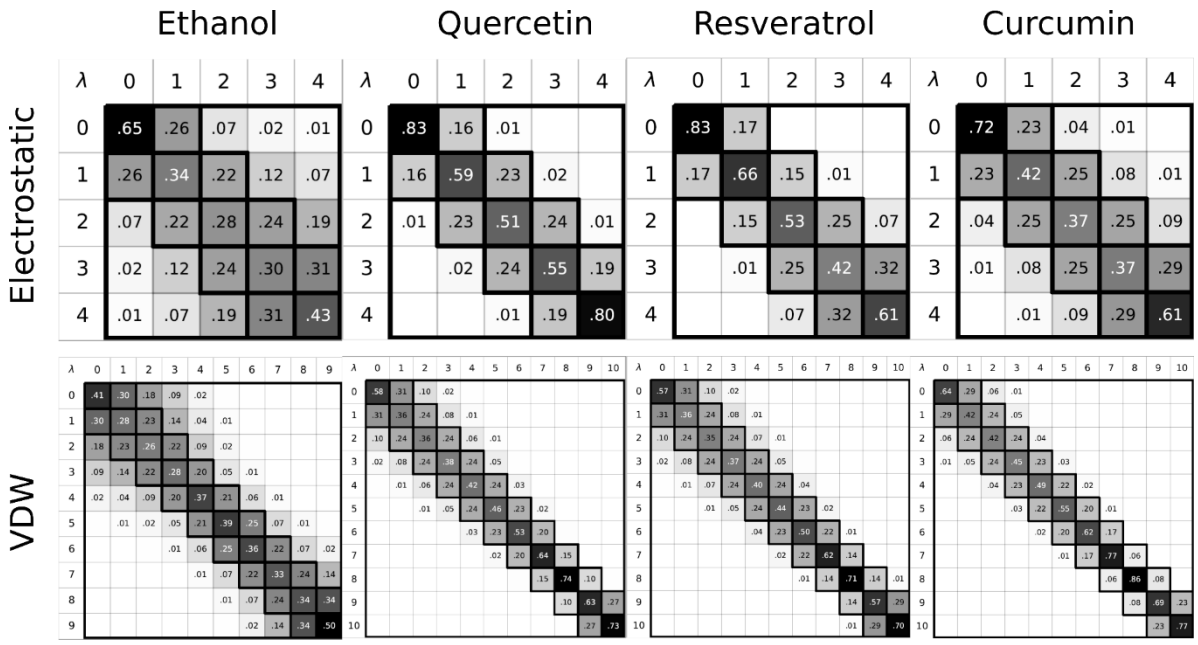


Figure S4: Overlap matrices plotted using MBAR from FEP calculations decoupling runs (electrostatic and VWD) for ethanol, quercetin, resveratrol and curcumin. They represent the overlap between the distributions of potential energy differences.

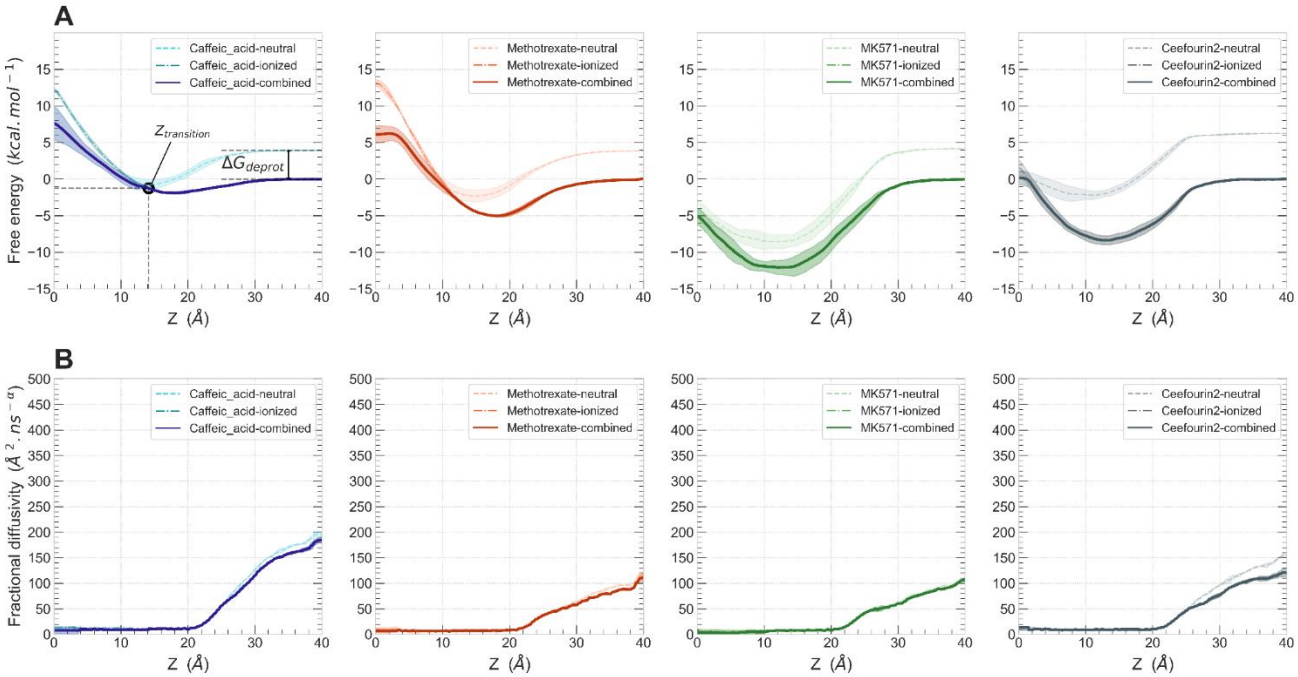


Figure S5: A) Computed PMF profiles, shifted by a the free energy of deprotonation in water (ΔG_{deprot}) and B) Fractional diffusivity profiles (neutral, ionized and combined) for caffeic acid, methotrexate, MK571 and ceefourin 2. Standard errors were calculated over three replicas.

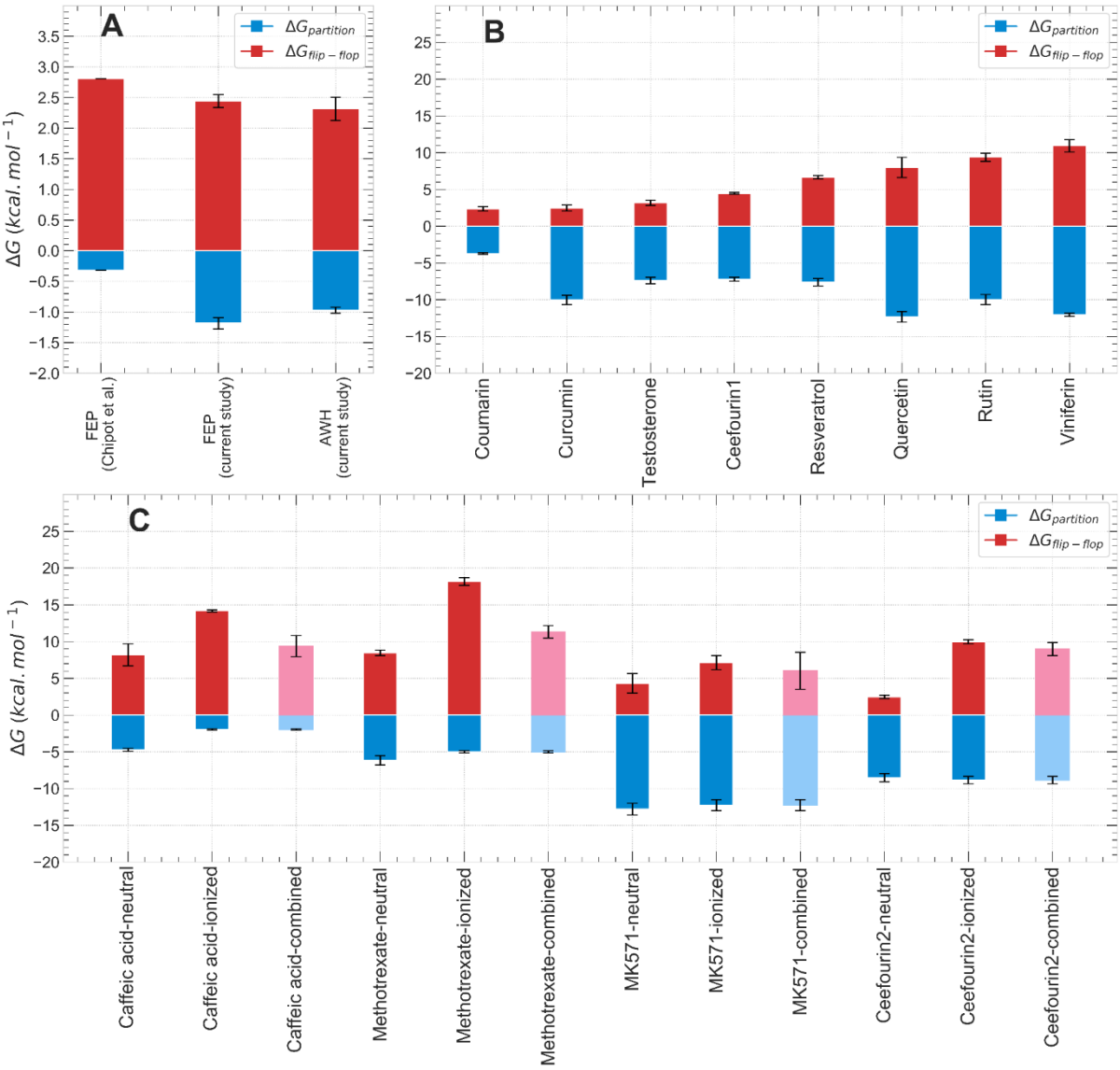
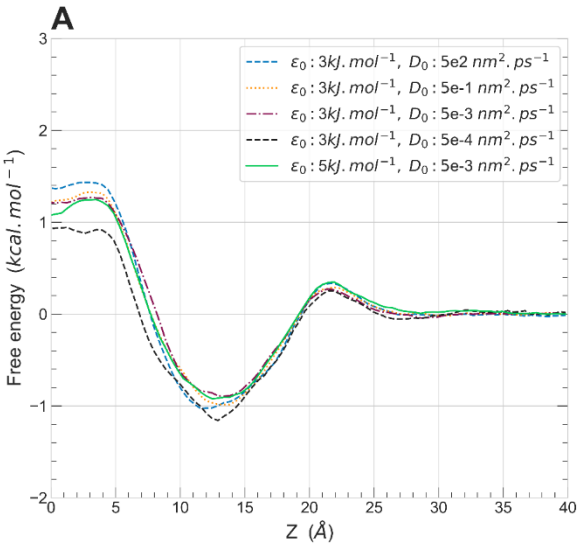


Figure S6: Comparison of free energy of partition (blue) and flip-flop (red). A) for ethanol between the FEP and AWH methods used in the current paper and the literature [70]. B) for neutral molecules. C) for ionizable molecules.



ε_0 (kJ. mol^{-1})	D_0 (nm^2. ps^{-1})	Initial stage time (ns)
3	5e2	40.7
3	5e-1	83.6
3	5e-3	67.36 ± 29.7
3	5e-4	>1000
5	5e-3	98.5

Figure S7: A) PMF profiles of ethanol using various values of initial diffusion constant (D_0) and error limit value (ε_0). B) Table showing initial stage times where the molecules has sampled all the reaction coordinate and after which the sampling is supposed to obey a asymptotic regime.

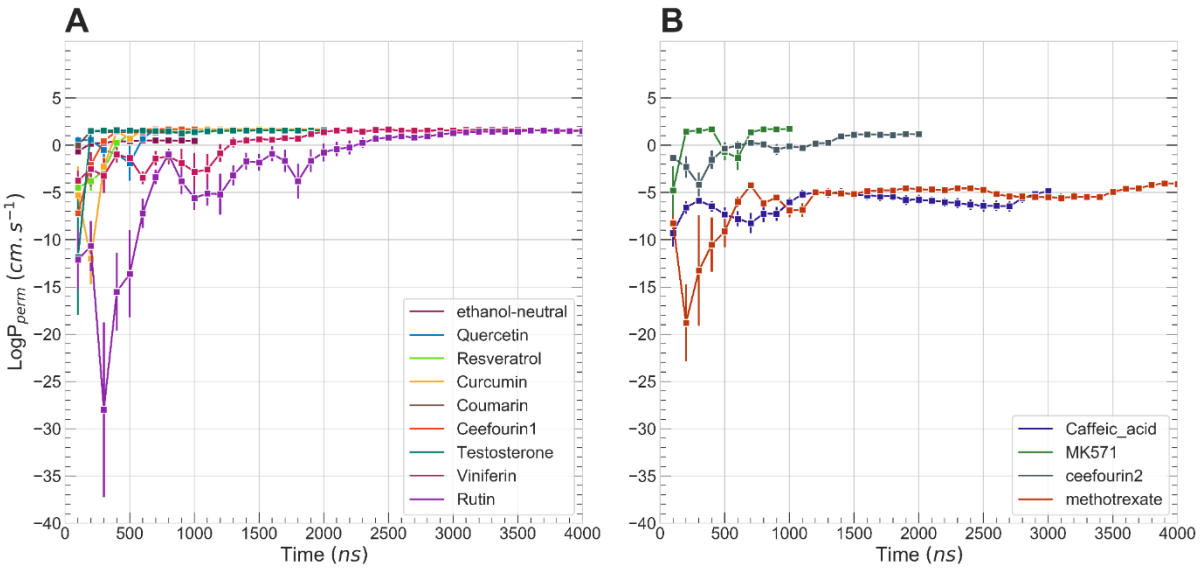


Figure S8: Convergence of computed permeation coefficient by MemCross (reported as $logP_{erm}^{calc}$) as a function simulation time with Δt = 100 ns.

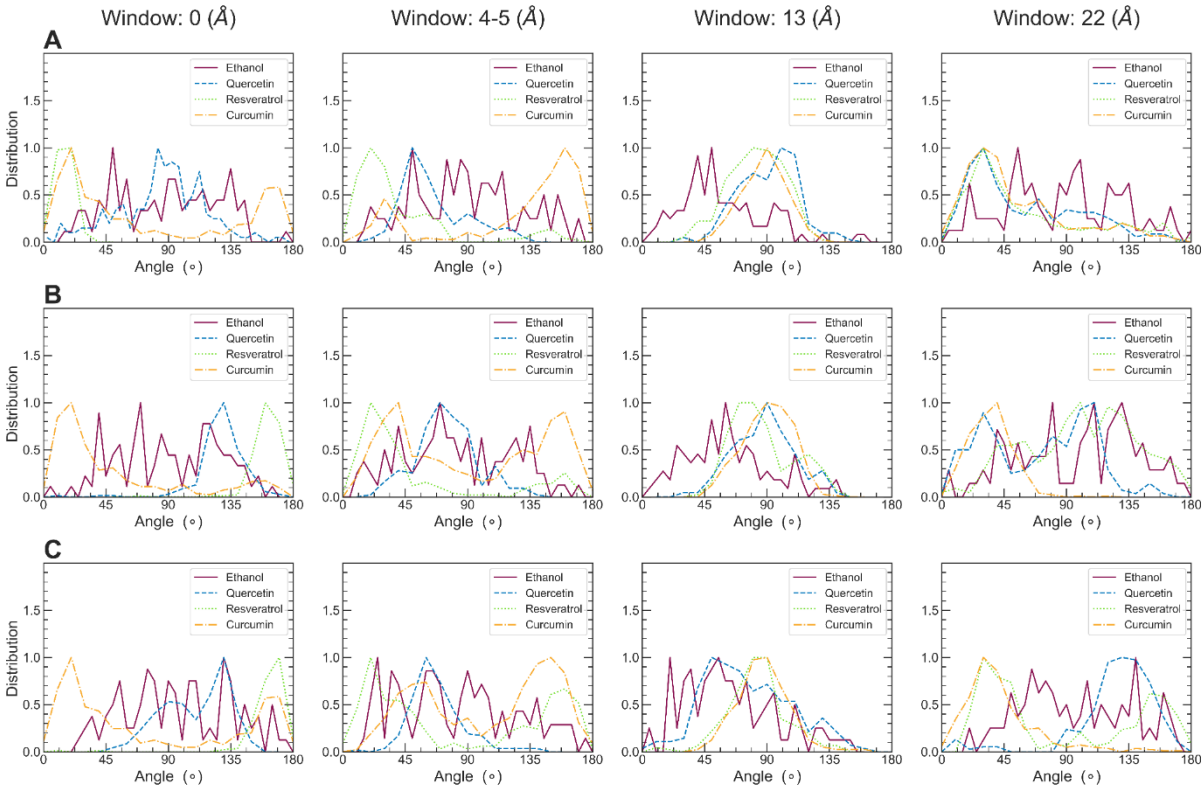


Figure S9: Normalized angle distributions of ethanol, quercetin, resveratrol and curcumin during FEP calculations at different windows. Only trajectories for the first λ-state were considered where the molecules maintain all their interactions with their environment.

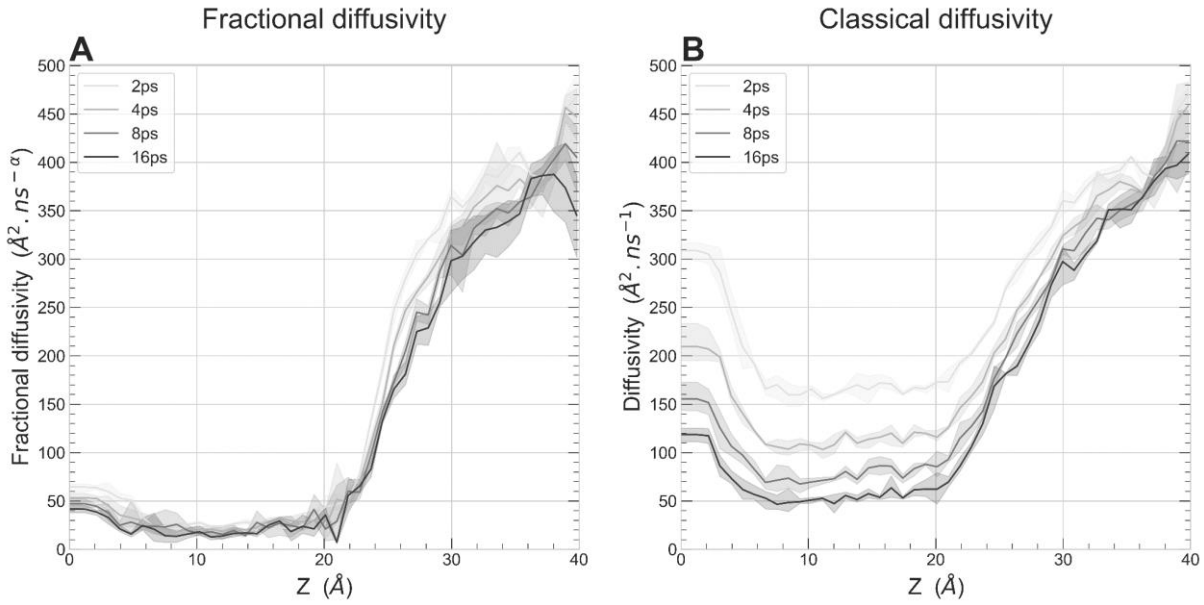


Figure S10: A) Fractional diffusivity (K_α) profiles of ethanol for different lag-times Δt compared to B) the diffusivity profiles calculated based on classical Smoluchowski equation.

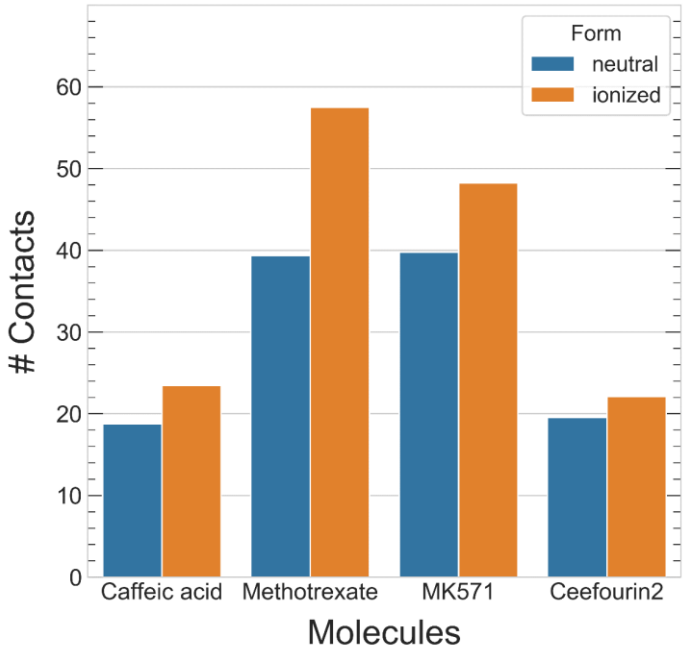


Figure S11: Number of contacts between water molecules and the permeant at Z-transition point (crossover point between neutral and ionized PMF profiles) at which the permeant neutralizes.

Table S1: Equilibration times required for angle convergence during equilibration runs before launching FEP calculations. * Window here reflects the depth where the molecules were placed, i.e., the distance between the membrane and the molecule center of masses.

Window (Å)*	Angle (°)	Equilibration time (ns)			
		Ethanol	Quercetin	Resveratrol	Curcumin
0	0	50	40	15	10
	90	50	50	15	15
	180	50	50	15	100
4	0	50	50	10	20
	90	50	30	10	20
	180	50	30	10	15
13	0	50	20	15	50
	90	50	10	10	35
	180	50	50	25	15
22	0	50	40	100	30
	90	50	25	50	25
	180	50	10	50	35

Appendix 2. Supplementary information for Chapter VI: "MemCross, a robust computational tool to evaluate permeability coefficients of drug-like molecules"

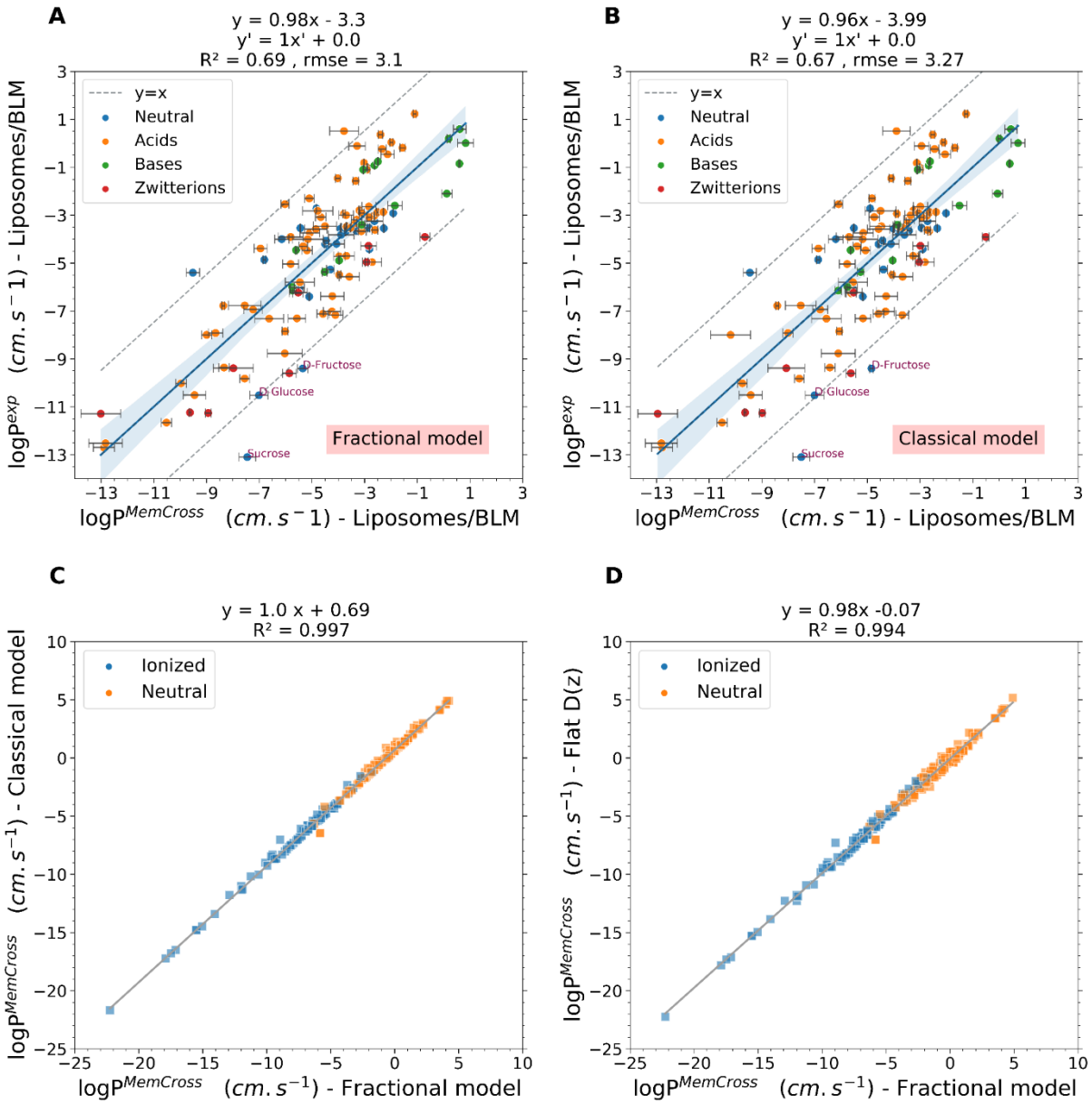


Figure S12: Fitted linear regression models upon correlations between $logP_{erm}^{MemCross}$ and $logP_{erm}^{exp}$ for the whole set of liposomes/BLM membrane measurements for (A) the fractional model for diffusivity calculations and (B) the fractional model for diffusivity calculations. We highlight all sugars (D-glucose, D-fructose and sucrose) as outliers as they all overpass the dotted lines with corresponds to twice the standard deviation of $logP_{erm}$ residuals where (residual = $logP_{erm}^{exp}$ - $logP_{erm}^{MemCross}$). (C) Correlation between $logP_{erm}^{MemCross}$ using fractional vs classical diffusivities. (D) Correlation between $logP_{erm}^{MemCross}$ using fractional vs flat (D = 10 Å².ns-1) diffusivities.

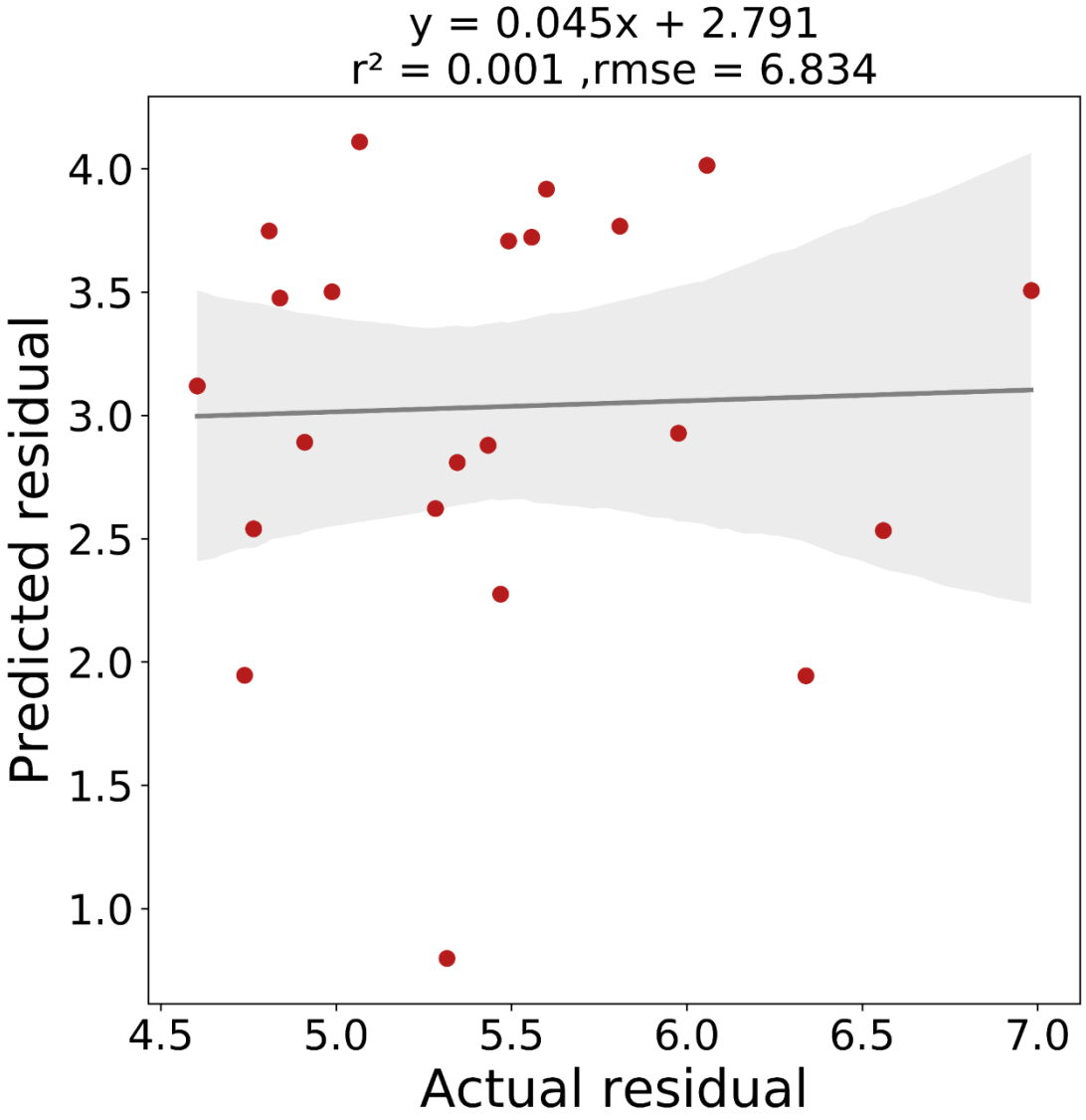


Figure S13: Predicted $logP_{erm}$ residual vs Actual $logP_{erm}$ residual calculated using a multivariate linear combination model based on different molecular descriptors and PMF-linked parameters.

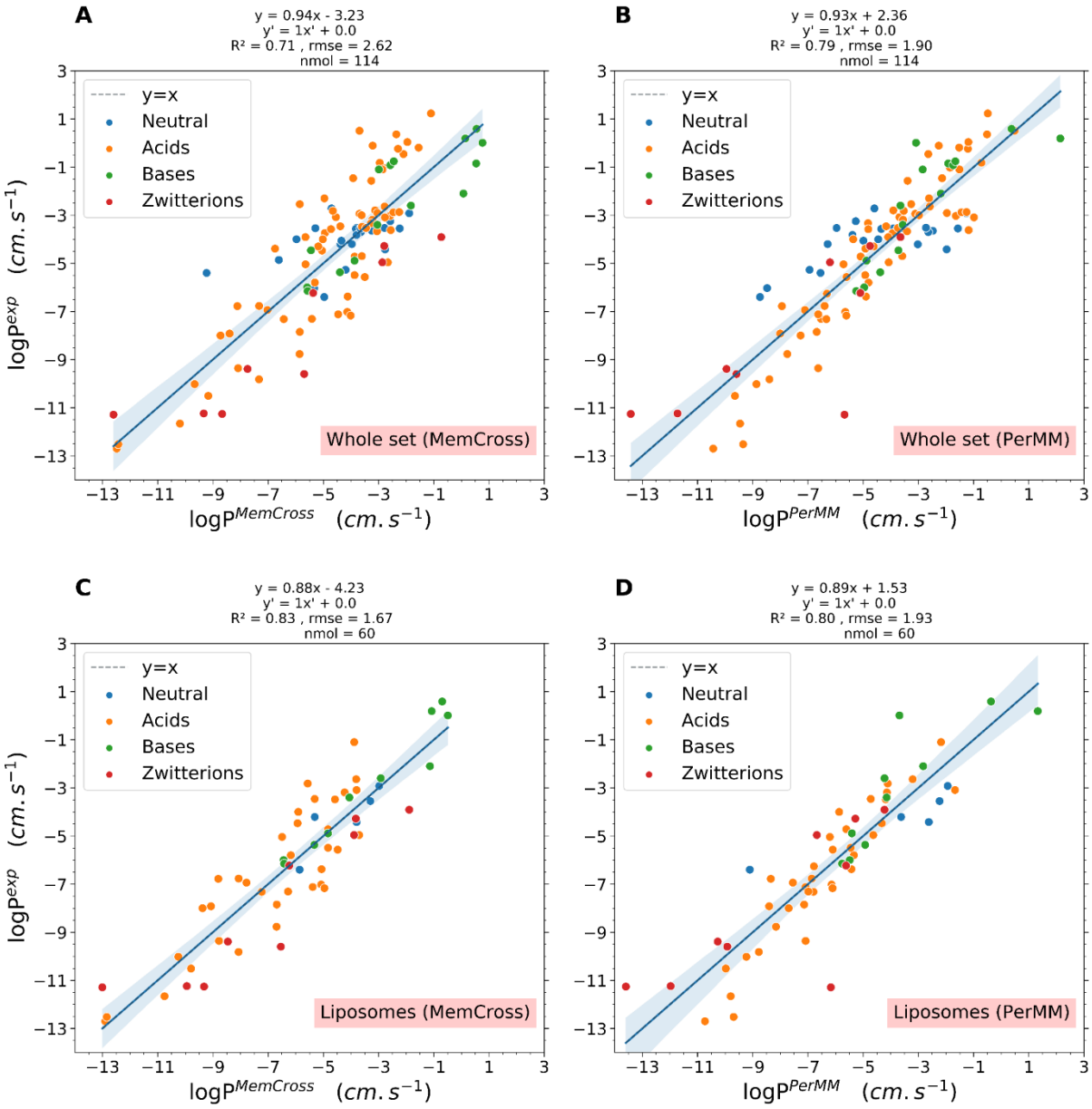


Figure S14: Comparison of of $logP_{erm}^{MemCross}$ (A) and $logP_{erm}^{PerMM}$ (B) with the mixed set of liposomes/BLM (C) and liposomes-only set (D).

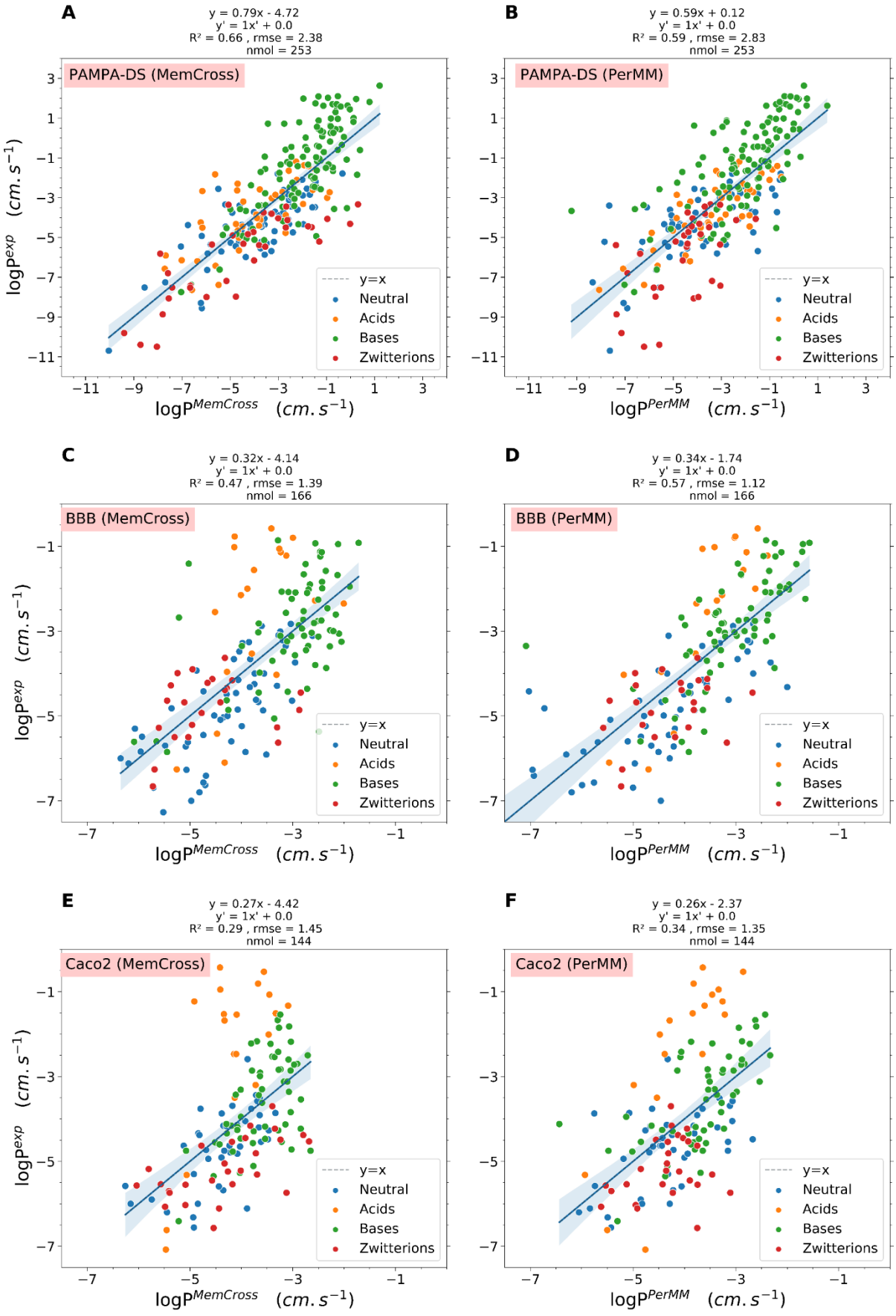


Figure S15: Comparison of $logP_{erm}^{MemCross}$ and $logP_{erm}^{PerMM}$ with the PAMPA-DS data set (A-B), with BBB data set (C-D), and with Caco2/MDCK data set (E-F), respectively.

Table S2: Molecules used in this study with their corresponding $\log P_{erm}^{MemCross}$ and $\log P_{erm}^{exp}$ values.

Molecules	*MemCross LogP$_{erm}$*	*Experimetal LogP$_{erm}$*	*Membrane type*
1,6-Hexanediol	-0.01624	-3.65	BLM
Butanol	1.41852	-2.92	Liposomes
Ethanol	0.48974	-4.42	Liposomes
Isobutyramide	-0.43765	-3.7	BLM
n-Butyramide	-0.60675	-3.52	BLM
Propanol	1.05386	-3.55	Liposomes
Propionamide	-1.22571	-4.21	Liposomes
Erythritol	-1.84684	-6.4	Liposomes
Theophylline	0.47019	-3.53	BLM
1,4-Butanediol	-0.63081	-3.57	BLM
2,3-Dideoxyadenosine	-0.78663	-4.2	BLM
2-Deoxyadenosine	-2.22795	-6.03	BLM
Acetamide	-2.18757	-3.54	BLM
Adenine	-3.59046	-4.86	BLM
Ethylene glycol	-1.18629	-4.06	BLM
Formamide	-2.9115	-4	BLM
Glycerol	-1.0259	-5.27	BLM
Hydrocortisone	0.68849	-3.25	BLM
Prednisolone	-0.6072	-3.82	BLM
Propylene glycol	-0.44829	-3.55	BLM
Urea	-6.36484	-5.4	BLM
Water	-1.56571	-2.72	BLM
D-Glucose	-3.8	-10.52	Liposomes
D-Fructose	-2.1	-9.4	Liposomes
Sucrose	-4.25	-13.1	Liposomes
1-Naphthoic acid	0.92247	0.36	BLM
2-Naphthoic acid	2.25314	1.23	BLM
4-Carboxymethylphenyl acetyl-NHMe	-1.93156	-4.47	Liposomes
4-Carboxymethylphenyl acetyl-Gly-NMe2	-2.20721	-5.8	Liposomes
4-Carboxymethylphenyl acetyl-NMe2	-0.39656	-3.48	Liposomes
p-Hydroxybenzoic acid	-1.399	-3.08	BLM
9-Anthroic acid	-0.49271	0.51	BLM
Acetic acid	-1.84131	-2.3	BLM
Carbamide methylhippuric acid	-5.82982	-8	Liposomes
Carbamoyl-p-Toluic acid	-3.73687	-4.39	BLM
Carboxy methylhippuric acid	-4.34812	-6.77	Liposomes
α-Carboxy p-Toluic acid	-1.82176	-3.74	BLM
Cl methylhippuric acid	-1.2294	-3.46	Liposomes
α-Chloro p-Toluic acid	1.77929	-0.19	BLM
CN methylhippuric acid	-2.56638	-5.04	Liposomes
α-Cyano p-Toluic acid	-0.03919	-1.57	BLM
Hydroxy methylhippuric acid	-2.28293	-6.26	Liposomes

α-Hydroxy p-Toluic acid	0.10707	-2.8	BLM
Methoxy methylhippuric acid	-1.90195	-4	Liposomes
α-Methoxy p-Toluic acid	1.19925	-0.46	BLM
Benzoic acid	0.98645	-0.24	BLM
Butyric acid	0.40472	-1.02	BLM
Hexanoic acid	1.34714	0.04	BLM
Lactic acid	-2.07662	-4.3	BLM
Propionic acid	-0.73673	-1.46	BLM
Salicylic acid	0.01393	-0.11	BLM
Tol GLY-GLYSar	-2.32206	-7.31	BLM
Tol GSar	-0.67994	-4.72	BLM
Tol-Ala	0.47334	-2.64	Liposomes
Tol-Ala-Ala	0.59732	-4.96	Liposomes
Tol-Ala-Ala-Ala	-0.9628	-7.02	Liposomes
Tol- Gly	-1.00E-05	-3.19	Liposomes
Tol Gly-Gly	-0.94718	-6.38	Liposomes
Tol-Gly-Gly-Gly	-2.79398	-8.77	Liposomes
Tol- Gly-Sar- Gly	-3.40142	-7.32	Liposomes
Tol-Sar- Gly	-0.68125	-5.49	Liposomes
Tol-Sar Gly-Gly	-2.7838	-7.85	Liposomes
Tol-Sar-Sar- Gly	-1.31115	-7.12	Liposomes
Tol-Sar	-1.51587	-2.82	Liposomes
2,4-Dihydroxybenzoic acid	-0.04675	-3.33	BLM
2-Hydroxybutyric acid	-2.56778	-3.91	BLM
2-Hydroxycaproic acid	0.17874	-3.67	BLM
2-Hydroxyvaleric acid	-0.2393	-3.53	BLM
Acetylsalicylic acid	0.28451	-0.82	BLM
BromoAcetic acid	0.19282	-2.91	BLM
ChloroAcetic acid	-0.4906	-2.94	BLM
Citric acid	-6.3002	-10.51	Liposomes
Epinephrine	-0.28743	-5.57	Liposomes
Formic acid	-2.78605	-2.54	BLM
Isobutyric acid	0.68119	-3.03	BLM
Isovaleric acid	0.80872	-2.88	BLM
Malic acid	-5.48207	-7.92	Liposomes
PhenylAcetic acid	0.48792	-3.09	Liposomes
Pivalic acid	0.70853	-3.62	BLM
Sebacic acid	-0.41166	-3	BLM
Tiglic acid	1.0331	-2.87	BLM
Valeric acid	0.39185	-1.1	Liposomes
Vanillic acid	-1.58868	-3.58	BLM
Tol-Gly	-5.15422	-9.36	Liposomes
12-Hydroxydodecanoic acid	-0.40215	-4.7	BLM
2-Hydroxynicotinic acid	-5.18328	-6.78	Liposomes
3,5-Dichlorobenzoic acid	-0.83058	-7.17	Liposomes
Carbamide methylhippuric acid	-9.81947	-12.7	Liposomes

Carboxy methylhippuric acid	-9.75713	-12.52	Liposomes
CN methylhippuric acid	-6.82312	-10.02	Liposomes
Hydroxy methylhippuric acid	-7.38991	-11.66	Liposomes
Methoxy methylhippuric acid	-4.35144	-9.82	Liposomes
Salicylic acid	-4.0263	-6.94	Liposomes
Codeine	3.98951	-0.85	BLM
Ethylamine	0.69211	-0.92	BLM
Methylamine	0.25298	-1.1	BLM
Chlorpromazine	4.00177	0.59	Liposomes
Domperidone	1.48224	-2.6	Liposomes
Ethanolamine	-0.6835	-4.89	Liposomes
Histamine	-2.35669	-4.46	BLM
Labetalol	3.49936	-2.1	Liposomes
Propranolol	3.57007	0.19	Liposomes
Tryptamine	0.82336	-0.76	BLM
Verapamil	4.23183	0.01	Liposomes
Methylamine	-2.50347	-6	Liposomes
Ethylamine	-2.48251	-6.15	Liposomes
Propylamine	-1.24315	-5.37	Liposomes
Pentylamine	0.19321	-3.4	Liposomes
Enrofloxacin	2.65011	-3.91	Liposomes
Fleroxacin	0.38521	-4.96	Liposomes
Norfloxacin	-2.27105	-6.23	Liposomes
Pefloxacin	0.45868	-4.28	Liposomes
Glycine	-6.47581	-11.24	Liposomes
L-Tryptophan	-4.7912	-9.39	Liposomes
L-Phenylalanine	-2.61806	-9.6	Liposomes
L-Serine	-5.76824	-11.26	Liposomes
L-Lysine	-9.94105	-11.29	Liposomes
11-b-Hydroxyprogesterone	2.88227	-3.46	L-
11-Dehydrocorticosterone	1.46552	-4.36	PAMPA-DS
11-Deoxycortisol	1.78603	-3.8	PAMPA-DS
Acetaminophen	-1.53702	-5.81	PAMPA-DS
Albendazole	2.75159	-3.12	PAMPA-DS
Alfuzosin	3.51092	-4.34	PAMPA-DS
Aminoglutethimide	-0.24223	-3.76	PAMPA-DS
Aniline	2.01529	-3.71	PAMPA-DS
Bendroflumethiazide	-0.3255	-5.27	PAMPA-DS
Benserazide	-6.75521	-10.7	PAMPA-DS
Budesonide	2.46789	-2.62	PAMPA-DS
Caffeine	1.25375	-5.55	PAMPA-DS
Carbamazepine	0.90506	-3.73	PAMPA-DS
Chloramphenicol	1.57653	-5.3	PAMPA-DS
Chlorthalidone	-1.56055	-6.64	PAMPA-DS
Clonidine	2.85891	-3	PAMPA-DS
Corticosterone	2.60259	-3.86	PAMPA-DS

Cortisone	1.31022	-4.46	PAMPA-DS
Creatinine	-4.8657	-7.52	PAMPA-DS
Cyclothiazide	-2.94835	-5.46	PAMPA-DS
Danazol	5.57251	-1.79	PAMPA-DS
Deoxycorticosterone	3.85085	-2.85	PAMPA-DS
Dexamethasone	1.45582	-4.05	PAMPA-DS
Diazepam	2.79862	-2.44	PAMPA-DS
Erythritol	-1.8474	-8.56	PAMPA-DS
Ethosuximide	0.12484	-5.2	PAMPA-DS
Famciclovir	0.74962	-4.59	PAMPA-DS
Felodipine	4.12118	-3.48	PAMPA-DS
Flavone	3.45306	-2.1	PAMPA-DS
Genistein	1.99184	-4.69	PAMPA-DS
Griseofulvin	3.20544	-3.61	PAMPA-DS
Guanfacine	-0.97434	-2.56	PAMPA-DS
Hydrochlorothiazide	-1.89534	-8.3	PAMPA-DS
Hydrocortisone	0.69603	-4.32	PAMPA-DS
Lamotrigine	1.46983	-5.87	PAMPA-DS
Lansoprazole	2.25359	-3.89	PAMPA-DS
Meprobamate	-0.81755	-5.71	PAMPA-DS
Methylprednisolone	-2.57974	-4.38	PAMPA-DS
Metolazone	0.20647	-4.85	PAMPA-DS
Minoxidil	0.25048	-4.62	PAMPA-DS
Nifedipine	-0.61202	-3.35	PAMPA-DS
Nitrendipine	2.15984	-1.8	PAMPA-DS
Pemoline	-1.84907	-4.93	PAMPA-DS
Phenytoin	0.26141	-4.37	PAMPA-DS
Praziquantel	2.60208	-2.78	PAMPA-DS
Prednisolone	-0.58889	-4.46	PAMPA-DS
Prednisone	2.75626	-4.33	PAMPA-DS
Primidone	-0.06058	-5.44	PAMPA-DS
Progesterone	3.69713	-2.55	PAMPA-DS
Propylthiouracil	-1.3749	-5.36	PAMPA-DS
Quintozene	4.15779	-2.2	PAMPA-DS
Resveratrol	1.26145	-4.38	PAMPA-DS
Testosterone	2.69058	-2.83	PAMPA-DS
Theophylline	0.47457	-5.99	PAMPA-DS
Thiabendazole	2.87504	-3.45	PAMPA-DS
Valdecoxib	2.71161	-4.1	PAMPA-DS
Alprazolam	2.45839	-3.73	PAMPA-DS
Antipyrine	1.82347	-5.69	PAMPA-DS
Etoposide	1.95707	-5.22	PAMPA-DS
Ganciclovir	-7.83583	-8.26	PAMPA-DS
Isocarboxazid	2.9566	-4.75	PAMPA-DS
Netivudine	-3.77393	-7.26	PAMPA-DS
Omeprazole	3.98469	-3.49	PAMPA-DS

Pantoprazole	0.14234	-3.4	PAMPA-DS
Triamcinolone	1.29987	-5.22	PAMPA-DS
Trioxsalen	3.48321	-2.89	PAMPA-DS
b-Naphthoic acid	2.40248	-2.72	PAMPA-DS
3,4-Dihydroxyphenylacetic acid	-2.90888	-6.15	PAMPA-DS
3-Hydroxyphenylacetic acid	-1.90329	-4.25	PAMPA-DS
3-Phenylpropionic acid	1.26349	-3.89	PAMPA-DS
Acetylsalicylic acid	0.29759	-4.45	PAMPA-DS
Carbamoyl-p-Toluic acid	-3.7512	-5.91	PAMPA-DS
α-Carboxy p-Toluic acid	-1.81884	-4.51	PAMPA-DS
α-Chloro p-Toluic acid	1.84601	-3.03	PAMPA-DS
α-Cyano p-Toluic acid	-0.04416	-4.33	PAMPA-DS
α-Hydroxy p-Toluic acid	0.10779	-5.02	PAMPA-DS
α-Methoxy p-Toluic acid	1.23872	-4.13	PAMPA-DS
Benzoic acid	1.01083	-3.94	PAMPA-DS
Benzthiazide	-1.12085	-6.43	PAMPA-DS
Cerivastatin	4.76792	-3.01	PAMPA-DS
Diclofenac	3.55177	-1.37	PAMPA-DS
Flufenamic acid	3.12798	-1.19	PAMPA-DS
Flumequine	1.0282	-3.85	PAMPA-DS
Flurbiprofen	2.26276	-1.78	PAMPA-DS
Fluvastatin	4.89625	-2.73	PAMPA-DS
Furosemide	1.07221	-4.03	PAMPA-DS
Gemfibrozil	4.1455	-1.59	PAMPA-DS
Glipizide	1.17516	-4.41	PAMPA-DS
Ibuprofen	3.16277	-2.11	PAMPA-DS
Indomethacin	2.74208	-1.65	PAMPA-DS
Isoxicam	0.34366	-3.2	PAMPA-DS
Kaempferol	0.11367	-5.66	PAMPA-DS
Ketoprofen	-1.81804	-2.67	PAMPA-DS
Lactic acid	-2.08076	-6.2	PAMPA-DS
Mefenamic acid	3.78755	-1.41	PAMPA-DS
Nalidixic acid	2.58264	-3.88	PAMPA-DS
Naproxen	0.04117	-2.3	PAMPA-DS
Oxolinic acid	0.0533	-4.66	PAMPA-DS
Phenylbutazone	3.94962	-1.96	PAMPA-DS
Piroxicam	3.49688	-3.32	PAMPA-DS
Probenecid	-1.14405	-1.83	PAMPA-DS
Quercetin	2.56898	-4.77	PAMPA-DS
Rosmarinic acid	-0.96183	-7.39	PAMPA-DS
Salicylic acid	0.01882	-2.64	PAMPA-DS
Warfarin	1.21298	-2.59	PAMPA-DS
Zomepirac	4.0084	-2.61	PAMPA-DS
Ceftibuten	-2.36902	-7.64	PAMPA-DS
Chlorothiazide	-3.80057	-6.58	PAMPA-DS
Meloxicam	2.03521	-2.86	PAMPA-DS

Amlodipine	4.92079	0.62	PAMPA-DS
4-Phenylbutylamine	2.56929	-0.59	PAMPA-DS
Acebutolol	1.26768	-3.39	PAMPA-DS
Albuterol	0.31352	-4.92	PAMPA-DS
Alfentanil	5.8105	-3.53	PAMPA-DS
Alprenolol	3.63416	0.02	PAMPA-DS
Amantadine	2.66254	-1.21	PAMPA-DS
Amiloride	-2.40586	-7.38	PAMPA-DS
Amitriptyline	5.2088	1.3	PAMPA-DS
Amoxapine	5.07623	-1.66	PAMPA-DS
Astemizole	6.16269	1	PAMPA-DS
Atenolol	0.50824	-5.06	PAMPA-DS
Bepridil	5.84237	1.63	PAMPA-DS
Bremazocine	5.0693	-1.49	PAMPA-DS
Brompheniramine	5.54741	-0.35	PAMPA-DS
Bupivacaine	2.26436	-2.07	PAMPA-DS
Buspirone	4.20922	-2.48	PAMPA-DS
Butacaine	4.95271	0.25	PAMPA-DS
Chlorpheniramine	5.32323	-0.6	PAMPA-DS
Chlorpromazine	5.23662	1.62	PAMPA-DS
Chlorprothixene	5.96525	1.44	PAMPA-DS
Citalopram	2.88082	-0.95	PAMPA-DS
Clemastine	6.2966	1.96	PAMPA-DS
Clotrimazole	2.78923	-1.31	PAMPA-DS
Clozapine	4.23818	-0.39	PAMPA-DS
Cyproheptadine	5.45069	0.44	PAMPA-DS
Desipramine	4.14644	1.74	PAMPA-DS
Dextromethorphan	4.92205	-0.18	PAMPA-DS
Dextrorphan	3.7634	-1.34	PAMPA-DS
Diltiazem	4.83928	-1.33	PAMPA-DS
Diphenhydramine	4.59085	-0.71	PAMPA-DS
Disopyramide	4.42037	-1.14	PAMPA-DS
Domperidone	1.49756	-2.78	PAMPA-DS
Doxepin	5.21672	0.44	PAMPA-DS
Ephedrine	2.89256	-2.9	PAMPA-DS
Famotidine	-2.93706	-7.75	PAMPA-DS
Fendiline	4.56654	1.62	PAMPA-DS
Guanabenz	2.31108	-1.34	PAMPA-DS
Haloperidol	5.64944	0.05	PAMPA-DS
Hydroxyzine	4.69895	-1.5	PAMPA-DS
Imipramine	4.96521	0.98	PAMPA-DS
Indinavir	5.13547	-3.57	PAMPA-DS
Labetalol	4.18911	-4.94	PAMPA-DS
Lidocaine	3.21448	-1.42	PAMPA-DS
Loxapine	5.79587	-1.09	PAMPA-DS
Meperidine	3.29378	0.79	PAMPA-DS

Methadone	4.50692	0.08	PAMPA-DS
Metipranolol	4.54594	0.3	PAMPA-DS
Metoclopramide	3.92224	-1.94	PAMPA-DS
Metoprolol	4.24169	-1.17	PAMPA-DS
Mexiletine	2.87936	-0.45	PAMPA-DS
Miconazole	3.62115	-0.45	PAMPA-DS
Midazolam	2.80229	-2.47	PAMPA-DS
Morantel	4.04093	-2.05	PAMPA-DS
Morphine	-0.33939	-3.59	PAMPA-DS
Nadolol	3.24761	-4.34	PAMPA-DS
Nalbuphine	3.02482	-2.56	PAMPA-DS
Naltrindole	1.47383	-0.94	PAMPA-DS
Nicotine	3.02892	-3.42	PAMPA-DS
Nortriptyline	4.80707	2.02	PAMPA-DS
Ondansetron	0.7504	-2.38	PAMPA-DS
Oxprenolol	3.70907	-0.6	PAMPA-DS
Papaverine	4.01434	-2.44	PAMPA-DS
Penbutolol	5.05142	1.7	PAMPA-DS
Pilocarpine	-0.73541	-4.88	PAMPA-DS
Pindolol	2.24678	-1.75	PAMPA-DS
Pirenzepine	2.40859	-3.46	PAMPA-DS
Practolol	1.14287	-3.4	PAMPA-DS
Procaine	4.09916	-2.46	PAMPA-DS
Procyclidine	5.26787	1.7	PAMPA-DS
Promethazine	5.32968	0.96	PAMPA-DS
Propafenone	2.50117	0.72	PAMPA-DS
Propoxyphene	1.64327	0.72	PAMPA-DS
Propranolol	4.3307	0.43	PAMPA-DS
Pumafentrine	4.94957	0.59	PAMPA-DS
Pyridoxine	-0.90603	-6.62	PAMPA-DS
Quetiapine	6.35019	-1.85	PAMPA-DS
Quinidine	1.49982	-1.56	PAMPA-DS
Quinine	4.35467	-1.05	PAMPA-DS
Ranitidine	1.44528	-5.14	PAMPA-DS
Risperidone	4.10991	-2.01	PAMPA-DS
Rizatriptan	3.24786	-2.76	PAMPA-DS
Scopolamine	2.37284	-3	PAMPA-DS
Sertraline	5.33341	2.1	PAMPA-DS
Sotalol	-0.01277	-4.83	PAMPA-DS
Sulpiride	0.8342	-4.57	PAMPA-DS
Sumatriptan	2.43265	-4.18	PAMPA-DS
Tamoxifen	3.51289	1.98	PAMPA-DS
Terfenadine	7.53512	2.63	PAMPA-DS
Thioridazine	7.00065	1.81	PAMPA-DS
Topotecan	0.93092	-5.12	PAMPA-DS
Trihexyphenidyl	4.30693	2.09	PAMPA-DS

Trimipramine	5.48201	1.58	PAMPA-DS
U69593	4.4204	0.37	PAMPA-DS
Venlafaxine	4.21827	-1.63	PAMPA-DS
Zimelidine	4.32949	-0.43	PAMPA-DS
Ziprasidone	6.49834	-0.61	PAMPA-DS
Zolpidem	3.03973	-3.06	PAMPA-DS
Doxorubicin	0.47934	-3.67	PAMPA-DS
Fentanyl	3.49252	-0.95	PAMPA-DS
Fluvoxamine	4.31464	0.88	PAMPA-DS
Hydralazine	-0.54341	-4.53	PAMPA-DS
Noscapine	2.94002	-2.25	PAMPA-DS
Perphenazine	6.15359	0.81	PAMPA-DS
Prazosin	4.4254	-2.58	PAMPA-DS
Primaquine	4.2301	0.56	PAMPA-DS
Timolol	4.18507	-0.97	PAMPA-DS
Triamterene	0.1727	-3.58	PAMPA-DS
Trimethoprim	1.5171	-3.38	PAMPA-DS
Verapamil	5.09213	0.26	PAMPA-DS
Zolmitriptan	1.25711	-1.71	PAMPA-DS
Chlortetracycline	1.17764	-5.39	PAMPA-DS
Enalaprilat	-2.44683	-7.43	PAMPA-DS
Acrivastine	4.777	-4.07	PAMPA-DS
Amoxicillin	-3.63343	-6.8	PAMPA-DS
Cefadroxil	-3.91008	-8.87	PAMPA-DS
Cefatrizine	-5.94374	-9.81	PAMPA-DS
Cephalexin	-2.46053	-7.53	PAMPA-DS
Cephradine	-0.04804	-7.98	PAMPA-DS
Cetirizine	2.67187	-4.13	PAMPA-DS
Ciprofloxacin	1.50219	-5.47	PAMPA-DS
CNV97100	0.34639	-5.32	PAMPA-DS
CNV97102	5.97687	-4.32	PAMPA-DS
CNV97103	1.8357	-3.81	PAMPA-DS
CNV97104	6.39353	-3.34	PAMPA-DS
Enoxacin	-0.21098	-4.9	PAMPA-DS
Gabapentin	-0.43832	-3.36	PAMPA-DS
Gatifloxacin	2.58379	-4.5	PAMPA-DS
Gly Pro	-5.08197	-10.4	PAMPA-DS
Gly Sar	-4.21536	-10.5	PAMPA-DS
Grepafloxacin	3.9485	-4.49	PAMPA-DS
L-Alanine	-3.59116	-8.07	PAMPA-DS
L-Arginine	-4.04139	-5.82	PAMPA-DS
L-Dopa	-3.4121	-7.52	PAMPA-DS
L-Leucine	-0.54627	-7.19	PAMPA-DS
Lomefloxacin	0.8757	-4.73	PAMPA-DS
L-Phenylalanine	-1.30861	-5.36	PAMPA-DS
L-Tryptophan	-1.59833	-8	PAMPA-DS

Me-ciprofloxacin	3.76121	-4.44	PAMPA-DS
Melphalan	2.60868	-3.45	PAMPA-DS
Norfloxacin	0.91219	-6.16	PAMPA-DS
Ofloxacin	4.35006	-5.21	PAMPA-DS
Sarafloxacin	0.58489	-4.84	PAMPA-DS
Sparfloxacin	2.18626	-4.04	PAMPA-DS
Trovafloxacin	1.41039	-3.75	PAMPA-DS
1,4-Butanediol	-0.62817	-5.03	BBB
Alfuzosin	3.51092	-4.64	BBB
Antipyrine	1.82347	-4	BBB
Arabinose	-1.7351	-6.63	BBB
Butanol	1.67449	-2.88	BBB
Caffeine	1.25375	-4	BBB
Carbamazepine	0.90506	-3.26	BBB
Corticosterone	2.60259	-4.29	BBB
Creatinine	-4.8657	-6.69	BBB
D-Fructose	-2.08754	-6.8	BBB
Dianhydrogalactitol	-0.1603	-5.6	BBB
Diazepam	2.79862	-3.01	BBB
Erythritol	-1.8474	-6.57	BBB
Estradiol	3.4353	-2.83	BBB
Ethanol	0.51607	-3.28	BBB
Ethosuximide	0.12484	-4.46	BBB
Ethylene glycol	-1.1921	-4.99	BBB
Etoposide	1.95707	-5.91	BBB
Formamide	-2.91436	-5.72	BBB
Galactitol	-1.69265	-6.41	BBB
Hispidulin	4.74867	-3.11	BBB
Hydrocortisone	0.69603	-5.85	BBB
Hydroxyurea	-6.84745	-6	BBB
Hypoxanthine	-5.55769	-5.49	BBB
Isocarboxazid	2.9566	-3.22	BBB
Isopropanol	0.97972	-3.66	BBB
Lamotrigine	1.46983	-4.67	BBB
Loratadine	0.32538	-3.48	BBB
Mannitol	-2.87104	-6.27	BBB
Meprobamate	-0.81755	-5.09	BBB
Methanol	-0.00183	-3.66	BBB
Methylurea	-4.15785	-5.7	BBB
Metronidazole	0.22732	-4.85	BBB
Pemoline	-1.84907	-5.45	BBB
Phenytoin	0.26141	-4.15	BBB
Procarbazine	1.33385	-4.62	BBB
Progesterone	3.69713	-3.74	BBB
Propylene glycol	-0.44513	-4.49	BBB
Ritonavir	1.75983	-4.87	BBB

Sucrose	-4.26908	-7.27	BBB
Temazepam	3.13066	-3.35	BBB
Testosterone	2.69058	-3.1	BBB
Theobromine	-0.38733	-5	BBB
Theophylline	0.47457	-5.24	BBB
Thiourea	-5.98672	-5.3	BBB
Thymidine	-5.6378	-5.84	BBB
Thymine	-2.25984	-3.93	BBB
Urea	-6.37329	-6.12	BBB
Xanthine	-2.78675	-5.62	BBB
Zaleplon	1.23219	-4.25	BBB
Adenosine	-1.8427	-4.42	BBB
D-Glucose	-3.77133	-4.82	BBB
Methylprednisolone	-2.57974	-7	BBB
Phenelzine	0.95798	-4.32	BBB
Butyric acid	0.41771	-2.15	BBB
Flurbiprofen	2.26276	-0.58	BBB
Fluvastatin	4.89625	-2.28	BBB
Ibuprofen	3.16277	-1.22	BBB
Kynurenic acid	-1.02061	-5.42	BBB
Naringenin	-0.42099	-3.96	BBB
Octanoic acid	2.82911	-1.14	BBB
Probenecid	-1.14405	-2.55	BBB
Quercetin	2.56898	-4.03	BBB
Quinolinic acid	-3.45653	-6.26	BBB
Taurocholic acid	-0.56335	-6.1	BBB
Thiothixene	6.63522	-2.35	BBB
Tolbutamide	1.08213	-3.53	BBB
Valproic acid	0.81017	-2	BBB
Warfarin	1.21298	-1.56	BBB
Ascorbic acid	-3.06831	-2.54	BBB
Chlorambucil	3.56424	-0.8	BBB
Indomethacin	2.74208	-1.06	BBB
Naproxen	0.04117	-0.77	BBB
Neotrofin	-5.05166	-3.8	BBB
Salicylic acid	0.01882	-1.02	BBB
Alfentanil	5.8105	-2.98	BBB
Amantadine	2.66254	-0.86	BBB
Aminoguanidine	-4.04399	-5.85	BBB
Amitriptyline	5.2088	-1.13	BBB
Amoxapine	5.07623	-2.75	BBB
Astemizole	6.16269	-2.61	BBB
Bremazocine	5.0693	-2.76	BBB
Brompheniramine	5.54741	-1.58	BBB
Bupropion	3.71539	-2.09	BBB
Buspirone	4.20922	-2.53	BBB

Chlorpheniramine	5.32323	-1.84	BBB
Chlorpromazine	5.23662	-1.23	BBB
Cimetidine	-6.03474	-5.61	BBB
Citalopram	2.88082	-2.07	BBB
Clemastine	6.2966	-0.95	BBB
Clozapine	4.23818	-3.11	BBB
Codeine	4.9126	-3.8	BBB
CP-141938	1.55664	-3.98	BBB
Diltiazem	4.83928	-2.81	BBB
Diphenhydramine	4.59085	-1.9	BBB
Domperidone	1.49756	-4.45	BBB
Donepezil	6.25877	-1.68	BBB
Dopamine	-3.32486	-2.68	BBB
Doxepin	5.21672	-1.24	BBB
Fentanyl	3.49252	-2.24	BBB
Fluoxetine	4.91776	-0.93	BBB
Fluphenazine	4.48729	-3.35	BBB
Guanidine	-4.67972	-5.6	BBB
Haloperidol	5.64944	-2.46	BBB
Hydroxyzine	4.69895	-3.04	BBB
Imatinib	2.32093	-3.7	BBB
Indinavir	5.13547	-5.37	BBB
Lidocaine	3.21448	-3.24	BBB
Loxapine	5.79587	-3.36	BBB
Meperidine	3.29378	-2.08	BBB
Mesoridazine	-2.73434	-1.41	BBB
Methadone	4.50692	-2.02	BBB
Metoclopramide	3.92224	-2.86	BBB
Midazolam	2.80229	-3.11	BBB
Mirtazapine	5.28522	-2.75	BBB
Morphine	-0.33939	-4.86	BBB
Naltrindole	1.47383	-3.03	BBB
Olanzapine	4.18562	-2.73	BBB
Oxycodone	2.98717	-3.48	BBB
Pergolide	5.30534	-1.14	BBB
Perphenazine	6.15359	-2.61	BBB
Pramipexole	3.14048	-2.57	BBB
Prazosin	4.4254	-4.36	BBB
Propranolol	4.3307	-1.42	BBB
Pyrimethamine	3.68917	-3.57	BBB
Quetiapine	6.35019	-3.06	BBB
Quinidine	1.49982	-2.82	BBB
Quinine	4.35467	-3.45	BBB
Rimantadine	3.53366	0.13	BBB
Risperidone	4.10991	-2.94	BBB
Rizatriptan	3.24786	-4.43	BBB

Saquinavir	-0.43986	-4.63	BBB
Selegiline	3.46389	-3.12	BBB
Sertraline	5.33341	-1.99	BBB
Sufentanil	5.57783	-3.87	BBB
Sumatriptan	2.43265	-5.06	BBB
Terfenadine	7.53512	-0.92	BBB
Thioridazine	7.00065	-1.95	BBB
Trazodone	5.10217	-3.13	BBB
Trifluoperazine	6.00862	-3	BBB
U69593	4.4204	-2.1	BBB
Venlafaxine	4.21827	-1.66	BBB
Verapamil	5.09213	-2.26	BBB
Ziprasidone	6.49834	-3.25	BBB
Doxorubicin	0.47934	-3.35	BBB
Gabapentin	-0.43832	-4.34	BBB
Glycine	-2.75814	-5.5	BBB
Grepafloxacin	3.9485	-4.86	BBB
L-Alanine	-3.59116	-5.5	BBB
L-Arginine	-4.04139	-4.64	BBB
L-Aspartic acid	-4.90294	-6.66	BBB
L-Dopa	-3.4121	-3.99	BBB
L-Glutamic acid	-4.79927	-6.26	BBB
L-Glutamine	-4.54	-5.28	BBB
L-Histidine	-3.81368	-4.28	BBB
L-Isoleucine	-0.10646	-4.16	BBB
L-Leucine	-0.54627	-3.63	BBB
L-Lysine	-1.954	-4.93	BBB
L-Methionine	-0.54663	-4.39	BBB
L-Ornithine	-2.99849	-4.68	BBB
L-Phenylalanine	-1.30861	-4.13	BBB
L-Threonine	-2.53356	-5.21	BBB
L-Tryptophan	-1.59833	-4.22	BBB
L-Tyrosine	-2.50359	-3.9	BBB
L-Valine	-0.82165	-4.68	BBB
Melphalan	2.60868	-5.27	BBB
Tiagabine	4.0367	-4.45	BBB
Cetirizine	2.67187	-5.63	BBB
Antipyrine	1.82347	-4.05	Caco2
Benserazide	-6.75521	-5.58	Caco2
Chloramphenicol	1.57653	-4.47	Caco2
Erythritol	-1.8474	-6.56	Caco2
Ethosuximide	0.12484	-4.91	Caco2
Guanfacine	-0.97434	-4.73	Caco2
Hydrochlorothiazide	-1.89534	-6.32	Caco2
Isocarboxazid	2.9566	-4.54	Caco2
Meprobamate	-0.81755	-4.94	Caco2

Minoxidil	0.25048	-5.68	Caco2
Netivudine	-3.77393	-6.2	Caco2
Pemoline	-1.84907	-5.3	Caco2
Praziquantel	2.60208	-3.44	Caco2
Primidone	-0.06058	-5.59	Caco2
Propylthiouracil	-1.3749	-3.76	Caco2
Thiabendazole	2.87504	-3.51	Caco2
Water	-1.56556	-6	Caco2
Acetaminophen	-1.53702	-4.34	Caco2
Alfuzosin	3.51092	-4.27	Caco2
Amprenavir	-1.4552	-4.38	Caco2
Caffeine	1.25375	-4.14	Caco2
Carbamazepine	0.90506	-3.69	Caco2
Clonidine	2.85891	-3.91	Caco2
Creatinine	-4.8657	-5.9	Caco2
Dexamethasone	1.45582	-4.65	Caco2
Diazepam	2.79862	-4.2	Caco2
Etoposide	1.95707	-6.11	Caco2
Famciclovir	0.74962	-4.79	Caco2
Flavone	3.45306	-4.48	Caco2
Ganciclovir	-7.83583	-6.99	Caco2
Genistein	1.99184	-2.59	Caco2
Hydrocortisone	0.69603	-4.63	Caco2
Lamotrigine	1.46983	-4.45	Caco2
Lansoprazole	2.25359	-3.76	Caco2
Loratadine	0.32538	-4.75	Caco2
Methylprednisolone	-2.57974	-4.63	Caco2
Omeprazole	3.98469	-3.86	Caco2
Phenytoin	0.26141	-4.16	Caco2
Resveratrol	1.26145	-4.82	Caco2
Testosterone	2.69058	-3.58	Caco2
Theophylline	0.47457	-4.17	Caco2
Urea	-6.37329	-6	Caco2
Pantoprazole	0.14234	-3.87	Caco2
Flumequine	1.0282	-2.47	Caco2
Acetylsalicylic acid	0.29759	-1.53	Caco2
Ampicillin	-3.85522	-7.08	Caco2
Ceftibuten	-2.36902	-5.32	Caco2
Chlorothiazide	-3.80057	-6.62	Caco2
Furosemide	1.07221	-3.5	Caco2
Warfarin	1.21298	-1.54	Caco2
Diclofenac	3.55177	-1.07	Caco2
Fluvastatin	4.89625	-1.33	Caco2
Glipizide	1.17516	-2.47	Caco2
Ibuprofen	3.16277	-0.53	Caco2
Indomethacin	2.74208	-0.81	Caco2

Isoxicam	0.34366	-1.68	Caco2
Ketoprofen	-1.81804	-1.23	Caco2
Naproxen	0.04117	-0.95	Caco2
Piroxicam	3.49688	-2.01	Caco2
Quercetin	2.56898	-3.2	Caco2
Salicylic acid	0.01882	-0.43	Caco2
Zomepirac	4.0084	-1.51	Caco2
Alfentanil	5.8105	-3.54	Caco2
Bremazocine	5.0693	-2.86	Caco2
Brompheniramine	5.54741	-2.7	Caco2
Chlorpheniramine	5.32323	-2.72	Caco2
Clemastine	6.2966	-2.5	Caco2
Dextromethorphan	4.92205	-2.6	Caco2
Dextrorphan	3.7634	-2.53	Caco2
Diphenhydramine	4.59085	-3.12	Caco2
Ephedrine	2.89256	-2.91	Caco2
Hydroxyzine	4.69895	-4.13	Caco2
Loxapine	5.79587	-4.23	Caco2
Nicotine	3.02892	-3.62	Caco2
Practolol	1.14287	-3.43	Caco2
Pumafentrine	4.94957	-3.36	Caco2
Rizatriptan	3.24786	-4.18	Caco2
Scopolamine	2.37284	-4.57	Caco2
Sulpiride	0.8342	-4.16	Caco2
Acebutolol	1.26768	-4.19	Caco2
Alprenolol	3.63416	-2.23	Caco2
Amantadine	2.66254	-2.17	Caco2
Amiloride	-2.40586	-4.75	Caco2
Amoxapine	5.07623	-3.84	Caco2
Atenolol	0.50824	-4.34	Caco2
Citalopram	2.88082	-2.99	Caco2
Desipramine	4.14644	-1.67	Caco2
Diltiazem	4.83928	-3.12	Caco2
Domperidone	1.49756	-4.46	Caco2
Famotidine	-2.93706	-6.41	Caco2
Guanabenz	2.31108	-2.86	Caco2
Imipramine	4.96521	-1.82	Caco2
Indinavir	5.13547	-4.72	Caco2
Labetalol	4.18911	-4.27	Caco2
Metoclopramide	3.92224	-2.54	Caco2
Metoprolol	4.24169	-1.85	Caco2
Midazolam	2.80229	-3.44	Caco2
Morphine	-0.33939	-4.55	Caco2
Nadolol	3.24761	-4.47	Caco2
Nalbuphine	3.02482	-3.3	Caco2
Pindolol	2.24678	-2.22	Caco2

Pirenzepine	2.40859	-5.11	Caco2
Prazosin	4.4254	-4.54	Caco2
Propranolol	4.3307	-1.54	Caco2
Quinidine	1.49982	-3.31	Caco2
Quinine	4.35467	-2.83	Caco2
Ranitidine	1.44528	-5.27	Caco2
Saquinavir	-0.43986	-5.35	Caco2
Sotalol	-0.01277	-4.6	Caco2
Sumatriptan	2.43265	-4.29	Caco2
Terfenadine	7.53512	-3.74	Caco2
Timolol	4.18507	-2.42	Caco2
Topotecan	0.93092	-4.77	Caco2
Trimethoprim	1.5171	-3.95	Caco2
Venlafaxine	4.21827	-2.84	Caco2
Verapamil	5.09213	-2.18	Caco2
Ziprasidone	6.49834	-4.75	Caco2
Zolmitriptan	1.25711	-4.26	Caco2
Doxorubicin	0.47934	-4.12	Caco2
CNV97100	0.34639	-5.05	Caco2
CNV97102	5.97687	-4.38	Caco2
CNV97103	1.8357	-4.45	Caco2
CNV97104	6.39353	-4.53	Caco2
Gly-Pro	-5.08197	-5.18	Caco2
Me-ciprofloxacin	3.76121	-3.7	Caco2
Sarafloxacin	0.58489	-5.24	Caco2
Amoxicillin	-3.63343	-5.7	Caco2
Cefadroxil	-3.91008	-6.07	Caco2
Cefatrizine	-5.94374	-5.57	Caco2
Cephalexin	-2.46053	-6.03	Caco2
Cephradine	-0.04804	-6.11	Caco2
Cetirizine	2.67187	-5.31	Caco2
Ciprofloxacin	1.50219	-5.22	Caco2
Enalaprilat	-2.44683	-5.55	Caco2
Gatifloxacin	2.58379	-4.81	Caco2
Gly-Sar	-4.21536	-5.54	Caco2
Grepafloxacin	3.9485	-4.23	Caco2
L-Alanine	-3.59116	-5.74	Caco2
L-Leucine	-0.54627	-5.45	Caco2
Lomefloxacin	0.8757	-4.54	Caco2
Phenylalanine	-1.30861	-4.63	Caco2
Norfloxacin	0.91219	-5.54	Caco2
Ofloxacin	4.35006	-4.49	Caco2
Sparfloxacin	2.18626	-4.16	Caco2
Acrivastine	4.777	-5.74	Caco2
Gabapentin	-0.43832	-6.57	Caco2

Appendix 3. Supplementary information for Chapter VII: “Thorough evaluation of binding sites in the human ABCC4/MRP4 Alphafold2 structure”

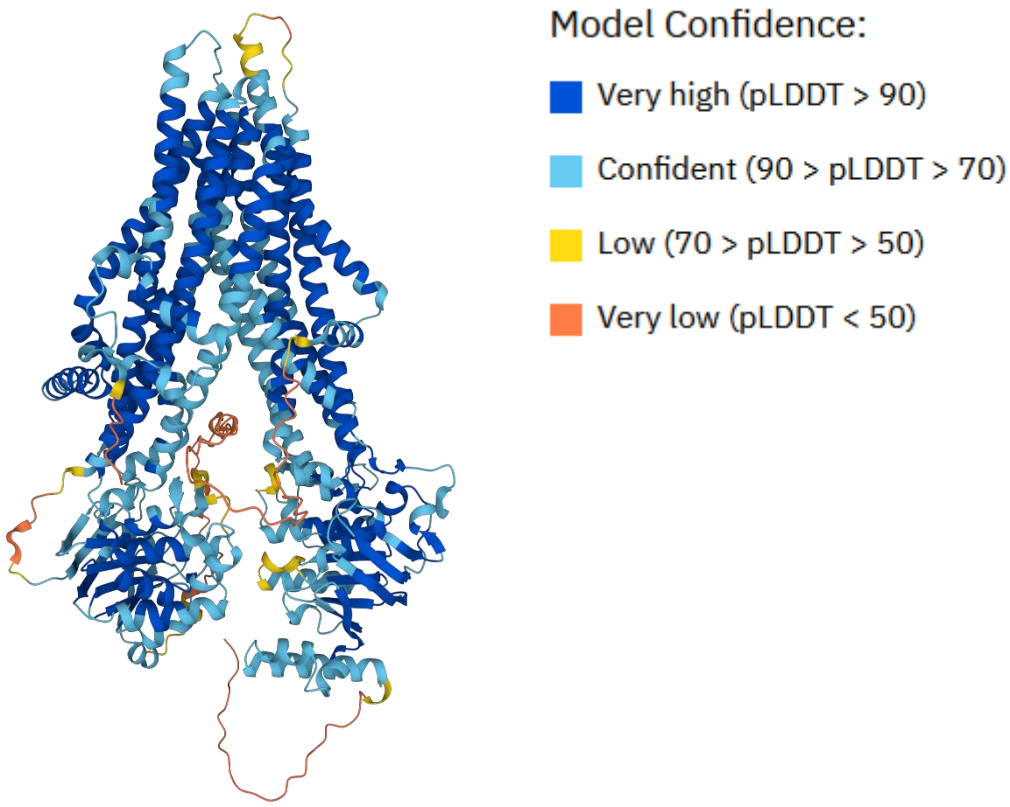


Figure S16: Structure of the human MRP4 from AlphaFold. Colors indicate the confidence of the model.

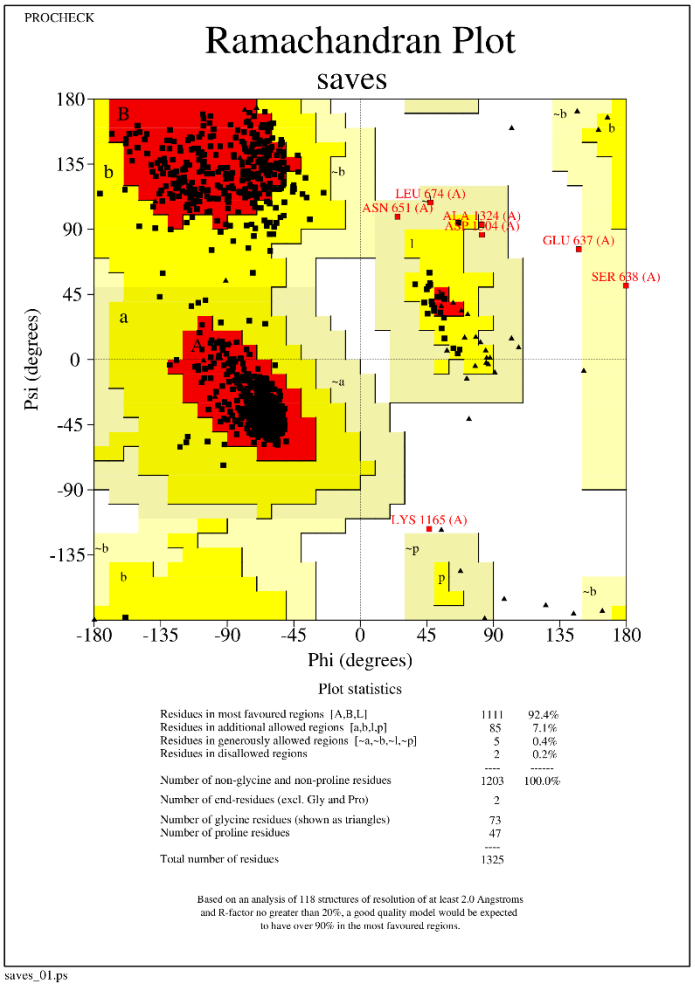


Figure S17: Ramachandran plot of hMRP4 structure from AlphaFold.

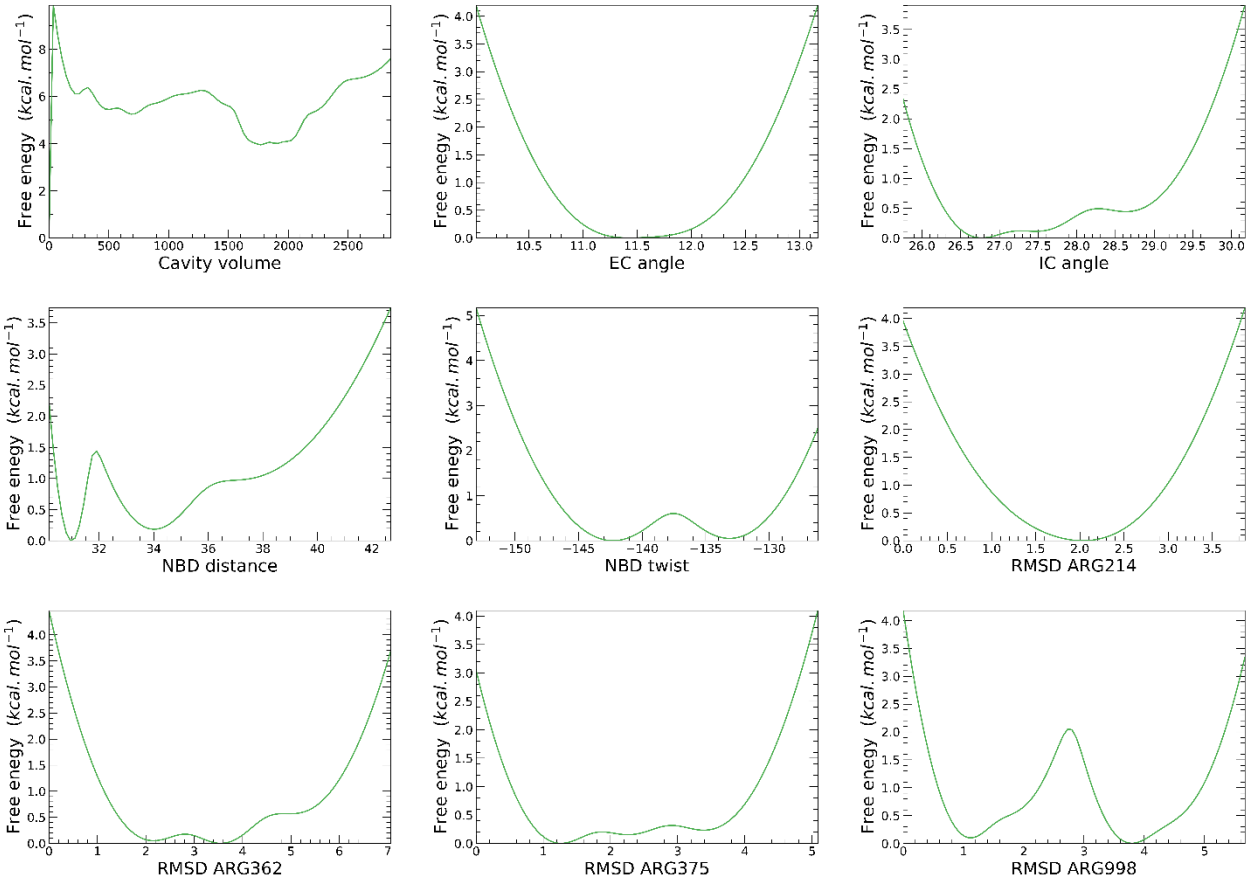


Figure S18: The one-dimentional clustering using InfleCS method for single parameters from equilibration MD simulations of hMRp4 structure.

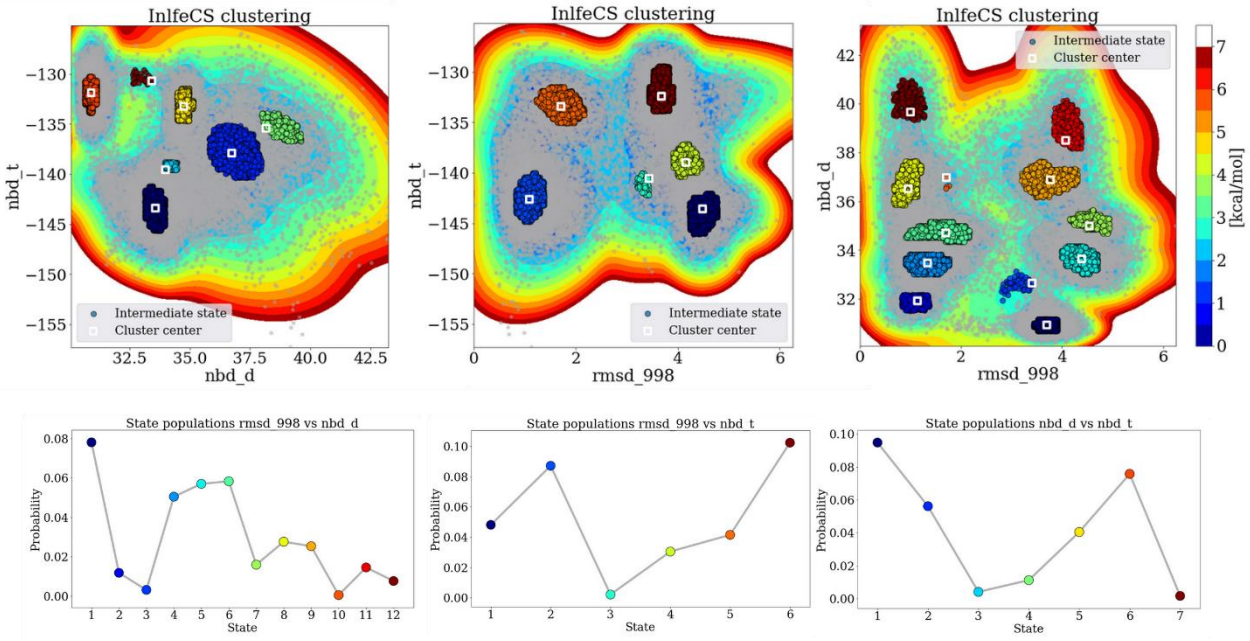


Figure S19: The 2-dimentional clustering using InfleCS method. The top panel represents the free energy landscapes sampled during MD simulations according to couples of selected parameters (i.e., NBD twist, NBD distance and ARG998 RMSD). The bottom panel depicts the cluster probabilities.

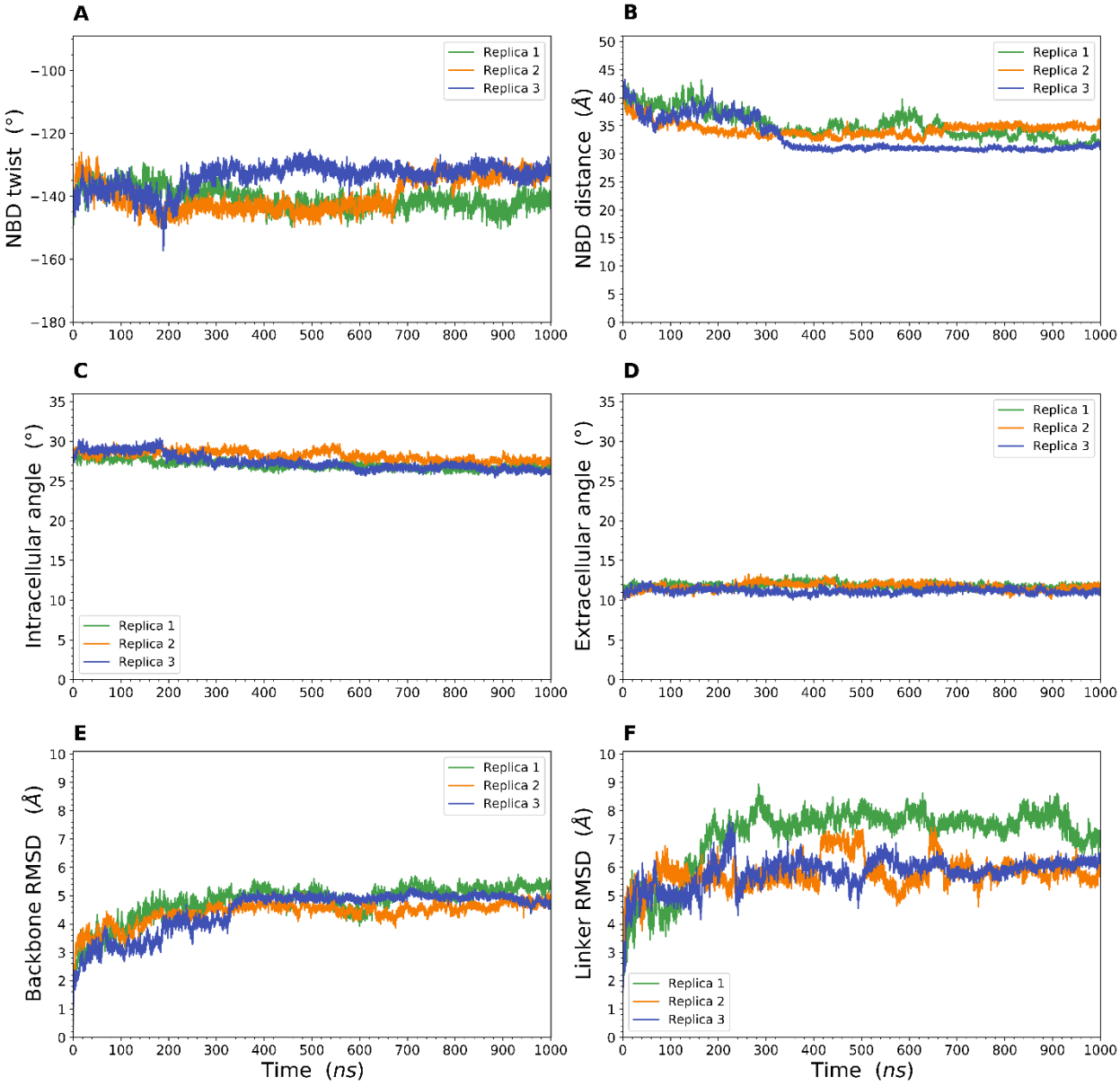


Figure S20: Evolution of different parameters during equilibration MD simulations over 3 replicas.

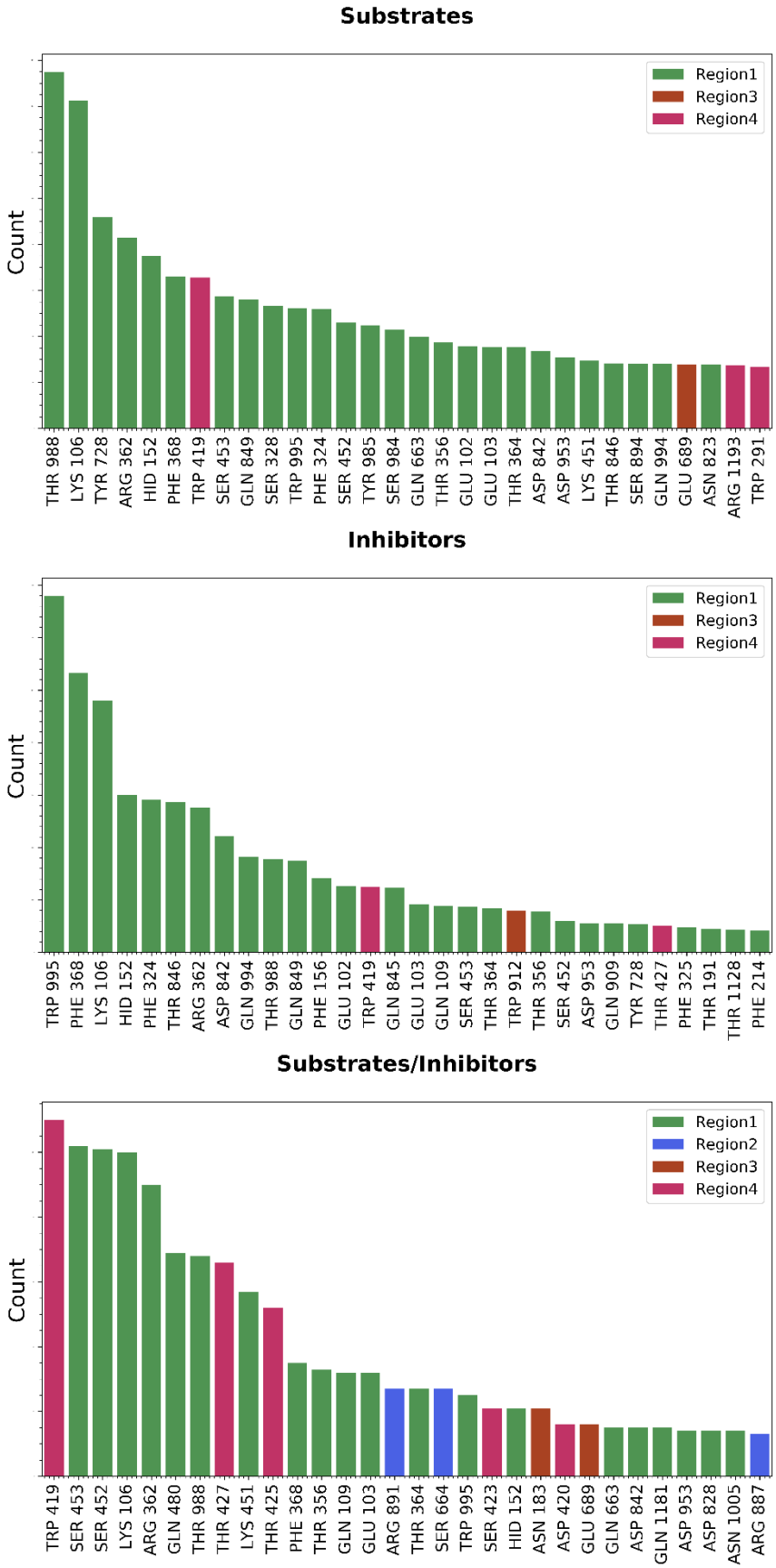


Figure S21: Count of the first 30 non-hydrophobic aminoacids in contact with the ligands (substrates, inhibitors and substrates/inhibitors) and their corresponding regions. Counts are normalized in each ligand class by the maximum number of counts among residues.

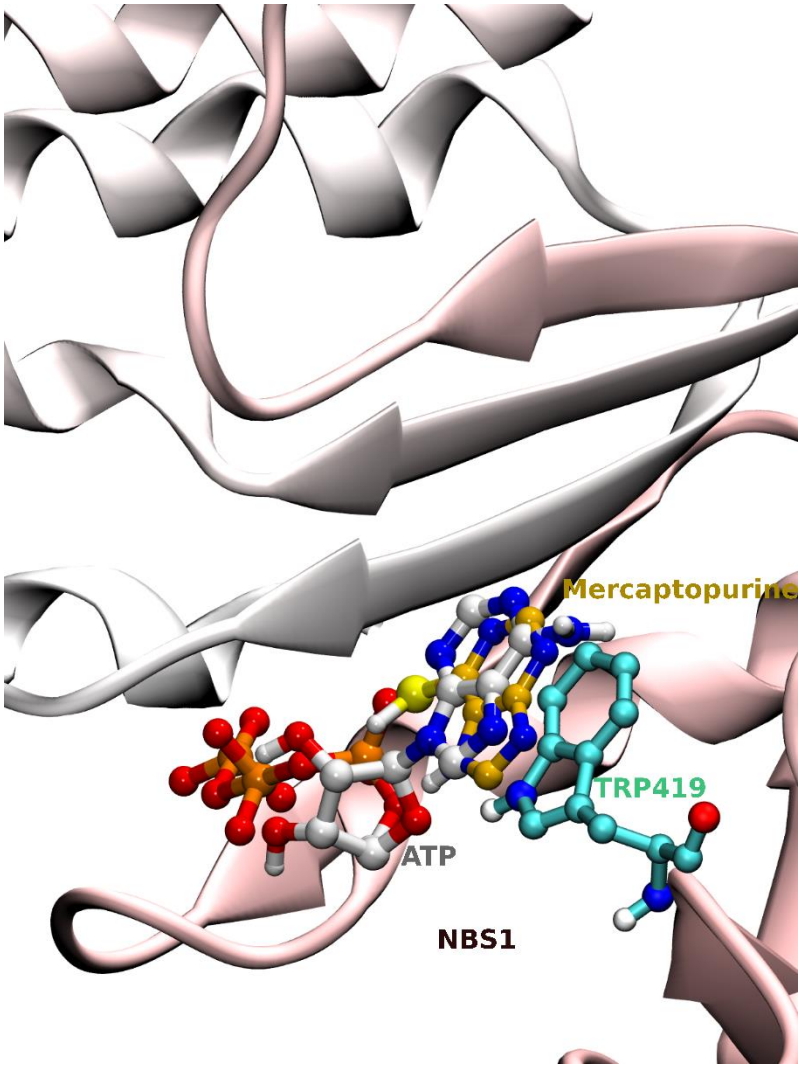


Figure S22: Poses of ATP and Mercaptopurine (substrate) in the nucleotide-binding site 1 (NBS1). Both ligands highlight a pi-stacking interaction with TRP419.

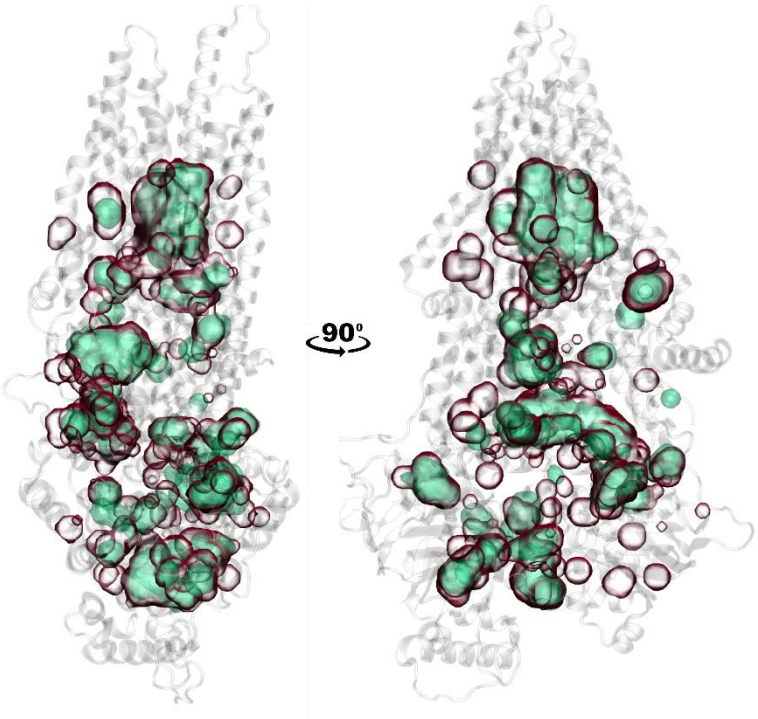


Figure S23: Comparison between substrate (green surfaces) and inhibitor (red transparent surfaces) COM locations in the hMRP4 structure.

Table S3: Coordinates of the grids used for the docking procedure

Grid name	X	Y	Z	Xmin	Xmax	Ymin	Ymax	Zmin	Zmax
Grid_1	66.5	87.5	140.5	52	81	73	102	126	155
Grid_2	66.5	87.5	121.5	52	81	73	102	107	136
Grid_3	66.5	87.5	102.5	52	81	73	102	88	117
Grid_4	66.5	87.5	83.5	52	81	73	102	69	98
Grid_5	66.5	87.5	64.5	52	81	73	102	50	79
Grid_6	66.5	68.5	140.5	52	81	54	83	126	155
Grid_7	66.5	68.5	121.5	52	81	54	83	107	136
Grid_8	66.5	68.5	102.5	52	81	54	83	88	117
Grid_9	*66.5*	*68.5*	*83.5*	*52*	*81*	*54*	*83*	*69*	*98*
Grid_10	66.5	68.5	64.5	52	81	54	83	50	79
Grid_11	47.5	87.5	140.5	33	62	73	102	126	155
Grid_12	47.5	87.5	121.5	33	62	73	102	107	136
Grid_13	47.5	87.5	102.5	33	62	73	102	88	117
Grid_14	47.5	87.5	83.5	33	62	73	102	69	98
Grid_15	47.5	87.5	64.5	33	62	73	102	50	79
Grid_16	47.5	68.5	140.5	33	62	54	83	126	155
Grid_17	47.5	68.5	121.5	33	62	54	83	107	136
Grid_18	47.5	68.5	102.5	33	62	54	83	88	117
Grid_19	47.5	68.5	83.5	33	62	54	83	69	98
Grid_20	47.5	68.5	64.5	33	62	54	83	50	79
Grid_21	79.5	85.5	78.5	65	94	71	100	64	93
Grid_22	79.5	85.5	64.5	65	94	71	100	50	79
Grid_23	79.5	71.5	78.5	65	94	57	86	64	93
Grid_24	79.5	71.5	64.5	65	94	57	86	50	79
Grid_25	20.5	80.5	73.5	6	35	66	95	59	88
Grid_26	20.5	71.5	73.5	6	35	57	86	59	88
Grid_ATP_NBS1	67.6	70	74.8	25	54	82	111	60	89
Grid_ATP_NBS2	39.9	96	74.5	53	82	56	85	60	89

Table S4: The different ligands used in the docking procedure of the current study and their corresponding binding regions.

CHEMBL_Code	Molecule_name	Type	Regions
CHEMBL34259	METHOTREXATE	Substrate	['Region1', 'Region2', 'Region3', 'Region5']
CHEMBL495	ALPROSTADIL	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7']
CHEMBL548	DINOPROSTONE	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7']
CHEMBL84	TOPOTECAN	Substrate	['Region1']
CHEMBL2074980	Bimane-glutathione	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CHEMBL35	FUROSEMIDE	Substrate	['Region1', 'Region5']
CHEMBL1425	MERCAPTOPURINE	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region9', 'Region10', 'Region11', 'Region12']
CHEMBL1543	GLUTATHIONE	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region9']
CHEMBL484	ADEFOVIR	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region9']
CHEMBL257695	AZIDOTHYMIDINEMONOPHOSPHATE	Substrate	Others
CHEMBL483	TENOFOVIR	Substrate	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region9']
CHEMBL495244		Inhibitor	['Region1', 'Region2', 'Region4', 'Region5']
CHEMBL213947		Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CHEMBL897	PROBENECID	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6']
CHEMBL139	DICLOFENAC	Inhibitor	['Region1', 'Region3', 'Region5', 'Region7']
CHEMBL932	DIPYRIDAMOLE	Inhibitor	['Region1']
CHEMBL571	KETOPROFEN	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5',

			'Region6', 'Region7', 'Region8']
CHEMBL521	IBUPROFEN	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region8', 'Region9', 'Region10', 'Region11', 'Region12']
CHEMBL6	INDOMETHACIN	Inhibitor	['Region5']
CHEMBL6966	VERAPAMIL	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6']
CHEMBL563	FLURBIPROFEN	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5', 'Region6', 'Region8']
CHEMBL15177	MK-571	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5']
CHEMBL2074695		Inhibitor	['Region1', 'Region5']
CHEMBL1478	LITHOCHOLICACID	Inhibitor	['Region1']
CHEMBL118	CELECOXIB	Inhibitor	['Region1', 'Region3', 'Region6']
CHEMBL411070	GLYCOCHOLICACID	Inhibitor	['Region1', 'Region2']
CHEMBL285913	TREQUINSIN	Inhibitor	['Region1', 'Region5']
CHEMBL185878	TAUROCHENODEOXYCHOLATE	Inhibitor	['Region1']
CHEMBL122	ROFECOXIB	Inhibitor	['Region1', 'Region3']
CHEMBL832	SULFINPYRAZONE	Inhibitor	['Region1', 'Region3', 'Region6']
CHEMBL272427	TAURURSODIOL	Inhibitor	['Region1']
CHEMBL28079	ZAPRINAST	Inhibitor	['Region1', 'Region3', 'Region5', 'Region6']
CHEMBL2074738		Inhibitor	['Region1']
CHEMBL15870	INDOPROFEN	Inhibitor	['Region1', 'Region3', 'Region5', 'Region6', 'Region11']
CHEMBL727	THIOGUANINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region9', 'Region10', 'Region11', 'Region12']
CHEMBL1552426	THROMBOXANE-B2	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region9']
CHEMBL258818	LITHOGLYCOCHOLATE	Inhibitor	['Region1', 'Region2', 'Region4']

CHEMBL1551	URSODIOL	Inhibitor	['Region1', 'Region2']
CHEMBL1325783	PROSTAGLANDINF1ALPHA	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region11']
CHEMBL388590	BENZBROMARONE	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7']
CHEMBL224867	TAUROCHOLICACID	Inhibitor	['Region1']
CHEMBL192	SILDENAFIL	Inhibitor	['Region1']
CHEMBL1084644	PROSTAGLANDINA1	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6', 'Region7']
CHEMBL5	NALIDIXICACID	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region11']
CHEMBL939	GEFITINIB	Inhibitor	['Region1', 'Region3']
CHEMBL193	NIFEDIPINE	Inhibitor	['Region1', 'Region2', 'Region5']
CHEMBL1726	NISOLDIPINE	Inhibitor	['Region1', 'Region2', 'Region5']
CHEMBL393220	ATORVASTATINCALCIUM	Inhibitor	['Region1', 'Region3', 'Region4']
CHEMBL1017	TELMISARTAN	Inhibitor	['Region1']
CHEMBL1200896	DINOPROSTTROMETHAMINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6', 'Region7']
CHEMBL1200799	TRAVOPROST	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6']
CHEMBL4063475		Inhibitor	Others
CHEMBL2074591		Inhibitor	['Region1', 'Region4', 'Region5']
CHEMBL415324	TRANILAST	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6']
CHEMBL383634	GLIQUIDONE	Inhibitor	['Region1', 'Region3', 'Region5']
CHEMBL611	TERAZOSIN	Inhibitor	['Region1', 'Region3']
CHEMBL3349339		Inhibitor	['Region1', 'Region3', 'Region4', 'Region5']
CHEMBL957	BOSENTAN	Inhibitor	['Region1']
CHEMBL2074644		Inhibitor	['Region1']
CHEMBL2074584		Inhibitor	['Region1', 'Region2', 'Region4']

CHEMBL1162492	ESTRADIOLDISULFATE	Inhibitor	Others
CHEMBL408701	TAUROLITHOCHOLICACID	Inhibitor	['Region1']
CHEMBL1200621	GEMIFLOXACINMESYLATE	Inhibitor	['Region1', 'Region2', 'Region4', 'Region5']
CHEMBL1200523	CEFAZOLINSODIUM	Inhibitor	['Region1']
CHEMBL685	MEBENDAZOLE	Inhibitor	['Region1', 'Region3']
CHEMBL1316	CARPROFEN	Inhibitor	['Region1', 'Region3']
CHEMBL77305	RUBITECAN	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5']
CHEMBL421	SULFASALAZINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CHEMBL1201124	KETOROLACTROMETHAMINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region7', 'Region9', 'Region10', 'Region11', 'Region12']
CHEMBL408	TROGLITAZONE	Inhibitor	['Region1', 'Region3', 'Region5']
CHEMBL260315		Inhibitor	['Region1', 'Region2', 'Region4']
CHEMBL1164729	FEBUXOSTAT	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region10']
CHEMBL1421	DASATINIB	Inhibitor	['Region1']
CHEMBL729	LOPINAVIR	Inhibitor	['Region1', 'Region3']
CHEMBL1324	TOLCAPONE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region11']
CHEMBL589973		Inhibitor	['Region1', 'Region3', 'Region5']
CHEMBL416146	ETORICOXIB	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region8']
CHEMBL126075	DILAZEP	Inhibitor	['Region1', 'Region3', 'Region5']
CHEMBL1010	CEFOTAXIMESODIUM	Inhibitor	['Region1', 'Region2', 'Region4', 'Region5']
CHEMBL413965	FELBINAC	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region8', 'Region12']

CHEMBL953	ENTACAPONE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region8']
CHEMBL603	ZAFIRLUKAST	Inhibitor	['Region1', 'Region5']
CHEMBL572	NITROFURANTOIN	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region9', 'Region11', 'Region12']
CHEMBL1095930	CEFUROXIMEAXETIL	Inhibitor	['Region1', 'Region2', 'Region4']
CHEMBL65	CAMPTOTHECIN	Inhibitor	['Region1', 'Region3']
CHEMBL295416	PIRINIXICACID	Inhibitor	['Region1', 'Region3', 'Region5', 'Region6']
CHEMBL1276010	UNDECYLENICACID	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region8', 'Region9', 'Region10', 'Region11', 'Region12']
CHEMBL1484	NICARDIPINE	Inhibitor	['Region1']
CHEMBL1294	QUINIDINE	Inhibitor	Others
CHEMBL1096	AMLEXANOX	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region10']
CHEMBL1737	SILDENAFILCITRATE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5']
CHEMBL367463	PRANOPROFEN	Inhibitor	['Region1', 'Region3']
CHEMBL1071	OXAPROZIN	Inhibitor	['Region1', 'Region3', 'Region5', 'Region6']
CHEMBL595	PIOGLITAZONE	Inhibitor	['Region1', 'Region3']
CHEMBL388978	STAUROSPORINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region6']
CHEMBL180022	NERATINIB	Inhibitor	['Region1']
CHEMBL1131	ACITRETIN	Inhibitor	['Region1', 'Region3', 'Region8']
CHEMBL19224	PAPAVERINE	Inhibitor	['Region1', 'Region3', 'Region5']
CHEMBL412272	DEOXYCHOLATE	Inhibitor	['Region1', 'Region4']
CHEMBL2074982		Inhibitor	['Region1', 'Region5']
CHEMBL1744447	ROSUVASTATIN-CALCIUM	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CHEMBL4101392		Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']

CHEMBL282724	SITAXENTAN	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CHEMBL1272	REPAGLINIDE	Inhibitor	['Region1', 'Region3']
CHEMBL56337	EPALRESTAT	Inhibitor	['Region1']
CHEMBL551978	DROTAVERINE	Inhibitor	['Region1']
CHEMBL21333	PRANLUKAST	Inhibitor	['Region1']
CHEMBL995	LOSARTANPOTASSIUM	Inhibitor	['Region1', 'Region3']
CHEMBL95	TACRINE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5', 'Region6', 'Region7', 'Region9', 'Region10', 'Region12']
CHEMBL1481	GLIMEPIRIDE	Inhibitor	['Region1']
CHEMBL1200692	OLMESARTANMEDOXOMIL	Inhibitor	['Region1']
CHEMBL4064157		Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6', 'Region7']
CHEMBL1456	MYCOPHENOLATEMOFETIL	Inhibitor	['Region1', 'Region3', 'Region4', 'Region5']
CHEMBL799	CILOSTAZOL	Inhibitor	['Region1', 'Region3']
CHEMBL1292	CLOFAZIMINE	Inhibitor	['Region1']
CHEMBL418509	NITROBENZYLMERCAPTOPURINERIBONU	Inhibitor	['Region4']
CHEMBL2218867	CLOPROSTENOLSODIUM	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region6']
CHEMBL121	ROSIGLITAZONE	Inhibitor	['Region1', 'Region2', 'Region3', 'Region5', 'Region12']
CHEMBL4097197		Inhibitor	['Region1']
CHEMBL2	PRAZOSIN	Inhibitor	['Region1', 'Region3', 'Region4']
CHEMBL623	NEFAZODONE	Inhibitor	['Region1', 'Region2', 'Region3']
CEE1	Ceefourin1	Inhibitor	['Region1', 'Region2', 'Region3', 'Region4', 'Region5']
CEE2	Ceefourin2	Inhibitor	['Region1', 'Region3', 'Region4', 'Region6']
CHEMBL1622	FOLICACID	Substrate/Inhibitor	['Region1', 'Region2', 'Region3', 'Region5']
CHEMBL259898	PRASTERONESULFURICACID	Substrate/Inhibitor	Others

CHEMBL395336		Substrate/Inhibitor	['Region2', 'Region3', 'Region4', 'Region5']
CHEMBL1697724	ESTRADIOL17BGLUCURONIDE	Substrate/Inhibitor	['Region1']
CHEMBL1679	LEUCOVORIN	Substrate/Inhibitor	['Region1', 'Region2', 'Region3', 'Region5']
CHEMBL316966	CYCLICADENOSINEMONOPHOSPHATE	Substrate/Inhibitor	['Region1', 'Region2', 'Region4', 'Region5']

Table S5: Coordinates of the defined regions after projecting ligand COMs on the hMRP4 structure from docking calculations.

	Grid center coordinates			Grid dimension		
	X	**Y**	**Z**	**X**	**Y**	**Z**
Region1	58.0	76.0	129.0	24.0	24.4	25.0
Region2	53.5	83.0	107.5	13.5	35.0	16.0
Region3	58.0	93.1	89.8	41.5	17.0	19.5
Region4	69.7	69.2	77.5	14.0	18.0	20.0
Region5	53.0	76.4	62.5	13.0	22.4	22.0
Region6	75.6	73.3	116.4	10.0	11.0	7.5
Region7	53.5	95.0	127.0	17.0	13.5	23.0
Region8	40.0	82.5	129.0	8.5	15.0	27.0
Region9	57.4	76.9	87.9	9.4	10.0	7.5
Region10	46.0	65.3	88.0	18.0	10.0	7.0
Region11	31.8	69.0	76.3	13.5	9.5	8.5
Region12	39.3	81.0	62.0	13.5	25.0	20.0

Table S6: Fraction of positively charged (ARG, LYS), negatively charged (GLU, ASP), polar (SER, THR, ASN, GLN), aromatic (TRP, PHE, TYR, HID, HIE, CYS) and hydrophobic residues in each of the identified regions.

	Positively charged	**Negatively charged**	**Polar**	**Aromatic**	**Hydrophobic**
Region1	0.1343	0.0896	0.2687	0.2239	0.2836
Region2	0.2766	0.2128	0.2340	0.1064	0.1702
Region3	0.1264	0.1264	0.3103	0.1379	0.2989
Region4	0.0968	0.1290	0.2903	0.0968	0.3871
Region5	0.2000	0.0889	0.2667	0.1333	0.3111
Region6	0.0000	0.0000	0.2000	0.4000	0.4000
Region7	0.2273	0.1364	0.2273	0.2727	0.1364
Region8	0.1250	0.0000	0.1250	0.5000	0.2500
Region9	0.4000	0.2000	0.4000	0.0000	0.0000
Region10	0.2727	0.1818	0.1818	0.1818	0.1818
Region11	0.5000	0.0000	0.0000	0.2500	0.2500
Region12	0.1364	0.1136	0.2727	0.0455	0.4318

Table S7: List of ligands that are present in Region 1, Region 2 and Region 3. Only ligands exhibiting a binding energy not greater than 1 kcal/mol compare to the best pose are considered.

CHEMBL_Code	Molecule_name	Type
CHEMBL34259	METHOTREXATE	Substrate
CHEMBL495	ALPROSTADIL	Substrate
CHEMBL548	DINOPROSTONE	Substrate
CHEMBL2074980	Bimane-glutathione	Substrate
CHEMBL1425	MERCAPTOPURINE	Substrate
CHEMBL1543	GLUTATHIONE	Substrate
CHEMBL484	ADEFOVIR	Substrate
CHEMBL483	TENOFOVIR	Substrate
CHEMBL213947		Inhibitor
CHEMBL897	PROBENECID	Inhibitor
CHEMBL521	IBUPROFEN	Inhibitor
CHEMBL6966	VERAPAMIL	Inhibitor
CHEMBL15177	MK-571	Inhibitor
CHEMBL118	CELECOXIB	Inhibitor
CHEMBL272427	TAURURSODIOL	Inhibitor
CHEMBL28079	ZAPRINAST	Inhibitor
CHEMBL15870	INDOPROFEN	Inhibitor
CHEMBL727	THIOGUANINE	Inhibitor
CHEMBL1552426	THROMBOXANE-B2	Inhibitor
CHEMBL1325783	PROSTAGLANDINF1ALPHA	Inhibitor
CHEMBL388590	BENZBROMARONE	Inhibitor
CHEMBL1084644	PROSTAGLANDINA1	Inhibitor
CHEMBL5	NALIDIXICACID	Inhibitor
CHEMBL939	GEFITINIB	Inhibitor
CHEMBL1200896	DINOPROSTTROMETHAMINE	Inhibitor
CHEMBL1200799	TRAVOPROST	Inhibitor
CHEMBL415324	TRANILAST	Inhibitor
CHEMBL611	TERAZOSIN	Inhibitor

CHEMBL77305	RUBITECAN	Inhibitor
CHEMBL421	SULFASALAZINE	Inhibitor
CHEMBL1201124	KETOROLACTROMETHAMINE	Inhibitor
CHEMBL1164729	FEBUXOSTAT	Inhibitor
CHEMBL1324	TOLCAPONE	Inhibitor
CHEMBL589973		Inhibitor
CHEMBL416146	ETORICOXIB	Inhibitor
CHEMBL126075	DILAZEP	Inhibitor
CHEMBL1010	CEFOTAXIMESODIUM	Inhibitor
CHEMBL953	ENTACAPONE	Inhibitor
CHEMBL572	NITROFURANTOIN	Inhibitor
CHEMBL1095930	CEFUROXIMEAXETIL	Inhibitor
CHEMBL1276010	UNDECYLENICACID	Inhibitor
CHEMBL1484	NICARDIPINE	Inhibitor
CHEMBL1096	AMLEXANOX	Inhibitor
CHEMBL1737	SILDENAFILCITRATE	Inhibitor
CHEMBL1071	OXAPROZIN	Inhibitor
CHEMBL388978	STAUROSPORINE	Inhibitor
CHEMBL1744447	ROSUVASTATINCALCIUM	Inhibitor
CHEMBL4101392		Inhibitor
CHEMBL282724	SITAXENTAN	Inhibitor
CHEMBL21333	PRANLUKAST	Inhibitor
CHEMBL95	TACRINE	Inhibitor
CHEMBL4064157		Inhibitor
CHEMBL2218867	CLOPROSTENOLSODIUM	Inhibitor
CHEMBL121	ROSIGLITAZONE	Inhibitor
CHEMBL623	NEFAZODONE	Inhibitor
CHEMBLCEE1	Ceefourin1	Inhibitor
CHEMBL1622	FOLICACID	Substrate/Inhibitor
CHEMBL395336		Substrate/Inhibitor

CHEMBL1679	LEUCOVORIN	Substrate/Inhibitor
CHEMBL316966	CYCLICADENOSINEMONOPHOSPHATE	Substrate/Inhibitor

Table S8: List of ligands that bind only Region 1. Only ligands exhibiting a binding energy not greater than 1 kcal/mol compare to the best pose are considered.

CHEMBL_Code	**Molecule_name**	**Type**
CHEMBL1697724	ESTRADIOL17BGLUCURONIDE	Substrate/Inhibitor
CHEMBL1292	CLOFAZIMINE	Inhibitor
CHEMBL4097197		Inhibitor
CHEMBL1200692	OLMESARTANMEDOXOMIL	Inhibitor
CHEMBL1481	GLIMEPIRIDE	Inhibitor
CHEMBL551978	DROTAVERINE	Inhibitor
CHEMBL56337	EPALRESTAT	Inhibitor
CHEMBL1294	QUINIDINE	Inhibitor
CHEMBL1421	DASATINIB	Inhibitor
CHEMBL1162492	ESTRADIOLDISULFATE	Inhibitor
CHEMBL957	BOSENTAN	Inhibitor
CHEMBL1017	TELMISARTAN	Inhibitor
CHEMBL192	SILDENAFIL	Inhibitor
CHEMBL2074738		Inhibitor
CHEMBL185878	TAUROCHENODEOXYCHOLATE	Inhibitor

Appendices

In silico exploration of membrane transport events: prediction of passive permeation coefficients and binding sites of human MRP4

La prédiction du transport membranaire des xénobiotiques est cruciale pour les industries pharmaceutique, cosmétique et agro-alimentaire. Le franchissement des membranes est en effet directement à la biodisponibilité des actifs utilisés dans ces différents domaines. Il repose sur deux processus majeurs : la perméation passive à travers les bicouches lipidiques et le transport à travers des protéines membranaires. Comprendre les interactions des xénobiotiques avec les différents composants des membranes cellulaires est un passage obligé pour rationaliser ces deux processus. Atteindre cette compréhension à l'échelle atomique est un réel défi au vu de la diversité des membranes cellulaires et de la diversité chimique des xénobiotiques.

Plusieurs méthodes expérimentales utilisent des lignées cellulaires pour mesurer la vitesse de ces processus. Des modèles membranaires plus simplistes ont également été utilisés pour déterminer les coefficients de perméation passive (premier processus). D'autre part, des modèles *in silico*, en particulier utilisant des simulations de dynamique moléculaire (MD), permettent de prédire les coefficients de perméation passive avec une précision grandissante, mais au prix d'un temps de calcul important. Dans cette thèse, nous avons utilisé des simulations de MD biaisées pour prédire les coefficients de perméation passive à travers les membranes. Nous avons développé le protocole MemCross, basé sur la méthode AWH, et l'avons testé sur une série de 13 molécules. Ensuite, nous avons étendu l'utilisation de MemCross à plusieurs centaines de xénobiotiques et nous avons comparé les coefficients de perméation calculés aux valeurs expérimentales. MemCross présente une bonne corrélation avec les données expérimentales provenant de modèles de membranes de phosphatidylcholine (R^2 = 0,83). Les coefficients de corrélation avec les modèles expérimentaux plus complexes utilisant des cellules sont plus faibles. MemCross est un outil prometteur car il est le premier à utiliser l'AWH pour prédire les coefficients de perméation avec des simulations de MD « tout-atome ». Le coût de calcul relativement faible de cette méthodologie a permis le premier benchmark sur des centaines de coefficients de perméation avec une description « tout-atome ».

Le processus de transport actif (deuxième processus) a également été abordé. Nous avons utilisé le récent modèle Alphafold 2 du transporteur humain MRP4. Des simulations de MD et de docking moléculaire, sur la protéine entière, nous ont permis de mettre en évidence les sites de liaison potentiels de tous les substrats et inhibiteurs connus de MRP4. Nous avons notamment mis en évidence plusieurs régions de liaison contiguës à la chambre principale de la protéine. Nos résultats suggèrent que trois de ces régions constituent un canal d'entrée iso-énergétique vers la chambre centrale. De plus, nous avons révélé de potentiels nouveaux résidus clés interagissant avec les ligands.

Mots-clés : transport membranaire, dynamique moléculaire, AWH, coefficient de perméation passive, transporteurs ABC, hMRP4, docking moléculaire

In silico exploration of membrane transport events: prediction of passive permeation coefficients and binding sites of human MRP4

Predicting membrane transport of xenobiotics is crucial to the pharmaceutical, cosmetic and food industries. Membrane crossing of active compounds used in these different fields is directly related to their bioavailability. There exist two major membrane crossing processes: passive permeation across lipid bilayers, and transport through membrane proteins. To reach a full understanding of these processes requires an atomic-scale rationalization of all intermolecular interactions between xenobiotics and membrane components. This is a delicate challenge because of the wide structural diversity of membranes and xenobiotics.

There exist several experimental protocols based on cultured cell-based membrane models to study both transport events. More simplistic membrane models have also been used to determine passive permeation coefficients (first process). Alternatively, theoretical tools, in particular molecular dynamics (MD) simulations, have been shown to predict the passive permeability with a growing accuracy, albeit for an important computational cost. Here, we used biased MD simulations to predict passive membrane permeation coefficients across membranes. We have developed a protocol named MemCross, based on the AWH method, and we tested it on a series of 13 molecules. Then, we extended our predictions to 350 xenobiotics and the predicted permeation coefficients were compared to experimental values. The results from MemCross correlated well with the experimental values obtained with phosphatidylcholine membrane models (R^2 = 0.83). Lower correlation coefficients were obtained with more complex cell-based assays. MemCross is a promising tool, as it is the first to use AWH to predict permeation coefficients based on all-atom MD simulations. Moreover, its relatively low computational cost allowed the first benchmark on hundreds of permeation coefficients with an all-atom description.

The active transport process (second process) was also tackled. Using the recent Alphafold 2 model of the human MRP4 transporter, we performed MD simulations and molecular docking on the whole protein. A comprehensive mapping of potential binding sites was established for all known substrates and inhibitors of MRP4. We highlighted several binding regions besides the main chamber of the protein. Our results have suggested that three of these regions constitute an iso-energetic entry channel to the central chamber. In addition, we revealed new potential key residues interacting with the ligands.

Key words: Membrane transport, MD simulations, AWH, passive permeation coefficient de, ABC transporters, hMRP4, molecular docking