Universidade de Aveiro Departamento de Ciências Médicas
2018

Igor Oliveira Marques

Transporte transmembranar de aniões por moléculas sintéticas: uma investigação por métodos de modelação molecular

Anion transmembrane transport by synthetic molecules: a molecular modelling investigation

Universidade de Aveiro Departamento de Ciências Médicas
2018

Igor Oliveira Marques

Transporte transmembranar de aniões por moléculas sintéticas: uma investigação por métodos de modelação molecular

Anion transmembrane transport by synthetic molecules: a molecular modelling investigation

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biomedicina, realizada sob a orientação científica do Doutor Vítor Manuel Sousa Félix, Professor Associado com Agregação do Departamento de Química da Universidade de Aveiro

Apoio financeiro da Fundação para a Ciência e a Tecnologia (FCT), do Ministério da Educação e Ciência, do Fundo Social Europeu no âmbito do III Quadro Comunitário de Apoio, através de bolsa de doutoramento SFRH/BD/87520/2012. Mais ainda, os estudos apresentados nesta tese foram realizados com recursos computacionais adquiridos sob o projeto P2020-PTDC/QEQ-SUP/4283/2014, financiado pela FCT e pelo Fundo Europeu de Desenvolvimento Regional (FEDER) através do COMPETE 2020 − Programa Operacional Competitividade e Internacionalização.

Aos meus afetos

o júri

presidente

Professora Doutora Ana Isabel Couto Neto da Silva Miranda
professora catedrática, Universidade de Aveiro

Professor Doutor Fernando Manuel Sebastião Silva Fernandes
professor catedrático, Faculdade de Ciências da Universidade de Lisboa

Professor Doutor Cláudio Manuel Simões Loureiro Nunes Soares
professor associado, Instituto de Tecnologia Química e Biológica da Universidade Nova de Lisboa

Professor Doutor Paulo José Garcia de Lemos Trigueiros de Martel
professor auxiliar, Universidade do Algarve

Professor Doutor Paulo Jorge de Almeida Ribeiro Claro
professor associado com agregação, Universidade de Aveiro

Professor Doutor Vítor Manuel Sousa Félix
professor associado com agregação, Universidade de Aveiro

agradecimentos

Ao Professor Vítor Félix, pela dedicação, paciência, exigência, tempo e oportunidades, sem os quais este trabalho não teria sido possível. Ao longo destes anos foram muitas as coisas que aprendi e que me fizeram crescer em várias valências, não só a nível científico. Sem o seu esforço continuado não teria tido as condições que tive para levar a cabo este projeto, nem tão pouco para alicerçar o meu futuro. Fica aqui o meu mais sincero obrigado!

Aos elementos que ao longo dos anos constituíram o Grupo de Modelação Molecular e Biofísica Computacional, por tudo o que convosco aprendi e por todos os desafios partilhados. Uma palavra de apreço ao Pedro e à Miranda pela ajuda que deram na concretização deste projeto.

Aos Professores Phil A. Gale, Paul D. Beer e Anthony P. Davis, pela oportunidade de desenvolver projetos científicos tão desafiantes na área da Química supramolecular.

À Secção Autónoma de Ciências da Saúde (atual Departamento de Ciências Médicas), CICECO e iBiMED, por me terem facultado as condições necessárias ao desenvolvimento do meu projeto de doutoramento.

À FCT, pelo financiamento e confiança depositada ao longo dos anos.

À comunidade do Crasto e aos amigos, um agradecimento especial por compreenderem a importância desta fase da minha vida, pelo vosso interesse e apoio.

À minha família, em especial aos meus pais, por todo o carinho e suporte dado ao longo dos anos. Sem vós eu não seria quem sou nem poderia estar onde estou. Obrigado, obrigado e obrigado!

À Rita, por me ter ajudado a manter a sanidade mental durante esta fase da minha vida e ser o meu porto de abrigo. Mil vezes obrigado!

palavras-chave

simulações de dinâmica molecular, fibrose cística, recetores sintéticos, transporte transmembranar de aniões

resumo

O transporte de aniões através de membranas celulares, com recurso a canais proteicos, é fundamental em vários processos biológicos. Assim, o funcionamento deficiente desta complexa maquinaria celular está relacionado com o aparecimento de várias canalopatias como a Fibrose Cística (FC), associada ao transporte deficiente de cloreto através do canal CFTR. Os tratamentos atuais para esta doença apenas minoram os seus sintomas, sendo necessário desenvolver tratamentos alternativos, como, por exemplo, as terapias de substituição de canal. Este facto, ao longo das últimas décadas, tem motivado o desenvolvimento de moléculas capazes de procederem ao reconhecimento e, posteriormente, ao transporte passivo de aniões em vesículas lipídicas e em modelos celulares de FC. No entanto, uma molécula com atividade de transporte depende de um delicado equilíbrio entre sua a lipofilia e a afinidade para o anião. Neste contexto, esta tese situa-se na interface entre Química supramolecular, medicinal e computacional.

Os estudos teóricos que aqui se reportam consistiram em cálculos de mecânica quântica conjugados com simulações de dinâmica molecular baseados em campos de forças clássicos. Foram investigadas duas séries de moléculas trípodes (capítulos II e V), uma de transportadores derivados de decalina (capítulo III), e quatro séries de esquaramidas (capítulo IV). Os resultados estruturais e energéticos obtidos contribuíram para compreender, ao nível atómico, a interação destas moléculas com modelos de membranas, bem como dos respetivos mecanismos de transporte de aniões.

As simulações de dinâmica molecular de difusão passiva foram realizadas com os complexos de cloreto colocados na fase aquosa ou dentro da bicamada fosfolipídica, permitindo o estudo da capacidade de um transportador permear a interface água/lípido e de se difundir no meio altamente empacotado da bicamada. Por outro lado, a avaliação da distribuição do potencial electroestático na superfície eletrónica das moléculas correlaciona-se com as suas constantes de associação com aniões. Nos capítulos IV e V descrevem-se também simulações de dinâmica molecular constrangidas, realizadas com esquaramidas lineares e derivados trípodes fluorinados, respetivamente. Estas simulações permitiram a estimativa das barreiras de energia livre associadas à difusão destas duas séries de moléculas através do modelo de membrana por reconstrução do potencial de força média ao longo da normal à bicamada. As barreiras energéticas para estas duas séries de moléculas são consistentes com os dados experimentais de transporte e lipofilia. Adicionalmente, uma simulação reportada no capítulo V mostra, pela primeira vez, um transportador neutro a facilitar o transporte de cloreto através de uma bicamada fosfolipídica.

No âmbito da Química supramolecular, no capítulo VI reporta-se o desenvolvimento de parâmetros de campo de forças para ligações de calcogénio. Estas ligações, tal como as ligações de halogénio, também permitem o reconhecimento e transporte transmembranar de aniões, surgindo como potenciais alternativas às ligações de hidrogénio convencionais estudadas nos capítulos anteriores.

keywords

molecular dynamics simulations, cystic fibrosis, synthetic receptors, anion transmembrane transport

abstract

Ion transport across cell membranes, *via* protein membrane channels, is crucial in several biological processes. Thus, the malfunctioning of this complex cellular machinery is linked with several channelopathies, such as cystic fibrosis (CF), associated with the deficient chloride transport through the CFTR channel. As present treatments only aim to manage the disease's symptoms, alternative treatments are needed, such as channel replacement therapies. Over the last decades, this fact has motivated the development of synthetic anion transporters able to recognise and later promote the passive anion transport in lipid vesicles or even in CF cell models. However, the design of small drug-like transporters is still not straightforward and depends on an intricate equilibrium between the transporters' binding affinity and lipophilicity. In this context, this thesis interfaces between the supramolecular, medicinal and computational fields of Chemistry.

The theoretical investigations reported in this thesis consisted on quantum calculations together with molecular dynamics simulations based on classical force fields. The following series of molecules were investigated: two series of tripodal molecules (chapters II and V), a series of decalin-based transporters (chapter III) and four series of squaramide-based compounds (chapter IV). The structural and energetic insights allowed to understand, at the atomistic level, the interaction of synthetic molecules with membrane models as well as the anion transport mechanisms.

The molecular dynamics simulations of passive diffusion were carried out with chloride complexes placed either in the water phase or inside the phospholipid bilayer, allowing the investigation of the transporters' ability to permeate the water/lipid interface and to diffuse within the highly packed bilayer medium. Moreover, the assessment of the electrostatic surface potential of the transporters yielded insights that generally correlate well with anion binding constants. In chapters IV and V, constrained molecular dynamics simulations with linear squaramides and fluorinated tripodal derivatives are reported, respectively. These simulations allowed to estimate the free energy barriers associated with the translocation of these two series of molecules across the membrane model, with the reconstruction of the potential of mean force along the bilayer normal. The energetic barriers assessed for both series of molecules agree well with their lipophilicities and experimental anion transport data. Furthermore, a simulation reported in chapter V shows, for the first time, a neutral transporter facilitating the translocation of chloride across a phospholipid bilayer.

Within the scope of supramolecular Chemistry, chapter VI reports the development of force field parameters for chalcogen bonding interactions. These bonding interaction, as well as halogen bonds, can also be used for the recognition and transmembrane transport of anions, becoming potential alternatives to the ubiquitous hydrogen bonds studied in the previous chapters.

Ultimately, you want to do something, even if it's not perfect.
As long as you make clear what the uncertainties are and
don't abuse the data to make grand claims, you'll probably be alright.

-- James C. Gumbart, 2016

Index

Figures Index

Tables Index

Schemes Index

Abbreviations

Item	*Definition*
[*I*]	Specific initial rate
^{1}H NMR	Proton Nuclear Magnetic Resonance
AA	All-atom
ADP	Adenosine DiPhosphate
AM1-BCC	Austin Model 1 – Bond Charge Correction
AMBER	Assisted Model Building with Energy Refinement
Apl	Area per lipid
ASL	Airway Surface Liquid
ATP	Adenosine TriPhosphate
Avg	Average
B3LYP	Becke 3-parameter hybrid functional combined with Lee–Yang–Parr correlation functional
Bt_1	Bilayer thickness measured between phosphorus atoms' average positions along the *z*-dimension
Bt_2	Bilayer thickness measured from the phosphorus peaks in electron density profiles
CCDC	Cambridge Crystallographic Data Centre
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CGenFF	CHARMM General Force Field
CHARMM	Chemistry at HARvard Macromolecular Mechanics
ChB	Chalcogen Bond(s)
CLCNKB	Chloride Voltage-Gated Channel Kb
*c*log*P*	BioByte's proprietary method to calculate log*P*
CNS	Central Nervous System
COM	Centre(s) of mass
CUDA	Compute Unified Device Architecture
D_2O	Deuterium oxide, heavy water
DFT	Density Functional Theory
D_{HH}	Bilayer thickness measured from the system's peaks in electron density profiles
DLiPC	1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine
DLPC	1,2-dilauroyl-*sn*-glycero-3-phosphocholine
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DMSO	Dimethyl sulfoxide
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
E^2	2nd Order Perturbation Theory Analysis
$EC_{50,270s}$	Concentration (mol% carrier to lipid) needed to obtain 50% anion efflux after 270 s
EMSL	Environmental Molecular Sciences Laboratory
ESP	Electrostatic Potential
FRT	Fisher Rat Thyroid
GABA	γ-Aminobutyric acid
GAFF	General AMBER Force Field
GLUT	Family of integral membrane proteins that mediate the transmembrane transport of monosaccharides such as glucose, fructose, or galactose
GPU	Graphics Processing Unit
Gr	Gramicidin
GROMOS	GROningen MOlecular Simulation
HB	Hydrogen Bond(s)
HF	Hartree–Fock molecular orbital theory
HPLC	High-Performance Liquid Chromatography
K_a	Affinity Constant
k_{ini}	Initial rate
LA	Lauroyl
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator

Item	*Definition*
LINCS	Linear Constraint Solver
log(*k'*)	Logarithm of retention factor *k* (in column chromatography)
log*P*	Logarithm of partition coefficient *P* between water and octanol
M06-2X	Minnesota 2006 density functional with 54% Hartree–Fock exchange
MARTINI	MARrink Toolkit INItiative
MD	Molecular Dynamics
MetaD	MetaDynamics
MLE	Maximum Likelihood Estimation
MM	Molecular Mechanics
MP2	second order Møller–Plesset perturbation theory
MY	Myristoyl
NAMD	Not (just) Another Molecular Dynamics program
NBO	Natural Bond Orbitals
NMR	Nuclear Magnetic Resonance
NPT	Isothermal-isobaric ensemble, with constant number of particles *N*, constant pressure *P* and constant temperature *T*
NPγT	Isothermal-isobaric ensemble, with constant number of particles *N*, constant pressure *P*, constant temperature *T* and γ lateral surface tension
NVE	Microcanonical ensemble, with constant number of particles *N*, constant volume *V* and constant energy *E*
NVT	Canonical ensemble, with constant number of particles *N*, constant volume *V* and constant temperature *T*
OA	Oleate Anion
OL	Oleoyl
OPLS	Optimized Potential for Liquid Simulations
PA	Palmitoyl
PBC	Periodic Boundary Conditions
PC	phosphatidylcholine
PE	phosphatidylethanolamine
PEPC	1-palmitoyl-2-elaidoyl-*sn*-glycero-3-phosphocholine
PME	Particle Mesh Ewald
PMEMD	Particle Mesh Ewald Molecular Dynamics
PMF	Potential of Mean Force
POPC	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine
PyMOL	Python-enhanced molecular graphics tool
R^2	Coefficient of determination
REMD	Replica–Exchange Molecular Dynamics
RESP	Restrained Electrostatic Potential
RMSD	Root-Mean-Square Deviation
S_{CD}	Deuterium order parameters
SD	Standard Deviation
SETTLE	Analytical algorithm to reset the positions and velocities to satisfy the holonomic constraints on the rigid water model
SHAKE	Analytical algorithm to impose constraints on internal coordinates such as bond lengths and bond angles
SMD	Steered Molecular Dynamics
sn-1	stereospecific number-1
sn-2	stereospecific number-2
SYRAH	Coarse-grained force field for biomolecules
TBA+	Tetrabutylammonium
TIP3P	Transferable Intermolecular Potential with 3 Points
tren	tris(2-aminoethyl)amine
UA	United atom
UFF	Universal Force Field
US	Umbrella Sampling
$V(\mathbf{r})$	Electrostatic Potential

Item	*Definition*
vdW	van der Waals
vFEP	variational Free Energy Profile
$V_S(\mathbf{r})$	Electrostatic Potential computed on a molecular surface
$V_{S,max}$	Most positive value of electrostatic potential on a molecular surface
$V_{S,min}$	Most negative value of electrostatic potential on a molecular surface
WBI	Wiberg Bond Indices
WHAM	Weighted Histogram Analysis Method
XB	Halogen Bond(s)
YFP-H148Q/I152L	I^--sensitive and genetically encodable yellow fluorescent protein
$\rho(\mathbf{r})$	Electron density at r

Chapter I.
Introduction

Summary

This chapter starts by focusing on the cell membrane constitution and its role as trafficking mediator between the inside of the cell and its surroundings. Afterwards, channelopathies – diseases derived from defective ion transmembrane transport through ionic channels – will be focused, with special interest on diseases related with chloride transport. This outlook will be followed by a brief take on the different synthetic transporters developed as potential channel replacement therapies. On the other hand, this chapter will also give a brief description of the basic concepts associated with classical force field Molecular Dynamics (MD) simulations of membrane systems, to allow for an easier understanding of the work reported in the subsequent chapters.

I.1. Ion transmembrane transport and channelopathies

I.1.1. Brief characterisation of biological membranes

The cell membrane defines the limits of the cell and allows it to concentrate its nutrients and synthesised products inside, while maintaining its by-products outside (see Figure I-1).[1, 2] For instance, potassium concentrations are higher in the cytosol, while outside the cell, the ions with higher concentration are sodium, calcium and chloride, to allow the normal cell functioning.[2, 3] The role of the cell membrane is heavily dependent on its composition: a phospholipid bilayer with embedded proteins (see Figure I-2), which control the in and out movement of substances (*vide infra*).[1, 2] Moreover, the cell membranes components are arranged in agreement with the fluid mosaic model, a two-dimensional solution of membrane proteins.[4] The composition and functions of the cell membrane are extensible to the organelles of the cell (*e.g.*, nucleus, lysosomes, endoplasmatic reticulum, mitochondria).[3]

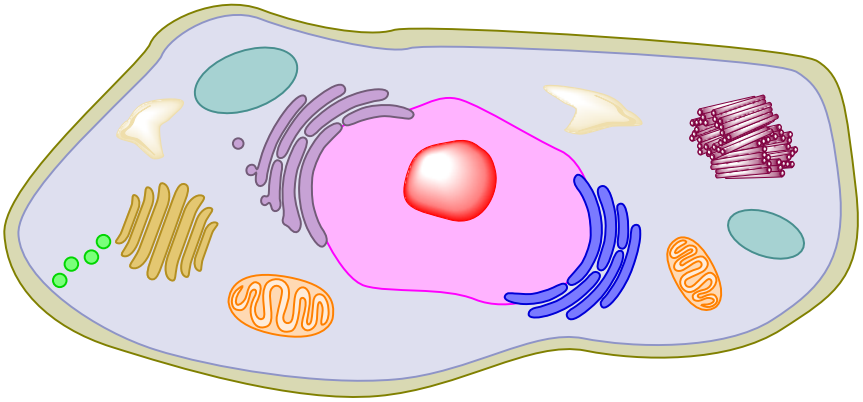


Figure I-1. Schematic representation of the cell, with the membrane in dark green, nucleus in pink, nucleolus in red, mitochondria in orange, smooth endoplasmic reticulum in purple, rough endoplasmic reticulum in blue, Golgi apparatus in light brown, vacuoles in white, lysosomes in teal, centriole in maroon and secretory vesicles in light green.

While the phospholipid molecules allow the self-assembly of the cell membrane and filter most of the water-soluble substances, the specialised membrane proteins facilitate the regulated transport of several solutes, and also allow for cell signalling.[1, 2] The lipid bilayer is composed of phospholipids, sphingolipids, and steroids (cholesterol and its derivatives).[2] Phospholipids, represented by the orange and grey structures in Figure I-2 are amphipathic molecules and are organised with their hydrophobic chains in the bilayer's core, while their hydrophilic heads are in the exterior of the bilayer, in contact with either the aqueous interstitial fluid or cytosol.[1, 2] These amphipathic molecules self-assemble as bilayers or micelles, due to van der Waals interactions between the phospholipid aliphatic tails and electrostatic interactions between the phospholipid head groups.[2] Cholesterol is evenly distributed between the inner and outer leaflets and the remaining bilayer components are distributed asymmetrically, as most of the sphingolipids and phosphatidylcholine (PC) lipids are found in the outside layer, whereas the phosphatidylethanolamine (PE) lipids are typically present in the inner leaflet.[1, 2] While some proteins associated with the membrane present some motility, diffusing along the surface of the membrane, other proteins have little or no motility at all, being fixed to the cell cytoskeleton or to an outer matrix. There are three types of proteins associated to the membrane:

- *Integral proteins* (or *intrinsic proteins*), which are embedded in the membrane and span it from side-to-side. The segments that cross the bilayer are arranged in either α-helices or β-sheets, typically composed of amino acids with nonpolar side-chains, while the cytosolic and exoplasmic portions have

amino acids with hydrophilic substituents. Ion channels are a typical example of this type of proteins, which are depicted by the magenta sketch in Figure I-2;[2, 3]

- *Peripheral proteins* (or *extrinsic proteins*), which are present in only one of the bilayer's leaflets, facing either the cytosol or the interstitial fluid. They can interact with the integral or lipid-anchored membrane proteins as well as with the lipid head groups. Internal peripheral proteins are often associated with the cytoskeleton, while the external ones are associated with the extracellular matrix. The blue and green structures in Figure I-2 represent, respectively, peripheral proteins present in the outside and in the inside leaflets of the cell membrane;[2, 3]
- *Lipid-anchored membrane proteins*, which are covalently bonded to one or more phospholipid molecules that anchor the protein to the membrane, as the hydrophobic moiety of the lipid is effectively embedded in a leaflet (yellow structure in Figure I-2). On the other hand, the peptide portion itself does not enter the bilayer (brown structure in Figure I-2) and can be attached or detached from the lipid portion, thus modulating the activity of the protein. Currently, there are four different lipid-anchoring motifs: amide-linked myristoyl, thioester-linked fatty acyl, thioether-linked prenyl, and amide-linked glycosyl phosphatidylinositol anchors.[2, 5]

Moreover, in the extracellular face of the membrane, most of these proteins are covalently bound to carbohydrate molecules, being designated glycoproteins, shown as the red structure in Figure I-2.[1]

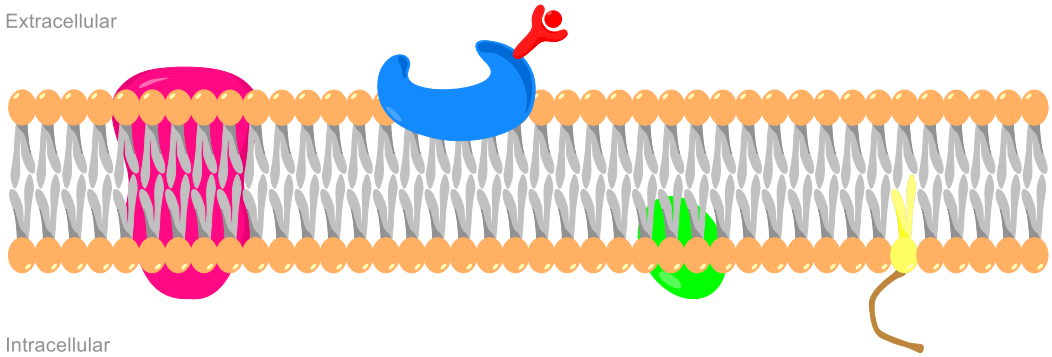


Figure I-2. Schematics of the bilayer structure of the cell membrane, with phospholipid molecules in orange and grey, an integral protein in magenta, a peripheric protein in green, a peripheric glycoprotein in blue with its carbohydrate moiety in red, and lipid-anchored protein in brown with its lipid anchor in yellow.

I.1.2. Transport across the membrane

Gases such as O_2 and CO_2, hydrophobic molecules like steroid hormones, and small uncharged polar molecules (*e.g.*, water and glycerol) can passively diffuse across cellular membranes without assistance, down their concentration gradient. The rate of diffusion of these molecules is dependent on their concentration, size and hydrophobicity. Moreover, passive diffusion only allows transport rates that are lower than what is needed to fulfil the cell's needs.[1, 2] On the other hand, the diffusion of charged molecules, regardless of their size, is energetically hindered by their charge and degree of hydration.[1] Therefore, charged entities depend on membrane proteins for their translocation,[5] and similar channels and transporters also exist to facilitate the transmembrane transport of neutral species.[1, 2] For instance, the GLUT family of membrane transporters facilitates the diffusion of monosaccharides (*e.g.*, glucose, fructose, galactose) down their concentration gradients at greater rates than passive diffusion allows.[2, 6, 7]

Three classes of membrane proteins facilitate the diffusion of substances across the bilayer, including channels, transporters and ATP-powered pumps. Channels facilitate the diffusion of water, some ions or hydrophilic molecules across the bilayer and down their concentration gradient, *i.e.*, without spending any energy.[2] These membrane proteins form continuous pores that span the membrane and interact poorly with the transported substance, allowing a great transport rate.[1] They can be open at all times (nongated channels) or close in response to electric or chemical stimuli (gated channels).[2] On the other hand, transporters span the membrane, but they experience conformational changes that lead to the alternative exposure of binding sites on either side of the phospholipid bilayer,[1] resulting in transport rates lower than channels.[2] Within this class of membrane proteins, there are uniporters, which move substances down their concentration gradient, similarly to channels,[2] and coupled transporters, which also move solutes down their concentration gradient, concomitantly with another species moving against its concentration gradient. Therefore, when the movement of the second solute is in the same direction, it is called a symporter or co-transporter, whereas when the movement of the second solute is in the opposite direction, that protein is called an exchanger or antiporter.[1] Alternatively, the transmembrane transport promoted by the expense of metabolic energy, obtained from the ATP hydrolysis to ADP in the ATP-powered pumps, allows the movement of solutes against their electric potential and/or chemical concentration gradient.[3] This process is called active transport.[6] These transmembrane transport mechanisms are illustrated in Figure I-3.

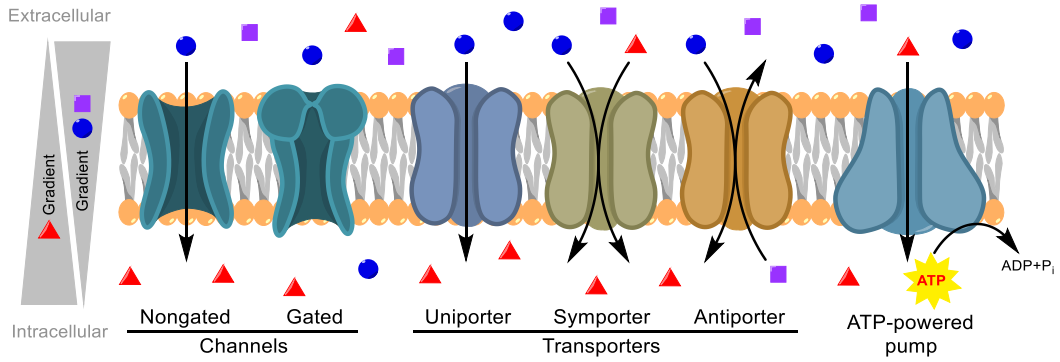


Figure I-3. Schematic representation of transport mechanisms assisted by membrane proteins.

Beyond these naturally occurring mechanisms, the transport of substances across the bilayer can also rely on smaller shuttle-like peptides that act as mobile carriers, such as valinomycin, a natural antibiotic that functions as an ion carrier. The proposed mechanism is as follows: the mobile carrier captures the ion at the water/lipid interface and then ferries it to the opposite interface, where it releases the guest ion.[8] This mechanism is further explored on section I.2. and is illustrated in Figure I-4 (*vide infra*).

I.1.3. Channelopathies

Normal functioning of ion channels

Ion channels, transmembrane proteins that span the cell membrane, carry out several functions in the cells and in the organism, such as electric excitability of muscular cells and neurons, pH, volume and osmotic regulation of the cell, transepithelial transport (gut, kidney, lung), hormone secretion (adrenals, gonads, pituitary, pancreas), and sensory transduction.[9-13]. Moreover, some functions of the cell membrane's ion channels are shared with the ones present in the membrane of organelles.[9]

Ion channels incorporate several polypeptide subunits that form a hydrophilic pore, with a 3.0 to 6.5 Å diameter, through which partially hydrated ions can move from one side of the membrane to the other.[11, 12] The most common ion channels are for Na^+, K^+, Ca^{2+}, H^+, and Cl^-.[14] Due to their homeostatic importance, both at cellular, tissue and organism level, ion channels have become a common drug target.[15] Based on this rationale, the following types of drugs have ion channels as therapeutic targets: antiepileptic drugs, local anaesthetics, migraine treatments, antipsychotics, sedatives, anti-anxiety agents, antiarrhythmics, antihypertensives, oral hypoglycaemics, or antiviral agents.[11, 12]

Ion channels dysfunctioning

Diseases associated with channel dysfunctioning are designated channelopathies, and may be attributed to toxic, autoimmune, or genetic causes.[15] Acquired channelopathies result from the interaction of venoms or antibodies with membrane receptors for neurotransmissors or ions, blocking or modulating the channels' activity, which leads to the presentation of signs or symptoms.[12] On the other hand, in inherited channelopathies the symptoms present themselves intermittently,[15] and can be a result of the loss or gain of channel function.[10] While the loss of function mutations in channels are often associated with recessive diseases (mutations in the Cystic Fibrosis Transmembrane Conductance Regulator, *CFTR*, or chloride channel Kb, *CLCNKB*, genes lead to Cystic Fibrosis (CF) or Barter's disease, respectively), the gain of function mutations are typically linked with dominant diseases (mutations in voltage dependent Na^+ channels lead to paramyotonia, cardiac arrhythmia and epilepsy).[10, 11] Moreover, different mutations can occur in the same gene, leading to dominant or recessive diseases.[11]

The first inherited channelopathies were identified in disorders of skeletal muscle excitability (myotonia and periodic paralysis),[12] and the list of affected tissues, organs and systems keeps expanding, exceeding more than 40 diseases.[10, 12] The ion channelopathies that particularly affect the central nervous system (CNS) and skeletal muscles are characterised by: *a)* sporadic recurrent manifestations of symptoms (paroxysms), with normal function between occurrences; *b)* being inherited mostly as autosomal dominant features; and *c)* typically affecting only a single organ, as a gene related with muscle excitability is expressed in skeletal muscle cells, no CNS symptoms manifest.[12] Beyond the skeletal muscles and the CNS, ion channelopathies also commonly affect the urinary, endocrine, cardiovascular, respiratory, and immune systems, the bones, as well as the eyes and ears, with disorders that may involve more than one organ or system.[10, 11, 16] Within the anion channelopathies, deficient chloride transmembrane transport is of paramount importance, as detailed below.

The chloride ion is essential for the maintenance of the membrane potential, transepithelial transport, regulation of the pH and volume of the cell or its organelles.[9] Moreover, besides being the most abundant physiological anion, along with phosphate and bicarbonate, it is also the most lipophilic one.[17] The best well-characterised chloride ion channels are the voltage-gated chloride channels, the CFTR, and the ligand-gated chloride channels operated by GABA and glycine.[9]

Several chloride related channelopathies were discovered over the last 25 years.[13] Some examples of chloride channelopathies are juvenile myoclonic epilepsy, startle disease (a neurologic disorder), deafness, blindness, lysosomal storage and degradation, Bartter's syndrome types III and IV (severe renal salt loss), Dent's disease (damage to the proximal tubules in the kidneys, leading to renal failure), nephrolithiasis (kidney stones), myotonia congenita and generalised myotonia (muscle stiffness), osteopetrosis (bone disease), pulmonary infections and fibrosis (CF), pancreas and liver afflictions (causing low nutrient absorption and bile duct blockage), and male infertility.[9, 10, 13]

CF is one of the most frequent inherited autosomal recessive channelopathies, with about one case per 3000 births in European or North American Caucasians.[18-20] Patients with this condition have a life expectancy between 30 to 40 years,[13] with problems such as chronic pulmonary infections (which

result in lung destruction, right heart insufficiency, and heart failure), fibrosis and atrophy of the pancreas or hepatobiliary complications (including biliary cirrhosis, portal hypertension or even hepatic failure).[13, 21] Moreover, over 95% of the male patients suffering from this condition are infertile due to the congenital bilateral absence of the vas deferens.[10, 13]

The *CFTR* gene encodes the CFTR transmembrane protein, which acts as a chloride channel as well as the regulator for other ion channels.[21] When the CFTR channel is defective, the NaCl and water secretion in the airways is reduced, with simultaneous NaCl absorption enhancement. In the lungs, this combination of events leads to the reduction of the Airway Surface Liquid (ASL), precluding the airway cleaning process, which allows pathogenic bacteria to colonise the airways, finally resulting in infection and destruction of the airways.[13, 21] Still, not all CTFR mutations lead to a similar phenotype, *i.e.*, gravity of the disease.[9] Mutations of the *CFTR* can be divided in five classes:

a) defective protein synthesis – nonsense mutations, deletions or insertions lead to premature truncation of the protein synthesis;[9]

b) defective processing – the trafficking of the protein to the cell membrane is affected, leading to its low expression on the cell surface;[9]

c) defective regulation – the regulation pathways of the channel are defective, so the channel cannot be activated;[9, 22]

d) impaired conduction – the conduction properties of the channel are diminished, limiting the rate of chloride transport;[9, 22]

e) reduced synthesis and/or trafficking – mutations in the promoter region of the CTFR gene and amino acid substitutions lead to the underproduction or defective production of normal CFTR channels.[22]

As no cure is currently available, the therapies for CF focus mainly in the management of the disease and in the treatment of the associated comorbidities,[9] while in the more severe cases, lung transplantation is needed.[23] Several areas of intervention are possible in CF, such as CFTR modulation, restoral of ASL, mucus alteration, nutrition, anti-inflammatory agents and anti-infective drugs as well as approaches that improve survival after lung transplantation.[24] Given the genetic nature of CF, genetic therapies are also envisaged,[9, 24] although limited by technical problems that hinder the efficiency of healthy gene transfection.[17] Alternatively, channel replacement therapies have been motivating the active search of small drug-like molecules to promote the anion transport, as summarised in the next section.

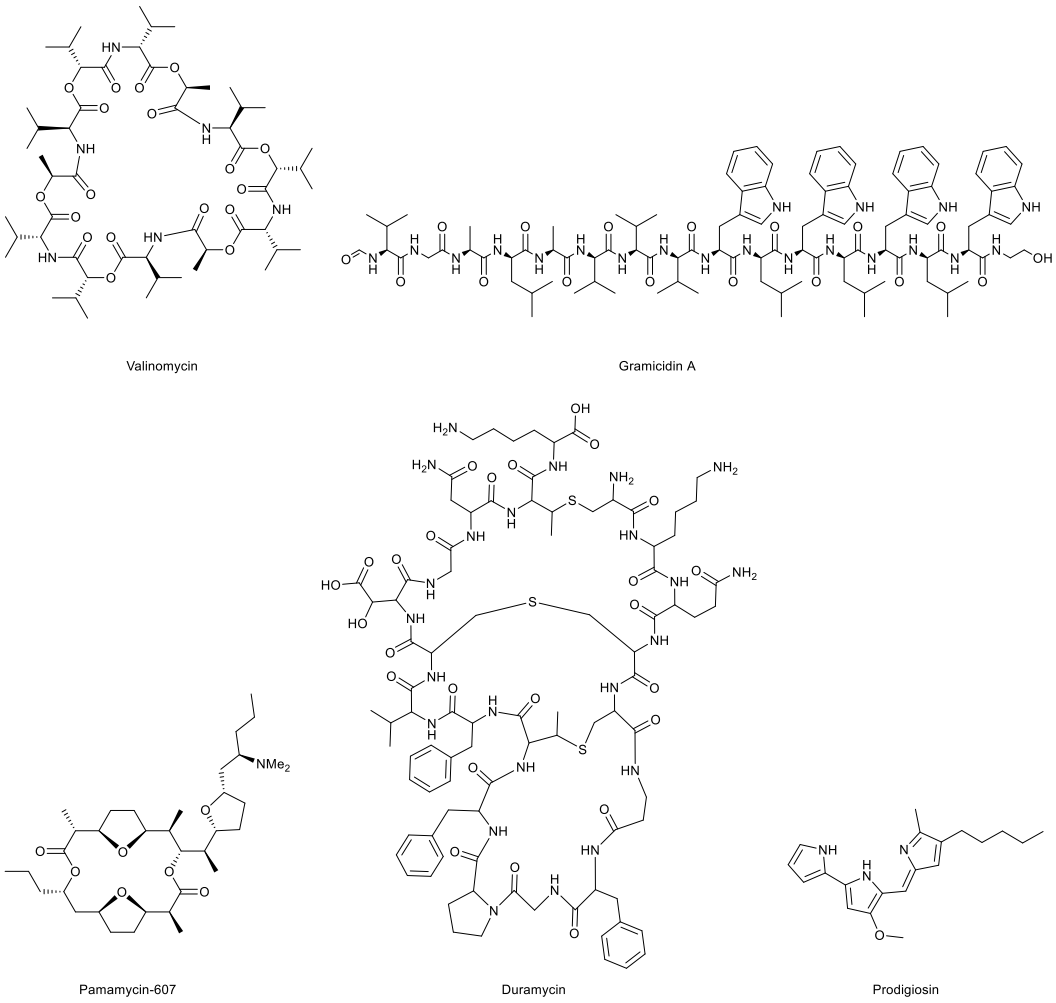
I.2. Synthetic ionophores

I.2.1. Outline

Over the last decades, inspired by the natural ionophores, a part of the supramolecular Chemistry community has been devoted to the design and development of synthetic receptors with the potential for ion transmembrane transport, *i.e.*, ionophores.[25] An ionophore is commonly defined as "a compound which can carry specific ions through membranes of cells or organelles".[26] Therefore, ionophores are small molecules able to carry ions or to facilitate their diffusion through a membrane, using a wide range of mechanisms (*vide infra* – carrier, channel, or relay). These compounds are naturally occurring as peptides, cyclic depsipeptides, macrotetrolides, and polyether antibiotics,[27, 28] but can also be designed and prepared in laboratory, being labelled as synthetic ionophores.[27-29]

While cation ionophores are quite frequent (*e.g.*, valinomycin and gramicidin A for K^+, sketched in Scheme I-1),[17] the natural occurring anion ionophores are quite rare. Natural occurring anion

receptors include pamamycins, macrodiolides with antibiotic activity;[30] duramycins, peptides that form weak selective anion channels;[31] and prodigiosins, which act as H^+/Cl^- symporters,[32] and also function as anion and proton transmembrane transporters.[17, 33] All these anion transporters are also sketched in Scheme I-1.



Scheme I-1. Structures of natural occurring cation ionophores (top) and anion ionophores (bottom).

The reduced number of natural anion ionophores has been a strong driving force in the development of synthetic anion receptors, especially amphiphilic ones, which may have the potential to be applied in the transport of anions through the phospholipid bilayer.[34]

In CF, the replacement of the malfunctioning ion channels with peptide-based ionophores has recently been proposed as a therapeutic alternative.[35] On the other hand, synthetic anion ionophores might also be applied as channel replacement therapeutics in CF and other channelopathies.[34] To develop a potential pharmaceutical application with an anion ionophore, the administration routes must also be considered: an anion ionophore applied to CF therapy should be amphiphilic: sufficiently hydrophilic for delivery in the lungs, and lipophilic enough to be internalised in the target epithelial membranes.[17, 36] Moreover, the pharmacokinetics and safety profiles of such drug-like molecules should not be overlooked.[37]

The hallmarks of the research in the field of synthetic ionophores are gathered in Table I-1. The development of synthetic ionophores was initially based on the modification of natural transporters, to

improve their affinity with anionic guests or their solubility properties.[17] Current work in this area relies in the synthesis of anion ionophores from scratch, with no natural components.[17]

Table I-1. Summary of advances in synthetic ionophores.

Year	Hallmark
1967	Synthesis of the first synthetic ion carrier[38]
1968	First report of chloride recognition and complexation by a macrocycle[39]
1978	First attempts to synthesise a natural ionophore[40, 41]
1982	Synthesis of the first synthetic ion channel[42, 43]
1998	First report of a crystal structure of a biological ion channel[44]
2005	Understanding of conductivity and selectivity based on architecture and chemistry of pathways[45]

I.2.2. Development of synthetic anion ionophores

This section will focus on some examples of synthetic anion ionophores, with emphasis on the types of transport mechanisms that those molecules employ as well as the types of recognition mechanisms that they present. Comprehensive reviews on the recent advances in synthetic transmembrane anion transporters can be found on refs. 34, 46-71 that illustrate the increasing number and diversity of synthetic receptors. These reviews also focus on other types of anion receptors with potential to be applied as health-related applications.

Ion transmembrane transport with resort to synthetic molecules mimics the natural mechanisms of ion carriers and ion channels, as depicted in Figure I-4.

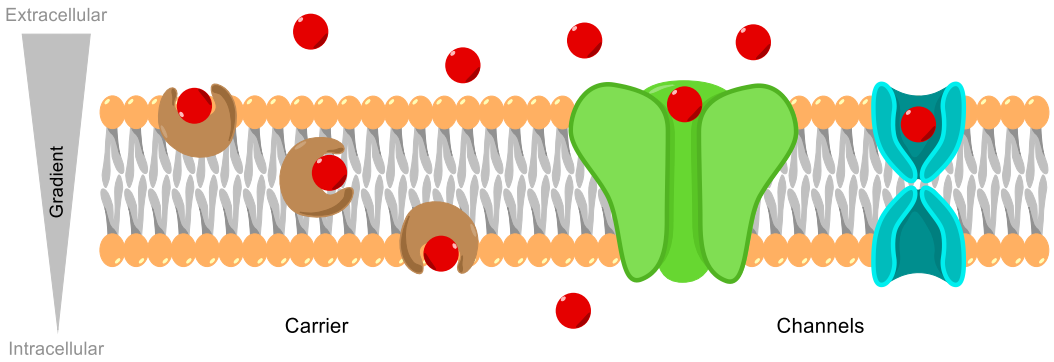


Figure I-4. Schematic representation of the carrier and channel mechanisms employed by the synthetic transporters.

Similarly to valinomycin, synthetic molecules that form receptor–ion complexes, recognising the ion and ferrying it in or out of the cell, are designated as carriers (brown structures in Figure I-4). The exterior of a carrier is typically hydrophobic and this mechanism is highly selective, but presents a low rate of transport.[17, 57] On the other hand, the stationary ion channels span the membrane, forming static polar pores that allow the in or out movement of anions, as well as some cations, at high rates.[17, 57] These channels may rely on one unit or on several self-assembled subunits and are respectively represented by the green and the blue structures in Figure I-4.[57] A third way for synthetic anion ionophores to promote the transmembrane transport is *via* a relay mechanism, in which the receptor molecules are aggregated inside the membrane and the anion is passed on to the receptors on the other side, as depicted in Figure I-5.[36] These mechanisms are exemplified below.

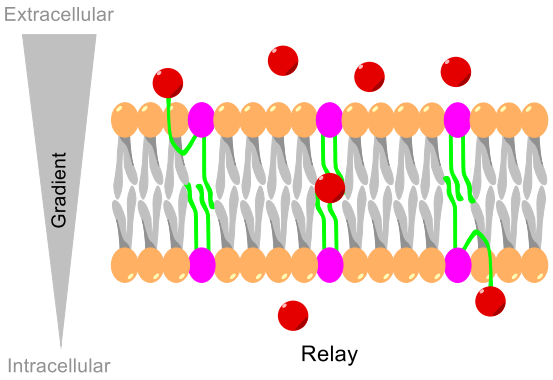
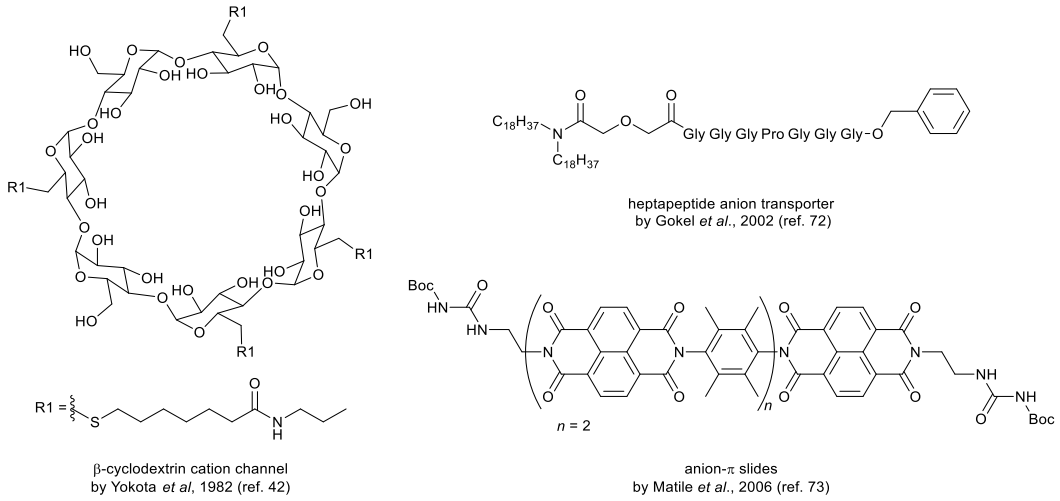


Figure I-5. Representation of the relay mechanism proposed by Smith and co-workers in ref. 36.

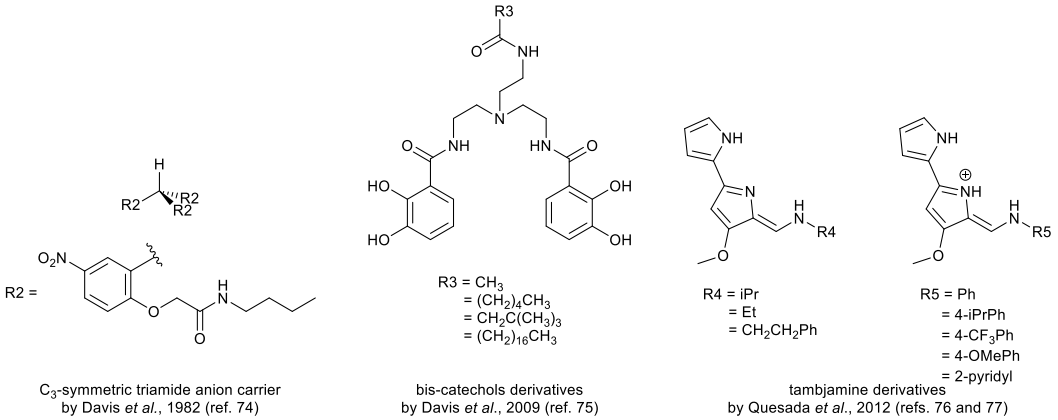
The first man-made cation channel, reported in 1982 by Yokota *et al*. (see Scheme I-2),[42] consisted of two β-cyclodextrins with four hydrophobic tails and three metal binding sites. Two β-cyclodextrins units, contacting via their hydrophobic tails, established a channel that allowed the high flow rate of Cu^{2+} and Co^{2+}.[42] Poly-glycine-based receptors for chloride transport (see Scheme I-2) were developed by Gokel *et al.*, based on the structural insights obtained from the crystal structures of transmembrane chloride channels. These peptides are anchored to the bilayer *via* interactions with the phospholipids acyl chains and form self-assembled pores that span the membrane. The presence of these synthetic channels allows the efflux of chloride from the phospholipid vesicles.[72] Matile *et al.* have synthesised oligonaphthalenediimide rods that span the membrane and form anion−π slides (see Scheme I-2). The anions are transported *via* a multi-ion hopping system, relying on several clustered receptors: as an anion interacts with a sub-unit, the previously anion is pushed to the next sub-unit, hopping its way through the membrane.[73]



Scheme I-2. Examples of synthetic ion transporters based on the channel mechanism.

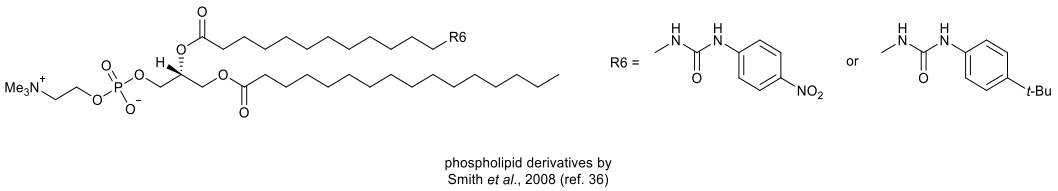
Davis *et al.* developed a C_3-symmetric triamide carrier (see Scheme I-3) that selectively promotes the efflux of nitrate over chloride. This transmembrane transport was promoted as a H^+/NO_3^- coupled-transport, effectively affecting the intravesicular pH of the studied systems.[74] Later, he reported a series of bis-catechols (see Scheme I-3) that promote the efflux of halides in agreement with the Hofmeister series. Although the lipophilicity of the amphiphilic transporters was related to their transport activity, the hydration energies of the anions were crucial to determine the transport efficiency.[75] Tambjamine derivatives (also sketched in Scheme I-3) are analogous to the natural

prodiginines and also present the ability to promote chloride transmembrane transport.[76] Quesada *et al.* have demonstrated for the first time that, beyond promoting the bicarbonate/chloride exchange in lipid vesicles at low concentrations, tambjamine derivatives are also able to promote the apoptosis of different cancer cell lines by changing their intracellular pH levels.[77]



Scheme I-3. Examples of synthetic anion transporters based on the carrier mechanism.

The relay mechanism was only observed with the phospholipid derivatives incorporating an urea binding group developed by Smith *et al.* (see Scheme I-4).[36] These molecules are compatible with the membrane environment due to the structural similarity with the phospholipids. However, upon the thickening of the bilayer, due to mismatch with the receptors length and inability of the aggregated molecules to span the bilayer, the transport rates decrease.[36]



Scheme I-4. Example of a synthetic anion transporters based on the relay mechanism.

The synthetic anion ionophores reported above rely on hydrogen bonds (*via* amide, urea or thiourea binding units) or electrostatic interactions (on the anion-π slides) to recognise the anions. As further shown in this thesis, hydrogen bonds in thiourea and squaramide binding units are one of the most common ion recognition mechanism. However, the non-conventional halogen and chalcogen bonds are able to recognise and, subsequently, transport anions across phospholipid bilayers, as elegantly demonstrated by the group of Stefan Matile (see Figure I-6).[78, 79]

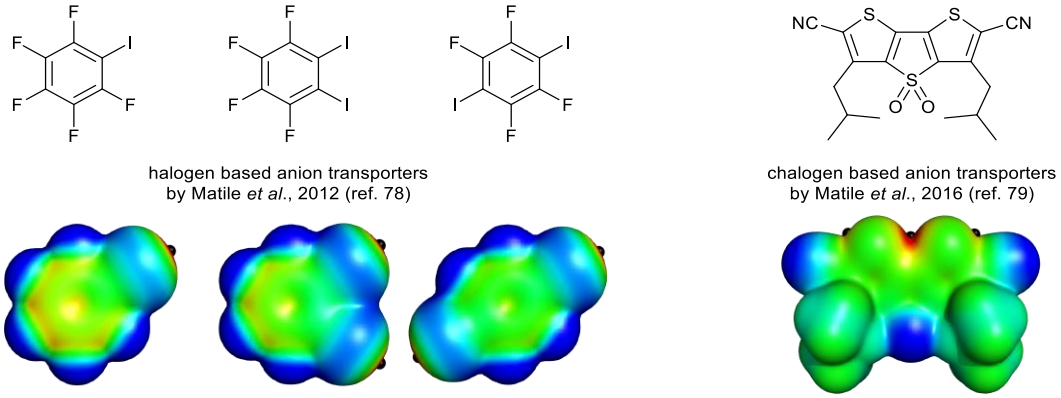


Figure I-6. Top: Chemical sketches of halogen- and chalcogen-based anion transporters developed by Matile *et al.*[78, 79] Bottom: Calculated electrostatic potentials on the 0.001 electrons Bohr^{-3} surfaces.* Colour ranges, in kcal mol^{-1}, for halogen-based molecules: red – greater than 25; blue – less than −5; and for the chalcogen-based molecule: red – greater than 45; blue – less than −25. The black dots correspond to the surface maxima near the iodine or sulfur atoms.

Halogen and chalcogen bonds are formed between anions that act as nucleophilic acceptors and a σ-hole – a highly localised positive-charge density on the surface of the halogen (chlorine, bromine or iodine) or chalcogen (sulfur, selenium or tellurium), as identified in Figure I-6. While only a σ-hole is present on the surface of the halogen atoms, on the opposite side to the covalent bond, two σ-holes are present on the surface of the chalcogen atoms, nearly opposite to their covalent bonds. Matile *et al.* have shown that the use of halogen bonds led to the size reduction of anion transporters, resulting in more hydrophobic compounds that do not need to have lipophilic groups as some hydrogen-bond-based transporters do.[78] On the other hand, the use of chalcogen-based transporters was assessed experimentally and corroborated by DFT calculations, opening new avenues in the anion transmembrane transport field with the use of these non-conventional recognition mechanisms.[79]

In this thesis, although the anion transport promoted by halogen or chalcogen bonds was not investigated, the anion recognition *via* chalcogen bonds is illustrated in Chapter VI. It comprises the development of classic force field parameters for anion recognition by a [2]rotaxane, with two mechanically interlocked molecules (*vide infra*).

The strength of binding is a key factor on how a receptor performs as a transmembrane transporter.[57, 80] The compound of interest must be able to reversibly associate with the anion to promote its transport, in agreement with the Goldilocks principle, where the transport activity follows a bell-shaped dependence on the anion binding affinity constants.[81, 82] In other words, a balance between strong binding constants (which prevent the release of the anion after the diffusion across the bilayer) and weak binding constants (which will preclude the uptake of the solvated anion) must be sought for each potential transmembrane transporter.

The same principle also applies to the lipophilic character of synthetic anion carriers, typically reflected by the logarithm of a molecule's octanol/water partition coefficient (log*P*). In a series of structurally related anion transporters, only by varying their lipophilicities, a too low or too high log*P* value is consistent with small to non-existent anion transport activity, while moderate values are associated with more pronounced anion effluxes.[83, 84]

A desirable synthetic anion transporter must, therefore, present a balance of lipophilicity and anion binding. In this context, molecular modelling methods have become increasingly useful in the

* These calculations were carried out as further detailed in section VI.2.2., Chapter VI.

development of synthetic anion transporters, with the prediction of several properties as well as the study of the interaction between anion carriers and membrane models.[83-91]

I.3. MD simulations of membranes

Within the scope of this thesis, MD simulations were a powerful tool to obtain structural and energetic insights at the atomistic level on the interactions between synthetic transporters and phospholipid bilayers. Thus, this section begins with a necessarily brief description of MD simulations based on classical force fields to contextualise the reader.

I.3.1. Basic concepts of MD simulations

In MD simulations, the movement of the system's particles is determined by solving Newton's equations of motion, while the forces between particles, as well as their potential energies, are estimated with classical Molecular Mechanics (MM) force fields.[92, 93]

MD simulations are a deterministic method, allowing to predict the configuration of a given system throughout time, while generating a trajectory (with all the positions of the system's particles along the simulation) that allows to assess how the system's properties evolve. In MD simulations, Newton's equations of motion are repeatedly integrated along the simulation time, generally every 1-10 femtoseconds − the duration of a "time step".[93] The time step choice will impact the quality of the MD simulation, as it should sample the fastest vibrations in the system (typically C–H, O–H, N–H bonds),[93-95] as well as its computational cost, as the forces on atoms are calculated and combined with the current coordinates and velocities to generate new coordinates, velocities and forces in the next time step.[93] Usually, to speed up MD simulations, the covalent bonds to hydrogen atoms are constrained using proper algorithms such as SHAKE,[96] SETTLE,[97] or LINCS,[98] that enable the use of a larger time step, typically 2.0 fs.[99]

Typically, MM force fields describe the systems of interest with the estimative of their bonded (bond lengths, bond angles and torsion angles) and nonbonded (van der Waals and electrostatic) interactions.[92, 100, 101] These interactions are estimated using a pre-determined functional form, or force field, to calculate the potential energy ($\mathcal{V}(\mathbf{r}^N)$) of the system, typically in the form presented in Eq. I-1.[92, 100, 101]

$$\mathcal{V}(\mathbf{r}^N) = \sum_{bonds} \frac{k_i}{2}\left(l_i - l_{i,0}\right)^2 + \sum_{angles} \frac{k_i}{2}\left(\theta_i - \theta_{i,0}\right)^2 + \sum_{torsions} \frac{V_n}{2}\left(1 + cos(n\omega - \gamma)\right) + \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right) \quad \text{Eq. I-1}$$

Thus, the potential energy of the system is a function of the positions ($\mathbf{r}$) of its N atoms or particles. The first two terms are modelled with resort to harmonic potentials and k_i stands for a force constant: the first one is the sum of the interactions between pairs of bonded atoms and increases in energy as the distance between atoms (l_i) deviates from the reference distance ($l_{i,0}$), while the second varies as the angle between three bonded atoms (θ_i) differs from the equilibrium value ($\theta_{i,0}$). In the third term, which assesses how the energy changes as a bond rotates, V_n is a force constant, n is the multiplicity of the torsion angle (ω) and γ is its phase angle. The final contribution in Eq. I-1 is for the non-bonded interactions, including the van der Waals and electrostatic interactions for the N atoms in the system that are separated by at least three bonds or in different molecules. The vdW interactions are modelled

with the following Lennard–Jones parameters for particles i and j: ε_{ij} is the potential well depth, σ_{ij} is the collision diameter at which the potential reaches zero, and r_{ij} is the distance between atoms i and j. The electrostatic interactions are described by Coulomb's law, with q_i and q_j standing for the charges of atoms i and j, while ε_0 is the permittivity in vaccum.[92, 100, 101]

Eq. I-1 is a general energy function followed by different atomistic force fields, apart of small variations or different functional forms that allow an adequate performance in the description of a given system, both in precision as well as in computational efficiency.[92] For instance, in the AMBER force fields the bond and angle force constants are not halved.[100, 102, 103] Moreover, even the force constants and equilibrium or reference values are the result of *a priori* choices. For instance, bond lengths and bond angles from other AMBER force fields, *ab initio* calculations and crystal structures were used as reference values in the initial version of the General AMBER Force Field (GAFF), while the corresponding force constants were empirically estimated.[102, 103] In the upcoming second generation of GAFF, most reference values will be a result of *ab initio* calculations or empirical values will be redeveloped to match *ab initio* high quality data.[104]

Another crucial characteristic of a force field is the "atom type" that is assigned to each atom of the system. This property allows to identify the element of each atom, its hybridisation state, and its local environment.[92] The different atom types will, most likely, lead to different force field parameters. Moreover, partial atomic charges are used to reproduce the distribution of the quantum electrostatic potential of a molecule.[92] These atomic charges can be derived from *ab initio* calculations, semi-empirical methods or empirical ones. For instance, the Restrained Electrostatic Potential (RESP) charges, used in the AMBER force fields, estimated at the HF/6-31G* theory level, following the Merz–Singh–Kollman scheme, typically using four concentric layers and six points per layer.[105] Alternatively, GAFF also allows the use of AM1–BCC charges, which consist on semi-empirical AM1 calculation to get Mulliken charges, followed by a bond charge correction.[106, 107] The Gasteiger–Marsili charges can be obtained empirically and only depend on the atoms connectivities.[108]

Three types of force fields exist: all-atom, united atom, and coarse-grained. All-atom force fields explicitly represent all the atoms in the system, including the nonpolar hydrogen atoms. On the other hand, united atom force fields explicitly represent polar hydrogen atoms with potential for hydrogen bonding, while nonpolar hydrogen atoms are united to their heavy atoms. Likewise, coarse-grained force fields do not represent all the atoms in a system, but use larger particles to represent elements (functional groups or other groups of atoms), being useful for larger systems and/or longer simulations. Along these three types of force fields, the performance increases with the loss of structural information.[109, 110] These types of force fields can also be combined in hybrids that take advantage of the atomistic level of detail for certain parts of the system while other parts, such as the solvents, are treated with a coarse graining approach, resulting in computational gains.[111]

Classical MD simulations can be carried out under the microcanonical or *NVE* ensemble (constant number of particles – *N*, volume – *V*, and energy – *E*), the canonical or *NVT* ensemble (constant *N*, *V* and temperature – *T*) or the isothermal-isobaric or *NPT* ensemble (constant *N*, *T* and pressure – *P*).[93] Moreover, in classical MD simulations, the total energy of a system comprises the potential energy of its particles, given by Eq. I-1, plus the kinetic energy associated with their motion. The kinetic energy $\mathcal{K}$ of the system is directly related to its temperature, as given in Eq. I-2:

$$\mathcal{K} = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m_i} = \frac{k_B T}{2}(3N - N_c) \qquad \text{Eq. I-2}$$

Where $\mathbf{p}_i$ is the total momentum of particle i, m_i is its mass, k_B is Boltzmann's constant and T corresponds to the system's temperature. The equipartition theorem establishes that each particle in an equilibrium system has an average kinetic energy of $k_BT/2$. In a system with N particles, each moving in the three dimensions, then the kinetic energy is given by $3Nk_BT/2$. N_c stands for the number of constrains, typically 3, enough to constraint the linear momentum of a particle to 0.[93]

The temperature of a simulated system and, therefore, its energy, can be maintained with resort to the Anderson,[112] Berendsen,[113] Nosé–Hoover thermostats,[114, 115] or to Langevin dynamics.[116] In the Anderson and Berendsen thermostats, the system is coupled to a heat bath (T_0) that imposes the desired temperature. This heat bath is represented by stochastic collisions, randomly occurring on selected particles, with the particles' velocities being periodically rescaled.[112, 113] In the Nosé–Hoover thermostat, an additional degree of freedom is added as an external system that acts over the explicit system. Additional virtual variables (coordinates, velocities and time) are also introduced and effectively affect the real ones, for instance, scaling the velocities of the system.[114, 115] In Langevin dynamics the system's particles are embedded within virtual particles, a friction bath. The collisions between the large real particles and smaller virtual ones determine how quickly the system exchanges energy with its environment, thus defining how the temperature fluctuates throughout time.[116]

The pressure in an MD simulation is typically calculated *via* the virial theorem of Clausius, *i.e.*, the sum of the products of the particles' coordinates and the respective forces acting on them.[93] The virial of the system, $\mathcal{W}$, is given by Eq. I-3:

$$\mathcal{W} = \sum_{\alpha=1}^{N} \boldsymbol{r}^{\alpha}\boldsymbol{f}^{\alpha} = -3Nk_BT \qquad \text{Eq. I-3}$$

Where $\boldsymbol{r}^{\alpha}$ and $\boldsymbol{f}^{\alpha}$ are the position of atom α and the forces acting on it, respectively.[117] In the case of an ideal gas, the only forces acting on the particles are those between the container and the gas and, therefore, considering equations Eq. I-3 and Eq. I-4, Eq. I-5 is obtained:

$$PV = Nk_BT \qquad \text{Eq. I-4}$$

$$\mathcal{W} = -3PV \qquad \text{Eq. I-5}$$

Where P corresponds to the pressure and V to the volume of a system.[93] However, in most thermodynamic systems, the virial is affected both by the interactions with the container (the ideal gas part), as well as by interactions between particles (the real gas part), thus affecting the pressure. In this case, the virial of the system is now given by Eq. I-6:[93]

$$\mathcal{W} = -3PV + \sum_{i=1}^{N}\sum_{j=i+1}^{N} r_{ij}\frac{dv(r_{ij})}{dr_{ij}} \qquad \text{Eq. I-6}$$

With $v(r_{ij})$ being the interaction between atoms i and j. If $dv(r_{ij})/dr_{ij}$ is written as f_{ij}, the force acting between atoms i and j, then the pressure is given by Eq. I-7:[93]

$$P = \frac{1}{V}\left[Nk_BT - \frac{1}{3}\sum_{i=1}^{N}\sum_{j=i+1}^{N} r_{ij}f_{ij}\right] \qquad \text{Eq. I-7}$$

As the forces are calculated throughout the MD simulation, the virial of the system and, consequently, its pressure, are easily accessible.[93] Therefore, for a system to maintain constant pressure, its temperature and volume are of utmost importance. In MD simulations, volume changes are possible *via* scaling of the simulation box (more on this matter below) or independently changing the *x*-, *y*- or *z*-dimensions of the system.[118] This feature can be accomplished *via* the Berendsen,[113] Andersen,[112] or Parrinello–Rahman barostats.[119] Likewise the corresponding thermostat, with the Berendsen barostat the system is linked to a pressure bath and the atoms' positions and the system's dimension are rescaled periodically to attain the reference pressure (P_0).[113] In the Andersen barostat, similarly to the Nosé–Hoover thermostat, an extended system also acts on the real system. In this case, the system's extra degree of freedom is a "piston" which will set how the system's volume changes.[112] The Parrinello–Rahman barostat is similar to the Andersen one, with the additional possibility of the simulation box being able to change shape to better fit the simulated system.[119]

MD simulations assess the motions of a limited number of atoms and molecules. While the number of particles possible to study in a system has been steadily increasing due to more efficient algorithms, as well as more powerful computational resources (from hundreds of atoms in the 1970s,[120] to thousands,[121] millions,[122] or even 4.125 trillion atoms[123] nowadays), the properties of such systems are still closer to a droplet, being affected by surface effects, than to a bulk fluid. This issue is mitigated with periodic boundary conditions (PBC), that make all particles in a simulation box equivalent, regardless of their relative position to the box edge, as the simulation box is replicated in all dimensions of the system. This way, the simulation box is then a small volume within a bulk fluid,[112] as it is surrounded by 26 images of itself. The coordinates of the particles in each surrounding box are obtained by addition or subtraction of the box sides dimensions. As a particle diffuses to a neighbouring box through a side, its image enters the box through the opposite side, as illustrated in the 2D example in Figure I-7, thus preserving the *N* particles within the simulation box. Several types of simulation boxes are possible within PBC, being the most common the cube/parallelepiped (easy to program and visualise) and the truncated octahedron (nearly spherical and allows for more efficient solvation).[93]

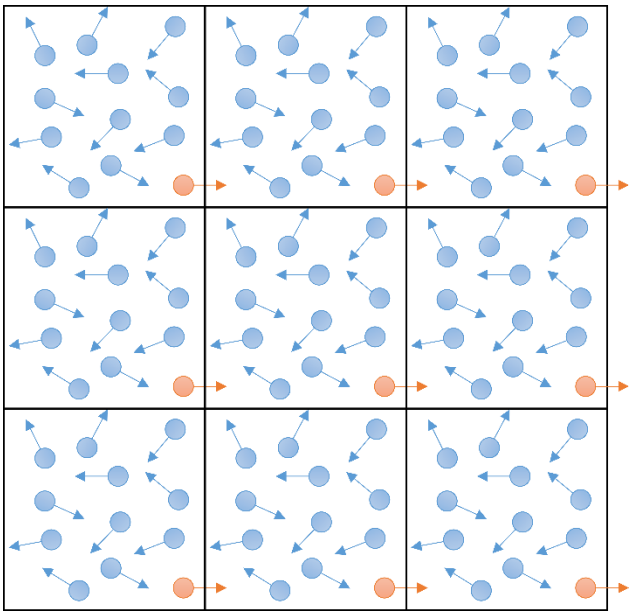


Figure I-7. Periodic boundary conditions in 2D.

I.3.2. Enhanced sampling techniques

MD simulations are an adequate method for the investigation of the dynamical behaviour changes of biological systems. However, an efficient sampling of these systems, characterised by complex potential energy surfaces with several minima and maxima, using stochastic MD simulations is unlikely.[124] These pitfall can be circumvented by enhanced sampling methods, such as Steered Molecular Dynamics (SMD),[125] Umbrella Sampling (US),[126] Replica–Exchange Molecular Dynamics (REMD),[127] or MetaDynamics (MetaD).[128] The first two methods, due to their application in Chapters IV and V, will be briefly explained below. These methods can also be used to ascertain how the free energy changes as the conformational space is sampled, being labelled potential of mean force (PMF) when the free energy changes are monitored along a reaction coordinate ξ.[93]

Steered Molecular Dynamics simulations

In SMD, a pseudo particle is linked to the real particle(s) *via* a "spring", and a steering force is applied to the pseudo particle along a reaction coordinate. Therefore, a system that was previously in equilibrium and is taken away from equilibrium will present a free energy difference. If a parameter λ is changed from λ_0 to λ_t between time zero and time t, then the work (W) done on the system cannot be smaller than the difference between the free energies associated with the initial and final values of λ, as shown by Eq. I-8:

$$\Delta F = F(\lambda_t) - F(\lambda_0) \leq \langle W \rangle \qquad \text{Eq. I-8}$$

According to Eq. I-8, a nonequilibrium process only provides an upper limit to the free energy difference.[129] However, Jarzynski found out that the free energy difference ΔF between the end states corresponding to λ_0 and λ_t can be related to W *via* Eq. I-9:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F} \qquad \text{Eq. I-9}$$

Where $\beta = 1/k_B T$. The average $\langle\ \rangle$ is obtained over several repetitions of the process.[130] The free energy difference between the two end states is finally given by:

$$\Delta F = -\frac{1}{\beta} \ln \langle e^{-\beta W} \rangle \qquad \text{Eq. I-10}$$

Umbrella Sampling simulations

In the US approach, the reaction coordinate is sampled in separated parallel "windows", thus increasing the efficiency of the overall simulation. Moreover, it is possible to enhance the sampling with increased simulation time or intermediate windows. In the MD simulation of each window, the system is only allowed to sample around a certain part of the reaction coordinate, which is attained by the application of a harmonic restraint to the reaction coordinate. The bias potential along the reaction coordinate $w_i(\xi)$ is given by:

$$w_i(\xi) = k_i \left(\xi - \xi_{0,i}\right)^2 \qquad \text{Eq. I-11}$$

Where k_i is the force constant that restrains the centre of the window i to the equilibrium position $\xi_{0,i}$. The force constant and distance between window centres should be chosen in a way that enables overlap of sampling in adjacent windows, an important requirement for the calculation of the PMF, which is obtained combining the data from each window, properly reweighted and unbiased (*vide infra*).[131] The force constant and distance between window centres affect the probability distribution $P_i^b(\xi)$ of the biased system, which is generally obtained through binning procedures to generate a histogram. $P_i^b(\xi_{0,i})$ is the fraction of sampled configurations relevant for the reaction coordinate $\xi_{0,i}$. The PMF for the biased system, along the reaction coordinate, is given by Eq. I-12, while the PMF for the unbiased system in each window is given by Eq. I-13, where F_i are constants that differ for each biasing potential or window,[132] and cannot be directly obtained from sampling.

$$A_i^b(\xi) = -\frac{1}{\beta}\ln P_i^b(\xi) \qquad \text{Eq. I-12}$$

$$A_i^u(\xi) = -\frac{1}{\beta}\ln P_i^b(\xi) - w_i(\xi) + F_i \qquad \text{Eq. I-13}$$

In this thesis, two methods were explored to reconstruct the PMF of the transmembrane diffusion of selected synthetic anion transporters: the Weighted Histogram Analysis Method (WHAM)[133, 134] and the variational Free Energy Profile (vFEP).[135]

Weighted Histogram Analysis Method

In WHAM the F_i constants can be iteratively calculated to combine the unbiased PMF of the independent windows i and j, solving the following equations:

$$P(\xi) = \frac{\sum_i^{windows} N_i P_i^b(\xi)}{\sum_j^{windows} N_j e^{[F_j - w_j(\xi)]\beta}} \qquad \text{Eq. I-14}$$

$$e^{-F_i\beta} = \int d\xi e^{-w_j(\xi)\beta} P(\xi) \qquad \text{Eq. I-15}$$

Where N_i corresponds to the total number of configurations sampled in window i to construct $P_i^b(\xi)$. Eq. I-14 and Eq. I-15 are coupled and are iteratively solved until self-consistency is attained, with the PMF $A(\xi)$ being obtained from $P(\xi)$ using Eq. I-12.[133, 134]

Variational Free Energy Profile

vFEP allows to restore free energy profiles under the maximum likelihood estimation (MLE),[136, 137] with cubic spline interpolation. When compared to WHAM, it is able to produce accurate PMF profiles with fewer data points and/or umbrella windows. Still, in vFEP the distinction between "probability" and "likelihood" is of paramount importance. Probability refers to the possible outcome of data and is usually modelled by a fixed functional form and a variable set of parameters, $P(\{\xi_n\}|\{\theta_m\})$, while likelihood refers to how likely a model can describe a set of observed outcome data, $\mathcal{L}(\{\theta_m\}|\{\xi_n\})$.[135]

The likelihood function of a set of data corresponds to the probability of obtaining that set of data given the probability distribution model function defined along with a set of trial model parameters. $P(\xi)$, the probability of observing a molecular system at a desired value ξ_i of a reaction coordinate ξ, is given by:

$$P(\xi) = \frac{e^{-F(\xi)}}{\int e^{-F(\xi_i)} d\xi_i} \qquad \text{Eq. I-16}$$

Where $F(\xi) \equiv \mathcal{F}(\xi)/k_B T$ corresponds to the unitless free energy profile, while $\mathcal{F}(\xi)$ is the energy profile. In a parametric model for the scaled free energy profile $F(\xi|\{\theta_m\})$, the probability distribution $P(\xi|\{\theta_m\})$ also contains the set of parameters, due to its relation to $F(\xi|\{\theta_m\})$. Therefore, the probability $P(\{\xi_n\}|\{\theta_m\})$ of a sampled data set $\{\xi_n\}$, with independent sampling points, is given by:

$$P(\{\xi_n\}|\{\theta_m\}) = P(\xi_1, \xi_2, \dots, \xi_N|\{\theta_m\}) = P(\xi_1|\{\theta_m\}) \cdot P(\xi_2|\{\theta_m\}) \cdot \dots \cdot P(\xi_N|\{\theta_m\}) \qquad \text{Eq. I-17}$$

The likelihood $\mathcal{L}$ of the trial free energy profile $F\{\theta_m\}$ with the observed data set $\{\xi_n\}$ becomes:

$$\mathcal{L}(F\{\theta_m\}|\xi_1, \xi_2, \dots, \xi_N) = \mathcal{L}(\{\theta_m\}|\xi_1, \xi_2, \dots, \xi_N) = \prod_{i=1}^{N} P(\xi_i|\{\theta_m\}) \qquad \text{Eq. I-18}$$

It is desirable to attempt to find an optimal solution to Eq. I-18 by defining a global function $F(\xi)$ with a set of parameters $\{\theta_m\}$, instead of dealing with individual values of ξ_i. The logarithm of the likelihood function, log-likelihood $\hat{l}$, is employed due to its convenience:

$$\hat{l}(\{\theta_m\}|\xi_1, \xi_2, \dots, \xi_N) = \frac{1}{N} \ln \mathcal{L} = \frac{1}{N} \sum_{n=1}^{N} \ln P(\xi_n|\{\theta_m\}) \qquad \text{Eq. I-19}$$

In the MLE method, the $\{\theta_m\}$ parameters are determined by maximising $\hat{l}$:

$$\hat{l}(\{\theta_m^*\}|\xi_1, \xi_2, \dots, \xi_N) = \underset{\{\theta_m\}\in\Theta}{\arg\max}\, \hat{l}(\{\theta_m\}|\xi_1, \xi_2, \dots, \xi_N) = \underset{\{\theta_m\}\in\Theta}{\arg\max}\, \frac{1}{N} \sum_{n=1}^{N} \ln P(\xi_n|\{\theta_m\}) \qquad \text{Eq. I-20}$$

In Eq. I-20, Θ defines the space that $\{\theta_m\}$ can span. When the biasing potential $w^\alpha(\xi)$ is applied to the α^{th} window in a set of US simulations, the probability of finding the system with a certain value of ξ can be obtained by Eq. I-21:

$$P^\alpha(\xi) = \frac{1}{Z^\alpha} e^{-[F(\xi)+w^\alpha(\xi)]} \qquad \text{Eq. I-21}$$

Where

$$Z^\alpha = \int_{-\infty}^{\infty} e^{-[F(\xi)+w^\alpha(\xi)]}\, d\xi \qquad \text{Eq. I-22}$$

In the simulation of the α^{th} window, with N^α observed points with $\{\xi_i^\alpha\}$ coordinate values, the probability of each point is $1/N^\alpha$. The likelihood of the whole system energy profile can be obtained from Eq. I-19 and Eq. I-21 as:

$$\hat{l}(F) \equiv \sum_{\alpha}^{windows} c^{\alpha}\, \hat{l}^{\alpha}(\{\theta_m\}|\{\xi_n^{\alpha}\}) = -\sum_{\alpha}^{windows} c^{\alpha}\left\{\ln Z^{\alpha} + \frac{1}{N^{\alpha}} \sum_{i}^{data\ points} [F(\xi_i^{\alpha}) + w^{\alpha}(\xi_i^{\alpha})]\right\} \quad \text{Eq. I-23}$$

Where $\{c^{\alpha}\}$ are the combination weights that define the relative contribution from the different windows when combining the local likelihoods into a global one. For windows with equal contributions, c^{α} may assume the same value for all windows. Finally, the F that maximises $\hat{l}(F)$ is determined as follows: *a)* define a trial function $F(\xi)$ with an initial parameter set $\{\theta_m\}$; *b)* assess the likelihood $\hat{l}(F)$ of the trial function $F(\xi)$ according to Eq. I-23; *c)* change the parameter set $\{\theta_m\}$ until reaching the maximum of $\hat{l}(F)$; *d)* the trial $F(\xi)$ with the maximal $\hat{l}(F)$ is the desired overall free energy profile.[135]

I.3.3. MD simulations of phospholipid bilayers

MD simulations of phospholipid bilayers produce structural insights at the atomic level, a feature that is hardly achievable trough experimental methods.[138] All MD simulations depend on the force fields that represent the interactions between atoms in the simulation system.[139] Due to the empirical origin of the encoded parameters, a careful selection of force fields for a specific type of simulation/system is necessary.[139] The simulation of phospholipid bilayer systems has seen great advances in the recent years, with the development of specific force fields. Table I-2 gathers a list of recently developed force fields.

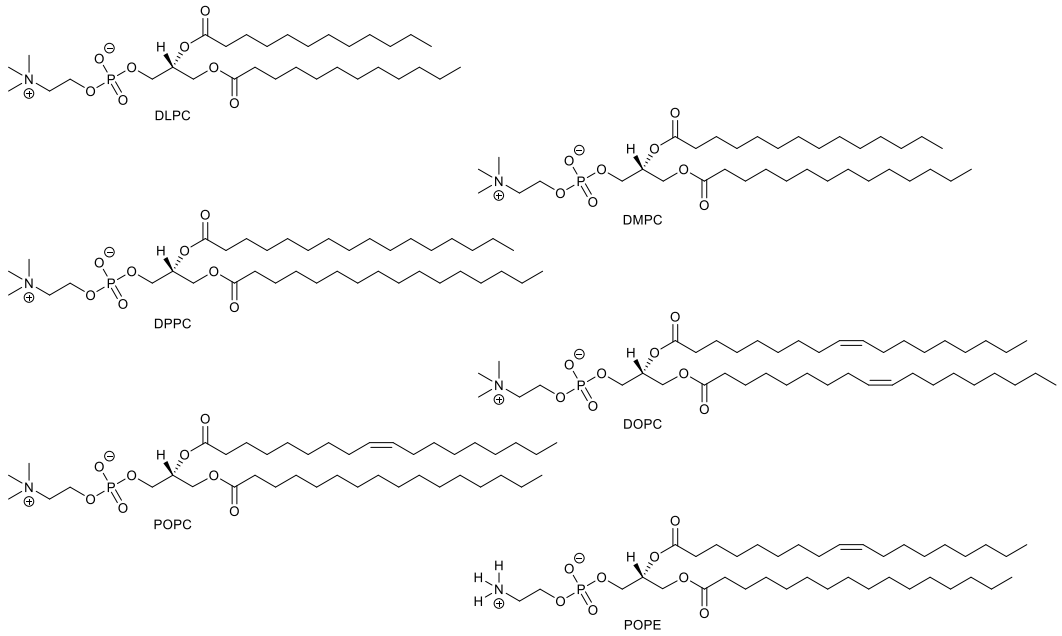
Table I-2. Examples of lipid specific force fields.

Type of force field	Force field	Lipids
All-atom	Lipid14	DLPC, DMPC, DPPC, DOPC, POPC, POPE,[140] and cholesterol[141]
	CHARMM36	DPPC, DMPC, DLPC, DOPC, POPC, POPE,[142] and cholesterol[143]
	OPLS-AA	DPPC,[144, 145] POPC, DOPC, PEPC, and cholesterol[145, 146]
United atom	G43A1-S3	DLPC, DMPC, DPPC, DOPC and cholesterol[147]
	G54A7	DPPC, DLPC, DMPC, DOPC, POPC[148, 149]
	CHARMM36-UA	DMPC, DPPC, POPC, DOPC, DMPC and cholesterol[150]
Coarse-grained	Dry MARTINI	POPC, DPPC, DMPC, DOPC, DLiPC, DOPE and cholesterol[151]
	SIRAH	DMPC[152]

DLiPC – 1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine; DLPC – 1,2-dilauroyl-*sn*-glycero-3-phosphocholine; DMPC – 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DOPC – 1,2-dioleoyl-*sn*-glycero-3-phosphocholine; DOPE – 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine; DPPC – 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine; PEPC – 1-palmitoyl-2-elaidoyl-*sn*-glycero-3-phosphocholine; POPC – 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; POPE – 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine.

From the force fields listed in Table I-2, Lipid14 is of special interest, as seen in the following chapters of this work. This lipid force field results from the evolution of the application of GAFF[102, 103] to lipid molecules in the Lipid11[153] and GAFFLipid[154] iterations. Lipid11 took the first steps in the transferability of lipid force fields,[92] splitting a phospholipid molecule in three components: a head group and two tails. These individually parameterised fragments can then be combined to simulate 32 different phospholipid molecules, such as DLPC, DMPC, DPPC, DOPC, POPC, and POPE, sketched in Scheme I-5. Noteworthy, other combinations are possible but require previous validation. Unfortunately, negatively charged phospholipids, important in the development of antimicrobial drugs,[155, 156] are only expected to be available in the upcoming expansion of the Lipid14 force field, dubbed Lipid17.[157] Still, likewise the initial works done with GAFF,[158-160] Lipid11 still required the use of a surface tension parameter to simulate the fluid phases of bilayers.[153] This issue was tackled in the parameterisation of GAFFLipid's Lennard–Jones parameters, allowing the simulation of membrane models in the fluid phases in a tensionless NPT ensemble.[154] Lipid14 results from these two-preceding

lipid force fields and allows the simulation of a combination of a phosphatidylcholine (PC) or a phosphatidylethanolamine (PE) head group with lauroyl (LA), myristoyl (MY), palmitoyl (PA), or oleoyl (OL) tails without the need to apply surface tension.[140] Moreover, cholesterol, one of the main components of cell membranes, affecting the phase and ordering of lipid bilayers, was recently parameterised, being fully compatible with Lipid14.[141] Furthermore, Lipid14 and CHARMM36 were able to reproduce the assembly of lipid bilayers from phospholipid molecules randomly distributed in aqueous solution.[157, 161] This feature used to be available only to coarse-grained[162] or united atoms[163] force fields, but with resort to the GPU acceleration[164-166] available in the recent versions of the AMBER package,[167-171] it is now possible with all-atom force fields.



Scheme I-5. Six examples of lipid types supported by the Lipid14 force field.[140]

As the experimental investigations that were used as the base to develop this thesis were carried out with POPC vesicles, the theoretical investigations herein presented use Lipid11 (Chapter II) and, later on, Lipid14 (Chapters III, IV and V) to describe the POPC lipids of the membrane models, as further discussed, due to their compatibility with GAFF, which was the classical the force field used to describe the anion transporters (*vide infra*).

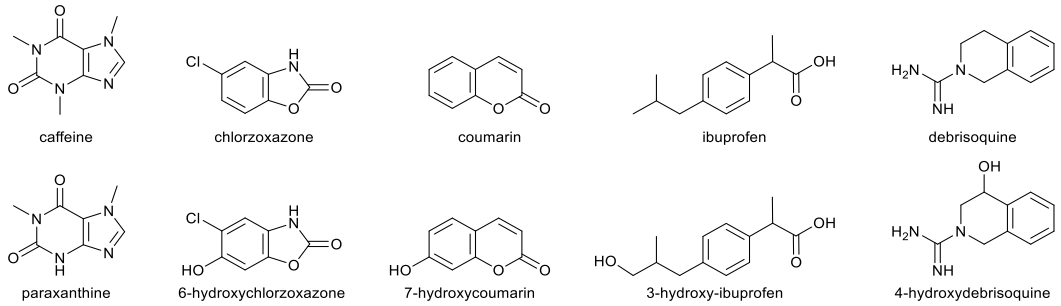
I.3.4. Applications in the movement of ions and molecules across membranes

The movement of ions and some small molecules, typically drugs, across lipid membranes has been studied with MD simulations, as briefly described below.

Toby W. Allen *et al.* have carried theoretical investigations of the mechanisms associated with the translocation of chloride, potassium and sodium across phospholipid bilayers.[172] This investigation was undertaken with US simulations[126] of the free ions, simulated along a reaction coordinate (the bilayer normal), along which the probability distribution was biased with the application of a 2.5 kcal mol^{-1} $Å^{-2}$ restraint. From these US simulations, it was possible to restore the free energy profile along the reaction coordinate or PMF associated with such translocation using WHAM**.** This method corrected for the biasing potential applied so that a particular region of the reaction coordinate could be sampled.[133] The free energy profiles obtained for these monoatomic ions revealed similar energy barriers for Cl^- (24.1±0.3 kcal mol^{-1}), Na^+ (25.6 ± 0.2 kcal mol^{-1}) and K^+ (25.8 + 0.1 kcal mol^{-1}), all located

at the core of the bilayer.[172] Moreover, from a structural point of view, it was observed that the hydrated ions, while permeating the bilayer, cause a water defect that drags the lipid head groups into the membrane.[172]

Michal Otyepka *et al*. have studied the transfer of five drug-like molecules that are cytochrome P450's substrates, as well as their metabolites (see Scheme I-6), from the water phase to the core of phospholipid bilayers of POPG or DOPC.[173] As several chemical and biological processes are governed by free energy differences, predicting how a molecular process will unfold (*e.g.*, the binding of an inhibitor) is an important asset in the development of drugs.[174] Still, other issues affect the pharmacokinetic profiles of drugs, such as their ability, or of their metabolites, to cross cell membranes until they reach their targets. Although simple MD simulations can provide important details on the interaction of small drug-like molecules and membrane models, these calculations are often insufficient to sample the conformational space associated with the traversing process of a bilayer. Hence, Michal Otyepka *et al*. have selected the bilayer normal as the reaction coordinate for their simulations, to compare the affinity and penetration capacities of the studied drugs and metabolites to both membrane models.[173] The membrane normal was properly sampled through US simulations,[126] and the free energy surfaces were reconstructed using WHAM.[133, 175] In their theoretical study, they found that the more hydrophobic drug molecules are more likely to permeate the phospholipid bilayer and have lower energy barriers to cross it, while their polar metabolites have lower affinities for the bilayer medium, showing higher energy barriers to cross the water/lipid interface. Moreover, when comparing the energy profiles across the neutral and negatively charged bilayers, it was found that the energy barriers are lower on the POPG membrane model, easing the permeation of this bilayer, being an important insight for future drug-design efforts.[173]



Scheme I-6. Drugs (top) and metabolites (bottom) studied by Michal Otyepka *et al*.[173]

I.4. Thesis scope

The work reported in this thesis has been developed in the scope of the long-term commitment of the Molecular Modelling and Computational Biophysics Group to the development and investigation of synthetic anion carriers. Indeed, our group has been a pioneer in the application of quantum calculations and classical MD simulations to the study of series of small drug-like potential transporters, such as bis-indolylureas,[85] tambjamines,[83] phenylthioureas,[86] *ortho*-phenylenediamine-based bis-ureas,[87] acylthioureas,[88] tris-thioureas,[89] bis-thioureidodecalins,[84] and squaramide derivatives[91] as well as a calix[4]arene derivative.[90] Part of these theoretical investigations were developed within the collaborative network with Prof. Philip A. Gale (University of Sydney) and Prof. Anthony P. Davis (University of Bristol), two lead researchers in the anion transport arena.

The theoretical investigations with tris-thiourea[89] and bis-thioureidodecalin derivatives[84] are herein featured as Chapters II and III, respectively. On the other hand, Chapter IV presents molecular modelling investigations on a series of squaramide derivatives, while Chapter V focus on a series tripodal

tris-thioureas with different degrees of fluorination. Both studies are currently submitted for publication. Although the tripodal molecules in Chapters II and V are structurally related, their anion affinity and lipophilicity were modulated resorting to different strategies (aryl moieties with different fluorination degrees in Chapter II, and alkyl chains with varying lengths and fluorination degrees in Chapter V). Moreover, each Chapter is complemented by its corresponding Appendix with additional details.

I.5. Objectives

Considering the paramount importance of the development of new synthetic anion transmembrane transporters towards channel replacement therapies, as demonstrated above, this thesis advanced with a comprehensive theoretical study on the anion transport promoted by synthetic molecules, as follows:

- Characterise, at the atomistic level, the interaction between synthetic transporters and the phospholipid bilayers;
- Study the mechanisms for the uptake and assisted transmembrane transport of chloride;
- Evaluate the structural impact resulting from membrane interaction and permeation by small synthetic molecules;
- Estimate the free energy profiles associated with anion complex/transporter translocation across the phospholipid bilayer.
- Understand the coupled-transport mechanisms associated with the synthetic molecules' transport activity.

Fulfilling these objectives will surely contribute for the development of more efficient anion transporters and potential anion channels replacement therapies.

I.6. References

1. B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, Garland Science, 6th edn., 2014.
2. H. Lodish, A. Berk, C. A. Kaiser, M. Krieger, A. Bretscher, H. Ploegh, A. Amon and K. C. Martin, *Molecular Cell Biology*, W. H. Freeman, 8th edn., 2016.
3. R. A. Rhoades and D. R. Bell, *Medical Physiology*, Lippincott Williams Wilkins, 4th edn., 2012.
4. S. J. Singer and G. L. Nicolson, *Science*, 1972, **175**, 720-731.
5. R. H. Garrett and C. M. Grisham, *Biochemistry*, 4th edn., 2008.
6. G. Karp, *Cell and Molecular Biology: Concepts and Experiments*, John Wiley and Sons, Inc, 7th edn., 2013.
7. M. Mueckler and B. Thorens, *Mol. Aspects Med.*, 2013, **34**, 121-138.
8. P. Lauger, *Science*, 1972, **178**, 24-30.
9. F. M. Ashcroft, *Ion Channels and Disease*, Academic Press, London, 1st edn., 2000.
10. C. A. Hubner and T. J. Jentsch, *Hum. Mol. Genet.*, 2002, **11**, 2435-2445.
11. F. M. Ashcroft, *Nature*, 2006, **440**, 440-447.
12. S. C. Cannon, *Neurotherapeutics*, 2007, **4**, 174-183.
13. R. Planells-Cases and T. J. Jentsch, *Biochim. Biophys. Acta*, 2009, **1792**, 173-189.
14. W. F. Boron and E. L. Boulpaep, *Medical Physiology: A Cellular and Molecular Approach*, Saunders, 2nd edn., 2008.
15. G. G. Celesia, *Clin. Neurophysiol.*, 2001, **112**, 2-18.
16. J. B. Kim, *Korean J. Pediatr.*, 2014, **57**, 1-18.
17. A. P. Davis, D. N. Sheppard and B. D. Smith, *Chem. Soc. Rev.*, 2007, **36**, 348-357.
18. B. P. O'Sullivan and S. D. Freedman, *The Lancet*, 2009, **373**, 1891-1904.
19. V. Scotet, I. Dugueperoux, P. Saliou, G. Rault, M. Roussey, M. P. Audrezet and C. Ferec, *Orphanet J. Rare Dis.*, 2012, **7**, 14.
20. M. R. Ahsan, H. s. K. Alam, A. R. Hasan, M. Z. Islam and A. Hossain, *Bangladesh Journal of Child Health*, 2017, **40**, 5.

21. F. Lehmann-Horn and K. Jurkat-Rott, in *Wiley Encyclopedia of Chemical Biology*, John Wiley & Sons, Inc., 2008, pp. 1201-1210.
22. A. Bush, E. W. F. W. Alton, J. C. Davies, U. Griesenbach and A. Jaffe, *Cystic Fibrosis in the 21st Century*, S. Karger, 2005.
23. M. D. Amaral, *J. Intern. Med.*, 2015, **277**, 155-166.
24. M. A. Ashlock, R. J. Beall, N. M. Hamblett, M. W. Konstan, C. M. Penland, B. W. Ramsey, J. M. Van Dalfsen, D. R. Wetmore and P. W. Campbell, 3rd, *Semin. Respir. Crit. Care Med.*, 2009, **30**, 611-626.
25. M. M. G. Antonisse and D. N. Reinhoudt, *Chem. Commun.*, 1998, 443-448.
26. A. D. McNaught and A. Wilkinson, *IUPAC. Compendium of Chemical Terminology*, Blackwell Scientific Publications, Oxford, 2nd edn., 1997.
27. T. M. Fyles, in *Bioorganic Chemistry Frontiers*, ed. H. Dugas, Springer Berlin Heidelberg, Berlin, Heidelberg, 1990, ch. Chapter 2, pp. 71-113.
28. *Encyclopedia of Science and Technology: I-Lev*, McGraw-Hill, New York, 10th edn., 2007.
29. T. M. Fyles, in *Inclusion Aspects of Membrane Chemistry*, eds. T. Osa and J. L. Atwood, Springer Netherlands, 1991, vol. 2, ch. Chapter 2, pp. 59-110.
30. E. J. Jeong, E. J. Kang, L. T. Sung, S. K. Hong and E. Lee, *J. Am. Chem. Soc.*, 2002, **124**, 14655-14662.
31. T. R. Sheth, R. M. Henderson, S. B. Hladky and A. W. Cuthbert, *Biochim. Biophys. Acta*, 1992, **1107**, 179-185.
32. S. Ohkuma, T. Sato, M. Okamoto, H. Matsuya, K. Arai, T. Kataoka, K. Nagai and H. H. Wasserman, *Biochem. J*, 1998, **334**, 731-741.
33. J. T. Davis, P. A. Gale, O. A. Okunola, P. Prados, J. C. Iglesias-Sanchez, T. Torroba and R. Quesada, *Nat Chem*, 2009, **1**, 138-144.
34. J. T. Davis, O. Okunola and R. Quesada, *Chem. Soc. Rev.*, 2010, **39**, 3843-3862.
35. G. A. Cook, O. Prakash, K. Zhang, L. P. Shank, W. A. Takeguchi, A. Robbins, Y. X. Gong, T. Iwamoto, B. D. Schultz and J. M. Tomich, *Biophys. J.*, 2004, **86**, 1424-1435.
36. B. A. McNally, E. J. O'Neil, A. Nguyen and B. D. Smith, *J. Am. Chem. Soc.*, 2008, **130**, 17274-17275.
37. A. J. Atkinson, S.-M. Huang, J. J. L. Lertora and S. P. Markey, *Principles of Clinical Pharmacology*, Academic Press, 3rd edn., 2012.
38. C. J. Pedersen, *J. Am. Chem. Soc.*, 1967, **89**, 7017-7036.
39. C. H. Park and H. E. Simmons, *J. Am. Chem. Soc.*, 1968, **90**, 2431-2432.
40. T. M. Cresp, C. L. Probert and F. Sondheimer, *Tetrahedron Lett.*, 1978, **19**, 3955-3958.
41. D. A. Evans, C. E. Sacks, R. A. Whitney and N. G. Mandel, *Tetrahedron Lett.*, 1978, **19**, 727-730.
42. I. Tabushi, Y. Kuroda and K. Yokota, *Tetrahedron Lett.*, 1982, **23**, 4601-4604.
43. A. J. M. van Beijnen, R. J. M. Nolte, J. W. Zwikker and W. Drenth, *Recl. Trav. Chim. Pays-Bas*, 2010, **101**, 409-410.
44. D. A. Doyle, J. Morais Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait and R. MacKinnon, *Science*, 1998, **280**, 69-77.
45. E. Gouaux and R. Mackinnon, *Science*, 2005, **310**, 1461-1465.
46. P. A. Gale, *Coord. Chem. Rev.*, 2000, **199**, 181-233.
47. P. A. Gale, *Coord. Chem. Rev.*, 2001, **213**, 79-128.
48. G. W. Gokel and A. Mukhopadhyay, *Chem. Soc. Rev.*, 2001, **30**, 274-286.
49. J. M. Boon and B. D. Smith, *Curr. Opin. Chem. Biol.*, 2002, **6**, 749-756.
50. P. A. Gale, *Coord. Chem. Rev.*, 2003, **240**, 191-221.
51. S. Matile, A. Som and N. Sorde, *Tetrahedron*, 2004, **60**, 6405-6435.
52. G. W. Gokel, W. M. Leevy and M. E. Weber, *Chem. Rev.*, 2004, **104**, 2723-2750.
53. N. Sakai, J. Mareda and S. Matile, *Acc. Chem. Res.*, 2005, **38**, 79-87.
54. P. A. Gale and R. Quesada, *Coord. Chem. Rev.*, 2006, **250**, 3219-3244.
55. T. M. Fyles, *Chem. Soc. Rev.*, 2007, **36**, 335-347.
56. G. W. Gokel and I. A. Carasel, *Chem. Soc. Rev.*, 2007, **36**, 378-389.
57. B. A. McNally, W. M. Leevy and B. D. Smith, *Supramol. Chem.*, 2007, **19**, 29-37.
58. N. Sakai, J. Mareda and S. Matile, *Mol Biosyst*, 2007, **3**, 658-666.
59. P. A. Gale, S. E. Garcia-Garrido and J. Garric, *Chem. Soc. Rev.*, 2008, **37**, 151-190.
60. S. Licen, F. De Riccardis, I. Izzo and P. Tecilla, *Curr Drug Discov Technol*, 2008, **5**, 86-97.
61. C. Caltagirone and P. A. Gale, *Chem. Soc. Rev.*, 2009, **38**, 520-563.
62. P. A. Gale, *Chem. Soc. Rev.*, 2010, **39**, 3746-3771.
63. P. R. Brotherhood and A. P. Davis, *Chem. Soc. Rev.*, 2010, **39**, 3633-3647.
64. S. O. Kang, J. M. Llinares, V. W. Day and K. Bowman-James, *Chem. Soc. Rev.*, 2010, **39**, 3980-4003.
65. S. Matile, A. Vargas Jentzsch, J. Montenegro and A. Fin, *Chem. Soc. Rev.*, 2011, **40**, 2453-2474.
66. C. J. Haynes and P. A. Gale, *Chem. Commun.*, 2011, **47**, 8203-8209.
67. M. Wenzel, J. R. Hiscock and P. A. Gale, *Chem. Soc. Rev.*, 2012, **41**, 480-520.
68. P. A. Gale, R. Perez-Tomas and R. Quesada, *Acc. Chem. Res.*, 2013, **46**, 2801-2813.
69. P. A. Gale, N. Busschaert, C. J. Haynes, L. E. Karagiannidis and I. L. Kirby, *Chem. Soc. Rev.*, 2014, **43**, 205-241.
70. N. Busschaert, C. Caltagirone, W. Van Rossom and P. A. Gale, *Chem. Rev.*, 2015, **115**, 8038-8155.
71. P. A. Gale, J. T. Davis and R. Quesada, *Chem. Soc. Rev.*, 2017, **46**, 2497-2519.

72. P. H. Schlesinger, R. Ferdani, R. Pajewski, J. Pajewska and G. W. Gokel, *Chem. Commun.*, 2002, 840-841.
73. V. Gorteau, G. Bollot, J. Mareda, A. Perez-Velasco and S. Matile, *J. Am. Chem. Soc.*, 2006, **128**, 14788-14789.
74. P. V. Santacroce, O. A. Okunola, P. Y. Zavalij and J. T. Davis, *Chem. Commun.*, 2006, 3246-3248.
75. S. K. Berezin and J. T. Davis, *J. Am. Chem. Soc.*, 2009, **131**, 2458-2459.
76. P. Iglesias Hernandez, D. Moreno, A. A. Javier, T. Torroba, R. Perez-Tomas and R. Quesada, *Chem. Commun.*, 2012, **48**, 1556-1558.
77. E. Hernando, V. Soto-Cerrato, S. Cortes-Arroyo, R. Perez-Tomas and R. Quesada, *Org Biomol Chem*, 2014, **12**, 1771-1778.
78. A. V. Jentzsch, D. Emery, J. Mareda, S. K. Nayak, P. Metrangolo, G. Resnati, N. Sakai and S. Matile, *Nat Commun*, 2012, **3**, 905.
79. S. Benz, M. Macchione, Q. Verolet, J. Mareda, N. Sakai and S. Matile, *J. Am. Chem. Soc.*, 2016, **138**, 9093-9096.
80. G. W. Gokel and N. Barkey, *New J. Chem.*, 2009, **33**, 947-963.
81. J. P. Behr, M. Kirch and J. M. Lehn, *J. Am. Chem. Soc.*, 1985, **107**, 241-246.
82. A. Vargas Jentzsch, D. Emery, J. Mareda, P. Metrangolo, G. Resnati and S. Matile, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 11675-11678.
83. V. Saggiomo, S. Otto, I. Marques, V. Felix, T. Torroba and R. Quesada, *Chem. Commun.*, 2012, **48**, 5274-5276.
84. S. J. Edwards, I. Marques, C. M. Dias, R. A. Tromans, N. R. Lees, V. Felix, H. Valkenier and A. P. Davis, *Chem. Eur. J.*, 2016, **22**, 2004-2011.
85. C. J. E. Haynes, S. J. Moore, J. R. Hiscock, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2012, **3**, 1436-1444.
86. N. Busschaert, S. J. Bradberry, M. Wenzel, C. J. E. Haynes, J. R. Hiscock, I. L. Kirby, L. E. Karagiannidis, S. J. Moore, N. J. Wells, J. Herniman, G. J. Langley, P. N. Horton, M. E. Light, I. Marques, P. J. Costa, V. Félix, J. G. Frey and P. A. Gale, *Chem. Sci.*, 2013, **4**, 3036-3045.
87. S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103-117.
88. C. J. Haynes, N. Busschaert, I. L. Kirby, J. Herniman, M. E. Light, N. J. Wells, I. Marques, V. Felix and P. A. Gale, *Org Biomol Chem*, 2014, **12**, 62-72.
89. I. Marques, A. R. Colaco, P. J. Costa, N. Busschaert, P. A. Gale and V. Felix, *Soft Matter*, 2014, **10**, 3608-3621.
90. P. J. Costa, I. Marques and V. Felix, *Biochim. Biophys. Acta*, 2014, **1838**, 890-901.
91. N. Busschaert, S. H. Park, K. H. Baek, Y. P. Choi, J. Park, E. N. W. Howe, J. R. Hiscock, L. E. Karagiannidis, I. Marques, V. Felix, W. Namkung, J. L. Sessler, P. A. Gale and I. Shin, *Nat Chem*, 2017, **9**, 667-675.
92. A. R. Leach, in *Molecular Modelling: Principles and Applications*, Pearson - Prentice Hall, 2nd edn., 2001, pp. 165-252.
93. A. R. Leach, in *Molecular Modelling: Principles and Applications*, Pearson - Prentice Hall, 2nd edn., 2001, pp. 301-352.
94. P. Auffinger, in *Handbook of RNA Biochemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2014, pp. 687-718.
95. V. Brázdová and D. R. Bowler, in *Atomistic Computer Simulations*, Wiley-VCH Verlag GmbH & Co. KGaA, 2013, pp. 193-211.
96. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
97. S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952-962.
98. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463-1472.
99. K. A. Feenstra, B. Hess and H. J. C. Berendsen, *J. Comput. Chem.*, 1999, **20**, 786-798.
100. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, 3rd, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman, *Comput. Phys. Commun.*, 1995, **91**, 1-41.
101. Y. Wang, S. A. Shaikh and E. Tajkhorshid, *Physiology (Bethesda)*, 2010, **25**, 142-154.
102. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
103. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
104. J. Wang, Development of the Second Generation of the General AMBER Force Field, https://www.researchgate.net/project/Development-of-the-Second-Generation-of-the-General-AMBER-Force-Field, (accessed Nov 15th, 2017).
105. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
106. A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2000, **21**, 132-146.
107. A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623-1641.
108. J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219-3228.
109. C. Chen, P. Depa, J. K. Maranas and V. Garcia Sakai, *J. Chem. Phys.*, 2008, **128**, 124906.
110. A. Liwo, *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*, Springer-Verlag Berlin Heidelberg, 2014.
111. A. Renevey and S. Riniker, *J. Chem. Phys.*, 2017, **146**, 124131.
112. H. C. Andersen, *J. Chem. Phys.*, 1980, **72**, 2384-2393.
113. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
114. S. Nose, *J. Chem. Phys.*, 1984, **81**, 511-519.
115. W. G. Hoover, *Phys Rev A Gen Phys*, 1985, **31**, 1695-1697.

116. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
117. E. B. Tadmor and R. E. Miller, in *Modeling Materials - Continuum, Atomistic and Multiscale Techniques*, Cambridge University Press, Cambridge, 2011, ch. 7, pp. 377-439.
118. S. M. Rogge, L. Vanduyfhuys, A. Ghysels, M. Waroquier, T. Verstraelen, G. Maurin and V. Van Speybroeck, *J. Chem. Theory Comput.*, 2015, **11**, 5583-5597.
119. M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
120. J. A. McCammon, B. R. Gelin and M. Karplus, *Nature*, 1977, **267**, 585-590.
121. A. Hospital, J. R. Goni, M. Orozco and J. L. Gelpi, *Adv Appl Bioinform Chem*, 2015, **8**, 37-47.
122. J. Gumbart, L. G. Trabuco, E. Schreiner, E. Villa and K. Schulten, *Structure*, 2009, **17**, 1453-1464.
123. W. Eckhardt, A. Heinecke, R. Bader, M. Brehm, N. Hammer, H. Huber, H.-G. Kleinhenz, J. Vrabec, H. Hasse, M. Horsch, M. Bernreuther, C. W. Glass, C. Niethammer, A. Bode and H.-J. Bungartz, in *Supercomputing*, eds. J. M. Kunkel, T. Ludwig and H. W. Meuer, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, ch. Chapter 1, pp. 1-12.
124. R. C. Bernardi, M. C. Melo and K. Schulten, *Biochim. Biophys. Acta*, 2015, **1850**, 872-877.
125. D. Kosztin, S. Izrailev and K. Schulten, *Biophys. J.*, 1999, **76**, 188-197.
126. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187-199.
127. Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, 1999, **314**, 141-151.
128. A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562-12566.
129. S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *J. Chem. Phys.*, 2003, **119**, 3559-3566.
130. C. Jarzynski, *Phys. Rev. Lett.*, 1997, **78**, 2690-2693.
131. J. Kastner, *Wires Comput Mol Sci*, 2011, **1**, 932-942.
132. D. K. Chakravorty, M. Kumarasiri, A. V. Soudackov and S. Hammes-Schiffer, *J. Chem. Theory Comput.*, 2008, **4**, 1974-1980.
133. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011-1021.
134. M. Souaille and B. Roux, *Comput. Phys. Commun.*, 2001, **135**, 40-57.
135. T. S. Lee, B. K. Radak, A. Pabis and D. M. York, *J. Chem. Theory Comput.*, 2013, **9**, 153-164.
136. C. Bartels, *Chem. Phys. Lett.*, 2000, **331**, 446-454.
137. Z. Q. Tan, *Journal of the American Statistical Association*, 2004, **99**, 1027-1036.
138. A. P. Lyubartsev and A. L. Rabinovich, *Biochim. Biophys. Acta*, 2016, **1858**, 2483-2497.
139. M. S. P. Sansom and P. C. Biggin, *Molecular Simulations and Biomembranes*, Royal Society of Chemistry, 1st edn., 2010.
140. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
141. B. D. Madej, I. R. Gould and R. C. Walker, *J. Phys. Chem. B*, 2015, **119**, 12424-12435.
142. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, Jr. and R. W. Pastor, *J. Phys. Chem. B*, 2010, **114**, 7830-7843.
143. J. B. Lim, B. Rogaski and J. B. Klauda, *J. Phys. Chem. B*, 2012, **116**, 203-210.
144. A. Maciejewski, M. Pasenkiewicz-Gierula, O. Cramariuc, I. Vattulainen and T. Rog, *J. Phys. Chem. B*, 2014, **118**, 4571-4581.
145. W. Kulig, M. Pasenkiewicz-Gierula and T. Rog, *Data Brief*, 2015, **5**, 333-336.
146. W. Kulig, M. Pasenkiewicz-Gierula and T. Rog, *Chem. Phys. Lipids*, 2016, **195**, 12-20.
147. S. W. Chiu, S. A. Pandit, H. L. Scott and E. Jakobsson, *J. Phys. Chem. B*, 2009, **113**, 2748-2763.
148. N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark and W. F. van Gunsteren, *Eur. Biophys. J.*, 2011, **40**, 843-856.
149. D. Poger and A. E. Mark, *J. Chem. Theory Comput.*, 2012, **8**, 4807-4817.
150. S. Lee, A. Tran, M. Allsopp, J. B. Lim, J. Henin and J. B. Klauda, *J. Phys. Chem. B*, 2014, **118**, 547-556.
151. C. Arnarez, J. J. Uusitalo, M. F. Masman, H. I. Ingolfsson, D. H. de Jong, M. N. Melo, X. Periole, A. H. de Vries and S. J. Marrink, *J. Chem. Theory Comput.*, 2015, **11**, 260-275.
152. E. E. Barrera, E. N. Frigini, R. D. Porasso and S. Pantano, *J. Mol. Model.*, 2017, **23**, 259.
153. A. A. Skjevik, B. D. Madej, R. C. Walker and K. Teigen, *J. Phys. Chem. B*, 2012, **116**, 11124-11136.
154. C. J. Dickson, L. Rosso, R. M. Betz, R. C. Walker and I. R. Gould, *Soft Matter*, 2012, **8**, 9617-9627.
155. J. N. Horn, J. D. Sengillo, D. Lin, T. D. Romo and A. Grossfield, *Biochim. Biophys. Acta*, 2012, **1818**, 212-218.
156. J. N. Horn, T. D. Romo and A. Grossfield, *Biochemistry*, 2013, **52**, 5604-5610.
157. A. A. Skjevik, B. D. Madej, C. J. Dickson, C. Lin, K. Teigen, R. C. Walker and I. R. Gould, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10573-10584.
158. B. Jojart and T. A. Martinek, *J. Comput. Chem.*, 2007, **28**, 2051-2058.
159. L. Rosso and I. R. Gould, *J. Comput. Chem.*, 2008, **29**, 24-37.
160. S. W. Siu, R. Vacha, P. Jungwirth and R. A. Bockmann, *J. Chem. Phys.*, 2008, **128**, 125103.
161. A. A. Skjevik, B. D. Madej, C. J. Dickson, K. Teigen, R. C. Walker and I. R. Gould, *Chem. Commun.*, 2015, **51**, 4402-4405.
162. W. Shinoda, R. DeVane and M. L. Klein, *J. Phys. Chem. B*, 2010, **114**, 6836-6849.
163. D. Poger, W. F. Van Gunsteren and A. E. Mark, *J. Comput. Chem.*, 2010, **31**, 1117-1125.

164. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
165. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
166. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
167. D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2010), AMBER 11, University of California, San Francisco.
168. D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2012), AMBER 12, University of California, San Francisco.
169. R. Salomon-Ferrer, D. A. Case and R. C. Walker, *Wires Comput Mol Sci*, 2013, **3**, 198-210.
170. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman (2014), AMBER 14, University of California, San Francisco.
171. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.
172. I. Vorobyov, T. E. Olson, J. H. Kim, R. E. Koeppe, 2nd, O. S. Andersen and T. W. Allen, *Biophys. J.*, 2014, **106**, 586-597.
173. M. Paloncyova, K. Berka and M. Otyepka, *J. Phys. Chem. B*, 2013, **117**, 2403-2410.
174. M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp and P. C. Biggin, *Chem Sci*, 2016, **7**, 207-218.
175. J. S. Hub, B. L. de Groot and D. van der Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713-3720.

Chapter II.

Tris-thiourea tripodal-based molecules as chloride transmembrane transporters: insights from molecular dynamics simulations

Summary

The interaction of six tripodal synthetic chloride transmembrane transporters with a POPC bilayer was investigated by means of Molecular Dynamics (MD) simulations using the General AMBER Force Field (GAFF) for the transporters and the Lipid11 force field for phospholipids. These transporters are structurally simple molecules, based on the tris(2-aminoethyl)amine scaffold, containing three thiourea binding units coupled with three *n*-butyl (**1**), phenyl (**2**), fluorophenyl (**3**), pentafluorophenyl (**4**), 4-trifluoromethylphenyl (**5**), or 3,5-bis(trifluoromethyl)phenyl (**6**) substituents (see Scheme II-1). The passive diffusion of **1**–**6**·Cl^- was evaluated with the complexes initially positioned either in the water phase or inside the bilayer. In the first scenario, the chloride ion is released in the water phase before the synthetic molecules achieve the water/lipid interface and permeate the membrane. In the latter one, only when the chloride complex reaches the interface it is able to release the anion to the water phase, with the transporter losing the initial *ggg* (*g* = gauche) tripodal shape. Independently of the transporter used in the membrane system, the bilayer structure is preserved and the synthetic molecules interact with the POPC lipids at the phosphate head group level, *via* N–H···O hydrogen bonds. Overall, the MD simulations' results indicate that the small tripodal molecules in this series have a low impact on the bilayer and are able to diffuse with chloride inside the lipid environment. Indeed, these are essential conditions for these molecules to promote transmembrane transport as anion carriers, in agreement with experimental efflux data. This chapter will focus in the optimisation of the tripodal chloride complexes, followed by the assessment of structural parameters (*i.e.*, area per lipid, bilayer thickness, order parameters and electron density profiles) of an overhydrated POPC bilayer described with Lipid11. This validated membrane model was further used to study the interaction between the six tripodal chloride complexes and the POPC bilayer.

The work reported in this chapter was published as: I. Marques, A. R. Colaço, P. J. Costa, N. Busschaert, P. A. Gale and V. Felix, *Soft Matter*, 2014, **10**, 3608-3621.

II.1. Introduction

Anion transport across phospholipid bilayers is essential to many cellular processes such as nerve conduction and maintenance of homeostasis.[1-4] Indeed, nature uses a combination of specialised proteins embedded in the membrane (*e.g.*, gated ion channels and ATP-driven transporters) to establish and control the anion concentrations in cellular compartments.[5, 6] The dysfunction of anion channels is currently linked with the occurrence of a multiplicity of serious pathologies,[1-3, 7, 8] including Cystic Fibrosis (CF) and male infertility, caused by defective transmembrane transport of chloride and bicarbonate anions.[1, 3, 4, 9, 10] On the other hand, there are only a few non-protein natural products able to operate anion transmembrane transport, such as the macrolide antibiotics pamamycin[11] and duramycin,[12] and the prodigiosin alkaloids.[13-15]

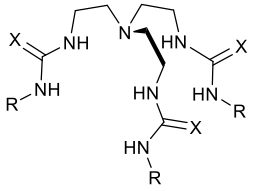
In this context, currently there is interest in the development of synthetic molecules able to promote passive anion transport and, consequently, with potential to be applied in the future as replacement therapeutics for malfunctioning transporter proteins.[16] Indeed, over the last decades, an increasing number of synthetic anion transporters have been designed using natural products or synthetic entities as platforms for assembly of the binding units.[17-25] The majority of these molecules mediate anion transport through lipid bilayers as mobile carriers or static channels.[14, 26]

On the other hand, MD simulations are a powerful tool to investigate, at the atomistic level, the energetics and mechanisms of anion transport, as well as the impact caused by the synthetic molecules on the structural and dynamic properties of phospholipid bilayers, from the first interactions with the membrane until their internalisation. Nevertheless, the permeation events through the membranes, addressed with MD studies, are typically focused on drug delivery,[27] protein ion channels,[28] and diffusion of small biological molecules across phospholipid bilayers,[29] leaving anion transport operated by synthetic molecules almost unexplored. Indeed, to the best of our knowledge, our recent reports on transmembrane chloride transport mediated by bis-indolylureas and bis-ureas are the first studies reporting MD simulations in this field.[30, 31] Furthermore, we have also shown, in a recent theoretical study, the ability of a charged calix[4]arene derivative to mediate the chloride transmembrane transport by means of steered MD.[32] Nowadays, the majority of membrane simulations are carried out using the all-atom CHARMM36[33] or the united-atom GROMOS96 (lipid specific parameter set G53A6$_L$) force fields.[34] Alternatively, large membrane systems may also be performed with coarse-grained force fields.[35] However, the two atomistic biomolecular force fields do not incorporate general parameters for arbitrary organic molecules. Indeed, the CHARMM36 can be used along with the CHARMM General Force Field (CGenFF)[36] for organic molecules, which, in its successive revised versions, includes parameters for a limited number of functional groups. For instance, the use of this force field to describe thiourea-based molecules still requires thorough development and subsequent validation of parameters, which is a challenge in itself.

By contrast, GAFF[37, 38] contains suitable parameters for most of the functional groups commonly used in the design of synthetic receptors. Furthermore, GAFF was also successfully tested in MD simulations of DOPC,[39, 40] POPC,[41] and DMPC bilayers.[39] Similar to other force fields, GAFF was able to successfully reproduce several structural and dynamic properties of these pure membrane models, namely when the surface tension coupling for surfaces parallel to the *xy*-plane was imposed. During the last years, two AMBER lipid specific force fields were made available: Lipid11[42] and GAFFlipid,[43] which was released when this work was in progress. While the former relies on transferable parameters per lipid head and lipid tail, making it compatible with several phospholipids, the latter relies on specific van der Waals interaction parameterisation for each type of phospholipid, yielding results closer to experimental data. However, both force fields have setbacks: likewise GAFF, Lipid11 still requires the use of surface tension to properly reproduce some structural parameters; on the other hand, GAFFlipid

specific parameterisation restricts its use on non-parameterised phospholipids. The core breakthroughs of each force field recently yielded Lipid14,[44] a transferable force field with no need for surface tension imposition, released after the conclusion of this computational research.

Recently, Gale *et al.*[45-47] synthesised a set of tris-thiourea and tris-urea receptors (see Scheme II-1) based on the tris(2-aminoethyl)amine (*tren*) tripodal scaffold aiming to make suitable transporters of relevant biological species such as chloride and bicarbonate anions. The binding affinities of this wide set of small structurally simple molecules were assessed in a DMSO-d_6/0.5% water solution by ^{1}H NMR experiments, revealing chloride association constants consistent with the use of these molecules in the transmembrane transport of this monatomic anion, as further demonstrated by efflux experiments carried out with POPC lipid vesicles. It was also reported that the molecules incorporating urea-binding groups were systematically less effective than the thiourea analogues in mediating chloride efflux from loaded POPC lipid vesicles, *via* anion antiport mechanisms.[45] Furthermore, the tripodal transporters with fluorinated phenyl substituents (**3**–**6** and **9**–**12**), which are less hydrophilic than the non-fluorinated analogues (**2** and **8**), are more potent anion transporters.[46] Moreover, assays carried out in cholesterol rich vesicles (70% POPC and 30% cholesterol) led to a significant decrease of the chloride transport rate, indicating that these small molecules operate as mobile carriers.[45, 46]



1 X=S, R=*n*-Bu
2 X=S, R=Ph
3 X=S, R=*p*-C_6H_4F
4 X=S, R=C_6F_5
5 X=S, R=4-CF_3Ph
6 X=S, R=3,5-$(CF_3)_2$Ph

7 X=O, R=*n*-Bu
8 X=O, R=Ph
9 X=O, R=*p*-C_6H_4F
10 X=O, R=C_6F_5
11 X=O, R=4-CF_3Ph
12 X=O, R=3,5-$(CF_3)_2$Ph

Scheme II-1. Series of tris-thiourea and tris-urea synthetic mobile carriers based on the *tren* scaffold.

Inspired by these experimental findings, the diffusion of **1**–**6** chloride complexes through a POPC membrane model was investigated by MD simulations. The results obtained are reported here, providing the first insights at the atomic level into how these small molecules interact with the membrane and mediate the chloride transmembrane transport. These studies rely on the same membrane model previously reported for the bis-ureas series of receptors,[31] which is further discussed.

II.2. Methods

The MD simulations were carried out with the AMBER 12 package[48] using Lipid11[42] for the POPC lipids and GAFF[37, 38] parameters and atomic RESP charges[49] for transporters **1**–**6** as follows. Structural analyses were performed with *cpptraj*,[50] except where mentioned otherwise. All structural diagrams were done with *PyMOL*.[51]

II.2.1. Force field parameters for thiourea tripod transporters

The atomic coordinates of **2** were obtained directly from the X-ray single crystal structure of the carbonate complex, deposited with the Cambridge Crystallographic Data Centre (CCDC)[52, 53] under RefCode FUXYED.[45] The initial structure of the remaining transporters was generated from **2** by the replacement of the phenyl substituents appended to the *tren* entity by butyl groups (**1**) or the adequate number of fluorine atoms (**3** and **4**) or trifluoromethyl groups (**5** and **6**). All structures were optimised at the HF/6-31G* level with *Gaussian 09*,[54] followed by a single point calculation to generate the electrostatic potential (ESP) at the same theory level using the Merz–Singh–Kollman scheme with 4 concentric layers per atom and 6 density points in each layer (IOp (6/33 = 2, 6/41 = 4, 6/42 = 6)). The

initial atomic charges of each transporter were then calculated by RESP fitting along with GAFF atom types using the *antechamber* module,[48] as implemented in the AMBER package.

To obtain atomic charges less dependent on molecular conformation or orientation, the calculation of the final RESP charges employed in the MD membrane simulations was preceded by conformational analyses on transporters **1**–**6**. The starting Molecular Mechanics (MM) energy minimised structures using the initial charges were heated in the gas-phase at 1000 K for 50 ps, followed by a collection run of 1.0 ns, using a time step of 1.0 fs. The use of this high temperature enables the stochastic search of the conformational space since the energetic barriers are easily surmounted.[55] Frames were saved every 0.1 ps leading to a trajectory file containing 10000 structures for each transporter. Subsequently, all these structures were minimised by MM using a steepest descendent gradient followed by the conjugate gradient algorithm, until the convergence criterion of 0.0001 kcal mol^{-1} $Å^{-1}$ was attained. Afterwards, the frames were energy sorted and the four lowest energy structures with different N–H thiourea binding group configurations were selected: two *syn-syn-syn*, one *syn-syn-anti* and one *syn-anti-anti*, leading to markedly different structures, thus avoiding an eventual charge derivation biased by using only *syn-syn-syn* structures. Conformations with an *anti-anti-anti* configuration were not observed. The selected structures were appointed for a new HF/6-31G* geometry optimisation as well as ESP calculations as described above. The individual ESP data were extracted from the corresponding *Gaussian 09*[54] outputs, concatenated, and subsequently used to generate the input files for the two-stage RESP fitting, using identical weights for all conformations.

II.2.2. Simulation systems

Pure POPC bilayer membrane model

System A, a pure membrane POPC system, composed of 128 POPC lipids and 6500 water molecules has been previously simulated at 303 K and reported by us,[31] using the Lipid11 force field for the lipids, the TIP3P water model[56] and a surface tension of 17 dyn cm^{-1}, as recommended for this particular force field to reproduce recent experimental data.[42, 57] Under these conditions, the biophysical parameters, area per lipid, bilayer thickness, electron density profiles and chain order parameters, were found in good agreement with those released with the original force field[42] for a system composed of 128 lipids and 31.6 water molecules per lipid. Therefore, system A differs from the one reported in ref. 42 only in the hydration degree, being sketched in Figure II-1.

System A has a total of 36652 atoms and two water slabs with a thickness of *ca.* 24 Å, which allow the insertion of the thiourea chloride complexes in the centre of a water slab. This was a requirement for the simulation of the internalisation of the transporters from the water phase into the phospholipid bilayer under periodic conditions with an 8.0 Å cut-off for the van der Waals and electrostatic interactions, as shown below.

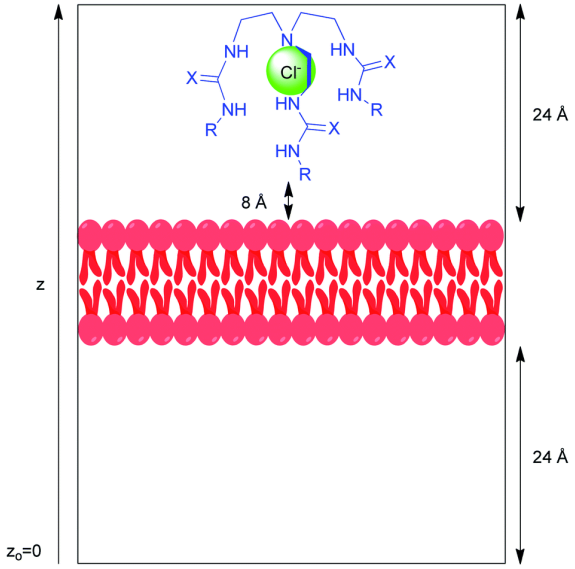


Figure II-1. Sketch representation of the over-hydrated membrane system able to immerse **1**–**6** chloride complexes in a water slab. The *z*-dimension runs along the longest edge of the orthorhombic simulation box, with the origin at 0 Å.

Insertion of chloride complexes in the POPC membrane model

To investigate whether the transporters are able to permeate the phospholipid membrane by passive diffusion, they were placed in the aqueous phase of a pre-equilibrated frame of system A,[31] using *Packmol*,[58] at a distance from the closest water/lipid interface that enabled the use of an 8.0 Å cut-off. Complexes **1**·Cl⁻ to **6**·Cl⁻ were left unrestrained to move around in the water slabs leading to systems B to G, respectively. On the other hand, to study if the chloride complexes are able to diffuse within the lipophilic environment provided by the phospholipid membrane, **3**·Cl⁻ was carefully placed at the core of a pre-equilibrated frame of a system containing 128 POPC lipids and 4040 water molecules (similar to the one of the Lipid11 original reference), using *Packmol*,[58] and left free of restraints, leading to system H. Thus, a total of eight systems were initially simulated, as given in Table II-1. Furthermore, the charge neutrality of both systems was achieved by the random addition of one sodium ion, which, along with the chloride ion, was described with van der Waals parameters developed to be used with the TIP3P water model.[59]

Table II-1. Simulated systems at 303 K using a surface tension of 17 dyn cm^{-1}.

System ID	Simulation ID[a]		Transporter	IP[b]	Simulation time (ns)
	Run 1	Run 2			
A	A1[c]	A2[d]	–	–	100
B	B1	B2	**1**	Water phase	100
C	C1	C2	**2**	Water phase	100
D	D1	D2	**3**	Water phase	100
E	E1	E2	**4**	Water phase	200
F	F1	F2	**5**	Water phase	100
G	G1	G2	**6**	Water phase	100
H	H1	H2	**3**	Bilayer core	150

[a)] For each system two runs were done, identified by the system ID and the corresponding run number; [b)] IP means the initial position of the chloride complex in the bilayer system; [c)] Run A1 was previously reported in ref. 31; [d)] The free membrane system sampling was increased with run A2; the reported values of system A correspond to both sampling periods being treated as one.

II.2.3. Simulation conditions

Systems A to H were simulated in orthorhombic boxes under periodic boundary conditions at 303 K. In addition, two independent runs using different random seeds for the initial velocities were performed for each system.

Simulation of pure membrane system A

The free membrane system was simulated under the conditions previously reported,[31] in two independent MD runs, to assess the effect of the increased hydration on the phospholipid bilayer.

Simulation of membrane systems with chloride complexes in the water phase – systems B–G

The unrestrained MD simulations of systems B to G were carried out as follows: the initial configuration of each system was submitted to 10000 steps of MM energy minimisation with a 500 kcal mol^{-1} $Å^{-2}$ positional restraint on the chloride complex and lipid molecules, through the steepest descent algorithm for 3000 steps plus 7000 steps of the conjugated gradient algorithm, followed by the relaxation of the entire system for another 10000 steps, with the same protocol. The equilibration of the system proceeded by heating it at 303 K in an NVT ensemble for 50 ps with a 10 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex and lipid molecules. The equilibration process proceeded with a 20 ns run using an NPγT ensemble with a surface tension of γ = 17 dyn cm^{-1}.[42] Then, the positional restraint was removed and the simulation continued for further 100 ns. In these two runs the long-range electrostatic interactions were described with the Particle Mesh Ewald (PME) algorithm[60] using a real-space cut-off at 8.0 Å. The cut-off for the Lennard–Jones interactions was also set at 8.0 Å. The temperature of the system was maintained at 303 K, using Langevin dynamics,[61] with a collision frequency γ of 1.0 ps^{-1}. The pressure was controlled by the Berendsen barostat[62] at 1 atm and a compressibility of 44.6×10^{-6} bar^{-1}, with a relaxation time of 1.0 ps. The covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm,[63] allowing the use of a 2.0 fs time step. This protocol was employed to run two independent runs of 100 ns for systems B to G, apart from system E, and the frames were saved every 1.0 ps. The evolution of several parameters (*vide infra*) along the simulation time for system E showed that it was not equilibrated along the first 100 ns, so the two runs were extended for another 100 ns. Throughout this last simulation period the equilibration of system E was achieved.

Simulation of membrane systems with chloride complexes at the bilayer core – system H

The unrestrained MD simulations of system H were preceded by a preparation stage in which the **3**·Cl^- complex was accommodated between the aliphatic chains of the phospholipids. The initial configuration of each system was submitted to 10000 steps of MM energy minimisation with a MM positional restraint on the chloride complex of 500 kcal mol^{-1} $Å^{-2}$, with 3000 steps of steepest descent algorithm, and 7000 steps of conjugated gradient algorithm. The whole system was then relaxed, with another set of 10000 steps. Subsequently, the system was heated to 303 K in an NVT ensemble for 50 ps with a 100 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex, followed by an NPγT run of 20 ns with a surface tension of γ = 17 dyn cm^{-1}. From this 20 ns NPγT run, a frame was selected according to the following criteria: *a)* no bad contacts between the POPC lipids and chloride complexes; and *b)* low water defects/permeation of the system. Thus, for the selected frame, the eventual permeating water molecules were repositioned in the aqueous phase and this final arrangement was used as a starting geometry for the multi-stage protocol, similar to the one used for systems B to G, except for a shorter equilibration period, with only 5.0 ns instead of 20 ns, and a collection time of 150 ns.

II.3. Results and Discussion

This section is divided into six subsections. The first one delves into the generation of the chloride complexes of **1**–**6**. In the second subsection, the simulation carried out with the over-hydrated POPC bilayer (system A) is briefly analysed and compared with experimental and theoretical biophysical data available. The third subsection reports the ability of the *tren*-based transporters **1**–**6** to diffuse through system A, starting from the water phase, while in the fourth the ability of these small molecules to diffuse across a lipid bilayer starting from the core of the membrane model is evaluated. The fifth subsection assesses the diffusion of a chloride complex starting from the water phase. The final subsection assesses the impact of **1**–**6** transporters on the structural and biophysical properties of the POPC phospholipid bilayer as a membrane model.

II.3.1. Generation of 1–6 chloride complexes for membrane MD simulations

For consistency, all structures of **1**–**6** chloride complexes were generated *via* quenched MD simulations as described above for the corresponding free transporters. The binding arrangements selected for the subsequent transmembrane diffusion studies were the lowest energy structures of the chloride complexes, illustrated in Figure II-2. Overall, in all complexes, the chloride ion is surrounded by the three thiourea binding groups establishing six N–H···Cl^- hydrogen bonds, with the N···Cl^- distances and N–H···Cl^- angles listed in Table II-2. In particular, the MM structure of **4**·Cl^-, with N···Cl^- distances and N–H···Cl^- angles ranging between 3.304 and 3.321 Å and 159.1 to 160.2°, respectively, compares well with its crystal structure as TBA^+ salt (CCDC RefCode SAJZEK),[46] where the N···Cl^- distances and N–H···Cl^- angles range from 3.226 to 3.393 Å and from 153.2 to 159.3°, respectively.

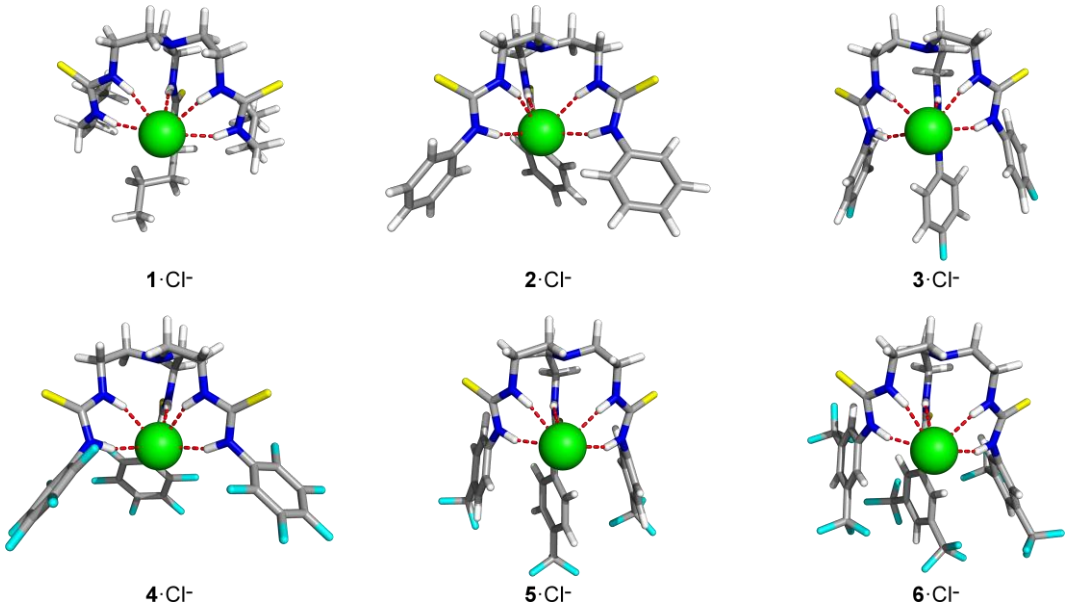


Figure II-2. MM lowest energy structure of the tripodal chloride complexes, with the N–H···Cl^- hydrogen bonds shown as red dashed lines. The carbon, nitrogen, sulfur, fluorine and hydrogen atoms are drawn in light grey, blue, yellow, cyan and white sticks, respectively. The chloride ion is represented as a green sphere.

Table II-2. Dimensions of N–H···Cl^- bonds found in the MM energy minimised structures of **1**–**6**·Cl^-.

Complex	Distance N···Cl^- (Å)		Angle N–H···Cl^- (°)	
	Min	Max	Min	Max
1·Cl^-	3.315	3.519	155.5	166.8
2·Cl^-	3.341	3.343	160.9	162.6
3·Cl^-	3.274	3.794	150.4	172.5
4·Cl^-	3.304	3.321	159.1	160.2
5·Cl^-	3.262	3.816	149.5	170.7
6·Cl^-	3.210	3.811	146.3	170.8

The conformation adopted by each transporter can be characterised by three N_{tren}–C–C–N individual torsion angles starting at the tertiary nitrogen atom. The values gathered in Table II-3 for all complexes show a *ggg* (*g* = gauche) conformation consistent with a tripodal conformational shape. The N_{tren}–C–C–N torsion angles for the **4**·Cl^- gas-phase structure (−50.2, −50.3 and −50.3°) are in excellent agreement with those found in the crystal structure (−51.4, −52.2 and −53.3°).[46] Therefore, these torsion angles will be used later to evaluate the subsequent conformational changes experienced by the transporters throughout the membrane MD simulations.

Table II-3. N_{tren}–C–C–N torsion angles (°) in gas-phase for **1-6** chloride complexes.

Chains	Transporter					
	1	**2**	**3**	**4**	**5**	**6**
1	52.9	52.8	−50.0	−50.3	50.2	45.3
2	64.7	52.8	−59.5	−50.2	59.0	59.0
3	47.8	52.8	−45.2	−50.3	45.7	49.6

II.3.2. Pure POPC bilayer membrane

The equilibration period of system A was established by monitoring the variation of the area per lipid and the related bilayer thickness throughout the course of the MD simulations. The first structural parameter was measured dividing the area of the *x*-*y* section of the system by the number of lipid molecules in each monolayer. In this work, the membrane bilayer thickness was assessed by two different methods: *a)* computing the average distance between the phosphorus atoms of each monolayer along the *z*-dimension (see Figure II-1); and *b)* measuring the distance between the peaks assigned to the phosphate head groups in the electron density profiles (see below), commonly called the D_{HH} distance.[34, 39, 40, 64] The evolution of the area per lipid and the bilayer thickness, given by the first method, along the 100 ns of MD simulation are shown in Figure II-3 for the two runs of system A. For the first run of this system, the area per lipid and bilayer thickness stabilised after the first 60 ns of simulation time, while for the second run these parameters stabilised after 40 ns of simulation. Therefore, those initial long periods were considered equilibration period and were discarded from the subsequent data analysis. The average values of both structural parameters calculated concatenating the last 40 ns of run 1 and the last 60 ns of run 2, corresponding to a total of 100 ns of sampling from these MD simulations, are gathered in Table II-4 for system A.

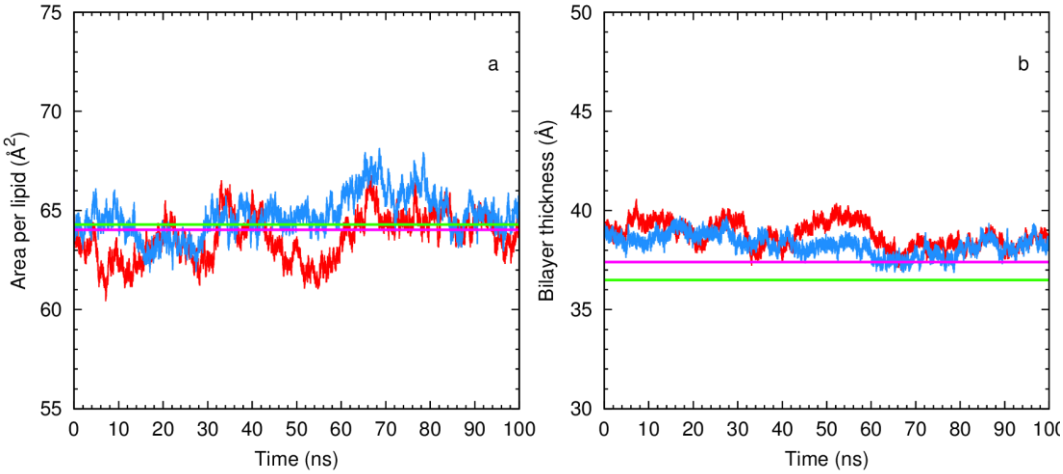


Figure II-3. Evolution of area per lipid (*a*) and bilayer thickness (*b*) in system A through the course of the MD simulation time. The reference values are plotted as green (experimental)[57] and magenta (theoretical)[42] lines, at 303 K. The red and blue lines correspond to each one of the two runs.

Table II-4. Structural parameters for area per lipid and bilayer thickness of system A, with the corresponding standard deviations, for 100 ns of sampling. Literature values are given for comparison purposes.

System	Apl (Å²)	Bt_1 (Å)	Bt_2 (Å)
A	64.86 ± 0.95	38.13 ± 0.42	36.8
Literature			
X-ray scattering	64.3[a]	–	36.5[a]
Computed	64.03 ± 1.16[b]	–	37.4[b]

Apl – Area per lipid; Bt_1 – Bilayer thickness measured between phosphorus atoms (method *a*); Bt_2 – Bilayer thickness from the electron density profiles (method *b*); [a)] ref. 57, at 303 K; [b)] ref. 42.

As mentioned above, these values were obtained with a surface tension of 17 dyn cm^{-1} , an optimal value previously reported by Skjevik *et al.*[42] to simulate a POPC bilayer with the Lipid11 force field.

It is noteworthy that the average area computed for system A, 64.86 ± 0.95 Å², is in close agreement with the experimental value of 64.3 Å² determined by X-ray scattering at 303 K.[57] Similar conclusions can be extracted with the analysis of the average bilayer thicknesses when assessed by method *a)*, as the yielded value (38.13 ± 0.42 Å) is close to the D_{HH} experimental value, 36.5 Å (5% deviation) obtained by X-ray scattering at 303 K,[57] since the D_{HH} value depends on more than just the phosphorus atoms in the lipid head groups, as demonstrated below.

The electron density profiles of the over-hydrated POPC bilayer were also evaluated, with recourse to a *ptraj* modification by Hannes Loeffler,[65] concatenating the simulation data from the last 40 ns of run 1 and the last 60 ns of run 2, thus yielding 100 ns of sampling. This parameter was computed dividing the *z*-dimension of the system into slices of 0.1 Å and calculating the number of atoms per slice, according to each atom's partial electronic charge. The normalisation of the profile was dependent on the division by *x*-*y* plane average area for the sampling time (average area per lipid of the sampling period times the number of lipids: ⟨Apl⟩×64). The electron density profile of system A is plotted in Figure II-4 with the individual density profiles for the water slabs, phospholipid bilayer and phosphate head groups represented by the phosphorus atoms. For comparison purposes, the experimental density profile for POPC obtained at 303 K by X-Ray diffraction scattering is also included (pink line in Figure II-4*a*).[57] The shape of each profile is consistent with the structure of the POPC membrane model along the *z*-dimension.

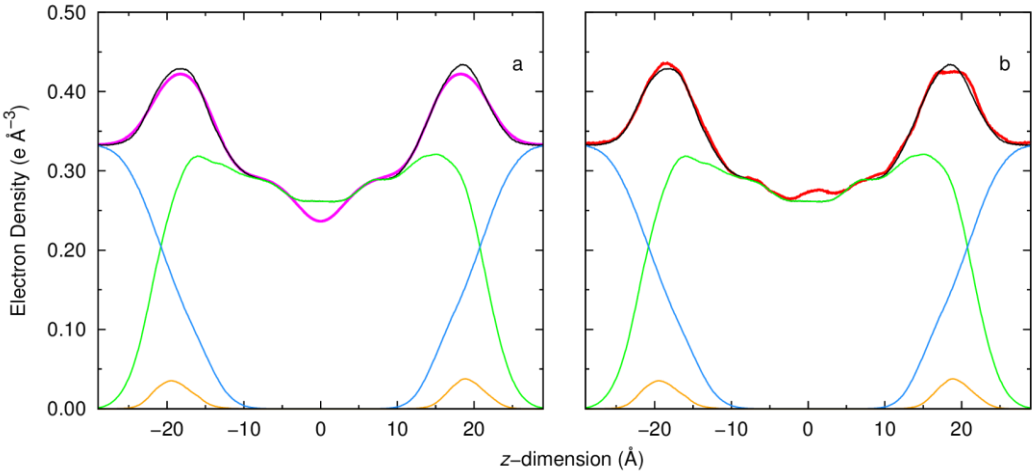


Figure II-4. Electron density profiles of system A, with the full system plotted in black, water in blue, phospholipids in green and phosphorus atoms in orange. z = 0 Å corresponds to the core of the POPC bilayer. In plot *a*, the X-ray scattering of the POPC bilayer profile is also shown as a pink line.[57] In plot *b*, the calculated full system electron density profile for the POPC bilayer on the Lipid11 paper is also shown as a red line.[42]

In fact, the electron density profile (black line) for system A has a symmetric shape around the centre of the POPC bilayer with two prominent peaks (distance between peaks: 36.8 Å, consistent with the D_{HH} 36.5 Å bilayer thickness obtained by experimental X-ray scattering at 303 K[57] and close to the 37.4 Å obtained by Skjevik *et al.*[42]), which are mainly determined by the phosphate groups and to a lesser extent by the choline moieties from the lipids and water molecules between them, as evident when this profile is compared with individual profiles of water slabs (blue line) and phospholipids (green line) and phosphorus atoms (orange line). In these circumstances, using the distance between these latter peaks, bilayer thickness was estimated at *ca.* 38.3 Å (method *b*), which is close to the value calculated using the distance between the opposite bilayer phosphorus heads throughout the course of the MD simulation (method *a*). Also, the small decrease of electron densities at z = 0 Å (the core of the bilayer) is indicative of potential intercalation between the phospholipid tails of each leaflet. This feature is also observed in the system profile reported by Skjevik *et al.* (red line in Figure II-4*b*).[42] The overall alignment of the black and pink lines in Figure II-4*a* shows a good fitting between the theoretical and experimental density profiles, while the alignment between the black and red lines in Figure II-4*b* also indicates a good fitting between the two theoretical profiles.

The ordering of the hydrophobic chains of the POPC lipid molecules, consisting on the relative orientation of the C–H bonds to the bilayer normal, was computationally estimated and compared with the experimental available deuterium order parameters, $|S_{CD}|$,[66-69] as well as with the values obtained by Skjevik *et al.*[42] The theoretical $|S_{CD}|$ values were also calculated with the *ptraj* modification by Hannes Loeffler,[65] using the same concatenated data as above through Eq. II-1:

$$|S_{CD}| = \frac{1}{2}\langle(3\cos^2\theta_i - 1)\rangle \qquad \text{Eq. II-1}$$

Where θ_i is the angle between the bilayer normal and the C_i–H bond. The identification of lipid chains and index i of the carbons in these chains are depicted in Figure II-5 (top).

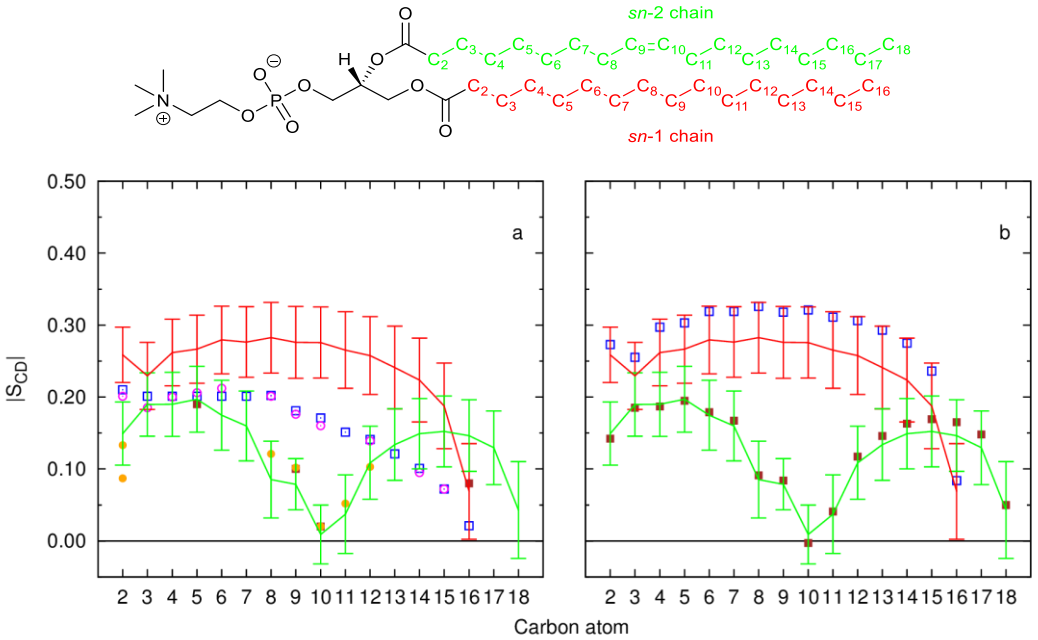


Figure II-5. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for 100 ns of sampling of system A. The $|S_{CD}|$ values calculated for the *sn*-1 chain are shown in red, while the values for the *sn*-2 chain are shown in green. The error bars associated with these results correspond to the SD. In plot *a*, the experimental values for the *sn*-1 chain were taken from refs. 68 (blue ▫) and 66 (magenta ○), while the values for the *sn*-2 chain were taken from refs. 67 (brown ▪) and 66 (orange •). In plot *b*, the computed $|S_{CD}|$ values from the Lipid11 paper[42] are presented as blue ▫ and brown ▪ for the *sn*-1 and *sn*-2 chains, respectively.

The computed $|S_{CD}|$ parameters are plotted *vs.* carbon index in Figure II-5*a* for system A, where the values for the *sn*-1 saturated and the *sn*-2 unsaturated chains are shown as red and green lines, respectively, along with the experimental values represented as points.[66-68] Regarding the experimental data, both chains display higher theoretical order parameters for the C–H bonds near the phospholipid heads. Furthermore, as would be expected, the order parameters for the *sn*-2 chain present the lowest values at the double bond between the carbon atoms C_9 and C_{10} (see Figure II-5).[66, 67] The computed *sn*-1 parameters for system A are systematically overestimated when compared with experimental ones, as expected for the Lipid11 force field.[42] Nonetheless, the inflexion of the calculated $|S_{CD}|$ values for the first two carbons on this chain matches the behaviour presented in the experimental data. In contrast with *sn*-1, the two sets of experimental data available for the *sn*-2 order parameters are composed of few points, with different values for the same carbon index (see Figure II-5), thus complicating the comparison with the computed values. Nevertheless, for system A, the theoretical values are close to the experimental ones. This is particularly evident for the C–H bonds involving the double bonded C_9 and C_{10} carbon atoms. The computed order parameters for both phospholipid aliphatic chains also follow the theoretical pattern previously reported by Skjevik *et al.*[42] (see Figure II-5*b*).

With the data above, it is demonstrated that the over-hydration of the 128 POPC lipids bilayer has a negligible effect on the structural parameters evaluated, which are mostly accurately reproduced as in the fully hydrated POPC bilayer model.[42] The small differences are ascribed to the different sampling and the effect of the surface tension on a larger periodic cell. Subsequently, this membrane model was used to investigate the passive diffusion of **1**–**6** chloride complexes across the lipid bilayer, being used as the reference for the evaluation of the structural parameters.

II.3.3. Passive diffusion of chloride complexes 1−6 from the water phase

The passive diffusion of the tripodal thiourea complexes through the POPC bilayer was evaluated by tracking their relative positions, given by the distance in the *z*-dimension (defined in Figure II-1) between the central tertiary nitrogen (N_{tren}) and the closest membrane interface (P_{int}), which is defined by the average position of the 64 phosphorus atoms in that monolayer along the *z*-dimension. The relative orientation of **1**−**6** towards the water/lipid interface was also assessed through the distances of P_{int} to the *tren* individual chains, respectively represented by the centre of mass (COM) of the four carbon atoms of each *n*-butyl chain of **1** or by the COM of the six carbon atoms of the aromatic rings of **2**−**6**. Henceforth, these four distances are designated as $N_{tren}\cdots P_{int}$ and $C_n\cdots P_{int}$, where *n* is the number of *tren* chains, having the value of 1, 2 or 3. Insights into the interaction of **1**−**6** with the POPC bilayer were obtained by the analysis of the N−H···O hydrogen bonds established between the N−H thiourea binding sites and the phosphate lipid heads or the ester group oxygen atoms of the *sn*-1 and *sn*-2 chains, using a cut-off of 3.5 Å for the N···O distances and 120° for the corresponding N−H···O angles. The putative N−H···Cl^-, N−H···OH_2 (with surrounding water molecules) and intramolecular N−H···S hydrogen bonds formed between the *tren* chains throughout the MD simulations were also evaluated using the same cut-offs.†

In the diffusion of **1**·Cl^- to **6**·Cl^- complexes from the water phase (systems B to G, as defined in Table II-1), the first event observed is the irreversible chloride release in the early stages of all MD runs, immediately after the removal of the positional restraint on the complex, which is consistent with the weak stability constants of 447, 191, 179, 128 and 156 M^{-1} measured in DMSO for the chloride complexes of **1**−**5**, in this order.[45, 46] Indeed, at this simulation stage, all complexes are quickly solvated by water molecules which substitute the chloride ion into the *tren* binding pocket as evident in Figure II-6, where the number of water molecules in the first solvation shell of the transporter is plotted throughout the simulation time. Afterwards, all small molecules migrate towards the water/lipid interface, permeating the membrane, as illustrated by the variation of $N_{tren}\cdots P_{int}$ and $C_n\cdots P_{int}$ distances for the 100 or 200 ns simulation length plotted in Figure II-7.

Hereafter, only the diffusion of **3** and **4**, the transporters with better results in the experimental studies of chloride efflux from POPC vesicles,[46] will be thoroughly discussed. However, equivalent passive diffusion behaviour was observed for the remaining transporters (**1**, **2**, **5** and **6**), as illustrated in Figure II-7, Figure II-8 and Figure A-1 (see Appendix A).

The permeation of the water/lipid interface by **3**, free of chloride, in simulation D1 begins with one fluorophenyl group (pink line in plot D1, Figure II-7) approaching the interface (*ca.* 20 ns) and then being slowly internalised for the following 30 ns. Subsequently, the transporter remains below the interface line until the end of the simulation. During the diffusion along the *z*-dimension, **3** establishes several hydrogen bonding interactions, both intra- and intermolecular, as shown in plot D1, Figure II-8. Indeed, with the release of the anion in the water phase the N−H···Cl^- interactions are replaced by N−H···S and N−H···OH_2 hydrogen bonds. As the transporter permeated through the interface, these types of interactions decreased. When **3** is below the interface, the interactions with the oxygen atoms of the POPC start replacing the intramolecular interactions as well as the interactions with the water molecules solvating the thiourea N−H binding units, as indicated in Figure II-8. This structural information is also evident from the evolution of different hydrogen bonding interactions through the simulation time plotted in Figure A-1 (see Appendix A). This sequence of events is depicted with the snapshots presented in Figure II-9.

† The distance and angle cut-offs were selected in agreement with the N−H···Cl^- dimensions found in single crystal X-ray structures deposited with CCDC[52, 53] of similar tris-thiourea chloride complexes.[46, 70-72]

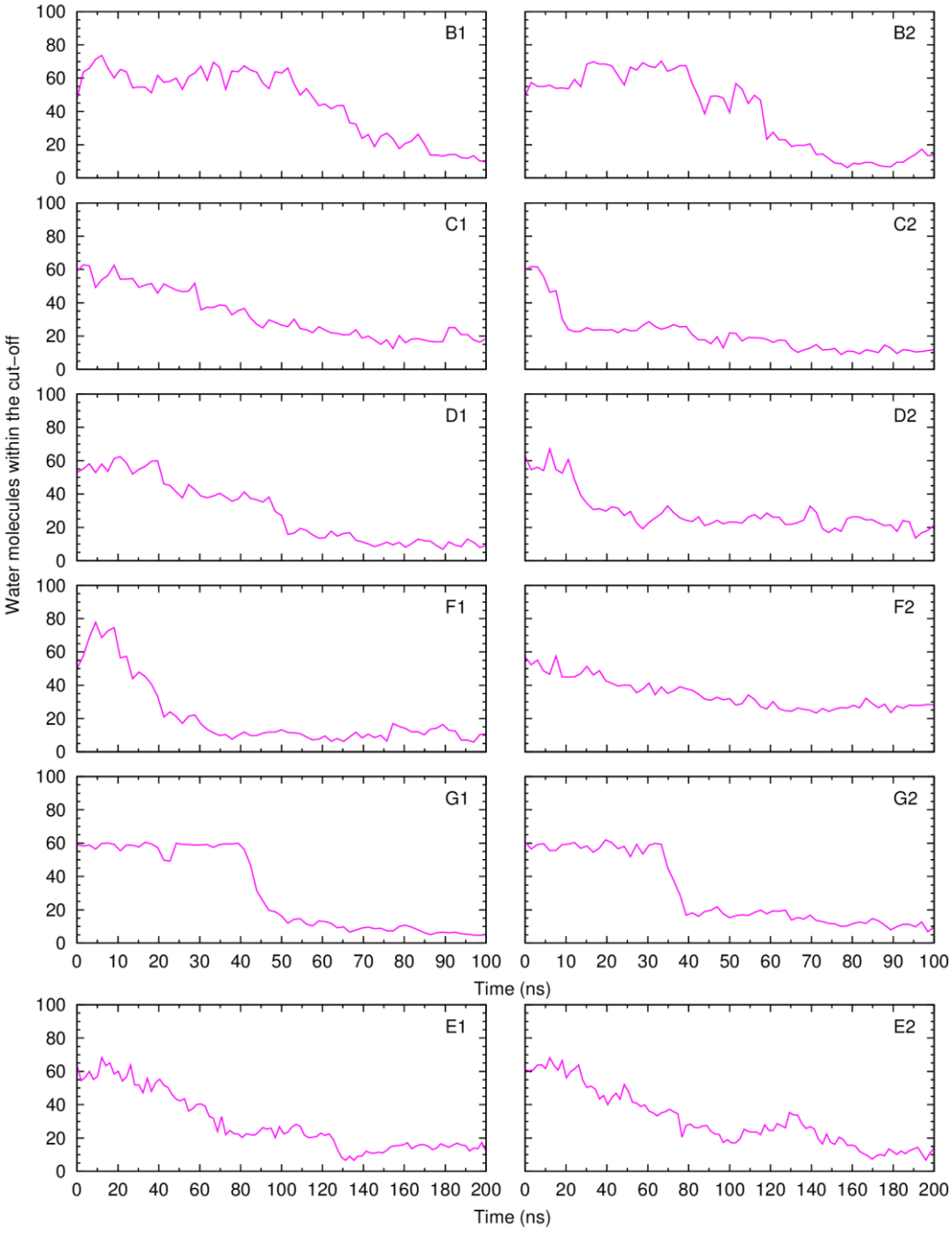


Figure II-6. Variation in the number of water molecules within the solvation shell defined by a cut-off of 3.5 Å‡ from **1** (simulations B1 and B2), **2** (simulations C1 and C2), **3** (simulations D1 and D2), **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2). Data were smoothed using Bézier curves.

‡ Typical radius for the first hydration shell of biomolecules.[73]

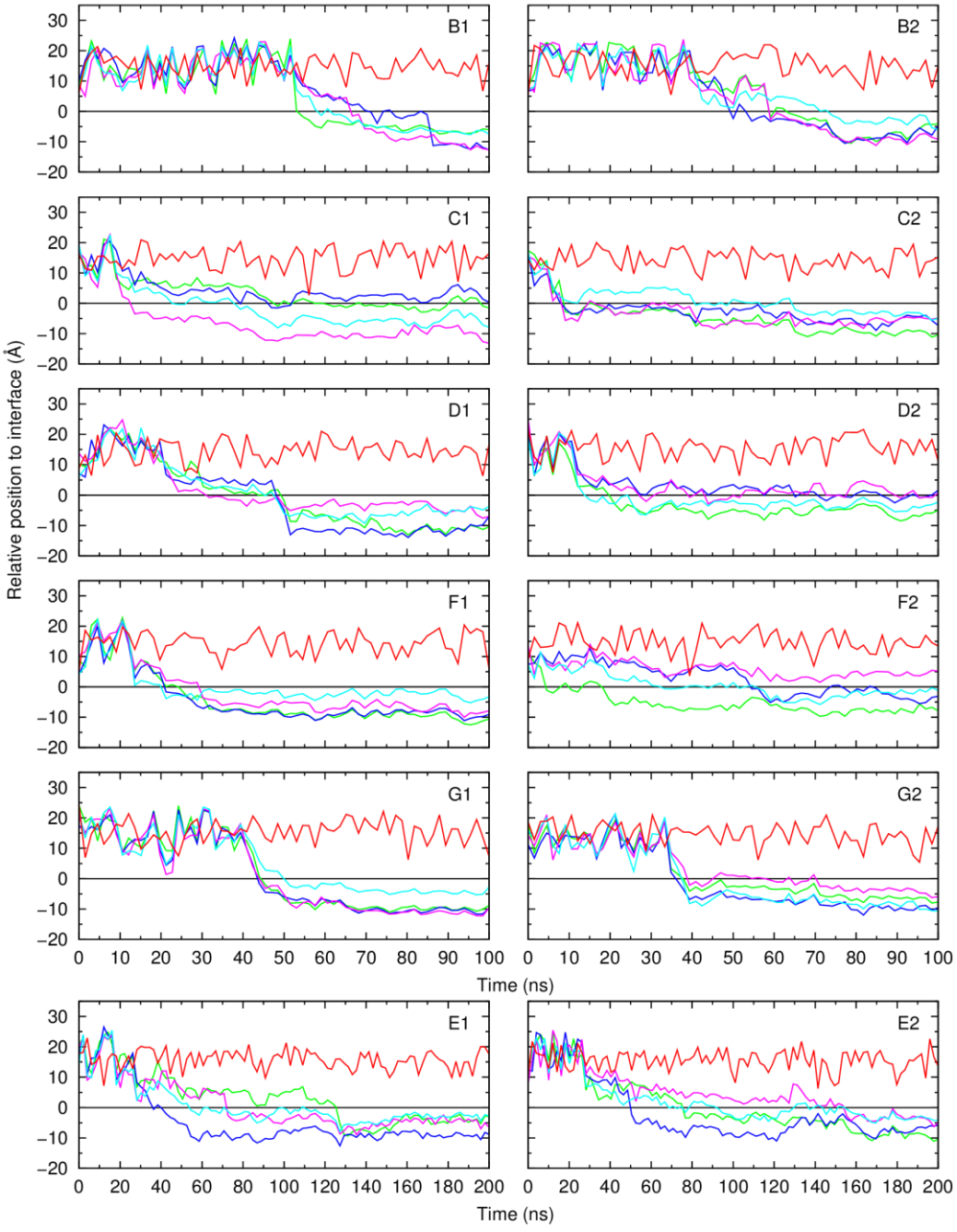


Figure II-7. Evolution of the $N_{tren}\cdots P_{int}$ and $C_n\cdots P_{int}$ distances for **1** (simulations B1 and B2), **2** (simulations C1 and C2), **3** (simulations D1 and D2), **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2). The C_1, C_2 and $C_3\cdots P_{int}$ distances are shown in dark green, light blue and pink lines, while $N_{tren}\cdots P_{int}$ is shown in cyan. The $Cl^-\cdots P_{int}$ distance is also shown, in red, while the water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

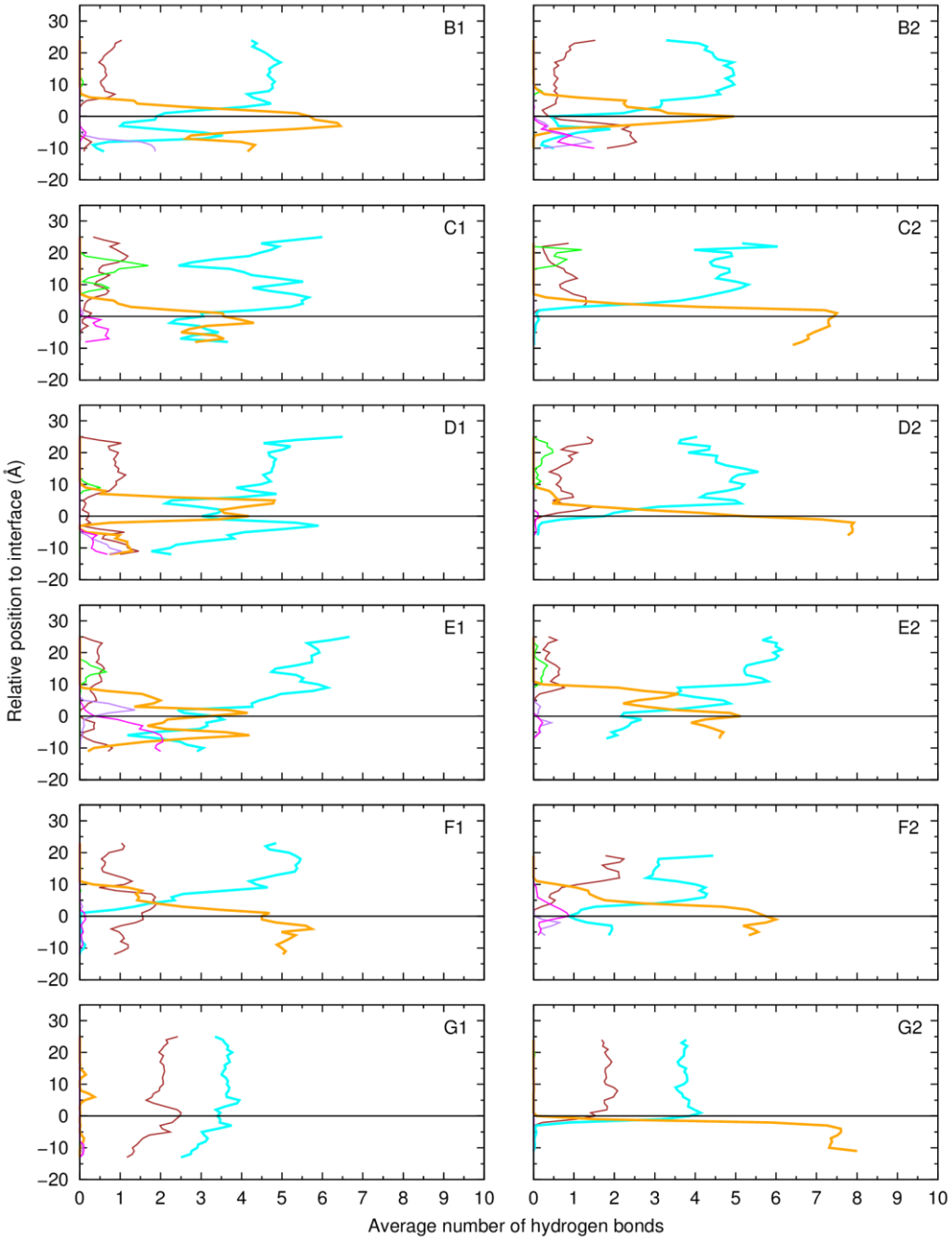


Figure II-8. Average number of hydrogen bonds *vs.* the relative position of the COM of **1** (simulations B1 and B2), **2** (simulations C1 and C2), **3** (simulations D1 and D2), **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2) to the interface. The following colour scheme was used for the interactions between the transporter and sulfur atoms (brown), chloride ions (green), water molecules (cyan), POPC head groups (orange), and ester groups (purple for the *sn*-1 chains and magenta for the *sn*-2 chains). Data were smoothed using Bézier curves.

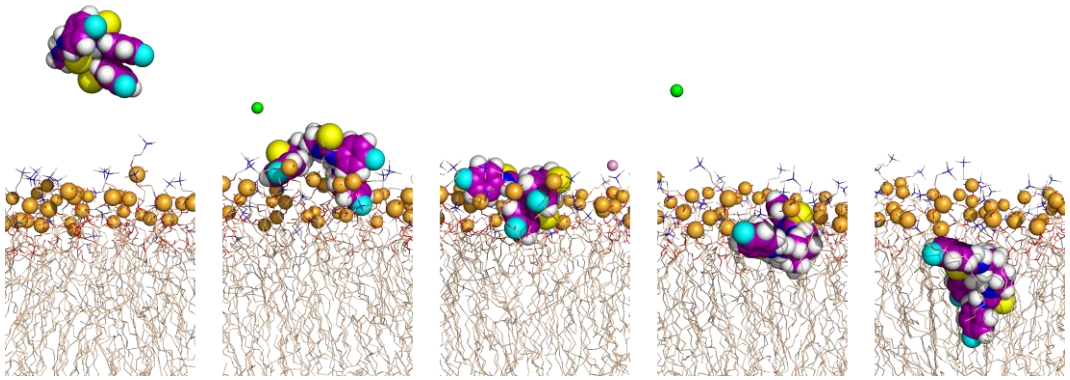


Figure II-9. The approach of **3** to the water/lipid interface of the bilayer is illustrated with five consecutive snapshots of simulation D1. The transporter and phosphorus atoms are represented in spheres with hydrogen atoms in white, oxygen atoms in red, nitrogen atoms in blue and carbon atoms in purple (transporter) or wheat (phospholipids) colour, while the chloride and sodium ions are shown as green and pink spheres, respectively. The lipids C–H atoms and the water slabs were omitted for clarity.

On the second run, simulation D2, the approach of **3** to the water/lipid interface, and its subsequent internalisation, differs from the former run. After the release of the anion in the water phase, **3** diffuses towards the interface, approaching it *ca*. 13 ns (see plot D2, Figure II-7), and remaining at the POPC head groups level until the end of simulation. As **3** diffuses along the *z*-dimension of the membrane system, it is evident that, while in the water phase, the transporter forms sporadic hydrogen bonds with chloride (see plot D2, Figure II-8), which are necessary for anion transport. However, in agreement with the low chloride association constants in competitive solvent,[45, 46] the anion capture event is unlikely in the MD simulations timescale. Furthermore, as expected, the most prominent interactions are intramolecular, with the sulfur atoms, or with the solvating water molecules. In addition, as **3** approaches the interface, the number of interactions with the oxygen atoms in the phospholipid bilayer starts to increase, while the remaining interactions decrease when the transporter is located at the interface level.

The diffusion of **4**·Cl^- along both runs (simulations E1 and E2) is comparable in the sequence of events, as shown in plots E1 and E2 of Figure II-7. After the release of the anion, the transporter diffuses towards the water/lipid interface, reaching it at 30 (E1) or 50 ns (E2) of simulation time. In both cases, the touchdown occurs through a pentafluorophenyl ring as indicated in Figure II-7 (plots E1 and E2, blue line). Likewise simulation D2, as **4** diffuses along the *z*-dimension, the most relevant hydrogen bonding interactions in the water phase are intramolecular and with the surrounding water molecules. As **4** approaches and permeates the interface, the intramolecular interactions disappear, the interactions with the water molecules wane, and concomitantly the hydrogen bonds with the lipids reach their maxima at the interface level.

In summary, the simulations performed with chloride complexes of tripodal transporters **1**–**6** initially located in the water phase show that all transporters are able to migrate to the water/lipid interface and interact with the phosphate head groups through multiple N–H···O=P hydrogen bonds. In addition, the distances between the COM of each transporter (given by all atoms) to the closest membrane interface, gathered in Table II-5 together with the time that each transporter took to achieve the interface, indicate that **1** and **6** are wholly internalised at the end of the simulation time in both runs (simulations B1, B2, G1, and G2). For transporters **2**, **3** and **5** the total internalisation was found in simulations C2, D1 and F1. By contrast, **4** remains at the interface level tightly hydrogen bonded to the surrounding phosphate head groups in both simulations of system E. It is noteworthy that the initial permeation of the bilayer was mainly achieved by the *tren* substituents, and subsequently the transporters remain at the interface or migrate towards the membrane core. However, in simulations

D2 and F1 with **3** and **5**, respectively, before the permeation of the bilayer, the transporters initially interacted with the interface *via* the *tren* head.

Table II-5. Summary of time (ns) necessary for transporters **1**–**6** to reach the interface in systems B to G, along with the average COM···P_{int} distances (Å) calculated for the last 25 ns (systems B, C, D, F and G) or 40 ns (system E) of simulation time.

System	Transporter	Run	Contact with the interface	COM···P_{int}
B	**1**	1	54	−7.11 ± 1.49
		2	50	−5.97 ± 1.23
C	**2**	1	14	−3.59 ± 1.50
		2	10	−5.66 ± 0.74
D	**3**	1	20	−7.95 ± 0.97
		2	13	−2.24 ± 1.49
E	**4**	1	30	−4.99 ± 0.78
		2	50	−5.99 ± 0.97
F	**5**	1	13	−7.54 ± 1.40
		2	4	−1.98 ± 0.83
G	**6**	1	42	−9.81 ± 0.68
		2	34	−7.02 ± 0.90

Moreover, the conformational changes experienced by **1**–**6** during the diffusion across the POPC bilayer were evaluated throughout the 100 or 200 ns simulation time using the N_{tren}–C–C–N torsion angles for the three individual *tren* chains as stated above. In the subsequent analysis, three torsion angles within the −70 to 70° range correspond to a *ggg* conformation consistent with a tripodal shape, whereas angle values outside such range lead to distinguishable *ggt*, *gtt* and *ttt* (*t* = trans) conformations indicative of non-tripodal conformations. Permutations in the sequence of these three torsion angles, as well as signal changes, are irrelevant for the description of the tripodal conformation shape and were not taken into account in the conformation definition. Furthermore, the conformations explored by **3** and **4** whilst in the water phase and after reaching the water/lipid interface are clearly accessed by plotting the distribution of N_{tren}–C–C–N torsion angles in histograms for both periods, as depicted in Figure II-10 and Figure II-11. The variations in N_{tren}–C–C–N torsion angles for all transporters (systems B to G) are plotted in Figure A-2 (see Appendix A). For clarity reasons, negative N_{tren}–C–C–N torsion angles, smaller than −120°, were shifted in the plots by the addition of 360° in Figure II-10, Figure II-11 and Figure A-2 (see Appendix A).

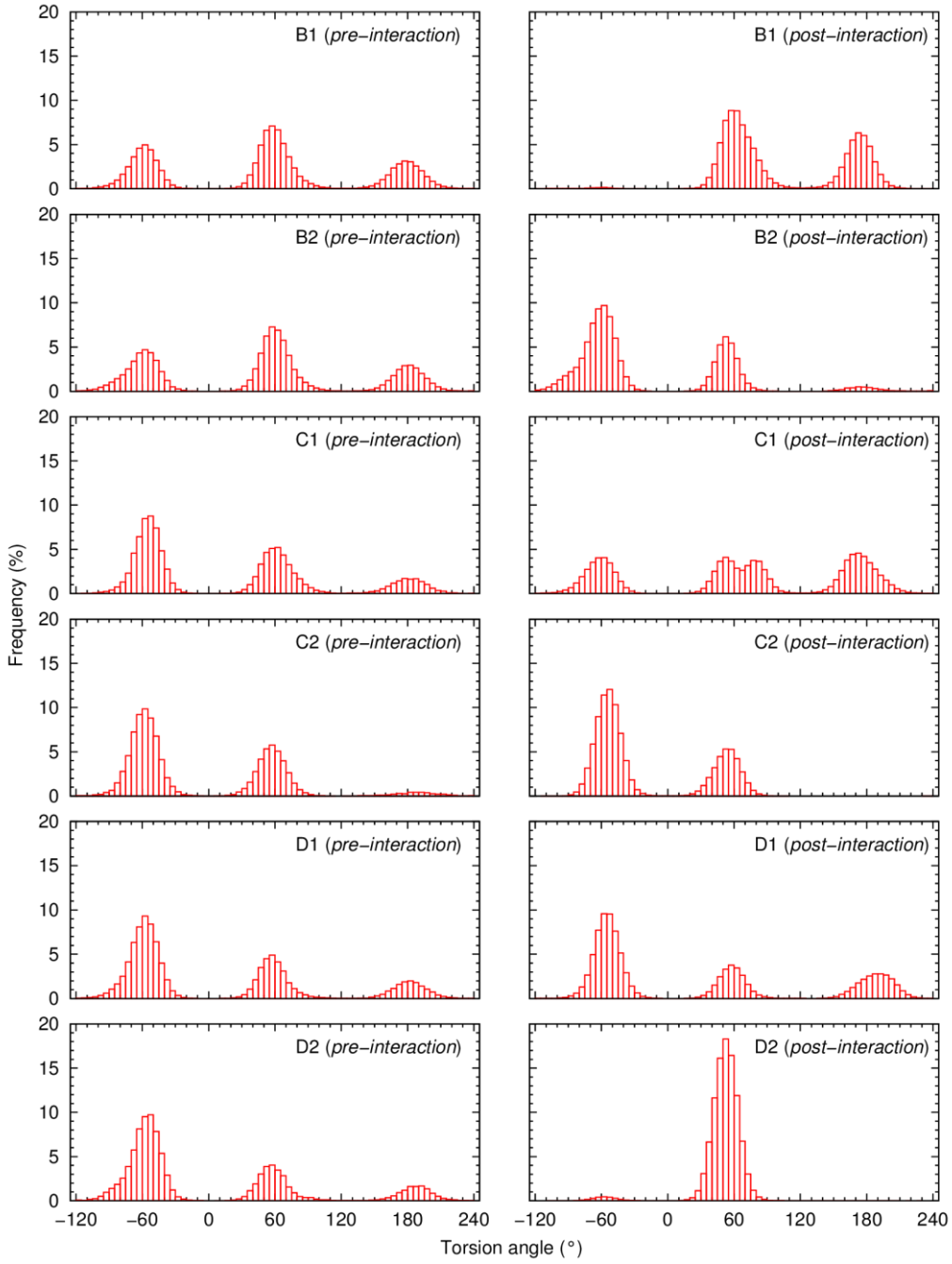


Figure II-10. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angle values of **1** (simulations B1 and B2), **2** (simulations C1 and C2), and **3** (simulations D1 and D2), before the interaction with the interface (*pre-interaction*) and after (*post-interaction*).

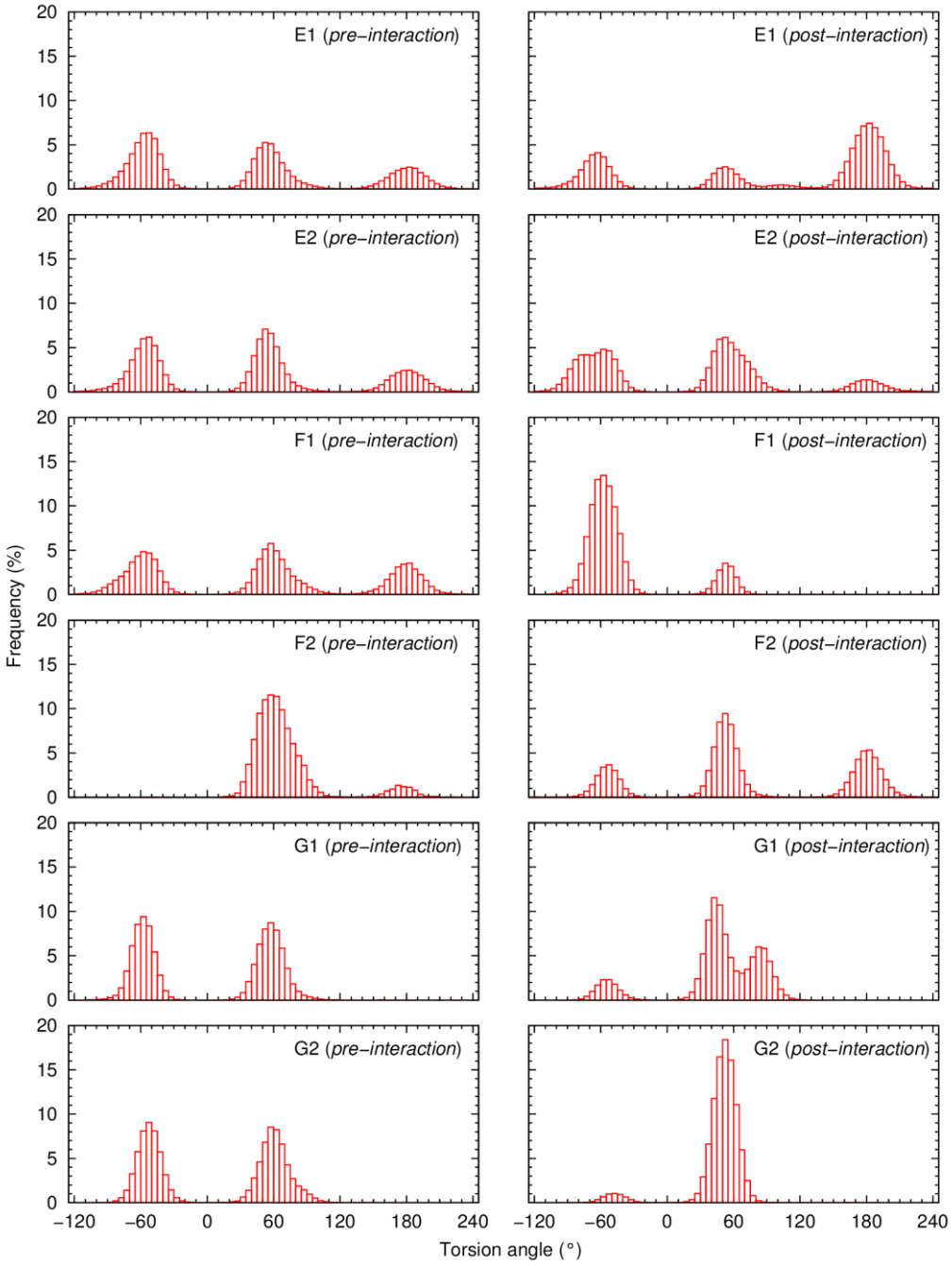


Figure II-11. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angle values of **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2), before the interaction with the interface (*pre-interaction*) and after (*post-interaction*).

Due to the similarities between the plots corresponding to the diffusion of **3** and **4** in the water phase of systems D and E, they are analysed together, whilst the plots concerning the remaining period of simulation after the contact with the water/lipid interface are discussed in a clustered fashion. Regardless of the period of time spent diffusing in the water phase, transporters **3** and **4** loose the *ggg* conformation consistent with the tripodal shape as evident in the *pre-interaction* plots D1, D2, E1 and E2 presented in Figure II-10 and Figure II-11. These histograms reveal that the N_{tren}–C–C–N torsion angles for the three *tren* chains experienced mostly values around −60 and 60° and to lesser extent values around 180°. Furthermore, the individual chain contribution for this distribution is achieved

assessing the variation of N_{tren}–C–C–N torsion angles throughout the simulation time presented in Figure A-2 (see Appendix A). Concerning the histograms *post*-interface *interaction*, the simulations can be divided in two groups: *a)* simulation D2 and *b)* the remaining simulations (D1, E1 and E2). In simulation D2, the distribution of torsion angles is mainly around 60°, with some records of values around −60°, consistent with a *ggg* conformation observed during most of the simulation time, as can be seen in Figure A-2 (see Appendix A). In the remaining *post-interaction* plots of D1, E1 and E2 (see Figure II-10 and Figure II-11), the values outside the range −70 to 70° were observed, therefore, even if the *ggg* conformation occurred after the interaction with the water/lipid interface, transporters **3** and **4** were mostly in non-tripodal conformations, as indicated by the variations in N_{tren}–C–C–N angles along the simulation time presented in Figure A-2 (see Appendix A). Overall, the analysis of N_{tren}–C–C–N torsion angles for transporters **1**–**5** shows that the *ggg* conformation adopted in the chloride complexes is lost with the anion release and in the subsequent events, including the diffusion through the water phase and subsequent interaction with the membrane interface, the transporters mainly adopt the *ggt* conformation, with a non-tripodal shape. The unique exception is **6** with three 3,5-bis(trifluoromethyl)phenyl substituents, for which the *ggg* conformation is retained in both phases of system G (see Figure A-2 in Appendix A).

Assessing the variation of the *z*-component of the dipole moment of a potential transmembrane transporter is a clear indicator of its orientation in the system.[32] This type of analysis was not sought for **1**–**6**, due to the conformational changes experimented by the transporters before the interaction with the water/lipid interface. Even after this event, as the tripodal shape was not always present, this sort of analyses was deemed irrelevant.

II.3.4. Passive diffusion of the chloride complex of 3 from the bilayer core

The passive diffusion of chloride complexes was also investigated with **3**·Cl^-, positioning its lowest energy binding arrangement at the core of the phospholipid bilayer, affording system H (see Table II-1).

Transporter **3** spontaneously migrates to the water/lipid interface releasing the anion to the aqueous phase when near the phospholipid head groups, as evident in Figure II-12. The diffusion of the complex from the core of the bilayer is illustrated by the sequence of snapshots given in Figure II-13.

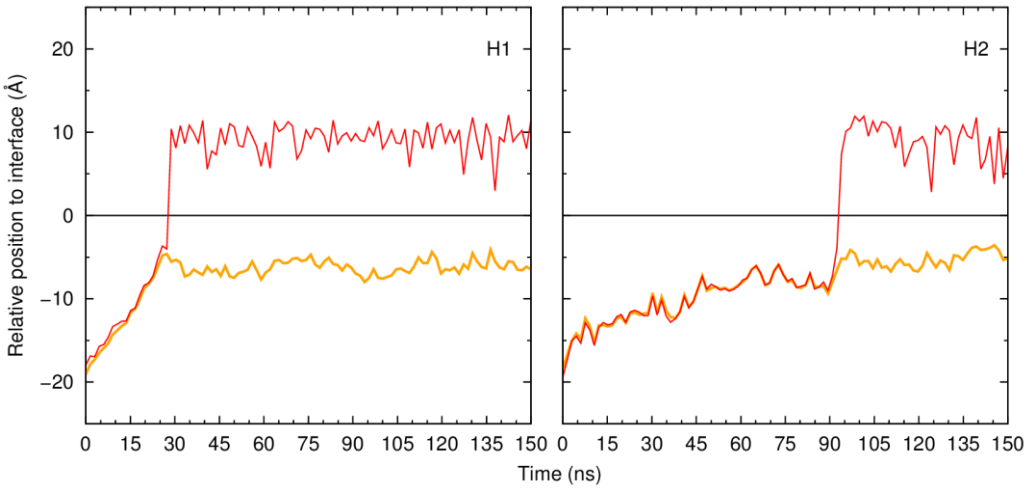


Figure II-12. Evolution of COM···P_{int} distances for **3**, as well as Cl^-···P_{int}, in simulations H1 and H2, during the 150 ns of simulation time. COM···P_{int} is shown in orange and Cl^-···P_{int} is shown in red, while the water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

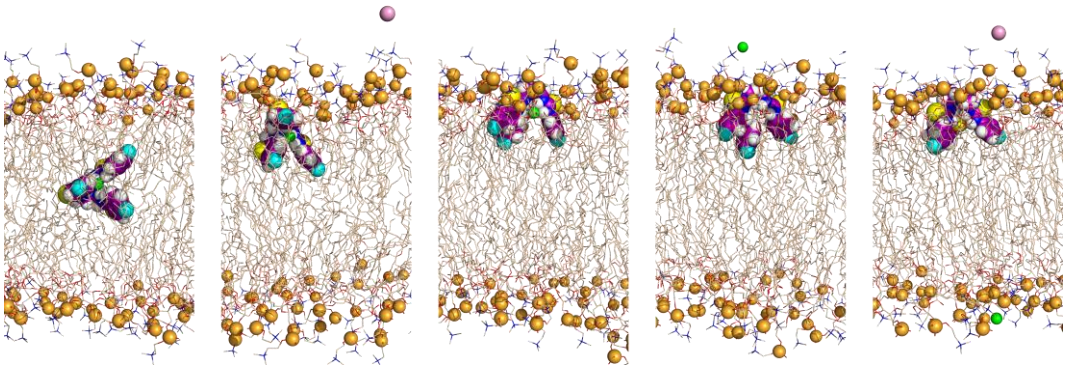


Figure II-13. The diffusion of **3** from the bilayer core and the release of the chloride ion is illustrated with five consecutive snapshots of simulation H1. Remaining details as given in Figure II-9.

3·Cl^- remains among the lipid tails during significant periods of simulation, 28 (H1) or 93 ns (H2). During this period, the complex diffuses to the water/lipid interface, releasing the chloride ion when it is at approximately 5 (H1) or 10 Å (H2) from the interface. The diffusion process is accompanied by a concomitant increase of the number of water molecules around the complex, which promotes the delivery of the anion to the aqueous phase, as indicated by Figure II-14. In addition, when the complex approaches the interface, the N–H···Cl^- hydrogen bonds are replaced by N–H···OH_2 or N–H···O=P and intramolecular N–H···S hydrogen bonds as depicted in Figure II-15 and Figure A-3 (see Appendix A), where the evolution of hydrogen bond counting is presented. As seen in Figure II-15, the interaction with the surrounding water molecules is more prominent in simulation H1, while hydrogen bonding interactions with phosphate head groups are more common in H2.

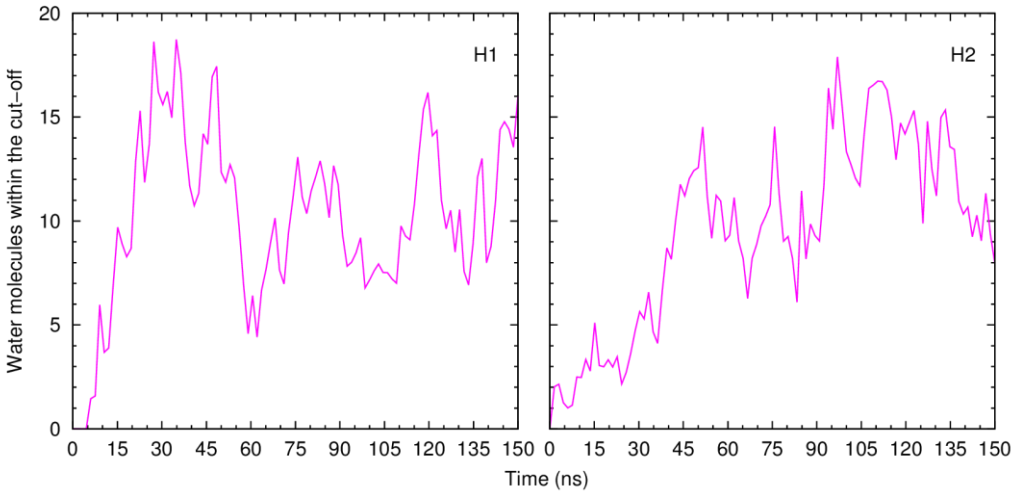


Figure II-14. Variation in the number of water molecules within the solvation shell defined by a cut-off of 3.5 Å in simulations H1 and H2. Data were smoothed using Bézier curves.

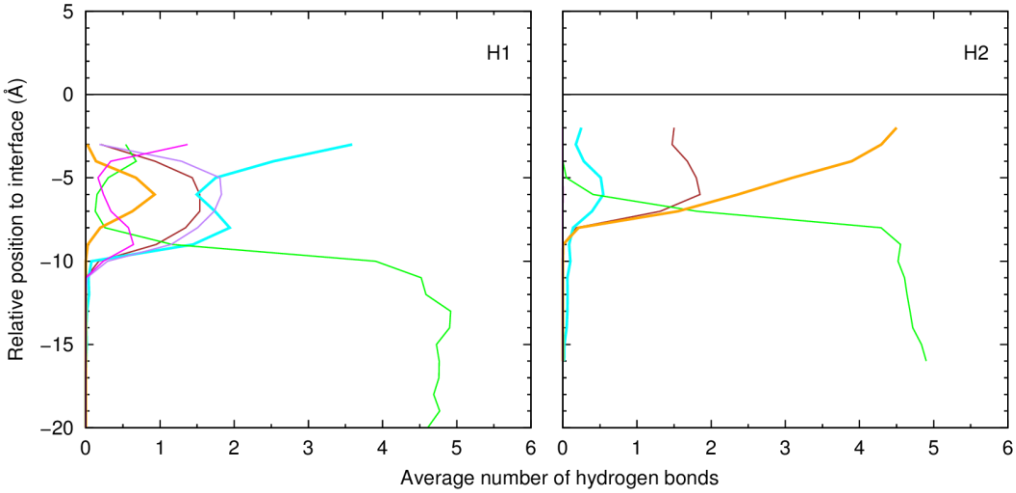


Figure II-15. Average number of hydrogen bonds *vs*. the relative position of the COM of **3** in simulations H1 and H2. Remaining details as given in Figure II-8.

During the last 50 ns of simulation, the COM···P_{int} distances for **3** were −6.04 ± 0.91 (H1) and −5.30 ± 0.99 Å (H2), indicating that the transporter diffused from the core of the bilayer, but not enough to achieve the water phase, in spite of its hydration (see Figure II-14).

The conformational freedom of the transporters inside the POPC bilayer is limited by the highly packed medium composed of hydrophobic lipid tails. Therefore, throughout the MD runs of system H, transporter **3** mostly preserves its initial *ggg* conformation, sporadically interrupted by *ggt* conformations, as inferred from the distribution of the N_{tren}–C–C–N torsion angle values of **3**, shown in Figure II-16, and their evolution through the 150 ns simulation time, as depicted in Figure A-4 (see Appendix A).

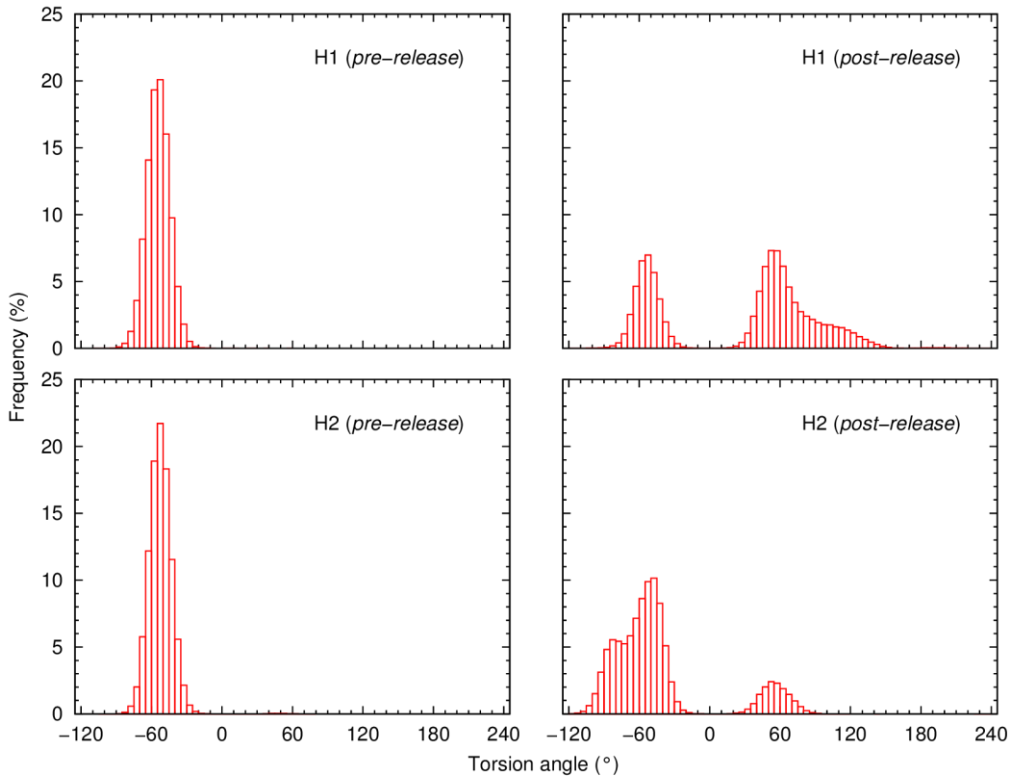


Figure II-16. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angle values of **3** in simulations H1 and H2, before the release of the anion (*pre-release*) and after (*post-release*).

The simulation findings aforementioned for system H indicate that the diffusion of the chloride complexes starting from the bilayer core is a likely event, which is consistent with the anion carrier mechanism as suggested by the experimental data.[45, 46] To obtain further theoretical insights into this hypothesis, a new collection run (simulation H') was carried out applying a positional restraint of 50 kcal mol^{-1} $Å^{-2}$ on the central tripodal nitrogen atom (N_{tren}) for 150 ns. The evolution of the relative position of chloride and COM of **3** throughout the course of the runs H'1 and H'2, given in Figure II-17, shows that the chloride association is maintained and remains at the core of bilayer. In other words, this result indicates that the receptor assists the chloride transmembrane transport as a shuttle like transporter.

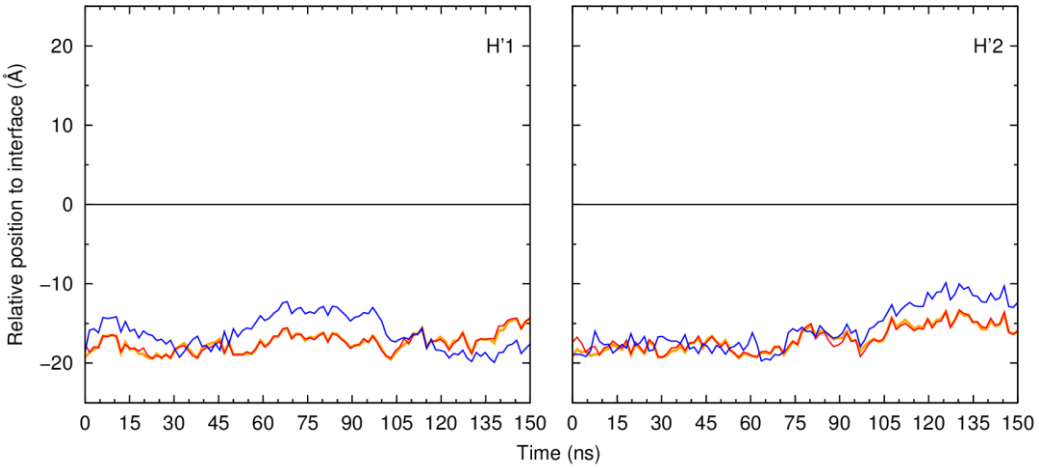


Figure II-17. Evolution of COM···P_{int} distances for **3**, Cl^-···P_{int}, and N_{tren}···P_{int} in simulations H'1 and H'2, during the 150 ns of simulation time. COM···P_{int} is shown in orange, Cl^-···P_{int} is shown in red, and N_{tren}···P_{int} is shown in blue, while the water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

II.3.5. Passive diffusion of the restrained chloride complex of 4 from the water phase

The MD simulations of systems B–G (*vide supra*) indicate that the thiourea-based receptors **1**–**6**, when initially positioned in the water phase, release the chloride ion before the corresponding complex achieves the water/lipid interface due to the prompt solvation of the anion by water molecules. On the other hand, the experimental transport studies also indicate that these molecules, in particular **4**,[46] are able to uptake an anion from water and subsequently transport it across the POPC bilayer. Indeed, the tripodal shape shields the anion from the water molecules and is able to bind the anion through the six N–H binding units, as shown by the X-ray single crystal structures of **3**·Cl^-, **4**·Cl^- and **9**·Cl^- (CCDC RefCodes SAKBUD, SAJZEK, and SAKCUE, in this order).[46] To ascertain the parameters that could allow us to reproduce similar behaviour in membrane simulations, preliminary MD simulations were carried out in water solution for chloride complexes of **1**–**6**. A torsional restraint of only 10 kcal mol^{-1} rad^{-2} was applied to the individual N_{tren}–C–C–N torsion angles, preserving the tripodal conformation, as described in Section A.4. of Appendix A. Although the conformational shape is kept, the anion is promptly solvated by the surrounding water molecules, definitively leaving the binding pocket, as shown in Figure A-5 and Figure A-6 (see Appendix A), in which the evolution of the distance from the receptors' COM to the anion and the inherent hydrogen bonds are depicted, respectively. This indicates that the tripodal shape is not sufficient to maintain the association in the water phase. It is important to note that this behaviour is not an artefact caused by hydrogen bonds in GAFF being mainly treated *via* electrostatic interactions, but derives from the low affinity constants of **1**–**5** for chloride, in competitive solvents such as DMSO (*vide supra*).[45, 46]

Therefore, in a subsequent approach, the torsion angle restraints were complemented with a single distance restraint between an N–H binding unit and the anion (see Section A.4. of Appendix A) to allow the chloride uptake and release in the simulation timescale. Several values were tested, and these events can be intermittently observed throughout the simulation course using a weak force constant of only 1.0 kcal mol^{-1} $Å^{-2}$ as depicted in Figure A-7 and Figure A-8 for complexes of **1**–**6** (see Appendix A).

Afterwards, these restraints were applied to **4**·Cl^- embedded in the water phase of the POPC bilayer (system E') and two new collection runs were performed. In both runs, E'1 and E'2, the anion complex is maintained throughout most of the simulation length, regardless of its location in the

membrane system, as seen in Figure II-18. Furthermore, when the complex crosses the water/lipid interface with the lipophilic C_6F_5 substituents first, the tripodal shape is unchanged (as seen in Figure II-19). In addition, the binding of chloride by the six N–H binding sites prevents these groups from interacting with the phosphate head groups, as observed in the unrestrained simulations of system E (for comparison purposes see Figure II-20 and Figure II-8).

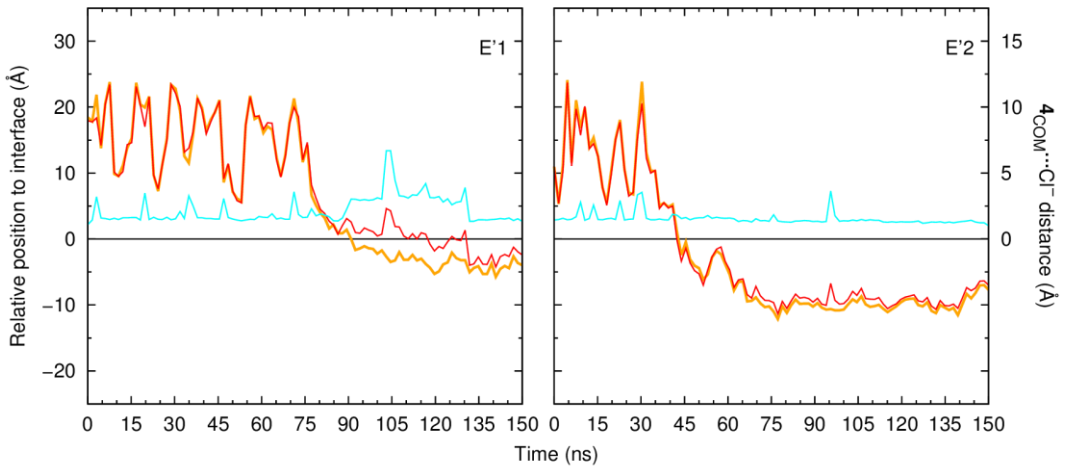


Figure II-18. Evolution of COM···P_{int} distances for **4** and Cl^-···P_{int}, in simulations E'1 and E'2, compared to the evolution of $\mathbf{4}_{COM}$···Cl^- distances during the 150 ns of simulation time. COM···P_{int} is shown in orange, Cl^-···P_{int} is shown in red and $\mathbf{4}_{COM}$···Cl^- is shown in cyan, while the water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

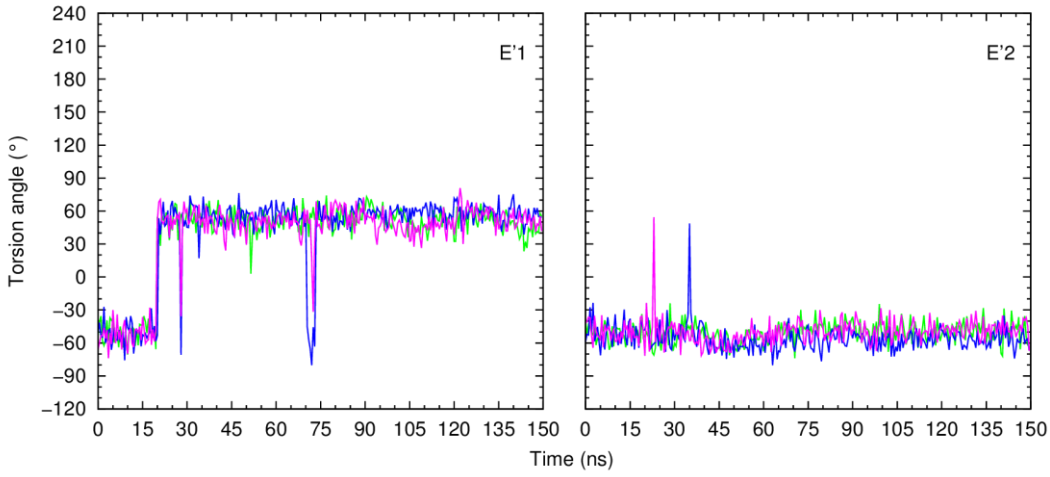


Figure II-19. Variations in the three N_{tren}–C–C–N tripodal torsion angles for 150 ns of MD simulations E'1 and E'2, extracted every 500 ps, for both runs, containing **4**·Cl^-.

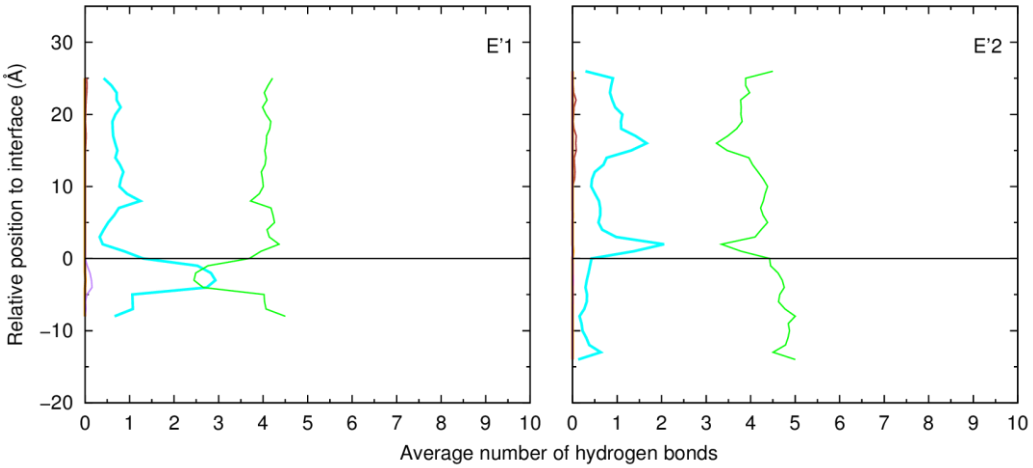


Figure II-20. Average number of hydrogen bonds *vs.* the relative position of the COM of **4** (simulations E'1 and E'2). Remaining details as given in Figure II-8.

In summary, these theoretical findings show that the simulation of an anion carrier transport mechanism using shuttle-like molecules, with low affinity for anions in the water phase, might require the use of weak restraints such as the ones used in simulation E'.

II.3.6. Impact of 1–6 in the POPC bilayer

The impact caused by **1**–**6** in the membrane structure was estimated for systems B to H through the evaluation of the corresponding area per lipid, bilayer thickness, electron density profiles and chain order parameters for the last 25 ns of each MD simulation in systems B, C, D, F and G, the last 40 ns for system E, and the last 50 ns in system H. These parameters were not assessed for systems E' and H'. The average area per lipid and bilayer thickness values found for each simulation are listed in Table II-6, while the electron density profiles and order parameters are shown in Appendix A, from Figure A-9 to Figure A-12. The structural parameters of systems B–G are compared with those obtained for the free POPC membrane, system A, while the parameters of the smaller system H are compared with data available from the Lipid11 original reference.[42]

Table II-6. Comparison between the area per lipid and bilayer thickness of systems A to H for the last 25 (systems B, C, D, F and G), 40 (system E) or 50 ns (system H) of simulation.

System	Area per lipid (Å²)		Bilayer thickness[a] (Å)	
	Run 1	Run 2	Run 1	Run 2
A[b]	64.86 ± 0.95	–	38.13 ± 0.42	–
B	64.80 ± 0.85	64.33 ± 1.23	38.15 ± 0.43	38.49 ± 0.53
C	63.81 ± 1.16	67.12 ± 0.95	38.70 ± 0.54	37.05 ± 0.49
D	65.25 ± 0.68	62.36 ± 1.16	37.70 ± 0.41	39.33 ± 0.50
E	65.38 ± 0.70	64.72 ± 0.63	37.42 ± 0.32	37.88 ± 0.35
F	64.99 ± 0.57	66.17 ± 0.69	37.95 ± 0.29	37.56 ± 0.30
G	60.14 ± 0.63	64.96 ± 1.23	40.90 ± 0.31	38.18 ± 0.59
Lipid11[c]	64.03 ± 1.16	–	37.4[d]	–
H	67.45 ± 0.99	66.16 ± 1.79	37.30 ± 0.86	36.61 ± 0.43

[a] Measured between phosphorus atoms; [b] Systems B–G are compared with system A; [c] System H is compared with data from ref. 42; [d] Value obtained measuring the distance between the peaks in the electron density profile.

Transporters **1**–**6** have a negligible impact on the area per lipid and bilayer thickness parameters. Indeed, the largest average differences assessed relatively to system A were caused by **3** and **6**, with only 7 (G1) and 4% (D2) decrease in area per lipid. Concomitant slight increases of 7 (G1) and 3% (D2) were observed in the average bilayer thickness.

The electron density profiles of systems B to G were obtained from the last 25 ns of MD collection data (apart of system E, 40 ns) and are plotted in Figure A-9 (see Appendix A), which also include the electron density profile for the corresponding transporter. Overall, there is good alignment between each system profile (black line) and the free membrane profile (pink line), system A, indicating that the bilayer structure in most systems is preserved. In fact, detectable misalignments were apparent only for simulations D2 and G1 with transporters **3** and **6** in agreement with the corresponding bilayer thickness differences found.

In line with the previous structural findings, the order parameters of these six systems seem mostly unaffected by the presence of **1**–**6**, as shown in Figure A-10 for systems B to G (see Appendix A). In most runs, a small ordering increase is observed for the *sn*-1 chains when compared with the *sn*-1 order parameters of the free membrane system. This effect is more pronounced in simulation E1, ran for 200 ns. The slight ordering trend is also apparent for the *sn*-2 chains and a slight intercalation between the phospholipid tails of each leaflet is observed, as found for system A, in agreement with the corresponding electron density profiles. However, it is important to note that the computed *sn*-1 parameters for POPC are systematically overestimated when simulated with the Lipid11 force field.[42]

Concerning the simulation of system H, in which **3**·Cl^- was initially positioned at the bilayer core, the impact of the presence of this transporter between the phospholipids was evaluated for the last 50 ns and compared with structural data from the Lipid11 paper.[42] As previously reported for systems B–G, the impact on the bilayer structural properties was also marginal, in agreement with the results of system D, as seen in Table II-6 (area per lipid and bilayer thickness) and in Figure A-11 and Figure A-12 (electron density profile and order parameters, see Appendix A). Therefore, the slight changes found for these biophysical parameters mainly reflect the necessary adjustments of the bilayer for the permeation and diffusion of transporters **1**–**6** in systems B–H.

II.4. Conclusions

In this work it was demonstrated that increasing the solvation of a 128 POPC bilayer has a marginal impact on the structure of this membrane model, which enabled further investigation of the diffusion and interaction of **1**–**6**·Cl^- in this large POPC bilayer system. The work herein reported portrayed an attempt to better understand the nature of the interaction between transporters **1**–**6** and the water/lipid interface of a POPC lipid membrane.

The unrestrained MD simulations showed that the six synthetic molecules are able to permeate the membrane without causing the disruption of the membrane structure, thus having a small impact on the biophysical parameters assessed. All transporters migrate spontaneously to the water/lipid interface penetrating the membrane through the phospholipids, regardless of their initial positions in the POPC bilayer system. Compounds **1**–**6** are preferentially located near the polar phospholipid head groups interacting with them *via* multiple N–H···O hydrogen bonds. Furthermore, when the chloride complex of **3** is initially positioned in the middle of the membrane core, the anion is released to the water phase upon approach of the complex to the interface. Also, partially restraining **3** within the bilayer prevents the diffusion of the anion complex and the release of the anion, showing that the receptor is needed for the anion transport inside the membrane. Furthermore, when the anion complexes of **1**–**6** are initially positioned in the water phase the chloride ion quickly leaves the tripodal binding pocket, being irreversibly solvated by the water molecules. However, when the N_{tren}–C–C–N chain torsion angles

and a single N–H···Cl^- hydrogen bond are appropriately restrained *via* weak force constants, the chloride complex is maintained and is able to permeate the bilayer. Albeit the uptake of chloride from the water phase, followed by the transmembrane transport and subsequent release in the water phase on the other side of the POPC bilayer was not observed, the computational findings reported partially corroborate the mobile carrier mechanism of the chloride synthetic transporters, as previously suggested by the experimental efflux studies.[45, 46] Indeed, the occurrence of this sequence of events is extremely unlikely on the MD simulation time scale, requiring the use of some biasing potential, eventually coupled with an artificial concentration gradient of chloride ions across the membrane, to prompt these events. Our results show that these transporters appear to have similar diffusion behaviour and a marginal effect on the bilayer structural properties, as demanded for a synthetic transmembrane anion carrier. The different transport abilities of these molecules are eventually related to the energetic barrier that each molecule needs to overcome across the bilayer.

II.5. References

1. F. M. Ashcroft, *Ion Channels and Disease*, Academic Press, London, 1st edn., 2000.
2. S. C. Cannon, *Neurotherapeutics*, 2007, **4**, 174-183.
3. R. Planells-Cases and T. J. Jentsch, *Biochim. Biophys. Acta*, 2009, **1792**, 173-189.
4. J. T. Davis, O. Okunola and R. Quesada, *Chem. Soc. Rev.*, 2010, **39**, 3843-3862.
5. H. Lodish, A. Berk, P. Matsudaira, C. A. Kaiser, M. Krieger, M. P. Scott, L. Zipursky and J. Darnell, *Molecular Cell Biology*, W. H. Freeman, 5th edn., 2003.
6. W. F. Boron and E. L. Boulpaep, *Medical Physiology - Updated Edition*, Saunders, 1st edn., 2004.
7. G. Bernard and M. I. Shevell, *Pediatr. Neurol.*, 2008, **38**, 73-85.
8. F. Lehmann-Horn and K. Jurkat-Rott, in *Wiley Encyclopedia of Chemical Biology*, John Wiley & Sons, Inc., 2008, pp. 1201-1210.
9. W. W. Shum, Y. C. Ruan, N. Da Silva and S. Breton, *J. Androl.*, 2011, **32**, 576-586.
10. B. Rode, T. Dirami, N. Bakouh, M. Rizk-Rabin, C. Norez, P. Lhuillier, P. Lores, M. Jollivet, P. Melin, I. Zvetkova, T. Bienvenu, F. Becq, G. Planelles, A. Edelman, G. Gacon and A. Toure, *Hum. Mol. Genet.*, 2012, **21**, 1287-1298.
11. E. J. Jeong, E. J. Kang, L. T. Sung, S. K. Hong and E. Lee, *J. Am. Chem. Soc.*, 2002, **124**, 14655-14662.
12. T. R. Sheth, R. M. Henderson, S. B. Hladky and A. W. Cuthbert, *Biochim. Biophys. Acta*, 1992, **1107**, 179-185.
13. S. Ohkuma, T. Sato, M. Okamoto, H. Matsuya, K. Arai, T. Kataoka, K. Nagai and H. H. Wasserman, *Biochem. J*, 1998, **334**, 731-741.
14. A. P. Davis, D. N. Sheppard and B. D. Smith, *Chem. Soc. Rev.*, 2007, **36**, 348-357.
15. J. T. Davis, P. A. Gale, O. A. Okunola, P. Prados, J. C. Iglesias-Sanchez, T. Torroba and R. Quesada, *Nat Chem*, 2009, **1**, 138-144.
16. P. A. Gale, *Chem. Commun.*, 2011, **47**, 82-86.
17. P. A. Gale, *Coord. Chem. Rev.*, 2001, **213**, 79-128.
18. P. A. Gale, *Coord. Chem. Rev.*, 2003, **240**, 191-221.
19. P. A. Gale and R. Quesada, *Coord. Chem. Rev.*, 2006, **250**, 3219-3244.
20. P. A. Gale, S. E. Garcia-Garrido and J. Garric, *Chem. Soc. Rev.*, 2008, **37**, 151-190.
21. C. Caltagirone and P. A. Gale, *Chem. Soc. Rev.*, 2009, **38**, 520-563.
22. P. A. Gale, *Chem. Soc. Rev.*, 2010, **39**, 3746-3771.
23. S. Matile, A. Vargas Jentzsch, J. Montenegro and A. Fin, *Chem. Soc. Rev.*, 2011, **40**, 2453-2474.
24. P. A. Gale, *Acc. Chem. Res.*, 2011, **44**, 216-226.
25. M. Wenzel, J. R. Hiscock and P. A. Gale, *Chem. Soc. Rev.*, 2012, **41**, 480-520.
26. B. A. McNally, W. M. Leevy and B. D. Smith, *Supramol. Chem.*, 2007, **19**, 29-37.
27. D. J. dos Santos, P. Saenz-Mendez, L. A. Eriksson and R. C. Guedes, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10174-10182.
28. R. O. Dror, M. O. Jensen, D. W. Borhani and D. E. Shaw, *J. Gen. Physiol.*, 2010, **135**, 555-562.
29. C. Wei and A. Pohorille, *J. Phys. Chem. B*, 2011, **115**, 3681-3688.
30. C. J. E. Haynes, S. J. Moore, J. R. Hiscock, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2012, **3**, 1436-1444.
31. S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103-117.
32. P. J. Costa, I. Marques and V. Felix, *Biochim. Biophys. Acta*, 2014, **1838**, 890-901.
33. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, Jr. and R. W. Pastor, *J. Phys. Chem. B*, 2010, **114**, 7830-7843.

34. D. Poger and A. E. Mark, *J. Chem. Theory Comput.*, 2010, **6**, 325-336.
35. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812-7824.
36. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell, Jr., *J. Comput. Chem.*, 2010, **31**, 671-690.
37. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
38. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
39. L. Rosso and I. R. Gould, *J. Comput. Chem.*, 2008, **29**, 24-37.
40. S. W. Siu, R. Vacha, P. Jungwirth and R. A. Bockmann, *J. Chem. Phys.*, 2008, **128**, 125103.
41. B. Jojart and T. A. Martinek, *J. Comput. Chem.*, 2007, **28**, 2051-2058.
42. A. A. Skjevik, B. D. Madej, R. C. Walker and K. Teigen, *J. Phys. Chem. B*, 2012, **116**, 11124-11136.
43. C. J. Dickson, L. Rosso, R. M. Betz, R. C. Walker and I. R. Gould, *Soft Matter*, 2012, **8**, 9617-9627.
44. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
45. N. Busschaert, P. A. Gale, C. J. Haynes, M. E. Light, S. J. Moore, C. C. Tong, J. T. Davis and W. A. Harrell, Jr., *Chem. Commun.*, 2010, **46**, 6252-6254.
46. N. Busschaert, M. Wenzel, M. E. Light, P. Iglesias-Hernandez, R. Perez-Tomas and P. A. Gale, *J. Am. Chem. Soc.*, 2011, **133**, 14136-14148.
47. N. Busschaert, L. E. Karagiannidis, M. Wenzel, C. J. E. Haynes, N. J. Wells, P. G. Young, D. Makuc, J. Plavec, K. A. Jolliffe and P. A. Gale, *Chem. Sci.*, 2014, **5**, 1118-1127.
48. D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2012), AMBER 12, University of California, San Francisco.
49. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
50. D. R. Roe and T. E. Cheatham, 3rd, *J. Chem. Theory Comput.*, 2013, **9**, 3084-3095.
51. The PyMOL Molecular Graphics System, Version 1.3r1 Schrödinger, LLC.
52. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
53. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171-179.
54. Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
55. A. K. Rappe and C. J. R. Casewit, *Molecular Mechanics Across Chemistry*, Univ Science Books 1st edn., 1997.
56. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
57. N. Kucerka, M. P. Nieh and J. Katsaras, *Biochim. Biophys. Acta*, 2011, **1808**, 2761-2771.
58. L. Martinez, R. Andrade, E. G. Birgin and J. M. Martinez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
59. I. S. Joung and T. E. Cheatham, 3rd, *J. Phys. Chem. B*, 2008, **112**, 9020-9041.
60. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
61. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
62. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
63. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
64. N. Kucerka, S. Tristram-Nagle and J. F. Nagle, *J. Membr. Biol.*, 2005, **208**, 193-202.
65. H. Loeffler, Handy Routines for Ptraj/Cpptraj. http://www.hecbiosim.ac.uk/ptraj-routines.
66. J. Seelig and N. Waespe-Sarcevic, *Biochemistry*, 1978, **17**, 3310-3315.
67. B. Perly, I. C. P. Smith and H. C. Jarrell, *Biochemistry*, 1985, **24**, 1055-1063.
68. T. Huber, K. Rajamoorthi, V. F. Kurze, K. Beyer and M. F. Brown, *J. Am. Chem. Soc.*, 2002, **124**, 298-309.
69. B. Perly, I. C. P. Smith and H. C. Jarrell, *Biochemistry*, 2002, **24**, 1055-1063.
70. S. K. Dey and G. Das, *Dalton transactions*, 2012, **41**, 8960-8972.
71. A. Gogoi and G. Das, *Supramol. Chem.*, 2013, **25**, 819-830.
72. X. Wu, L. W. Judd, E. N. W. Howe, A. M. Withecombe, V. Soto-Cerrato, H. Li, N. Busschaert, H. Valkenier, R. Perez-Tomas, D. N. Sheppard, Y. B. Jiang, A. P. Davis and P. A. Gale, *Chem*, 2016, **1**, 127-146.
73. D. Laage, T. Elsaesser and J. T. Hynes, *Chem. Rev.*, 2017, **117**, 10694-10725.

Chapter III.

Tilting and tumbling in anion carriers tuned through n*-alkyl substitution: theoretical insights*

Summary

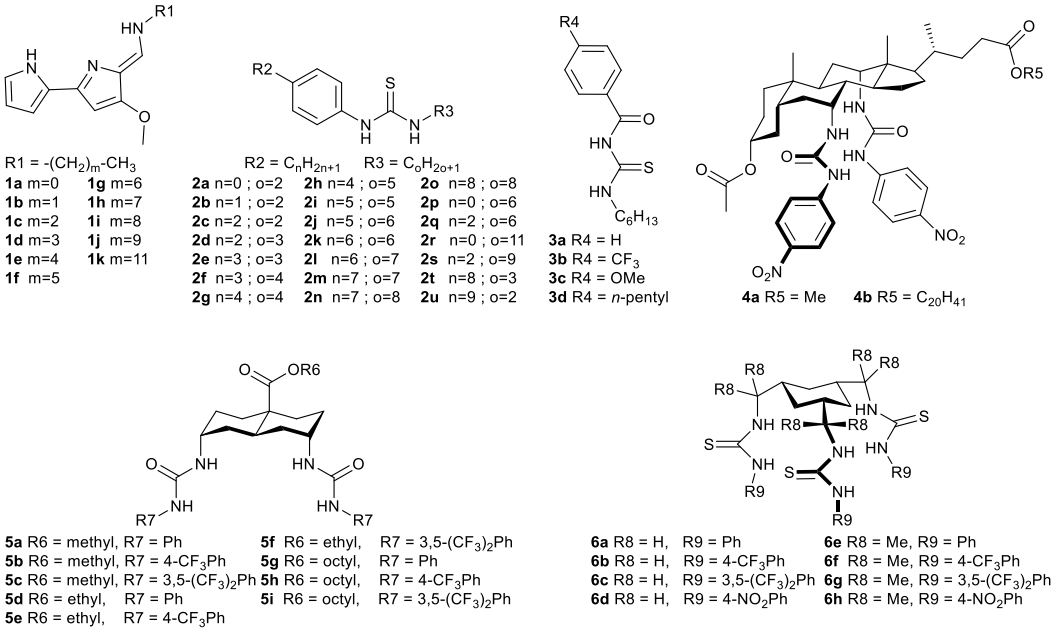
Anion transport by synthetic carriers (anion ionophores) holds promise for medical applications, especially the treatment of Cystic Fibrosis (CF). Among the factors which determine carrier activity, the size and disposition of alkyl groups have proven to be remarkably important. Herein, a series of bis-thioureidodecalin anion ionophores is described, in which alkyl substituents on one face are varied from C_0 to C_{10} in two-carbon steps (see Scheme III-2). Molecular Dynamics (MD) simulations carried out at the atomistic level allowed to rationalise the experimental anion transport data, as follows. When placed inside phospholipid bilayers, the chloride complexes release the anion as they approach the water/lipid interface. The free transporter then stays at the interface, adopting an orientation that depends on the alkyl substituent. Furthermore, longer chains tilt the binding site away from the interface, potentially freeing the transporter or complex to move through the membrane. However, chains which are too long can also slow transport by inhibiting movement, and especially reorientation, within the phospholipid bilayer.

The work reported in this chapter was developed within the scope of the collaboration established with the experimental group of Professor Anthony P. Davis (University of Bristol), resulting in the publication of the paper: S. J. Edwards, I. Marques, C. M. Dias, R. A. Tromans, N. R. Lees, V. Félix, H. Valkenier and A. P. Davis, *Chem. Eur. J.*, 2016, **22**, 2004-2011. Herein, only the computational studies carried out for this project are presented in detail.

III.1. Introduction

The transport of anions across phospholipid bilayers has become an active field of supramolecular Chemistry.[1-7] There is particular interest in developing small molecules for use in biological research and medicine. For example, a number of genetic conditions are caused by malfunctioning chloride channels. These include Bartter syndrome, Best disease and, most importantly, cystic fibrosis.[8] Synthetic anion transporters could potentially be used to treat such conditions by replacing the activity of the missing or malfunctioning chloride channels.[9]

The activities of synthetic anion carriers are determined by various factors. Early work is focused on anion affinities, and there are many examples of series within which transport rates correlate with binding constants.[10-16] However, these correlations do not extend beyond closely related molecules, and it is clear that other considerations are relevant.[17-19] Interactions of receptors and complexes with the membrane environment are especially significant. On one hand, it is clearly important that transporters must partition into the membrane from water, and, on the other hand, if added to pre-existing membranes (*e.g.*, cells or aqueous vesicle suspensions) they must be able to find the interiors of the bilayers. Molecules which are too hydrophilic may fail on the first condition, whereas molecules which are too hydrophobic may fail on the second (due to precipitation from water). These factors may underlie the observation of optimum lipophilicity in studies of transporters such as tambjamines **1**,[20] phenylthioureas **2**,[21, 22] and acylthioureas **3**[23] (see Scheme III-1).

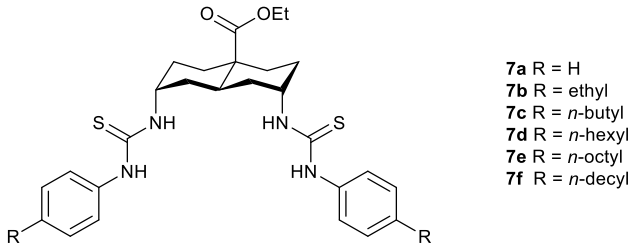


Scheme III-1. Series of molecules where lipophilicity tuning has been previously studied: tambjamines **1**,[20] phenylthioureas **2**,[21, 22] acylthioureas **3**,[23] steroid-based cholapods **4**,[11] bis-ureidodecalins **5**,[24] and tris-thioureas **6**.[15]

However, other investigations show that lipophilic groups can also affect the intrinsic transport activities of anion ionophores, that is, the rate at which they operate once present in the membranes. To study these effects, it is necessary to use assay methods in which the transporter concentration within the membranes is known reliably. In practice, this can be achieved by: *a)* employing transporters which are too lipophilic to leach into the aqueous phase; and *b)* ensuring that these molecules are in the membranes by pre-incorporating them during vesicle production. Implementation of such methodology is non-trivial, because the presence of anion ionophores in the membranes means that the vesicles are

anion permeable at all stages of the experiment. Thus, for the steroid-based cholapods, such as **4**, the methyl and eicosyl esters **4a** and **4b** possess very similar activities.[11] However, for other systems, the effects can be large. For example, in bis-ureidodecalins **5**, extending side-chain R6 gave substantial benefits (*e.g.*, a four-fold increment in activity for R6 = methyl → octyl),[24] while the methyl groups in the tris-thioureas **6g** resulted in a *ca.* 14-times greater activity than analogous **6c**.[15] Moreover, these effects do not just depend on the overall lipophilicity of the transporters. Simple thioureas **2,** for which n+m=11,[21] are equally lipophilic, but their transport activities were found to vary by a factor of four. Anion transport activity was optimal when the side-chains were roughly equal in length (*e.g.*, *n*=5, *m*=6), suggesting that "lipophilic balance" is advantageous for anion ionophore activity.

The series of transporters alkyl-substituted transporters **7**, structurally related with **5** and presented in Scheme III-2,[25] likewise **2**, provides a continuous structural variation over a range of chain lengths, allowing the fine tuning of the lipophilic substituent effects.



Scheme III-2. Series of bis-thioureidodecalins studied in this work.

The chloride binding affinities (K_a) of bis-thioureidodecalins **7a**–**7f** were experimentally assessed in chloroform and in wet dimethyl sulfoxide (DMSO) and are listed in Table III-1, together with the *c*log*P* values for these six molecules.[25] As expected, the increasing chain length has little to no effect on the binding affinity strength across this series in both aprotic solvents. In contrast, the initial rates (k_{ini})[§] and specific initial rates ([*I*])[**] of **7a**–**7f** (also summarised in Table III-1) show that the anion transport activity of these alkyl-decorated receptors is modulated by the chain length. Indeed, when the chain length is extended from C_0 to C_6, the transport rates increase approximately three times, however, when the chain length reaches C_8 and C_{10} the rates diminish by almost the same amount. Thus, activity passes through a maximum, similarly to **2**.[21] The observation of a maximum suggests that the alkyl groups determine the anion transport activity, probably involving the details of interactions with membrane lipids.[25]

Thus, to rationalise how the lipophilic groups of **7a**–**7f** affect the movement and positioning within the membrane, extensive MD simulations were carried out with their chloride complexes in a POPC bilayer, obtaining structural and energetic insights on the anion transport activity.

§ Being F_0 and F_t the fluorescence at time 0 and time *t*, k_{ini} is obtained from fits (0–500 s) of F_0/F_t to a double exponential decay function.[26]

** [*I*] is the initial slope of F_0/F_t *vs*. time *t*, divided by the transporter/lipid ratio in the vesicles, and averaged over a range of varying ratios.[16]

Table III-1. Chloride binding affinities and transport data for bis-thioureidodecalins **7a**–**7f**.[a]

Transporter	R	*c*log*P*	K_a (M^{-1})		k_{ini} (s^{-1})		[*I*] (s^{-1})
			$CHCl_3$	DMSO-d_6	1:2500	1:1000	
7a	H	5.2	4.3×10^5	7.3×10^2	0.0010	0.0026	2.6
7b	ethyl	6.4	1.6×10^5	7.0×10^2	0.0012	0.0033	3.2
7c	*n*-butyl	7.9	1.6×10^5	7.2×10^2	0.0027	0.0043	5.5
7d	*n*-hexyl	9.5	1.5×10^5	7.2×10^2	0.0031	0.0068	7.3
7e	*n*-octyl	11.0	1.7×10^5	7.1×10^2	0.0018	0.0034	4.0
7f	*n*-decyl	12.6	1.8×10^5	6.7×10^2	0.0015	0.0024	3.1

[a] Table reproduced from ref. 25.

III.2. Methods

The MD simulations were carried out with the AMBER 14 package,[27-29] using the Lipid14 force field[30] for the POPC lipids, and General AMBER Force Field (GAFF)[31, 32] parameters and atomic RESP charges[33] for transporters **7a**–**7f** as follows.

III.2.1. Calculation of atomic RESP charges for transporters 7a–7f

The initial structure of **7a** was generated from a crystal structure of a decalin derivative featuring a carboxylate group at the C4a position, deposited with the Cambridge Crystallographic Data Centre (CCDC)[34, 35] under RefCode MXRCDC.[36] Two phenyl thiourea substituents were appended at both C2 and C7 decalin skeleton positions. The remaining five molecules of the diaxial thiourea series were generated from **7a** by addition of the appropriate alkyl chain at the *para* position of both phenyl rings. Subsequently, all structures were optimised at the HF/6-31G* level with *Gaussian 09*.[37] The optimisation of each single structure of **7a**–**7f** was followed by a single point calculation to generate the electrostatic potential (ESP) at the same theory level, using the Merz–Singh–Kollman scheme with 4 concentric layers per atom and 6 density points in each layer (IOp (6/33 = 2, 6/41 = 4, 6/42 = 6)). The initial atomic charges of each molecule were then calculated by RESP fitting, along with the attribution of GAFF atom types, using the *antechamber* module,[29] as implemented in the AMBER package.

To obtain atomic charges less dependent of the molecular conformation or orientation, the calculation of the final RESP charges employed in the MD simulations reported was preceded by conformational analyses on receptors **7a**–**7f**. The initial Molecular Mechanics (MM) energy minimised structures, using the initial RESP charges, were heated in the gas-phase at 500 K for 50 ps, followed by a collection run of 5.0 ns, using a time step of 1.0 fs. The use of this high temperature allows the stochastic search of the conformational space, since the energetic barriers are easily surmounted.[38] Frames were saved every 0.1 ps leading to a trajectory file containing 50000 structures for each molecule. All these structures were further minimised by MM using a steepest descendent gradient followed by the conjugate gradient algorithm, until the convergence criterion of 0.0001 kcal mol^{-1} $Å^{-1}$ was attained. Afterwards, the MM minimised structures were clustered with *UCSF Chimera*.[39] Five representative conformations with substantially different RMSD values, with both thiourea binding units adopting a *syn* configuration but displaying a different spatial disposition, were selected and underwent new HF/6-31G* geometry optimisations and ESP calculations as described above. The individual ESP data were extracted from the corresponding *Gaussian 09*[37] outputs, concatenated, and subsequently used to generate the input files for the two-stage RESP fitting, using identical weights for all conformations.

III.2.2. Generation of chloride complexes

The structures of **7a**–**7f** chloride complexes were obtained in gas-phase *via* conformational analyses as described above for the free molecules. However, the production runs were carried out only for 1.0 ns, leading, at the end of the quenched dynamics protocol, to 10000 MM minimised structures. The most representative structure of each complex, obtained by cluster analysis with *UCSF Chimera*,[39] is depicted in Figure III-1 and was selected for the subsequent MD simulations.

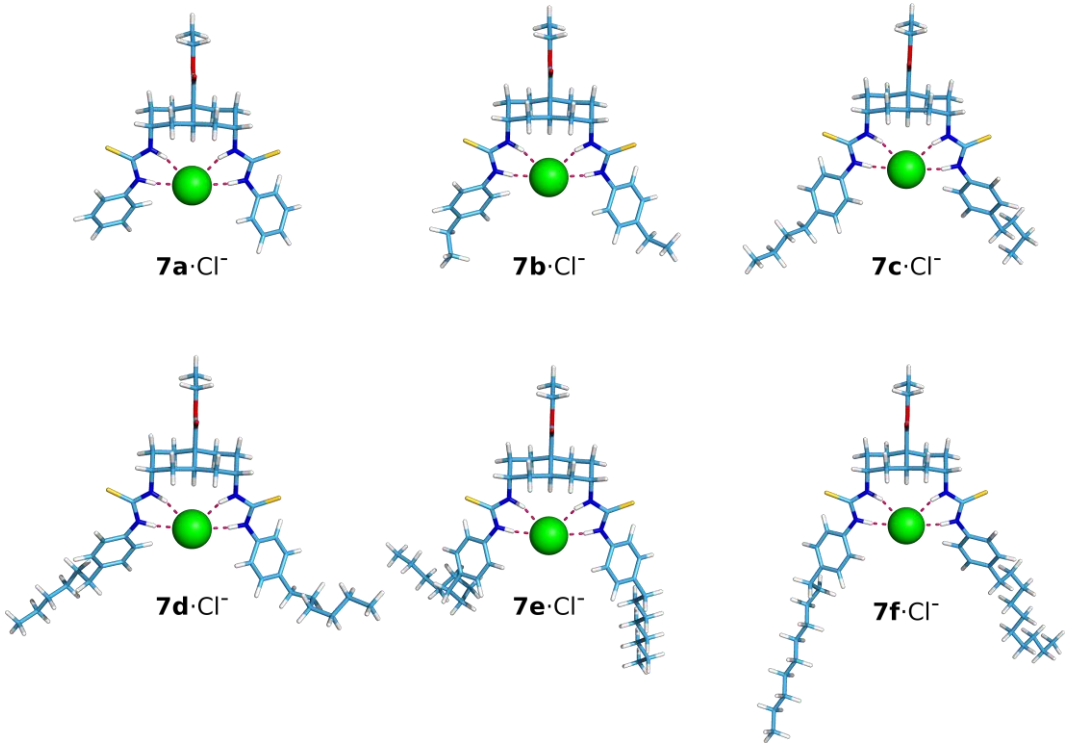


Figure III-1. Chloride complexes of **7a**–**7f** obtained from quenched MD simulations. The transporter is shown in sticks, with the hydrogen atoms shown in white, oxygen atoms in red, nitrogen atoms in blue, sulfur atoms in yellow, and carbon atoms in light blue. The chloride ion is shown as a green sphere and the N–H···Cl^- bonds are shown as magenta dashes.

III.2.3. Simulations in a POPC bilayer

Simulation of the initial POPC bilayer

Before the simulation of the chloride complexes inserted into a POPC membrane model, a membrane system with an orthorhombic shape (dubbed *free membrane*) composed of 128 phospholipids, 6500 TIP3P water molecules,[40] 18 Cl^- and 18 Na^+ ions (0.15 M) was built with a water per lipid ratio of 50.1. This free membrane was simulated for 150 ns, under periodic boundary conditions at 303 K with force field parameters taken from Lipid14 and the ions described with van der Waals parameters developed to be used along with the TIP3P water model.[41] The simulation protocol followed is equivalent to the one detailed below for the membrane systems with chloride complexes. The structural parameters area per lipid, bilayer thickness, order parameters and electron density profiles estimated for the last 70 ns of sampling, are similar to those originally reported for the POPC bilayer system containing a water per lipid ratio of 31.6.[30] Therefore, increasing the number of water molecules had no impact in the bilayer system, and this equilibrated *free membrane* was subsequently

used in the MD simulations carried out with the chloride complexes either inserted into the core of the bilayer (setup A – an approximation to the experimental chloride transmembrane transport studies) or immersed in the water slab (setup B) with a random orientation. This initial positioning of the chloride complexes of **7a**–**7f** within the free membrane was performed with *Packmol*,[42] and yielded systems A.**7a** to A.**7f**, as well as systems B.**7a** to B.**7f**, as given in Table III-2.

Simulation of membrane systems with chloride complexes at the bilayer core – setup A

The MD simulations A_1.**7a** to A_4.**7f** were carried out as follows: the initial configuration of each system was submitted to 10000 steps of MM energy minimisation with a 500 kcal mol^{-1} $Å^{-2}$ positional restraint on the chloride complex and lipid molecules, through the steepest descent algorithm for 3000 steps plus 7000 steps of the conjugated gradient algorithm, followed by the relaxation of the entire system for another 10000 steps, with the same protocol. The equilibration of the system proceeded by heating it to 303 K in an NVT ensemble for 100 ps with a 10 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex and lipid molecules. The equilibration stage proceeded with a 5.0 ns run using an NPT ensemble with a 5.0 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex. Then, the positional restraint was removed and the simulation continued for further 150 ns. The long-range electrostatic interactions were described with the Particle Mesh Ewald (PME) algorithm[43] using a real-space cut-off at 10 Å. The cut-off for the Lennard–Jones interactions was also set at 10 Å. The temperature of the system was maintained at 303 K, using Langevin dynamics,[44] with a collision frequency γ of 1.0 ps^{-1}. The pressure was controlled by the Berendsen barostat[45] at 1 atm and compressibility of 44.6×10^{-6} bar^{-1}, with a relaxation time of 1.0 ps. The covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm,[46] allowing the use of a 2.0 fs time step. This protocol was employed to run four independent MD runs of 150 ns for each setup A membrane system (see Table III-2), with the frames saved every 10 ps.

Simulation of membrane systems with chloride complexes at the bilayer core – setup A'

From the end of the equilibration period of simulations A_1.**7a**–A_1.**7f**, a single production run of 150 ns was undertaken with each chloride complex maintained through the use of four distance restraints between the nitrogen atoms of each thiourea binding units and the chloride ion. Harmonic restraints of 5.0 kcal mol^{-1} $Å^{-2}$ were applied to obtain N···Cl^- distances of 3.5 Å.††

Simulation of membrane systems with chloride complexes in the water slab – setup B

The MD simulations B_1.**7a**–B_2.**7f**, listed in Table III-2, were carried out using the aforementioned protocol given for simulations of setup A, except for the length of the production runs, which were 200 ns long.

†† The distance restraints were selected in agreement with the N–H···Cl^- dimensions found in a single crystal X-ray structure deposited with CCDC[34, 35] of a bis-ureidodecalin chloride complex.[47]

Table III-2. Membrane systems simulated at 303 K.

Setup	Transporter	System ID	Simulation ID				Initial Position	Simulation time (ns)	N⋯Cl^- distance restraints
			Run 1	Run 2	Run 3	Run 4			
A	**7a**	A.**7a**	A_1.**7a**	A_2.**7a**	A_3.**7a**	A_4.**7a**	Bilayer core	150	No
A	**7b**	A.**7b**	A_1.**7b**	A_2.**7b**	A_3.**7b**	A_4.**7b**	Bilayer core	150	No
A	**7c**	A.**7c**	A_1.**7c**	A_2.**7c**	A_3.**7c**	A_4.**7c**	Bilayer core	150	No
A	**7d**	A.**7d**	A_1.**7d**	A_2.**7d**	A_3.**7d**	A_4.**7d**	Bilayer core	150	No
A	**7e**	A.**7e**	A_1.**7e**	A_2.**7e**	A_3.**7e**	A_4.**7e**	Bilayer core	150	No
A	**7f**	A.**7f**	A_1.**7f**	A_2.**7f**	A_3.**7f**	A_4.**7f**	Bilayer core	150	No
A'	**7a**	A'.**7a**	A'_1.**7a**	–	–	–	Bilayer core	150	Yes
A'	**7b**	A'.**7b**	A'_1.**7b**	–	–	–	Bilayer core	150	Yes
A'	**7c**	A'.**7c**	A'_1.**7c**	–	–	–	Bilayer core	150	Yes
A'	**7d**	A'.**7d**	A'_1.**7d**	–	–	–	Bilayer core	150	Yes
A'	**7e**	A'.**7e**	A'_1.**7e**	–	–	–	Bilayer core	150	Yes
A'	**7f**	A'.**7f**	A'_1.**7f**	–	–	–	Bilayer core	150	Yes
B	**7a**	B.**7a**	B_1.**7a**	B_2.**7a**	–	–	Water phase	200	No
B	**7b**	B.**7b**	B_1.**7b**	B_2.**7b**	–	–	Water phase	200	No
B	**7c**	B.**7c**	B_1.**7c**	B_2.**7c**	–	–	Water phase	200	No
B	**7d**	B.**7d**	B_1.**7d**	B_2.**7d**	–	–	Water phase	200	No
B	**7e**	B.**7e**	B_1.**7e**	B_2.**7e**	–	–	Water phase	200	No
B	**7f**	B.**7f**	B_1.**7f**	B_2.**7f**	–	–	Water phase	200	No

III.3. Results and Discussion

III.3.1. Structural impact of 7a–7f on the bilayer model

The structural impact of **7a**–**7f** in the membrane model was examined through evaluation of the area per lipid, bilayer thickness, electron density profiles and order parameters $|S_{CD}|$ during the last 50 ns of each MD run of setup A with each transporter (see Table III-2). The average values for area per lipid and bilayer thickness are gathered in Table III-3 and the electron density profiles and order parameters are plotted from Figure B-1 to Figure B-12 (see Appendix B). The area per lipid and bilayer thickness average values are quite similar to the ones reported for the membrane system without any transporter. The small variations observed in all these structural parameters can be attributed to the necessary adjustments for the nesting of the transporters at the water/lipid interface level.

Table III-3. Comparison between the area per lipid and bilayer thickness of simulations A_1.**7a**–A_4.**7f** (50 ns of sampling) to the free membrane system (70 ns of sampling).

Simulation ID	Area per lipid (Å²)		Bilayer thickness (Å)		*N*
	Avg ± SD	Range	Avg ± SD	Range	
Free membrane	65.58 ± 1.19	[61.98:68.98]	37.75 ± 0.57	[36.05:39.60]	7000
A_1.**7a**	66.26 ± 1.12	[63.22:69.29]	37.50 ± 0.52	[36.28:39.22]	5000
A_2.**7a**	65.04 ± 1.34	[62.00:68.51]	38.15 ± 0.68	[36.57:40.02]	5000
A_3.**7a**	66.19 ± 0.93	[63.16:68.80]	37.57 ± 0.46	[36.21:38.95]	5000
A_4.**7a**	65.38 ± 0.84	[62.77:68.21]	37.93 ± 0.47	[36.40:39.48]	5000
A_1.**7b**	66.07 ± 0.86	[63.64:68.98]	37.69 ± 0.39	[36.22:38.73]	5000
A_2.**7b**	65.68 ± 1.26	[61.21:69.21]	37.84 ± 0.56	[36.44:40.09]	5000
A_3.**7b**	66.04 ± 1.00	[63.12:68.31]	37.65 ± 0.47	[36.46:39.26]	5000
A_4.**7b**	66.10 ± 1.00	[63.34:68.96]	37.66 ± 0.42	[36.46:38.78]	5000
A_1.**7c**	66.76 ± 1.67	[63.11:70.61]	37.39 ± 0.75	[35.47:39.33]	5000
A_2.**7c**	65.48 ± 0.93	[62.66:68.40]	37.98 ± 0.45	[36.49:39.45]	5000
A_3.**7c**	65.73 ± 1.45	[61.95:69.61]	37.90 ± 0.70	[36.02:39.82]	5000
A_4.**7c**	66.40 ± 1.15	[63.47:70.31]	37.61 ± 0.55	[35.94:38.99]	5000
A_1.**7d**	65.66 ± 1.34	[62.69:69.17]	37.88 ± 0.71	[35.87:39.79]	5000
A_2.**7d**	65.65 ± 1.19	[62.10:70.29]	37.88 ± 0.56	[35.79:39.77]	5000
A_3.**7d**	65.59 ± 1.23	[62.05:68.73]	38.01 ± 0.52	[36.60:39.49]	5000
A_4.**7d**	65.39 ± 1.01	[62.55:68.42]	37.98 ± 0.50	[36.59:39.55]	5000
A_1.**7e**	65.76 ± 1.00	[63.11:68.80]	37.92 ± 0.52	[36.27:39.47]	5000
A_2.**7e**	66.00 ± 0.97	[63.78:69.30]	37.82 ± 0.50	[36.20:39.19]	5000
A_3.**7e**	65.37 ± 1.20	[61.96:69.07]	38.05 ± 0.58	[36.32:39.91]	5000
A_4.**7e**	66.52 ± 1.13	[63.49:69.80]	37.54 ± 0.50	[36.13:38.93]	5000
A_1.**7f**	66.08 ± 1.27	[63.18:69.30]	37.77 ± 0.58	[36.22:39.36]	5000
A_2.**7f**	66.03 ± 1.45	[62.65:70.70]	37.71 ± 0.64	[35.86:39.46]	5000
A_3.**7f**	65.17 ± 1.09	[61.97:67.87]	38.16 ± 0.51	[36.79:39.61]	5000
A_4.**7f**	65.34 ± 0.97	[62.90:68.36]	38.09 ± 0.44	[36.54:39.34]	5000

III.3.2. Diffusion of 7a–7f within the bilayer model

The relative position of chloride towards the closest membrane interface was evaluated throughout the course of the MD runs using the $P_{int}\cdots Cl^-$ distance, in which P_{int} represents the centre of mass (COM) determined by the phosphorus atoms in that monolayer (which is considered as the water/lipid interface). Likewise, the position of the transporter was evaluated by the distance between P_{int} and the COM defined by the carbon atoms of the decalin skeleton ($P_{int}\cdots decalin_{COM}$). The $P_{int}\cdots N{-}H_{COM}$, $P_{int}\cdots p{-}C_{COM}$, $P_{int}\cdots tail_{COM}$ distances were also determined, in which $N{-}H_{COM}$, $p{-}C_{COM}$ and $tail_{COM}$ represent the COM determined by the nitrogen atoms of the two thiourea binding sites, the phenylene carbon atoms *para* to the thiourea groups, and the two terminal carbon atoms in each of the aliphatic groups R in **7b**–**7f**, respectively. All these reference points are identified in the sketch given in Figure III-2, with colour coding. Taken together, the $P_{int}\cdots decalin_{COM}$, $P_{int}\cdots N{-}H_{COM}$ and $P_{int}\cdots p{-}C_{COM}$ distances give the position and the orientation of the receptor core relative to the membrane interface. Meanwhile, the $P_{int}\cdots tail_{COM}$ distance provides information on the extension and orientation of the alkyl substituents.

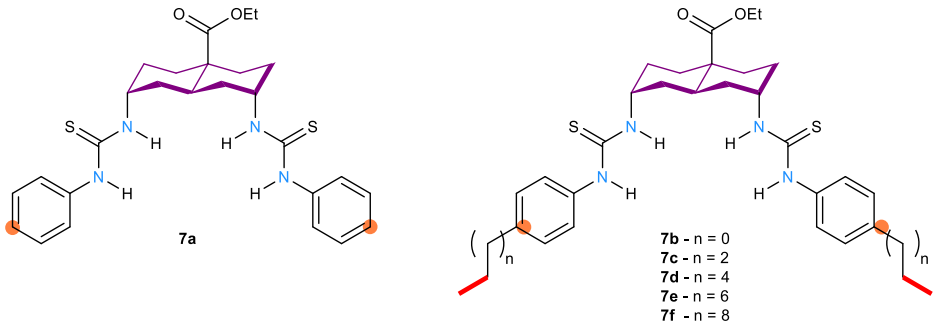


Figure III-2. Atoms used to define the different referential COM monitored throughout the MD simulations: decalin$_{COM}$ (purple), N–H$_{COM}$ (blue), *p*–C$_{COM}$ (orange), and tail$_{COM}$ (red).

The evolution of these four distances in simulations A$_1$.**7a** to A$_4$.**7f** is shown from Figure III-3 to Figure III-8. In most cases, chloride release occurs within the first 15 ns of the 150 ns simulation. The snapshots of the simulations, given in Figure III-9 and Figure III-10, revealed that when the complex moves towards the water/lipid interface, it is met by water molecules permeating from the aqueous phase. The water molecules solvate the chloride ion and promote the breakup of the complex, so that chloride release occurs before the complex reaches the membrane interface. The diffusion of the chloride complex of **7d** in simulation A$_1$.**7d** is featured in Movie III-1, which also presents the anion release assisted by water molecules. The time taken for chloride release varies between runs, and does not show meaningful differences between receptors.

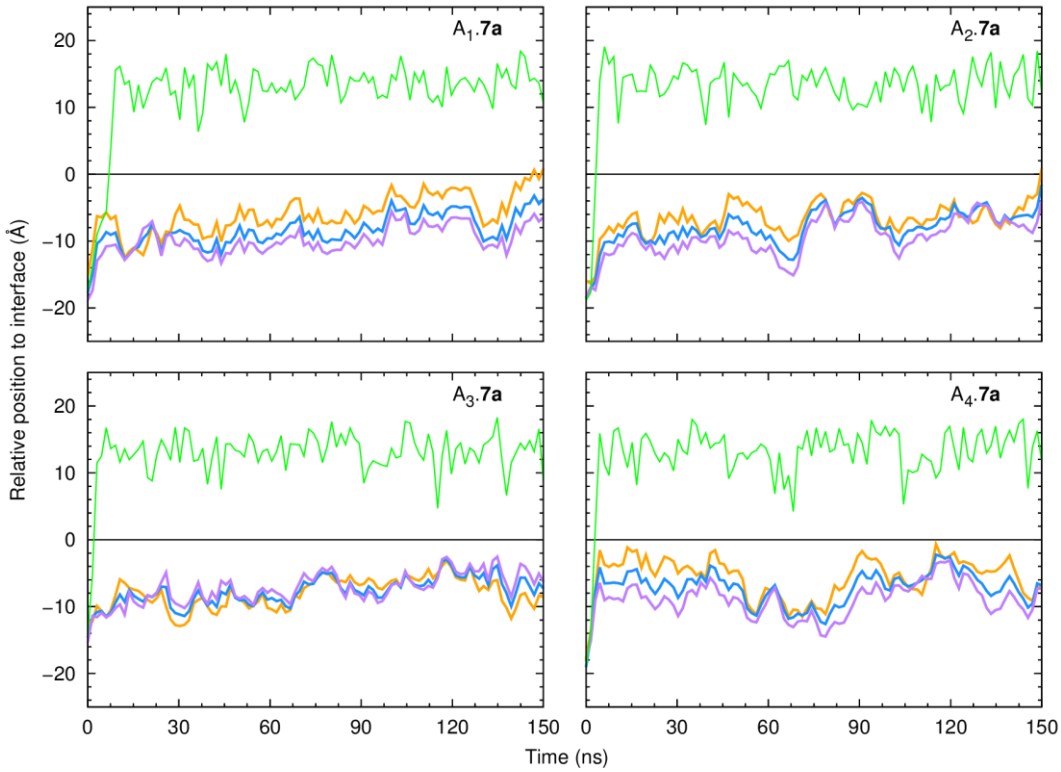


Figure III-3. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$ and *p*–C$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7a**–A$_4$.**7a**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and Cl$^-$···P$_{int}$ are shown in purple, blue, orange and green lines, in this order. The line colouring is in agreement with the sketch given in Figure III-2. The water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

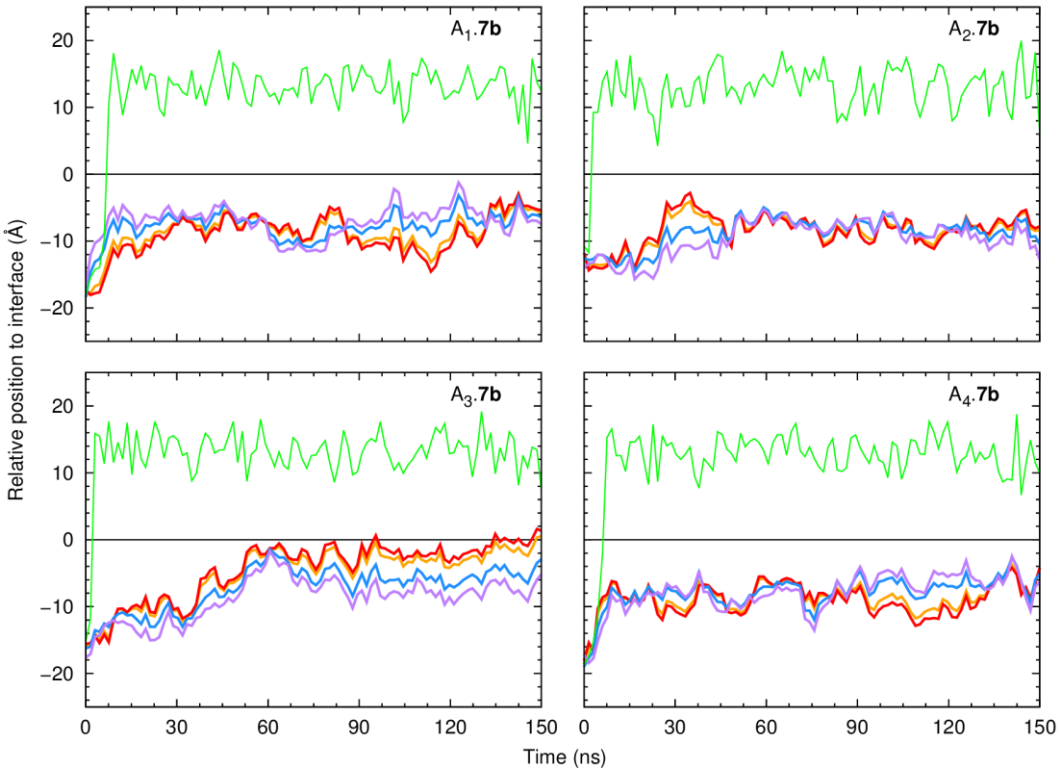


Figure III-4. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7b**–A$_4$.**7b**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. tail$_{COM}$···P$_{int}$ is shown a red line. Remaining details as given in Figure III-3.

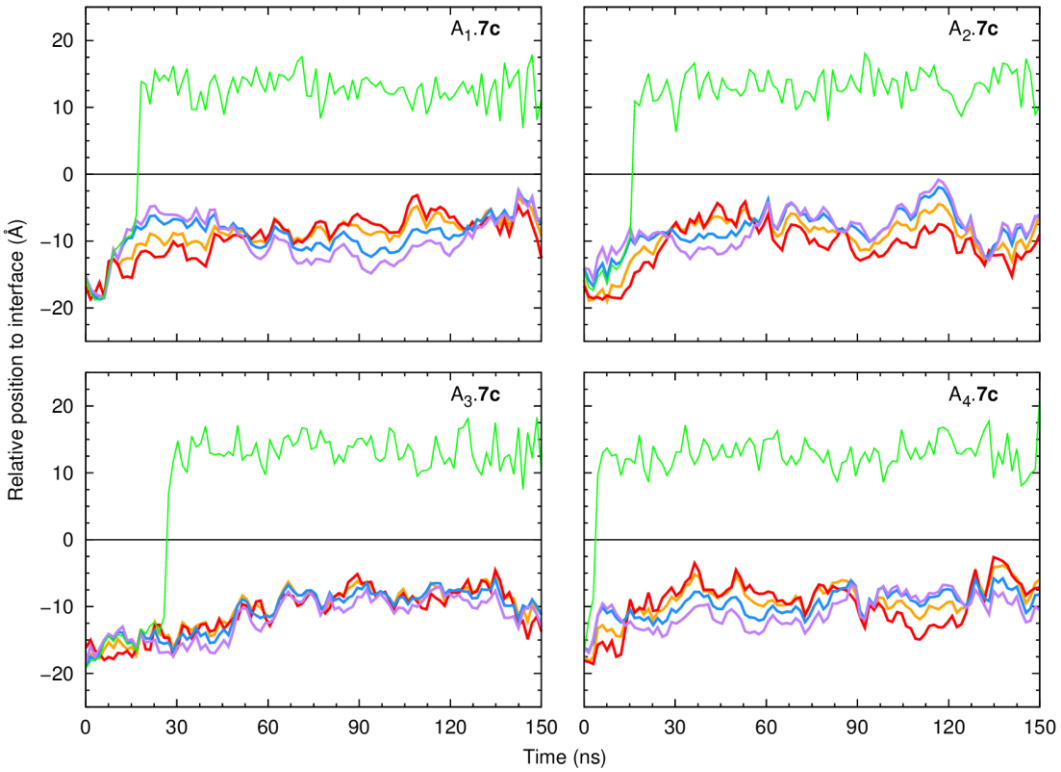


Figure III-5. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7c**–A$_4$.**7c**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. Remaining details as given in Figure III-4.

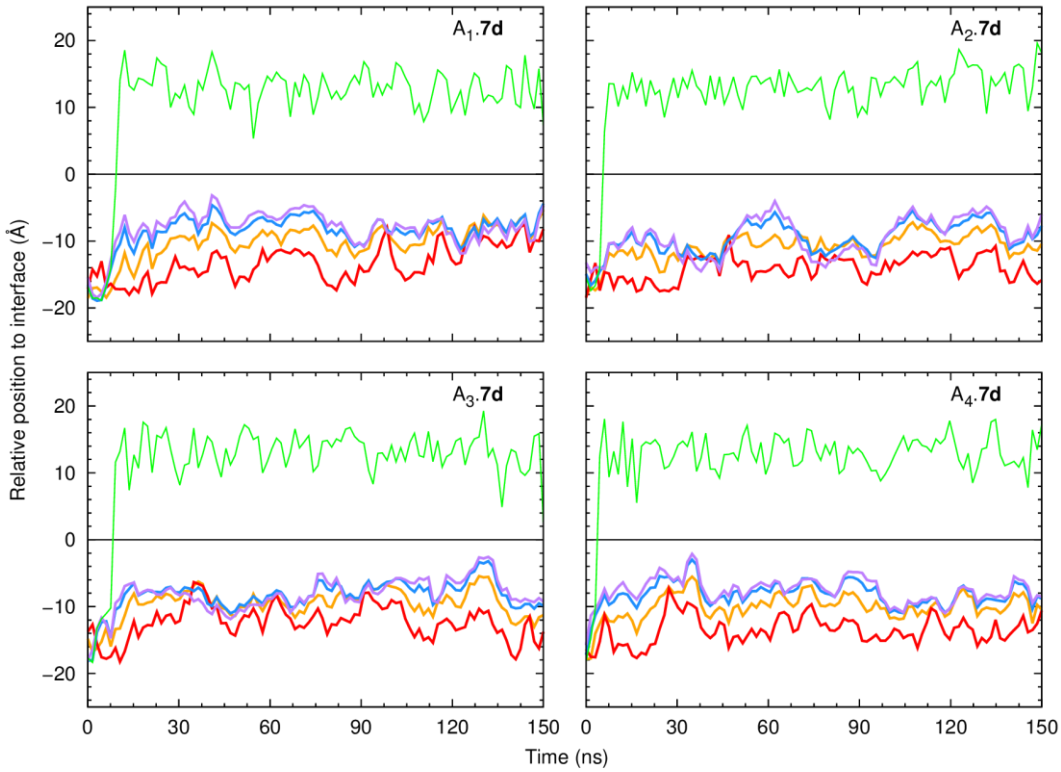


Figure III-6. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7d**–A$_4$.**7d**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. Remaining details as given in Figure III-4.

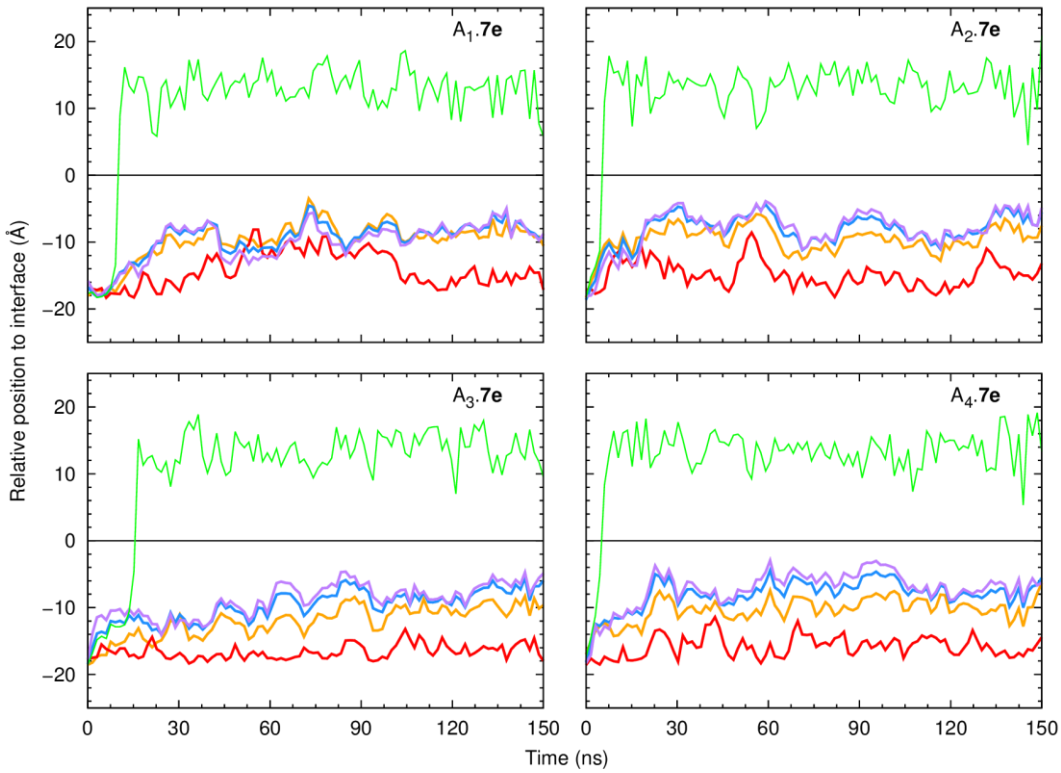


Figure III-7. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7e**–A$_4$.**7e**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. Remaining details as given in Figure III-4.

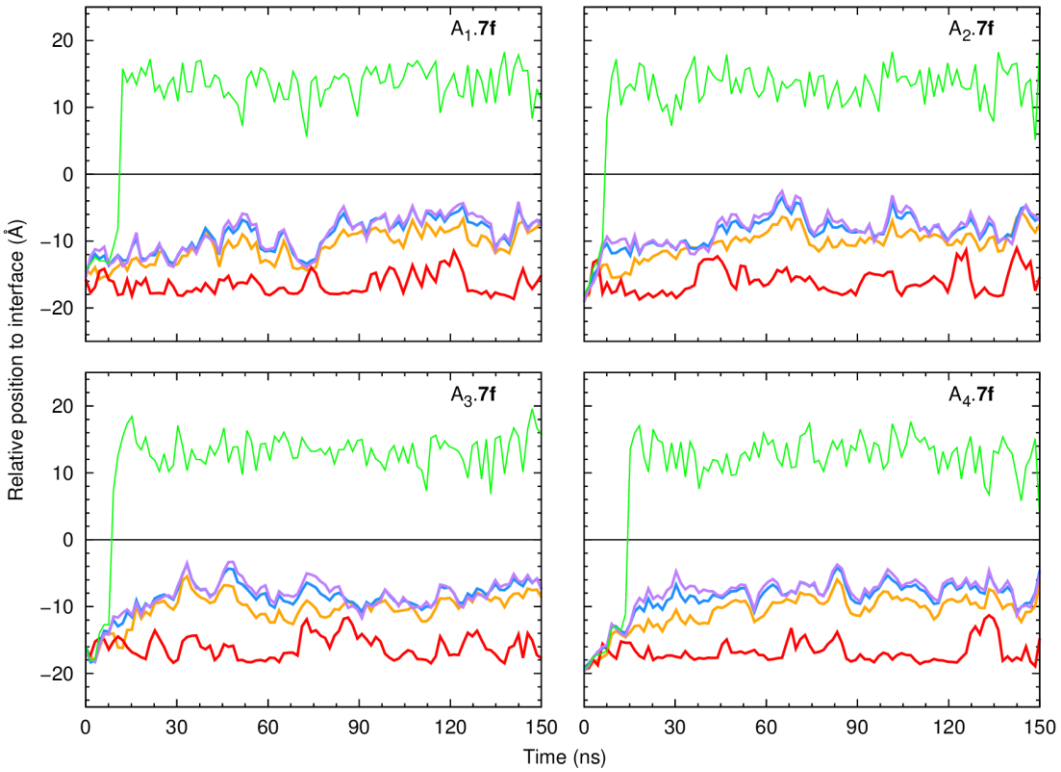


Figure III-8. Evolution of decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances for MD simulations A$_1$.**7f**–A$_4$.**7f**, as well as Cl$^-$···P$_{int}$, during 150 ns of simulation time. Remaining details as given in Figure III-4.

Once the chloride ions have departed, the simulations showed the transporter molecules continuing their movement towards the water/lipid interface and settling into equilibrium positions. In these later parts of the simulations, consistent differences emerge between **7a**–**7f**. In the case of **7a**, the transporter adopts an orientation in which the aromatic rings (see Figure III-3, orange line) are clearly closer to the interface than the decalin moiety and thiourea-binding units (purple and blue lines, respectively). This preferential orientation, with the phenyl groups angled towards the interface, is also evident in the last two consecutive snapshots presented in Figure III-9. In contrast, the medium-sized transporters **7b** and **7c** (R=ethyl, *n*-butyl) do not display a clear preferential orientation (see Figure III-4 for **7b**, Figure III-5 for **7c**, and Figure III-9 for snapshots), whereas the largest transporters **7d**–**7f** (R=*n*-hexyl, *n*-octyl, *n*-decyl) are positioned with the aromatic groups angled away from the interface (see Figure III-6 for **7d**, Figure III-7 for **7e**, Figure III-8 for **7f**, and Figure III-10 for snapshots).

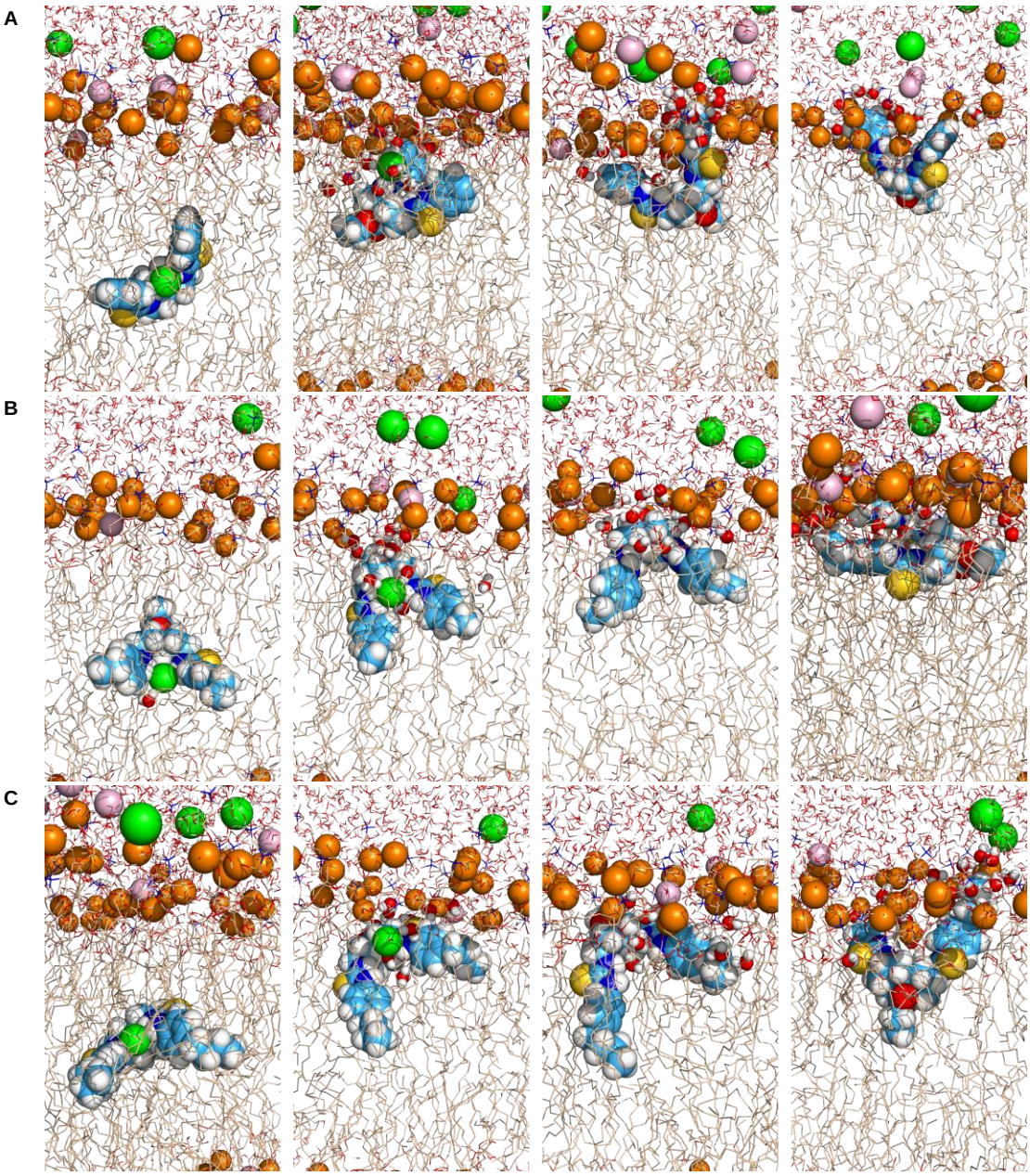


Figure III-9. Consecutive snapshots depicting the diffusion of the chloride complex of **7a**, **7b** and **7c** in simulations A_1.**7a** (**A**), A_1.**7b** (**B**), and A_1.**7c** (**C**), respectively. The transporter, the phosphorus atoms and ions are represented as spheres. The hydrogen atoms are shown in white, oxygen atoms in red, nitrogen atoms in blue, sulfur atoms in yellow, phosphorus atoms in orange and carbon atoms in light blue (transporters) or wheat (phospholipids), whereas the chloride and sodium ions are shown in green and pink, respectively. The chloride decomplexation assisted by water is emphasised with the depiction of water molecules within 3.5 Å from the transporters as spheres. The lipid C–H bonds are omitted for clarity.

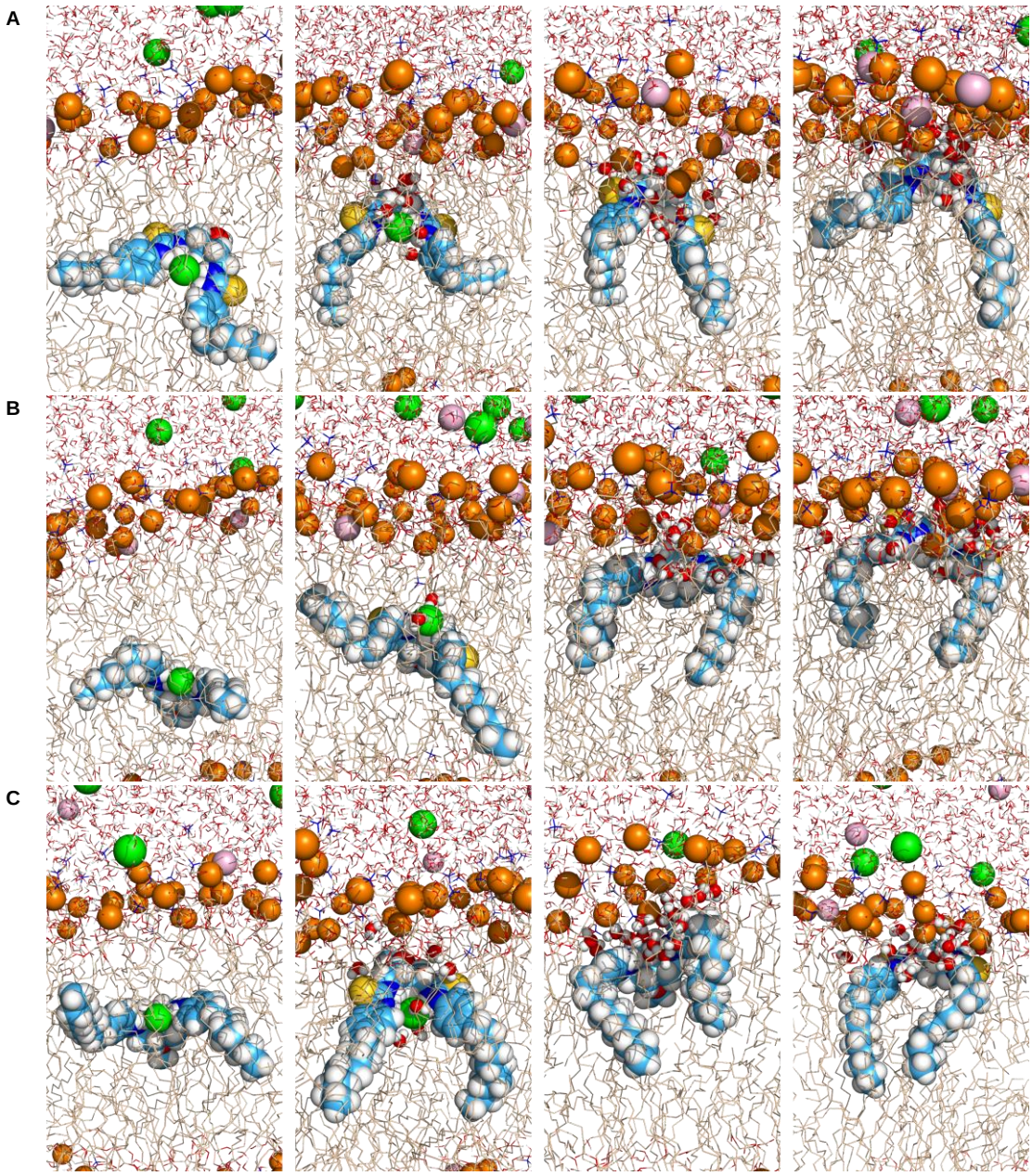


Figure III-10. Consecutive snapshots depicting the diffusion of the chloride complex of **7d**, **7e** and **7f** in simulations A_1.**7d** (**A**), A_1.**7e** (**B**), and A_1.**7f** (**C**), respectively. Remaining details as given in Figure III-9.

The differences in orientation are highlighted in Figure III-11, which shows the average position of the reference points (decalin$_{COM}$, N–H$_{COM}$, *p*–C$_{COM}$, and tail$_{COM}$) over the 200 ns of sampling for each system (see Table III-4 and Table B-1 in Appendix B for the independent MD runs). For compound **7a** the reference points *p*–C$_{COM}$ (◆), N–H$_{COM}$ (■), and decalin$_{COM}$ (▲) lie at increasing distances from the water/lipid interface. For compounds **7b** and **7c**, they are close to each other, and for **7d–7f**, they appear in the reverse order.

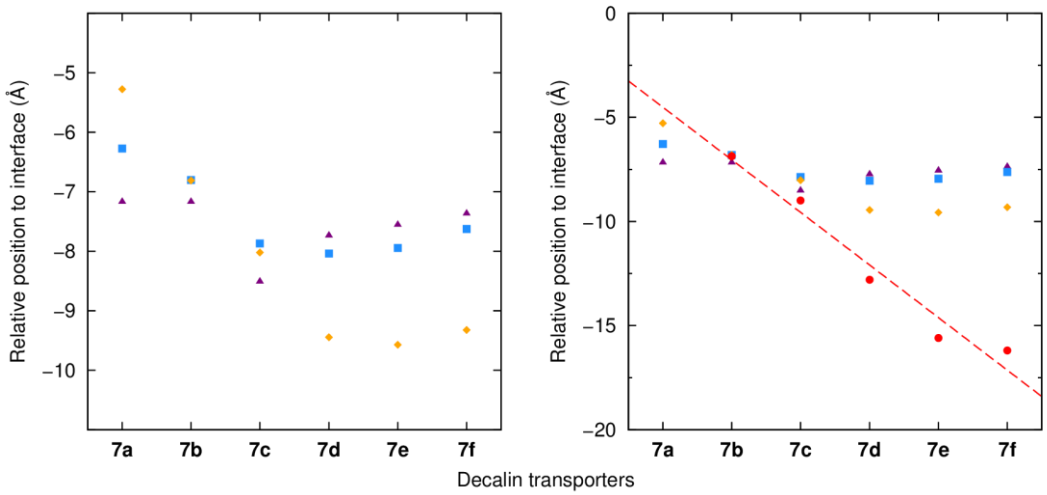


Figure III-11. Average positions of decalin$_{COM}$ (▲), N−H$_{COM}$ (■), *p*−C$_{COM}$ (◆), and tail$_{COM}$ (●) thiourea reference points relative to the closest interface. Each point was calculated averaging the last 50 ns of the four independent setup A MD runs, resulting in 200 ns of total sampling time.

When the simulations were performed with the decalin complexes positioned initially in the water phase (setup B), the chloride ion was also released quickly, after which the transporters entered the membrane. Their final positions and orientations (see Figure III-12) resemble those found when the bilayer core was the starting position, indicating that these results are independent of the initial setup of the experiments.

Table III-4. decalin$_{COM}$···P$_{int}$, N−H$_{COM}$···P$_{int}$, *p*−C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances (Å) averaged (Avg) over all MD runs for each system with the corresponding standard deviations (SD).[a,b]

System ID	decalin$_{COM}$···P$_{int}$	N−H$_{COM}$···P$_{int}$	*p*−C$_{COM}$···P$_{int}$	tail$_{COM}$···P$_{int}$
	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD
A.**7a**	7.17 ± 2.36	6.27 ± 1.92	2.12 ± 3.04	–
A.**7b**	7.17 ± 2.20	6.80 ± 1.79	5.40 ± 1.54	6.86 ± 3.99
A.**7c**	8.51 ± 2.79	7.87 ± 2.26	7.86 ± 2.10	9.00 ± 3.08
A.**7d**	7.73 ± 2.07	8.04 ± 1.88	8.01 ± 1.99	12.80 ± 2.40
A.**7e**	7.55 ± 1.72	7.94 ± 1.49	8.52 ± 1.83	15.60 ± 1.63
A.**7f**	7.37 ± 1.93	7.63 ± 1.66	7.98 ± 1.66	16.19 ± 2.21
A'.**7a**	6.07 ± 2.26	4.85 ± 2.52	2.12 ± 3.04	–
A'.**7b**	6.34 ± 1.77	5.70 ± 1.38	5.40 ± 1.54	5.32 ± 1.82
A'.**7c**	9.42 ± 1.64	8.30 ± 1.43	7.86 ± 2.10	7.97 ± 3.31
A'.**7d**	6.66 ± 2.93	6.66 ± 2.49	8.01 ± 1.99	12.01 ± 1.71
A'.**7e**	8.07 ± 1.70	8.47 ± 1.59	8.52 ± 1.83	13.65 ± 2.53
A'.**7f**	6.81 ± 2.74	7.42 ± 2.27	7.98 ± 1.66	14.07 ± 2.03
B.**7a**	7.05 ± 2.13	5.27 ± 2.07	2.12 ± 3.04	–
B.**7b**	8.60 ± 1.71	8.21 ± 1.50	5.40 ± 1.54	8.22 ± 2.64
B.**7c**	8.15 ± 2.01	8.40 ± 2.01	7.86 ± 2.10	10.51 ± 3.46
B.**7d**	7.53 ± 1.66	7.78 ± 1.64	8.01 ± 1.99	12.77 ± 2.17
B.**7e**	6.81 ± 2.36	7.50 ± 2.00	8.52 ± 1.83	14.31 ± 2.19
B.**7f**	9.22 ± 2.24	8.59 ± 1.79	7.98 ± 1.66	15.28 ± 2.71

[a] *N* =20000, 5000 and 10000 for systems derived from setups A, A' and B, respectively; [b] Values for each independent MD run can be found in Table B-1 and Table B-2 (see Appendix B).

To assess the effect of chloride binding on transporter orientation, a single MD run for each complex was also undertaken with four distance restraints between the chloride ion and the nitrogen atoms of the thiourea binding units, setup A', as described above. The six chloride complexes migrate from the bilayer core to the water/lipid interface, staying at $P_{int}\cdots decalin_{COM}$ distances comparable to those calculated for setup A (see Figure III-12). The average $P_{int}\cdots N{-}H_{COM}$, $P_{int}\cdots p{-}C_{COM}$ and $P_{int}\cdots tail_{COM}$ distances are also comparable in both setups suggesting that the preferred orientations for complex and transporter are similar in each case.

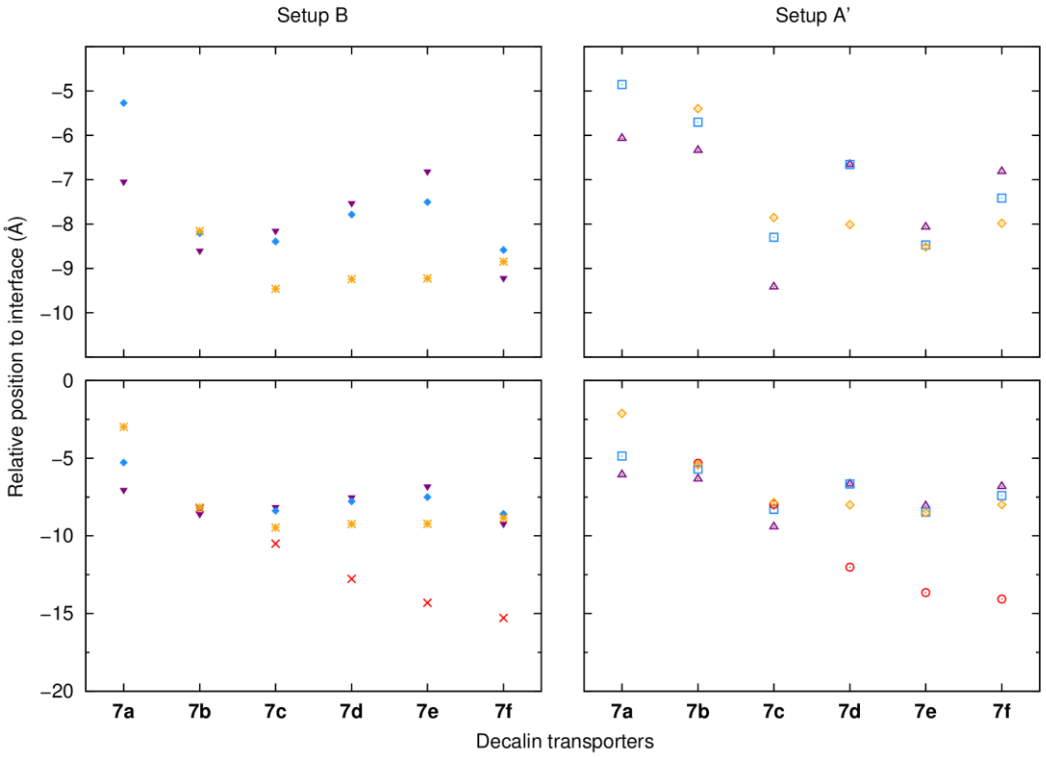


Figure III-12. Average positions of $decalin_{COM}$, $N{-}H_{COM}$, $tail_{COM}$ and $p{-}C_{COM}$ thiourea reference points relatively to the closest interface. Each point was calculated averaging 100 ns for setup B simulations ($decalin_{COM}$ – ▾; $N{-}H_{COM}$ – ◆; $p{-}C_{COM}$ – ✳; and $tail_{COM}$ – ×), and 50 ns for setup A' simulations ($decalin_{COM}$ – ▵; $N{-}H_{COM}$ – ▫; $p{-}C_{COM}$ – ◇; and $tail_{COM}$ – ○).

The change of orientation on passing from **7a** through to **7d** suggests an explanation for the enhanced activity along this sequence (see Table III-1). For compound **7a**, both transporter and chloride complex lie with their polar regions directed towards the water/lipid interface. They are thus able to form effective polar interactions, either with the phospholipid head groups (in the case of **7a** itself) or with water molecules that are bound to bulk aqueous phase (for both **7a** and **7a**·Cl^-). These polar interactions must be broken before the transporter/complex can pass through the membrane, and therefore represent potential barriers to transport. When hydrocarbon chains are added to the *para* positions of the phenyl groups, they tend to turn the transporter/complex, pulling the polar region away from the interface. The interactions with the interface are now easier to break, so that transport is facilitated. Although this logic can be applied to both transporter and chloride complex, the movement of the complex is probably more relevant, because it likely serves as the rate-determining step for the overall transport process. A related argument was previously used to explain the reason why "lipophilic balance" favours transport in thioureas **2**.[21] In that case, a polar unit at one end of a chain was expected to interact especially well with the membrane interface (although, unlike the present work, the hypothesis was not supported by calculations).

III.3.3. Interaction energies between transporters 7a–7f and POPC bilayer

The decrease in activity from **7d** to **7f** (see Table III-1) does not seem to result from the positioning of the receptors or complexes in the membranes. As shown in Figure III-11, the distance parameters for *p*-C_{COM} (◆), N–H_{COM} (■), and decalin$_{COM}$ (▲) remain fairly constant along the series. However, it is notable that the $P_{int}\cdots tail_{COM}$ distance (●) increases steadily from **7b** through to **7f**, implying that the hydrocarbon tails are extended into the membrane interior, generally aligned with the phospholipids (see also Figure III-10, right). Considering the geometry of the transporter/complex at the water/lipid interface, it is apparent that a 180° rotation must occur at some point to allow the complex formation or chloride release at the opposite interface. Such inversions were observed during the simulations, but only occasionally and much more often for transporters with shorter side-chains (especially, **7a**–**7c**, R=H, ethyl, *n*-butyl). The interaction energies between the POPC membrane model and the thiourea molecules were energetically evaluated through Eq. III-1:

$$\Delta E_{TOTAL} = \Delta E_{ELEC} + \Delta E_{vdW} + \Delta E_{INT} \qquad \text{Eq. III-1}$$

ΔE_{ELEC} and ΔE_{vdW} correspond to the non-bonded electrostatic and van der Waals energy terms, while ΔE_{INT} corresponds to the sum of bond, angle, and torsion energies. The three individual MM energy components are given by Eq. III-2 to Eq. III-4:

$$\Delta E_{ELEC} = E_{ELEC_{System}} - \left(E_{ELEC_{POPC}} + E_{ELEC_{7m}}\right) \qquad \text{Eq. III-2}$$

$$\Delta E_{vdW} = E_{vdW_{System}} - \left(E_{vdw_{POPC}} + E_{vdw_{7m}}\right) \qquad \text{Eq. III-3}$$

$$\Delta E_{INT} = E_{INT_{System}} - \left(E_{INT_{POPC}} + E_{INT_{7m}}\right) \qquad \text{Eq. III-4}$$

Here *System* represents the transporter and the phospholipid membrane discarding the water molecules and all chloride ions and sodium counterions. *POPC* stands for the 128 phospholipid molecules and **7*m*** (***m*** = **a**, **b**, **c**, **d**, **e** or **f**) for the isolated transporter. These three energies were estimated with snapshots extracted every 100 ps from the last 50 ns of each MD simulation carried out with setup A.

The bonded term (ΔE_{INT}) amounts to zero, given that the individual terms INT_{System}, INT_{POPC} and INT_{7m} were calculated using the same MD simulation. In these conditions, through Eq. III-1, the contribution of both non-bonded energy terms for the intermolecular interactions between the phospholipids and decaline transporters was evaluated. The average values of ΔE_{TOTAL}, ΔE_{ELEC} and ΔE_{vdW} for 200 ns of sampling are given in Table III-5 for each system and in Table B-3 (see Appendix B) for each MD run, and can be visualised in Figure III-13.

Table III-5. Average MM energy terms (kcal mol^{-1}) with the corresponding standard deviations.[a]

System ID	ΔE_{TOTAL}	ΔE_{ELEC}	ΔE_{vdW}
A.**7a**	−95.64 ± 15.90	−33.95 ± 12.45	−61.69 ± 6.45
A.**7b**	−108.21 ± 16.56	−38.84 ± 15.65	−69.37 ± 6.73
A.**7c**	−112.94 ± 14.59	−33.41 ± 11.34	−79.53 ± 6.51
A.**7d**	−113.26 ± 23.45	−29.06 ± 14.24	−84.20 ± 11.58
A.**7e**	−134.24 ± 14.78	−36.87 ± 11.28	−97.37 ± 6.64
A.**7f**	−127.66 ± 14.02	−25.10 ± 11.25	−102.55 ± 7.17

[a] N = 2000 (100 ps interval between snapshots).

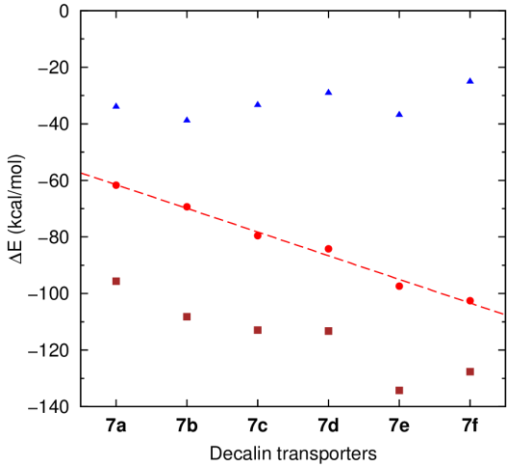


Figure III-13. Average ΔE_{TOTAL} (▪), ΔE_{ELEC} (▲), and ΔE_{vdW} (●) energies. Each point was calculated averaging the 200 ns of sampling for setup A simulations. A linear fit of the ΔE_{vdW} energies is shown as red dashed line (R^2 = 0.99), with a slope of −8 kcal mol^{-1} per ethylene unit.

The ΔE_{ELEC} values are almost independent of the transporter. This term also includes the hydrogen bonding interactions between the phospholipid heads and N−H binding sites, which were not found to differ over the range of transporters (see Table III-5), in agreement with the similar decalin$_{COM}$···P$_{int}$ and N−H$_{COM}$···P$_{int}$ distances reported in Figure III-11. In contrast, the ΔE_{vdW} term is the major contribution for the stabilisation of the transporters inserted into the phospholipid bilayers. As expected, a linear trend between the number of carbon atoms of alkyl chain substituent (and thus the total size of the molecule) and the ΔE_{vdW} was found, with the most negative stabilising energy between **7f**, bearing the longest alkyl substituents, and the phospholipids.

Thus, although the alkyl side-chains are flexible, they will tend to pack with the phospholipid in a manner that maximises the van der Waals interactions, providing a significant length-dependent barrier to rotation. Although the calculations focus on the free receptors, it is likely that the chloride complexes would behave similarly. When the side-chain extends, it is possible that reorientation of the receptor or (more likely) complex could become rate-determining, accounting for the trend from **7d** to **7f**. The increase in chain length could also hinder other types of movement, but the reorientation would appear to be especially critical.

III.4. Conclusions

MD simulations have provided detailed information on the movements and positions of the transporters **7a**–**7f** in POPC membrane models. For short alkyl substituents, the polar regions of the carriers/complexes point towards the water/lipid interface, in which they can interact with water or phosphate groups, hindering transport. For long alkyl substituents, the transporter/complex finds difficulty reorienting within the membrane when it passes between interfaces. The anion transport activity maximum occurs when the alkyl chains are long enough to help the transporter/complex escape the interface, but not so long as to hinder rotation.

Given enough time, all transporters/complexes locate at the water/lipid interface, but with different orientations. Thus, for **7a**, without an alkyl chain, the thiourea binding sites are directed towards the aqueous phase, interacting mainly with water molecules or phospholipid head groups, preventing the chloride uptake from the water phase. In contrast, the addition of alkyl chains of intermediate length (ethyl, **7b**, or *n*-butyl, **7c**) results in a not well-defined preferential orientation. In contrast, the *n*-hexyl alkyl chains of **7d**, with an almost parallel orientation regarding the phospholipid

tails, result in a vdW interaction energy consistent with the maximum anion transport rate experimentally observed. Further extending the alkyl chains (*n*-octyl, **7e**, and *n*-decyl, **7f**) leads to higher vdW interaction energies, resulting in more difficult reorientation of the complex in the membrane as well as the release of chloride.

III.5. References

1. A. P. Davis, D. N. Sheppard and B. D. Smith, *Chem. Soc. Rev.*, 2007, **36**, 348-357.
2. G. W. Gokel and N. Barkey, *New J. Chem.*, 2009, **33**, 947-963.
3. J. T. Davis, O. Okunola and R. Quesada, *Chem. Soc. Rev.*, 2010, **39**, 3843-3862.
4. S. Matile, A. Vargas Jentzsch, J. Montenegro and A. Fin, *Chem. Soc. Rev.*, 2011, **40**, 2453-2474.
5. H. Valkenier and A. P. Davis, *Acc. Chem. Res.*, 2013, **46**, 2898-2909.
6. P. A. Gale, R. Perez-Tomas and R. Quesada, *Acc. Chem. Res.*, 2013, **46**, 2801-2813.
7. N. Busschaert, C. Caltagirone, W. Van Rossom and P. A. Gale, *Chem. Rev.*, 2015, **115**, 8038-8155.
8. F. M. Ashcroft, *Ion Channels and Disease*, Academic Press, London, 1st edn., 2000.
9. J. R. Broughman, L. P. Shank, W. Takeguchi, B. D. Schultz, T. Iwamoto, K. E. Mitchell and J. M. Tomich, *Biochemistry*, 2002, **41**, 7350-7358.
10. A. V. Koulov, T. N. Lambert, R. Shukla, M. Jain, J. M. Boon, B. D. Smith, H. Li, D. N. Sheppard, J. B. Joos, J. P. Clare and A. P. Davis, *Angew. Chem. Int. Ed. Engl.*, 2003, **42**, 4931-4933.
11. B. A. McNally, A. V. Koulov, T. N. Lambert, B. D. Smith, J. B. Joos, A. L. Sisson, J. P. Clare, V. Sgarlata, L. W. Judd, G. Magro and A. P. Davis, *Chem. Eur. J.*, 2008, **14**, 9599-9606.
12. L. W. Judd and A. P. Davis, *Chem. Commun.*, 2010, **46**, 2227-2229.
13. N. Busschaert, I. L. Kirby, S. Young, S. J. Coles, P. N. Horton, M. E. Light and P. A. Gale, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 4426-4430.
14. S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103-117.
15. J. A. Cooper, S. T. Street and A. P. Davis, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 5609-5613.
16. H. Valkenier, L. W. Judd, H. Li, S. Hussain, D. N. Sheppard and A. P. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 12507-12512.
17. A. Vargas Jentzsch, D. Emery, J. Mareda, P. Metrangolo, G. Resnati and S. Matile, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 11675-11678.
18. S. Bahmanjah, N. Zhang and J. T. Davis, *Chem. Commun.*, 2012, **48**, 4432-4434.
19. S. J. Edwards, H. Valkenier, N. Busschaert, P. A. Gale and A. P. Davis, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 4592-4596.
20. V. Saggiomo, S. Otto, I. Marques, V. Felix, T. Torroba and R. Quesada, *Chem. Commun.*, 2012, **48**, 5274-5276.
21. H. Valkenier, C. J. E. Haynes, J. Herniman, P. A. Gale and A. P. Davis, *Chem. Sci.*, 2014, **5**, 1128-1134.
22. M. J. Spooner and P. A. Gale, *Chem. Commun.*, 2015, **51**, 4883-4886.
23. C. J. Haynes, N. Busschaert, I. L. Kirby, J. Herniman, M. E. Light, N. J. Wells, I. Marques, V. Felix and P. A. Gale, *Org Biomol Chem*, 2014, **12**, 62-72.
24. S. Hussain, P. R. Brotherhood, L. W. Judd and A. P. Davis, *J. Am. Chem. Soc.*, 2011, **133**, 1614-1617.
25. S. J. Edwards, I. Marques, C. M. Dias, R. A. Tromans, N. R. Lees, V. Felix, H. Valkenier and A. P. Davis, *Chem. Eur. J.*, 2016, **22**, 2004-2011.
26. H. Valkenier, C. M. Dias, K. L. Porter Goff, O. Jurcek, R. Puttreddy, K. Rissanen and A. P. Davis, *Chem. Commun.*, 2015, **51**, 14235-14238.
27. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
28. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
29. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman (2014), AMBER 14, University of California, San Francisco.
30. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
31. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
32. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
33. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
34. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
35. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171-179.
36. D. J. Chadwick and J. D. Dunitz, *J Chem Soc Perk T 2*, 1979, 276-284.

37. Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
38. A. K. Rappe and C. J. R. Casewit, *Molecular Mechanics Across Chemistry*, Univ Science Books 1st edn., 1997.
39. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
40. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
41. I. S. Joung and T. E. Cheatham, 3rd, *J. Phys. Chem. B*, 2008, **112**, 9020-9041.
42. L. Martinez, R. Andrade, E. G. Birgin and J. M. Martinez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
43. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
44. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
45. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
46. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
47. H. Li, H. Valkenier, L. W. Judd, P. R. Brotherhood, S. Hussain, J. A. Cooper, O. Jurcek, H. A. Sparkes, D. N. Sheppard and A. P. Davis, *Nat Chem*, 2016, **8**, 24-32.

Chapter IV.

Atomistic characterisation of squaramide derivatives as anion transmembrane transporters by in silico *investigations*

Summary

A comprehensive experimental and theoretical investigation on chloride transmembrane transport promoted by four series of squaramide derivatives (see Scheme IV-1), with different fluorination degrees, number of convergent N−H binding units and conformational shapes is reported. The experimental chloride binding and transport abilities of these small synthetic molecules were straightforwardly rationalised with quantum descriptors (E^2 energy values, derived from the 2nd Order Perturbation Theory Analysis, and $V_{S,max}$ values, the most positive values on the distribution of the surface electrostatic potential) estimated via DFT calculations, followed by Molecular Dynamics (MD) simulations in POPC bilayers. The *tren*-based molecules **D**, with three squaramide binding motifs, present the highest values for E^2 and $V_{S,max}$, resulting in high affinity for chloride and strong interactions with phospholipid head groups. These tripodal molecules are able to form transient chloride complexes in water and permeate the phospholipid bilayer. On the other hand, when inside of the membrane model, the encapsulated chloride ion is sheltered from water molecules and is not released to the aqueous phase, in agreement with the absent transport activity experimentally assessed. In contrast, the linear molecules **A**, with a single squaramide binding motif, having moderate values of E^2 and $V_{S,max}$, are able to uptake and release chloride either in water phase or at the water/lipid interface level. Therefore, these molecules enable the chloride transmembrane transport, in line with the experimentally found high transport activity. The potential of mean force (PMF) profiles associated with the diffusion of free transporters **A** and their chloride complexes across phospholipid bilayers were estimated, showing that the entrance to the bilayer is energetically favoured, while the exit through the opposite interface is thermodynamically disfavoured. Thus, in the anion carrier mechanism, the free or complexed squaramides must only shuttle between interfaces, without leaving the phospholipoid bilayer.

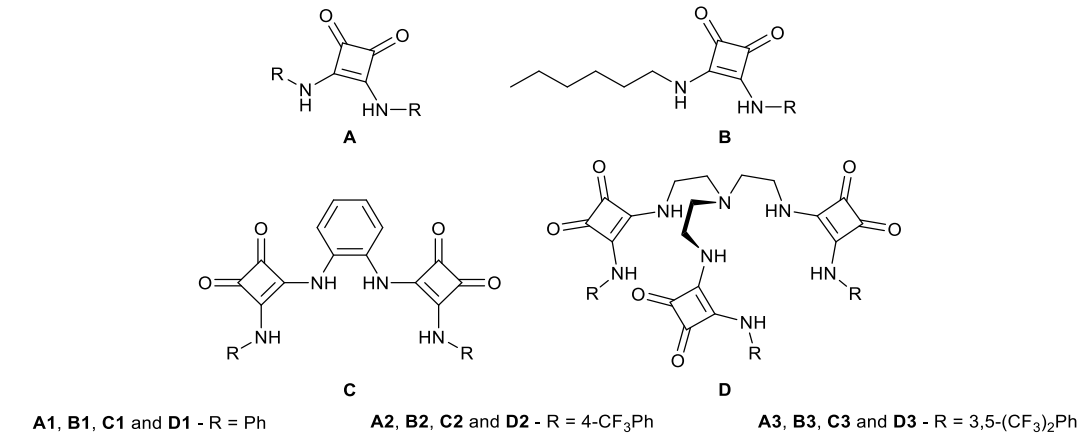
The work reported in this chapter is currently submitted for publication.

IV.1. Introduction

The transport of relevant biological ions across of membrane phospholipid bilayers is a vital process that relies in membrane channels.[1, 2] A wide range of disorders involving membrane channels occurs when they are deficient or absent in the cell membrane.[1, 2] For instance, the impairment of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), an important transmembrane channel, affects both the transport of chloride and bicarbonate throughout the organism, resulting in chronic pancreatitis, meconiumileus, male infertility and Cystic Fibrosis (CF).[2] The symptoms associated with chloride's deficient transmembrane transport further extend to epilepsy, Startle disease, deafness, blindness, lysosomal storage and neurodegeneration, osteopetrosis, lung infections, renal salt loss, and kidney stones.[3] Channel replacement therapies, that rely on anion ionophores to promote the anion transmembrane transport, are currently being developed. Faced with this challenge, the supramolecular Chemistry community has been developing, over the last decades, an expanding number of putative anion transporters.[4-16] These molecules exploit different combinations of binding units and scaffolds (*e.g.*, cholapods, and *tren*-based molecules, using (thio)ureas; tambjamines; anion-π slides; or even more recently molecules that rely on non-conventional halogen[17-20] and chalcogen bonds[21]) resulting in anion transporters that function either as mobile carriers or as artificial anion channels.

Squaramide derivatives are well characterised as ditopic receptors, both experimentally and computationally.[22-34] Squaramide-based receptors, decorated with aromatic motifs, are able to establish strong hydrogen bonding interactions with halides, which are assisted by the convergent hydrogen bonds from the neighbouring C–H groups of the aryl substituents.[32] Consequently, these squaramide derivatives show a greater binding affinity for halides then the analogous urea-based receptors.[32] Given their abilities as anions receptors, as well as their applicability in the field of medicinal Chemistry,[35-40] squaramide derivatives have been studied as anion transporters.[41]

Squaramides **A1**, **A2** and **A3** (see Scheme IV-1) were shown to be superior anion transporters than the corresponding urea and thiourea analogous, given that it was the possible to improve their binding affinities without significantly changing their lipophilicities.[41] The application of the squaramide derivatives has also been expanded to the transport of amino-acids.[42] On the other hand, the incorporation of the squaramide motifs in a steroidal framework has produced exceptional anion receptors, but poor anion transporters.[43] Further tuning of the simple oxosquaramide derivatives **A1**–**A3** involved their thionation to afford analogous thiosquaramides, which revealed the ability to enable/disable the anion transport as a function of the medium's pH,[44] further increasing the interest of applications containing the squaramide binding motifs.



Scheme IV-1. Squaramide derivatives investigated in this work.

MD simulations are a powerful tool to obtain the energetic and structural insights on transmembrane transport,[45] including drug delivery and ion traffic through protein channels.[46, 47] While the *in silico* studies on these topics are quite common, the MD studies on anion transport promoted by small synthetic molecules are still scarce. Indeed, our group has pioneered this field with the seminal paper on the diffusion of bis-indolylureas in a POPC bilayer,[48] showing the importance of the interactions between these molecules and the phospholipid head groups in the anion transmembrane transport. Furthermore, a proof of concept study with a cationic calix[4]arene in a DOPC bilayer allowed us to observe the passive transport of chloride assisted by a synthetic anion carrier for the first time.[49] We have extended the MD investigations with membrane models to the study of the interaction of *ortho*-phenylenediamine-based bis-urea molecules[50] and *tren*-based tris-thiourea compounds[51] with POPC bilayers. More recently, MD studies on a series of bis-thioureidodecalin transporters with different alkyl chains lengths allowed to establish a relation between their transport abilities and their tilting and tumbling within the bilayer, which are slowed down by extending the alkyl chains *via* van der Waals interactions with phospholipid tails.[52]

This chapter presents a comprehensive investigation on the anion transport properties of twelve structurally related neutrogenic squaramide-based molecules, sketched in Scheme IV-1, including the already reported **A1**, **A2** and **A3**.[41] While this series contain a single squaramide motif flanked by either two aryl substituents, in the **B** series, one of the aromatic groups is replaced by an *n*-hexyl alkyl substituent. The **C** series' molecules are bis-squaramides based on an *ortho*-phenylenediamine scaffold, while the **D** series have three squaramide binding motifs appended to a *tren* scaffold. Overall, these series allow the systematic study of the effects of fluorination on the lipophilicity and anion binding strength, and how they can affect anion transport. Furthermore, the **C** and **D** molecules offer the possibility to investigate how the different number of binding units and their conformational dispositions affect anion binding and transport.

The transport properties of six of these molecules (**A2**, **B2**, **C2**, **C3**, **D2** and **D3**) have been previously related with their ability to induce apoptosis in cancer cell-lines.[53] In this work, we report further insights on the squaramide derivatives' transport properties, derived from a synergetic theoretical and experimental approach, with a clear focus on the structural and energetic features of these four series of molecules as anion carriers.

IV.2. Methods

IV.2.1. Quantum mechanics calculations

All quantum calculations herein reported were carried out with *Gaussian 09*,[54] using different theory levels, as described along this work. These calculations include the geometry optimisations of the free squaramide derivatives and their chloride complexes, calculation of the E^2 interaction energies and calculation of the distribution of electrostatic potential $V(\mathbf{r})$, as thoroughly detailed below.

$V(\mathbf{r})$ were computed on the molecular surfaces of the squaramide derivatives from previously optimised structures of their chloride complexes, after removal of the anion, using the same level of theory. Subsequently, the $V(\mathbf{r})$ was evaluated on the 0.001 electrons Bohr^{-3} contour of $\rho(\mathbf{r})$, and is henceforth labelled $V_S(\mathbf{r})$. The electrostatic potential surface ranges, including the most negative and most positive values ($V_{S,min}$ and $V_{S,max}$, respectively), were ascertained using *Multiwfn*.[55, 56]

The E^2 energy derived from the 2nd Order Perturbation Theory Analysis was ascertained from the DFT optimised structures of the chloride squaramide complexes, using the *Natural Bond Orbital* (*NBO*) Version 6.0 program,[57-59] at the same level of theory.

IV.2.2. MD simulations

The MD simulations were carried out with AMBER 14[60] (passive diffusion), NAMD 2.9[61] (Steered Molecular Dynamics – SMD) or AMBER 2016[62] (Umbrella Sampling – US), with resort to GPU acceleration.[63-65] The Lipid14[66] force field was employed for the POPC lipids, while parameters from the General AMBER Force Field (GAFF)[67, 68] and atomic RESP charges[69] were used for the squaramide derivatives, as follows.

The initial structures of **A1**, **A2**, **A3** and **D1** were obtained directly from the crystal structures deposited with the Cambridge Crystallographic Data Centre (CCDC),[70, 71] with the RefCodes FAXBUD, FAXCEO, FAWZAG,[41] and WELXUI, respectively.[72] The structures of the **B** subset were obtained from the corresponding **A** analogues, with the replacement of a phenyl moiety by an alkyl chain. **C1** was generated by atomic manipulation of the crystal structure with the RefCode DETYAD,[73] with the inclusion of necessary phenyl moieties. The structures of **C2**, **C3**, **D2** and **D3**, due to their complexity, were generated at a later stage (*vide infra*).

Subsequently, these structures were optimised at the HF/6-31G* level with *Gaussian 09*.[54] The optimisation of these eight individual structures was followed by a single point calculation to generate the electrostatic potential (ESP) at the same theory level, using the Merz–Singh–Kollman scheme with 4 concentric layers per atom and 6 density points in each layer (IOp (6/33 = 2, 6/41 = 4, 6/42 = 6)). The initial atomic charges of each molecule were then calculated by RESP fitting, along with the attribution of GAFF atom types, using the *antechamber* module,[74] as implemented in the AMBER package.

To obtain atomic charges less dependent of the molecular conformation or orientation, the calculation of the final RESP charges employed in all membrane MD simulations was preceded by conformational analyses on the squaramide derivatives, as follows: The initial Molecular Mechanics (MM) energy minimised structures, using the initial RESP charges, were heated at high temperature in the gas-phase for 50 ps, followed by collection runs of 0.1 ns for the individual molecules of series **A** and **B** and 2.0 ns for molecules **C1** and **D1**, in agreement with the structural complexity of each subset, and using a time step of 1.0 fs, allowing the stochastic search of the conformational space. Frames were saved every 0.1 ps leading to trajectory files containing 1000 structures for series **A** and **B** and 20000 structures for molecules **C1** and **D1**. All these structures were further minimised by MM using a steepest descendent gradient followed by the conjugate gradient algorithm, until the convergence criterion of 0.0001 kcal mol^{-1} $Å^{-1}$ was attained. Afterwards, the MM minimised structures were clustered with *UCSF Chimera*.[75] From the resulting clusters, three representative conformations with substantially different RMSD values were selected for **A1**, while for the remaining mono-squaramide derivatives four conformations were selected. Five and eight individual conformations were selected for **C1** and **D1**, respectively. The fluorinated analogous of these two subsets were obtained from **C1** and **D1**, adding the suitable number of $-CF_3$ groups. All selected conformations had the N–H binding units of the squaramide moieties adopting a *syn* configuration, but different spatial dispositions of the phenyl substituents. All these structures underwent further HF/6-31G* geometry optimisations and ESP calculations as described above. The individual ESP data were extracted from the corresponding *Gaussian 09*[54] outputs, concatenated, and subsequently used to generate the input files for the two-stage RESP fitting, using identical weights for all conformations.

The structures of the chloride complexes of the individual squaramide derivatives in subsets **A**, **B** and **C**, as well as **D1**, were obtained in gas-phase *via* conformational analyses as described above for the free receptors. However, the production runs were carried out for 5.0 ns, leading, at the end of the quenched dynamics protocol, to 50000 MM minimised structures. Moreover, for series **A** and **B** (apart **A3**, with four $-CF_3$ moieties), a 5.0 kcal mol^{-1} $Å^{-2}$ distance restraint between the nitrogen atoms of the N–H binding units and the anion was employed. The lowest energy structure of each complex was

selected for DFT optimisation at the M06-2X/6-31+G** level of theory. The initial geometries of the chloride complexes of **D2** and **D3** were generated from **D1**, as described above. The quantum optimised structures of the squaramide chloride complexes were used in further single point quantum calculations (such as the determination of the E^2 energy values or the distribution of the electrostatic potential, *vide supra*), as well as starting binding arrangements in MD simulations in POPC bilayer model.

The MD membrane simulations were carried out with the chloride complexes immersed in the water slab of a free membrane with *Packmol*,[76] affording a starting scenario W. This previously equilibrated membrane system is a POPC membrane model, with an orthorhombic shape, and composed of 128 phospholipids, 6500 TIP3P model water molecules,[77] 18 Cl^- and 18 Na^+ ions (*ca.* 0.15 M). The ions were described with van der Waals parameters developed to be used along with the TIP3P water model.[78] Due to the insertion of the chloride complexes, an additional Na^+ counterion was also added to the system, to neutralise the net charge, without significant impact on the NaCl concentration.

The MD simulations were carried out as follows: the initial configuration of each system was submitted to 20000 steps of MM energy minimisation with a 500 kcal mol^{-1} $Å^{-2}$ positional restraint on the chloride complex and lipid molecules, through the steepest descent algorithm for 10000 steps plus 10000 steps of the conjugated gradient algorithm, followed by the relaxation of the entire system for another 20000 steps, with the same protocol. The equilibration of the system proceeded by heating it to 303 K in an NVT ensemble for 100 ps with a 10 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex and lipid molecules. The equilibration stage proceeded with a 5.0 ns run using an NPT ensemble with a 5.0 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex. Then, the positional restraint was removed and the simulation continued for further 200 ns. The long-range electrostatic interactions were described with the Particle Mesh Ewald (PME) algorithm[79] using a real-space cut-off at 10 Å. The cut-off for the Lennard−Jones interactions was also set at 10 Å. The temperature of the system was maintained at 303 K, using Langevin dynamics,[80] with a collision frequency γ of 1.0 ps^{-1}. The pressure was controlled by the Berendsen barostat[81] at 1 atm and compressibility of 44.6×10^{-6} bar^{-1}, with a relaxation time of 1.0 ps. The covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm,[82] allowing the use of a 2.0 fs time step. The MD simulation trajectory frames were saved every 10 ps.

The MD simulations were also carried with the chloride complexes inserted into the core of the bilayer, affording scenario M. In these simulations, the membrane model accounts only 4040 water molecules and a total of 11 Na^+ and 11 Cl^- ions, maintaining the ionic strength. The simulation protocol was the same given for simulations of scenario W. At least two independent runs were carried out for the planar squaramides (series **A**, **B** and **C**), while for the tripodal chloride complexes two independent runs were carried out for each initial orientation of the complex, either vertically or horizontally to the membrane normal.

Moreover, for selected molecules, from the end of the equilibration period of the first MD run in both scenarios, an independent MD production run of at least 200 ns was undertaken with the chloride complex maintained with 3.5 Å distance restraints between the nitrogen atoms of the squaramide moieties and the chloride ion,‡‡ using harmonic restraints of 5.0 kcal mol^{-1} $Å^{-2}$. Table IV-1 summarises the number of MD runs, as well as the time lengths, for all these simulations.

‡‡ The distance restraints were selected in agreement with the N−H···Cl^- dimensions found in single crystal X-ray structures deposited with CCDC[70, 71] of squaramide-based chloride complexes.[32, 41, 83, 84]

Table IV-1. Summary of the MD simulations carried out with squaramide chloride complexes in POPC membrane models.

Squaramide	Starting Scenario							
	W		W'		M		M'	
	Simulation IDs[a]	Time	Simulation IDs[a]	Time	Simulation IDs[a]	Time	Simulation IDs[a]	Time
A1	W_1.**A1**, W_2.**A1**	200	W'$_1$.**A1**	200	M_1.**A1**, M_2.**A1**	200	M'$_1$.**A1**	200
A2	W_1.**A2**, W_2.**A2**	200	W'$_1$.**A2**	200	M_1.**A2**, M_2.**A2**	200	M'$_1$.**A2**	200
A3	W_1.**A3**, W_2.**A3**	200	W'$_1$.**A3**	200	M_1.**A3**, M_2.**A3**	200	M'$_1$.**A3**	200
B1	W_1.**B1**, W_2.**B1**	200	W'$_1$.**B1**	200	M_1.**B1**, M_2.**B1**	200	M'$_1$.**B1**	200
B2	W_1.**B2**, W_2.**B2**	200	W'$_1$.**B2**	200	M_1.**B2**, M_2.**B2**	200	M'$_1$.**B2**	200
B3	W_1.**B3**, W_2.**B3**	200	W'$_1$.**B3**	200	M_1.**B3**, M_2.**B3**	200	M'$_1$.**B3**	200
C1	W_1.**C1**, W_2.**C1**	200	–	–	M_1.**C1**, M_2.**C1**	200	–	–
C2	W_1.**C2**, W_2.**C2**	200	–	–	M_1.**C2**, M_2.**C2**	200	–	–
C3	W_1.**C3**, W_2.**C3**	200	–	–	M_1.**C3**, M_2.**C3**	200	–	–
D1	W_1.**D1**, W_2.**D1**	200	–	–	M_1.**D1**$_V$, M_2.**D1**$_V$, M_1.**D1**$_H$, M_2.**D1**$_H$[b]	500	–	–
D2	W_1.**D2**, W_2.**D2**	200	–	–	M_1.**D2**$_V$, M_2.**D2**$_V$, M_1.**D2**$_H$, M_2.**D2**$_H$[b]	500	–	–
D3	W_1.**D3**, W_2.**D3**	200	–	–	M_1.**D3**$_V$, M_2.**D3**$_V$, M_1.**D3**$_H$, M_2.**D3**$_H$[b]	500	–	–

[a] Each simulation ID results from the combination of the initial position of the transporter in the membrane system, the run number and the transporter. For instance, W_1.**A1** stands for the first MD run carried out with **A1** in setup W; [b] The chloride complexes of **D1**, **D2** and **D3** were positioned within the membrane core either perpendicular to the bilayer normal (subscript H) or parallel to it (subscript V).

The final frame of the equilibration period of the W scenario of molecules **A1**, **A2** and **A3** was used as the starting frame for the SMD simulations. Two approaches were employed: diffusion of the free receptor or diffusion of the anion complex of each squaramide derivative. The free receptors were dragged across the lipid bilayer normal (*viz*, *z*-dimension, the reaction coordinate) at 2.5 Å ns^{-1} in 32 ns long MD simulations, while the anion complexes were dragged at 5.0 Å ns^{-1} throughout a MD simulation with 16 ns, in the NPT ensemble. Throughout the travelled 80 Å along the membrane system, a force constant of 5.0 kcal mol^{-1} Å^{-2} was applied to the non-hydrogen atoms of the squaramide derivatives, while the anion complex was preserved with the use of two harmonic distance restraints between the nitrogen atoms and the chloride ion (*vide supra*), a feature that was only possible with the use of collective variables in NAMD 2.9.[61] In a frame from the equilibrated free membrane, a solvated chloride ion was pulled along the bilayer normal at 5.0 Å ns^{-1} in a 16 ns long MD simulation in the NPT ensemble. A force constant of 5.0 kcal mol^{-1} Å^{-2} was applied to the anion along the 80 Å of the membrane system.

From the SMD simulations of the free or anion associated **A** molecules, as well as the free chloride ion, several starting configurations were selected as starting points for the independent US simulations. The spacing between the centre of mass (COM) of the squaramide derivative or of the free anion in the US simulations is *ca.* 1.0 Å, with a total of 67 evenly spaced independent starting points (ranging from +33 to −33 Å along the *z*-dimension). For the squaramide derivatives, each one of the 67 windows was simulated for 50 ns, with the first 30 ns being discarded as equilibration period, with a positional restraint of 3.0 kcal mol^{-1} Å^{-2} applied only in the *z*-dimension, a feature that has only become available in AMBER 16.[62] Moreover, in the US simulations of the squaramide complexes, two 5.0 kcal mol^{-1} Å^{-2} harmonic distance restraints between the anion and the nitrogen atoms were applied. Given that the initial structures for the US windows of the squaramide complexes had been generated in NAMD,[61] each independent window MD simulation was preceded by a minimisation and heating stage,

with the same distance restraints of the production run. The remaining details are as given for the passive diffusion simulations, apart of the positional restraints. In the US simulations of the free anion, also restrained with a 3.0 kcal mol^{-1} $Å^{-2}$ force applied along the *z*-dimension, the 67 independent windows were simulated for 25 ns, and only the final 10 ns were considered for further analysis. The production trajectory frames were saved every 10 ps, while the distance between the COM of the squaramide derivative or chloride and the COM defined by phosphorus atom of the lipid head groups and the terminal CH_3 carbon atoms of the POPC bilayer was saved every 50 steps.

IV.3. Results and Discussion

IV.3.1. DFT calculations

Figure IV-1 shows the structures of the squaramide chloride complexes optimised in gas-phase using the M06-2X functional and the 6-31+G** level of theory with *Gaussian 09*.[54] The calculated N···Cl^- distances and N–H···Cl^- angles are listed in Table IV-2, as well as the values of the individual N_{tren}–C–C–N torsion angles, starting at the tertiary nitrogen atom (N_{tren}) of the tripodal receptors **D**.

The **A** series chloride complexes exhibit symmetric structures, with equal N–H···Cl^- hydrogen bonding dimensions, which are similar to those found in the crystal structures of these complexes.[41] On the other hand, the C–C covalent bond lengths within the squaramide motif are in close agreement with the crystallographic ones, apart from the C–C distance between the carbonyl groups, which is 0.031 Å longer in the DFT optimised structures. A detailed crystal packing analysis of these three chloride complexes shows the existence of intermolecular C–H···O=C bonding contacts which can induce slight distortions on the bond lengths and angles within the squaramide motif (see Appendix C).[41]

Table IV-2. Hydrogen bonds dimensions in the DFT optimised chloride complexes of the squaramide series, along with the N_{tren}–C–C–N torsion angles of the **D** subset.

Squaramide	N···Cl⁻ distance (Å)	N–H···Cl⁻ angle (°)	N_{tren}–C–C–N torsion angle (°)
A1	3.147 ; 3.147	175.2 ; 175.2	–
A2	3.120 ; 3.120	174.9 ; 174.9	–
A3	3.091 ; 3.091	174.4 ; 174.4	–
B1	3.168 ; 3.155	164.9 ; 172.9	–
B2	3.185 ; 3.130	163.9 ; 173.7	–
B3	3.173 ; 3.088	163.2 ; 173.3	–
C1	3.264 ; 3.257	177.0 ; 175.9	–
	3.264 ; 3.257	177.1 ; 175.9	–
C2	3.267 ; 3.247	176.2 ; 176.1	–
	3.266 ; 3.248	176.2 ; 176.1	–
C3	3.319 ; 3.199	169.2 ; 177.7	–
	3.235 ; 3.211	166.8 ; 175.1	–
D1	3.367 ; 3.297	163.9 ; 170.4	64.5
	3.359 ; 3.294	163.7 ; 170.6	64.9
	3.372 ; 3.300	163.2 ; 171.2	64.8
D2	3.376 ; 3.288	162.9 ; 170.3	64.3
	3.372 ; 3.285	162.6 ; 170.3	64.1
	3.368 ; 3.289	162.5 ; 170.6	64.5
D3	3.376 ; 3.262	161.7 ; 167.9	64.1
	3.381 ; 3.262	161.6 ; 169.2	64.0
	3.382 ; 3.264	160.7 ; 168.9	64.3

For the **B** series, due to its asymmetric structure, the lengths of the N–H···Cl⁻ hydrogen bonds are slightly different, with the N–H unit closer to the phenyl substituent establishing the stronger interaction, as it always has the shorter length and the more linear angle. Concerning the series with the bis-squaramide units (**C**), the two N–H binding units adjacent to the peripheral aromatic substituents have shorter N···Cl⁻ distances than the ones near the central bridging 1,2-phenylene moiety. In the optimised structures of the **D** tripodal complexes, the chloride ion is surrounded by the three-squaramide moieties, establishing six hydrogen bonds, with the three N···Cl⁻ distances adjacent to the phenyl substituents being systematically shorter than the remaining ones. Moreover, the three N_{tren}–C–C–N torsion angles, with values of *ca.* 64.4°, are consistent with a tripodal *ggg* (*g* = gauche) conformation.[51]

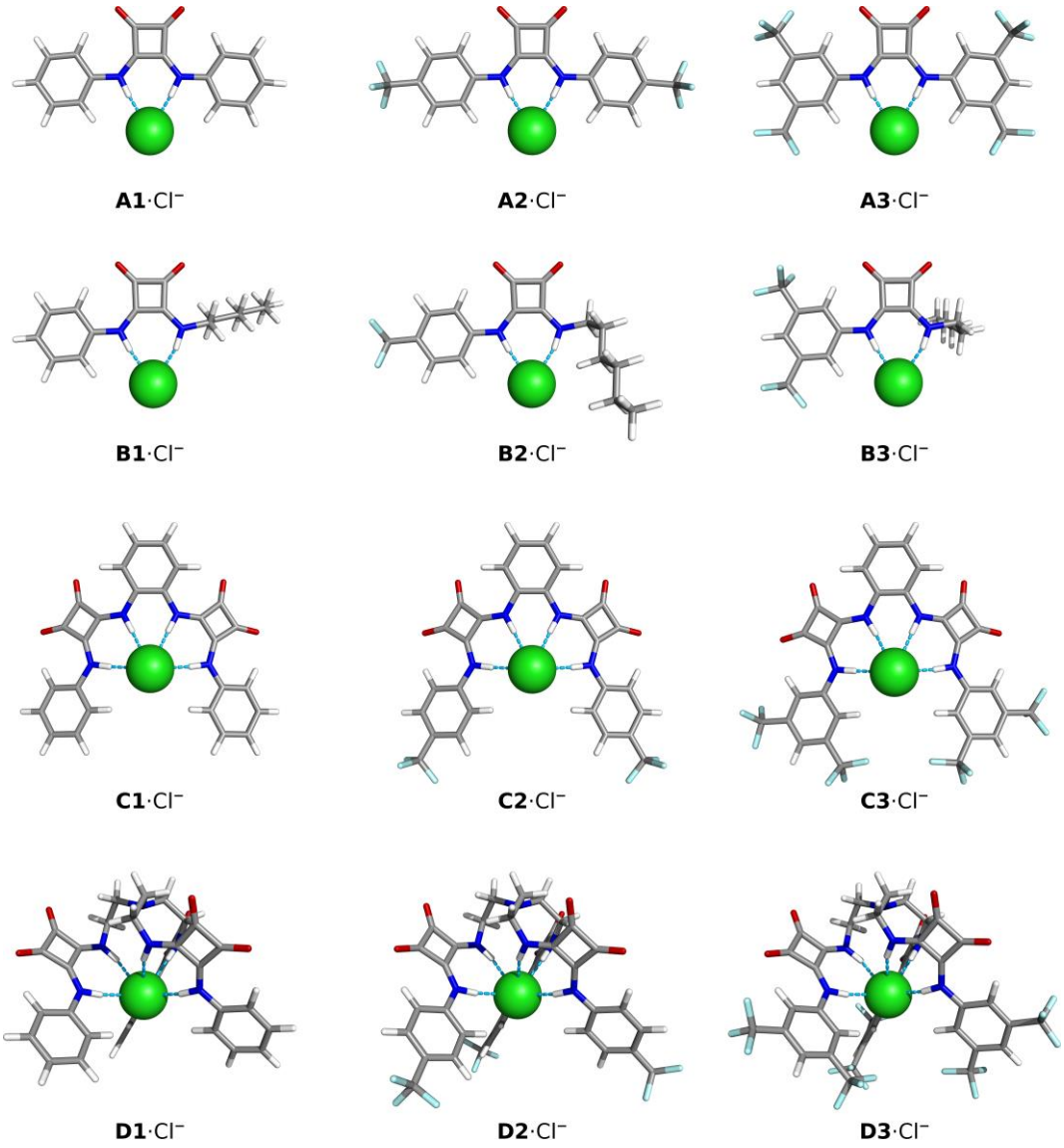


Figure IV-1. DFT optimised structures of the twelve squaramide chloride complexes. The hydrogen, carbon, oxygen, nitrogen, fluorine atoms and chloride ion are shown in white, grey, red, blue and cyan sticks, respectively, while the anion is shown as a green sphere.

Across each individual series, the N–H···Cl⁻ interactions become stronger with the growing fluorination of the phenyl substituents. Subsequently, the effect of the different degrees of fluorination were ascertained using the E^2 energy values as well as $V_S(\mathbf{r})$. The E^2 values, listed in Table IV-3, show that the stabilisation energy of the squaramide complexes increases with the fluorination of the phenyl substituents, as a function of their electron withdrawing nature. The $V_S(\mathbf{r})$ of the squaramide ligands in their chloride complexes (*vide supra*) was determined at the same level of theory and is shown in Figure IV-2 for all molecules. The linear squaramides of the **A** and **B** series display a well-defined positive region (red) centred at the N–H binding units while the most negative region (blue) covers the oxygen atoms of the carbonyl groups. In the molecules of the **C** series the red region encompasses the four acidic N–H binding units of the two squaramide motifs, while the carbonyl groups are surrounded by blue areas of negative potential. In the tripodal molecules, the most positive region is located inside the binding pocket enclosing all six N–H binding units. On the other hand, in **D2** and **D3**, the areas of negative potential, in addition to the carbonyl groups, also include the $-CF_3$ moieties. The negative potentials (blue area) around the oxygen atoms of the carbonyl groups enable the squaramide motif to establish

hydrogen bonds with putative hydrogen bond acceptors, such as the water molecules in a membrane system (*vide infra*) or other squaramide motifs, as observed in crystalline state.[41,85] The maximum values of $V_S(\mathbf{r})$, $V_{S,max}$, listed in Table IV-3, mainly reflect the number and the position of the $-CF_3$ electron withdrawing groups on the squaramide's phenyl substituents. Oddly enough, for **C1** and **D1** the $V_{S,max}$ values are comparable, despite the presence of another binding unit in the tripodal receptor.

Table IV-3. E^2 interaction energy values (kcal mol^{-1}) for the N–H···Cl$^-$ hydrogen bonds in the DFT optimised structures of the chloride complexes of the squaramide series, along with $V_{S,max}$ (kcal mol^{-1}) values ascertained for the squaramide ligands in their chloride complexes.

Squaramide	E^2	$V_{S,max}$
A1	65.60	77.27
A2	72.74	88.32
A3	81.23	94.78
B1	60.06	72.62
B2	60.92	78.22
B3	68.38	81.58
C1	79.25	99.66
C2	80.44	107.30
C3	84.38	110.49
D1	85.77	99.71
D2	85.92	110.46
D3	87.86	114.93

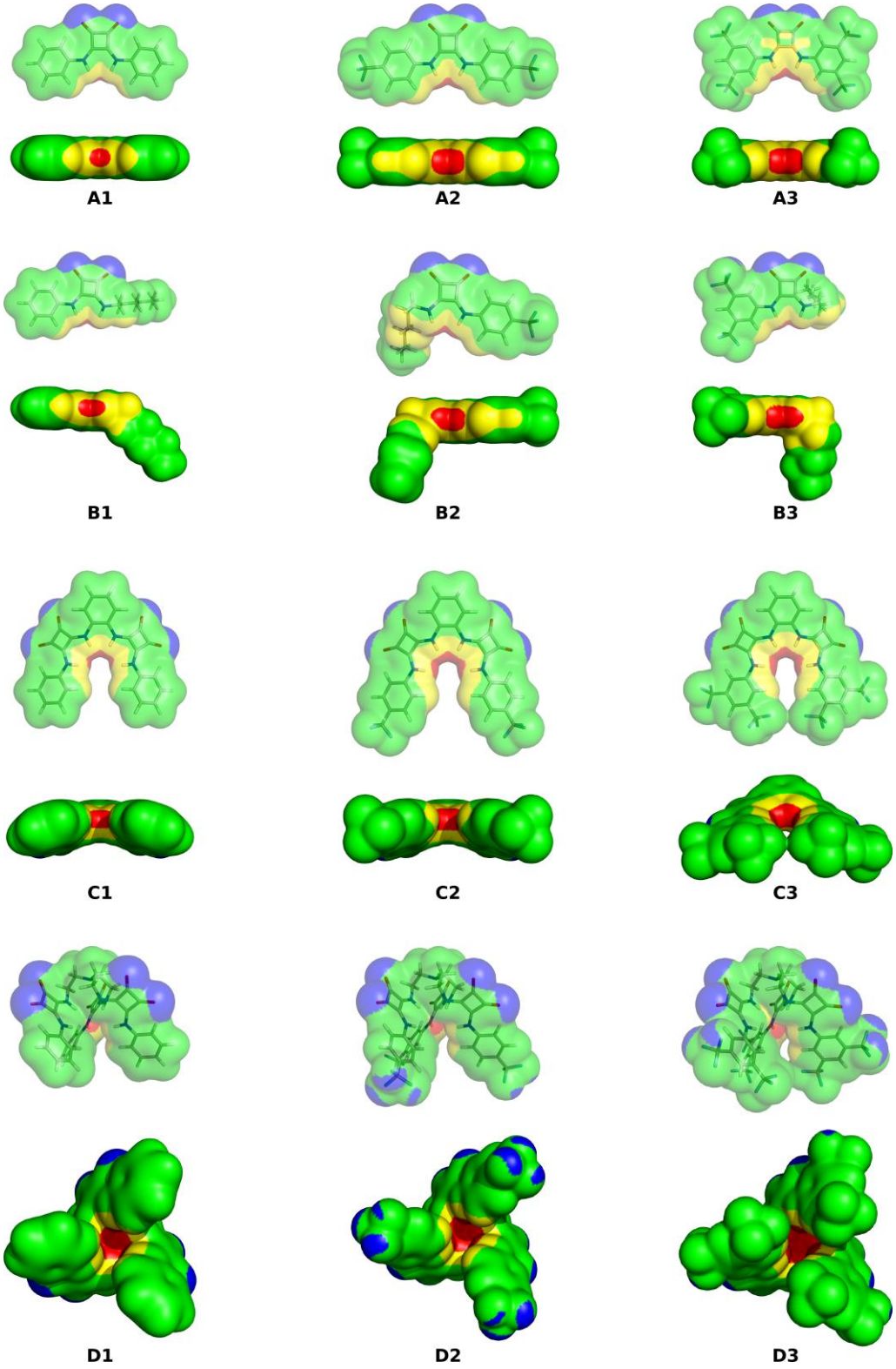


Figure IV-2. Distribution of the electrostatic potential mapped on the 0.001 electrons Bohr^{-3} isodensity surface of the squaramide derivatives top and front views. The colour scales, in kcal mol^{-1}, are as follows: **A** series: blue – lower than −15.0; green – from −15.0 to 25.0; yellow – from 25.0 to 65.0; and red – greater than 65.0; **B** series: blue – lower than −17.5; green – from −17.5 to 20.0; yellow – from 20.0 to 57.5; and red – greater than 57.5; **C** series: blue – lower than −12.5; green – from −12.5 to 35.0; yellow – from 35.0 to 82.5; and red – greater than 82.5; **D** series: blue – lower than −7.5; green – from −7.5 to 40.0; yellow – from 40.0 to 87.5; and red – greater than 87.5.

As expected, the E^2 values of the chloride complexes increase along with the $V_{S,max}$ of the free receptors, being linearly related (R^2 = 0.90). On the other hand, only the E^2 or the $V_{S,max}$ values of the **A** series linearly increase with the natural logarithm of the corresponding binding affinity constants (K_a, see Table IV-4), as can be seen in Figure IV-3 (R^2 = 0.96 for E^2 and R^2 = 1.00 for $V_{S,max}$). These quantum descriptors are also correlated with the anion transport properties of the **A** series (ln($EC_{50,270s}$[§§]), with R^2 = 0.80 for E^2 and R^2 = 0.93 for $V_{S,max}$), as previously observed for a series of simple acylthioureas.[86] In stark contrast, such correlations were not found for the **C** series.

Table IV-4. Summary of the chloride binding (K_a) and anion transport (k_{ini} and $EC_{50,270s}$) properties of the twelve squaramide derivatives studied.

Squaramide	K_a (M^{-1})[a]	k_{ini} (s^{-1})[b]	$EC_{50,270s}$ (mol%)[b]
A1	240	0.1680	1.379
A2	433	2.1450	0.062
A3	602	1.4540	0.010
B1	119	0.0890	–
B2	729	0.3800	0.209
B3	776	0.8950	0.120
C1	871	0.9930	0.357
C2	4742	0.9800	0.085
C3	2259	0.6520	0.164
D1	821	0.0380	–
D2	8021	0.0021	–
D3	1627	0.0038	1.379

[a] Determined in DMSO-d_6/0.5% water at 298 K;[41, 53] [b] Determined in Cl^-/NO_3^- transport assays.[41, 53]

[§§] $EC_{50,270s}$ is the mol% transporter to lipid concentration of transporter necessary to promote 50% chloride efflux after 270 s.

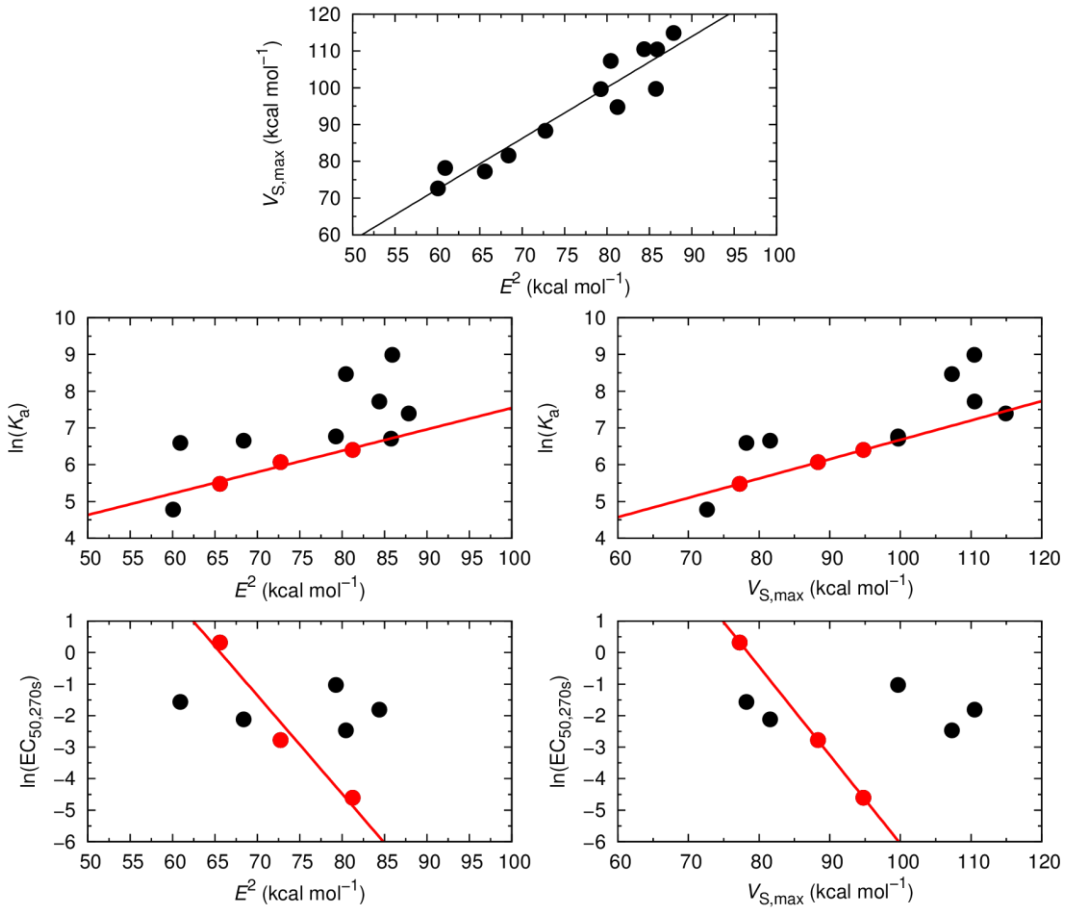


Figure IV-3. Top: $V_{S,max}$ values of the twelve squaramide derivatives plotted as a function of the E^2 energy values (R^2 = 0.90). Middle: ln(K_a) values plotted as function of the E^2 (left) or $V_{S,max}$ values (right). The red points stand for the **A** series squaramides, with the R^2 of 0.96 and 1.00, respectively. Bottom: ln($EC_{50,270s}$) values plotted as function of the E^2 (left) or $V_{S,max}$ values (right). The red points stand for the **A** series squaramides, with the R^2 of 0.96 and 1.00, respectively.

This comparison indicates that these quantum descriptors *per se* are insufficient for the accurate description of these squaramide derivatives as anion transporters. Thus, MD simulations were undertaken to gain further insights on the anion transport abilities of these squaramide-based molecules at the atomistic level.

IV.3.2. Passive diffusion MD simulations

The MD simulations of the squaramide complexes on membrane systems were preceded by a force field parameterisation to more accurately reproduce the structure of the squaramide core in the phospholipid bilayer environment. This parameterisation was based in crystallographic data and is reported in section C.1. in Appendix C.

The DFT optimised structures of their chloride complexes were either inserted into the water slab (scenario W) or at the core (scenario M) of a POPC bilayer composed of 128 phospholipids (*vide supra*). In addition, the anion carrier ability of these receptors was evaluated using distance restraints between the squaramide's binding sites and the chloride ion, leading to scenarios W' and M'. Scenarios W and W' allowed to assess of the passive diffusion of squaramide derivatives from the aqueous phase to the water/lipid interface, while the M and M' scenarios enabled the assessment of diffusion events

that the chloride complexes experience within the highly packed phospholipid bilayer. Furthermore, the W and M starting scenarios allowed the monitoring, at the atomistic level, of the eventual chloride release and uptake events, intrinsically associated with the anion carrier mechanism. These MD simulations also permitted the evaluation of the preferential orientation of the transporters relatively to the bilayer normal as well as the competitive hydrogen bonding interactions between the squaramide binding sites and the water molecules or phospholipids over the N–H···Cl^- bonding interactions.

Most systems were simulated in two independent MD runs of 200 ns each, asunder of systems with tripodal molecules (**D** series), which were simulated for 250 ns (W' scenario) or 500 ns (M scenario). Henceforth, these MD simulations are identified following a notation scheme derived from the initial position of the chloride complex in the membrane system, the number of the MD run and the squaramide molecule, as gathered in Table IV-1. For instance, W_1.**A1** stands for the first MD run carried out with **A1** in scenario W.

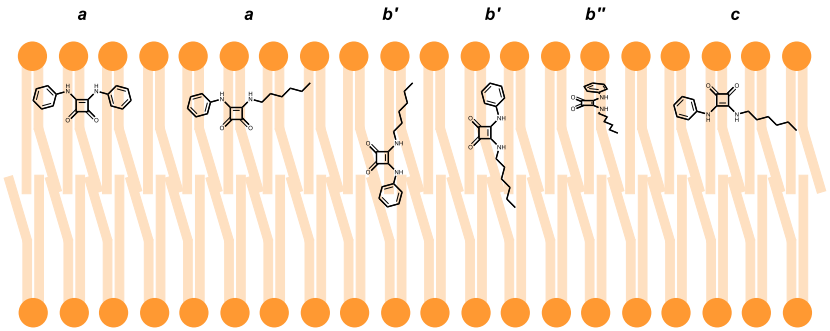
MD simulations with chloride complexes of linear squaramides A and B

In agreement with the chloride efflux studies,[41, 53] MD simulations of the chloride complexes of the **A** and **B** were carried out in scenario W for 200 ns. The position of each squaramide derivative in the bilayer system was assessed throughout the production runs using the relative distance from the COM defined by the two carbonyl groups (C=O_{COM}) to the closest water/lipid interface (P_{int}, defined by the average position of the phosphorus atoms in that monolayer) measured along the *z*-dimension (the membrane normal, C=O_{COM}···P_{int} distance). Moreover, this distance, together with the relative distance between the COM of the nitrogen atoms from the two binding units and P_{int} (N–H_{COM}···P_{int}) gives the relative positioning of these carbonyl groups and N–H binding sites towards the interface.

The evolution of the N–H_{COM}···P_{int} and C=O_{COM}···P_{int} distances along the entire production run of simulation W_1.**A2** is plotted in Figure IV-4, together with the number of hydrogen bonds that **A2** establishes with any of the nineteen chloride ions present in the system.*** The initially coordinated chloride and the N–H binding sites are quickly solvated by the water molecules and the anion is released to the aqueous phase, with the disruption of the N–H···Cl^- hydrogen bonds (see Figure IV-4, green line). Concomitantly, these hydrogen bonding interactions are replaced by N–H···OH_2, as evident in Figure IV-5 (blue line), with the plot of the count of these hydrogen bonds *vs.* the relative position of **A2** along the bilayer normal (*z*-dimension). Afterwards, as **A2** approaches the water/lipid interface, some of the N–H···OH_2 interactions are replaced by hydrogen bonds with the phosphate head groups (N–H···O_4P^-, orange line in Figure IV-5). The diffusion of **A2** from the water phase, illustrated in Figure IV-6 with six selected snapshots taken from MD run W_1.**A2**, shows the squaramide derivative entering the phospholipid bilayer with a hydrophobic $-CF_3$ group first. After 25 ns of simulation, the squaramide derivative is found completely embedded into the phospholipid bilayer. In Figure IV-4 the evolution of the N–H_{COM}···P_{int} distance (blue line) shows that the binding units are closer to the interface, while the C=O_{COM}···P_{int} relative distance (red line), for a systematically longer time than the former one, indicating that the polar carbonyl groups point towards the lipophilic bilayer core – henceforth, this spatial disposition is type ***a*** (see Scheme IV-2). In addition, both 4–CF_3Ph substituents adopt a roughly perpendicular disposition to the plane of the interface, enabling the rigid molecule **A2** to establish N–H···O hydrogen bonding interactions, mainly with water molecules, followed by interactions with carbonyl and phospholipid head groups (see Figure IV-6). Equivalent results were observed for the second run and for the MD simulations carried out with **A1** and **A3**. In the MD simulations of **A1** and **A3**,

*** The hydrogen bonding interactions were assessed with cut-offs of 3.5 Å and 120° in agreement with the N–H···Cl^- dimensions found in single crystal X-ray structures deposited with CCDC[70, 71] of squaramide-based chloride complexes.[32, 41, 83, 84]

the squaramides' permeation is accompanied by a decrease of the N–H···OH_2 hydrogen bonds and concomitant increase of the N–H···O_4P^- interactions at the interface level (see Figure IV-5).



Scheme IV-2. Possible spatial orientations for squaramide series **A** and **B** (illustrated with **A1** and **B1**) observed in the MD simulations in membrane models.

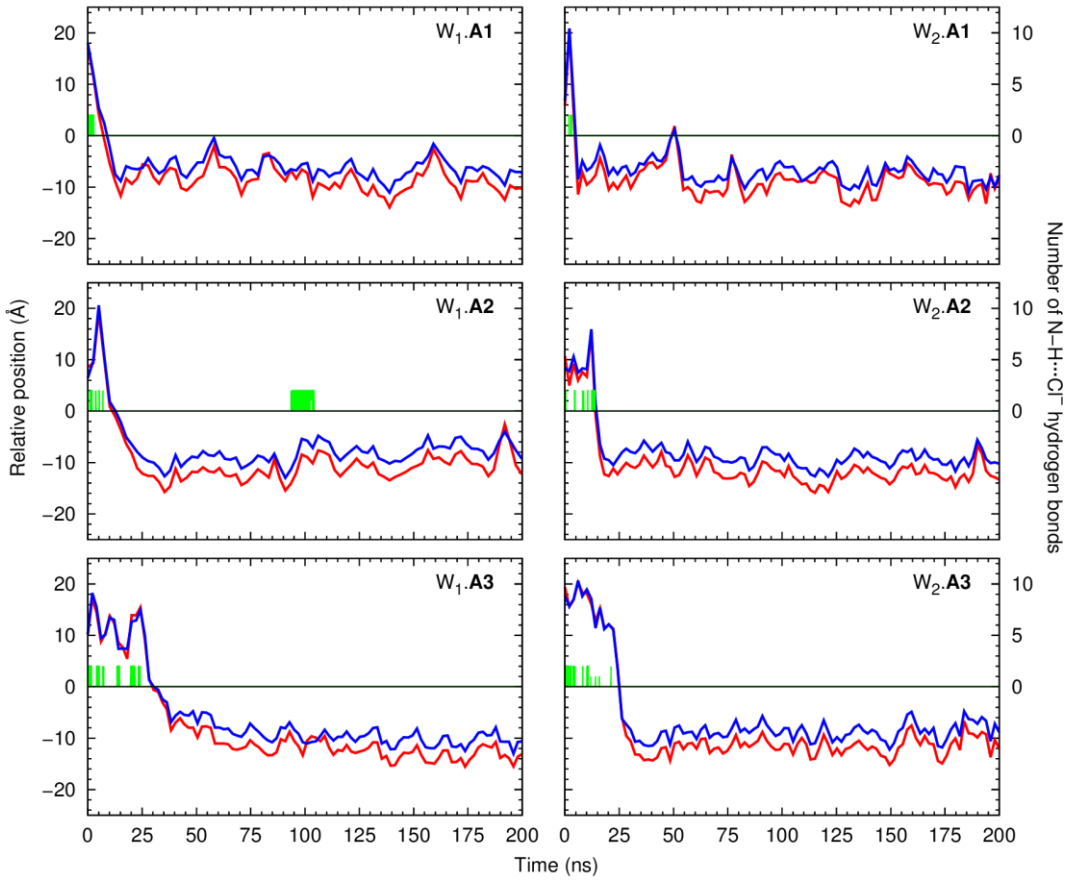


Figure IV-4. Evolution of the C=O_{COM}···P_{int} (red line) and N–H_{COM}···P_{int} (blue line) distances throughout 200 ns for the MD runs of the **A** series chloride complexes in scenario W. The evolution of the number of N–H···Cl^- hydrogen bonds is also plotted as a green line. The water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

Overall, the evolution of the N–H···Cl^- interactions (green lines in Figure IV-4) shows that the **A** series squaramides can recognise chloride throughout the passive diffusion in the bilayer system. Indeed, it is noteworthy that **A2**, in MD simulation W_1.**A2**, uptakes a solvated chloride ion, among the nineteen present in the water phase, for 10 ns (between the 93rd and 104th ns), while embedded in the

membrane. This feature can be perceived by the increase of N–H···Cl^- bonding interactions in Figure IV-4, for the aforementioned period, as well as in the 4th to 6th panels of Figure IV-6. However, in the second MD run of **A2** in scenario W, and in both MD runs of **A1** and **A3** (see Figure IV-4), no chloride recognition events are observed after the internalisation of the small molecules. Noteworthy, in all these MD simulations, sporadic chloride recognition events in the water phase are monitored prior to the permeation of the water/lipid interface. The chloride uptake and release observed events are remarkable and show that these molecules have the potential to function as anion carriers. Moreover, the MD simulations suggest that the anion binding can either occur in the water phase or in the interface (*vide infra*).

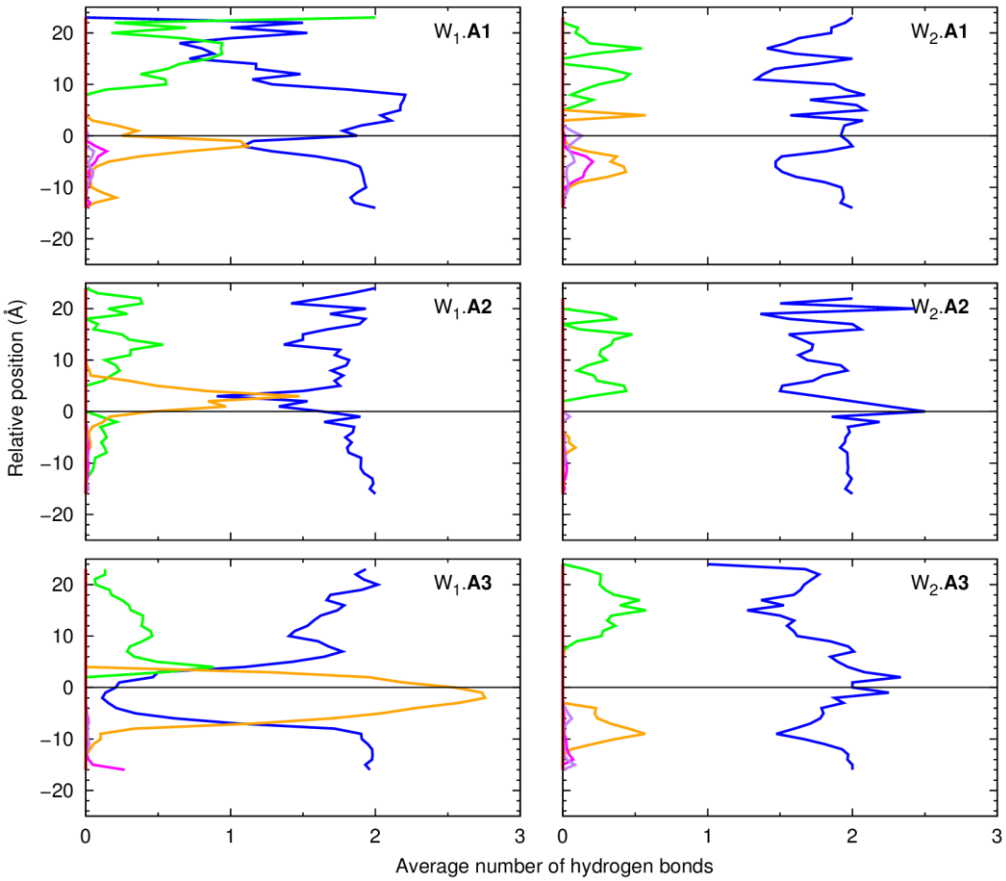


Figure IV-5. Average number of hydrogen bonds *vs*. the relative position of the COM of the **A** series molecules in the MD runs carried out in scenario W. The following colour scheme was used for the interactions between the squaramide derivative and chloride ions (green), water molecules (blue), POPC head groups (orange), and ester groups (magenta for the *sn*-1 chains and purple for the *sn*-2 chains). The water/lipid interface is represented as a black line at z = 0 Å. Data were smoothed using Bézier curves.

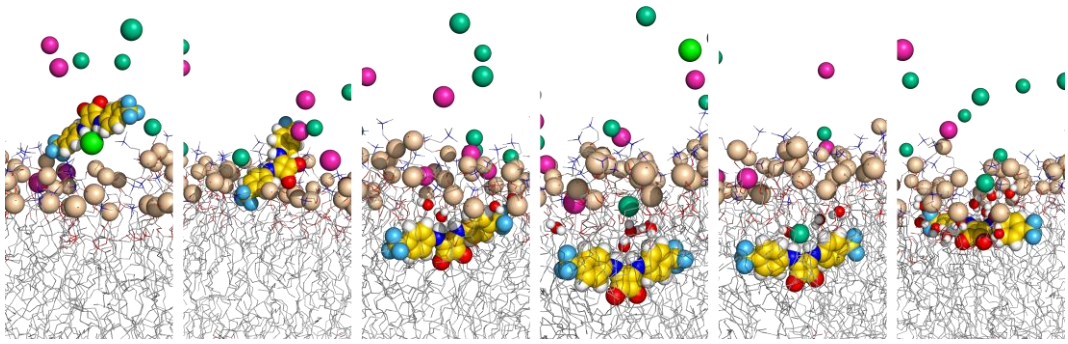


Figure IV-6. Consecutive snapshots depicting the diffusion of the chloride complex of **A2** throughout MD simulation W_1.**A2**. The squaramide derivative, sodium and chloride ions together with the phosphorus atoms are represented in spheres. The remaining atoms of the bilayer are shown as lines. The hydrogen atoms are shown in white, oxygen atoms in red, nitrogen atoms in blue, fluorine atoms in light blue, phosphorus atoms in wheat and carbon atoms in yellow (squaramide derivative) or grey (phospholipids). The sodium ions are shown in pink and the chloride ions are shown in green (complex) or green marine (counterions). The interaction of the internalised transporter with the solvent molecules is emphasised with the depiction of water molecules within 3.5 Å as spheres. The lipids' C–H bonds and most water molecules were omitted for clarity.

In stark contrast with the **A** series molecules, the **B** series squaramides have an asymmetric structure, with the squaramide core decorated with an *n*-hexyl chain and a phenyl substituent accounting for different degrees of fluorination (*vide* Scheme IV-1) and lipophilicities. These two structural features dictate that additional dispositions to type ***a*** are adopted by these molecules in the phospholipid packed environment, as follows.

In the MD simulations of the asymmetric squaramides **B**, in addition to the $C{=}O_{COM}\cdots P_{int}$ and $N{-}H_{COM}\cdots P_{int}$ distances, the distance between the COM of the aromatic ring (Ph_{COM}) and the closest P_{int} ($Ph_{COM}\cdots P_{int}$ distance), was evaluated to ascertain the relative orientation of **B1**, **B2** or **B3** towards the water/lipid interface. The evolution of these distances throughout the two independent MD runs is plotted in Figure IV-7 for **B1**, **B2** and **B3**. In the MD runs of the **B** series in scenario W, after the interface permeation, the squaramide core disposition of the free squaramide derivatives interchanges between type ***a*** and other two spatial dispositions, characterised by similar $C{=}O_{COM}\cdots P_{int}$ and $N{-}H_{COM}\cdots P_{int}$ distances (type ***b***) or a shorter $C{=}O_{COM}\cdots P_{int}$ distance than the $N{-}H_{COM}\cdots P_{int}$ distance (type ***c***). Furthermore, the MD simulations have shown that type ***b*** disposition has to be subdivided in function of the $Ph_{COM}\cdots P_{int}$ distance: in type ***b'*** the squaramide core can be roughly perpendicular to the plane of the interface, while in type ***b''*** the squaramide core is almost parallel to the interface with three referential points approximately at the same distance to the interface. Furthermore, in type ***b'*** the Ph_{COM} is either above or below the $C{=}O_{COM}$ and $N{-}H_{COM}$ referential points. These four relevant spatial dispositions of the **B** molecules are also sketched in Scheme IV-2.

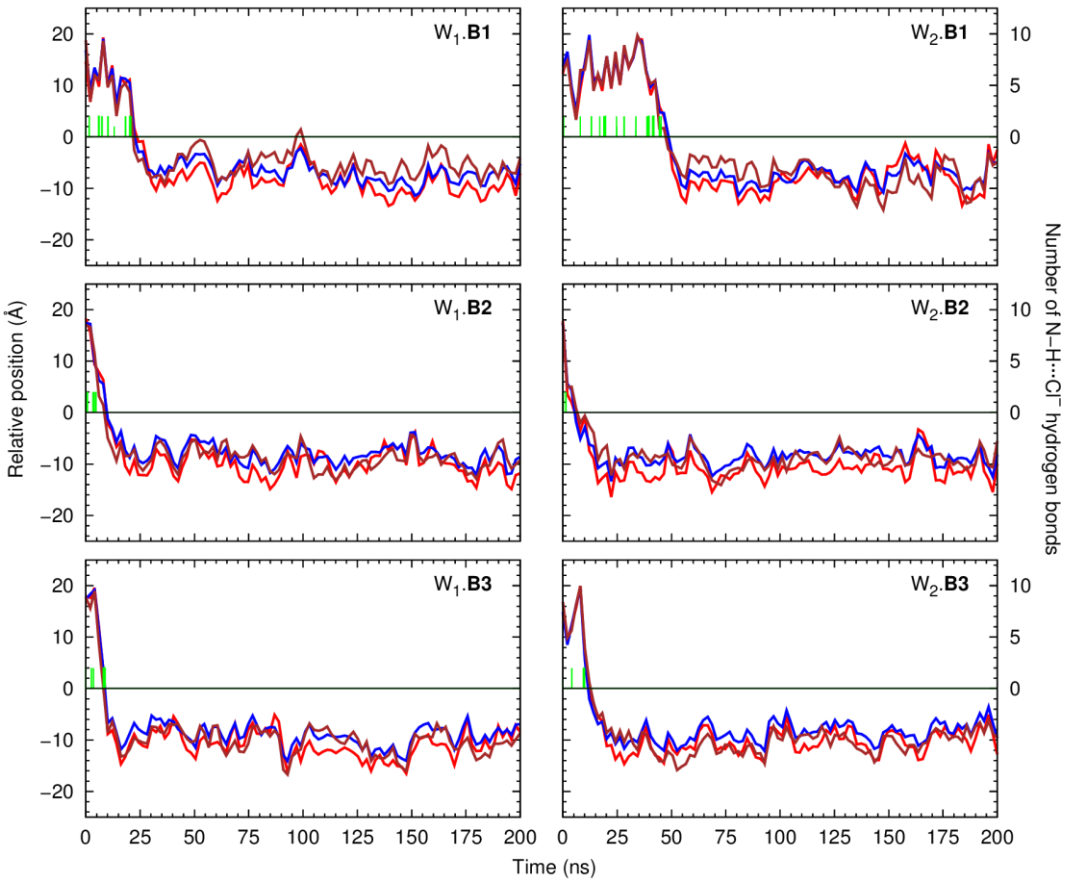


Figure IV-7. Evolution of the C=O_{COM}···P_{int} (red line), N–H_{COM}···P_{int} (blue line) and Ph_{COM}···P_{int} (brown line) distances throughout 200 ns for the MD runs of the **B** series chloride complexes in scenario W. Remaining details as given in Figure IV-4.

For instance, in simulation W_1.**B1**, after the permeation, the receptor oscillates mainly between dispositions ***a*** and ***b'***, with brief visits of ***b''***. Noteworthy, when the molecule adopts the ***b'*** disposition, the aromatic group is found closer to the water/lipid interface, while the *n*-hexyl chain is almost vertically aligned with the surrounding phospholipids (see Figure IV-8). In the second run of this system, for *ca.* 15 ns (between the 130th and the 145th ns), **B1** adopts a ***b'*** disposition, but with the phenyl group pointing to the bilayer core. In this MD run, **B1** also adopts a type ***c*** disposition between the 140th and the 170th ns. In the MD simulations with **B2**, decorated with a 4–CF_3Ph, the most usual disposition is type ***a***, followed by both ***b*** subtypes. For the ***b'*** disposition, the fluorinated substituent is frequently found closer to the bilayer core. As observed for the system with **B2**, in simulations W_1.**B3** and W_2.**B3**, the most common spatial dispositions are ***a*** and ***b'***, with the 3,5–$(CF_3)_2$Ph substituent preferentially pointing to the bilayer core in ***b'***. This is particularly evident in the second MD run (see Figure IV-7).

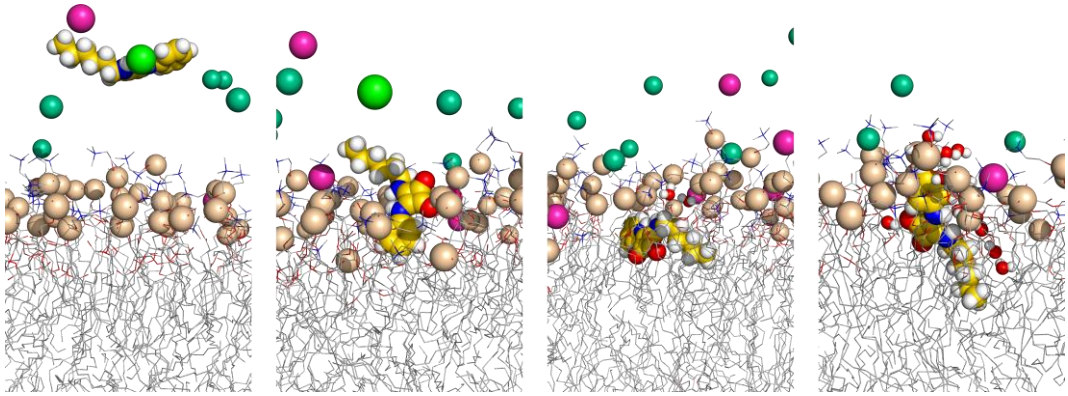


Figure IV-8. Consecutive snapshots depicting the diffusion of the chloride complex of **B1** throughout MD simulation W_1.**B1**. Remaining details as given in Figure IV-6.

The analysis of the spatial dispositions can be related with the lipophilic character of the substituents present along these two series. Indeed, for the symmetric **A** series, with equally lipophilic aromatic substituents, the main spatial disposition adopted is type ***a***. On the other hand, in the asymmetric **B** series, the different lipophilic nature of the substituents leads to a frequent exchange between spatial dispositions with the prevalence of type ***b***. With the increase of the lipophilic character of the aromatic substituents (log*P* values: Ph = 2.05 < 4–CF_3Ph = 2.93 < 3,5–$(CF_3)_2$Ph = 3.81),[87] the *n*-hexyl substituent (log*P* = 2.88)[87] is closer to the water/lipid interface, rather than being aligned with the phospholipid tails. Throughout their diffusion along the membrane system, likewise the **A** series, the squaramide core of the **B** molecules interacts with the POPC lipids and water molecules *via* hydrogen bonds, as can be seen in Figure IV-9.

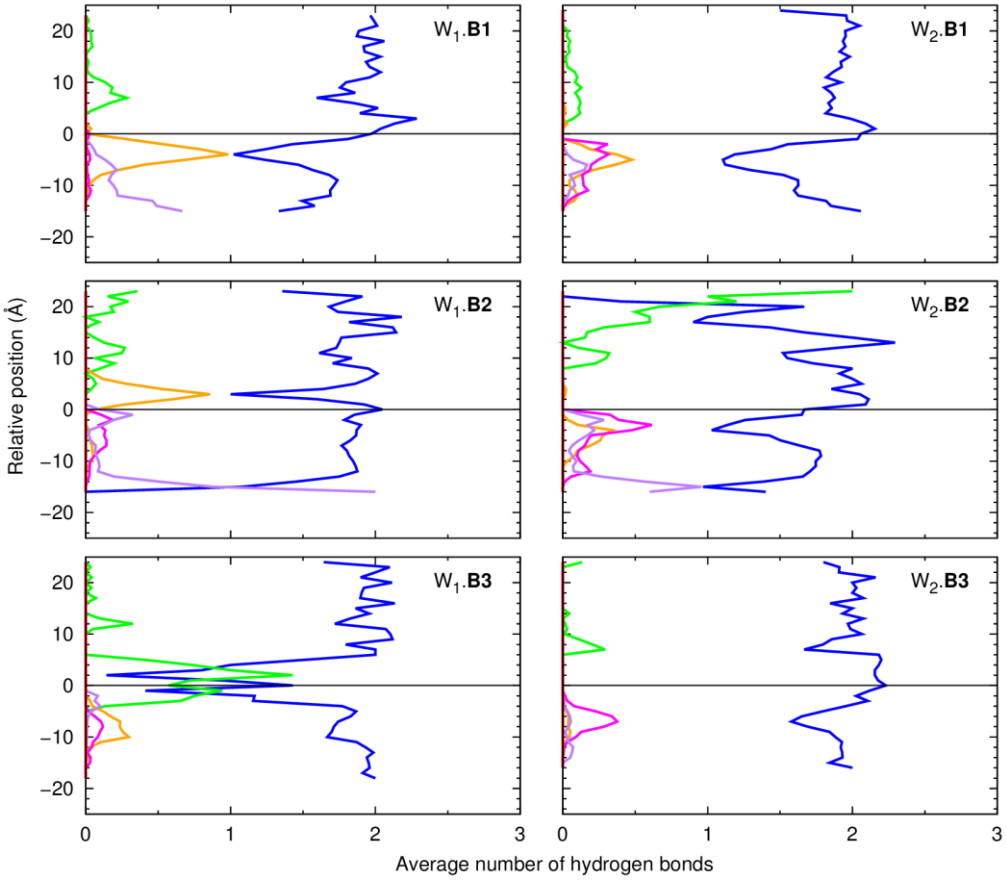


Figure IV-9. Average number of hydrogen bonds *vs.* the relative position of the COM of the **B** series molecules in the MD runs carried out in scenario W. Remaining details as given in Figure IV-5.

Having observed that the free **A** and **B** squaramides are able to permeate the membrane, the passive diffusion of their chloride complexes was investigated for 200 ns using two N···Cl^- distances restraints (scenario W'). As shown in Figure IV-10, all complexes are able to permeate the water/lipid interface, staying below it until the end of simulation time, with the chloride ion widely exposed to a variable number of water molecules ranging between 0 and 9, as summarised in Table IV-5 and Table C-3. On the other hand, the carbonyl moieties of the squaramide motifs, found deeper in the POPC bilayer, have a negligible solvation.

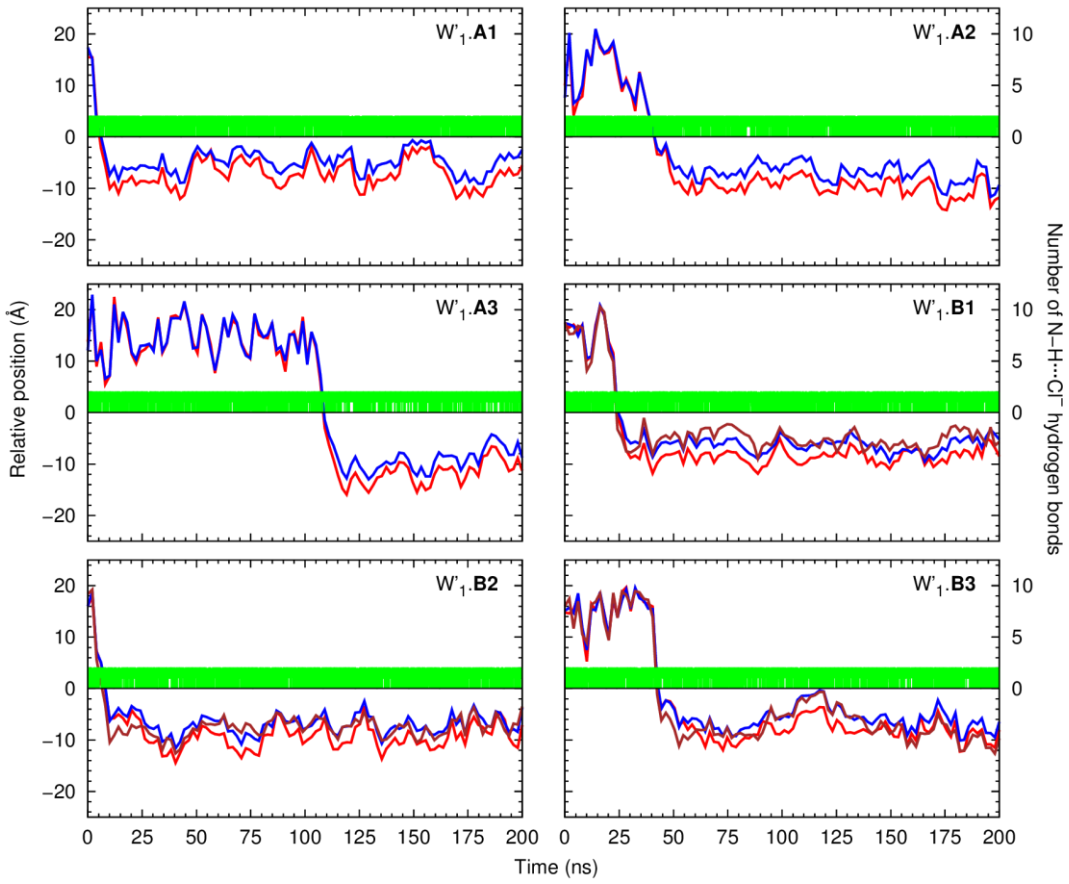


Figure IV-10. Evolution of the $C{=}O_{COM}\cdots P_{int}$ (red line), $N{-}H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (brown line) distances throughout 200 ns for the MD runs of the **A** and **B** series chloride complexes in scenario W'. Remaining details as given in Figure IV-4.

Table IV-5. Average number of water molecules solvating the squaramide chloride complexes within the first solvation shell (cut-off of 3.4 Å[a]).[b]

MD simulation	Solvated entity		
	Squaramide derivative	Carbonyl groups[c]	Complexed chloride
W'$_1$.**A1**[d]	9.2 ± 3.6	0.4 ± 0.8	3.7 ± 1.0
W'$_1$.**A2**[d]	7.2 ± 2.5	0.2 ± 0.5	3.9 ± 0.9
W'$_1$.**A3**[d]	9.7 ± 10.4	0.6 ± 1.9	3.7 ± 1.1
W'$_1$.**B1**[d]	8.3 ± 3.0	0.5 ± 0.8	3.5 ± 1.2
W'$_1$.**B2**[d]	8.1 ± 3.1	0.3 ± 0.7	4.1 ± 1.1
W'$_1$.**B3**[d]	10.1 ± 3.6	0.4 ± 0.7	3.8 ± 0.9
M'$_1$.**A1**[d]	8.8 ± 3.0	0.4 ± 0.8	3.8 ± 1.0
M'$_1$.**A2**[d]	6.6 ± 2.6	0.1 ± 0.5	3.5 ± 1.3
M'$_1$.**A3**[d]	6.0 ± 2.7	0.0 ± 0.3	3.3 ± 1.0
M'$_1$.**B1**[d]	10.0 ± 3.2	0.5 ± 0.8	4.2 ± 1.0
M'$_1$.**B2**[d]	9.3 ± 2.7	0.6 ± 0.8	4.1 ± 0.9
M'$_1$.**B3**[d]	8.9 ± 3.0	0.4 ± 0.6	3.9 ± 1.0
M$_{all}$.**C1**[e]	10.6 ± 3.2	1.5 ± 1.5	2.6 ± 0.9
M$_{all}$.**C2**[e]	9.8 ± 3.2	2.2 ± 1.8	2.4 ± 0.9
M$_{all}$.**C3**[e]	9.3 ± 3.2	2.1 ± 1.7	2.1 ± 0.8
M$_{all}$.**D1**[f]	9.5 ± 3.2	4.5 ± 2.3	0.1 ± 0.3
M$_{all}$.**D2**[f]	9.5 ± 2.9	4.2 ± 2.1	0.0 ± 0.2
M$_{all}$.**D3**[f]	9.1 ± 3.4	4.2 ± 2.3	0.1 ± 0.3

[a] Approximate radius for the first hydration shell of biomolecules;[88] [b] These parameters were evaluated throughout the last 100 ns of simulation time for the **A**, **B** and **C** series, and throughout the last 400 ns for the **D** series; [c] Sum of the number of water molecules around the independent C=O groups; [d] N = 10000; [e] The data from the two individual runs was concatenated and treated together, resulting in N = 20000; [f] The data from the four individual runs (vertical and horizontal initial orientations) was concatenated and treated together, resulting in N = 160000. The complete statistics for the individual MD runs can be found in Table C-3 (see Appendix C).

The ability of these two linear squaramide series to assist the chloride transport by passive diffusion across the phospholipid bilayer was theoretically investigated with the anion complexes inserted into the bilayer core with a random spatial disposition. The evolution of the C=O$_{COM}$···P$_{int}$ and N−H$_{COM}$···P$_{int}$ relative distances, plotted in Figure IV-11 for squaramides **A**, shows that during the first ns of MD production runs, the complexes of this series reorient within the packed phospholipid environment adopting mainly a type ***a*** disposition (see Scheme IV-2). Subsequently, the chloride ion is released before the complex reaches the water/lipid interface, as indicated by the interruption of the N−H···Cl$^-$ interactions, also shown in Figure IV-11 (green line). Indeed, in the first ns of MD sampling, few water molecules permeate the interface and diffuse towards the anion, solvating it, and then promoting its release from the squaramide receptor to the water phase. Afterwards, the free **A** squaramides nest between the phospholipids preserving the ***a*** type disposition, as depicted for the MD simulations performed in scenario W. This sequence of diffusion events is illustrated in Figure IV-12 with selected snapshots obtained from simulation M$_1$.**A3**. It is also noteworthy that in this MD run, a chloride ion is uptaken by embedded **A3**, near the end of the MD simulation for a period of *ca.* 20 ns (see Figure IV-11 and Figure IV-12). This is in line with the uptake event reported for simulation W$_1$.**A2** (*vide supra*), emphasising the hypothesis that the anion binding can occur at the interface, rather than in the water phase.

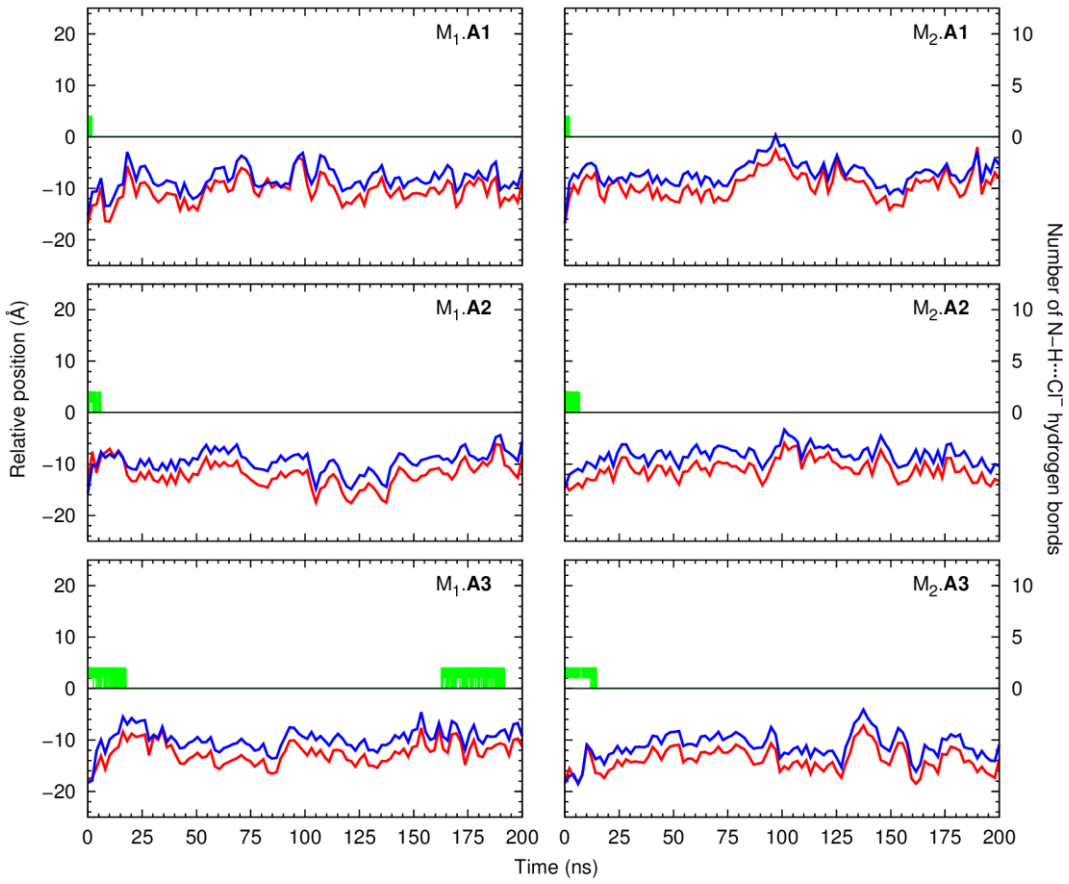


Figure IV-11. Evolution of the C=O_{COM}···P_{int} (red line) and N–H_{COM}···P_{int} (blue line) distances throughout 200 ns for the MD runs of the **A** series chloride complexes in scenario M. Remaining details as given in Figure IV-4.

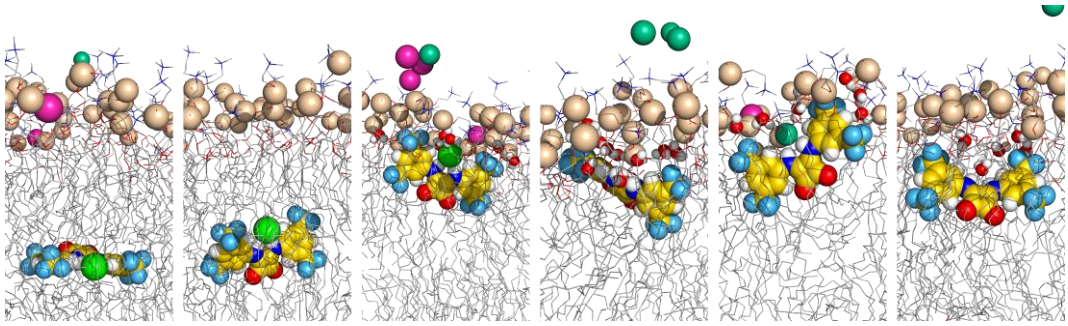


Figure IV-12. Consecutive snapshots depicting the diffusion of the chloride complex of **A3** throughout MD simulation M_1.**A3**. Remaining details as given in Figure IV-6.

Likewise the **A** squaramide complexes, in the MD runs carried with complexes of **B** series squaramides in scenario M the chloride ion leaves the squaramide derivative assisted by water molecules towards the aqueous phase. Furthermore, the spatial disposition adopted by the **B** molecules exchanges between types ***a***, ***b*** and ***c*** (see Figure IV-13), corroborating the insights reported for the MD simulations where their complexes are initially positioned in the water phase (scenario W).

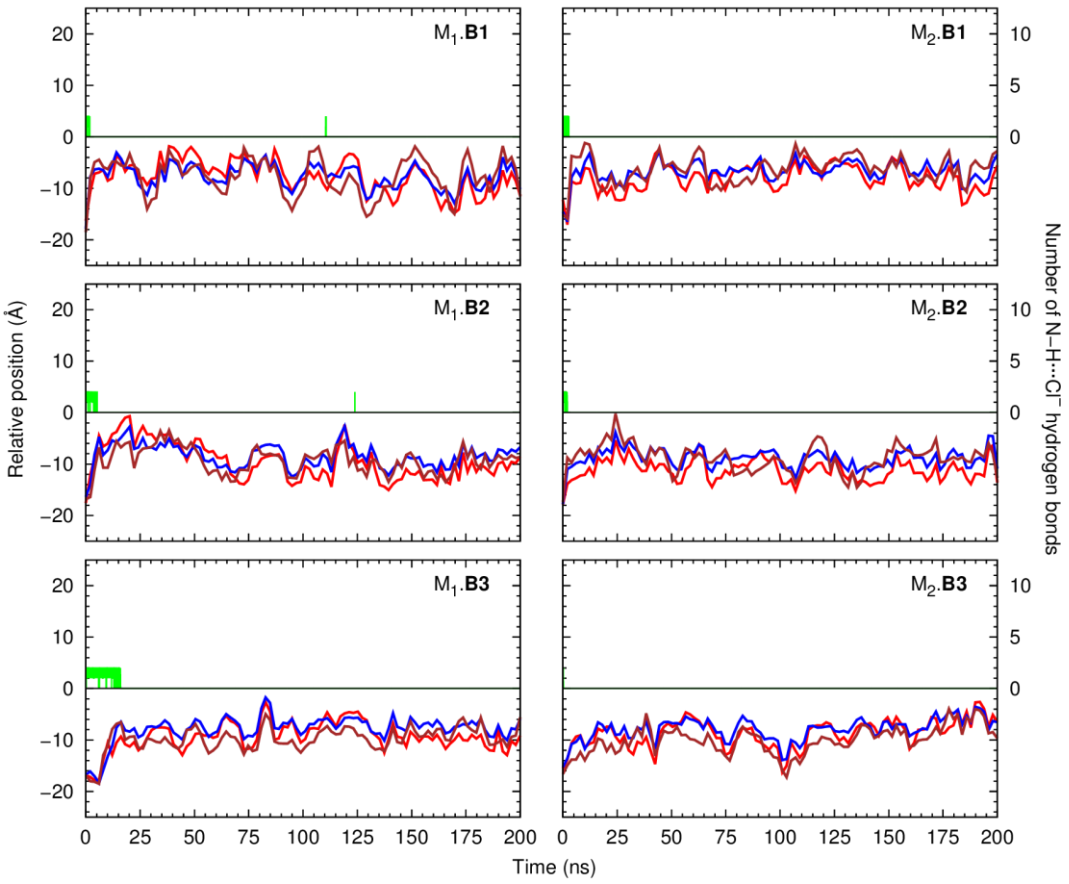


Figure IV-13. Evolution of the C=O_{COM}···P_{int} (red line), N–H_{COM}···P_{int} (blue line) and Ph_{COM}···P_{int} (brown line) distances throughout 200 ns for the MD runs of the **B** series chloride complexes in scenario M. Remaining details as given in Figure IV-4.

In the simulations carried out in the M' scenario, the complexes promptly migrate to the water/lipid interface, with the N–H binding units close to the interface (see Figure IV-14). Consequently, the anion is exposed to the water molecules that would facilitate its release to the aqueous phase, while the C=O groups of the **A** and **B** squaramide derivatives have a lower solvation, as summarised in Table IV-5.

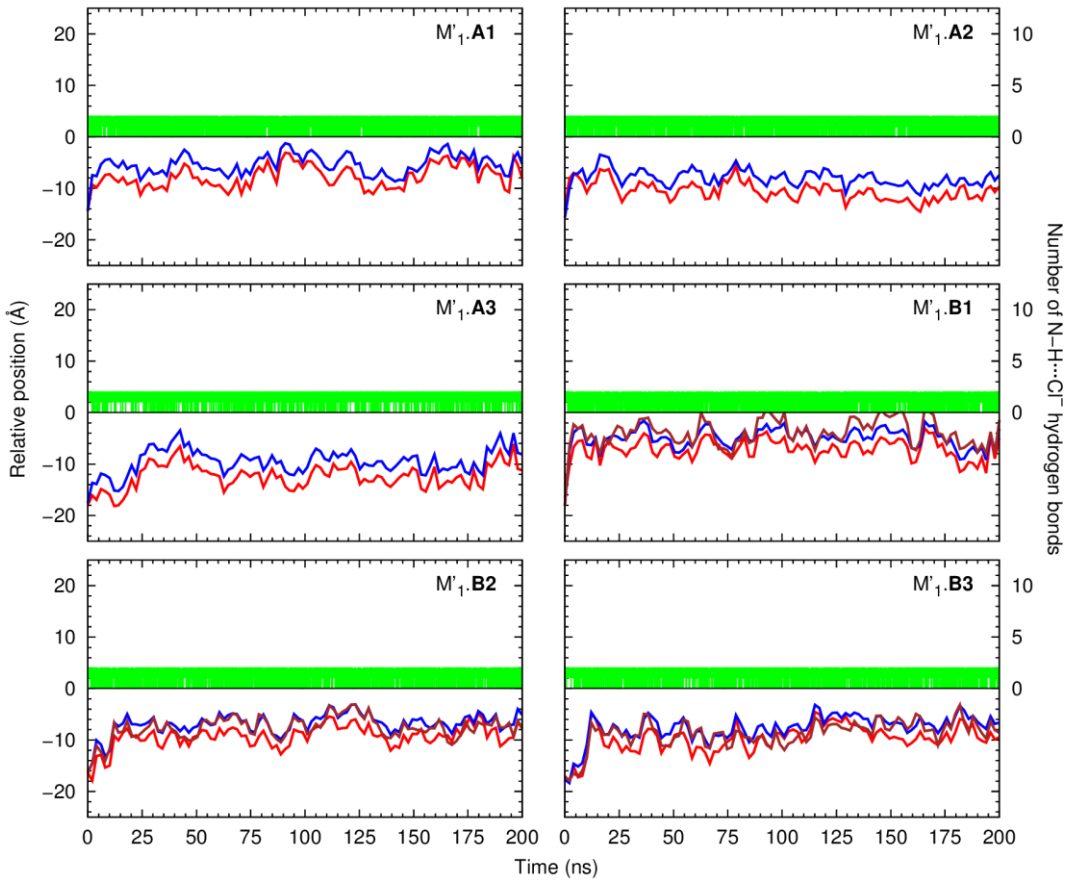


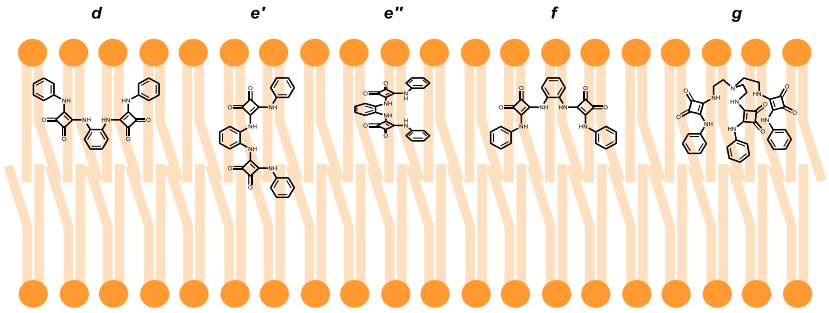
Figure IV-14. Evolution of the C=O_{COM}···P_{int} (red line), N–H_{COM}···P_{int} (blue line) and Ph_{COM}···P_{int} (brown line) distances throughout 200 ns for the MD runs of the **A** and **B** series chloride complexes in scenario M'. Remaining details as given in Figure IV-4.

MD simulations with chloride complexes of squaramides C

Given the two-dimensional scaffold of the **C** series squaramides, different referential points were necessary to assess the passive diffusion of their complexes in the membrane model. Therefore, N–H_{COM} stands for the COM of the four nitrogen atoms of the binding units, Sp_{COM} is the COM of the aromatic ring in the 1,2-phenylenediamine spacer, and Ph_{COM} corresponds to the COM defined by the two aromatic substituents.

The evolution of the relative distances between the N–H_{COM}, Sp_{COM} and Ph_{COM} and the closer water/lipid interface (P_{int}) for simulation W_1.**C1** is plotted in Figure IV-15, along with the number of hydrogen bonds between **C1** with any chloride present in the system. After the permeation, **C1** adopts a spatial disposition of type ***d*** (see Scheme IV-3), characterised by the aromatic substituents closer to the water phase and the 1,2-phenylenediamine spacer pointing to the bilayer core, for a period of *ca.* 50 ns. Subsequently, for a longer period, the squaramide adopts a roughly parallel disposition to the interface plane, type ***e'***, with the overlap of the three lines corresponding to the N–H_{COM}···P_{int} (blue), Sp_{COM}···P_{int} (orange) and Ph_{COM}···P_{int} (pink) distances. In the second run of **C1** in scenario W, after the internalisation of the squaramide derivative, two additional spatial dispositions arose: type ***e''***, also characterised with the overlap of the three lines corresponding to the three relative distances (as shown in Figure IV-15), but with the molecule adopting a perpendicular disposition towards the interface; and type ***f***, with the aromatic substituents pointing to the bilayer core and the 1,2-phenylenediamine spacer near the

interface. All these spatial dispositions are depicted in Figure IV-16, with snapshots taken from both MD runs. Noteworthy, in MD simulation W_1.**C1**, the N–H···Cl^- hydrogen bonds are interrupted twice, with the subsequent uptake of an anion, definitively restoring the chloride complex prior to the internalisation (as also seen in Figure IV-17), while in W_2.**C1** the chloride complex is maintained throughout the 200 ns simulation time.



Scheme IV-3. Possible spatial orientations for squaramide series **C** and **D** (illustrated with **C1** and **D1**) observed in the MD simulations in membrane models.

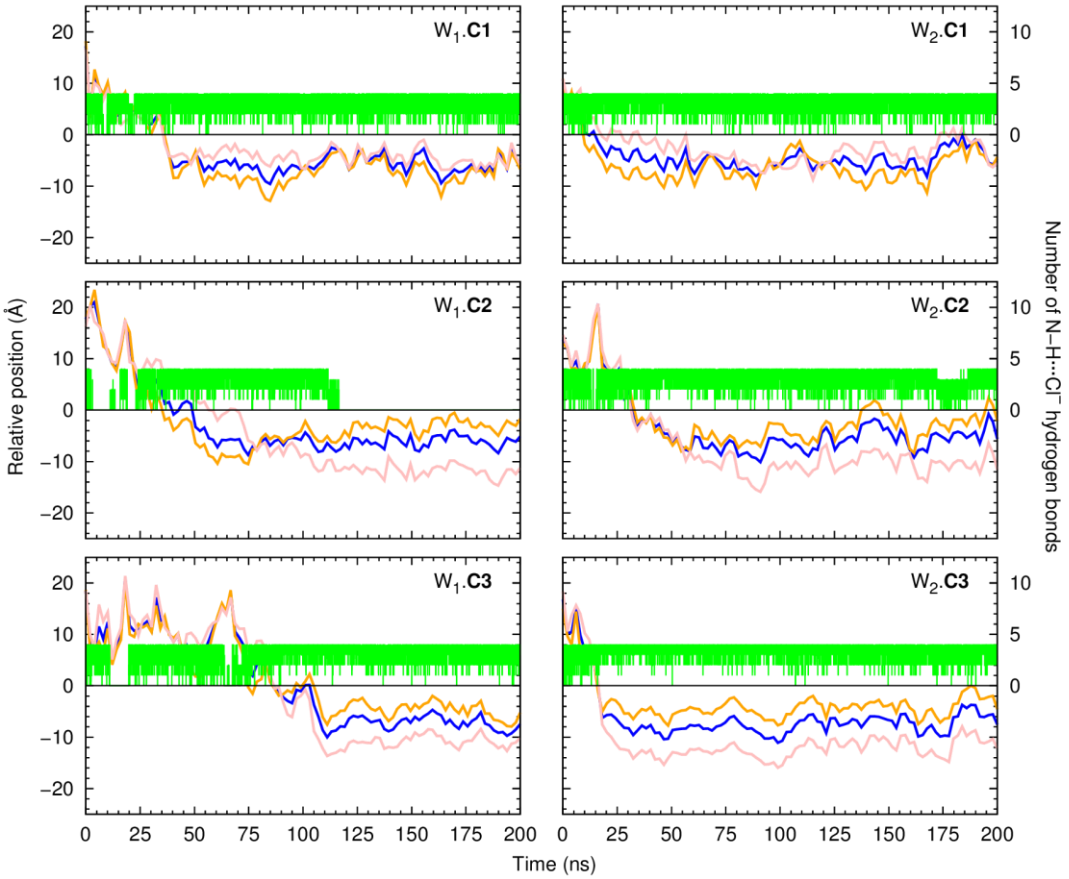


Figure IV-15. Evolution of the $Sp_{COM}\cdots P_{int}$ (orange line), $N–H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (pink line) distances throughout 200 ns for the MD runs of the **C** series chloride complexes in scenario W. Remaining details as given in Figure IV-4.

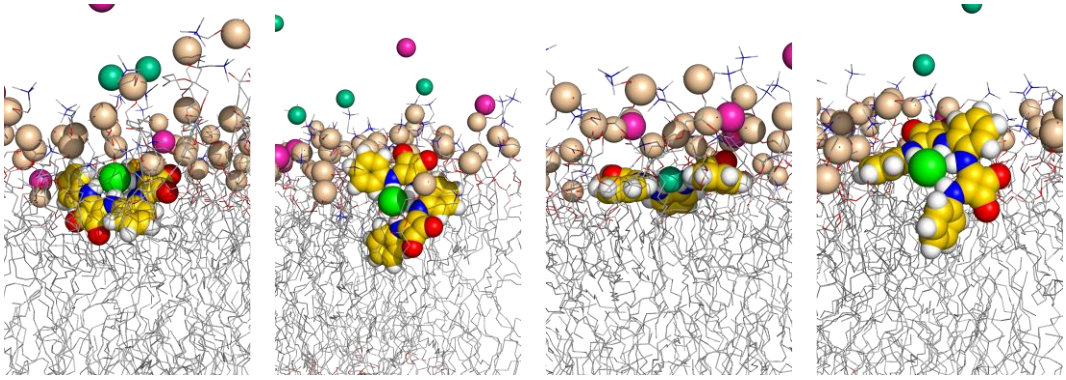


Figure IV-16. Illustrative snapshots depicting the orientations of the chloride complex of **C1** throughout MD simulations in scenario W. Remaining details as given in Figure IV-6.

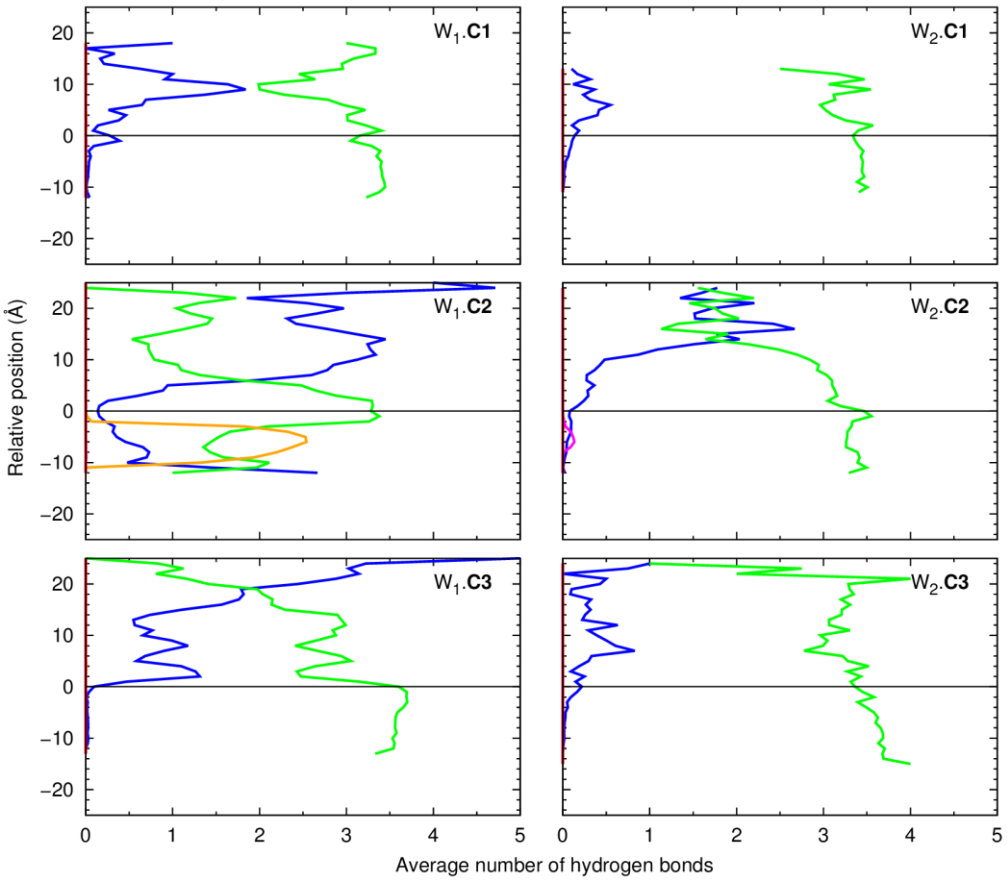


Figure IV-17. Average number of hydrogen bonds *vs*. the relative position of the COM of the **C** series molecules in the MD runs carried out in scenario W. Remaining details as given in Figure IV-5.

For molecules **C2** and **C3**, with two 4−CF_3Ph or two 3,5−$(CF_3)_2$Ph substituents, respectively, the evolution of the N−$H_{COM}\cdots P_{int}$, $Sp_{COM}\cdots P_{int}$ and $Ph_{COM}\cdots P_{int}$ distances (see Figure IV-15) indicates that, after internalising and nesting between the phospholipids, these molecules adopt a type ***f*** spatial disposition with the hydrophobic substituents deeply embed into the bilayer. Apart of MD simulation W_1.**C2**, the chloride complexes are maintained throughout most of the simulation time. In W_1.**C2** the anion definitively leaves **C2** to the water phase *ca.* 70 ns after the internalisation of the complex (see

Figure IV-15), with the replacement of the N–H···Cl^- interactions by strong hydrogen bonds to a phosphate head group and water molecules (see Figure IV-17).

The diffusion of the chloride complexes of the **C** series within the phospholipid bilayer was also investigated (scenario M), initially positioning the three anion complexes in the middle of the phospholipid molecules' tails. Moreover, due to the bulk of these molecules, the complexes were lodged almost parallel to the water/lipid interface plane (type ***e'*** disposition). In the beginning of the MD simulations, all complexes promptly diffuse towards the interface, as seen in Figure IV-18, and concomitantly reorient themselves, acquiring the same types of spatial dispositions as observed in MD simulations from the water phase (scenario W, *vide supra*). The final spatial dispositions of **C2** and **C3** anion complexes are unequivocally type ***f***, whereas non-fluorinated **C1** interchange its disposition between types ***d***, ***e'*** and ***f***. Overall, the complexes are highly hydrated when they reach the interface (see Table IV-5, with the solvation of squaramide derivative, its carbonyl groups and complexed chloride), but the anion is firmly bonded by the four cooperative N–H···Cl^- interactions, which preclude the anion release to the water phase (see Figure IV-17). Thus, the MD simulations with the **C** series in scenarios W' and M' were not undertaken. Moreover, these structural insights, obtained at the atomistic level, are entirely consistent with the moderate transport ability experimentally observed for this series of molecules.

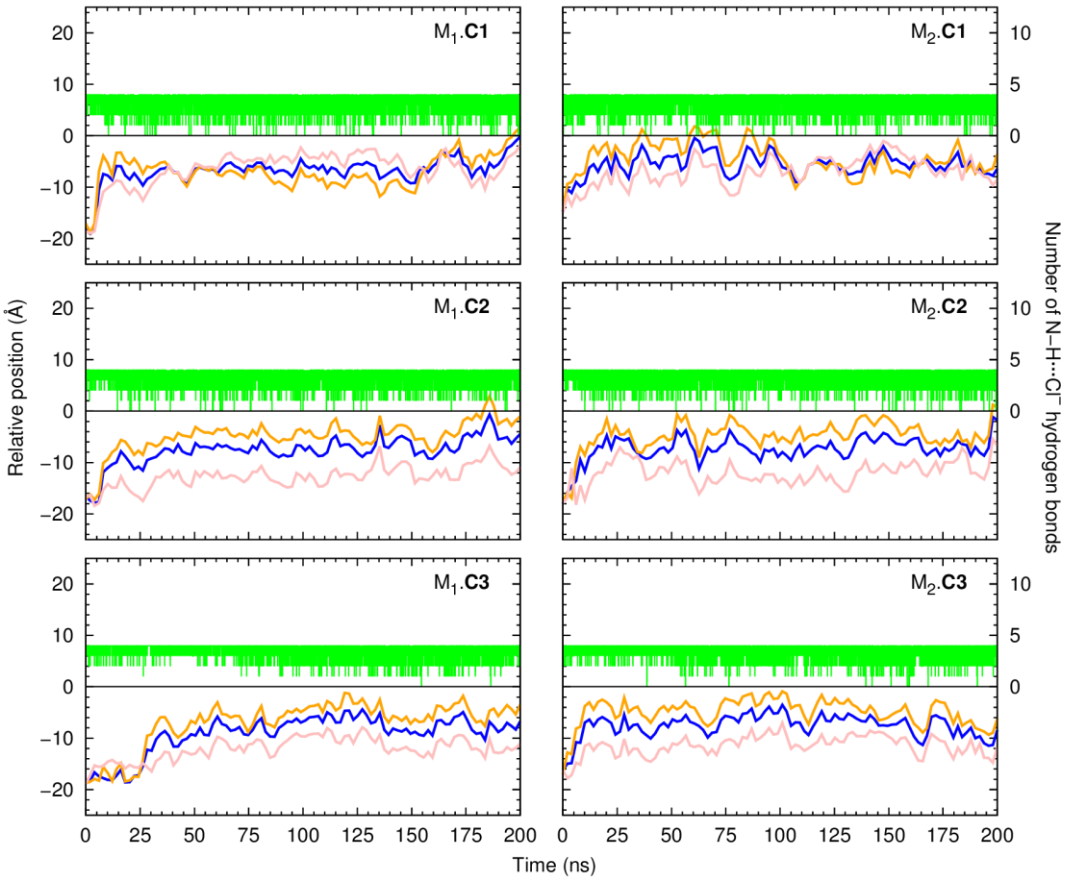


Figure IV-18. Evolution of the $Sp_{COM}\cdots P_{int}$ (orange line), $N–H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (pink line) distances throughout 200 ns for the MD runs of the **C** series chloride complexes in scenario M. Remaining details as given in Figure IV-4.

MD simulations with chloride complexes of tripodal squaramides D

Throughout the MD simulations performed with the **D** series in scenario W, the position of the tripodal squaramide-based molecules was monitored using the relative distance between the central tripodal nitrogen (N_{tren}) and P_{int} ($N_{tren}\cdots P_{int}$ distance), measured along the membrane normal, as described above. The orientation of the **D** series squaramides was followed using that relative distance together with the $N{-}H_{COM}\cdots P_{int}$ and $Ph_{COM}\cdots P_{int}$ distances, in which Ph_{COM} is now the COM defined by the aromatic carbon atoms from the three substituents and $N{-}H_{COM}$ stands for the COM defined by the nitrogen atoms of the N–H six binding units.

The three tripodal molecules exhibit equivalent diffusion behaviour, being able to permeate the water/lipid interface and residing below it in a well-defined spatial disposition until the end of the simulation time (*vide infra*). This spatial disposition of type ***g*** (see Scheme IV-3) is characterised by the N_{tren} and N–H binding units near of the interface level (orange and blue lines) and the aromatic substituents between the lipophilic tails (magenta line), as illustrated in Figure IV-19, with the evolution of the three relative distances for the MD simulations of the **D** series molecules in the W scenario, plotted along the number of hydrogen bonds to chloride (green line). In addition, the passive diffusion of this tripodal complex is depicted in Figure IV-20 with six snapshots taken from the first MD simulation of **D3**·Cl^- in scenario W.

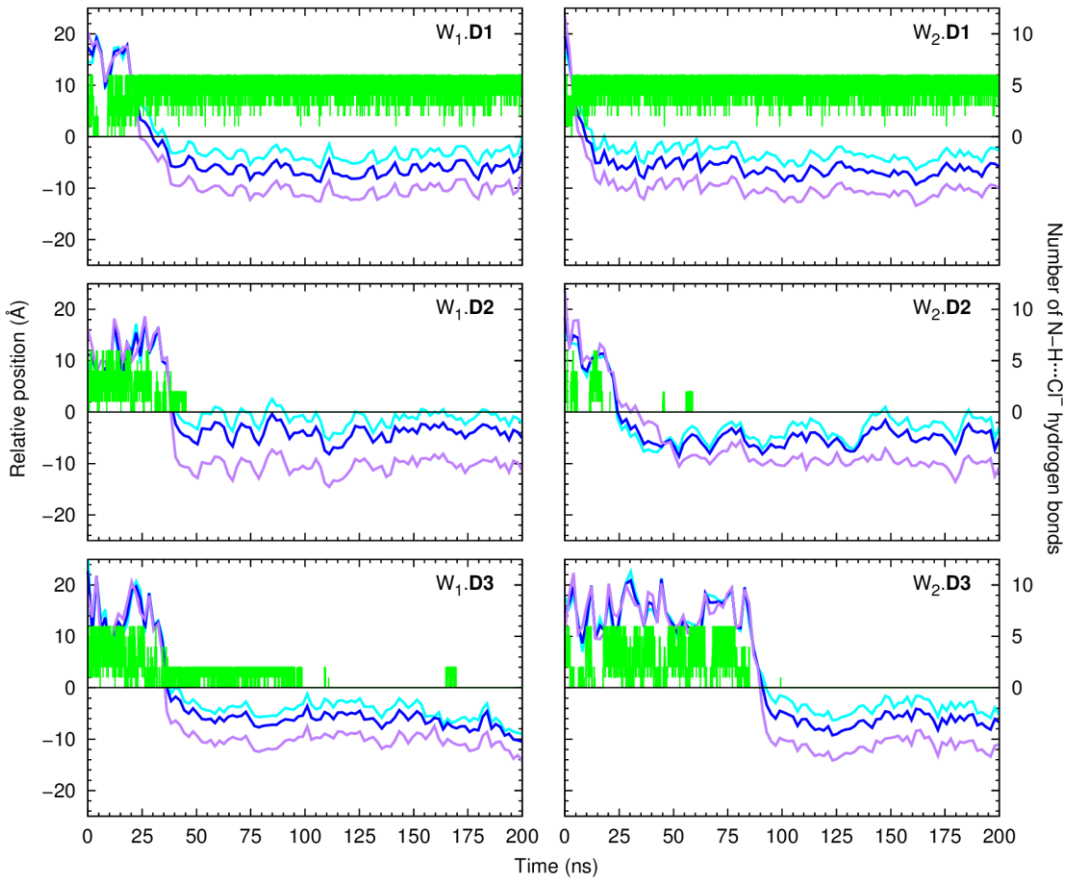


Figure IV-19. Evolution of the $N_{tren}\cdots P_{int}$ (cyan line), $N{-}H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (purple line) distances throughout 200 ns for the MD runs of the **D** series chloride complexes in scenario W. Remaining details as given in Figure IV-4.

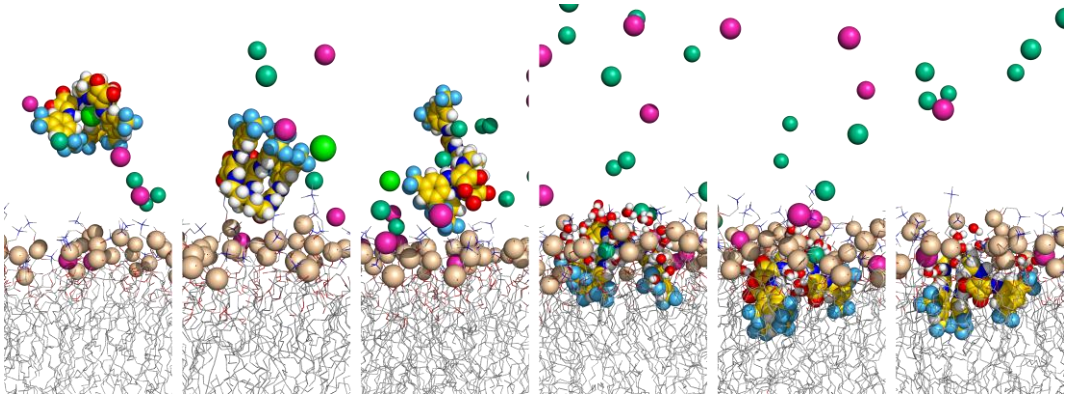


Figure IV-20. Consecutive snapshots depicting the diffusion of the chloride complex of **D3** throughout the first MD run in scenario W. Remaining details as given in Figure IV-6.

Overall, the MD simulations of the **D** series molecules are characterised by multiple chloride release and uptake events, prior to the internalisation of the squaramide derivatives. However, due to the six convergent binding units (and, consequently, higher value of $V_{S,max}$) the initially complexed chloride is replaced by other hydrated chloride ions present in the system. Thus, in both MD runs of **D1** in scenario W, the receptor enters the bilayer coordinating a chloride and this complex is preserved until the end of the MD simulation with up to six N–H···Cl^- hydrogen bonds (see Figure IV-19 and Figure IV-21). In contrast, in both MD runs of **D2** and in simulation W_2.**D3**, the tripodal molecule loses its chloride before permeating the bilayer. On the other hand, in the first MD run of **D3**, the initial anion complex is maintained for a period of *ca.* 61 ns, as the receptor lodges itself within the membrane, followed by the definitive anion release to the water phase. By the end of the MD sampling periods, all binding units of **D2** and **D3** are locked by hydrogen bonds established with the lipid head groups, carbonyl groups of neighbouring squaramide moieties, or water molecules, as shown in Figure IV-21, where the number of hydrogen bonds is plotted *vs.* the relative position of the receptor along the membrane normal.

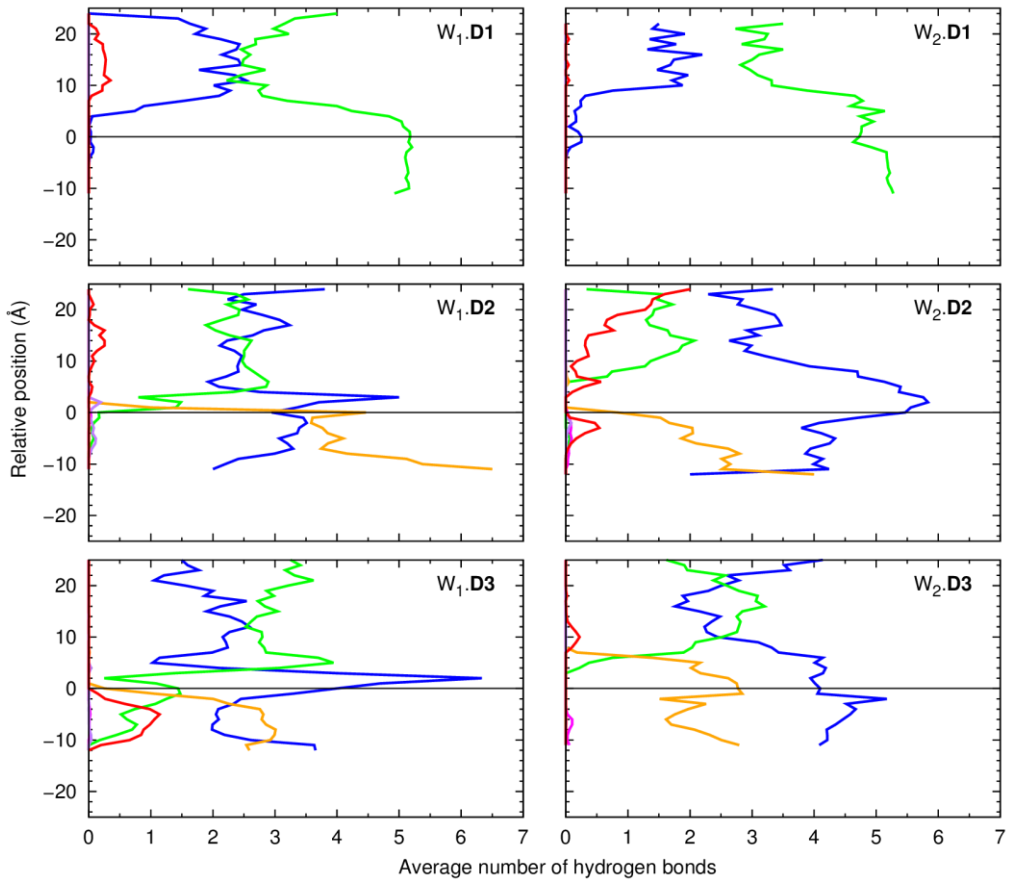


Figure IV-21. Average number of hydrogen bonds *vs*. the relative position of the COM of the **D** series molecules in the MD runs carried out in scenario W. Remaining details as given in Figure IV-5.

To evaluate the conformational changes that the **D** tripodal molecules experience during the passive diffusion process, the evolution of the three N_{tren}–C–C–N torsion angles was monitored throughout the 200 ns MD simulations. When the three individual torsion angles assume values of *ca.* ±60° the squaramide derivative adopts a typical tripodal shape (a *ggg* conformation – *vide supra*). Moreover, if any N_{tren}–C–C–N torsion angle has a value closer to ±180°, the tripodal shape is lost and three non-tripodal conformations are possible: *ggt*, *gtt* and *ttt* (*t* = trans), regardless of the torsion angle position and the angle value. The N_{tren}–C–C–N torsion angles for the MD simulations with the tripodal molecules were histogrammed before (*pre*) and after (*post*) internalisation and are plotted in Figure IV-22. These parameters can be related with the hydrogen bonds established between the N–H binding units and the surrounding water molecules, phosphate head groups or carbonyl moieties as follows:

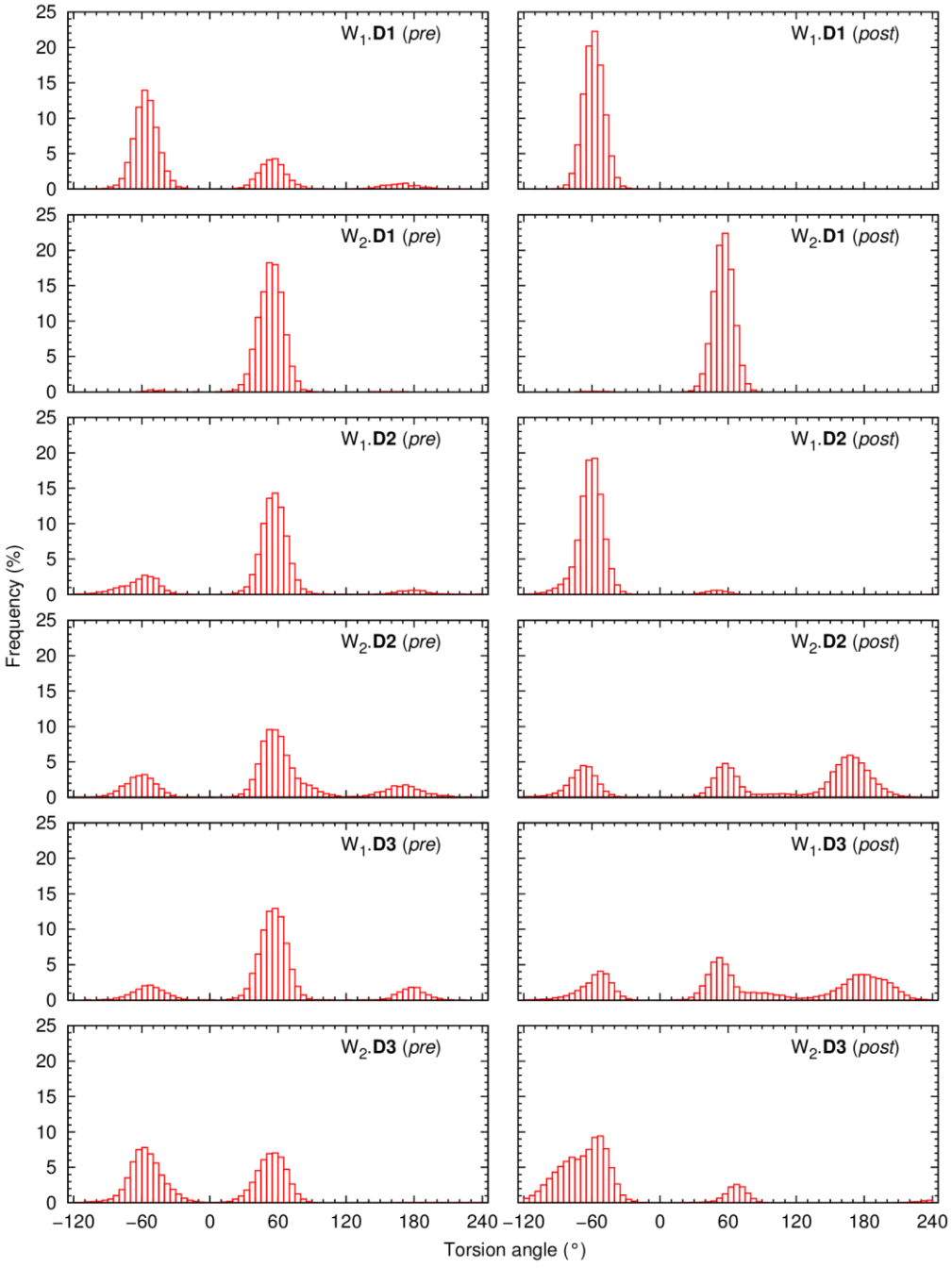


Figure IV-22. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angles values of the **D1** and **D2** molecules in the MD simulations carried out in scenario W, before the interaction with the interface (*pre*) and after (*post*).

The *pre* internalisation histogram of W_1.**D1** exhibits three well defined peaks (centred at −60, 60 and 180°), consistent with the existence of a dominant *ggg* conformation, followed by a transitory *ggt* one, resulting from the replacement of the complexed chloride ion by water molecules. In contrast, the *post* internalisation histogram only shows a single peak (centred at −60°), undoubtedly corresponding to **D1** in tripodal *ggg* shape complexed with chloride. In the second MD run with **D1**, the anion complex is preserved throughout the whole simulation as well as the *ggg* conformation (see Figure IV-21 and Figure IV-22).

Likewise in simulation W_1.**D1**, the internalisation of **D2** in simulation W_1.**D2** is preceded by the replacement of the complexed anion by water molecules with the occurrence of *ggg* and *ggt* conformations. However, after the internalisation, the binding pocket is shared by water molecules and a phosphate head group, as shown in Figure IV-21. Moreover, when the number of hydrogen bonds to the water molecules increases, the interactions with the phosphate head groups decrease and *vice-versa*, although the *ggg* conformation is maintained until the end of the simulation, as indicated by the *post* histogram in Figure IV-22, with a major peak at −60° and a smaller one at 60°. In the second MD run, **D2** presents a more intriguing conformational behaviour. Indeed, after the definitive anion departure, the water molecules, carbonyl moieties and phosphate groups occupy the binding units as the squaramide derivative permeates the bilayer, resulting in a new *gtt* conformation for a substantial simulation length (see Figure IV-21 and Figure IV-22). Afterwards, this conformation is definitively converted into *ggt*, with the N−H binding units interacting only with water molecules and phosphate head groups.

After the anion release, **D3** presents different conformational behaviours along the simulation length of the two runs (see Figure IV-22). Remarkably, in the first MD run, both before and after the internalisation, **D3** assumes mainly a *ggt* conformation followed by a *ggg*. During the passive diffusion, the number of hydrogen bonds to the chloride ions decreases as the receptor binds to water molecules and POPC head groups. Concomitantly, the carbonyl moieties of the squaramide motifs are hydrogen bonded by the N−H of adjacent binding units. Noteworthy, in the second MD run, the number of hydrogen bonding interactions indicates that **D3** recognises chloride ions and water molecules, with the maintenance of the *ggg* conformation (see Figure IV-21 and Figure IV-22). With the internalisation of **D3**, the anion complex is definitely disrupted, with the binding sites recognising water molecules as well as phosphate head groups, maintaining the *ggg* conformation.

Likewise the **A**, **B** and **C** series of squaramide molecules, the ability of the tripodal molecules to assist the chloride ion diffusion within the phospholipid bilayer was also ascertained in scenario M. Initially, the anion complexes were inserted into the POPC bilayer with the C_3 symmetry axis roughly parallel to the membrane normal. The variations on the $N_{tren}\cdots P_{int}$, $N{-}H_{COM}\cdots P_{int}$ and $Ph_{COM}\cdots P_{int}$ distances, along with the number of $N{-}H\cdots Cl^-$ hydrogen bonds, for the long production runs of 500 ns are plotted in Figure IV-23 for the **D** series molecules. The histograms of the $N_{tren}{-}C{-}C{-}N$ torsion angles are plotted in Figure IV-24. Equivalent results were obtained for all runs, as discussed for the M_1.**D1** run.

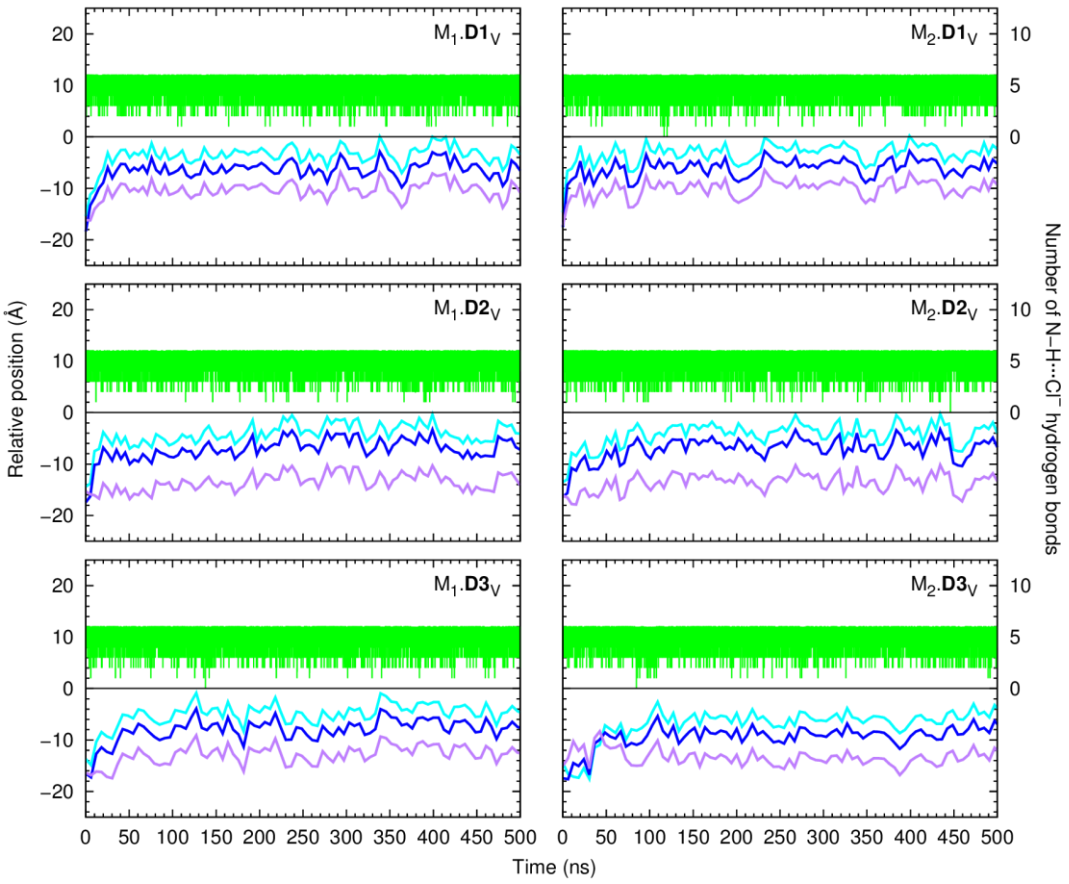


Figure IV-23. Evolution of the $N_{tren}\cdots P_{int}$ (cyan line), $N{-}H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (purple line) distances throughout 500 ns for the MD runs of the **D** series chloride complexes in scenario M in a vertical orientation. Remaining details as given in Figure IV-4.

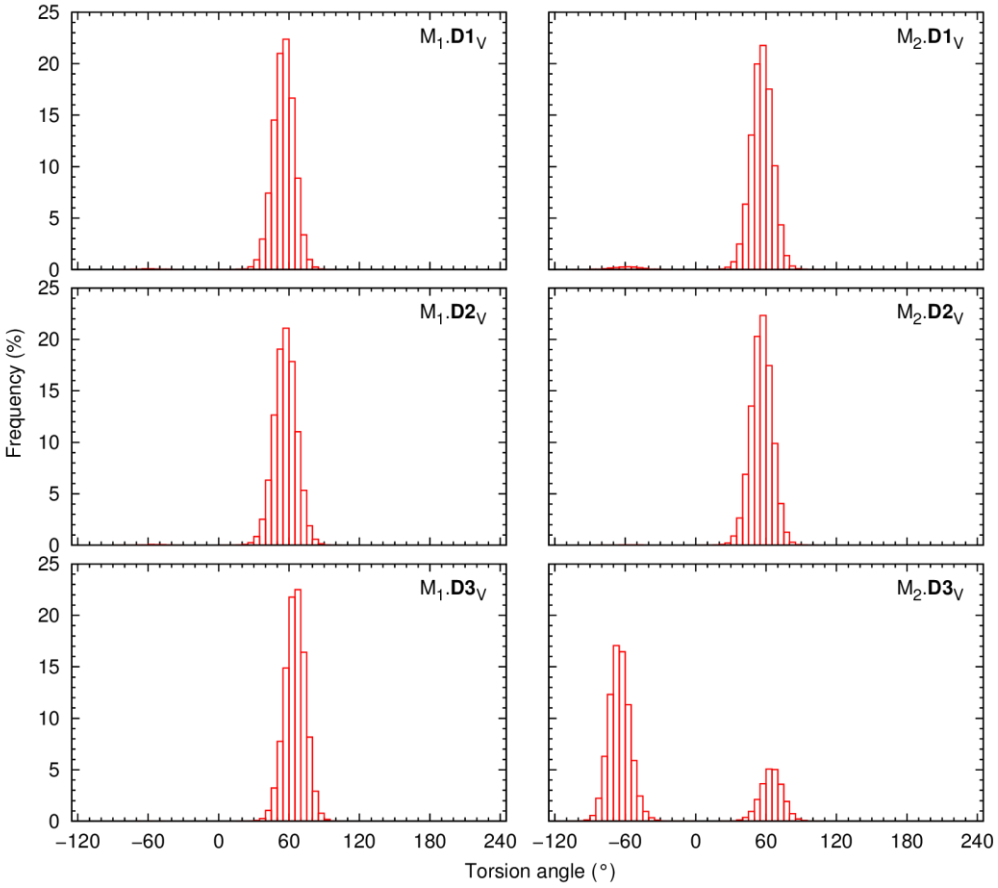


Figure IV-24. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angles values of the **D** series molecules in the MD simulations carried out in scenario W in a vertical orientation.

The complex **D1**·Cl^- promptly diffuses towards the membrane interface with the phenyl groups pointing to the bilayer core, while the tripodal nitrogen atom and the carbonyl groups are closer to the water phase. This preferential orientation is maintained until the end of the simulation time, in excess of 400 ns, with the *ggg* tripodal **D1** receptor tightly bonded to the chloride ion under the water/lipid interface. The pertinent diffusion events are illustrated in Figure IV-25 (left) with three snapshots extracted from the first MD run of **D1**·Cl^- with the receptor in a vertical orientation. Table IV-5 summarises the number of water molecules within a 3.4 Å cut-off of the complexed chloride and squaramide carbonyl groups, which were calculated for the last 400 ns of simulation time, after the complexes reached the interface. The carbonyl groups of the **D** series molecules are exposed to the aqueous phase, being solvated by an average number of water molecules ranging between 3.7 and 4.9. In contrast, in average, there are almost no water molecules around the complexed chloride throughout the last 400 ns of simulation time. In other words, the combination of the tripodal shape and the orientation of the complex shelters the chloride ion from the water molecules. The final orientation of **D1** is entirely consistent with the negative electrostatic potential around the polar carbonyl groups (*vide supra*) and the inherent lipophilic nature of the aromatic substituents of this tripodal molecule.

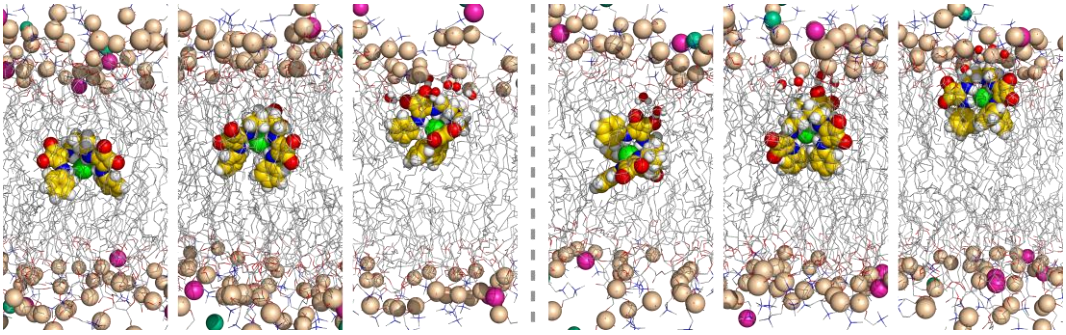


Figure IV-25. Consecutive snapshots depicting the diffusion of the chloride complex of **D1** throughout the first MD run in scenario M: left – initial vertical orientation; right – initial horizontal orientation. Remaining details as given in Figure IV-6.

Having observed that the anion complexes of **D** have the ability to diffuse within the phospholipid bilayer, we decided to ascertain if the results reported above are dependent on the starting orientation of the tripodal anion complexes. Thus, the tripodal chloride complex was alternatively placed at the bilayer core with the C_3 symmetry axis perpendicular to bilayer normal. The current discussion will focus on the M_1.**D1**$_H$ run, given that similar results were observed for the remaining MD run and simulated systems of **D2** and **D3**. The evolution of the $N_{tren}\cdots P_{int}$, $Ph_{COM}\cdots P_{int}$ and $N{-}H_{COM}\cdots P_{int}$ distances, as well as the $N_{tren}{-}C{-}C{-}N$ torsion angles histograms are plotted in Figure IV-26 and Figure IV-27, respectively, for the MD runs with the **D** series complexes starting with the horizontal orientation in scenario M.

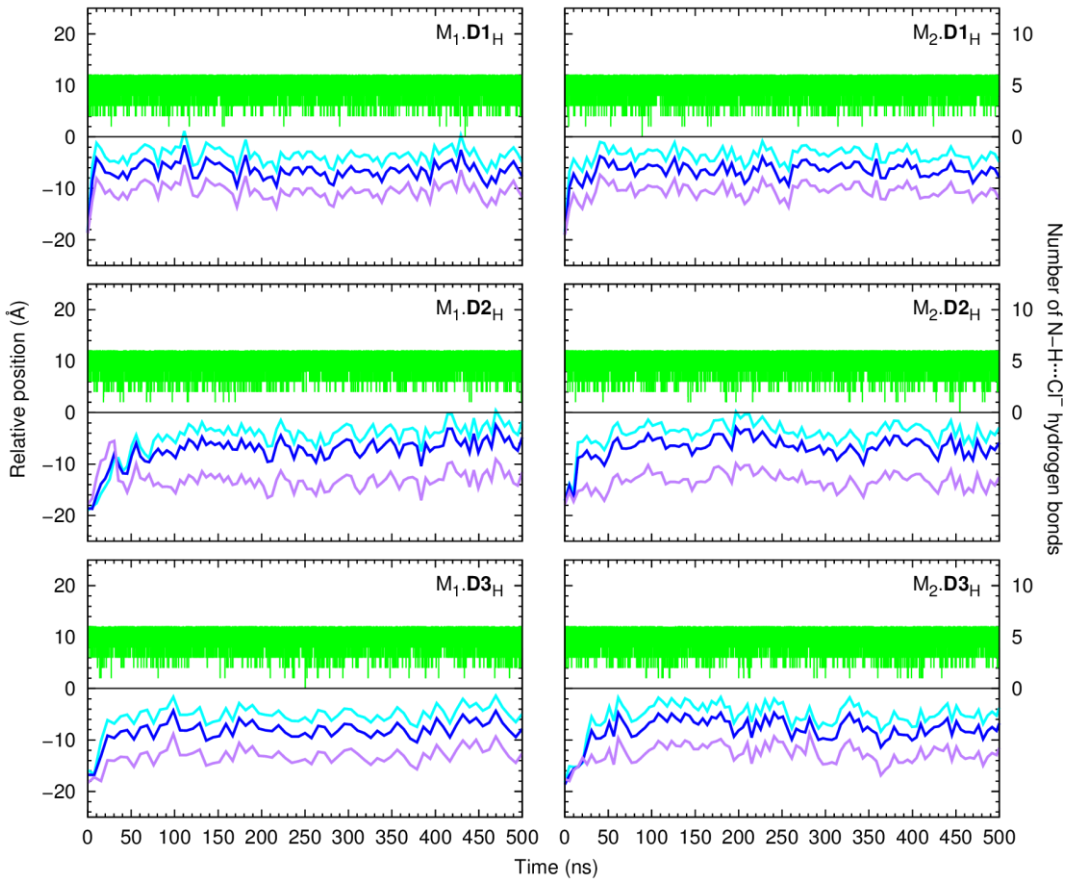


Figure IV-26. Evolution of the $N_{tren}\cdots P_{int}$ (cyan line), $N{-}H_{COM}\cdots P_{int}$ (blue line) and $Ph_{COM}\cdots P_{int}$ (purple line) distances throughout 500 ns for the MD runs of the **D** series chloride complexes in scenario M in a horizontal orientation. Remaining details as given in Figure IV-4.

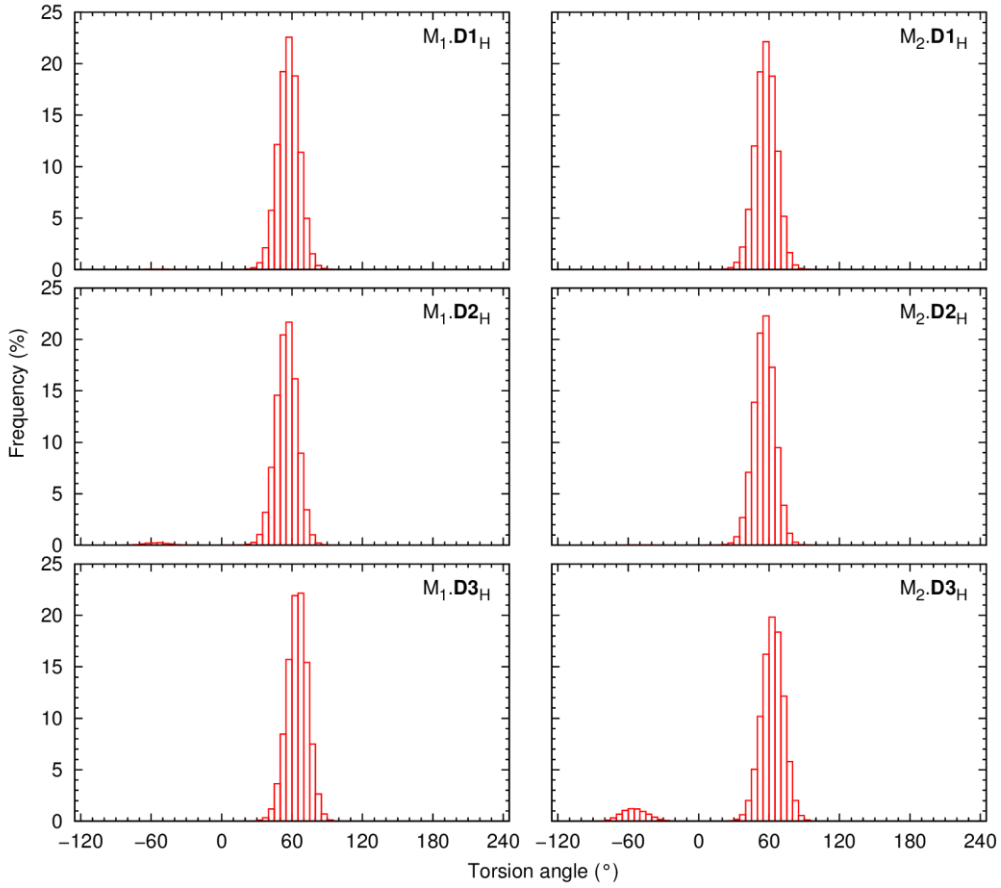


Figure IV-27. Frequency histograms showing the distribution of the three N_{tren}–C–C–N torsion angles values of the **D** series molecules in the MD simulations carried out in scenario W in a horizontal orientation.

D1·Cl^- is able to reorient itself during the first ns of the MD production run, concomitantly with the diffusion along the bilayer normal towards the water/lipid interface. In other words, equivalent diffusion events were observed independently of the initial orientation of the complex inside of the phospholipid bilayer, as can be seen with selected snapshots of this MD simulation in Figure IV-25 (right). As the **D** series are able to permeate the interface complexed with chloride and are unable to release it when placed in the membrane core, the MD simulations in scenarios W' and M' were not undertaken.

In summary, the **D** series anion complexes are preserved during the extensive 6 µs sampling period of scenario M simulations, which is entirely consistent with the inability of this series to operate the chloride transport as experimentally observed. Therefore, if these tripodal molecules are able to uptake a chloride at the water/lipid interface level (as observed in some MD simulations of scenario W), it is very unlikely that the chloride release to the water phase can occur, as requested by the anion carrier mechanism. Furthermore, the MD insights suggest that lack of transport activity is mainly due to the sheltering tripodal architecture synergistically associated with a superior binding affinity of the **D** molecules for chloride ion, which is characterised by higher $V_{S,max}$ values inside of its tripodal conformation.

IV.3.3. Free energy calculations: SMD and US simulations

The energetic barriers associated with the anion transmembrane transport were estimated from the Potential of Mean Force (PMF) profiles obtained by combining US simulations and the Weighted Histogram Analysis Method (WHAM).[89-91] The US simulations were preceded by SMD simulations to generate the starting positions of the three **A** free squaramides and their chloride complexes along the membrane normal (*vide supra*). The diffusion process of the squaramide derivatives along the *z*-dimension was monitored, with special attention to the changes in hydration and hydrogen bonding interactions.

SMD simulations

Throughout the SMD simulations, several water molecules accompany the diffusing anion complexes or free receptors, as plotted in Figure IV-28 and Figure IV-29, respectively. The translocation of these entities across the lipid bilayer drags many water molecules. The hydration of the squaramide derivatives decreases drastically as they are closer to the bilayer core (z = 0 Å), from *ca*. 36 water molecules (for the smaller **A1**) to a minimum of *ca*. 3 at the centre of the POPC membrane. In the SMD simulations of the free transporters, these water molecules are hydrogen bonded to the N–H binding sites, while in the chloride complexes they also surround the anion.

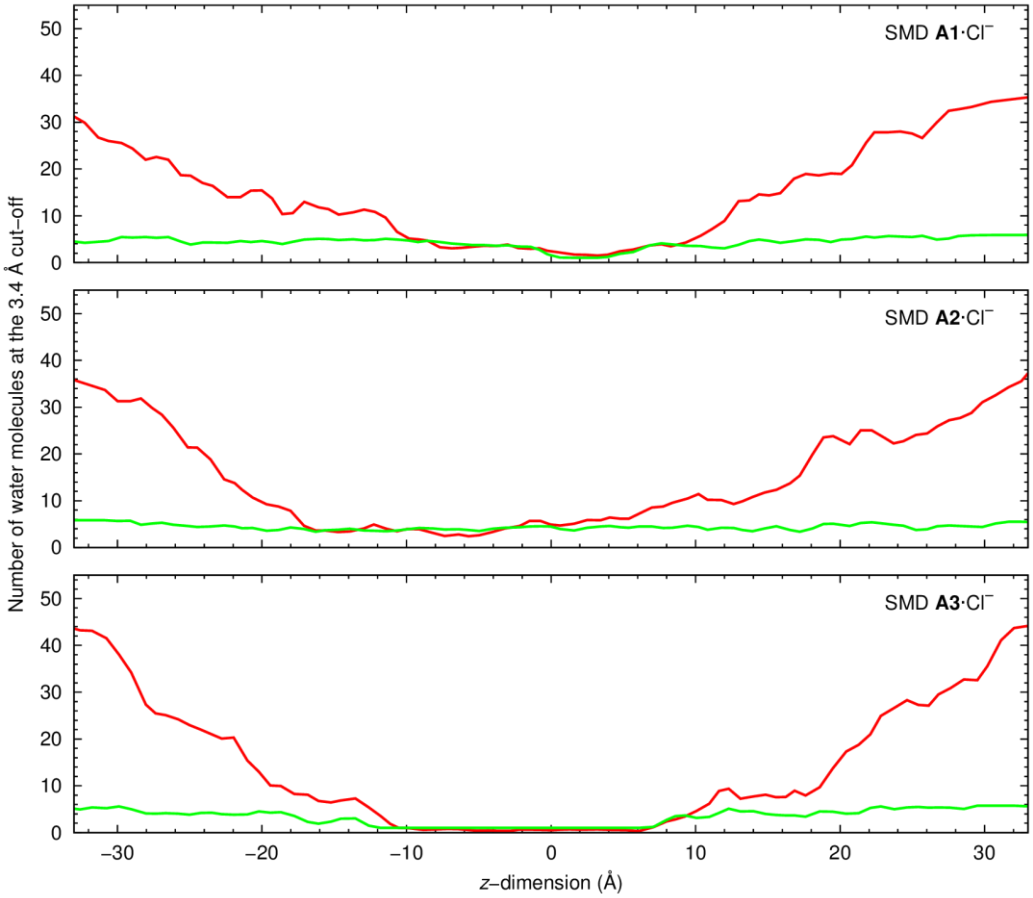


Figure IV-28. Variation of the number of water molecules within the solvation shell defined by a cut-off of 3.4 Å from the **A** series squaramides (red line) and their complexed chloride (green line), in the SMD simulations of the chloride complexes from one side of the POPC to the opposite side. Data were smoothed using Bézier curves.

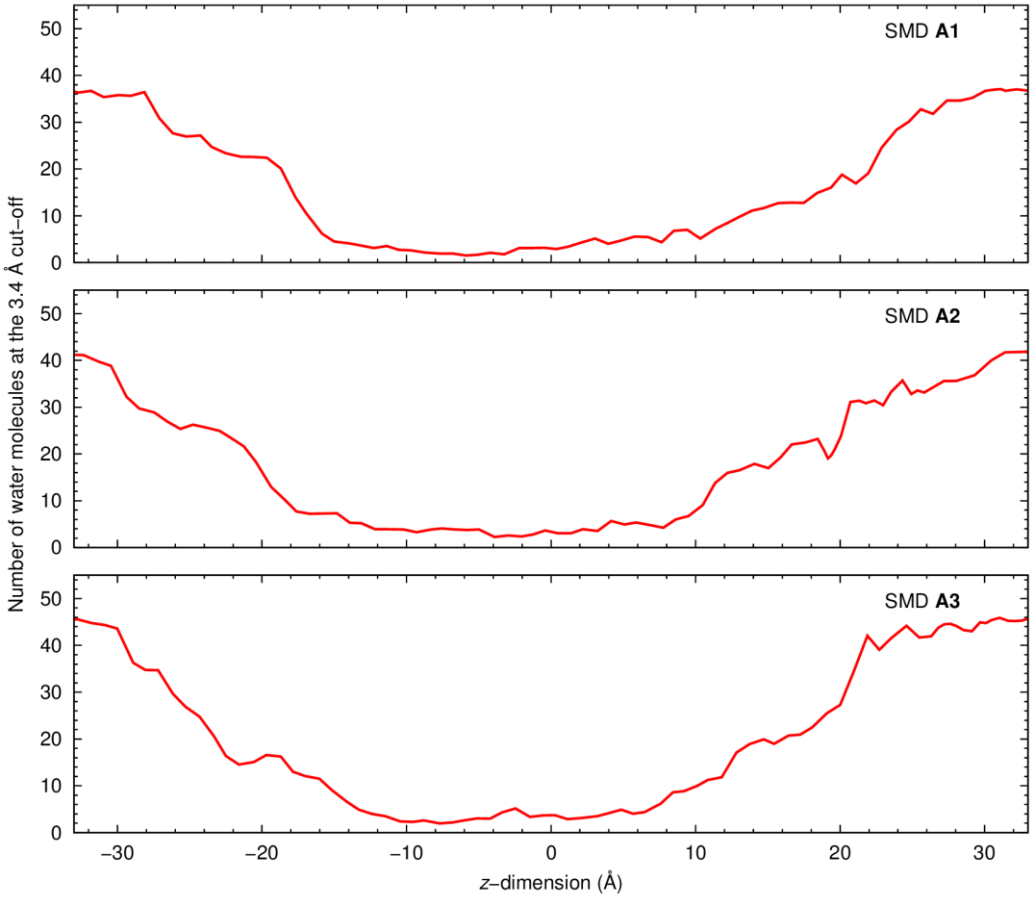
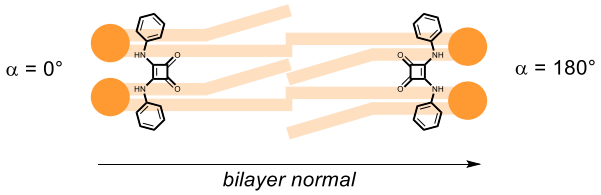


Figure IV-29. Variation of the number of water molecules within the solvation shell defined by a cut-off of 3.4 Å from the **A** series squaramides (red line), in the SMD simulations of the free squaramide derivatives from one side of the POPC to the opposite side. Data were smoothed using Bézier curves.

US simulations

In the 67 US simulations used to restore the PMF profile of each studied system, the orientation of the squaramide derivatives was monitored ascertaining the tilt angle α between the bilayer normal and the vector bisecting the N–H binding sites and the carbonyl groups. An angle of 0° or of 180° indicates that the vector and the bilayer normal are aligned; however, when the N–H binding units are pointing in the negative direction of the membrane normal, α assumes a value of 0°, while an angle of 180° indicates that the N–H binding units of the squaramide motif point towards the positive direction of the *z*-dimension, as sketched in Scheme IV-4. The α angle was histogrammed along the *z*-dimension for all systems and is plotted in Figure IV-30 and Figure IV-31. Overall, in water phase (for *z*-dimension values under *ca.* −20 and above +20 Å), as expected, the squaramide derivatives adopt an undefined orientation in either free or complexed form. On the other hand, in the US windows comprised between −20 and +20 Å, where the transporters are found within the POPC bilayer and restricted by the highly packed lipid molecules, the α angles assume values close to 0° or to 180°. In other words, during the US simulations' equilibration period, the molecules reoriented themselves, with the N–H binding units always pointing to one of the water/lipid interfaces. This spatial orientation corresponds to a type ***a*** disposition observed in the passive diffusion MD simulations. (*vide supra*). The snapshots shown in Figure IV-32 were taken from the US windows at −30, −15, 0, +15 and +30 Å and illustrate the

preferential orientations of **A1**·Cl^- during the diffusion trough the phospholipid bilayer. Consequently, they also enlighten the putative anion carrier mechanism proposed for these synthetic drug-like molecules.



Scheme IV-4. Orientations of the **A** series squaramides (represented by **A1**) for α angle values of 0 or 180°.

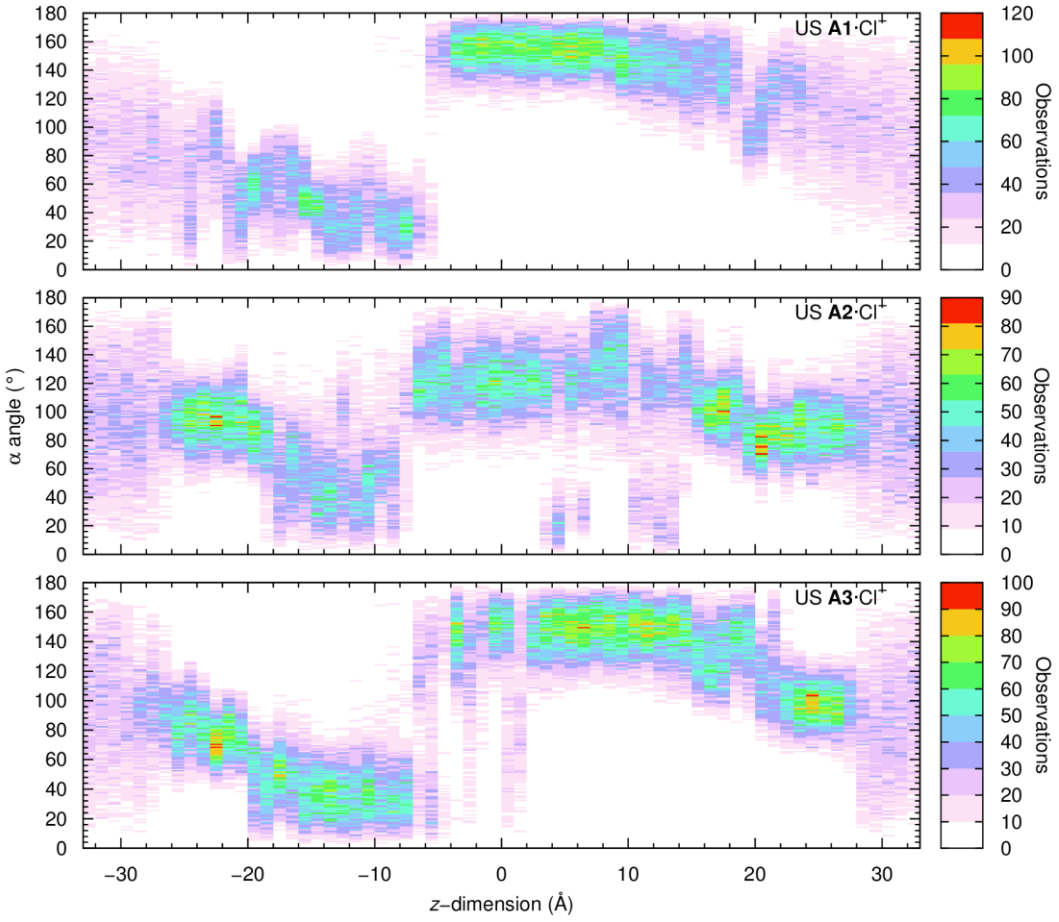


Figure IV-30. 2D histograms created from the α angle values monitored along the *z*-dimension positions of the chloride complexes of the **A** series squaramides throughout the 67 independent US windows. The colour ranges from white (no occurrence) to red (several occurrences).

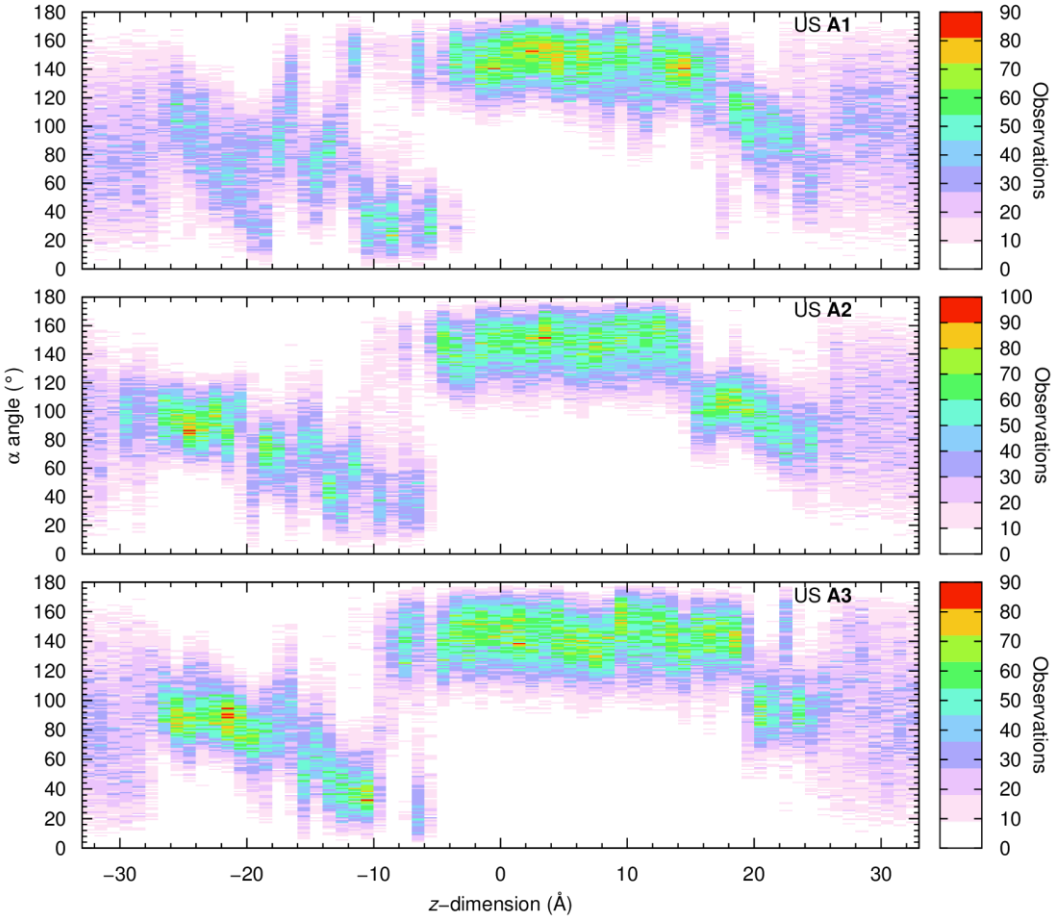


Figure IV-31. 2D histograms created from the α angle values monitored along the *z*-dimension positions of the free **A** series squaramides throughout the 67 independent US windows. The colour ranges from white (no occurrence) to red (several occurrences).

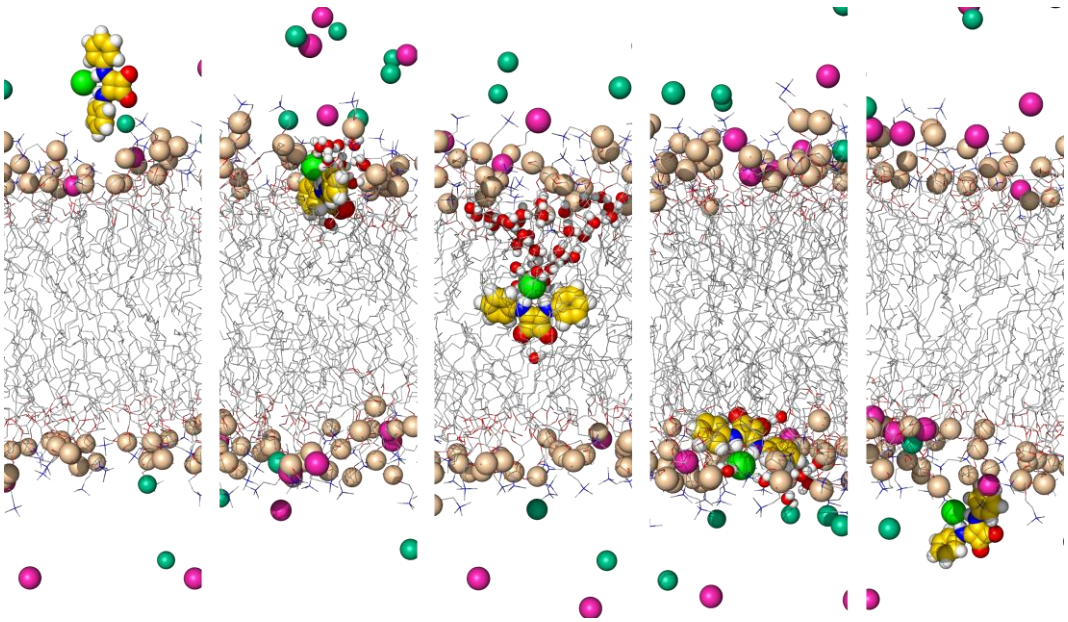


Figure IV-32. Illustrative snapshots depicting the orientation of **A1**·Cl^- in the US windows at z = −30, −15, 0, +15 and +30 Å (from left to right). Remaining details as given in Figure IV-6.

The PMF profiles along the membrane normal were estimated for the three molecules of the **A** series, in their free and chloride complexed forms. For energetic comparison purposes, the PMF of a single free chloride was also estimated in the pure bilayer model used in this work. The energy profiles, normalised to zero in bulk water and estimated spanning the whole membrane model are shown in Figure IV-33, for the free transporters (left) and for their chloride complexes (right), together with the free chloride. The individual profiles with the calculated Monte Carlo Bootstrap error analysis,[92] determined with 1000 Monte Carlo trials, are shown in Figure IV-34, Figure IV-35 and Figure IV-36. The PMF profiles, spanning the membrane, are asymmetric indicating that the entry and the exit of the POPC bilayer is accompanied by different interactions with water molecules and/or phospholipids,[93] regardless of the orientation of the squaramide derivatives. The interactions of the small synthetic transporters with the water/lipid interface are discussed below. The US windows with the squaramide derivatives were simulated for 50 ns, with the last 20 ns being used to estimate the PMF profiles of the six systems. The initial 30 ns were discarded as equilibration period, given that the PMF profiles only converged after this period as shown in Figure IV-37. On the other hand, for the free chloride, the US simulations were 25 ns long with the initial 15 ns corresponding to the equilibration period.

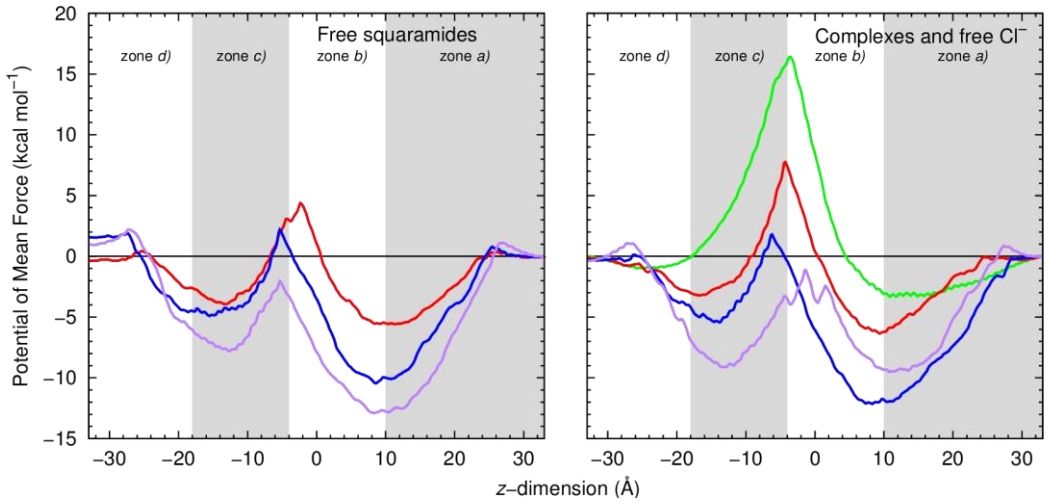


Figure IV-33. PMF as a function of the **A** squaramides' distance to the membrane COM (z = 0 Å), for the free squaramides (left) and for the chloride complexes (right). The red, blue and purple lines correspond to **A1**, **A2** and **A3**, in this order, while the green line corresponds to the PMF of a free chloride.

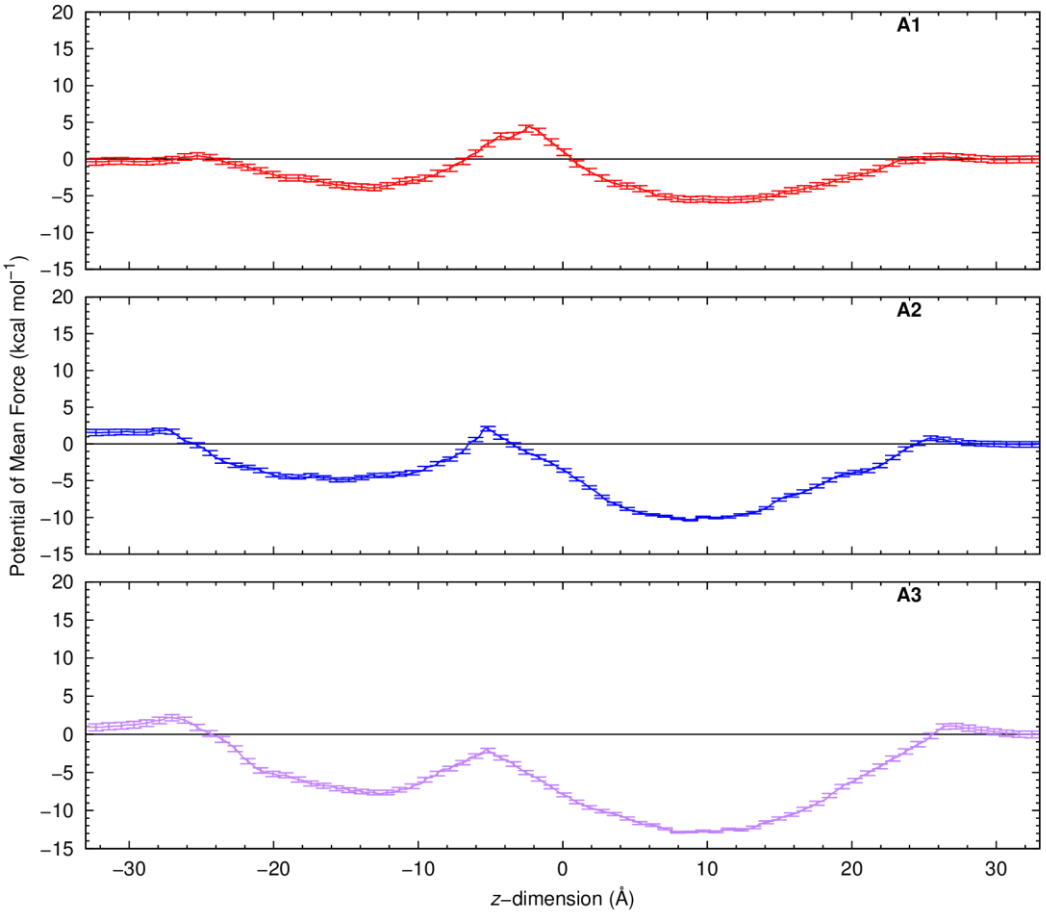


Figure IV-34. PMF as a function of the **A** series free squaramides' distance to the membrane COM (z = 0 Å). The error bars correspond to the bootstrap errors calculated from 1000 Monte Carlo trials and are upscaled 20 times.

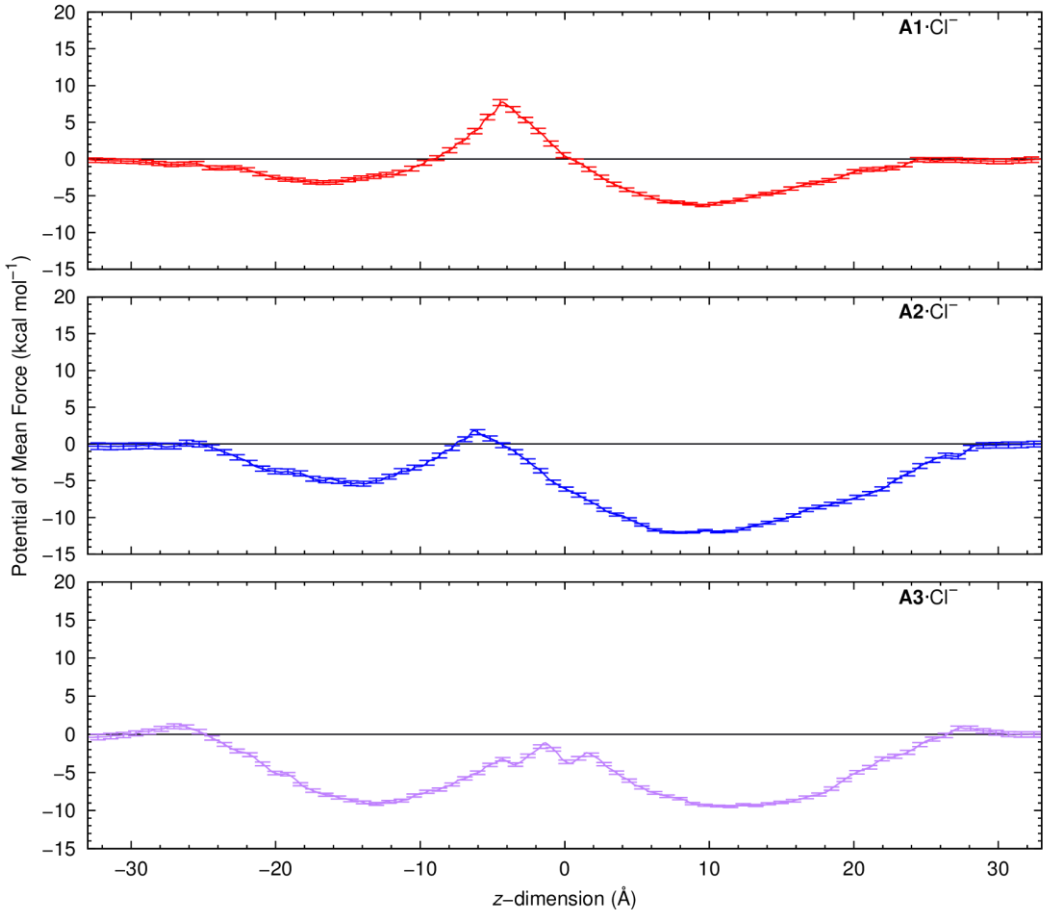


Figure IV-35. PMF as a function of the **A** series chloride complexes' distance to the membrane COM (z = 0 Å). The error bars correspond to the bootstrap errors calculated from 1000 Monte Carlo trials and are upscaled 20 times.

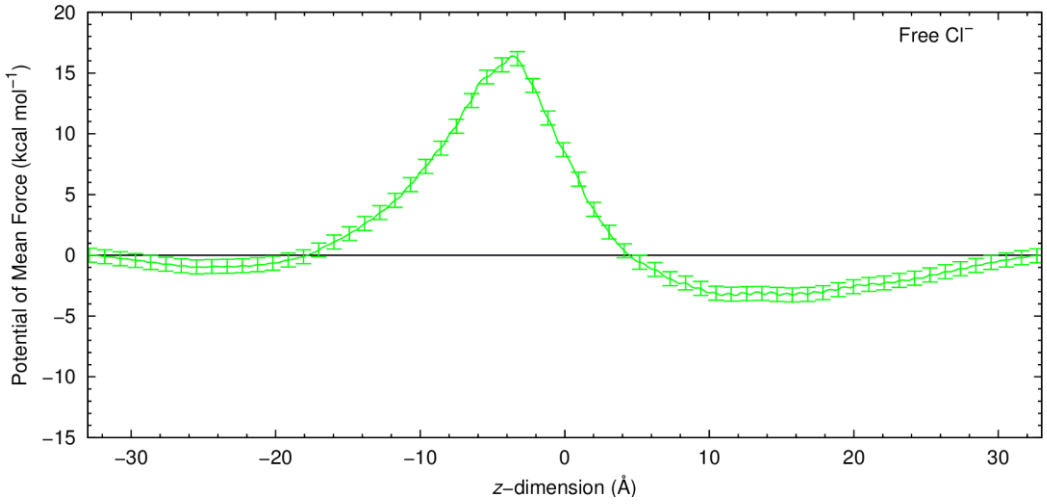


Figure IV-36. PMF as a function of the free chloride distance to the membrane COM (z = 0 Å). The error bars correspond to the bootstrap errors calculated from 1000 Monte Carlo trials.

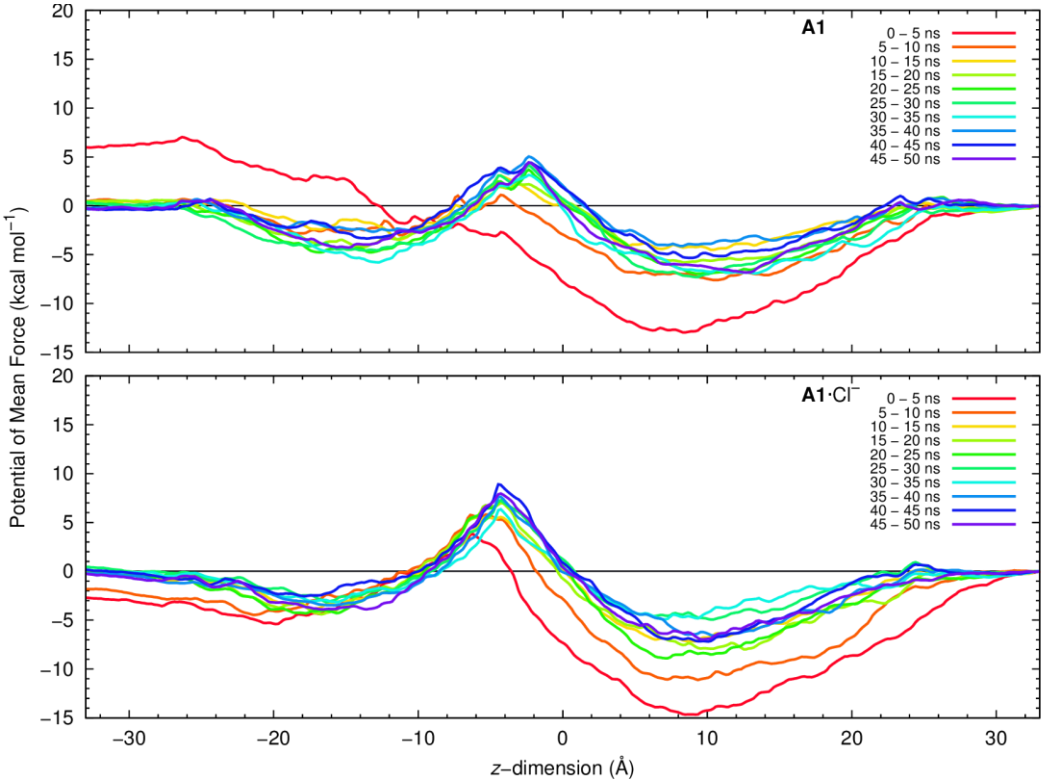


Figure IV-37. The equilibration and convergence of the US simulation windows were assessed in 5.0 ns intervals (coloured according to time period), taken through the trajectory and used to calculate the PMF of **A1** and **A1**·Cl⁻. The PMF curves for **A1** and **A1**·Cl⁻ converged within the initial 30 ns.

Overall, the profiles of the three complexes show that the three receptors lower the energetic barrier for the transmembrane transport of chloride, when compared with the free anion profile, thus being able to facilitate the chloride transport as anion carriers. Furthermore, the energy profiles are divided in four zones, identified in Figure IV-33: *a)* recognition and permeation of the bilayer, leading to the first basin; *b)* the energetic barrier associated with the diffusion within the highly packed bilayer medium; *c)* the second basin, after the core of the bilayer; and *d)* before the protrusion of the opposite water/lipid interface.

The permeation of the **A1** chloride complex is favoured by 6.4 kcal·mol^{-1} at z = 9.4 Å, *i.e.*, when the complex is nested just below the water/lipid interface and surrounded by few water molecules (zone *a*). The diffusion of the complex within the phospholipids' aliphatic tails (zone *b*), towards the bilayer core, naturally results in the increase of the energy barrier to a maximum of *ca*. 7.7 kcal·mol^{-1} at z = −4.4 Å. The approach to the opposite water/lipid interface (zones *c* and *d*) is accompanied by the decrease of energy to −3.2 kcal·mol^{-1} at z = −16.6 Å, as the complex is solvated with water molecules coming from the closest aqueous phase (see Figure IV-32). The energy profile of **A2**'s chloride complex mirrors the PMF of **A1**·Cl⁻, with the entry being lowered by 12.1 kcal·mol^{-1}, at z = 8.3 Å. The energy increases from −12.1 to 1.8 kcal·mol^{-1} when the complex is positioned at z = −6.1 Å. In contrast, the PMF of **A3**'s chloride complex displays an energetic profile with two deep basins (zones *a* and *c*), at similar distances from the bilayer core (*ca*. 13 Å), with energy minima of *ca*. −9.3 kcal·mol^{-1} relative to the water phase. Between these basins, the maximum of −1.1 kcal·mol^{-1} occurs at z = −1.4 Å (zone *b*). This odd energy profile should be intimately linked with the higher lipophilic character of **A3** when compared with the other two analogous. This complex feels better within the phospholipid environment and needs to overcome an energetic barrier of *ca*. 10 kcal·mol^{-1} to leave the bilayer from z = −13.2 Å (zone *d*). At a similar position, the less lipophilic complexes need only to overcome barriers of 3.2 (**A1**·Cl⁻) or 5.4 (**A2**·Cl⁻) kcal·mol^{-1}.

In other words, while **A1**·Cl^- and **A2**·Cl^- have low energetic barriers to exit the bilayer, **A3**·Cl^- has to surmount a larger energy barrier to reach the water phase. In addition, **A2**·Cl^- has a smaller energetic barrier to cross the bilayer core, most likely due to its lipophilic character. These results indicate that **A2** should be energetically favoured to assist the chloride transport when compared with the analogous **A1** and **A3**, which is in line with the experimental transport studies.

The energy profiles associated with the diffusion of the free squaramide derivatives across the POPC bilayer are quite similar. In zone *a*, it is evident that all receptors' permeation of the bilayer is energetically driven, being favoured by 5.6 kcal·mol^{-1} for **A1**, by 10.4 kcal·mol^{-1} for **A2** and by 12.9 kcal·mol^{-1} for **A3**. These results are in line with the MD simulation carried out in scenario W, in which the permeation of the water/lipid interface is accompanied by hydrogen bonding interactions to the phospholipid head groups and water molecules. Free **A1**, **A2** and **A3** face comparable energy barriers of 9.9, 12.6, and 10.8 kcal·mol^{-1}, in this order, to surmount the bilayer core. After the energetic barrier, the energy drop to the second basin is of 8.3 kcal·mol^{-1} for **A1**, while **A2** and **A3** are favoured by 7.1 and 5.8 kcal·mol^{-1}, respectively, to stay in the basin of zone *c*. Despite the different energetic profiles in zones *a* to *c*, receptors **A1** and **A2** show low energetic barriers to exit the bilayer, 3.6 and 6.5 kcal·mol^{-1} respectively, as also observed for their anion associations. On the other hand, free **A3** has to overcome a higher energy barrier of *ca*. 8.7 kcal·mol^{-1} to leave the bilayer. These profiles show that the differences between free **A1** and **A2** to span the bilayer are quite small, while **A3**, the most lipophilic molecule, faces several energetic barriers to go from one side of the bilayer to the other.

In summary, these energetic outcomes suggest that squaramides **A1**-**A3** should operate the anion transport event without leaving the phospholipid bilayer, shuttling between both water/lipid interfaces. In this context, the anion transport by **A2** is thermodynamically favoured when compared with **A1** and **A3**.

IV.4. Conclusions

Among the squaramides series, **C** and **D**, with four and six convergent binding sites, respectively, have higher E^2 and $V_{S,max}$ values, and therefore, higher binding affinities for chloride. In agreement, the MD simulations carried out in scenario W showed that these series' molecules are able to recognise and release chloride ion in water phase. In addition, the **C** series, **D1** and **D3** are able to permeate the water/lipid interface complexed with a chloride ion. On the other hand, after permeating the interface, only the molecules of the **D** series, with a more flexible tripodal shape, intermittently bind chloride from the water phase. In stark contrast, the MD simulations of series **C** and **D** squaramide derivatives show that chloride complexes initially positioned in the bilayer core (scenario M) promptly diffuse towards the interface and are maintained throughout the simulation length. Given that the chloride uptake and release events are *condiciones sine quibus non* for a small drug-like molecule to operate as an anion transmembrane transporter, this structural result is entirely consistent with the absence of experimental transport activity for squaramide derivatives **D**. The encapsulated chloride is sheltered from the water molecules by the tripodal scaffold architecture of these receptors, as shown by the negligible number of solvent molecules surrounding the anion (see Table IV-5). The **C** series molecules also tightly held the chloride ion, but its hydration is higher than in the **D** series, hinting that the release of the anion is more likely, although it has not been observed in the reported MD simulations.

The **A** series molecules, with more moderate E^2 and $V_{S,max}$ values, when located below the water/lipid interface, adopt well-defined spatial dispositions with the N–H binding units pointing to the water phase. This orientation enables the sporadic exchange of the hydrogen bonding interactions established to the phosphate head groups or water molecules with N–H···Cl^- hydrogen bonds formed with a chloride ion uptaken from the water phase. In contrast, the **B** series molecules, with overall lower

E^2 and $V_{S,max}$ values, are also able to bind chloride in water (scenario W), but, after the permeation, due to their asymmetric nature, are unable to preserve a well-defined orientation consistent with occurrence of further chloride binding events. This comparison suggests that the **A** series molecules are better anion carriers than **B**, corroborating the experimental transport data.

Our PMF calculations have unequivocally indicated that the chloride transport across the phospholipid bilayer is energetically favoured by the **A** series' molecules when compared with its unassisted translocation. Moreover, the PMF profiles have shown that the free and anion complexed squaramides are preferentially located below the water/lipid interface level, supporting the passive diffusion MD simulations. The energy profiles of free **A1**–**A3** on positive *z*-dimension values follow their lipophilicities, with **A3**, the most lipophilic squaramide, being thermodynamically favoured over **A1** and **A2**. For the chloride complexes of these three linear transporters, there is no significant energetic differences for their diffusions until the membrane core (z = 0 Å). However, the chloride complex of **A3** has to overcome the lowest energy barrier to reach the interface on the negative side of the *z*-dimension. The asymmetric nature of the energy profiles relatively to z = 0 Å is a consequence of the reorientation of the linear squaramide transporters within the phospholipid bilayer coupled with the solvation chains preserved along the US simulation windows.

Overall, this comprehensive theoretical study shows that our results are entirely consistent with the anion carrier mechanism proposed for the **A** series,[41] which are able to easily shuttle within the hydrophobic medium of the phospholipid bilayer as free transporters or as chloride complexes. Furthermore, the chloride uptake and release events can occur in the water or, alternatively and more likely, with the squaramides nested at the water/lipid interface level, without surmounting the energy barrier associated with the transporter's membrane exit.

IV.5. References

1. G. G. Celesia, *Clin. Neurophysiol.*, 2001, **112**, 2-18.
2. J. LaRusch, J. Jung, I. J. General, M. D. Lewis, H. W. Park, R. E. Brand, A. Gelrud, M. A. Anderson, P. A. Banks, D. Conwell, C. Lawrence, J. Romagnuolo, J. Baillie, S. Alkaade, G. Cote, T. B. Gardner, S. T. Amann, A. Slivka, B. Sandhu, A. Aloe, M. L. Kienholz, D. Yadav, M. M. Barmada, I. Bahar, M. G. Lee, D. C. Whitcomb and G. North American Pancreatitis Study, *PLoS Genet.*, 2014, **10**, e1004376.
3. R. Planells-Cases and T. J. Jentsch, *Biochim. Biophys. Acta*, 2009, **1792**, 173-189.
4. P. A. Gale, *Coord. Chem. Rev.*, 2000, **199**, 181-233.
5. P. D. Beer and P. A. Gale, *Angew. Chem. Int. Ed.*, 2001, **40**, 486-516.
6. P. A. Gale, *Coord. Chem. Rev.*, 2001, **213**, 79-128.
7. P. A. Gale, *Coord. Chem. Rev.*, 2003, **240**, 191-221.
8. P. A. Gale and R. Quesada, *Coord. Chem. Rev.*, 2006, **250**, 3219-3244.
9. P. A. Gale, S. E. Garcia-Garrido and J. Garric, *Chem. Soc. Rev.*, 2008, **37**, 151-190.
10. C. Caltagirone and P. A. Gale, *Chem. Soc. Rev.*, 2009, **38**, 520-563.
11. P. A. Gale, *Chem. Soc. Rev.*, 2010, **39**, 3746-3771.
12. S. Matile, A. Vargas Jentzsch, J. Montenegro and A. Fin, *Chem. Soc. Rev.*, 2011, **40**, 2453-2474.
13. M. Wenzel, J. R. Hiscock and P. A. Gale, *Chem. Soc. Rev.*, 2012, **41**, 480-520.
14. N. Sakai and S. Matile, *Langmuir*, 2013, **29**, 9031-9040.
15. P. A. Gale, N. Busschaert, C. J. Haynes, L. E. Karagiannidis and I. L. Kirby, *Chem. Soc. Rev.*, 2014, **43**, 205-241.
16. P. A. Gale, E. N. W. Howe and X. Wu, *Chem*, 2016, **1**, 351-422.
17. A. V. Jentzsch, D. Emery, J. Mareda, S. K. Nayak, P. Metrangolo, G. Resnati, N. Sakai and S. Matile, *Nat Commun*, 2012, **3**, 905.
18. A. Vargas Jentzsch, A. Hennig, J. Mareda and S. Matile, *Acc. Chem. Res.*, 2013, **46**, 2791-2800.
19. A. Vargas Jentzsch and S. Matile, *J. Am. Chem. Soc.*, 2013, **135**, 5302-5303.
20. A. V. Jentzsch and S. Matile, *Top. Curr. Chem.*, 2015, **358**, 205-239.
21. S. Benz, M. Macchione, Q. Verolet, J. Mareda, N. Sakai and S. Matile, *J. Am. Chem. Soc.*, 2016, **138**, 9093-9096.
22. S. Tomas, M. C. Rotger, J. F. Gonzalez, P. M. Deya, P. Ballester and A. Costa, *Tetrahedron Lett.*, 1995, **36**, 2523-2526.
23. J. O. Morley, *Journal of Molecular Structure-Theochem*, 1995, **357**, 49-57.
24. R. Prohens, S. Tomas, J. Morey, P. M. Deya, P. Ballester and A. Costa, *Tetrahedron Lett.*, 1998, **39**, 1063-1066.

25. D. Quiñonero, A. Frontera, G. A. Suñer, J. Morey, A. Costa, P. Ballester and P. M. Deyà, *Chem. Phys. Lett.*, 2000, **326**, 247-254.
26. D. Quinonero, A. Frontera, P. Ballester and P. M. Deya, *Tetrahedron Lett.*, 2000, **41**, 2001-2005.
27. R. Prohens, M. C. Rotger, M. N. Pina, P. M. Deya, J. Morey, P. Ballester and A. Costa, *Tetrahedron Lett.*, 2001, **42**, 4933-4936.
28. D. Quiñonero, R. Prohens, C. Garau, A. Frontera, P. Ballester, A. Costa and P. M. Deyà, *Chem. Phys. Lett.*, 2002, **351**, 115-120.
29. M. Neus Piña, M. Carmen Rotger, A. Costa, P. Ballester and P. M. Deyà, *Tetrahedron Lett.*, 2004, **45**, 3749-3752.
30. V. Ramalingam, M. E. Domaradzki, S. Jang and R. S. Muthyala, *Org. Lett.*, 2008, **10**, 3315-3318.
31. A. Rostami, A. Colin, X. Y. Li, M. G. Chudzinski, A. J. Lough and M. S. Taylor, *J. Org. Chem.*, 2010, **75**, 3983-3992.
32. V. Amendola, G. Bergamaschi, M. Boiocchi, L. Fabbrizzi and M. Milani, *Chem. Eur. J.*, 2010, **16**, 4368-4380.
33. A. Rostami, C. J. Wei, G. Guerin and M. S. Taylor, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 2059-2062.
34. B. W. Gung and S. C. Schlitzer, *Tetrahedron Lett.*, 2015, **56**, 5043-5047.
35. J. A. Butera, M. M. Antane, S. A. Antane, T. M. Argentieri, C. Freeden, R. F. Graceffa, B. H. Hirth, D. Jenkins, J. R. Lennox, E. Matelan, N. W. Norton, D. Quagliato, J. H. Sheldon, W. Spinelli, D. Warga, A. Wojdan and M. Woods, *J. Med. Chem.*, 2000, **43**, 1187-1202.
36. J. A. Butera, D. J. Jenkins, J. R. Lennox, J. H. Sheldon, N. W. Norton, D. Warga and T. M. Argentieri, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2495-2501.
37. R. I. Storer, C. Aciro and L. H. Jones, *Chem. Soc. Rev.*, 2011, **40**, 2330-2346.
38. P. M. Gloria, J. Gut, L. M. Goncalves, P. J. Rosenthal, R. Moreira and M. M. Santos, *Bioorg. Med. Chem.*, 2011, **19**, 7635-7642.
39. T. S. Elliott, A. Slowey, Y. L. Ye and S. J. Conway, *Medchemcomm*, 2012, **3**, 735-751.
40. S. P. Kumar, P. M. C. Glória, L. M. Gonçalves, J. Gut, P. J. Rosenthal, R. Moreira and M. M. M. Santos, *MedChemComm*, 2012, **3**, 489.
41. N. Busschaert, I. L. Kirby, S. Young, S. J. Coles, P. N. Horton, M. E. Light and P. A. Gale, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 4426-4430.
42. X. Wu, N. Busschaert, N. J. Wells, Y. B. Jiang and P. A. Gale, *J. Am. Chem. Soc.*, 2015, **137**, 1476-1484.
43. S. J. Edwards, H. Valkenier, N. Busschaert, P. A. Gale and A. P. Davis, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 4592-4596.
44. N. Busschaert, R. B. P. Elmes, D. D. Czech, X. Wu, I. L. Kirby, E. M. Peck, K. D. Hendzel, S. K. Shaw, B. Chan, B. D. Smith, K. A. Jolliffe and P. A. Gale, *Chem. Sci.*, 2014, **5**, 3617-3626.
45. M. Chavent, A. L. Duncan and M. S. Sansom, *Curr. Opin. Struct. Biol.*, 2016, **40**, 8-16.
46. S. M. Loverde, *J. Phys. Chem. Lett.*, 2014, **5**, 1659-1665.
47. W. Khuntawee, P. Wolschann, T. Rungrotmongkol, J. Wong-Ekkabut and S. Hannongbua, *J. Chem. Inf. Model.*, 2015, **55**, 1894-1902.
48. C. J. E. Haynes, S. J. Moore, J. R. Hiscock, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2012, **3**, 1436-1444.
49. P. J. Costa, I. Marques and V. Felix, *Biochim. Biophys. Acta*, 2014, **1838**, 890-901.
50. S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103-117.
51. I. Marques, A. R. Colaco, P. J. Costa, N. Busschaert, P. A. Gale and V. Felix, *Soft Matter*, 2014, **10**, 3608-3621.
52. S. J. Edwards, I. Marques, C. M. Dias, R. A. Tromans, N. R. Lees, V. Felix, H. Valkenier and A. P. Davis, *Chem. Eur. J.*, 2016, **22**, 2004-2011.
53. N. Busschaert, S. H. Park, K. H. Baek, Y. P. Choi, J. Park, E. N. W. Howe, J. R. Hiscock, L. E. Karagiannidis, I. Marques, V. Felix, W. Namkung, J. L. Sessler, P. A. Gale and I. Shin, *Nat Chem*, 2017, **9**, 667-675.
54. Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
55. T. Lu and F. Chen, *J. Mol. Graph. Model.*, 2012, **38**, 314-323.
56. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
57. E. D. Glendening, C. R. Landis and F. Weinhold, *J. Comput. Chem.*, 2013, **34**, 1429-1437.
58. ***NBO 6.0.*** E. D. Glendening, J, K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, C. R. Landis, and F. Weinhold, Theoretical Chemistry Institute, University of Wisconsin, Madison (2013).
59. E. D. Glendening, C. R. Landis and F. Weinhold, *J. Comput. Chem.*, 2013, **34**, 2134-2134.

60. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman (2014), AMBER 14, University of California, San Francisco.
61. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781-1802.
62. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.
63. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
64. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
65. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
66. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
67. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
68. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
69. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
70. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
71. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171-179.
72. C. Jin, M. Zhang, L. Wu, Y. Guan, Y. Pan, J. Jiang, C. Lin and L. Wang, *Chem. Commun.*, 2013, **49**, 2025-2027.
73. G. Ambrosi, M. Formica, V. Fusi, L. Giorgi, A. Guerri, M. Micheloni, P. Paoli, R. Pontellini and P. Rossi, *Chem. Eur. J.*, 2007, **13**, 702-712.
74. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247-260.
75. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
76. L. Martinez, R. Andrade, E. G. Birgin and J. M. Martinez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
77. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
78. P. Li, L. F. Song and K. M. Merz, Jr., *J. Chem. Theory Comput.*, 2015, **11**, 1645-1657.
79. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
80. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
81. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
82. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
83. R. B. Elmes, P. Turner and K. A. Jolliffe, *Org. Lett.*, 2013, **15**, 5638-5641.
84. Y. Li, G. H. Yang, Y. Y. Shen, X. S. Xue, X. Li and J. P. Cheng, *J. Org. Chem.*, 2017, **82**, 8662-8667.
85. C. E. Silva, H. F. Dos Santos, N. L. Speziali, R. Diniz and L. F. de Oliveira, *J. Phys. Chem. A*, 2010, **114**, 10097-10109.
86. C. J. Haynes, N. Busschaert, I. L. Kirby, J. Herniman, M. E. Light, N. J. Wells, I. Marques, V. Felix and P. A. Gale, *Org Biomol Chem*, 2014, **12**, 62-72.
87. Calculator Plugin was used for structure property prediction and calculation, MarvinSketch 16.5.16.0, 2016, ChemAxon Ltd. (http://www.chemaxon.com)
88. D. Laage, T. Elsaesser and J. T. Hynes, *Chem. Rev.*, 2017, **117**, 10694-10725.
89. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011-1021.
90. B. Roux, *Comput. Phys. Commun.*, 1995, **91**, 275-282.
91. A. Grossfield, "WHAM: the weighted histogram analysis method", version 2.0.9.1, http://membrane.urmc.rochester.edu/content/wham
92. B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall and CRC, New York, 1993.
93. H. A. Filipe, M. J. Moreno, T. Rog, I. Vattulainen and L. M. Loura, *J. Phys. Chem. B*, 2014, **118**, 3572-3581.

Chapter V.
How alkyl fluorination modulates the efficacy & mechanism of synthetic anion ionophores: molecular modelling insights

Summary

A molecular modelling study on chloride transmembrane transport promoted by a series of *tren*-based molecules (see Scheme V-1) is presented, comprising quantum and classical force field calculations. As the most positive value on the molecular electrostatic potential surface ($V_{S,max}$) quantum descriptor was found insufficient to rationalise the experimental chloride binding and transport abilities of these synthetic molecules, Molecular Dynamics (MD) simulations in a POPC membrane model were undertaken. The passive diffusion MD simulations showed that, when initially positioned at the core of the bilayer, the tripodal transporters are able to release the encapsulated anion to the water phase after reaching the water/lipid interface. On the other hand, in the water phase, these tripodal molecules are able to form transient chloride complexes and permeate the phospholipid bilayer, either free or complexed. Constrained MD simulations of the diffusion of the **C3x** series and **C5F**, as well as of their chloride complexes, across the phospholipid bilayer, allowed to estimate the Potential of Mean Force (PMF) profiles associated with this traversing movement. The energetic barriers ascertained are straightforwardly related with the lipophilic character (log*P*) of each transporter as well as with the experimentally assessed transport data in lipid vesicles and cell lines. Moreover, MD simulations of the **C3F** and oleate anion (OA) association showed that this carboxylate anion assists the recycling of the free synthetic transporter after the chloride release, allowing the gathering of further insights on the rate-limiting step associated with the transport mechanisms.

The computational studies reported in this chapter, along with the experimental results, are currently submitted for publication. Although these ten tripodal molecules are structurally related with the six *tren*-based tris-thioureas of Chapter II, in this work a more recent and robust all-atom force field was used (Lipid14). In addition, more advanced MD simulations methods were also applied and required computational resources that only became available in the last year of this thesis.

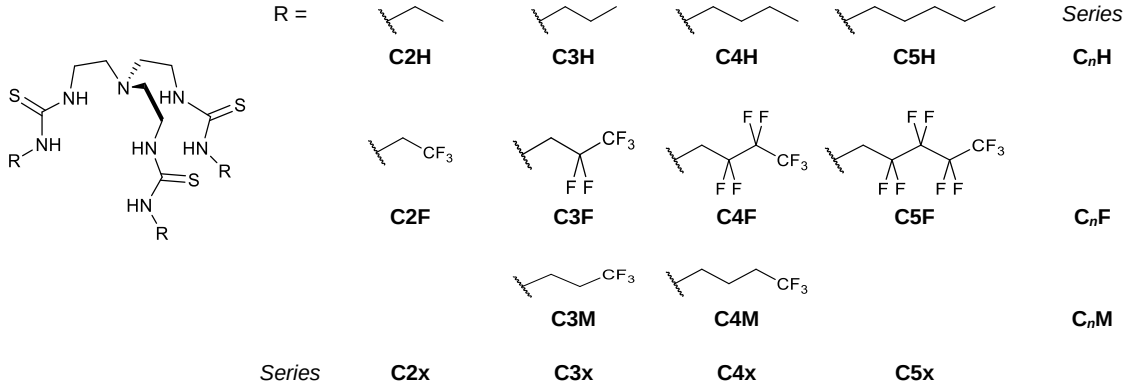
V.1. Introduction

The design and study of small synthetic transmembrane anion transporters have become an important target in supramolecular Chemistry due to the potential biological activity of these compounds, for example in the treatment of Cystic Fibrosis (CF).[1, 2] Numerous classes of anion transporters have now been reported[3] and a number of systematic studies has demonstrated various structure-activity relationships to optimise transporter design.[4-9]

Fluorination of therapeutically active compounds is commonly used in medicinal Chemistry to modulate lipophilicity or acidity.[10] The ability to tune lipophilicity *via* small structural changes allows to establish a well-defined relationship between this property and transport efficacy.[5, 6, 9, 11] Indeed, fluorination of aryl substituents has been shown to enhance the efficacy of anion carriers, due to the increase of lipophilicity and anion binding strength.[12-14] Alkyl fluorination has been reported to increment the Cl^-/NO_3^- exchange efficacy of monoacylglycerol transporters,[15] but there has been little systematic study of the effects of compounds with such substituents, therefore the precise mechanistic effects are still unclear.

Due to the highly polarised nature of the C–F bond,[16] naturally resulting from the high electronegativity of fluorine, compounds with perfluorinated substituents are very non-polarisable.[17] Perfluorinated substituents become interesting structural motifs for the rational design of anion transporters, given the possibility to modulate and control their lipophilicity as well as their van der Waals (vdW) interactions with the phospholipid tails, as further demonstrated in this chapter.

The series of molecules in Scheme V-1 has been experimentally studied as chloride receptors as well as anion transporters.[18] The **C*n*H**, **C*n*F** and **C*n*M** series, with $n = 2 - 5$, allow the study of the effect of extending the length of the alkyl chain. As for the **C2x**, **C3x**, **C4x** and **C5x** series, x = H, F or M enables the evaluation of the effect of the degree of fluorination. The experimental results necessary for the understanding of this theoretical investigation are summarised below.



Scheme V-1. Series of *tren*-based molecules with varying degrees of fluorination and alkyl chain lengths studied in this work.

Table V-1 lists the retention factor experimentally assessed (expressed as log(*k*')), obtained from retention time data measured in a octadecyl carbon chain (C18)-bonded silica reverse phase HPLC column after isocratic elution with 30% CH_3CN/H_2O of the free transporters.[18] In these conditions, log(*k*') is proportional to log*P*.[19] Thus, the perfluorination of the alkyl chains increases the lipophilicity, as clearly illustrated in compounds of the **C3x** series with the same chain length, with log(*k*') following the trend **C3H** (−0.87) < **C3M** (−0.28) < **C3F** (0.27). In contrast, in spite of the potential to enhance the acidity of the thiourea binding units with fluorination of alkyl substituents (*vide infra*), the chloride

association constants (K_a) determined in wet dimethyl sulfoxide (DMSO), also listed in Table V-1, appear to be unaffected by this structural modification.[18]

Table V-1. Selected data for the tris-thiourea series present in Scheme V-1, including log(*k*') and chloride binding and transport properties.[a]

Compound	log (*k*')	Cl^- K_a (M^{-1})	k_{ini} in vesicles (s^{-1})			k_{ini} in cells (s^{-1})
			Only T[b]	T and Gr[c]	T and OA[d]	
C2H	−1.82	557	0.0012	0.0016	0.0013	−0.0004
C3H	−0.87	677	0.0014	0.0065	0.0032	0.0006
C4H	−0.14	596	0.0014	0.0139	0.0065	0.0035
C5H	0.45	648	0.0016	0.0217	0.0073	0.0024
C2F	−0.44	470	0.0357	0.0359	0.0812	0.0174
C3F	0.27	565	0.2030	0.1559	0.2314	0.0308
C4F	0.92	733	0.0455	0.0412	0.0898	0.0228
C5F	1.53	669	0.0028	0.0034	0.0053	0.0013
C3M	−0.28	466	0.0172	0.0145	0.1041	0.0154
C4M	−0.03	575	0.0080	0.0102	0.0322	0.0138

[a] Table adapted from ref. 18; [b] 0.01 mol% transporter; [c] 0.01 mol% transporter + 0.5 mol% gramicidin; and [d] 0.01 mol% transporter + 2 mol% oleate anion.

In stark contrast, the fluorination of the thiourea tripods has a significant impact on the anion transport ability of these small molecules, as shown by the transport initial rate (k_{ini}) assessed both in lipid vesicles and cell lines and also gathered in Table V-1.[18] In summary, the chloride transport ability of the *tren* compounds was measured in a pH-driven vesicle-based assay. At the beginning of the transport activity assessment, the intra-vesicular pH was 7.0, while the external pH was 8.0, with a small chloride imbalance ($[Cl^-]_{internal} > [Cl^-]_{external}$), with the transporters being added externally. In these experimental conditions these *tren*-based molecules are able to dissipate the pH gradient by either an electroneutral H^+/Cl^- symport, Cl^-/OH^- antiport, or a combination of both processes. In lipid vesicles, the pure alkyl compounds of the **C*n*H** series are less active than their **C*n*F** perfluorinated analogues, with the anion transport activity increasing along with the chain length, while the compounds of the **C*n*M** series, with only a terminal $-CF_3$, present intermediate activity.[18]

Given that other structurally related tripodal thioureas exhibit Cl^-/OH^- selectivity,[20] then all tripodal molecules sketched in Scheme V-1 were further experimentally screened along with two protonophores coupled to the chloride transport process: gramicidin (Gr), a bacterial proton channel which allows fast efflux of protons; or OA, which is known to promote proton efflux through a flip-flop mechanism facilitated by anion ionophores.[21] If the compounds of Scheme V-1 are rate-limited by their H^+, or functionally equivalent OH^-, transport, they should show an enhancement in the presence of Gr or OA, with the electrogenic Cl^- transport coupled to the H^+ efflux.[18] The values of the initial transport rates in these conditions are also reported in Table V-1.

Indeed, the transport activity of the **C*n*H** series is greatly enhanced in the presence of Gr and, to a lesser extent, in the presence of OA, indicating that the alkyl substituted molecules are selective for Cl^- over H^+/OH^- and able to facilitate a degree of electrogenic Cl^- transport. In contrast, while no enhancement in the presence of Gr is observed for the fluorinated **C*n*F** series and only slight increments are observed for **C*n*M**, the presence of OA leads to a transport activity enhancement of all fluorinated tripodal thioureas.[18] Therefore, the distinct enhancement of overall Cl^- transport mediated by fluorinated compounds in the presence of OA but not Gr suggests that the rate-limiting step of the transport process is different between the alkyl **C*n*H** series and the fluorinated analogues **C*n*F** and **C*n*M**. Two plausible mechanisms for the overall chloride transport process are depicted in Figure V-1.[18]

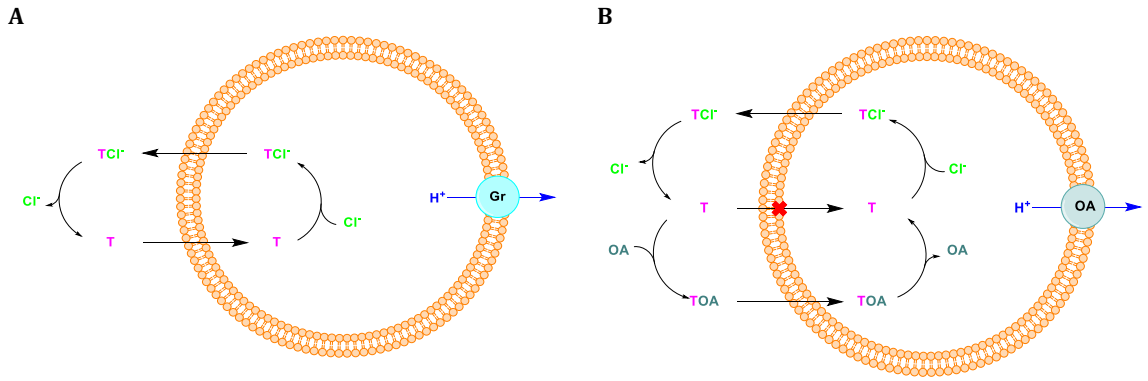


Figure V-1. Rationalisation of chloride transport coupling to Gr and OA protonophores. T = tripodal transporter. **A**: Anion transport enhanced by proton efflux of Gr. **B**: Enhancement of fluorinated transporters by OA, which allows the back diffusion of T into the vesicle. Figure adapted from ref. 18.

In the Gr-coupled mechanism (see Figure V-1A), the anion ionophore must perform four sequential steps to complete the electrogenic chloride transport process: (1) chloride uptake inside the vesicle; (2) complex translocation across the membrane; (3) chloride release outside the vesicle; (4) back-diffusion of the free transporter into the vesicle. If the H^+ or OH^- transport process facilitated by the anion ionophore is rate-limiting, then the presence of Gr will facilitate H^+ efflux and enhance the overall chloride transport, as observed for the **C*n*H** series (see Table V-1). On the other hand, no enhancement of the transport activity indicates that the whole electrogenic chloride transport process is rate-limiting, as found for the fluorinated **C*n*F** and **C*n*M** series.[18]

The presence of OA (see Figure V-1B) opens new pathways to complete recycling of the anion transporter. The anion binding of the deprotonated OA to the tripodal molecules facilitates their flip-flop motion within the highly packed phospholipid medium, thus avoiding the back-diffusion of the free transporter and concomitantly increasing the H^+ efflux. In agreement with this proposed mechanism, the overall chloride transport is enhanced for the **C*n*H** series (see Table V-1), but this enhancement is smaller when compared with the mechanism with Gr. However, one must keep in mind that the H^+ efflux process facilitated by OA (ferried with the tripodal molecule) is expected to be slower than with Gr (a H^+ transmembrane channel). On the other hand, OA enhances the transport activity of the **C*n*F** and **C*n*M** series, contrary to Gr. Given that the two mechanisms differ in the back-diffusion of the free transporter, this seems to be the rate-limiting step for the fluorinated thiourea tripodal transporters when coupled with Gr (see Figure V-1A).[18]

The chloride transport activity facilitated by the *tren* compounds was also assessed in cell membranes, using Fisher rat thyroid (FRT) epithelial cells expressing the halide-sensitive yellow fluorescent protein YFP-H148Q/I152L.[22, 23] It is important to note that the FRT cells do not have chloride channels, ruling out the possibility of this anion efflux without the assistance of external anion ionophores, such as the small anion carriers.[22, 23] Overall, the transport data obtained in cell lines follow the same trend of the data obtained in the pH-driven vesicle assays (see Table V-1).[18]

This plethora of experimental anion transport data was the driving force for comprehensive molecular modelling studies undertaken and herein reported in detail. Structural and energetic insights into the interaction of the **C*n*H**, **C*n*F** and **C*n*M** series of transporters with the POPC membrane model were obtained as well as into the proposed anion transport mechanism (see Figure V-1).

V.2. Methods

The quantum calculations were carried out with *Gaussian 09*,[24] using different theory levels, as further detailed. These calculations include the geometry optimisations of OA, the free tripodal derivatives and their chloride complexes and calculation of the distribution of electrostatic potential ($V(\mathbf{r})$). The $V(\mathbf{r})$ on the molecular surface of the tripodal compounds was computed from previously optimised structures of their chloride complexes, after removal of the anion, using the same level of theory. Subsequently, $V(\mathbf{r})$ was evaluated on the 0.001 electrons Bohr^{-3} contour of $\rho(\mathbf{r})$, and is henceforth labelled $V_S(\mathbf{r})$. The electrostatic potential surface ranges, including the most negative and most positive values ($V_{S,min}$ and $V_{S,max}$, respectively), were ascertained using *Multiwfn*.[25, 26]

The theoretical investigation on membrane systems comprised passive diffusion MD simulations, Steered Molecular Dynamics (SMD) and Umbrella Sampling (US) simulations. All MD simulations were carried out with AMBER 2016,[27] with resort to GPU acceleration.[28-30] The Lipid14[31] force field was employed for the POPC lipids, while parameters from the General AMBER Force Field (GAFF)[32, 33] and atomic RESP charges[34] were used for the tripodal molecules and OA, as follows. The initial structures of the free tripodal molecules were generated by atomic manipulation of an unpublished crystal structure of **C4H**·Cl$^-$.[35] Subsequently, these structures were optimised at the HF/6-31G* level with *Gaussian 09*.[24] The optimisation of these ten individual structures was followed by a single point calculation to generate the electrostatic potential (ESP) at the same theory level, using the Merz–Singh–Kollman scheme with 4 concentric layers per atom and 6 density points in each layer (IOp (6/33 = 2, 6/41 = 4, 6/42 = 6)). The initial atomic charges of each molecule were then calculated by RESP fitting, along with the assignment of GAFF atom types, using the *antechamber* module,[36] as implemented in the AMBER package.

To obtain atomic charges less dependent of the molecular conformation or orientation, the calculation of the final RESP charges employed in all membrane MD simulations was preceded by conformational analyses on the tripodal compounds, using the following protocol: The initial Molecular Mechanics (MM) energy minimised structures, using the initial RESP charges, were heated up to 500 K in the gas-phase for 50 ps, followed by collection runs of 0.5 ns, and using a time step of 1.0 fs, allowing the stochastic search of the conformational space. Frames were saved every 0.1 ps leading to trajectory files containing 5000 structures. All these structures were further minimised by MM using a steepest descendent gradient followed by the conjugate gradient algorithm, until the convergence criterion of 0.0001 kcal mol^{-1} Å^{-1} was attained. Afterwards, the MM minimised structures were clustered with the *UCSF Chimera*.[37] From the resulting clusters, three representative conformations with substantially different RMSD values were selected for each tripodal molecule. All selected conformations had the N–H binding units of the thiourea moieties adopting a *syn* configuration, but different spatial dispositions of the thiourea appended chains. All these structures were further HF/6-31G* geometry optimised, followed by their ESP calculations as described above. Then, the individual ESP data were extracted from the corresponding *Gaussian 09*[24] outputs, concatenated, and subsequently used to generate the input files for the two-stage RESP fitting, using identical weights for all conformations. On the other hand, for OA, eight structures were generated from the POPC lipids taken from an equilibrated bilayer. Afterwards, these randomly selected structures were optimised at the HF/6-31G* level followed by ESP calculations. As previously described, the ESP data was used to generate the final multi-RESP atomic charges.

The structures of the tripodal chloride complexes, also generated from the X-ray crystal structure of **C4H**·Cl$^-$, were DFT optimised at the M06-2X/6-31+G** level of theory. The quantum optimised structures of the tripodal chloride complexes were used in further single point quantum calculations as well as the starting binding arrangements in MD simulations in POPC bilayer model. The

structure of **C3F**·OA was generated with the carboxylate group positioned inside the tripodal binding pocket. This initial geometry was optimised *via* DFT at the same level of theory, and the final structure was used in the ensuing MD simulations.

The chloride or OA complexes were inserted in a free POPC bilayer with *Packmol*,[38] at its core (scenario A) or in its water slab (scenario B). This previously equilibrated membrane system is a POPC membrane model, with an orthorhombic shape composed of 128 phospholipids, 6500 TIP3P model water molecules,[39] and 17 Cl^- and 18 Na^+ solvated ions (*ca.* 0.15 M). The monoatomic ions were described with van der Waals parameters developed to be used along with the TIP3P water model.[40]

The MD simulations with the chloride complexes were carried out as follows: the initial configuration of each system was submitted to 20000 steps of MM energy minimisation with a 500 kcal mol^{-1} $Å^{-2}$ positional restraint on the chloride complex and lipid molecules, through the steepest descent algorithm for 10000 steps plus 10000 steps of the conjugated gradient algorithm. Subsequently, the restraints were removed and the entire system was allowed to relax for another 20000 steps, using the same algorithms. The equilibration of the system proceeded by heating it to 303 K in an NVT ensemble for 100 ps with a 10 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex and lipid molecules. The equilibration stage proceeded with a 5.0 ns run using an NPT ensemble with a 5.0 kcal mol^{-1} $Å^{-2}$ restraint on the chloride complex. Then, the positional restraint was removed and the simulation continued for further 300 ns (scenario A, a single MD run) or 200 ns (scenario B, two independent MD runs). The long-range electrostatic interactions were described with the Particle Mesh Ewald (PME) algorithm[41] using a real-space cut-off at 10 Å. The cut-off for the Lennard–Jones interactions was also set at 10 Å. The temperature of the system was maintained at 303 K, using Langevin dynamics,[42] with a collision frequency γ of 1.0 ps^{-1}. The pressure was controlled by the Berendsen barostat[43] at 1 atm and compressibility of 44.6×10^{-6} bar^{-1}, with a relaxation time of 1.0 ps. The covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm,[44] allowing the use of a 2.0 fs time step. The MD simulation trajectory frames were saved every 10 ps. The **C3F**·OA complex was simulated using the same multi-stage equilibration protocol, but with shorter production runs of 100 ns in both scenarios. The post processing and analysis of the MD simulation trajectory files were performed with *cpptraj*.[45]

The starting configurations necessary for the US simulations of free **C3H**, **C3F**, **C3M** and **C5F** were generated from previous SMD simulations, as follows: the complexed chloride in scenario A was relocated to the water phase and then the system was minimised, heated, and equilibrated as described above. Afterwards, the final frame of the equilibration period was used as the starting frame for the SMD. Subsequently, the free receptors were dragged from the bilayer core, along the membrane normal (*viz*, *z*-dimension, the reaction coordinate) at 2.5 Å ns^{-1} in MD simulations of 17 ns, in the NPT ensemble. Throughout the travelled 42.5 Å along the membrane system, a force constant of 5.0 kcal mol^{-1} $Å^{-2}$ was applied to the non–hydrogen atoms of the tripodal derivatives. A similar procedure was also applied to **C3F** in scenario B: the complexed anion was relocated to the water phase and the system was minimised, heated and equilibrated as above, prior to the SMD simulation where the tripodal molecule was dragged along the bilayer system for 170 ns at 0.5 Å ns^{-1}, with a force constant of 5.0 kcal mol^{-1} $Å^{-2}$ as well.

While the starting configurations for the US simulations regarding the free tripodal molecules were selected from the SMD simulations, the starting geometries for the tripodal complexes (both chloride and OA) were gathered from the passive diffusion MD simulations carried out in scenarios A and B. The spacing between the centre of mass (COM) of the tripodal molecules in the US simulations is *ca.* 1 Å, with a total of 31 evenly spaced independent starting points (ranging from 0 to +30 Å along the *z*-dimension). A distance restraint of 5.0 kcal mol^{-1} $Å^{-2}$ was applied along the *z*-dimension, between the COM of the non-hydrogen atoms of the tripodal receptor and the COM defined by phosphorus atom of the lipid head groups and the terminal CH_3 carbon atoms. Each one of the 31 windows was simulated for 80 ns, except for the chloride complexes of **C3F** and **C5F**, which were run for 100 ns, and the US

simulations of **C3F**·OA, which were 130 ns long. Only the last 50 ns were considered as sampling time, with the initial MD simulation period being discarded as equilibration period. In the US simulations of the tripodal complexes, six 5.0 kcal mol^{-1} Å^{-2} harmonic distance restraints between the nitrogen atoms and the chloride ion or carboxylate carbon atom were applied. Each independent US window underwent a minimisation and heating stage, with the same distance restraints of the production run. The remaining details are as given for the passive diffusion simulations, apart of the distance restraints. The trajectory frames were saved every 10 ps, while the distance between the COM of the tripodal derivative and the COM of the POPC bilayer, defined by the phosphorus atoms of the lipid head groups and the terminal CH_3 carbon atoms, was saved every 50 steps.

The US simulations previously carried out for the free chloride were re-visited in the scope of this work, considering only the 31 windows from z = 0 to 30 Å (see Chapter IV, section IV.2.2.).

V.3. Results and Discussion

V.3.1. DFT calculations

The chloride complexes of the ten tripodal transporters were subjected to DFT optimisation at the M06-2X/6-31+G** theory level. After obtaining the optimised complexes, shown in Figure V-2 for the **C3x** series, the anion was removed and the $V(\mathbf{r})$ was computed on the molecular surface ($V_S(\mathbf{r})$), to determine the position and magnitude of the electrostatic potential maximum ($V_{S,max}$), intimately related with the anion recognition ability of anion receptors.[6] The $V_{S,max}$ values for the ten molecules are gathered in Table V-2, and the distribution of the electrostatic potential for the **C3x** series is also shown in Figure V-2.

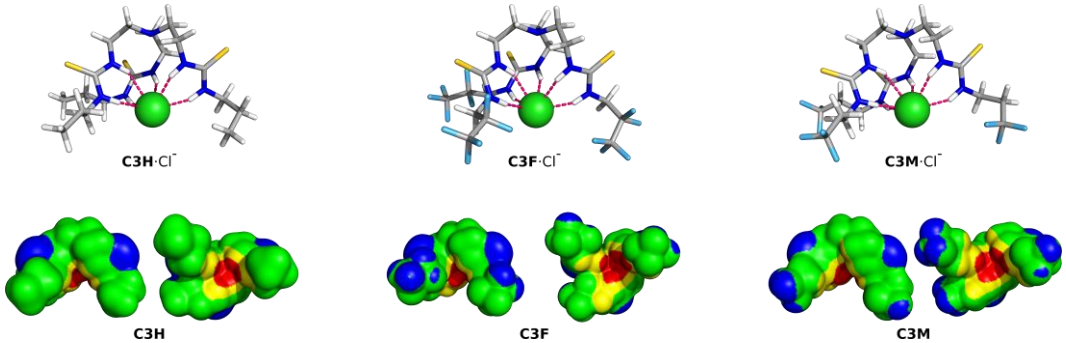


Figure V-2. DFT optimised structures of complexes **C3H**·Cl$^-$, **C3F**·Cl$^-$, and **C3M**·Cl$^-$ (top) and $V(\mathbf{r})$ mapped on the 0.001 electrons Bohr^{-3} isodensity surface of **C3H**, **C3F**, and **C3M** in side and bottom views (bottom). The hydrogen, carbon, sulfur, nitrogen, fluorine atoms and chloride ion are shown in white, grey, yellow, blue and cyan sticks, respectively, while the anion is shown as a green sphere. The N–H···Cl$^-$ hydrogen bonds are draw as pink dashes. The colour scales range from blue to red, in kcal mol^{-1}, as follows: blue – lower than −5.0; green – between −5.0 and 30.0; yellow – between 30.0 and 65.0; and red – greater than 65.0.

Table V-2. $V_{S,max}$ (kcal mol^{-1}) values ascertained for the tripodal compounds in their chloride complexes.

Compound	**C2H**	**C3H**	**C4H**	**C5H**	**C2F**	**C3F**	**C4F**	**C5F**	**C3M**	**C4M**
$V_{S,max}$	82.03	81.72	81.44	81.31	92.27	93.03	92.76	92.86	90.36	91.40

As the optimised compounds have the three thiourea binding units synergistically recognising the chloride ion, the natural location of the $V_{S,max}$ is within the binding pocket of the tripodal ligand, as evident from the areas of higher electrostatic potential, depicted in red in Figure V-2. Along the four series, the fluorination of the **C$_n$H** series consistently leads to the increase of the $V_{S,max}$ values. For

instance, the non-fluorinated **C3H** has a $V_{S,max}$ of 81.72 kcal mol^{-1}, while its fluorinated analogous **C3M** and **C3F**, with $-CH_2-CH_2-CF_3$ and $-CH_2-CF_2-CF_3$ chains, have values of 90.36 and 93.03 kcal mol^{-1}, respectively. On the other hand, independently of the chain length, the $V_{S,max}$ values are similar along the non-fluorinated **C$_n$H** series, as well as in **C$_n$M**, which present only a terminal $-CF_3$ group. In addition, molecules of the **C$_n$F** series, which contain a single methylene group between the thiourea binding units and the terminal fluorinated moieties, present similarly higher $V_{S,max}$ values. In other words, these molecules should have comparable high binding affinities for chloride (K_a). Indeed, molecule **C4F**, has the highest anion binding constant (see Table V-1). In stark contrast, molecule **C3F** only presents a moderate anion association constant, albeit it has the highest $V_{S,max}$ value. These puzzling results indicate that other factors, such as solvent effects and/or conformational changes, play an important role in the anion recognition in DMSO solution. Moreover, it is noteworthy that **C3F**, with the most acidic thiourea binding units presents the best chloride transmembrane transport activity. However, the $V_{S,max}$ and the transport data are not straightforwardly related, prompting us to further undertake a comprehensive MD investigation with POPC membrane models.

V.3.2. Passive diffusion MD simulations

The passive diffusion of the tripodal chloride complexes in the POPC bilayer was evaluated considering two distinct initial scenarios: the complexes were positioned with a random orientation in the bilayer core (A) or in the water phase (B). Afterwards, the unrestrained complexes were permitted to freely diffuse for 300 ns along a single MD run in scenario A, and for 200 ns in two independent runs in scenario B. The distance between the tripodal nitrogen (N_{tren}) atom and the closest water/lipid interface (P_{int}, defined by the COM of the 64 phosphorus atoms in that monolayer) was evaluated along the simulation time, together with the distance between P_{int} and the COM defined by the terminal carbon atoms of the three thiourea chains (C_{ter}). These distances are plotted in Figure V-3 for all simulated systems of scenario A, together with the counting of the N–H hydrogen bonding interactions with chloride, phosphate head groups or water molecules, assessed using cut-offs of 3.5 Å and 120°.†††

††† The distance and angle cut-offs were selected in agreement with the N–H···Cl$^-$ dimensions found in single crystal X-ray structures deposited with CCDC[46, 47] of similar tris-thiourea chloride complexes.[13, 20, 48, 49]

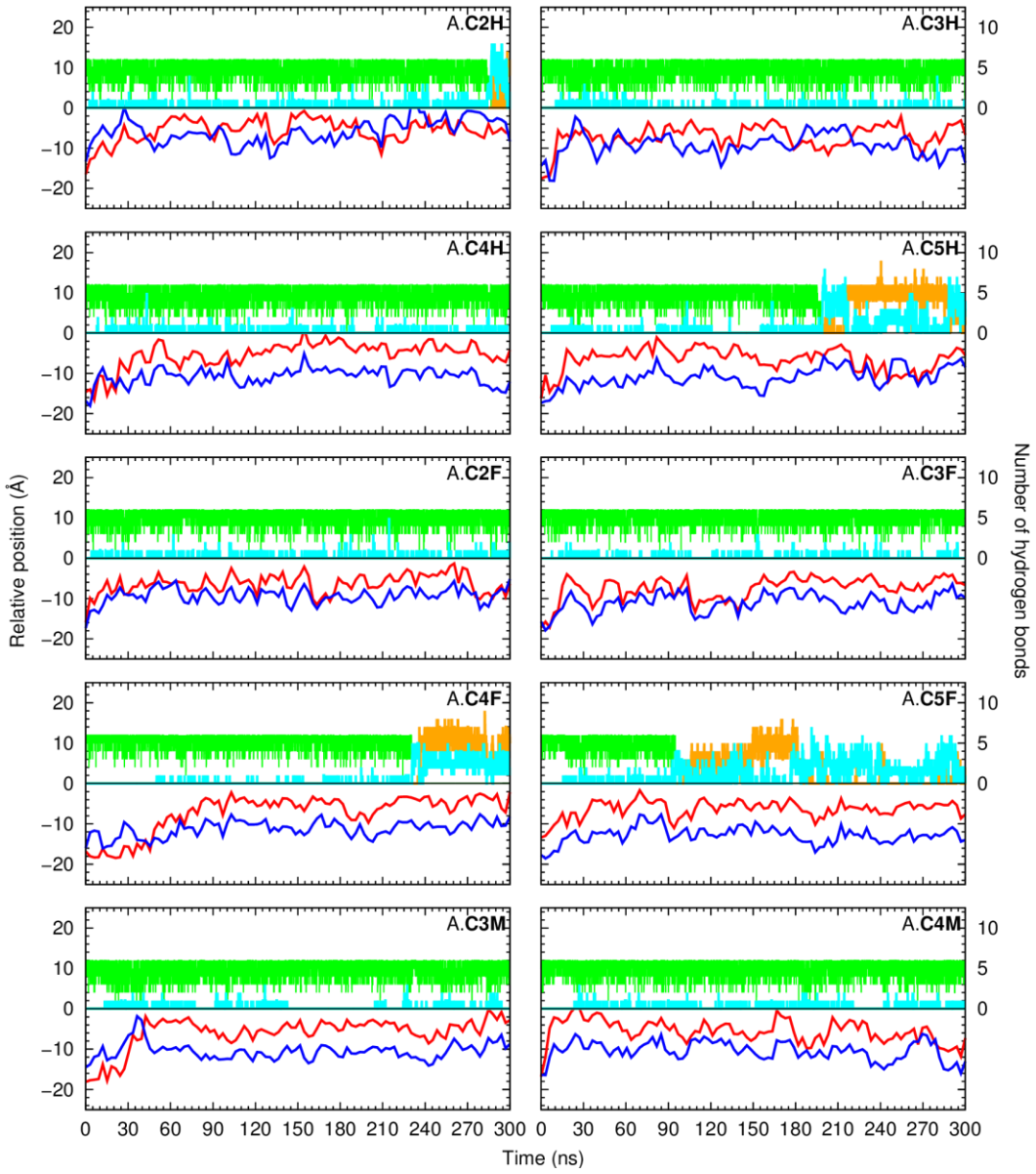


Figure V-3. Evolution of the $N_{tren}\cdots P_{int}$ (red line) and $C_{ter}\cdots P_{int}$ (blue line) distances throughout 300 ns of the MD runs of the tripodal chloride complexes in scenario A. Data were smoothed with Bézier curves. The evolution of the number of N–H···Cl^- hydrogen bonds is also plotted as a green line, while the hydrogen bonds with the water molecules or phosphate head groups are shown in cyan and orange, respectively. The water/lipid interface is represented as a black line at z = 0 Å. Each MD simulation is identified with the initial scenario and tripodal molecule studied.

Overall, in the MD simulations in scenario A, the anion tripodal complexes diffuse towards the water/lipid interface, with the chloride ion sheltered from the water molecules by the tripodal conformational shape. Moreover, throughout the MD simulation time, transporters **C2H**, **C4F**, **C5H** and **C5F** end up releasing the anion to the water phase (green line), concomitantly binding to water molecules (cyan line) or phosphate head groups (orange line). Moreover, it is evident that with the growing fluorination degrees and thiourea chains length, the molecules adopt well-defined orientations within the highly packed phospholipid medium. This is particularly evident in the MD simulations with **C4H**, **C5F**, **C3F**, **C4F, C5F**, **C3M** and **C4M**, where the *tren* moiety is closer to the water phase and the chains are nearly aligned with the phospholipid tails. In contrast, **C2H**, **C3F** and **C2F**, with the shorter

chains and lower fluorination degree, adopt random orientations along their MD runs. These features are clearly illustrated in Figure V-4, with selected snapshots for the MD runs of **C3H** and **C3F** (top and bottom, respectively). Furthermore, the non-fluorinated **C5H**, upon release of the anion adopts an undefined orientation, possibly a result of the interaction with phosphate head groups and water molecules at the interface. However, the anion release which was also observed in the 300 ns MD runs with **C2H**, **C4F** and **C5F** did not lead to the reorientation of the transporters.

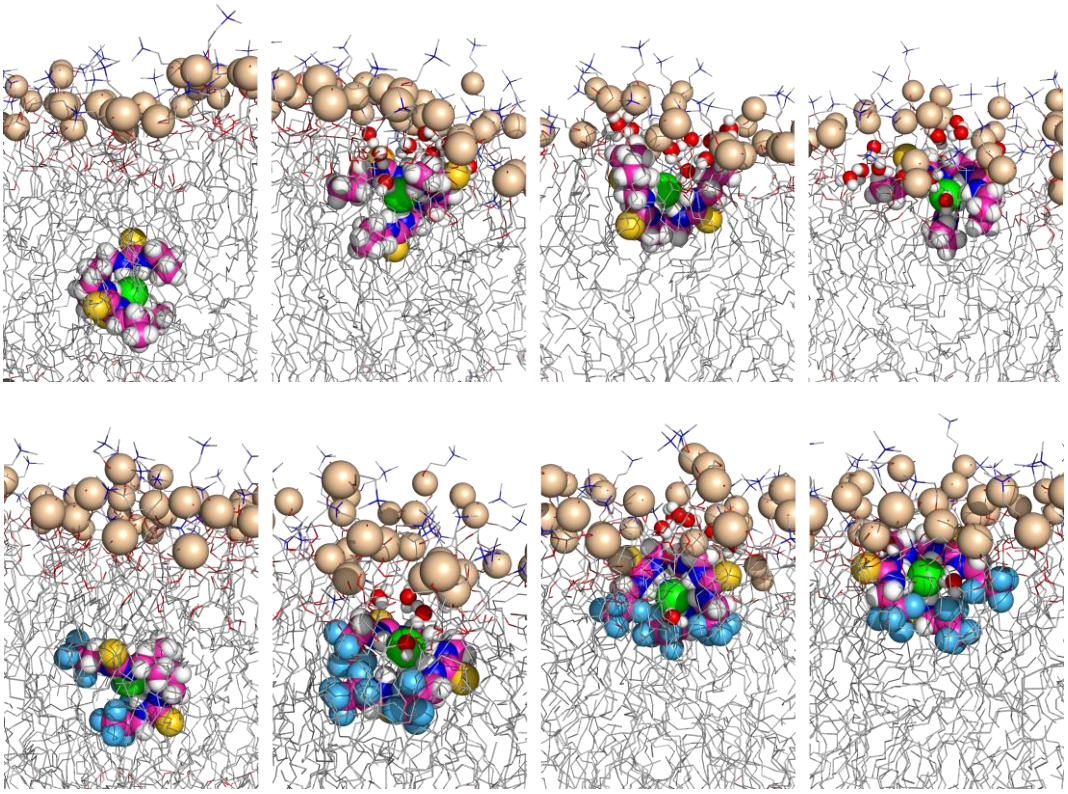


Figure V-4. Consecutive snapshots depicting the diffusion of the chloride complex of **C3H** (top) and **C3F** (bottom) throughout the MD simulation in scenario A. The tripodal derivative, the complexed chloride together with the phosphorus atoms are represented in spheres. The remaining atoms of the bilayer are shown as lines. The hydrogen atoms are shown in white, oxygen atoms in red, sulfur atoms in yellow, nitrogen atoms in blue, fluorine atoms in light blue, phosphorus atoms in wheat, chloride in green and carbon atoms in pink (**C3H** and **C3F**) or grey (phospholipids). The interaction of the internalised tripodal transporters with the solvent molecules is emphasised with the depiction of water molecules within 3.5 Å from them as spheres. The lipids' C–H bonds, most water molecules, and the solvated chloride and sodium ions were omitted for clarity.

As illustrated in Figure V-5 and Figure V-6, in the MD simulations in scenario B, all tripodal molecules are able to permeate the membrane, as free transporters or associated with a chloride, staying below the water/lipid interface. Moreover, the transporters can uptake the anion, in the aqueous phase or at interface, and subsequently enter the POPC bilayer with it, as observed in the first MD run of **C2F**. The transporters with the higher fluorination degree, *i.e.*, with higher $V_{S,max}$ values (see Table V-2), establish stable anion associations throughout most of the MD simulation time, as evident in both MD runs of **C3F**·Cl^-, which tripodal host presents a $V_{S,max}$ value of 93.03 kcal mol^{-1}. The spatial disposition of each tripodal molecule below the interface is equivalent to the one observed when the complex is initially positioned in the core of the lipid bilayer (scenario A).

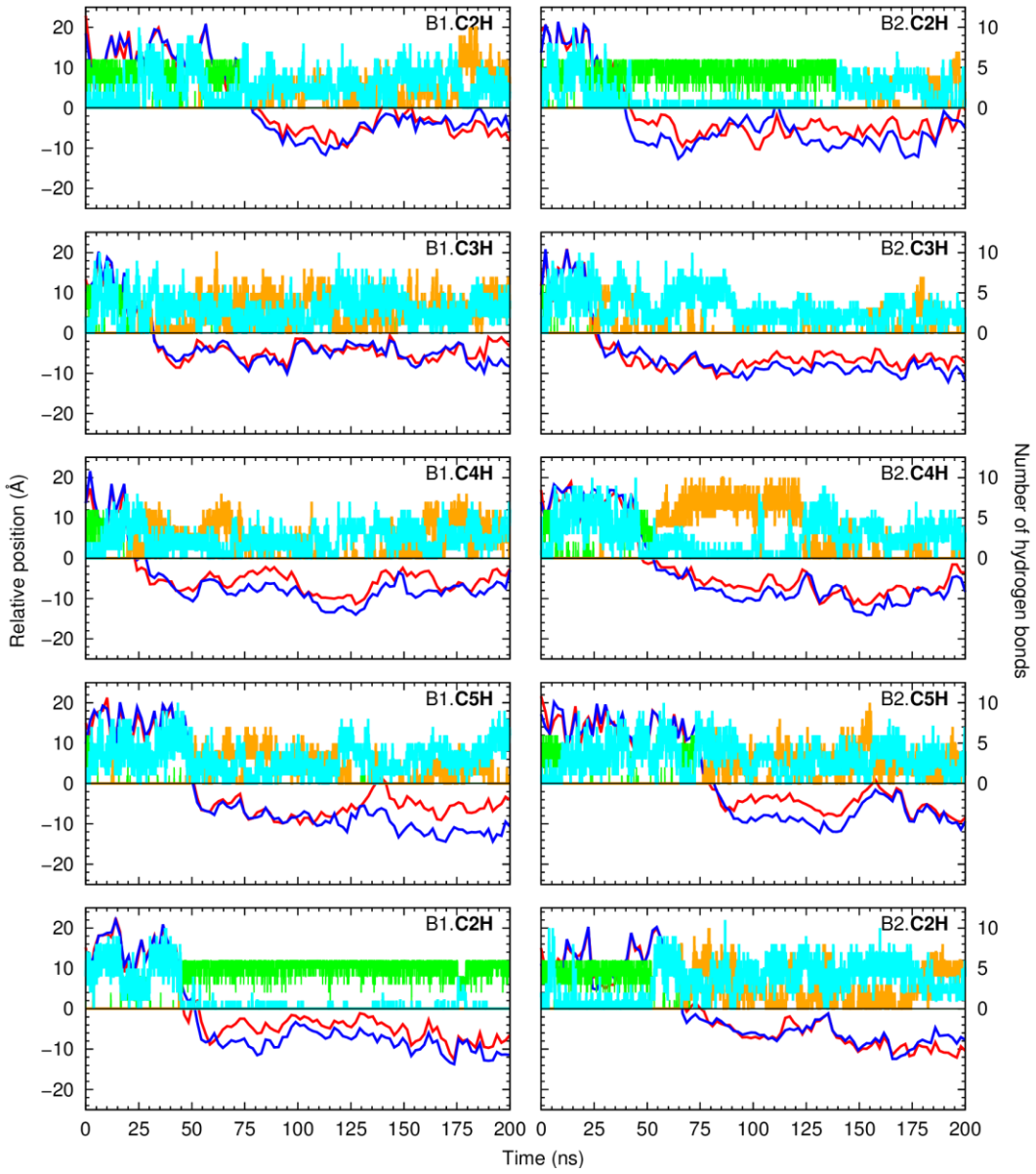


Figure V-5. Evolution of the $N_{tren}\cdots P_{int}$ (red line) and $C_{ter}\cdots P_{int}$ (blue line) distances throughout 200 ns of the MD runs of chloride complexes of **C2H**, **C3H**, **C4H**, **C5H** and **C2F** in scenario B. Each MD simulation is identified with the initial scenario, MD run number and tripodal molecule studied. Remaining details as given in Figure V-3.

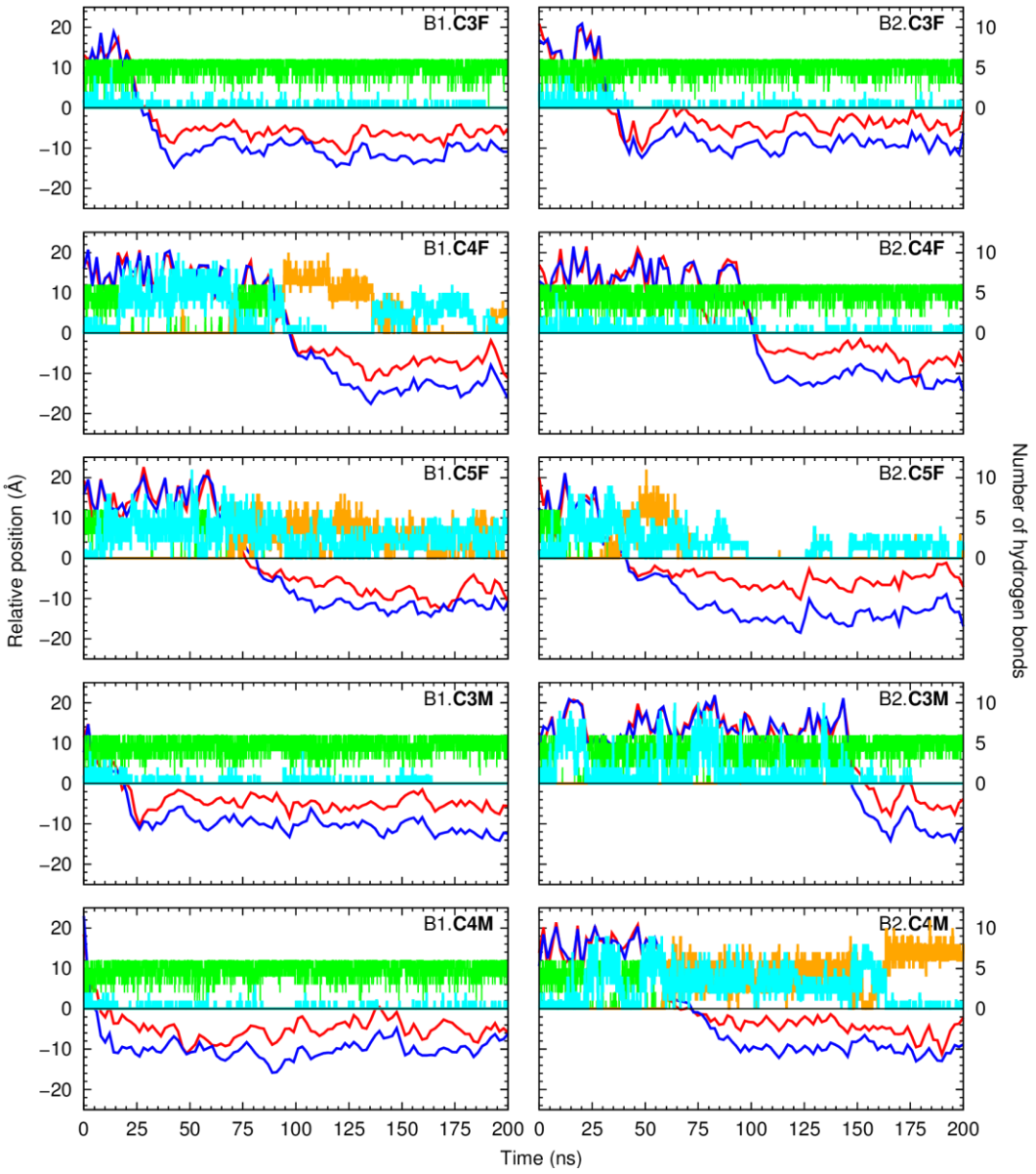


Figure V-6. Evolution of the $N_{tren}\cdots P_{int}$ (red line) and $C_{ter}\cdots P_{int}$ (blue line) distances throughout 200 ns of the MD runs of the chloride complexes of **C3F**, **C4F**, **C5F**, **C3M** and **C4M** in scenario B. Each MD simulation is identified with the initial scenario, MD run number and tripodal molecule studied. Remaining details as given in Figure V-3.

V.3.3. Constrained MD simulations

Further structural insights into the anion transport ability of these tripodal molecules were obtained through US simulations. Furthermore, the free energy profiles associated with the transmembrane translocation of the free tripodal molecules and their complexes were also estimated from the US simulations computing the PMF. These calculations were preceded by preliminary SMD simulations of free **C3F** being dragged from one aqueous phase, across the POPC bilayer, to the other water slab in a 170 ns long simulation. A force constant of 5.0 kcal mol^{-1} Å^{-2} was applied to the receptor along the membrane normal (the *z*-dimension), allowing it to move through the membrane model with a constant velocity of 0.5 Å ns^{-1}. Noteworthy, during this slow movement, **C3F** was able to uptake a chloride from the water phase, permeate the POPC bilayer (as observed in the passive MD simulations

in scenario B), traverse it and, subsequently, release the chloride ion in the opposite water phase. This sequence of events, illustrated in Figure V-7, proves that this tripodal fluorinated molecule is able to promote the chloride transmembrane transport in agreement with the anion carrier mechanism. To the best of our knowledge, this is the first time that this sequence of events is observed at the atomistic level for a neutral transporter. We have previously observed comparable mechanistic findings, but for a highly charged calix[4]arene.[50] In spite of the slow velocity, the permeation of the transient bulky **C3F**·Cl^- in the POPC bilayer is accompanied by a depression of the entry leaflet and a large number of solvating water molecules (see Figure V-8). Still, the slow diffusion velocity allowed the tilting and tumbling of **C3F**, as it adopted the same preferential orientation near of both water/lipid interfaces, as observed in the passive MD simulations, as detailed above.

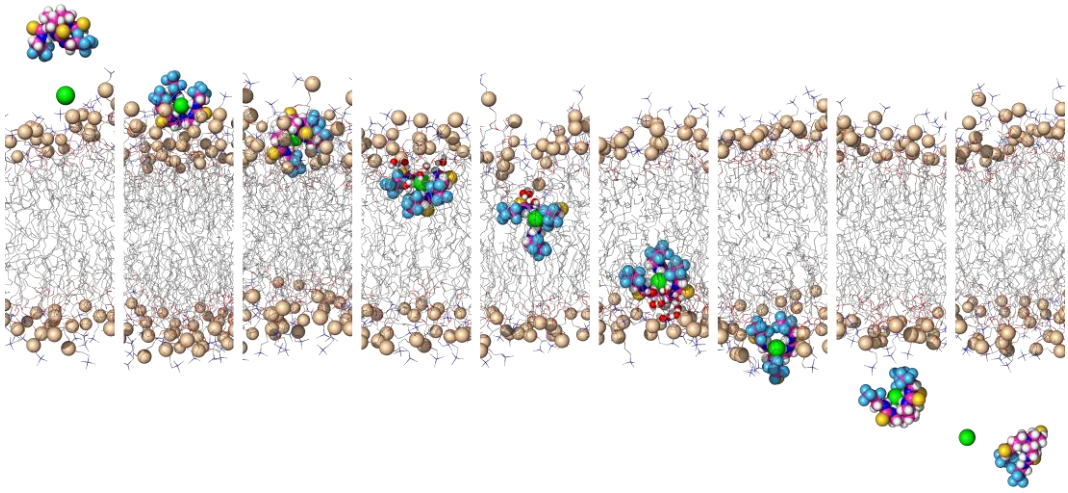


Figure V-7. Consecutive snapshots depictive the movement of **C3F** across the POPC bilayer in a 170 ns long SMD simulation. The free tripodal transporter has uptaken a chloride before permeating the bilayer and dragged it to the opposite side of the membrane model, where they dissociated. Remaining details as given in Figure V-4.

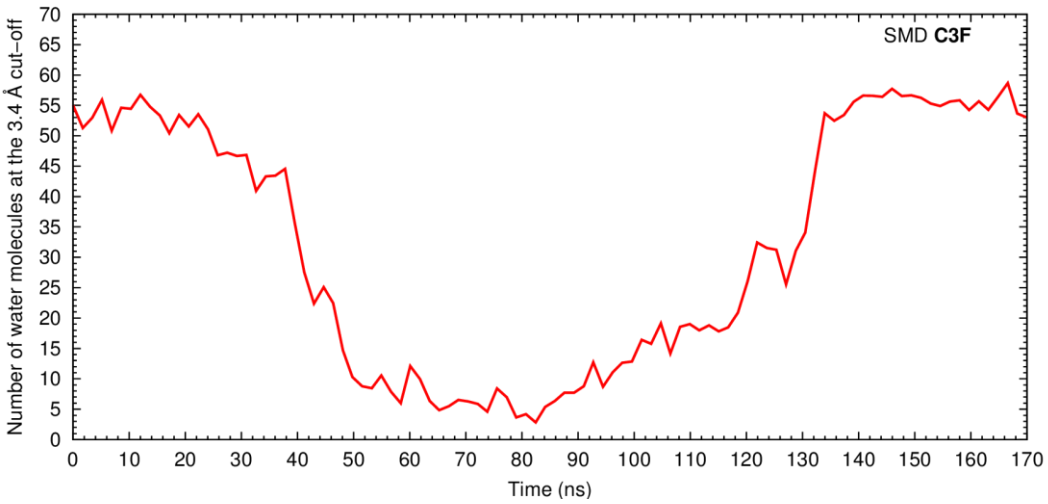


Figure V-8. Variation of the number of water molecules within the solvation shell defined by a cut-off of 3.4 Å‡‡‡ from **C3F**, in the SMD simulation from one side of the POPC to the opposite side. Data were smoothed using Bézier curves.

‡‡‡ Approximate radius for the first hydration shell of biomolecules.[51]

However, the structural defects caused by the entrance of the complex in the POPC bilayer have precluded the use of this type of SMD simulations to generate the starting geometries for the subsequent US simulations. Indeed, these residual effects from the SMD would translate to increased PMF errors that could only be diminished with longer equilibration times of the initial states, namely for the translocation of charged molecules.[52] Thus, in these circumstances, we decided to use the passive diffusion MD simulations, in both scenarios, to generate the starting structures for the further US simulations, as the passively diffusing tripodal chloride complexes have negligible effects in the membrane structure. On the other hand, the starting structures for the US simulations of the free tripodal molecules were generated from SMD simulations with the free transporter initially positioned in the bilayer core, given that its exit to the water phase had a low structural effect on the POPC bilayer, avoiding the presence of an eventual chloride complex, at least below the water/lipid interface.

The effect of the fluorination degree on the chloride transport ability was energetically characterised with the PMF calculations carried out for the free and chloride complexes of the **C3x** series, while the effect of extending the chain length was evaluated with tripodal molecules **C3F** and **C5F**. Indeed, while the $V_{S,max}$ values increase in the order **C3H** < **C3M** < **C3F**, as discussed above, for **C3F** and **C5F** they are similar and, therefore, the different transport activities of these two molecules must be a result of their different lipophilicities and interactions with the phospholipids.

31 evenly spaced windows (*ca.* 1.0 Å), with the tripodal molecules successively positioned between z = 0 Å and z = 30 Å, were simulated with a z-dimension distance restraint with a force constant of 5.0 kcal mol^{-1} Å^{-2}. Each US window was simulated for 80 ns apart of the chloride complexes of **C3F** and **C5F**, which were run for 100 ns. The convergence and equilibration of the PMF profiles was ascertained assessing the data in sequential intervals of 20 ns, until it was apparent that the curves were continuously overlapped (see Figure V-9). For instance, for **C3F**·OA, longer simulation windows were needed, due to the flexibility of OA.

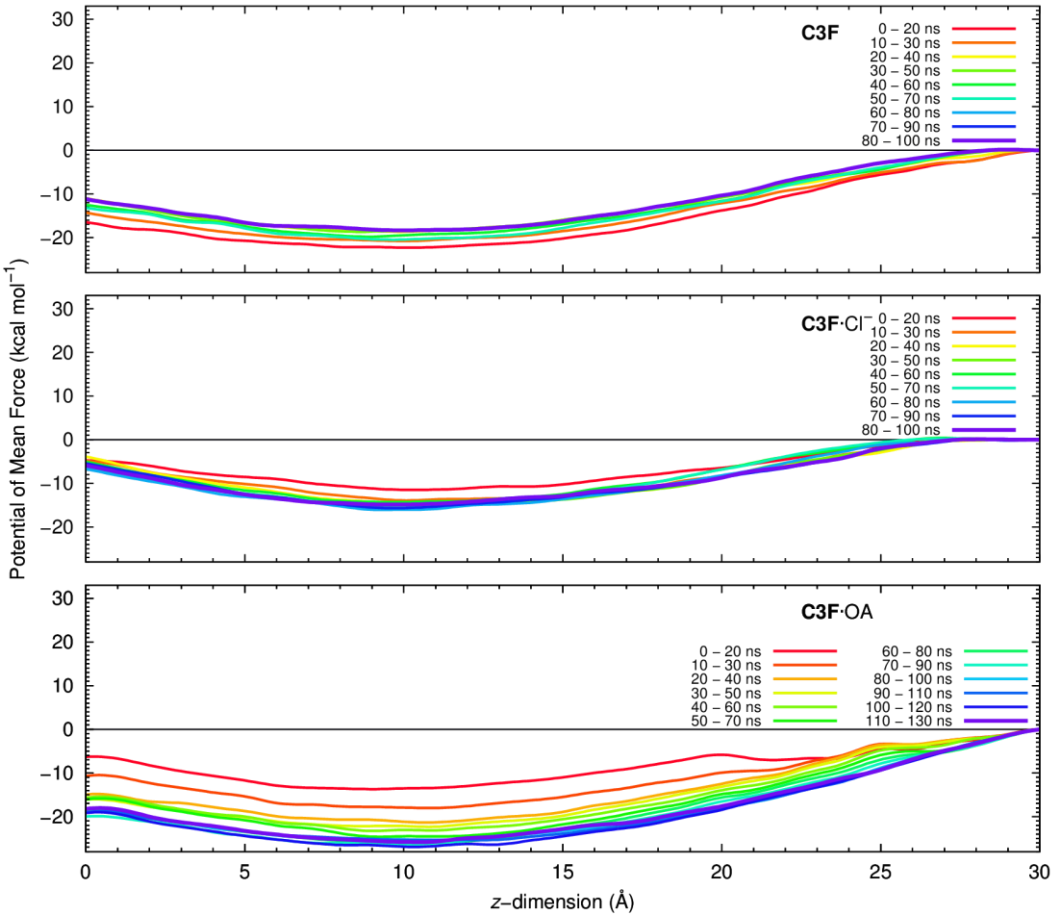


Figure V-9. The equilibration and convergence of the US simulation windows were assessed in 20 ns intervals (coloured according to time period), taken through the trajectory and used to calculate the PMF of **C3F**, **C3F**·Cl^- or **C3F**·OA. The PMF curves for **C3F** and **C3F**·Cl^- converged within the initial 50 ns. However, for **C3F**·OA extended simulation time was needed to achieve the sampling time of 50 ns.

The hydration of the free tripodal molecules and their chloride complexes was evaluated throughout the last 50 ns of the US simulations and is presented in Figure V-10. It is clear that, as expected, the number of water molecules pronouncedly decreases as the tripodal receptors are located closer to the bilayer core. At this position, the complexes and free transporters are surrounded by an equivalent number of water molecules, ranging from 4 to 6. This result indicates that the translocation of the transporters is accompanied by a drift of water molecules, given that no solvent molecules were present in the starting geometries around the bilayer core. In addition, this number of water molecules is comparable with the solvation found in the SMD simulation of **C3F**·Cl^- when the complex is at the bilayer core (see Figure V-8).

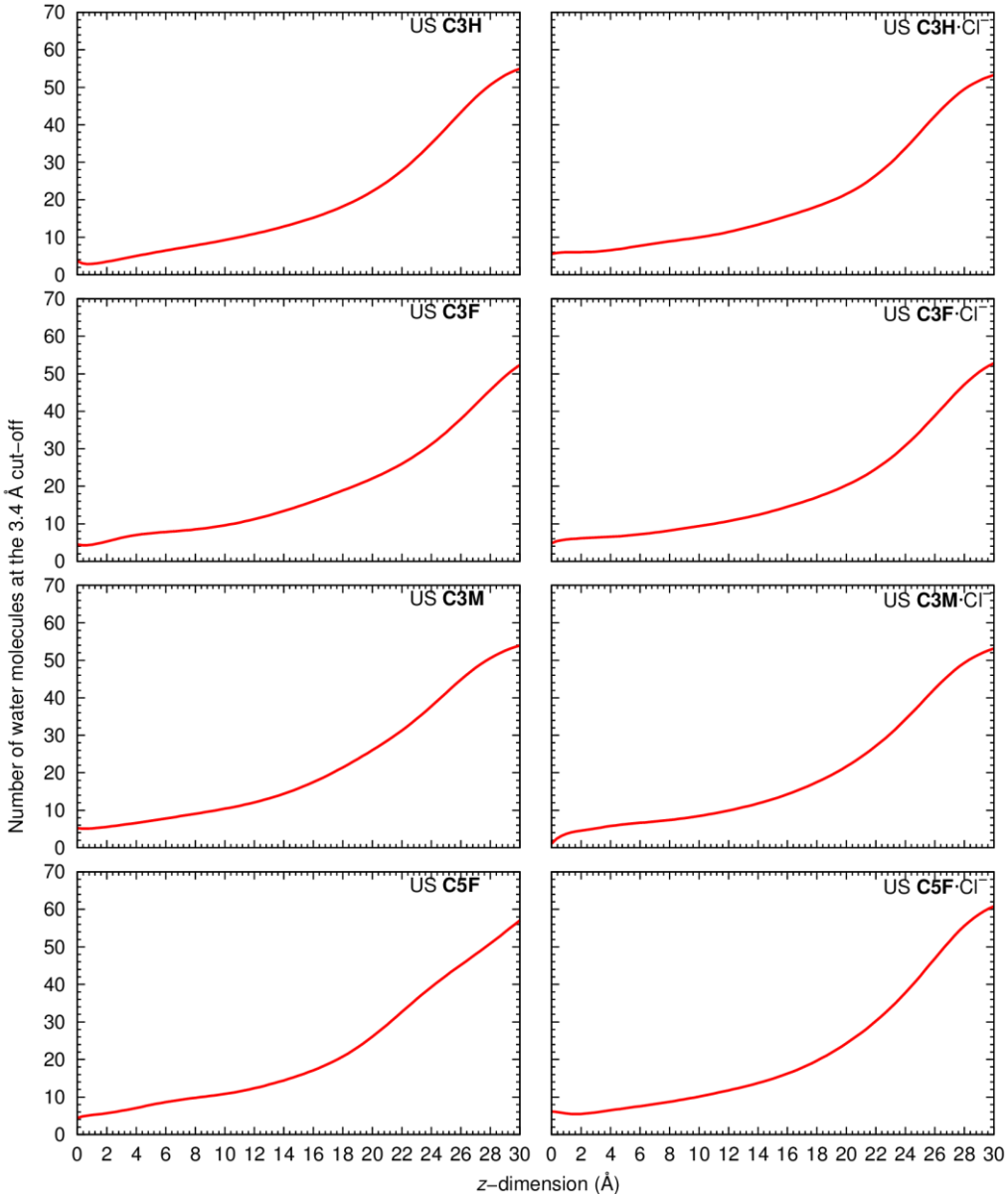


Figure V-10. Variation of the number of water molecules within the solvation shell defined by a cut-off of 3.4 Å from **C3H**, **C3F**, **C3F** and **C5F**, during the last 50 ns of the US simulations of the free transporters (left) or their chloride complexes (right), between z = 0 Å and z = 30 Å. Data were smoothed using Bézier curves.

The interactions between the tripodal molecules and the phospholipids were primarily assessed throughout the entire US simulations time, using the α angle, defined between the bilayer normal (z-dimension) and the vector $\vec{t}$, determined by N_{tren} and C_{ter}, as given above. An α angle of 0° corresponds to a tripodal molecule nearly aligned with the phospholipid chains, with the *tren* moiety closer to the water/lipid interface, while an angle of 180° indicates that the molecule is orientated in the opposite way, with the *tren* moiety closer to bilayer core. The α angle values were histogrammed together with the positions occupied by the tripodal receptors throughout the US windows, and the corresponding 2D histograms are shown in Figure V-11 for the free molecules and in Figure V-12 for the chloride complexes. In the windows where the free molecules or the chloride complexes are found in water phase (typically, z-dimension values above 20 Å) the α angles are scattered, consistent with a random orientation of the transporters. In contrast, inside the POPC bilayer, the range of α angles experienced by the tripodal molecules is confined to a narrow interval between *ca.* 20 and 60°.

Moreover, the number of hits decreases along with the lipophilicity of the molecule, as **C5F** and **C3F** have broader red areas than **C3M** and **C3H**. The chloride complexes follow the same trend, although the α angles are more spread as consequence of the presence of the anion. In addition, this preferential orientation was also observed in the passive diffusion MD simulations in scenarios A and B (see above).

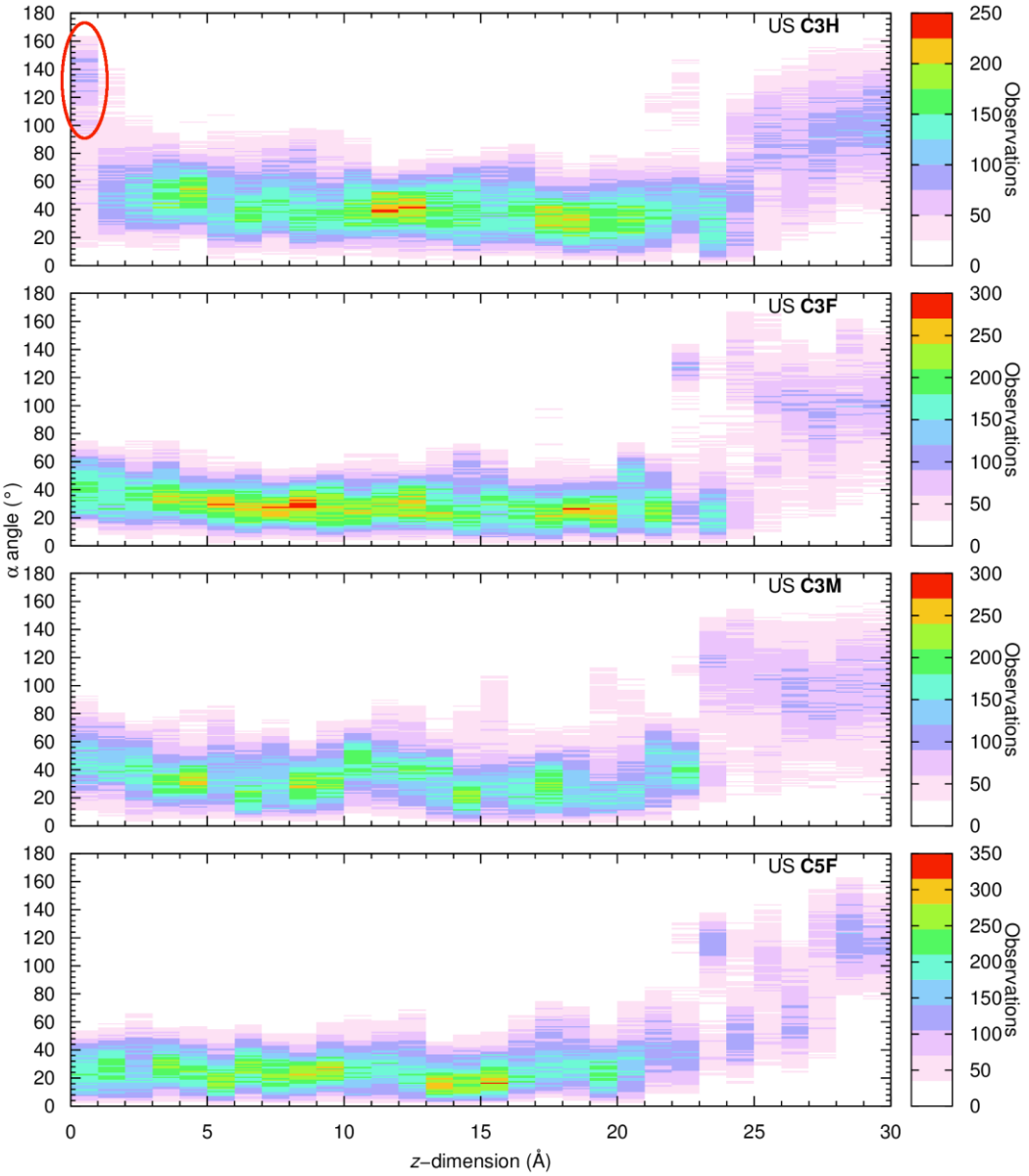


Figure V-11. 2D histogram created from the α angle values monitored along the *z*-dimension positions of the free transporters **C3H**, **C3F**, **C3M** and **C5F**, throughout the 31 independent US windows. The colour ranges from white (no occurrence) to red (several occurrences).

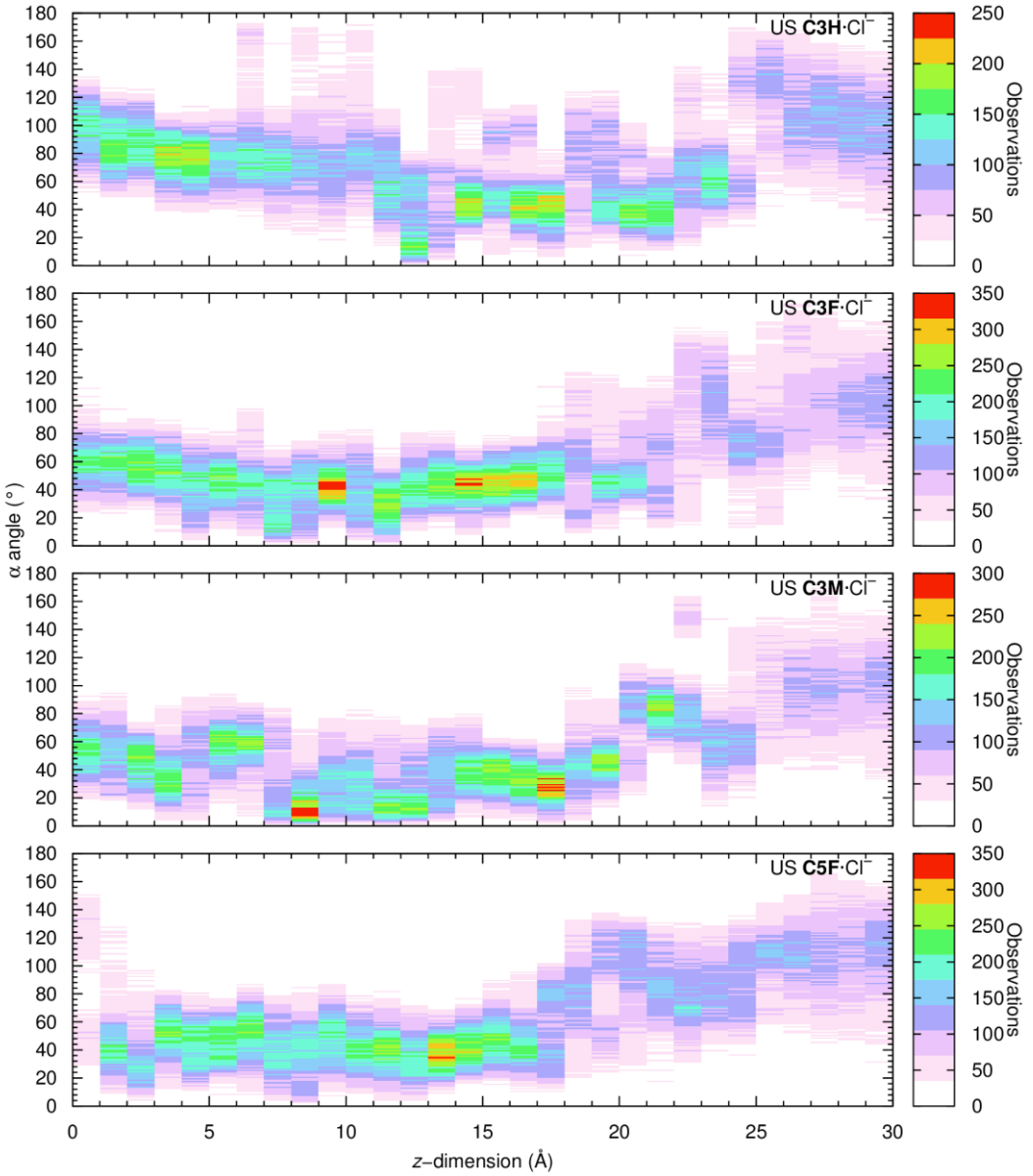


Figure V-12. 2D histogram created from the α angle values monitored along the *z*-dimension positions of the chloride complexes of **C3H**, **C3F**, **C3M** and **C5F**, throughout the 31 independent US windows. The colour ranges from white (no occurrence) to red (several occurrences).

The free energy profiles were estimated using the variational free energy profile (vFEP) method,[53, 54] from the last 50 ns of the US simulations with **C3H**, **C3F**, **C3M** and **C5F**, and are plotted in Figure V-13 for their chloride complexes (left) and the free transporters (right; the individual profiles, plotted with the bootstrap errors calculated from 100 random data sets with the same size, are shown in Figure V-14). All profiles were normalised for z = 30 Å (bulk water) and the energy profile for the translocation of a free chloride, assessed during the last 10 ns of 25 ns long US simulations, is also plotted in Figure V-14 for comparison purposes (its individual profile, with the bootstrap errors, is shown in Figure V-15). As the free tripodal molecules permeate the POPC bilayer the free energy drops until $z \approx$ 10 Å, below the water/lipid interface, where most profiles reach their minima (for **C5F**, it is at $z \approx$ 8.0 Å). Subsequently, the free energy rises until the bilayer core (z = 0 Å), indicating that an energy barrier must be surmounted to cross the lipid tails towards the opposite leaflet, as follows: 5.45 (**C3H**), 7.27 (**C3F**), 8.81 (**C3M**), and 7.02 (**C5F**) kcal mol^{-1}. On the other hand, the entrance in the POPC bilayer is favoured by 11.02 ± 0.14 (**C3H**), 18.80 ± 0.20 (**C3F**), 13.00 ± 0.16 (**C3M**), and 24.10 ± 0.16 (**C5F**) kcal mol^{-1}, which

are the same energy barriers, in absolute values, associated with the exit of the free molecules from the phospholipid membrane. Furthermore, the profile of **C3H** shows a small inflexion around the bilayer core, which is the result of the transporter's orientation change relatively to the bilayer normal (circled area in Figure V-11).

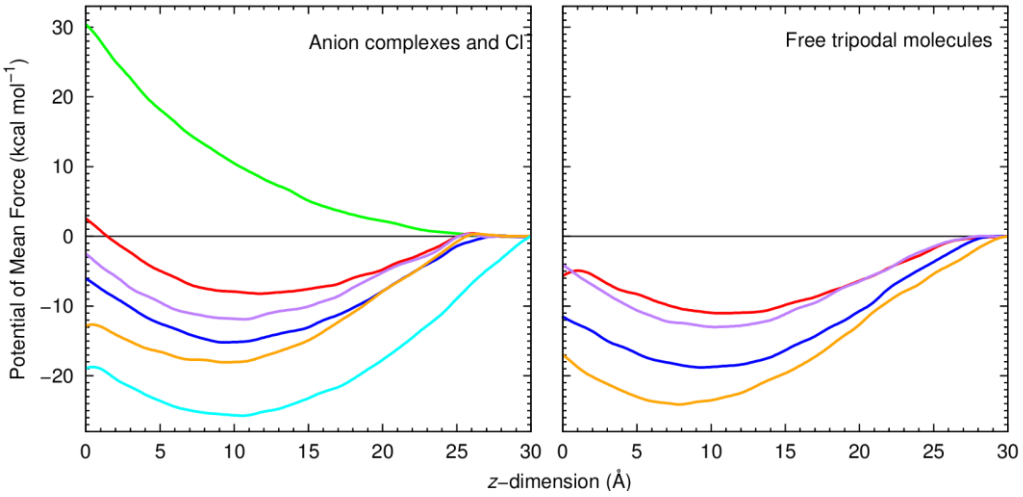


Figure V-13. PMF as a function of the *tren* thiourea derivatives' distance to the membrane COM (z = 0 Å), for the anion complexes (left) and for the free tripodal molecules (right). The red, blue, purple and orange lines correspond to **C3H**, **C3F**, **C3M** and **C5F**, in this order. The green line corresponds to the PMF of a free chloride and the cyan line corresponds to the PMF of **C3F**·OA.

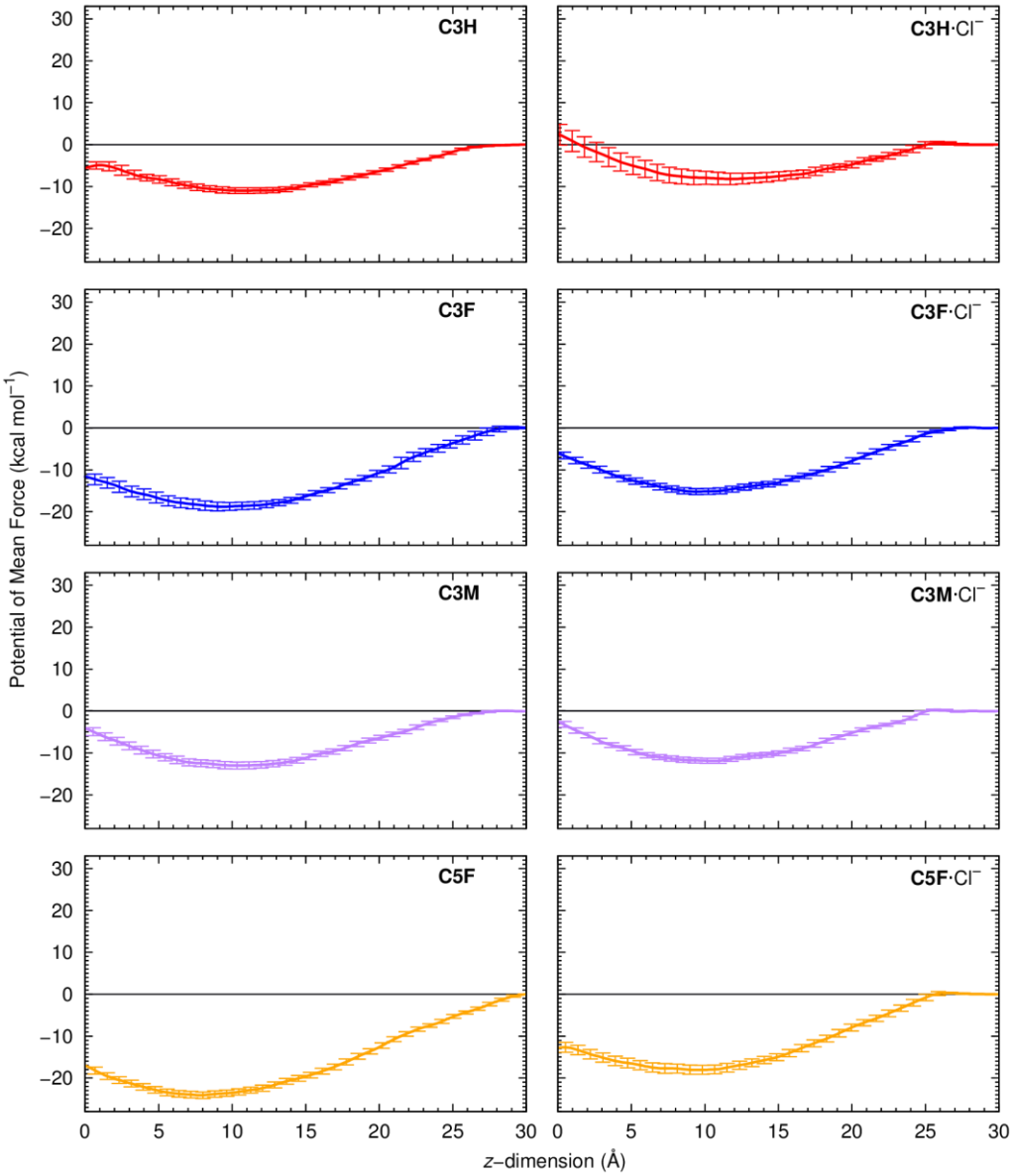


Figure V-14. PMF as a function of the *tren* thiourea derivatives' distance to the membrane COM (z = 0 Å), for the free **C3H**, **C3F**, **C3M** and **C5F** as well as their chloride complexes. The error bars correspond to the bootstrap errors calculated from 100 random data sets with the same size and are upscaled 5 times.

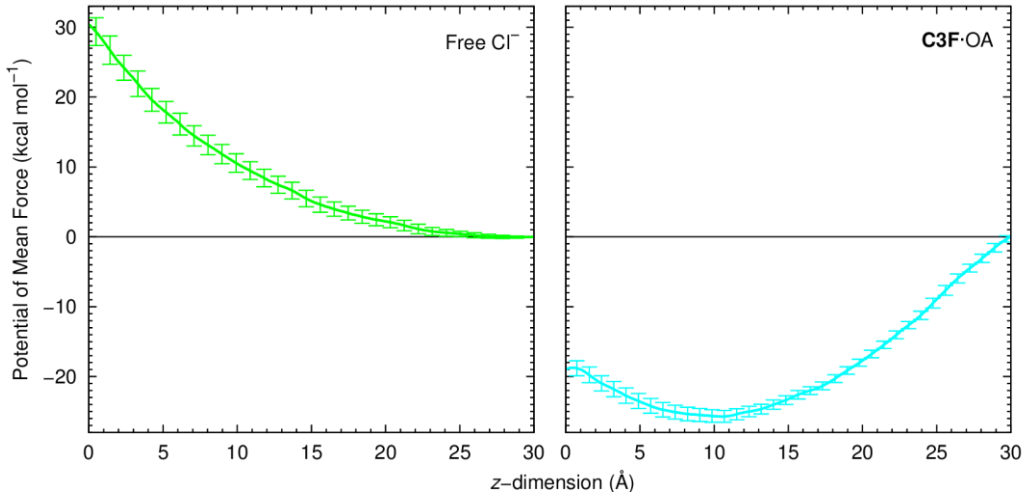


Figure V-15. PMF as a function of the distance to the membrane COM (z = 0 Å), for the free chloride and **C3F**·OA. The error bars correspond to the bootstrap errors calculated from 100 random data sets with the same size and are upscaled 5 times.

The energy profiles of the chloride complexes mirror the same behaviour of the free molecules, although the global minima are higher due to the presence of the coordinated hydrophilic anion: −8.21 ± 0.27 (**C3H**), −15.21 ± 0.14 (**C3F**), −11.89 ± 0.12 (**C3M**), and −18.06 ± 0.21 (**C5F**) kcal mol^{-1}. As in the free molecules, these minima also occur at $z \approx 10$ Å, in close agreement with the equilibrium positions attained by the chloride complexes in the passive diffusion MD simulation carried out in scenario A. The presence of the anion is also translated to the slightly higher energy barriers associated with the translocation of the less lipophilic complexes across the bilayer core: 10.55 (**C3H**), 9.08 (**C3F**), and 9.27 (**C3M**) kcal mol^{-1}. In contrast, **C5F**·Cl$^-$ has to overcome an energy barrier of 5.19 kcal mol^{-1} at the bilayer core, which is lower than for the free molecule. However, the necessary energy for the chloride complexes to enter or leave the POPC bilayer follow the same trend of the free tripodal molecules, as evident when these energies are plotted against each other, leading to a straight relation (R^2 = 0.94, as depicted in Figure V-16). Finally, when the energy barriers of the chloride complexes are compared with the free anion (see Figure V-13A), it is clear that the unassisted translocation of the anion is highly disfavoured, meaning that the presence of the tripodal transporters is essential to shuttle the chloride ion across the POPC bilayer.

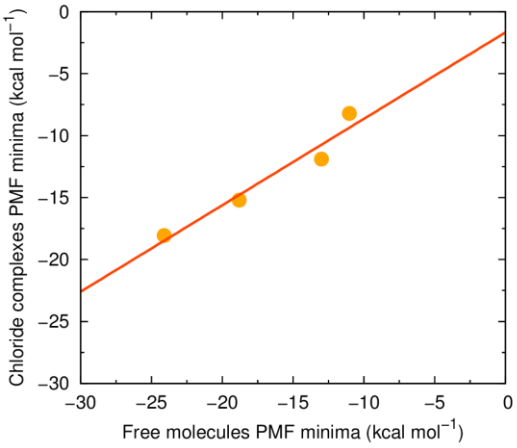


Figure V-16. Comparison of the PMF minima between the free tripodal molecules **C3H**, **C3F**, **C3M** and **C5F** and their chloride complexes (R^2 = 0.94).

Noteworthy, when the retention times are plotted against the global minima of the individual PMF profiles (see Figure V-17), highly linear correlations appear for the free transporters (R^2 = 0.96) as well as for their chloride complexes (R^2 = 0.93). These results undoubtedly show that the lipophilicity, determined by the fluorination degree and thiourea chain length, plays an important role on the interaction of **C3H**, **C3M**, **C3F** and **C5F** with the phospholipids and, consequently, in their transport activities. This has prompted us to estimate the MM interaction energies between the transporters and the POPC lipids at the global minima, given that the tripodal molecules were found to adopt a well-defined orientation within the highly packed lipid medium.

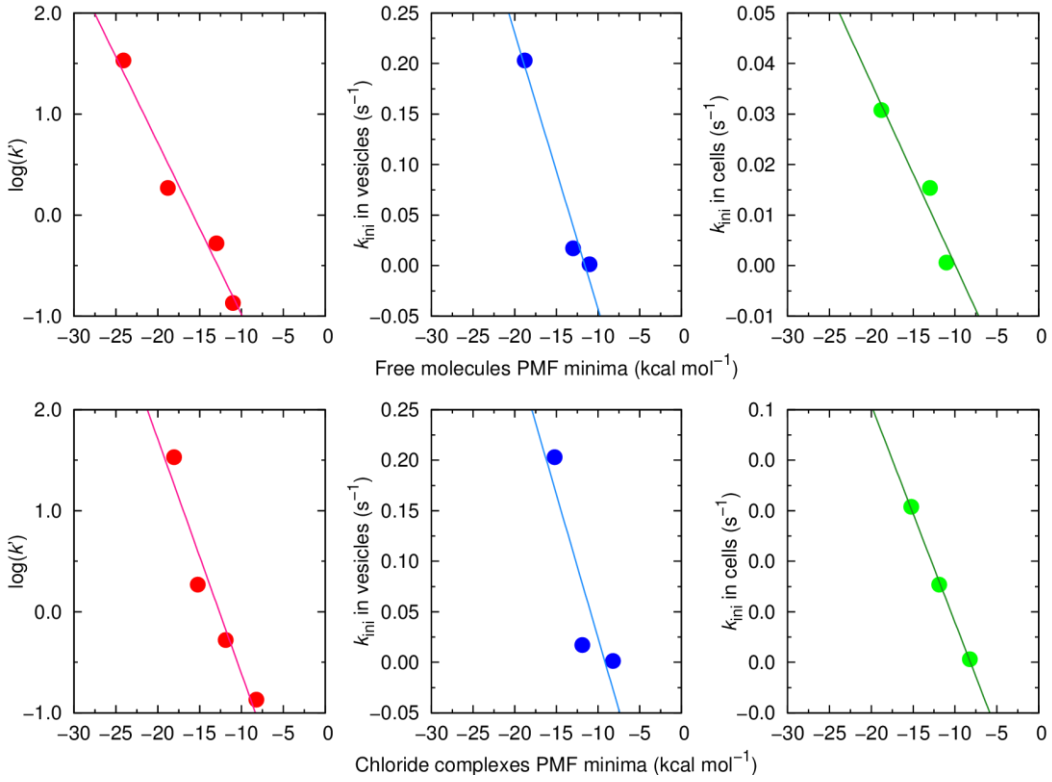


Figure V-17. Left: log(*k'*) as a function of the PMF minima calculated for the free tripodal molecules **C3H**, **C3F**, **C3M**, and **C5F** (top, R^2 = 0.96) and their chloride complexes (bottom, R^2 = 0.93). Middle: Chloride efflux k_{ini} in vesicles as a function of the PMF minima calculated for the free tripodal molecules of the **C3x** series (top, R^2 = 0.97) and their chloride complexes (bottom, R^2 = 0.78). Right: Chloride efflux k_{ini} in cells as a function of the PMF minima calculated for the free tripodal molecules of the **C3x** series (top, R^2 = 0.93) and their chloride complexes (bottom, R^2 = 1.00).

The interaction energies between the POPC bilayer and the tripodal molecules were energetically evaluated through Eq. V-1:

$$\Delta E_{TOTAL} = \Delta E_{ELEC} + \Delta E_{vdW} + \Delta E_{INT} \qquad \text{Eq. V-1}$$

ΔE_{ELEC} and ΔE_{vdW} correspond to the non-bonded electrostatic and van der Waals energy terms, while ΔE_{INT} corresponds to the sum of bond, angle and torsion energies. The three individual MM energy components are given by Eq. V-2 through Eq. V-4:

$$\Delta E_{ELEC} = E_{ELEC_{System}} - \left(E_{ELEC_{POPC}} + E_{ELEC_T}\right) \qquad \text{Eq. V-2}$$

$$\Delta E_{vdW} = E_{vdW_{System}} - \left(E_{vdW_{POPC}} + E_{vdW_T}\right) \qquad \text{Eq. V-3}$$

$$\Delta E_{INT} = E_{INT_{System}} - \left(E_{INT_{POPC}} + E_{INT_T}\right) \qquad \text{Eq. V-4}$$

Here *System* represents the transporter and the phospholipid membrane, excluding the water molecules and all chloride and sodium ions. *POPC* stands for the 128 phospholipid molecules and *T* for the isolated tripodal molecule (**C3H**, **C3M**, **C3F** or **C5F**). These three energies were estimated with snapshots extracted every 100 ps from the last 50 ns of the US window corresponding to the *z*-dimension value of the minimum in energy profile of the corresponding transporter.

The bonded term (ΔE_{INT}) amounts to zero, given that the individual terms $E_{INT_{System}}$, $E_{INT_{POPC}}$ and E_{INT_T} were calculated using the same MD simulation. In these conditions, through Eq. V-1, only the contribution of both non-bonded energy terms is evaluated in the intermolecular interactions between the phospholipids and the tripodal transporters. The average values of ΔE_{TOTAL}, ΔE_{ELEC} and ΔE_{vdW} for the relevant US simulations are given in Table V-3. The vdW interactions are the main contributor for the total MM interaction energy, showing that the transport activity is strongly dependent on the interactions between the phospholipids and the tripodal transporters, also leading to a direct relation for the free transporters between the vdW contributions and the global minima in the individual free energy profiles (R^2 = 0.90), as plotted in Figure V-18.

Table V-3. Average MM energy terms (kcal mol^{-1}), with the corresponding standard deviations, assessed during the last 50 ns of the US window at z = 10 Å for **C3H**, **C3F** and **C3M** or z = 8.0 Å for **C5F**.[a]

Molecule	ΔE_{TOTAL}	ΔE_{ELEC}	ΔE_{vdW}
C3H	−77.47 ± 12.04	−21.03 ± 9.33	−56.44 ± 5.38
C3F	−103.57 ± 15.10	−37.21 ± 13.15	−66.37 ±4.98
C3M	−110.23 ± 24.18	−52.01 ± 21.68	−58.23 ± 5.51
C5F	−84.71 ± 15.10	−17.53 ± 11.00	−67.18 ± 5.81

[a] N = 500.

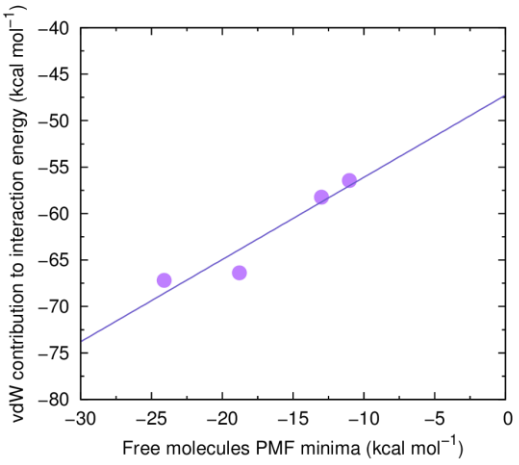


Figure V-18. vdW contribution for the total MM interaction energy between the free tripodal transporters **C3H**, **C3F**, **C3M** and **C5F** as a function of the PMF minima calculated for them (R^2 = 0.90)

Moreover, the transport activity in the series with the same thiourea chain lengths, the **C3x** series, is straightforwardly related with the energy associated with the entry or exit of the free or anion complexed tripodal molecules, as illustrated in Figure V-17. In other words, in this molecules' set, **C3F** with the higher lipophilicity and chloride affinity (the $V_{S,max}$ values discussed above) is the best transporter. In stark contrast, although **C5F** has a comparable anion affinity, it has to overcome a larger energy barrier to reach the water/lipid interface to release the anionic guest to the water phase. This means that the transport activity of these tripodal molecules results from a delicate balance between the lipophilic character and the anion binding affinity.

As aforementioned, the addition of OA considerably enhances the anion efflux promoted by the fluorinated molecules, namely **C3F**, leading us to propose that the return of the transporter after the translocation of chloride occurs associated with OA rather than in its free form (see Figure V-1). Thus, to investigate this possible pathway, the interaction of **C3F**·OA with a POPC bilayer was also ascertained through US simulations, followed by the reconstruction of the corresponding PMF, using the methodology adopted above for the free transporters and their chloride complexes. Given that the carboxylate group of OA was maintained encapsulated in the tripodal binding pocket of **C3F** throughout the 31 MD simulations of 130 ns,[§§§] the orientation of the long OA alkyl chain relatively to the membrane normal was assessed monitoring the position of the terminal CH_3 carbon atom (OA_{CH3}), together with the position of the tripodal nitrogen atom (N_{tren}). The evolution of these positions along the sampling period (the last 50 ns) is plotted in Figure V-19 and shows that in the middle of the POPC bilayer (z = 0 Å), characterised by the high disorder degree of the lipid chains, the OA chain is randomly orientated. As the position of the complex is moved toward the water/lipid interface, the OA chain becomes progressively aligned with the phospholipid alkyl tails. Noteworthy, when the transporter is at the interface level or in the water phase, the OA's tail is partially immersed between the phospholipids and presents an equivalent orientation, as illustrated in Figure V-20.

[§§§] The MD simulation time was extended due to the flexibility of OA.

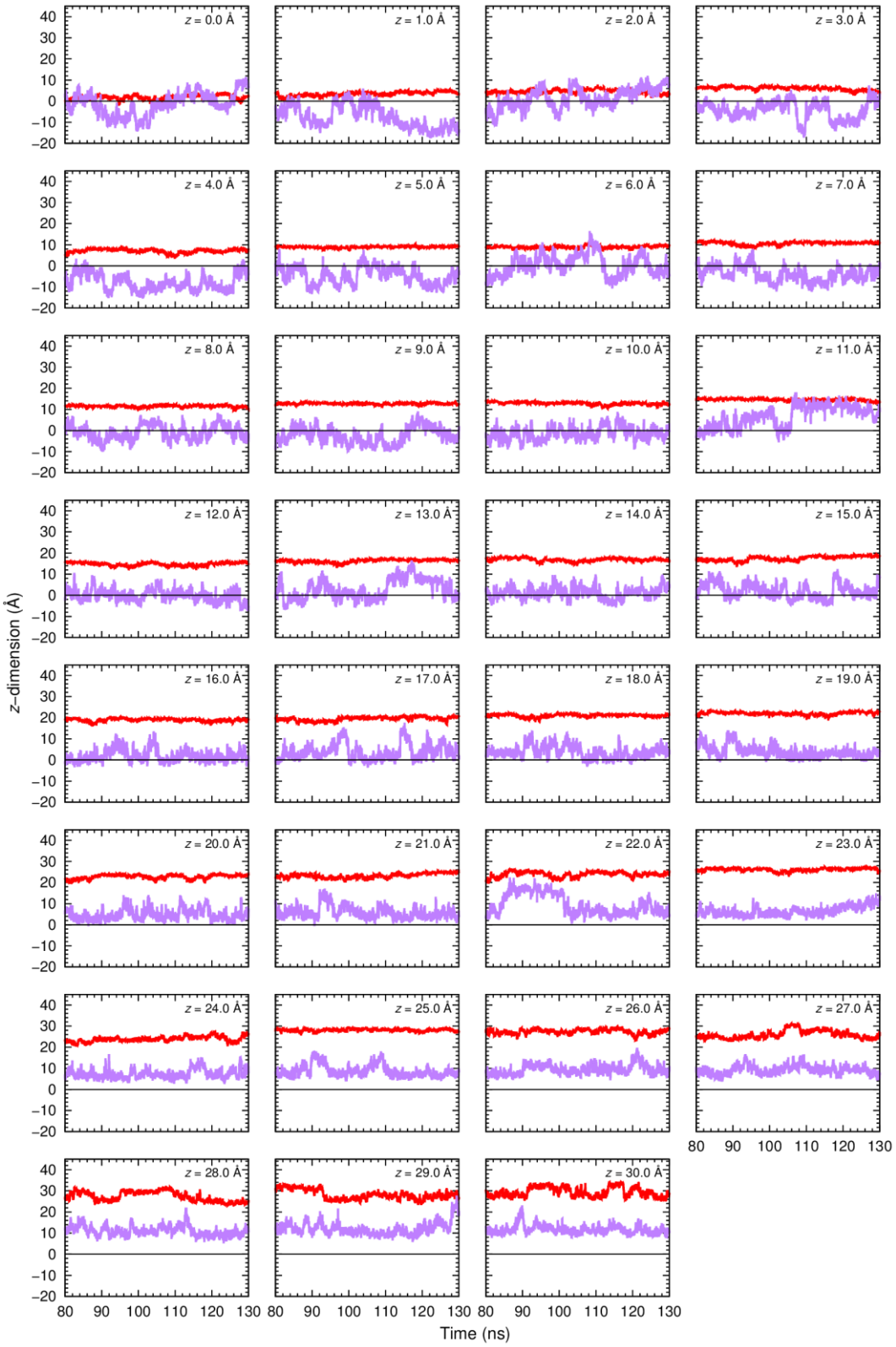


Figure V-19. Monitorisation of the N_{tren} (red) and OA_{CH3} (purple) referential points in **C3F** and OA along the final 50 ns of the corresponding 31 US windows. The centre of the bilayer is represented as a black line at z = 0 Å.

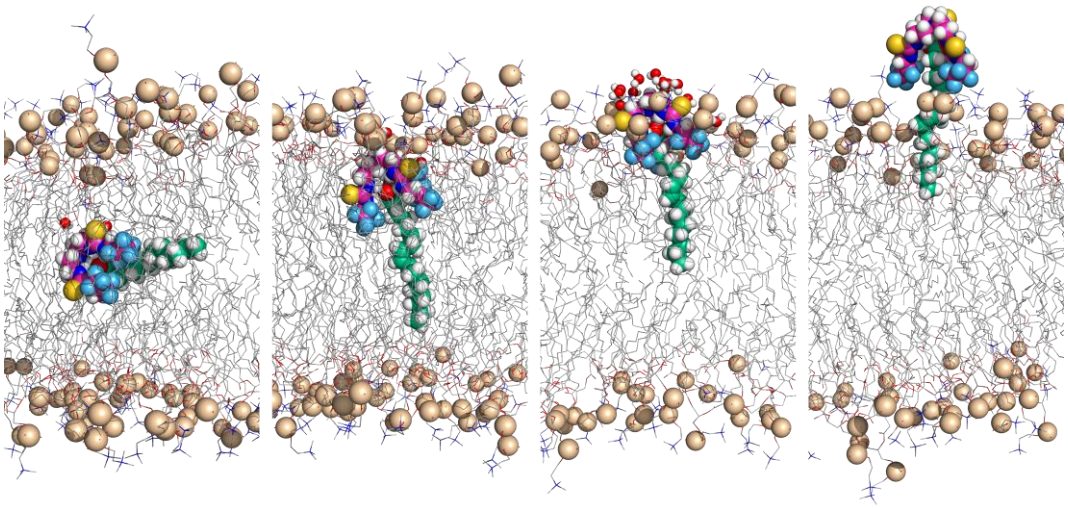


Figure V-20. Illustrative snapshots depicting the orientation of **C3F**·OA in the US windows at z = 0, 10, 20, and 30 Å (from left to right). OA is shown in spheres, with its carbon atoms in light green. Remaining details as given in Figure V-4.

The PMF reconstructed from all these MD simulations is plotted in Figure V-15 and has a lower energy minimum of −24.92 ± 0.17 kcal mol^{-1}, occurring at a distance along the z-dimension similar to the ones found for **C3F** (minimum of −18.80 kcal mol^{-1}) and **C3F**·Cl$^-$ (minimum of −15.21 kcal mol^{-1}). On the other hand, to cross the bilayer core, an energy barrier of 8.42 kcal mol^{-1} was found, which is only *ca.* 1 kcal mol^{-1} higher than for free **C3F**. In other words, the reinternalisation of the transporter after the release of the chloride ion is energetically favoured when associated with OA rather than being free, due to the interactions between the long alkyl chain of OA and the phospholipid tails. Therefore, these energetic findings suggest that OA can facilitate the translocation of **C3F** after it has promoted the chloride efflux. On the other hand, the high energy barrier of 24.92 kcal mol^{-1} for **C3F**·OA to reach the water/lipid interface suggests that this event is very unlikely and consequently, the complex dissociation should occur at the interface, freeing the transporters to uptake a solvated chloride. Two further MD simulations were carried without distance restraints between the carboxylate and the tripodal molecule to evaluate how OA dissociates from **C3F** when the receptor is above the interface (z = 30 Å). In the first one, the association was maintained throughout the 100 ns, with OA being passively internalised and dragging **C3F** along with it, as depicted in Figure V-21, path A. In the second MD simulation, with a distance restraint applied to **C3F** along the z-dimension, the tripodal receptor remains in the water phase, free to uptake and transport solvated chloride ions, while OA is recycled into the bilayer, as depicted in Figure V-21, path B. Both MD simulations are in agreement with the mechanism proposed in Figure V-1.

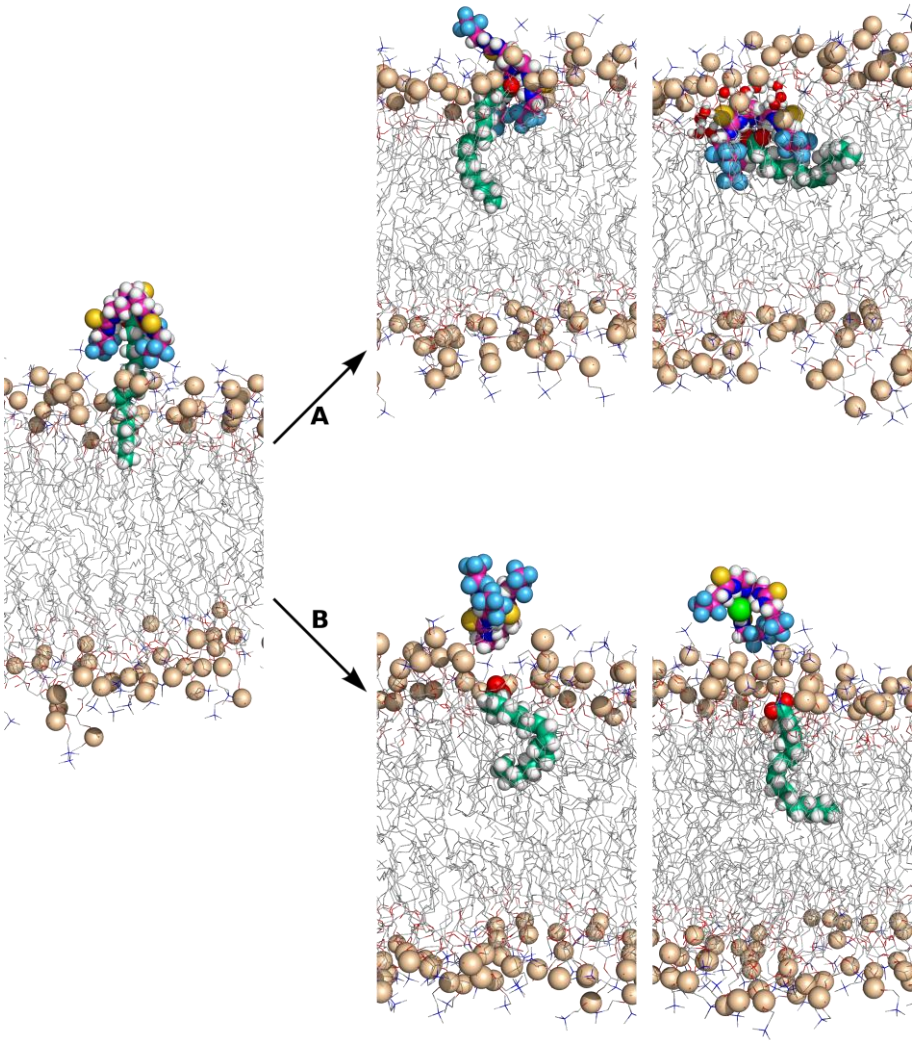


Figure V-21. From the same starting geometry, two MD simulations were carried out to assess the dissociation of **C3F**·OA. A: No distance restraints were applied to the system; B: a distance restraint of 30 Å along the *z*-dimension was applied between **C3F** and the COM of the bilayer.

V.4. Conclusions

The experimentally assessed anion transport properties of ten tripodal molecules with different fluorination degrees were rationalised with resort to molecular modelling methods. Fluorination of alkyl chains appears to allow a fine tuning of the lipophilic character of putative anion transporters while maintaining moderate anion binding properties. Indeed, the $V_{S,max}$ values ascertained were found to be consistent with the degree of fluorination of the **C*n*H**, **C*n*F** and **C*n*M** series.

Concerning the passive diffusion MD simulations, it was shown that extending the alkyl chains, as well as their degree of fluorination, influences the orientation of the tris-thioureas. While **C2H**, **C2F** and **C3H** are more prone to have an undefined pose at the water/lipid interface level, the remaining molecules have a clearly defined orientation, with the *tren* moiety pointing to bulk water and the thiourea substituents closer to the bilayer core.

For the first time, we have observed, through MD simulations, a neutral molecule facilitating the chloride transmembrane transport, as follows: anion uptake on the water phase of one side of the bilayer, followed by the diffusion of the complex across the POPC membrane model, and, subsequently, release of chloride to the opposite aqueous phase. In other words, this is the definitive proof of the anion carrier mechanism at the atomistic level.

Furthermore, the US simulations, coupled with the corresponding PMF profiles, show that the more lipophilic molecules partition easily into the phospholipid bilayer, characterised by significant energy gains (**C3F** and **C5F**), which are reduced for the less lipophilic molecules (**C3H** and **C3M**). These energetic gains have the same magnitude of the penalty that each molecule must overcome to exit the bilayer. In addition, the energy values of the PMF minima are intrinsically related with the vdW interactions between the thiourea chain substituents and the phospholipid tails. Moreover, plotting the free energy profiles minima *vs.* anion transport activity data in vesicles or cell lines results in nearly linear trends.

Finally, the US simulations carried out with the **C3F**·OA complex, followed by the estimate of the associated free energy profile, suggest that carboxylate anion with a long aliphatic tail can facilitate the passive back-diffusion of the transporter through the phospholipid bilayer after the release of chloride.

V.5. References

1. J. M. Tomich, U. Bukovnik, J. Layman and B. D. Schultz, *Channel Replacement Therapy for Cystic Fibrosis, Cystic Fibrosis – Renewed Hopes Through Research*, InTech, 1st edn., 2012.
2. N. Busschaert and P. A. Gale, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 1374-1382.
3. N. Busschaert, C. Caltagirone, W. Van Rossom and P. A. Gale, *Chem. Rev.*, 2015, **115**, 8038-8155.
4. B. A. McNally, A. V. Koulov, T. N. Lambert, B. D. Smith, J. B. Joos, A. L. Sisson, J. P. Clare, V. Sgarlata, L. W. Judd, G. Magro and A. P. Davis, *Chem. Eur. J.*, 2008, **14**, 9599-9606.
5. V. Saggiomo, S. Otto, I. Marques, V. Felix, T. Torroba and R. Quesada, *Chem. Commun.*, 2012, **48**, 5274-5276.
6. N. Busschaert, S. J. Bradberry, M. Wenzel, C. J. E. Haynes, J. R. Hiscock, I. L. Kirby, L. E. Karagiannidis, S. J. Moore, N. J. Wells, J. Herniman, G. J. Langley, P. N. Horton, M. E. Light, I. Marques, P. J. Costa, V. Félix, J. G. Frey and P. A. Gale, *Chem. Sci.*, 2013, **4**, 3036-3045.
7. H. Valkenier, C. J. E. Haynes, J. Herniman, P. A. Gale and A. P. Davis, *Chem. Sci.*, 2014, **5**, 1128-1134.
8. S. J. Edwards, I. Marques, C. M. Dias, R. A. Tromans, N. R. Lees, V. Felix, H. Valkenier and A. P. Davis, *Chem. Eur. J.*, 2016, **22**, 2004-2011.
9. N. J. Knight, E. Hernando, C. J. E. Haynes, N. Busschaert, H. J. Clarke, K. Takimoto, M. Garcia-Valverde, J. G. Frey, R. Quesada and P. A. Gale, *Chem. Sci.*, 2016, **7**, 1600-1608.
10. S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320-330.
11. M. J. Spooner and P. A. Gale, *Chem. Commun.*, 2015, **51**, 4883-4886.
12. N. Busschaert, P. A. Gale, C. J. Haynes, M. E. Light, S. J. Moore, C. C. Tong, J. T. Davis and W. A. Harrell, Jr., *Chem. Commun.*, 2010, **46**, 6252-6254.
13. N. Busschaert, M. Wenzel, M. E. Light, P. Iglesias-Hernandez, R. Perez-Tomas and P. A. Gale, *J. Am. Chem. Soc.*, 2011, **133**, 14136-14148.
14. S. Hussain, P. R. Brotherhood, L. W. Judd and A. P. Davis, *J. Am. Chem. Soc.*, 2011, **133**, 1614-1617.
15. S. Bahmanjah, N. Zhang and J. T. Davis, *Chem. Commun.*, 2012, **48**, 4432-4434.
16. D. O'Hagan, *Chem. Soc. Rev.*, 2008, **37**, 308-319.
17. B. E. Smart, *J. Fluorine Chem.*, 2001, **109**, 3-11.
18. M. J. Spooner, H. Lib, X. Wu, I. Marques, P. M. R. Costa, V. Félix, D. N. Sheppard and P. A. Gale, *Manuscript submitted for publication.*, 2017.
19. OECD, *Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method*, OECD Publishing.
20. X. Wu, L. W. Judd, E. N. W. Howe, A. M. Withecombe, V. Soto-Cerrato, H. Li, N. Busschaert, H. Valkenier, R. Perez-Tomas, D. N. Sheppard, Y. B. Jiang, A. P. Davis and P. A. Gale, *Chem*, 2016, **1**, 127-146.
21. X. Wu and P. A. Gale, *J. Am. Chem. Soc.*, 2016, **138**, 16508-16514.
22. L. V. Galietta, S. Jayaraman and A. S. Verkman, *Am. J. Physiol. Cell Physiol.*, 2001, **281**, C1734-1742.
23. H. Li, H. Valkenier, L. W. Judd, P. R. Brotherhood, S. Hussain, J. A. Cooper, O. Jurcek, H. A. Sparkes, D. N. Sheppard and A. P. Davis, *Nat Chem*, 2016, **8**, 24-32.
24. Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J.

Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

25. T. Lu and F. Chen, *J. Mol. Graph. Model.*, 2012, **38**, 314-323.
26. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
27. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.
28. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
29. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
30. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
31. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
32. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
33. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
34. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
35. E. N. W. Howe, Private communication, 2016.
36. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247-260.
37. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
38. L. Martinez, R. Andrade, E. G. Birgin and J. M. Martinez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
39. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
40. P. Li, L. F. Song and K. M. Merz, Jr., *J. Chem. Theory Comput.*, 2015, **11**, 1645-1657.
41. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
42. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
43. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
44. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
45. D. R. Roe and T. E. Cheatham, 3rd, *J. Chem. Theory Comput.*, 2013, **9**, 3084-3095.
46. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
47. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171-179.
48. S. K. Dey and G. Das, *Dalton transactions*, 2012, **41**, 8960-8972.
49. A. Gogoi and G. Das, *Supramol. Chem.*, 2013, **25**, 819-830.
50. P. J. Costa, I. Marques and V. Felix, *Biochim. Biophys. Acta*, 2014, **1838**, 890-901.
51. D. Laage, T. Elsaesser and J. T. Hynes, *Chem. Rev.*, 2017, **117**, 10694-10725.
52. C. T. Lee, J. Comer, C. Herndon, N. Leung, A. Pavlova, R. V. Swift, C. Tung, C. N. Rowley, R. E. Amaro, C. Chipot, Y. Wang and J. C. Gumbart, *J. Chem. Inf. Model.*, 2016, **56**, 721-733.
53. T. S. Lee, B. K. Radak, A. Pabis and D. M. York, *J. Chem. Theory Comput.*, 2013, **9**, 153-164.
54. T. S. Lee, B. K. Radak, M. Huang, K. Y. Wong and D. M. York, *J. Chem. Theory Comput.*, 2014, **10**, 24-34.

Chapter VI.
Chalcogen bonding macrocycles and [2]rotaxanes for anion recognition: an in silico *approach*

Summary

Electron-deficient heavy chalcogen atoms contain Lewis acidic σ-holes, able to form attractive supramolecular interactions, known as chalcogen bonding (ChB). Still, their potential in solution-phase anion binding applications is just in its infancy. Macrocycle **Mac** and rotaxane **Rot** hosts (see Scheme VI-1) contain a 5-(methylseleno)-1,2,3-triazolium motif, whose binding potential in solution was explored both experimentally and theoretically. This chapter presents, in detail, the molecular modelling investigations carried out in this stimulating project, comprising DFT calculations and Molecular Dynamics (MD) simulations of the ChB interactions, that allowed to rationalise the experimental binding data. Thus, this chapter illustrates how the molecular modelling tools can be used to investigate the anion recognition by interlocked supramolecular structures mediated by the non-conventional ChB interactions.

The work reported in this chapter was developed within the scope of the collaboration established with the experimental group of Professor Paul D. Beer (University of Oxford), resulting in the publication of the paper: J. Y. Lim, I. Marques, A. L. Thompson, K. E. Christensen, V. Felix and P. D. Beer, *J. Am. Chem. Soc.*, 2017, **139**, 3122-3133. Herein, only the computational studies carried out for this project are presented in detail.

VI.1. Introduction

Chalcogen bonds, the attractive noncovalent interactions between an electrophilic element of group 16 (sulfur, selenium and tellurium) and a Lewis base, are a type of σ-hole-based interactions, similar to the well-known halogen bonding (XB).[1, 2] Electron withdrawing groups covalently attached to the chalcogen atom anisotropically redistribute the electron density on the atom itself, forming electron-deficient σ-holes which can act as a Lewis acid. Consequently, either one or two σ-holes may form on a sp^2 or sp^3-hybridised chalcogen atom, respectively, affecting the geometry of the resulting chalcogen···Lewis base interaction.[3, 4] Due to their unique highly directional geometric requirements and comparable strength to ubiquitous hydrogen bonding (HB) interactions,[5, 6] ChB has been recently applied in organic reactivity,[7-9] materials and crystal engineering,[10, 11] anion transport,[12] self-assembly processes,[13-15] and pharmaceutics.[16] Nevertheless, compared to XB interactions,[17, 18] recognition studies in solution involving ChB systems are extremely rare, with their potential for anion binding only recently realised.[19-21] This is, in part, due to the inherent chemical instability of compounds bearing the heavier chalcogens, which, in spite of their enhanced ChB-donor properties, are often highly prone to oxidative decomposition and hydrolysis,[20, 22] which limits their incorporation in more complex host molecules and limits the study of their binding properties in more competitive protic/aqueous solvents. The incorporation of ChB donor groups into elaborate host structures such as macrocycles and mechanically interlocked molecules for anion recognition applications is unprecedented.[23]

The synthesis of readily accessible 5-(methylseleno)-1,2,3-triazolium ChB motifs has led to the design of chemically robust and stable ChB hosts macrocycle **Mac** and rotaxane **Rot** for anion recognition (see Scheme VI-1).[23] Other than strongly polarising the chalcogen atoms for potential ChB–anion interactions, the potent electron withdrawing nature of the triazolium unit also helps to stabilise the chalcogen atoms by reducing their electron density and, thus, propensity for oxidation and hydrolysis,[24] with the ChB donor motifs being able to withstand the various chemical manipulations required for integration into macrocyclic structures. The crystal structure of the iodide complex of **Mac**, shown in Figure VI-1, provided the first evidence for the importance of the σ-holes on the sp^3-hybridised Se atoms for anion recognition in solid state.[23]

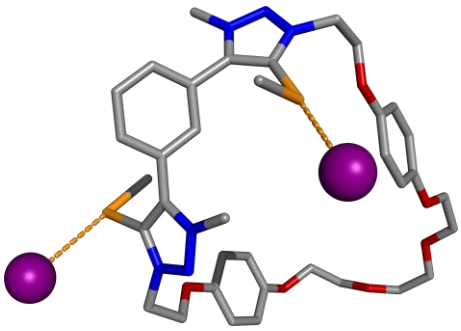
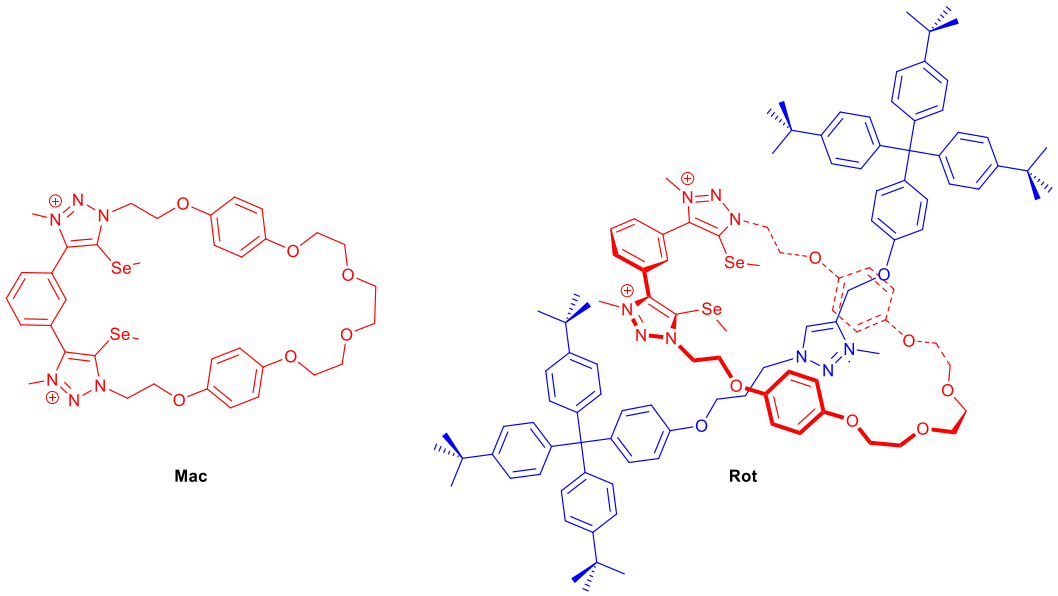


Figure VI-1. Crystal structure of the **Mac**·$(I^-)_2$ complex, showing the ChB interactions (orange dashed lines) between the Se atoms (orange sticks) and the anions (purple spheres). The carbon, nitrogen and oxygen atoms are depicted as grey, blue and red sticks. The hydrogen atoms, solvent molecules, and the minor components of the disorder are omitted for clarity.



Scheme VI-1. Structures of ChB hosts macrocycle **Mac** and [2]rotaxane **Rot**.

^{1}H NMR anion binding studies were carried out with **Mac** and **Rot** complexes of chloride, bromide and iodide, in aqueous solvent mixtures of either 2% D_2O for the macrocycle or 20% D_2O water for the rotaxane in acetone-d_6 (v/v).[23] Table VI-1 summarises the 1:1 stoichiometric association constants (K_a) determined from the titration data. Both hosts present a distinct Hofmeister bias for anion binding, with anion affinities decreasing in the order $I^- > Br^- > Cl^-$, similar to other σ-hole-based halogen bonding hosts in the presence of water.[25-27]

Table VI-1. Anion association constants (M^{-1}) for **Rot** and **Mac**.[23]

		Host	Cl^-	Br^-	I^-
K_a	acetone/D_2O 4:1	**Rot**	696 ± 28	983 ± 18	2084 ± 109
	acetone/D_2O 98:2	**Mac**	1180 ± 47	1298 ± 66	1534 ± 115

This theoretical work comprises two different parts: the preliminary parameterisation of the Se bonding parameters in the macrocycle, followed by the parameterisation of the ChB intermolecular interactions; and the evaluation of the halide recognition mediated by ChB interactions through DFT calculations and MD simulations in water/acetone solvent mixtures used in the experimental NMR binding studies.

VI.2. Methods

The theoretical approach used in this study comprised the assembly of the necessary starting structures, preliminary quantum calculations, development of the specific Se force field parameters and parameterisation of the ChB interactions, followed by MD simulations of the free ChB-based receptors and their halide complexes in competitive aqueous solvent mixtures.

VI.2.1. Starting structures

The structure of the macrocyclic component of **Rot** was directly taken from the single crystal X-ray structure of **Mac**.[23] The structure of the axle was generated from the crystal structure of a rotaxane analogous deposited with the Cambridge Crystallographic Data Centre (CCDC)[28, 29] under RefCode TACPUK,[30] replacing the iodine atom of the triazolium with an hydrogen atom. Subsequently, the ChB **Rot** rotaxane was obtained by the assembly of these two components in an interlocked fashion with the C–H triazolium binding unit and the methylseleno (SeMe) units of the macrocyclic cleft adopting an almost orthogonal relative disposition. The halide complexes were further generated by the insertion of each anion into the **Rot** binding cavity establishing two putative chalcogen bonds with the SeMe groups and a putative C–H···A interaction with the C–H triazolium (A = Cl^-, Br^-, or I^-). In addition, the axial component's triazolium methyl group points towards the oxygen atoms of the macrocyclic ether loop.

The force field parametrisation of the chalcogen bonds was carried out using the model compound **Mac**$_{methyl}$, which was generated by capping the benzene-1,3-bis(triazolium) chalcogen binding motif of **Mac** with two methyl groups.

VI.2.2. Quantum calculations

All quantum calculations were carried out with *Gaussian 09*,[31] as follows: *a)* optimisation of the **Mac**$_{methyl}$ halide complexes in gas-phase followed by the analysis of their ChB interactions; and *b)* the derivatisation of Restrained Electrostatic Potential (RESP) charges for the individual rotaxane components, **Mac** and the axle.

All calculations involving the **Mac**$_{methyl}$ model ligand were performed with M06-2X functional and using the 6-311++G** basis set for the H, C and N atoms, and the aug-cc-pVDZ-PP basis set for the Se centres, which was obtained from the EMSL database.[32, 33] In addition, in the optimisation of the halide complexes of **Mac**$_{methyl}$ the bromine and iodine centres were described with the aug-cc-pVDZ-PP basis set, and the aug-cc-pVDZ basis set for chlorine. The distribution of the electrostatic potential, $V(\mathbf{r})$, mapped onto the molecular surface ($V_S(\mathbf{r})$) of **Mac**$_{methyl}$ was computed through a single point on the previously optimised structures of **Mac**$_{methyl}$, free or chalcogen bonded to halides, after removal of the anions. In this work, $V(\mathbf{r})$ was evaluated on the 0.001 electrons $Bohr^{-3}$ contour of $\rho(\mathbf{r})$. The $V_S(\mathbf{r})$ ranges, between the most negative ($V_{S,min}$) and the most positive values ($V_{S,max}$), were calculated using *Multiwfn*.[34, 35] The Wiberg Bond Indices (WBI),[36] which consist in the sum of squares of off-diagonal density matrix elements between atoms, were ascertained from the DFT optimised structures of the halide complexes of **Mac**$_{methyl}$. These calculations were carried out using the *Natural Bond Orbital* (*NBO*) Version 3.1 program,[37, 38] within *Gaussian 09*.

VI.2.3. Classical force field calculations

Molecular Mechanics (MM) calculations and MD simulations were carried out with AMBER 2016.[39] The rotaxane components (axle and macrocycle) were described with parameters taken from the General AMBER Force Field (GAFF),[40, 41] except those involving the selenium centres, and RESP atomic charges.[42] The water molecules were described with the TIP3P model,[43] while the all-atoms model of the acetone solvent molecules was described with RESP atomic charges, GAFF bonding parameters and van der Waals (vdW) parameters taken from ref. 44. The halides were described with a discrete charge of −1 and vdW parameters developed for the TIP3P water model.[45] The force field parameters and charges of the PF_6^- counterion were taken from ref. 46. The post-processing of trajectory files to obtain the structural data was performed with *cpptraj*.[47]

VI.2.4. Selenium force field parameters

Due to the inexistence of Se parameters within the AMBER family of force fields, its bonding parameters were derived using the *MCPB.py*[48] and *paramfit*[49] utilities, available within the AMBER 2016 package. The selenium vdW parameters were taken from the Universal Force Field (UFF),[50] with the original radius (r_{vdw}) converted to the AMBER formalism as $r_{vdw} = r_{min}/2$, where r_{min} is 4.205 Å. The determination of the Se bonding parameters was carried out with the fragment **TrzMe$_3$SeMe** depicted in Figure VI-2, together with its corresponding GAFF atom types, attributed with *antechamber*,[51] apart of Se.

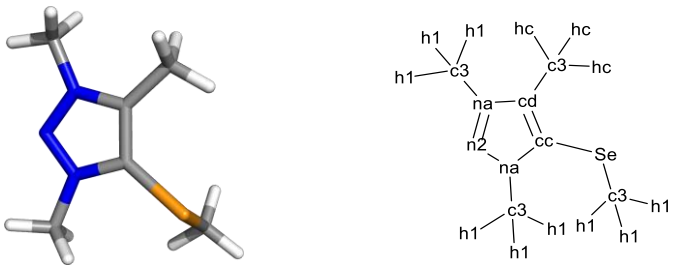


Figure VI-2. Fragment **TrzMe$_3$SeMe** (1,3,4-trimethyl-5-(methylselanyl)-1H-1,2,3-triazol-3-ium) used in the parametrisation of Se bonding terms along with the default GAFF atom types.

TrzMe$_3$SeMe was optimised at the B3LYP level, with the H, C and N atoms being described with the 6-311++G** basis set, while the Se atom was described with the aug-cc-pVDZ-PP basis set, obtained from the EMSL database. The force field parameters for the bond lengths and bond angles involving the Se–C bonds were generated from this optimised structure, using the Seminario method**** implemented in *MCPB.py*, and are listed in Table VI-2.

The *MCPB.py* utility does not generate torsion angle terms. Thus, the parameters for the torsion angles of the types Se–C–N–C, Se–C–N–N, Se–C–C–N, Se–C–C–C, C–Se–C–H, were directly obtained from GAFF considering the ss atom type instead of Se and are gathered in Table VI-2, while the torsion angle terms centred at the Se-cc bond were obtained with *paramfit*, following the standard workflow for the development of this kind of parameters. **TrzMe$_3$SeMe** was re-optimised at the MP2 level, with the H, C, and N atoms being described with the cc-pVDZ basis set, and Se was described with the aug-cc-pVDZ-PP basis set. Subsequently, a relaxed Potential Energy Surface scan was undertaken, with the torsion angles c3-Se-cc-na and c3-Se-cc-cd being shifted 1.8° 199 times, resulting in a 360° scan around that bond, allowing to discover the energy barriers depicted in Figure VI-3. Subsequently, the genetic and simplex algorithm of *paramfit*†††† was used to fit MM energies to the quantum data and obtain the torsion force field parameters listed in Table VI-2.

**** In the Seminario method, the bond and angle force constants are systematically derived from the sub matrices of the Cartesian Hessian matrix.[48]

†††† In the genetic and simplex algorithm, a population of possible parameters is created, fitted to quantum data and ranked. The best parameters are combined to create a new generation in a recombination operation, while some parameters in this new generation are randomly altered by a mutation operation. This new generation of parameters is then fitted and ranked. As this process is repeated over several generations, the best parameters are "evolved".[49]

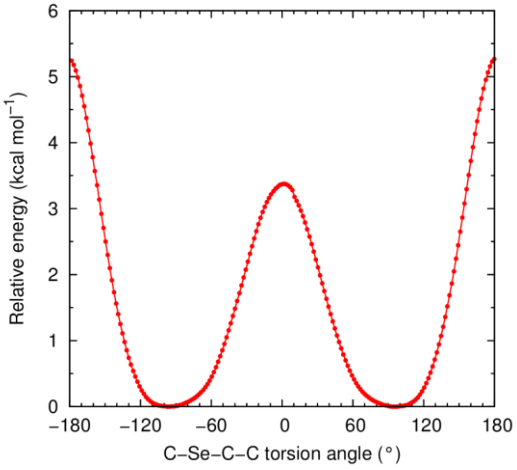


Figure VI-3. Molecular energy variation of **TrzMe$_3$SeMe** as a function of the variation of the C–Se–C–C torsion angle. The minima were found at torsion angles of ± 95.5°.

The parametrisation process that led to the development of the force field parameters incorporating the Se centre used in the classical force field calculations of **Mac** and **Rot** (see Table VI-2) is depicted in the flowchart of Figure VI-4.

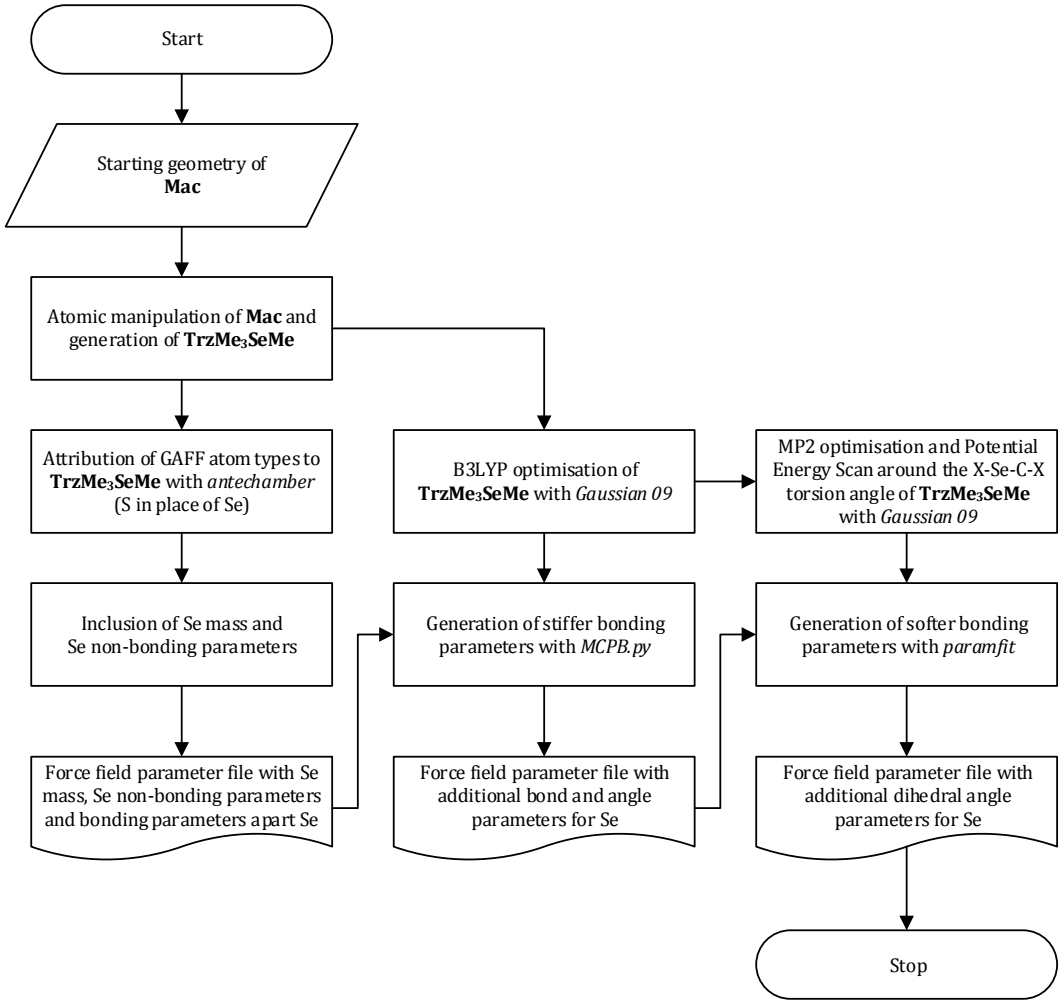


Figure VI-4. Flowchart depicting the most important steps in the development of Se bonding parameters.

Table VI-2. Selenium force field parameters used in this work.

Mass	Atomic weight (A_r)		Notes
Se	78.96		Taken from ref. 52
Bond length	K_r (kcal mol^{-1} Å^{-2})	r_{eq} (Å)	Notes
Se-c3	132.7	1.9874	Determined with *MCPB.py*
Se-cc	169.6	1.9160	Determined with *MCPB.py*
Se-cd	169.6	1.9160	Determined with *MCPB.py*
Bond angle	K_θ (kcal mol^{-1} rad^{-2})	θ_{eq} (°)	Notes
Se-c3-h1	39.96	108.28	Determined with *MCPB.py*
Se-cc-cd	104.37	129.50	Determined with *MCPB.py*
Se-cd-cc	104.37	129.50	Determined with *MCPB.py*
Se-cc-na	107.78	125.57	Determined with *MCPB.py*
Se-cd-na	107.78	125.57	Determined with *MCPB.py*
cc-Se-c3	147.13	98.98	Determined with *MCPB.py*
cd-Se-c3	147.13	98.98	Determined with *MCPB.py*

Torsion angles	Scaling factor	$V_n/2$ (kcal mol^{-1})	γ (°)	periodicity N	Notes
Se-cc-na-c3	1	1.70000000	180.000	2.0	GAFF parameters for ss-cc-na-c3
Se-cd-na-c3	1	1.70000000	180.000	2.0	GAFF parameters for ss-cd-na-c3
Se-cc-na-n2	1	1.70000000	180.000	2.0	GAFF parameters for ss-cc-na-n2
Se-cd-na-n2	1	1.70000000	180.000	2.0	GAFF parameters for ss-cd-na-n2
Se-cc-cd-na	1	4.00000000	180.000	2.0	GAFF parameters for ss-cc-cd-na
Se-cd-cc-na	1	4.00000000	180.000	2.0	GAFF parameters for ss-cd-cc-na
Se-cc-cd-ca	1	4.00000000	180.000	2.0	GAFF parameters for ss-cc-cd-ca
Se-cd-cc-ca	1	4.00000000	180.000	2.0	GAFF parameters for ss-cd-cc-ca
cc-Se-c3-h1	1	0.33333333	0.000	3.0	GAFF parameters for cc-ss-c3-h1
cd-Se-c3-h1	1	0.33333333	0.000	3.0	GAFF parameters for cd-ss-c3-h1
c3-Se-cc-na	1	0.62620000	180.000	2.0	Determined with *paramfit*
c3-Se-cd-na	1	0.62620000	180.000	2.0	Determined with *paramfit*
c3-Se-cc-cd	1	0.08700000	180.000	−2.0	Determined with *paramfit*
c3-Se-cc-cd	1	1.30670000	180.000	3.0	Determined with *paramfit*
c3-Se-cd-cc	1	0.08700000	180.000	−2.0	Determined with *paramfit*
c3-Se-cd-cc	1	1.30670000	180.000	3.0	Determined with *paramfit*

Improper torsion	K_Φ (kcal mol^{-1})	Φ_0 (°)	periodicity N	Notes
Se-cc-cd-na	1.10000000	180.000	2.0	GAFF parameters for ss-cc-cd-na
Se-cd-cc-na	1.10000000	180.000	2.0	GAFF parameters for ss-cd-cc-na

vdW	r (Å)	ε (kcal mol^{-1})	Notes
Se	2.1025	0.2910	Adapted from UFF.[50]

VI.2.5. Calculation of RESP charges of Rot components

The structure of **Mac**, obtained from its single crystal X-ray diffraction structure, was optimised using the B3LYP functional, followed by a single point calculation in HF, with 4 concentric layers per atom and 6 density points in each layer (IOp (6/33 = 2, 6/41 = 4, 6/42 = 6)), in agreement with GAFF's development.[40, 41] In both quantum calculations, the Se centres were treated with the aug-cc-pVDZ-PP basis set, while the remaining atoms were treated with the 6-31G* basis. The RESP charges of the axial component of **Rot** were obtained following the same methodology.

VI.2.6. Parametrisation of the ChB interactions

The optimised structures of **Mac**$_{methyl}$ with halides are shown in Figure VI-5 and the corresponding ChB dimensions are listed in Table VI-7 (*vide infra*). The parametrisation of the ChB interactions was based on the computed Se···A distances (A = Cl^-, Br^-, I^-) and C–Se···A angles of these optimised **Mac**$_{methyl}$ halide complexes. In addition, V(**r**), mapped onto the electron density surface of **Mac**$_{methyl}$ halide complexes, revealed the position of the $V_{S,max}$ with respect to each Se atom, *i.e.*, which is enclosed in the σ-hole. The Se···$V_{S,max}$ distances and the C–Se···$V_{S,max}$ angles are given in Table VI-3.

Table VI-3. Internal coordinates of each SeMe binding unit's $V_{S,max}$ in the optimised halide complexes of **Mac**$_{methyl}$.

Halide guest	Se···$V_{S,max}$ (Å)	$C_{triazolium}$–Se···$V_{S,max}$ (°)	C_{methyl}–Se···$V_{S,max}$ (°)
Cl^-	2.071 ; 2.118	166.1 ; 160.0	73.3 ; 69.6
Br^-	2.068 ; 2.094	164.9 ; 161.2	72.7 ; 69.6
I^-	2.063 ; 2.087	163.3 ; 161.5	71.4 ; 70.0

Similarly to our previous classical MD investigations on anion recognition by XB,[25, 27, 53, 54] the ChB interactions were represented by the addition of a massless extra point (EP) of charge to GAFF,[55] effectively representing the σ-hole found in front of each Se atom. Moreover, to accurately reproduce the slightly bent $C_{triazolium}$–Se···A angles observed in the DFT optimised models, the EP cannot make a 180° $C_{triazolium}$–Se–EP angle as it would be expected in XB interactions. Given that the equilibrium angle computed by *MCPB.py* for the $C_{triazolium}$–Se–C_{methyl} in fragment **TrzMe$_3$SeMe** of Figure VI-2 is of 98.98° (see Table VI-2), and the $C_{triazolium}$–Se···$V_{S,max}$ maximum value is 166° (see Table VI-3), then, considering that the $C_{triazolium}$, Se, C_{methyl} and $V_{S,max}$ are co-planar, the $C_{triazolium}$–Se–EP and the C_{methyl}–Se–EP angles were set to 166 and 67° (166 – 99 = 67°), respectively. This approach allowed to reduce the degrees of freedom in this parametrisation effort, limiting it to the optimisation of suitable Se–EP distances for subsequent MM and MD simulations, as follows.

Several Se–EP distances were systematically evaluated in gas-phase *via* MM energy minimisations of the halide complexes of **Rot**. For a given Se–EP distance, RESP atomic charges were re-calculated for the macrocyclic component of the rotaxane, using the following approach: in the previously B3LYP optimised structure of the macrocyclic component of **Rot**, an EP was positioned in front of each $C_{triazolium}$–Se bond at the fixed angles given above and at the desired Se–EP distances and then the corresponding atomic charges were obtained *via* a two-stage RESP fitting (*vide supra*). The ranges of Se–EP distances tested for the halide complexes are gathered in Table VI-4 (**Rot**·Cl^-), Table VI-5 (**Rot**·Br^-), and Table VI-6 (**Rot**·I^-), together with the MM optimised Se···A distances and the atomic charges for Se and EP.

Table VI-4. Summary of MM optimised distances for **Rot**·Cl− as function of the Se−EP distance, along with EP and Se RESP charges.

Distances (Å)		Charges *q(e)*	
Se−EP	Se···Cl⁻	EP	Se
No EP	3.609 ; 3.648	–	0.074607
2.57	3.420 ; 3.406	0.039976	−0.000261
2.58	3.418 ; 3.401	0.039410	0.000851
2.59	3.415 ; 3.396	0.038846	0.001964
2.60	3.412 ; 3.390	0.038284	0.003079
2.61	3.408 ; 3.384	0.037722	0.004195
2.62	3.404 ; 3.377	0.037162	0.005312
2.63	3.400 ; 3.369	0.036603	0.006430
2.64	3.306 ; 3.456	0.036044	0.007550
2.65	3.306 ; 3.436	0.035487	0.008671
2.66	3.384 ; 3.335	0.034930	0.009793
2.67	3.294 ; 3.360	0.034374	0.010917
2.68	3.262 ; 3.292	0.033819	0.012041
2.69	Atomic clash	0.033265	0.013165

Table VI-5. Summary of MM optimised distances for **Rot**·Br− as function of the Se−EP distance, along with EP and Se RESP charges.

Distances (Å)		Charges *q(e)*	
Se−EP	Se···Br⁻	EP	Se
No EP	3.656 ; 4.031	–	0.074607
2.79	3.649 ; 3.643	0.027828	0.024263
2.80	3.649 ; 3.642	0.027308	0.025322
2.81	3.649 ; 3.641	0.026798	0.026359
2.82	3.649 ; 3.640	0.026299	0.027371
2.83	3.648 ; 3.638	0.025814	0.028350
2.84	3.648 ; 3.636	0.025346	0.029292
2.85	3.647 ; 3.634	0.024899	0.030188
2.86	3.646 ; 3.631	0.024474	0.031031
2.87	3.645 ; 3.627	0.024077	0.031813
2.88	3.643 ; 3.622	0.023711	0.032525
2.89	3.512 ; 3.749	0.023381	0.033160
2.90	3.501 ; 3.751	0.023089	0.033708
2.91	3.483 ; 3.753	0.022840	0.034165
2.92	3.436 ; 3.754	0.022634	0.034525
2.93	Atomic clash	0.022474	0.034785

Table VI-6. Summary of MM optimised distances for **Rot**·I^- as function of the Se–EP distance, along with EP and Se RESP charges.

Distances (Å)		Charges *q(e)*	
Se–EP	Se···I^-	EP	Se
No EP	4.003 ; 4.060	–	0.074607
3.05	3.922 ; 3.912	0.022963	0.032667
3.06	3.919 ; 3.905	0.023046	0.032417
3.07	3.915 ; 3.896	0.023114	0.032208
3.08	3.910 ; 3.887	0.023164	0.032048
3.09	3.906 ; 3.878	0.023194	0.031940
3.10	3.901 ; 3.867	0.023203	0.031891
3.11	3.749 ; 4.014	0.023189	0.031901
3.12	3.727 ; 4.013	0.023153	0.031973
3.13	3.680 ; 4.011	0.023093	0.032108
3.14	Atomic clash	0.023010	0.032305

Table VI-4 through Table VI-6 show that several Se–EP distances can be applied in the MM optimisation of the halide complexes of **Rot**. The use of the longer 2.69, 2.93 and 3.14 Å Se–EP distances lead to atomic clashes in the **Rot**·Cl^-, **Rot**·Br^-, and **Rot**·I^- complexes, in this order. Thus, the subsequent values of 2.68, 2.92 and 3.13 Å Se–EP distances allow the optimisation of these complexes in gas-phase, leading to Se···A distances closer to the ones assessed in the **Mac**$_{methyl}$ model complexes by DFT calculations (see Table VI-7). Unfortunately, further MD simulations of these complexes, carried out in gas-phase at 300 K for 5.0 ns, were shown to be instable, resulting in the exclusion of those Se–EP distances, leading to the testing of shorter values until stable MD simulations were accomplished. The 2.57, 2.79 and 3.05 Å Se–EP distances were found to allow stable MD simulations with **Rot**·Cl^-, **Rot**·Br^-, and **Rot**·I^-, in this order. The average Se···A distances 3.455 ± 0.133 and 3.484 ± 0.131 Å were calculated for **Rot**·Cl^-, while for the other complexes the following values were assessed: 3.701 ± 0.141 and 3.721 ± 0.148 Å for **Rot**·Br^-, and 3.948 ± 0.139 and 3.990 ± 0.144 Å for **Rot**·I^-. Therefore, the 2.57, 2.79 and 3.05 Å Se–EP distances were definitively appointed for the subsequent MD simulations in the acetone/water solvent mixtures.

VI.2.7. General MD simulation methods

The starting geometries of the halide complexes of **Mac** or **Rot**, as well as the corresponding free hosts, were minimised in gas-phase by MM until the convergence criterion of 0.0001 kcal mol^{-1} $Å^{-1}$ was achieved. Afterwards, the optimised structures were solvated in cubic boxes with different number of solvent molecules, in agreement with the solvent mixtures used in ^1H NMR experimental binding studies (see Table VI-1): for **Mac** – 1005 acetone molecules and 84 water molecules, for a 98:2 *v*/*v* acetone/water solvent mixture; for **Rot** – 1802 acetone molecules and 1836 water molecules, corresponding to a 4:1 *v*/*v* acetone/water solvent mixture. A variable number of PF_6^- counterions was added to neutralise each system's net charge.

Subsequently, each solvated system was equilibrated under periodic boundary conditions using the following multistage protocol. The system was relaxed by MM minimisation of solvent molecules and by keeping the solutes fixed with a positional restraint of 500 kcal mol^{-1} $Å^{-2}$. The restraint was then removed, allowing the entire system to relax. These two minimisation stages comprised an initial set of 10000 steepest descent algorithm steps, followed by 10000 steps of conjugated gradient algorithm. The equilibration proceeded with heating up the system to 300 K for 100 ps using a NVT ensemble and a weak positional restraint (10 kcal mol^{-1} $Å^{-2}$) on the solutes. Afterwards, each system's density was allowed to equilibrate in a NPT ensemble at 1 atm for 1.0 ns, at the same temperature, followed by a NPT

data collection run of 50 ns. The collection run's trajectory frames were saved every 1.0 ps. Three independent runs were performed for each system. The CUDA version of the *PMEMD* executable was used for the simulation of all solvated systems.[56-58] The bond lengths involving all hydrogen atoms were constrained with the SHAKE algorithm, allowing the usage of a 2.0 fs time step.[59] The Particle Mesh Ewald (PME) method was used to treat the long-range electrostatic interactions.[60] The non-bonded van der Waals interactions were truncated with a 10 Å cut-off. The temperature of the system was maintained at 300 K, using Langevin dynamics,[61] with a collision frequency γ of 1.0 ps^{-1}. The pressure was controlled by the Berendsen barostat[62] at 1 atm and compressibility of 44.6×10^{-6} bar^{-1}, with a relaxation time of 1.0 ps.

VI.3. Results and Discussion

The DFT optimised structures of the free receptor model **Mac**$_{methyl}$ (the macrocycle's benzene-1,3-bis(methylseleno-triazolium) binding core capped with two methyl groups) and its halide complexes are presented in Figure VI-5, along with their $V_S(\mathbf{r})$. The relevant ChB dimensions for the **Mac**$_{methyl}$·A (A = Cl^-, Br^-, or I^-) complex models and their individual $V_{S,max}$ values are summarised in Table VI-7.

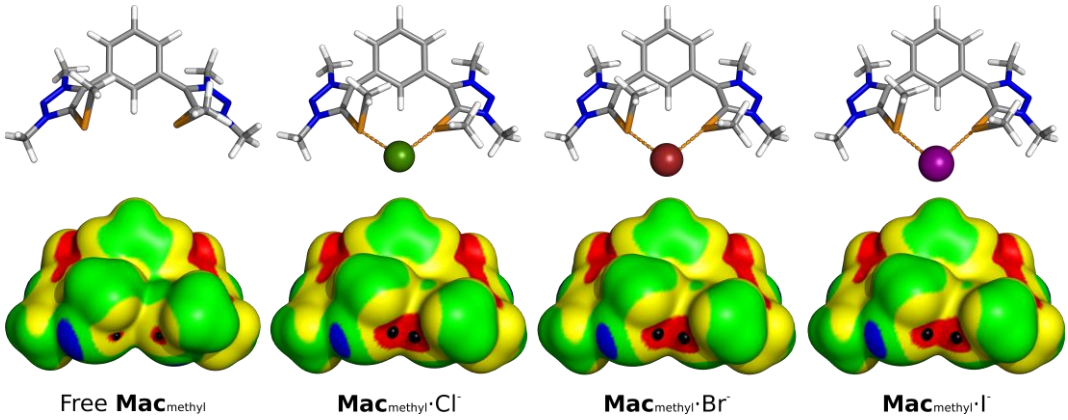


Figure VI-5. Structural and electronic features obtained by DFT calculations on the free **Mac**$_{methyl}$ and on its halide complexes, with the optimised structures (top) and the $V(\mathbf{r})$ on the $\rho(\mathbf{r})$ = 0.001 electrons $Bohr^{-3}$ surface of **Mac**$_{methyl}$ free or after removal of the halide (bottom). The chalcogen bonds are drawn as orange dashed lines, and the halides are drawn as a green (Cl^-), brown (Br^-), or purple (I^-) spheres. The location of each $V_{S,max}$ is represented as a black dot, and corresponds to the σ-hole in front of each SeMe binding unit. The colour ranges, in kcal mol^{-1}, are as follows: blue – below 112.5; green – between 112.5 and 127.5; yellow –between 127.5 and 142.5; red – above 142.5.

Table VI-7. ChB dimensions and $V_{S,max}$ values of the SeMe binding units of **Mac**$_{methyl}$ in the unbound and halide complexes.

Complex	Se···A (Å)[a]	$C_{triazolium}$–Se···A (°)[a]	$V_{S,max}$ (kcal mol^{-1})
Cl^-	2.804 ; 2.870	175.3 ; 168.5	154.76 ; 154.47
Br^-	2.963 ; 3.015	176.3 ; 170.2	153.78 ; 153.46
I^-	3.166 ; 3.244	177.1 ; 171.3	152.69 ; 152.18
Unbound	–	–	146.07 ; 145.07

[a] A stands for Cl^-, Br^- or I^-.

The model receptor **Mac**$_{methyl}$ binds to each halide anion *via* two cooperative ChB interactions differing slightly in lengths that increase with the size of the halide (Cl^-: 2.804, 2.870 Å; Br^-: 2.963, 3.015 Å; and I^-: 3.166, 3.244 Å). In addition, the $C_{triazolium}$–Se···A bond angles range from 168.5 to 177.1°, which deviate only slightly from linearity, akin to XB, the sister σ-hole interaction. For free **Mac**$_{methyl}$, the distribution of the electrostatic potential mapped onto the electron density surface shows two narrow separated positive regions (in red) which surround the most electron-deficient site (in black) located at the termini of each $C_{triazolium}$–Se bond (see Figure VI-5, top right). Interestingly, the presence of the halide induces a greater electron deficiency at the binding site, resulting in an increase in the $V_{S,max}$ of approximately 7.99 kcal mol^{-1} on the Se atoms compared to the free receptor. This structural analysis clearly suggests that the ChB interactions can be treated using a classical force field approach equivalent to the one extensively used by our group for the anion recognition mediated by the sister σ-hole-based XB interaction.[25, 27, 53, 54]

The covalent nature of the ChB–halide interactions was evaluated through the calculation of the WBI[36] from the DFT optimised structures of the **Mac**$_{methyl}$·A complexes. Generally, heavier halides gave larger independent WBI values, growing in the order Cl^- (0.189; 0.153) < Br^- (0.199; 0.167) < I^- (0.213; 0.167), with the values in parentheses representing the WBI for each ChB–halide interaction, which suggests that the ChB interactions with the heavier halide have a larger degree of covalent character. This result is entirely consistent with the greater anion affinities of **Mac** and **Rot** observed for the heavier halides (see Table VI-1).

After demonstrating the linearity of the ChB–anion interactions, classical MD simulations were performed to investigate the anion complexes with receptors **Mac** and **Rot** in their respective solvent mixtures. The ChB interactions were simulated using an EP of charge to represent the σ-hole found in front of each C–Se activated bond as shown in Figure VI-5, in a similar way as our previous work on XB anion recognition by interlocked host systems.[25, 27, 53, 54] In addition, force field bonding parameters involving the Se centre were developed to be used together with GAFF.[40, 41]

The interaction between the SeMe binding units and the halide (A) guests was evaluated throughout the simulation time with the assessment of the individual Se···A distances and $C_{triazolium}$–Se···A angles. In addition, the distance between each halide and the centre of mass (COM) of the macrocycle, considering only the non-hydrogen atoms (MAC_{cent}), was monitored to evaluate the position of the anion relatively to the host's binding cavity. The average values of these structural parameters for three independent MD runs are listed in Table VI-8. Moreover, in the rotaxane complexes, the recognition of halide guests is assisted by a hydrogen bonding interaction with the axle triazolium proton, which was simulated using weak harmonic restraints on the C···A distance and C–H···A angle.[23] Therefore, the dimensions of this singular HB interaction, also listed in Table VI-8, are not discussed.

Figure VI-6 shows representative snapshots of the **Rot**·A complexes taken from the MD runs. The average A···MAC_{cent} distances for these complexes indicate that the anion binding of the three halides occurs above the plane of the macrocycle, with chloride closest to the rotaxane's binding pocket, and the larger bromide and iodide anions being progressively further away. Moreover, the halide anions are kept bonded to the interpenetrated host by two convergent ChB interactions along the three MD runs, which are occasionally interrupted, as suggested by the small standard deviations of the Se···A distances gathered in Table VI-8. The two ChB–halide interactions are nearly linear throughout the simulation time of the three independent runs, with the average $C_{triazolium}$–Se···A angles around 170°. Generally, the average Se···A distances mirror the trend calculated from the DFT optimised structures of the halide model complexes with **Mac**$_{methyl}$ (see Table VI-7).

In the three MD runs of the macrocycle complexes, the two simultaneous ChB interactions to the halide guests were also maintained (see Table VI-8), with larger average Se···A distances and

standard deviations than those observed in the interlocked complexes. This comparison suggests that the chalcogen bonds of the **Mac**·A complexes were more frequently broken. Moreover, the average A···MAC$_{cent}$ distances show that the anions are clearly positioned above the macrocycle plane throughout most of the simulation time, exposed to the solvent molecules, as illustrated in Figure VI-6.

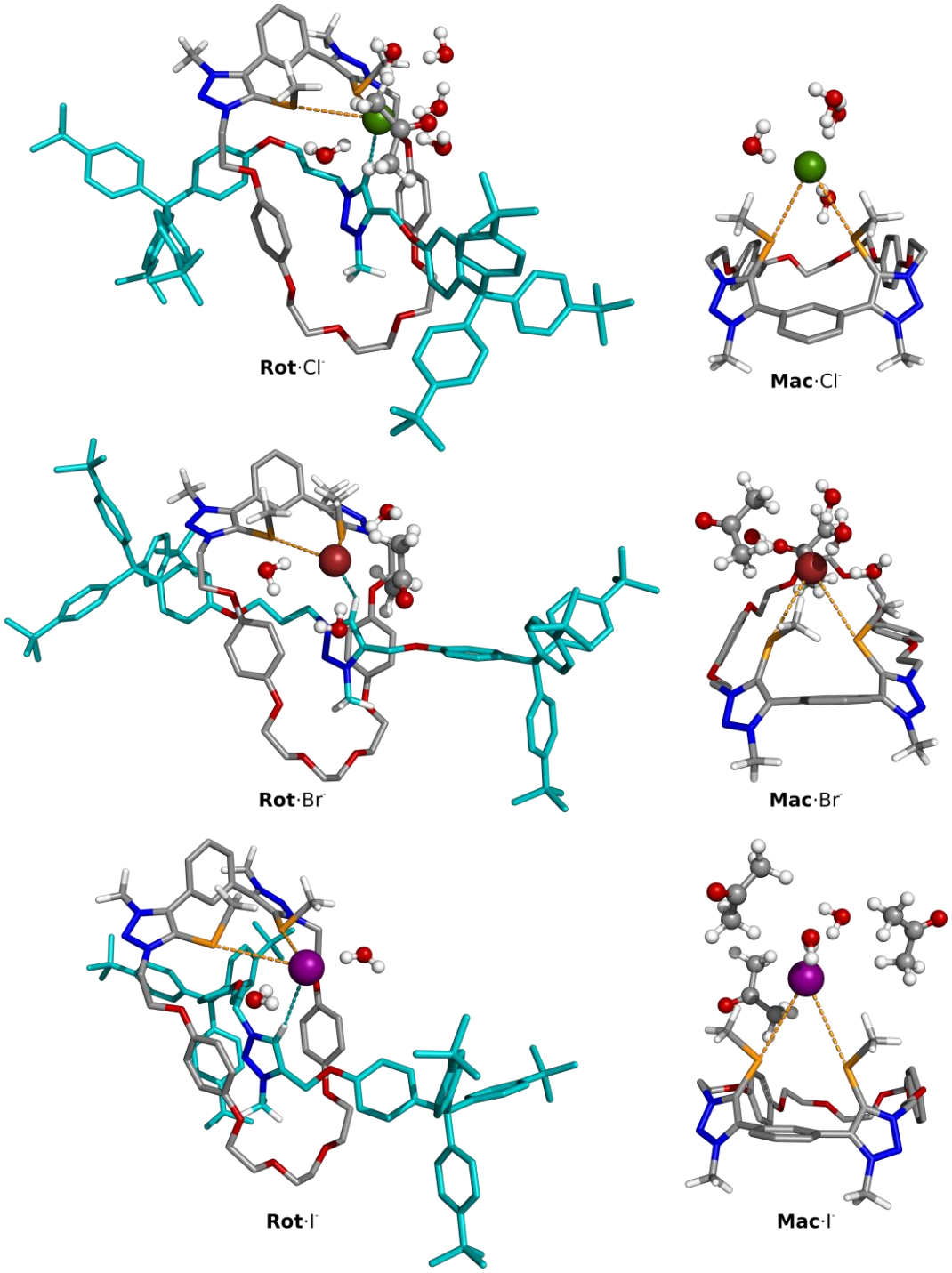


Figure VI-6. Illustrative snapshots of the **Rot**·A complexes (left) and **Mac**·A complexes (right), showing the anions surrounded by several solvent molecules. The anions are bound by two chalcogen bonds (orange dashed lines), which are assisted by a single $C_{triazolium}$–H···A hydrogen bond (teal dashed line) in the rotaxane.

Table VI-8. Average ChB and HB dimensions obtained from MD simulations of the halide complexes of **Mac** and **Rot**, along with the distances between the halide guests and the centre of mass of the macrocycle in either system.[a,b,c]

Anions	Rot					Mac		
	Se···A	$\angle C_{triazolium}$–Se···A	C_{axle}···A	$\angle C_{axle}$–H···A	A···MAC$_{cent}$	Se···A	$\angle C_{triazolium}$–Se···A	A···MAC$_{cent}$
Cl^-	3.712 ± 0.214 ; 3.700 ± 0.213	170.3 ± 4.0 ; 170.7 ± 3.9	3.765 ± 0.188	134.5 ± 15.2	4.462 ± 0.520	3.759 ± 0.240 ; 3.740 ± 0.232	169.0 ± 4.6 ; 169.3 ± 4.5	5.714 ± 0.582
Br^-	4.017 ± 0.261 ; 4.027 ± 0.268	170.3 ± 4.6 ; 170.2 ± 4.6	3.941 ± 0.191	135.8 ± 14.1	4.805 ± 0.536	4.070 ± 0.313 ; 4.088 ± 0.318	168.2 ± 5.6 ; 167.9 ± 5.7	6.040 ± 0.673
I^-	4.317 ± 0.274 ; 4.316 ± 0.280	169.6 ± 5.1 ; 169.8 ± 5.1	4.141 ± 0.193	135.9 ± 15.9	5.163 ± 0.644	4.422 ± 0.332 ; 4.403 ± 0.323	167.1 ± 6.3 ; 167.5 ± 6.1	6.380 ± 0.676

[a)] The values for the individual MD runs are given in Table D-1 to Table D-4 (see Appendix D). All distances and angles are given in Å and deg, respectively; [b)] A stands for Cl^-, Br^- or I^-; [c)] N = 150000.

Table VI-9. Number of solvent (water and acetone) molecules around the SeMe binding units and halides in MD simulations of free and anion-Complexed **Rot** or **Mac**.[a,b]

Complex	Rot				Mac			
	SeMe binding units		Halide guest		SeMe binding units		Halide guest	
	No of water molecules	No of acetone molecules	No of water molecules	No of acetone molecules	No of water molecules	No of acetone molecules	No of water molecules	No of acetone molecules
	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD
Unbound	0.3 ± 0.6	1.1 ± 0.9	–	–	0.1 ± 0.2	2.4 ± 1.1	–	–
Cl^-	0.8 ± 0.8	0.3 ± 0.5	4.0 ± 0.9	0.5 ± 0.7	0.9 ± 0.8	1.5 ± 0.9	3.3 ± 1.0	2.1 ± 1.1
Br^-	0.5 ± 0.7	0.5 ± 0.6	3.6 ± 1.0	0.9 ± 0.8	0.5 ± 0.6	1.6 ± 0.9	2.4 ± 1.0	2.5 ± 1.1
I^-	0.2 ± 0.5	0.7 ± 0.6	2.4 ± 1.1	1.0 ± 0.9	0.2 ± 0.4	1.8 ± 0.9	1.0 ± 0.9	2.3 ± 1.2

[a)] The values for the individual MD runs are given in Table D-5 and Table D-6(see Appendix D); [b)] N = 150000.

The solvation spheres of the anions and of the SeMe binding units were also evaluated within a 3.5 Å radius‡‡‡‡ to ascertain the role of the water and acetone molecules on the binding selectivity of **Mac** and **Rot** towards the halide guests (see Table VI-9). The ChB-donor SeMe groups of both free host molecules are preferentially solvated by acetone molecules, with very few water molecules (≤0.3) surrounding them. In stark contrast, upon binding of the halide anions, the average number of water molecules around the SeMe groups rises in both host systems. In the complex of the smaller and more hydrophilic chloride, the anion leads to a larger number of water molecules around the SeMe groups, accompanied by few acetone molecules. Consequently, this increases the competition between the more hydrated halides and water molecules for the SeMe binding sites, making the disruption of the ChB–anion interactions, due to solvating water molecules, more likely. The hydration of the halide anions, in agreement with their lipophilic character, follows the trend $Cl^- > Br^- > I^-$, which are inverse to the trend of experimental anion binding constants reported in Table VI-1. Indeed, when the hydration of the SeMe binding units or of the halide guests is plotted against the logarithm of the anion association constants in Figure VI-7, highly linear relations of $R^2 \geq 0.94$ appear.

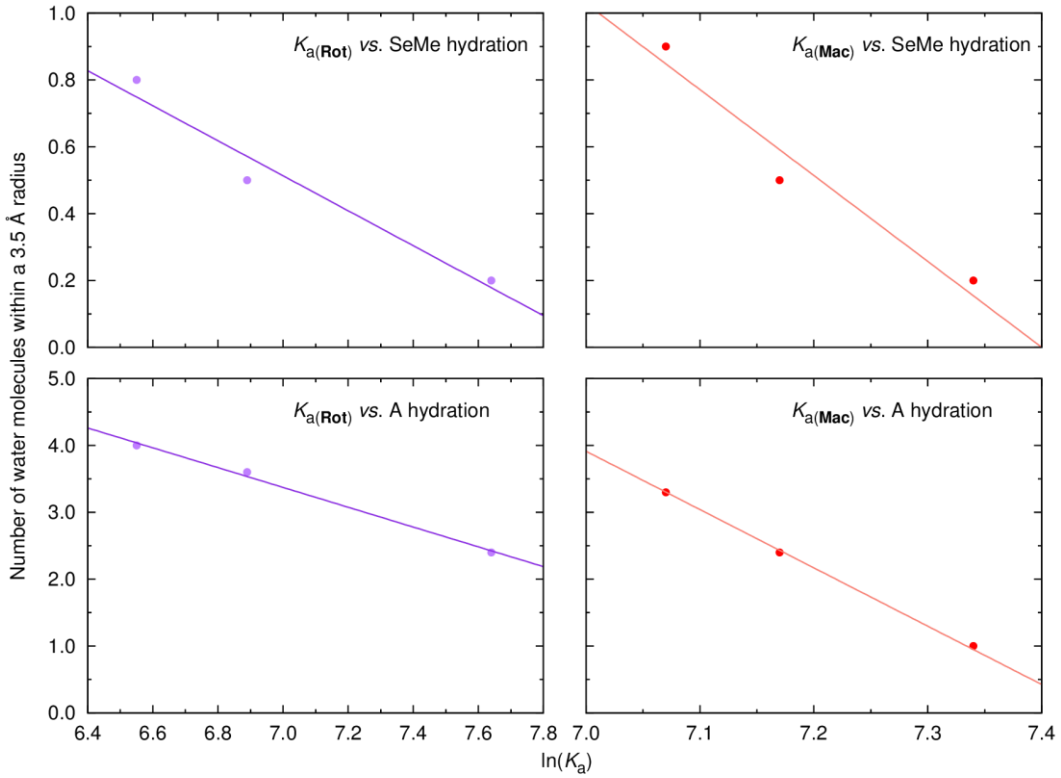


Figure VI-7. Hydration of the SeMe binding units (top) and halide guests (bottom) in the MD simulations of **Rot**·A (left) and **Mac**·A (right) complexes as a function of the logarithm of the experimental anion binding associations (see Table VI-1). The linear regressions yielded the following R^2 values: 0.96 ($K_{a(Rot)}$ *vs.* SeMe hydration), 0.94 ($K_{a(Mac)}$ *vs.* SeMe hydration), 1.00 ($K_{a(Rot)}$ *vs.* anion hydration), and 1.00 ($K_{a(Mac)}$ *vs.* anion hydration).

Finally, the MD simulation results also offer insights on the observed superiority of rotaxane **Rot** over macrocycle **Mac** for anion binding, seen in Table VI-1 where comparable binding affinities were obtained for both ChB hosts despite the much greater competitiveness of the solvent used for **Rot**. As aforementioned, **Rot** is able to host the anions within its three-dimensional binding pocket, with shorter A···MAC$_{cent}$ distances as compared to **Mac** (see Table VI-8). As shown in Table VI-9, this not only results in more extensive halide dehydration upon binding, as the anions are less exposed to the water

‡‡‡‡ Typical radius for the first hydration shell of biomolecules.[63]

molecules, but also in a smaller resulting degree of hydration of each host's binding cavity (due to a smaller number of water molecules around the SeMe groups). Indeed, the average number of water molecules enclosing each free anion in the two solvent mixtures was also assessed by MD simulations, being 4.7 ± 0.9 (Cl^-), 3.5 ± 1.2 (Br^-), and 1.6 ± 1.0 (I^-) in the acetone/water 98:2 *v*/*v* solvent mixture used in the binding studies of **Mac.** On the other hand, the ChB–anion interactions of **Rot** are subjected to less extensive interruptions by the water molecules in close proximity, and hence lead to higher anion affinities compared to **Mac**, as the free anion guests in the acetone/water 4:1 *v*/*v* solvent mixture are solvated by an average of 6.9 ± 1.1 (Cl^-), 6.3 ± 1.2 (Br^-), and 4.8 ± 1.3 (I^-) water molecules. Naturally, other structural factors also contribute to the superior halide binding of **Rot**, including the higher net charge, the hydrogen bond between the axle triazolium proton and halides, as well as the preorganised binding cavity of this rotaxane-host system. It is notable, however, that these computational findings provide the first glimpse of the important roles played by the solvent, particularly water, in determining the anion binding affinity and selectivity of ChB host systems.

VI.4. Conclusions

The theoretical results reported in this chapter illustrate how classical MD simulations in solution, associated with adequate DFT calculations, can be used as a supportive tool to rationalise the experimental binding data. Indeed, DFT calculations and MD simulations in aqueous solvent mixtures indicate that the selectivity is determined by the different hydrophilic characters of the anions allied to the hydration of the binding units in the presence of the anions.

To allow the simulation of the specific systems studied in this work, an unprecedented parameterisation effort was undertaken. This was accomplished employing recent methodologies to derive the bond stretching and bond angle bending (stiffer) parameters – using *MCPB.py*,[48] and relevant torsional angle (softer) parameters – using *paramfit*,[49] pertinent to the Se centres. Furthermore, the vdW parameters for Se were taken and adapted from the well-established UFF.[50] Afterwards, the parameterisation of the ChB interactions resorted to the EP of charge approach, which has been successfully applied to the more directional XB interactions.[25, 27, 53, 54]

This pioneering work in the development of bonding parameters for Se, as well as the approach devised to model its ChB interactions, opens a new avenue on the modelling of the increasingly important systems incorporating Se centres, both biological (selenoproteins in diseases afflicting the immune, endocrine or central nervous systems[64]) and synthetic (molecules with anti-oxidant,[65] anti-cancer,[66] anti-microbial,[67, 68] or anion transport properties[12]).

VI.5. References

1. L. C. Gilday, S. W. Robinson, T. A. Barendt, M. J. Langton, B. R. Mullaney and P. D. Beer, *Chem. Rev.*, 2015, **115**, 7118-7195.
2. G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati and G. Terraneo, *Chem. Rev.*, 2016, **116**, 2478-2601.
3. A. Bauza, I. Alkorta, A. Frontera and J. Elguero, *J. Chem. Theory Comput.*, 2013, **9**, 5201-5210.
4. A. Bauza, D. Quinonero, P. M. Deya and A. Frontera, *Crystengcomm*, 2013, **15**, 3137-3144.
5. P. Nziko Vde and S. Scheiner, *J. Phys. Chem. A*, 2014, **118**, 10849-10856.
6. A. Bauza, T. J. Mooibroek and A. Frontera, *Chemphyschem*, 2015, **16**, 2496-2517.
7. V. B. Birman and X. Li, *Org. Lett.*, 2006, **8**, 1351-1354.
8. C. A. Leverett, V. C. Purohit and D. Romo, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 9479-9483.
9. S. Fukumoto, T. Nakashima and T. Kawai, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 1565-1568.
10. A. F. Cozzolino, P. J. W. Elder and I. Vargas-Baca, *Coord. Chem. Rev.*, 2011, **255**, 1426-1438.
11. J. Fanfrlik, A. Prada, Z. Padelkova, A. Pecina, J. Machacek, M. Lepsik, J. Holub, A. Ruzicka, D. Hnyk and P. Hobza, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 10139-10142.

12. S. Benz, M. Macchione, Q. Verolet, J. Mareda, N. Sakai and S. Matile, *J. Am. Chem. Soc.*, 2016, **138**, 9093-9096.
13. A. F. Cozzolino, P. S. Whitfield and I. Vargas-Baca, *J. Am. Chem. Soc.*, 2010, **132**, 17265-17270.
14. A. Kremer, A. Fermi, N. Biot, J. Wouters and D. Bonifazi, *Chem. Eur. J.*, 2016, **22**, 5665-5675.
15. P. C. Ho, P. Szydlowski, J. Sinclair, P. J. Elder, J. Kubel, C. Gendy, L. M. Lee, H. Jenkins, J. F. Britten, D. R. Morim and I. Vargas-Baca, *Nat Commun*, 2016, **7**, 11299.
16. Y. Nagao, T. Hirata, S. Goto, S. Sano, A. Kakehi, K. Iizuka and M. Shiro, *J. Am. Chem. Soc.*, 1998, **120**, 3104-3110.
17. T. M. Beale, M. G. Chudzinski, M. G. Sarwar and M. S. Taylor, *Chem. Soc. Rev.*, 2013, **42**, 1667-1680.
18. A. Brown and P. D. Beer, *Chem. Commun.*, 2016, **52**, 8645-8658.
19. H. Zhao and F. P. Gabbai, *Nat Chem*, 2010, **2**, 984-990.
20. G. E. Garrett, G. L. Gibson, R. N. Straus, D. S. Seferos and M. S. Taylor, *J. Am. Chem. Soc.*, 2015, **137**, 4126-4133.
21. G. E. Garrett, E. I. Carrera, D. S. Seferos and M. S. Taylor, *Chem. Commun.*, 2016, **52**, 9881-9884.
22. T. Chivers and R. S. Laitinen, *Chem. Soc. Rev.*, 2015, **44**, 1725-1739.
23. J. Y. Lim, I. Marques, A. L. Thompson, K. E. Christensen, V. Felix and P. D. Beer, *J. Am. Chem. Soc.*, 2017, **139**, 3122-3133.
24. H. Park, L. J. Edgar, M. A. Lumba, L. M. Willis and M. Nitz, *Org Biomol Chem*, 2015, **13**, 7027-7033.
25. M. J. Langton, S. W. Robinson, I. Marques, V. Felix and P. D. Beer, *Nat Chem*, 2014, **6**, 1039-1043.
26. J. Y. Lim and P. D. Beer, *Chem. Commun.*, 2015, **51**, 3686-3688.
27. M. J. Langton, I. Marques, S. W. Robinson, V. Felix and P. D. Beer, *Chem. Eur. J.*, 2016, **22**, 185-192.
28. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
29. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2016, **72**, 171-179.
30. N. L. Kilah, M. D. Wise, C. J. Serpell, A. L. Thompson, N. G. White, K. E. Christensen and P. D. Beer, *J. Am. Chem. Soc.*, 2010, **132**, 11893-11895.
31. Gaussian 09, Revision A.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
32. D. Feller, *J. Comput. Chem.*, 1996, **17**, 1571-1586.
33. K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li and T. L. Windus, *J. Chem. Inf. Model.*, 2007, **47**, 1045-1052.
34. T. Lu and F. Chen, *J. Mol. Graph. Model.*, 2012, **38**, 314-323.
35. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
36. K. B. Wiberg, *Tetrahedron*, 1968, **24**, 1083-1096.
37. NBO Version 3.1, E. D. Glendening, A. E. Reed, J. E. Carpenter, and F. Weinhold.
38. A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899-926.
39. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, 3rd, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.
40. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
41. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
42. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
43. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
44. S. K. Burger and G. A. Cisneros, *J. Comput. Chem.*, 2013, **34**, 2313-2319.
45. P. Li, L. F. Song and K. M. Merz, Jr., *J. Chem. Theory Comput.*, 2015, **11**, 1645-1657.
46. Z. P. Liu, S. P. Huang and W. C. Wang, *J. Phys. Chem. B*, 2004, **108**, 12978-12989.
47. D. R. Roe and T. E. Cheatham, 3rd, *J. Chem. Theory Comput.*, 2013, **9**, 3084-3095.
48. P. Li and K. M. Merz, Jr., *J. Chem. Inf. Model.*, 2016, **56**, 599-604.
49. R. M. Betz and R. C. Walker, *J. Comput. Chem.*, 2015, **36**, 79-87.
50. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024-10035.
51. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247-260.

52. J. R. de Laeter, J. K. Böhlke, P. De Bièvre, H. Hidaka, H. S. Peiser, K. J. R. Rosman and P. D. P. Taylor, *Pure Appl. Chem.*, 2003, **75**.
53. T. A. Barendt, A. Docker, I. Marques, V. Felix and P. D. Beer, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 11069-11076.
54. J. Y. Lim, I. Marques, L. Ferreira, V. Felix and P. D. Beer, *Chem. Commun.*, 2016, **52**, 5527-5530.
55. M. A. Ibrahim, *J. Comput. Chem.*, 2011, **32**, 2564-2574.
56. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
57. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
58. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
59. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
60. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
61. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
62. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
63. D. Laage, T. Elsaesser and J. T. Hynes, *Chem. Rev.*, 2017, **117**, 10694-10725.
64. B. Moghadaszadeh and A. H. Beggs, *Physiology (Bethesda)*, 2006, **21**, 307-315.
65. G. K. Azad and R. S. Tomar, *Mol. Biol. Rep.*, 2014, **41**, 4865-4879.
66. Y. C. Chen, K. S. Prabhu and A. M. Mastro, *Nutrients*, 2013, **5**, 1149-1168.
67. P. A. Tran and T. J. Webster, *Nanotechnology*, 2013, **24**, 155101.
68. E. Cremonini, E. Zonaro, M. Donini, S. Lampis, M. Boaretti, S. Dusi, P. Melotti, M. M. Lleo and G. Vallini, *Microb Biotechnol*, 2016, **9**, 758-771.

Chapter VII.
Conclusions & Future perspectives

VII.1. Main findings

The complexity of the investigations presented increased along the thesis, from the straightforward passive diffusion Molecular Dynamics (MD) simulations, in which the interactions between the transporters and the POPC membrane models were investigated, to the more complex constrained MD simulations (Steered Molecular Dynamics, SMD, and Umbrella Sampling, US, simulations) carried out to estimate the energy barriers associated with the translocation of chloride complexes across phospholipid bilayers as well as to bring critical insights to the putative anion transport mechanisms.

The work in Chapter II focused on the interaction of six tripodal molecules with a POPC bilayer and how the chloride transporters impacted the biophysical properties of the membrane model. The N–H binding units were found to be crucial for the recognition of chloride as well as the POPC phosphate head groups. Moreover, the receptors adopted well-defined orientations in the bilayer, consistent with the hydrophobicity of their components. The structural findings obtained by a combination of quantum calculations and MD simulations are consistent with the experimental transport activity trends previously reported for this series of tripod molecules.

In Chapter III, the work developed with the six bis-thioureidodecalin derivatives produced valuable insights on how extending the length of the alkyl chains modulated transport activity. The substituents with different lipophilicities led to different interactions with the phospholipid tails and these interactions mediated the (re)orientation of the anion transporters at the water/lipid interface level. These factors, more than the binding affinity, explained the "bell-shaped" curve obtained when the experimental anion transport activity is plotted against the size of the decalin's alkyl substituents. Indeed, the energetic and structural data obtained with this theoretical investigation unequivocally showed that the transport activity of the small anion carriers depends on a delicate balance between anion affinity and the transporters' lipophilicity.

The squaramide binding motif was comprehensively studied in Chapter IV, with four different scaffolds and substituents with three degrees of fluorination, in twelve transporters. The passive diffusion MD simulations showed that molecules with two or three binding motifs hardly release the anion when the complex is initially placed within the bilayer, while molecules with a single binding unit were shown to be able to release and uptake chloride at the water/lipid interface. The energy profiles associated with the diffusion of the simpler squaramide derivatives across the bilayer were also estimated and found to be consistent with the experimental anion transport data. Moreover, taken together, all these insights suggest that the anion transport mediated *via* a carrier mechanism occurs with the transporter shuttling between opposite interfaces, without diffusion to bulk water.

The ten compounds presented in Chapter V allowed a detailed study of lipophilic balance in a series of tripodal molecules. Passive diffusion MD simulations showed that the length and fluorination degree of the alkyl chains have a greater influence over the orientation of the tripodal receptor than over its binding affinity. Moreover, constrained MD simulations showed at the atomistic level, for the first

time, a neutral transporter facilitating the translocation of chloride across a POPC bilayer. On the other hand, the energy profiles estimated for several analogous molecules, with different lipophilicities or binding affinities, are in perfect agreement with the anion transport data, both in vesicles and in cells. Finally, constrained MD simulations produced the first structural and energetic insights into the anion carrier mechanism assisted by a protonophore.

Overall, the work reported in this thesis produced valuable contributions into the transmembrane anion transport promoted by synthetic molecules, thus fully fulfilling the main objectives initially devised for this PhD thesis. The findings reported will surely be useful for the design of new compounds with potential to become channel replacement therapies.

VII.2. Direction of future work

As part of the on-going research line of the Molecular Modelling and Computational Biophysics Group, the know-how acquired with the pioneering theoretical work reported on this thesis on the arena of anion transmembrane transport should be extended to other series of structurally related putative anion carriers. Indeed, the MD methodologies employed along this PhD project, complemented by suitable quantum calculations, can be used to rationalise experimental transport data or, alternatively, to design putative transporters, thus steering the experimental effort.

Still, novel methods in the field of molecular modelling are continuously being developed and some might be worth considering in the study and understanding of anion transport promoted by synthetic carriers. Some examples of molecular modelling technics that might be applied to the study of anion transporters are presented below. On the other hand, considering that synthetic anion carriers are tested in a multitude of experimental conditions (*vide infra*), the theoretical studies should take steps to more faithfully reproduce the experimental conditions used in the anion efflux studies.

For instance, the experimental mechanism proposed for the anion carriers include either the symport of a proton, the antiport of another anion, or even both coupled transport processes.[1] These molecular entities can be included in the theoretical setup design to ascertain if the proton follows the anionic complex across the bilayer or if there is anion competition for the anion transporter. The existence of another anionic species might also imply that the one compartment systems reported in this thesis are insufficient for the study of concentration gradients or separation of said ionic species.[2] A way to tackle this matter is the simulation of a system with two independent phospholipid bilayers, which creates a two-compartment system.[2] This would also be a closer representation of the lipid vesicles employed in the experimental efflux studies. Thanks to the GPU acceleration currently available in MD suites such as AMBER,[3-5] the study of such systems is now within grasp.

On the other hand, the use of cholesterol is also used as an approach to verify if an anion transporter functions as an anion carrier or as a channel.[6] Cholesterol, which increases the ordering of the phospholipid tails,[7] makes the permeation of anion carriers more difficult, which, in turn, lowers the anion transport activity when compared with a pure phospholipid bilayer.[6] Assessing how growing ratios of cholesterol:phospholipid affect the interactions and energetic profiles of traversing anion transporters could be an interesting avenue to explore, given the roles that cholesterol plays in cellular membranes.[7] With the development of cholesterol parameters to be used along Lipid14,[7] such theoretical investigations can also be undertaken.

Along with the increasing use of Cystic Fibrosis cell models,[8] anion transporters have also been demonstrated to present cytotoxic activity in cancer cells,[9] as well as antibacterial properties,[10, 11] both linked with anion transport activity. This poses a challenge due to the complex structure of cell membranes, comprising (glyco)proteins, (glyco)lipids, and cholesterol. While the inclusion of different

protein components could seriously hamper the performance of atomistic MD simulations, the use of a broader number of lipid components is more viable, as several phospholipid parameters are available in Lipid14 (neutral head groups, typical of eukaryotic cells)[12] or expected to be available in the upcoming Lipid17 (with negatively charged head groups, typical of prokaryotic cells).[13-15]

On the other hand, the malfunction of protein channels for other anions, such as sulfate, phosphate, bicarbonate, and iodide are also linked with the occurrence of several diseases and conditions,[16-19] thus prompting the development of synthetic transmembrane transporters for these anions.[20-24] However, in contrast with chloride transporters, the number of synthetic molecules able to promote the selective transport of iodide or oxyanions is still scarce. As the number of systems reported in the literature is expected to grow, the theoretical investigations herein reported can also be applied to those series of transporters.

In the context of the cell models described above, complex systems comprising several lipids types, membrane proteins of interest and synthetic anion transporters might also be studied in a hybrid setup, *i.e.*, a combination of all-atom and coarse grained force fields.[25] This approach would allow to efficiently simulate the system and assess how the small molecules interact with relevant membrane proteins. This option might imply that other MD software other than AMBER must be considered, such as LAMMPS, which also benefits from GPU acceleration.[26]

Finally, the study of other reaction coordinates other than the distance between the anion carrier and the centre of mass of the bilayer along the membrane normal can also be explored. For instance, the study of aspirin, diclofenac and ibuprofen diffusion across a lipid bilayer was complemented with modulation of their intramolecular hydrogen bonding interactions.[27] These studies were carried out *via* MetaDynamics (MetaD) simulations and showed that these internal reaction coordinates affected the diffusion and partition of these compounds throughout the bilayer.[27] Therefore, the study of reaction coordinates internal to the anion complex (*e.g.*, distance between anion and receptor's binding units) could be of interest to bring further insights into the anion carrier mechanism. The implementation of MetaD[28] in the AMBER software suite might be worth testing to explore if this kind of MD simulations can be applied to the study of anion transmembrane transporters.

VII.3. References

1. J. T. Davis, O. Okunola and R. Quesada, *Chem. Soc. Rev.*, 2010, **39**, 3843-3862.
2. S. A. Kirsch and R. A. Bockmann, *Biochim. Biophys. Acta*, 2016, **1858**, 2266-2277.
3. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
4. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
5. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
6. N. Busschaert, P. A. Gale, C. J. Haynes, M. E. Light, S. J. Moore, C. C. Tong, J. T. Davis and W. A. Harrell, Jr., *Chem. Commun.*, 2010, **46**, 6252-6254.
7. B. D. Madej, I. R. Gould and R. C. Walker, *J. Phys. Chem. B*, 2015, **119**, 12424-12435.
8. N. Busschaert, S. H. Park, K. H. Baek, Y. P. Choi, J. Park, E. N. W. Howe, J. R. Hiscock, L. E. Karagiannidis, I. Marques, V. Felix, W. Namkung, J. L. Sessler, P. A. Gale and I. Shin, *Nat Chem*, 2017, **9**, 667-675.
9. S. J. Moore, C. J. E. Haynes, J. González, J. L. Sutton, S. J. Brooks, M. E. Light, J. Herniman, G. J. Langley, V. Soto-Cerrato, R. Pérez-Tomás, I. Marques, P. J. Costa, V. Félix and P. A. Gale, *Chem. Sci.*, 2013, **4**, 103-117.
10. C. R. Elie, G. David and A. R. Schmitzer, *J. Med. Chem.*, 2015, **58**, 2358-2366.
11. A. I. Share, K. Patel, C. Nativi, E. J. Cho, O. Francesconi, N. Busschaert, P. A. Gale, S. Roelens and J. L. Sessler, *Chem. Commun.*, 2016, **52**, 7560-7563.
12. C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
13. J. N. Horn, J. D. Sengillo, D. Lin, T. D. Romo and A. Grossfield, *Biochim. Biophys. Acta*, 2012, **1818**, 212-218.
14. J. N. Horn, T. D. Romo and A. Grossfield, *Biochemistry*, 2013, **52**, 5604-5610.

15. A. A. Skjevik, B. D. Madej, C. J. Dickson, C. Lin, K. Teigen, R. C. Walker and I. R. Gould, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10573-10584.
16. J. Hästbacka, A. de la Chapelle, M. M. Mahtani, G. Clines, M. P. Reeve-Daly, M. Daly, B. A. Hamilton, K. Kusumi, B. Trivedi, A. Weaver, A. Coloma, M. Lovett, A. Buckler, I. Kaitila and E. S. Lander, *Cell*, 1994, **78**, 1073-1087.
17. A. Superti-Furga, A. Rossi, B. Steinmann and R. Gitzelmann, *Am. J. Med. Genet.*, 1996, **63**, 144-147.
18. F. M. Ashcroft, *Ion Channels and Disease*, Academic Press, London, 1st edn., 2000.
19. J. J. Wine, *Curr. Biol.*, 2001, **11**, R463-R466.
20. I. Izzo, S. Licen, N. Maulucci, G. Autore, S. Marzocco, P. Tecilla and F. De Riccardis, *Chem. Commun.*, 2008, 2986-2988.
21. J. T. Davis, P. A. Gale, O. A. Okunola, P. Prados, J. C. Iglesias-Sanchez, T. Torroba and R. Quesada, *Nat Chem*, 2009, **1**, 138-144.
22. N. Busschaert, L. E. Karagiannidis, M. Wenzel, C. J. E. Haynes, N. J. Wells, P. G. Young, D. Makuc, J. Plavec, K. A. Jolliffe and P. A. Gale, *Chem. Sci.*, 2014, **5**, 1118-1127.
23. P. Y. Liu, S. T. Li, F. F. Shen, W. H. Ko, X. Q. Yao and D. Yang, *Chem. Commun.*, 2016, **52**, 7380-7383.
24. B. P. Benke, P. Aich, Y. Kim, K. L. Kim, M. R. Rohman, S. Hong, I. C. Hwang, E. H. Lee, J. H. Roh and K. Kim, *J. Am. Chem. Soc.*, 2017, **139**, 7432-7435.
25. S. Genheden and J. W. Essex, *J. Chem. Theory Comput.*, 2015, **11**, 4749-4759.
26. S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1-19.
27. J. P. Jambeck and A. P. Lyubartsev, *J. Phys. Chem. Lett.*, 2013, **4**, 1781-1787.
28. M. Mahmoud, B. Volodymyr, R. Christopher and S. Celeste, *Journal of Physics: Conference Series*, 2015, **640**, 012020.

Appendix A.
Additional data and extended discussion of Chapter II

A.1. Additional figures for simulations B–G

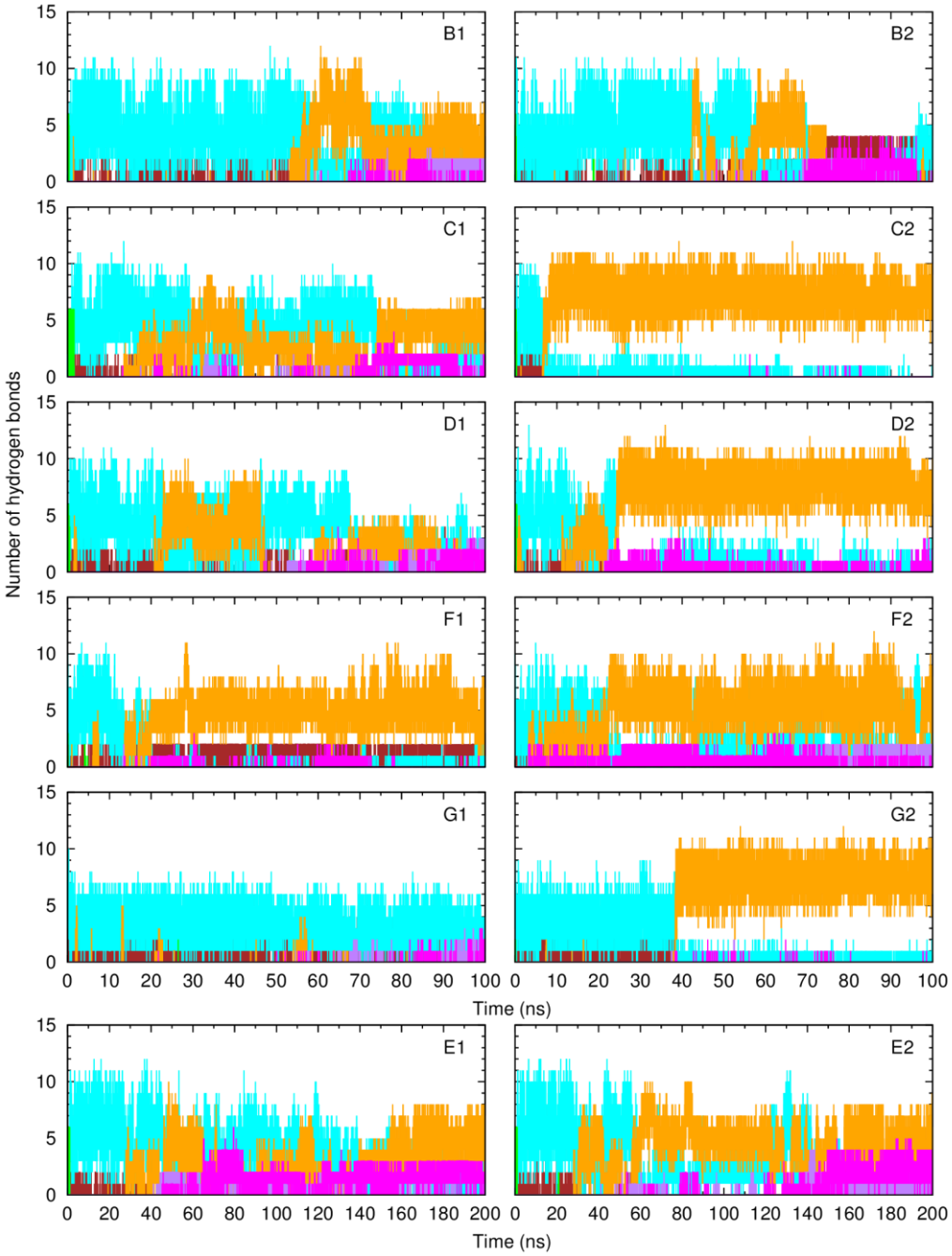


Figure A-1. Hydrogen bonds counting for the interactions established by the N–H groups of **1** (simulations B1 and B2), **2** (simulations C1 and C2), **3** (simulations D1 and D2), **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2). The following colour scheme was used for the interactions between the receptor and sulfur atoms (brown), chloride ion (green), water molecules (cyan), POPC head groups (orange), ester groups (purple for the *sn*-1 chains and magenta for the *sn*-2 chains).

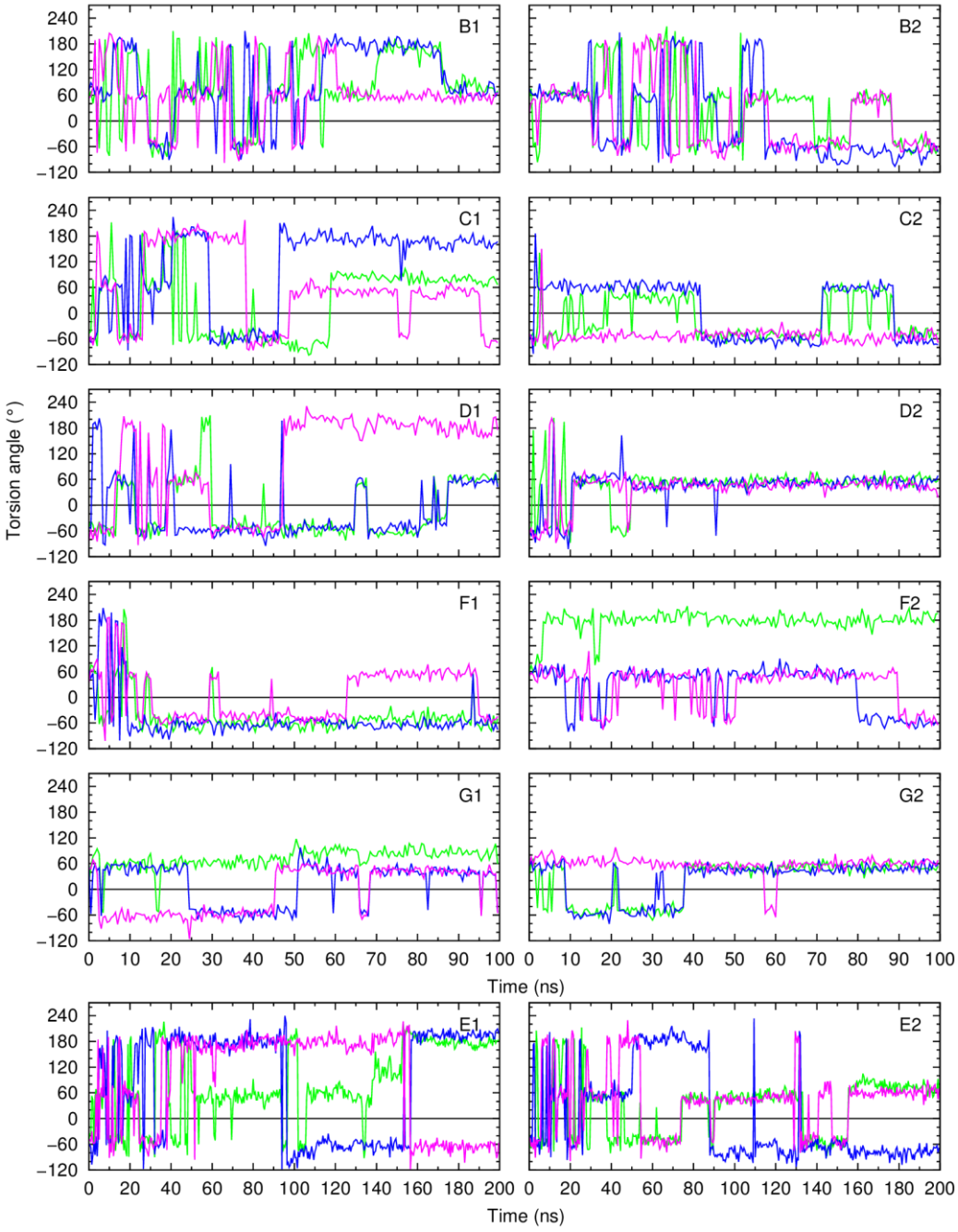


Figure A-2. Variations in the three N_{tren}–C–C–N tripodal torsion angles of **1** (simulations B1 and B2), **2** (simulations C1 and C2), **3** (simulations D1 and D2), **4** (simulations E1 and E2), **5** (simulations F1 and F2) and **6** (simulations G1 and G2), along the MD simulations, extracted every 500 ps.

A.2. Additional figures for simulation H

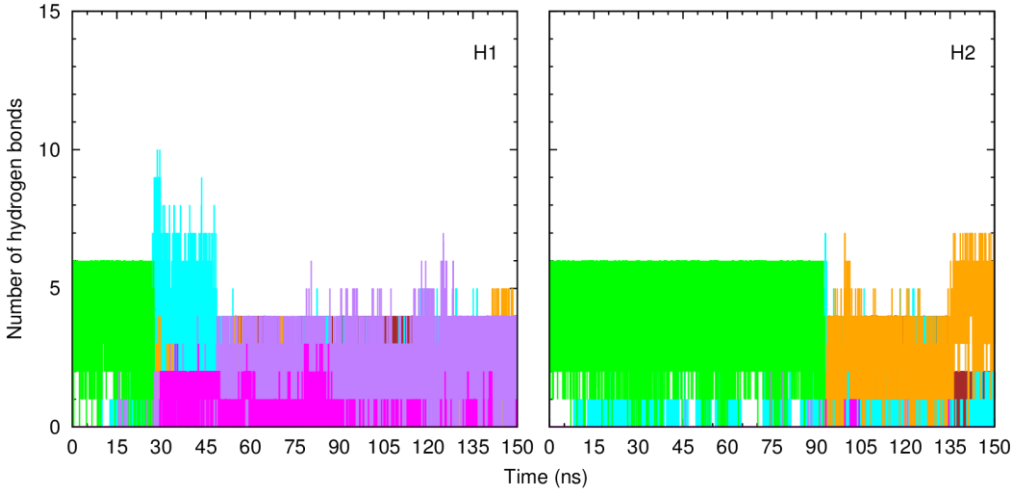


Figure A-3. Hydrogen bonds counting for the interactions established by the N–H groups of **3** (simulations H1 and H2). The following colour scheme was used for the interactions between the receptor and sulfur atoms (brown), chloride ion (green), water molecules (cyan), POPC head groups (orange), ester groups (purple for the *sn*-1 chains and magenta for the *sn*-2 chains).

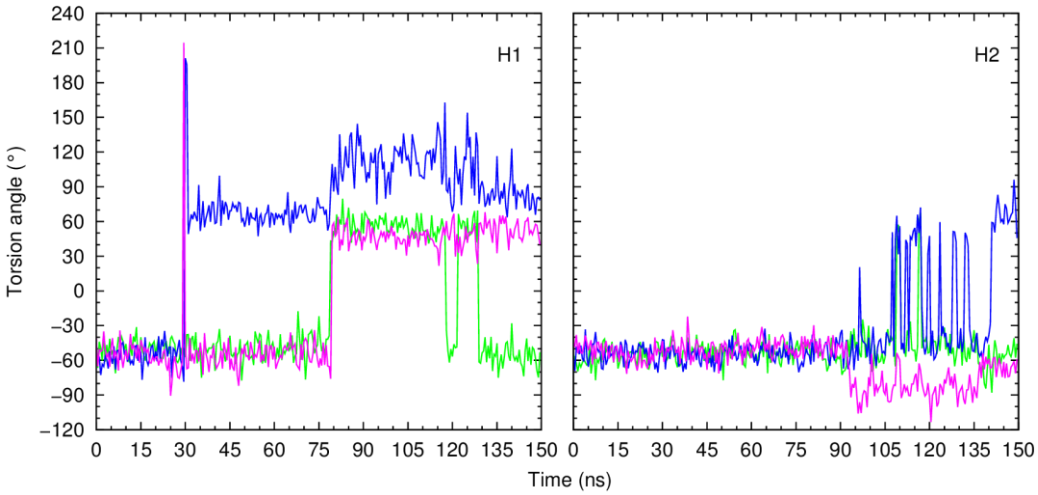


Figure A-4. Variations in the three N_{tren}–C–C–N tripodal torsion angles for 150 ns of MD simulation, extracted every 500 ps, for simulations H1 and H2, containing **3**.

A.3. Methods of preliminary MD simulation in water solution and additional figures for simulation E'

Chloride complexes of **1**–**6** were immersed in cubic boxes containing 3009 TIP3P water molecules, affording systems **a**–**f** and **g**–**l**. In the six former systems, the individual N_{tren}–C–C–N torsion angles of the tripodal molecule were restrained with a 10 kcal mol^{-1} rad^{-2}, while in systems **g**–**l** an additional distance restraint was also applied, as summarised in Table A-1. The following multi-stage equilibration process was carried out before the 30 ns production.

Initially, the solvent was relaxed by a MM minimisation with a harmonic restraint of 500 kcal mol^{-1} Å^{-2} on the anion complexes, followed by a MM minimisation of all system. The system was then heated to 303 K during 50 ps using Langevin dynamics[1] with a collision frequency of 1.0 ps^{-1} in an NVT ensemble, with a 10 kcal mol^{-1} Å^{-2} restraint on the chloride complexes. The density of the system was then adjusted for 1.0 ns in a NPT ensemble at 1 atm with isotropic pressure scaling using a relaxation time of 1.0 ps, with the anion complexes still restrained with a 10 kcal mol^{-1} Å^{-2} force constant. The 30 ns collection runs were performed for simulations **a**–**f** and **g**–**l**, with the restraints given in Table A-1.

The SHAKE algorithm[2] was used to constrain all bonds involving hydrogen atoms, allowing the use of a 2.0 fs time step. An 8.0 Å cut-off was used for the van der Waals and non-bonded electrostatic interactions. Frames were saved every 1.0 ps leading to trajectory files containing 30000 structures for each simulation. These simulations were performed with the *PMEMD* AMBER 12 executable,[3] which allowed to accelerate explicit solvent Particle Mesh Ewald (PME)[4] calculations, using GPUs.[5-7]

Table A-1. Water solution simulated systems at 303 K for 30 ns.

System ID	Transporter	N_{tren}–C–C–N torsion angle restraints (kcal mol^{-1} rad^{-2})	Cl$^-$···N distance restraints (kcal mol^{-1} Å^{-2})
a	**1**	10[a]	–
b	**2**		
c	**3**		
d	**4**		
e	**5**		
f	**6**		
g	**1**	10[a]	1.0[b]
h	**2**		
i	**3**		
j	**4**		
k	**5**		
l	**6**		

[a)] This force constant was applied to each one of the three N_{tren}–C–C–N torsion angles; [b)] This force constant was applied between the anion and the nitrogen atom of an N–H binding unit.

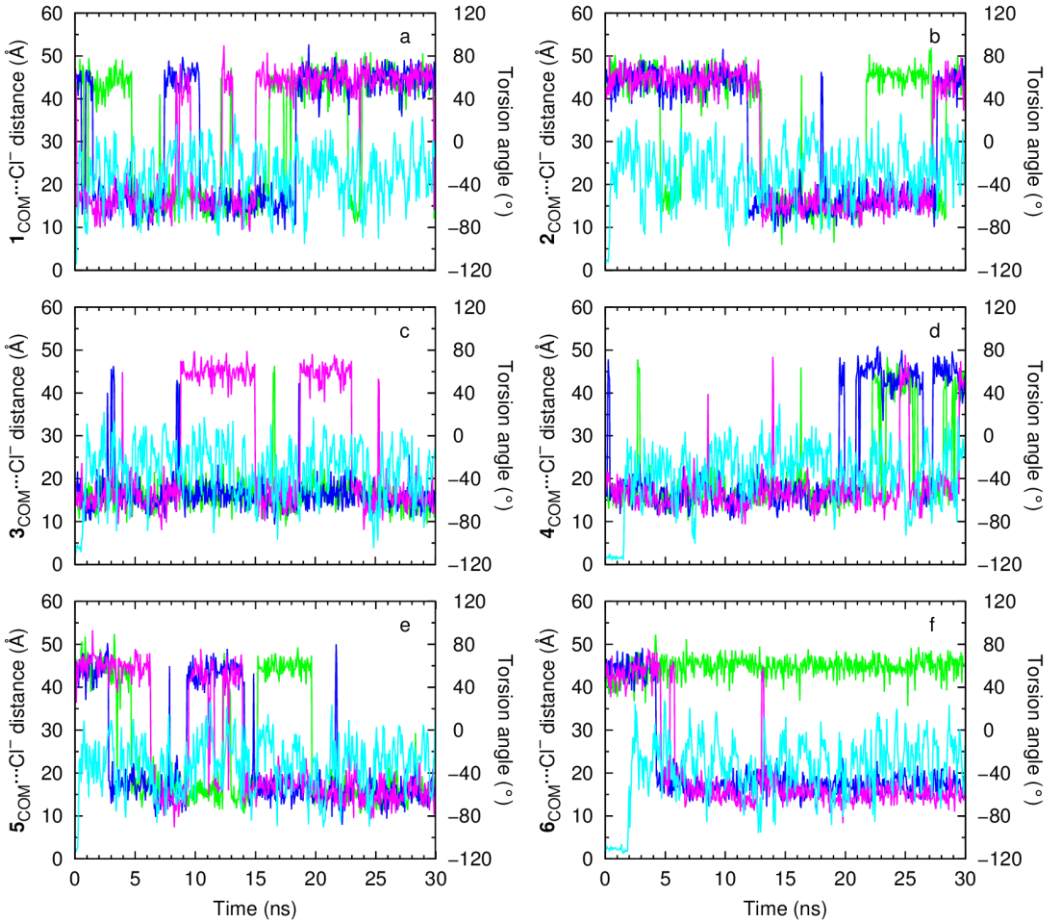


Figure A-5. Evolution of the **1**–**6**$_{COM}$···Cl^- distances (cyan line) and variations in the three N_{tren}–C–C–N tripodal torsion angles (blue, green and magenta lines), in simulations **a**–**f** for 30 ns. All parameters were extracted every 50 ps.

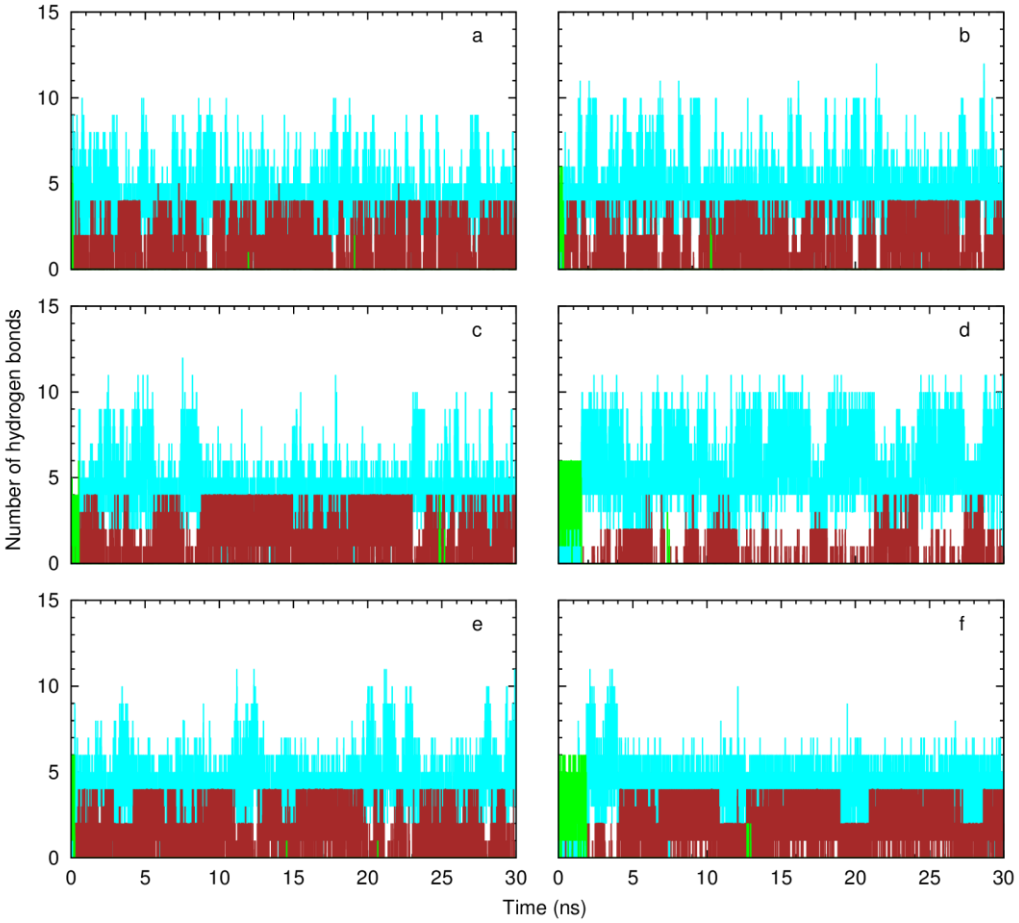


Figure A-6. Hydrogen bonds counting for the interactions established by the N–H groups of **1**–**6** in simulations **a**–**f**. The following colour scheme was used for the interactions between the receptor and sulfur atoms (brown), chloride ion (green), and water molecules (cyan).

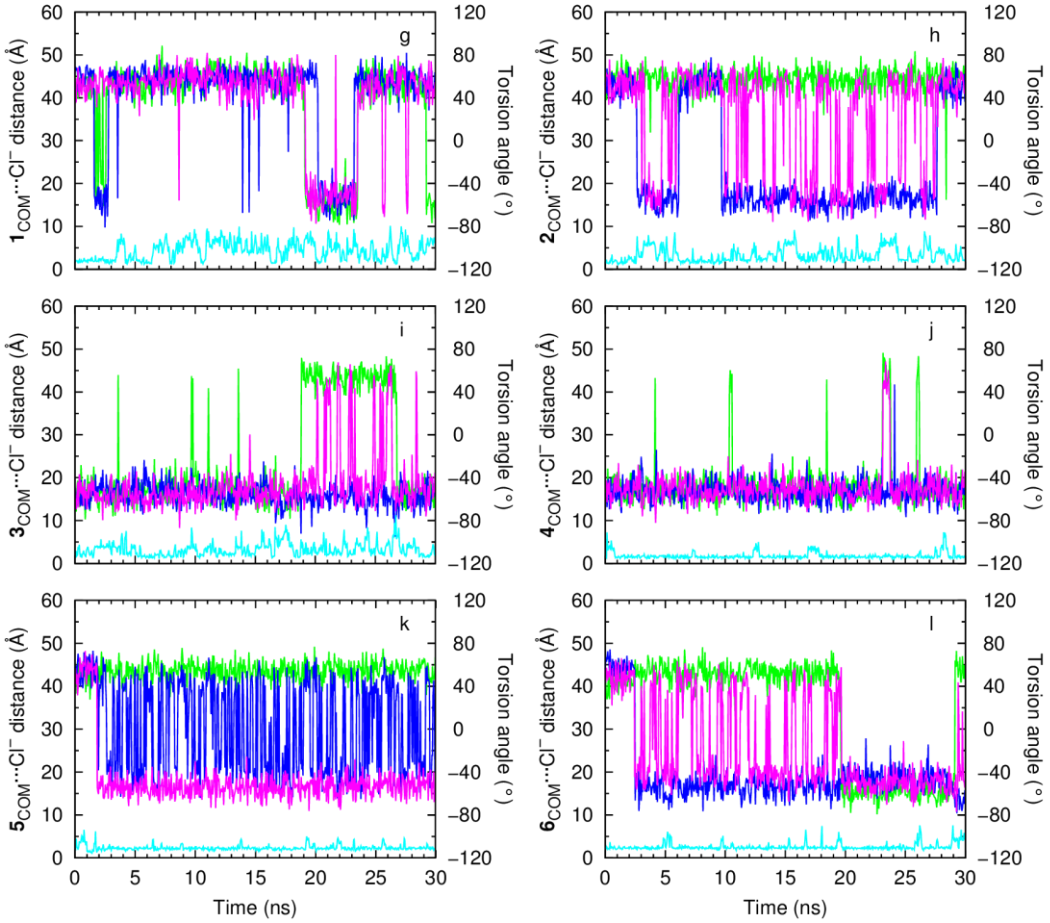


Figure A-7. Evolution of the **1**–**6**$_{COM}$···Cl^- distances (cyan line) and variations in the three N_{tren}–C–C–N tripodal torsion angles (blue, green and magenta lines), in simulations **g**–**l** for 30 ns. All parameters were extracted every 50 ps.

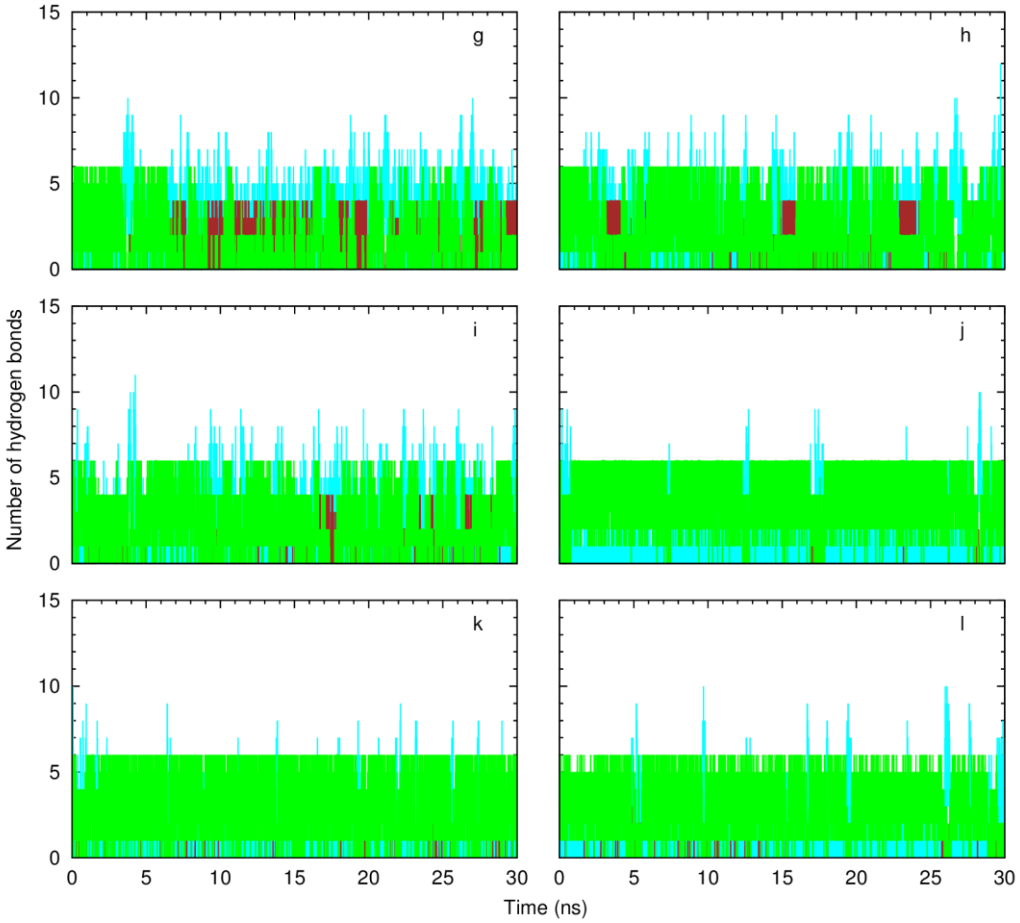


Figure A-8. Hydrogen bonds counting for the interactions established by the N–H groups of **1**–**6** in simulations **g**–**l**. The following colour scheme was used for the interactions between the receptor and sulfur atoms (brown), chloride ion (green), and water molecules (cyan).

A.4. Additional figures for the evaluation of the impact of 1–6 onto the POPC membrane

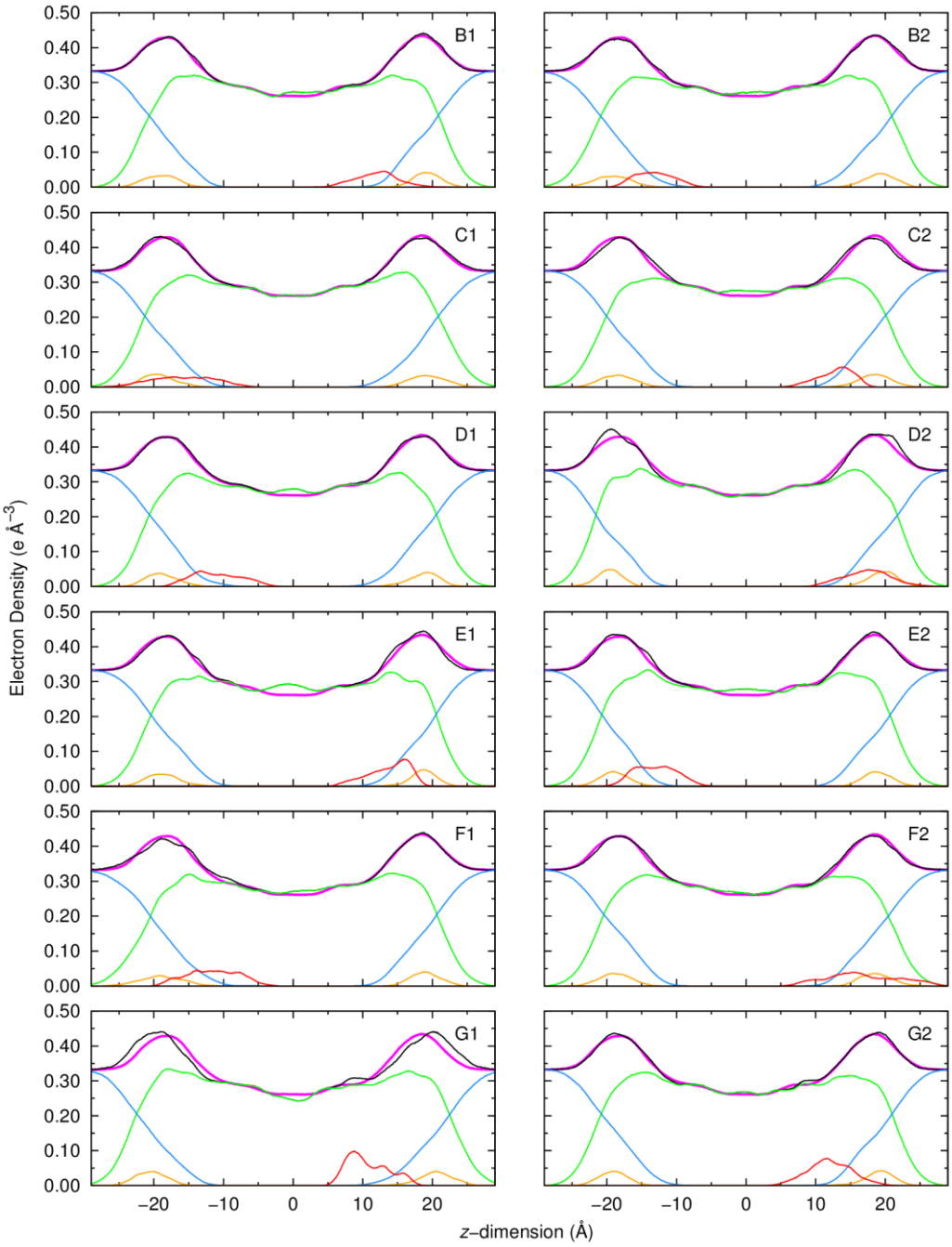


Figure A-9. Electron density profiles of simulations B1–G2 with the full system plotted in black, water in blue, phospholipids in green, phosphorus atoms in orange and the corresponding transporter profile in red (scaled 5 times). z = 0 Å corresponds to the core of the POPC bilayer. The POPC bilayer profile of system A is also shown as a pink line.

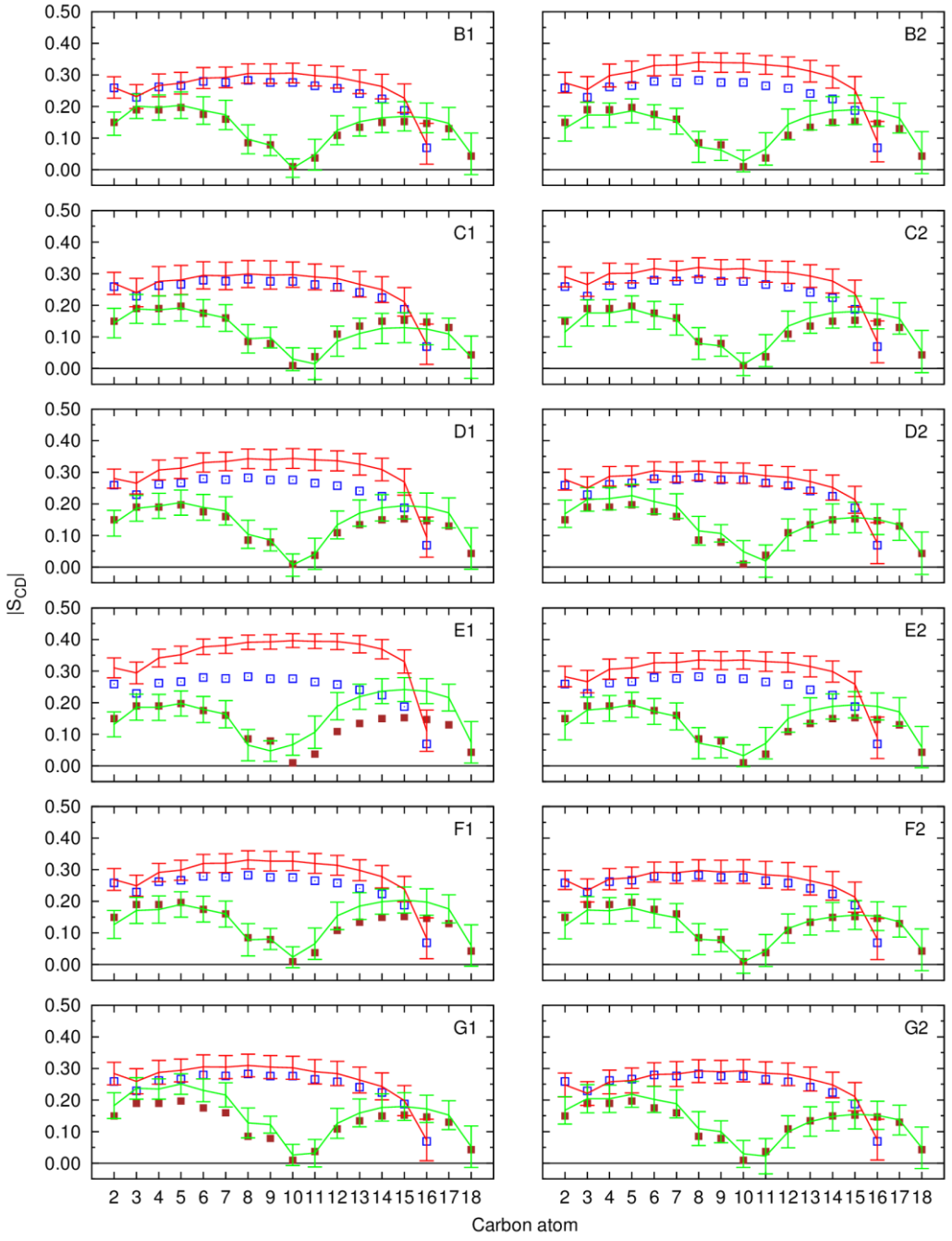


Figure A-10. Computed $|S_{CD}|$ for palmitoyl and oleyl chains for 25 ns of sampling of simulations B1–G2. The $|S_{CD}|$ values calculated for the *sn*-1 chain are shown in red, while the values for the *sn*-2 chain are shown in green. The error bars associated with these results correspond to the SD. The computed $|S_{CD}|$ values from system A are presented as blue ▫ (*sn*-1 chain), and brown ▪ (*sn*-2 chain).

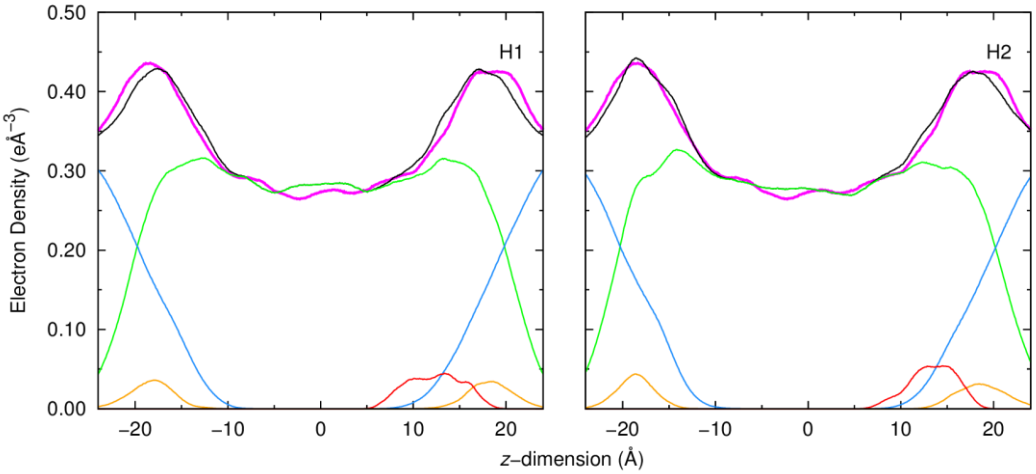


Figure A-11. Electron density profiles of simulations H1 and H2 with the full system plotted in black, water in blue, phospholipids in green, phosphorus atoms in orange and **3**'s profile in red (scaled 5 times). z = 0 Å corresponds to the core of the POPC bilayer. The POPC bilayer profile reported on the Lipid11 paper is also shown as a pink line.

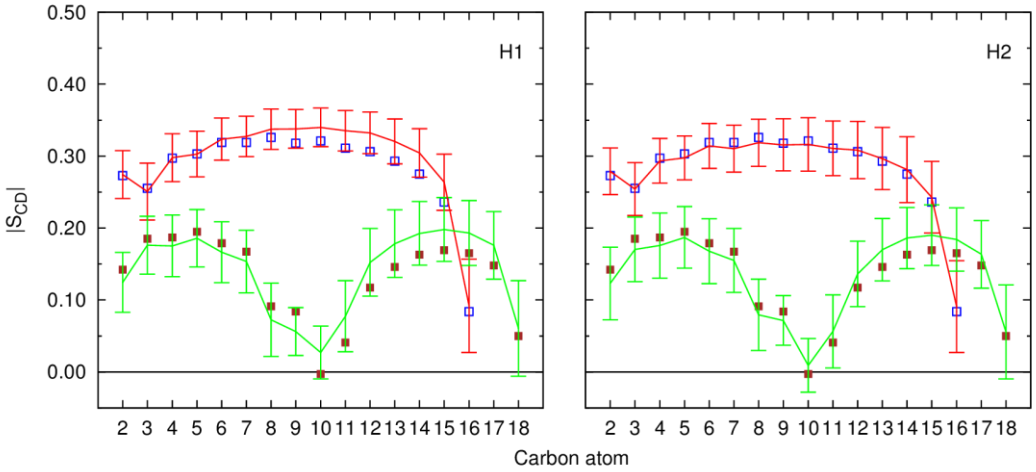


Figure A-12. Computed |S_{CD}| for palmitoyl and oleyl chains for 50 ns of sampling of simulations H1 and H2. The |S_{CD}| values calculated for the *sn*-1 chain are shown in red, while the values for the *sn*-2 chain are shown in green. The error bars associated with these results correspond to the SD. The computed |S_{CD}| values from the Lipid11 paper are presented as blue ▫ (*sn*-1 chain), and brown ▪ (*sn*-2 chain).

A.5. References

1. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
2. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
3. D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2012), AMBER 12, University of California, San Francisco.
4. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
5. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
6. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
7. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.

Appendix B.
Additional data for Chapter III

B.1. Movie III-1 caption

Movie III-1. Passive diffusion of the chloride complex of **7d** in simulation A_1.**7d**. This movie shows the passive diffusion of the chloride complex of **7d** from the membrane core towards the water/lipid interface (simulation A_1.**7d**, between the 6th and the 10th ns of MD simulation time). This movie also features the chloride release below the interface assisted by water molecules coming from the aqueous phase. The chloride decomplexation assisted by water molecules is emphasised with the depiction of water molecules within 3.5 Å from the anion. All atoms are presented in spheres apart from those of phospholipids, which are drawn in stick fashion. The hydrogen atoms are in white, oxygen atoms in red, nitrogen atoms in blue, sulfur atoms in yellow, phosphorus atoms in orange, sodium counterions in pink, and carbon atoms in light blue (transporter) or wheat (phospholipids). The coordinated chloride is shown in light green, while the remaining aqueous phase chloride ions are shown in dark green. The lipid C–H bonds are omitted for clarity.

B.2. Additional figures for simulations A_1.7a to A_4.7f

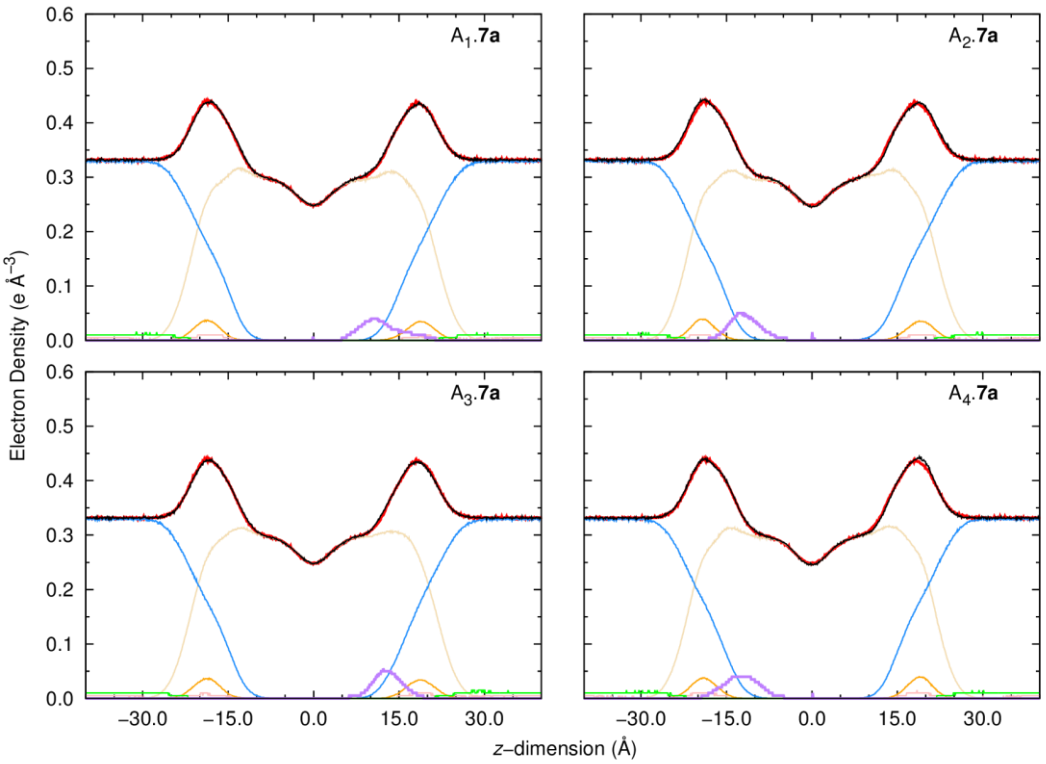


Figure B-1. Electron density profiles of simulations A_1.**7a**–A_4.**7a** with the full system plotted in black, water in blue, phospholipids in wheat, phosphorus atoms in orange, chloride ions in green, sodium ions in pink and **7a** in purple. The ions' and the transporter's profiles are scaled 5 times. z = 0 Å corresponds to the core of the POPC bilayer. The free membrane profile is also shown as a red line.

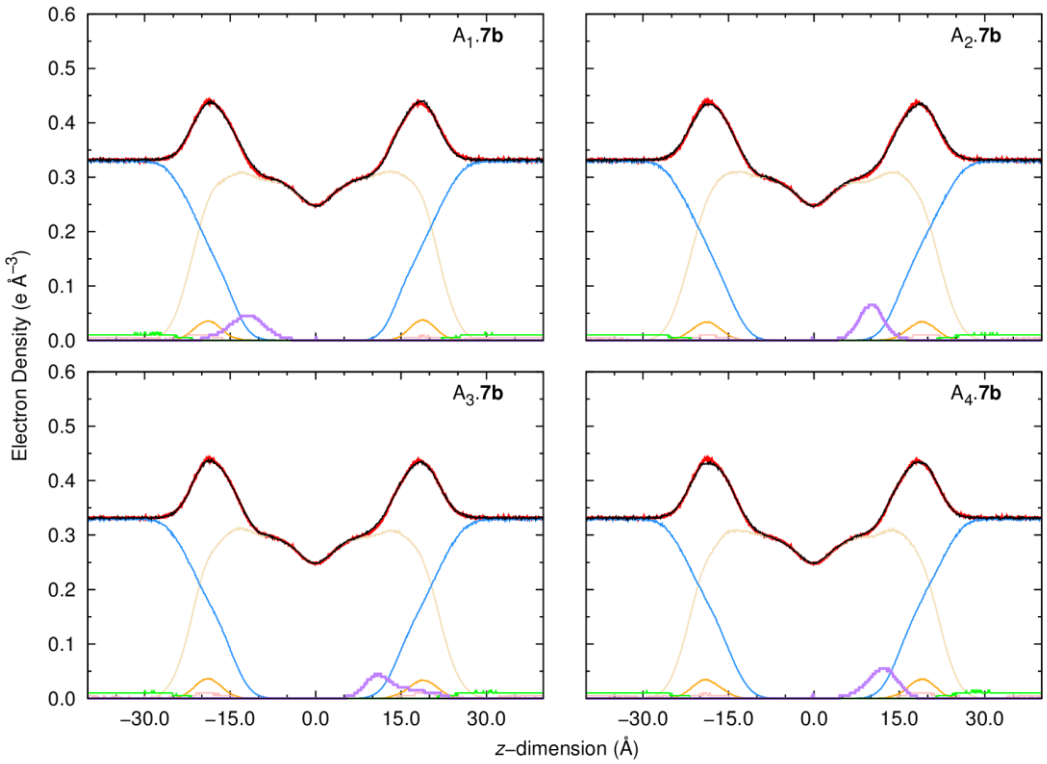


Figure B-2. Electron density profiles of simulations A1.**7b**–A4.**7b**, with **7b** plotted in purple. Remaining details as given in Figure B-1.

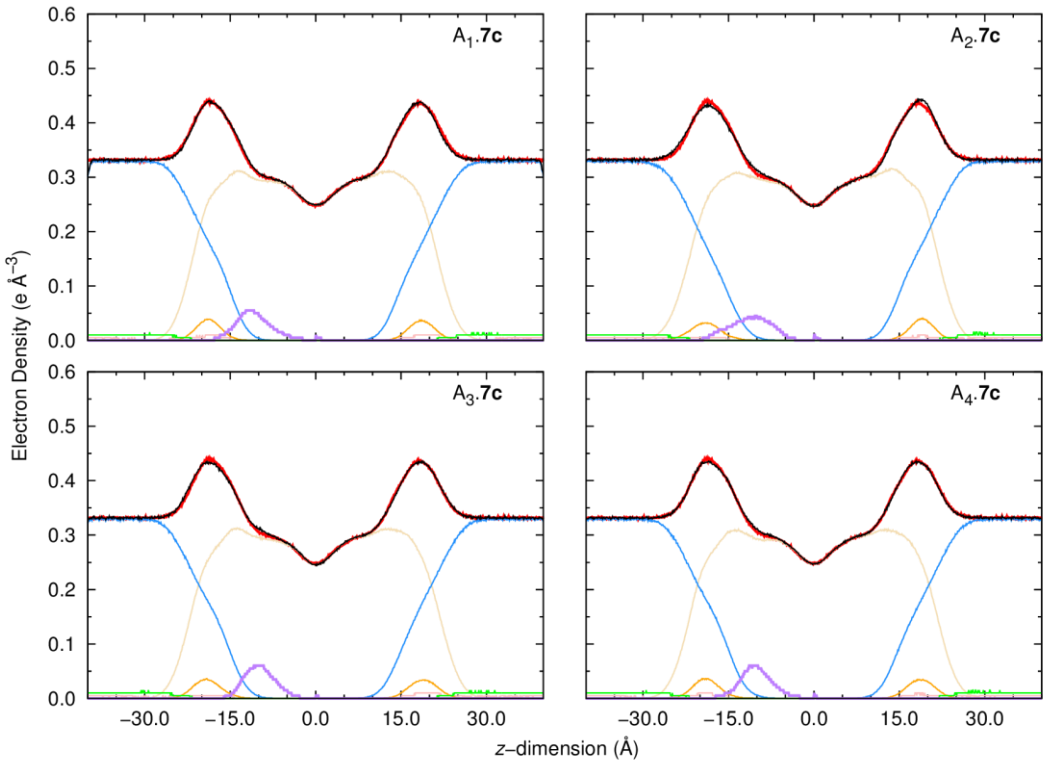


Figure B-3. Electron density profiles of simulations A1.**7c**–A4.**7c**, with **7c** plotted in purple. Remaining details as given in Figure B-1.

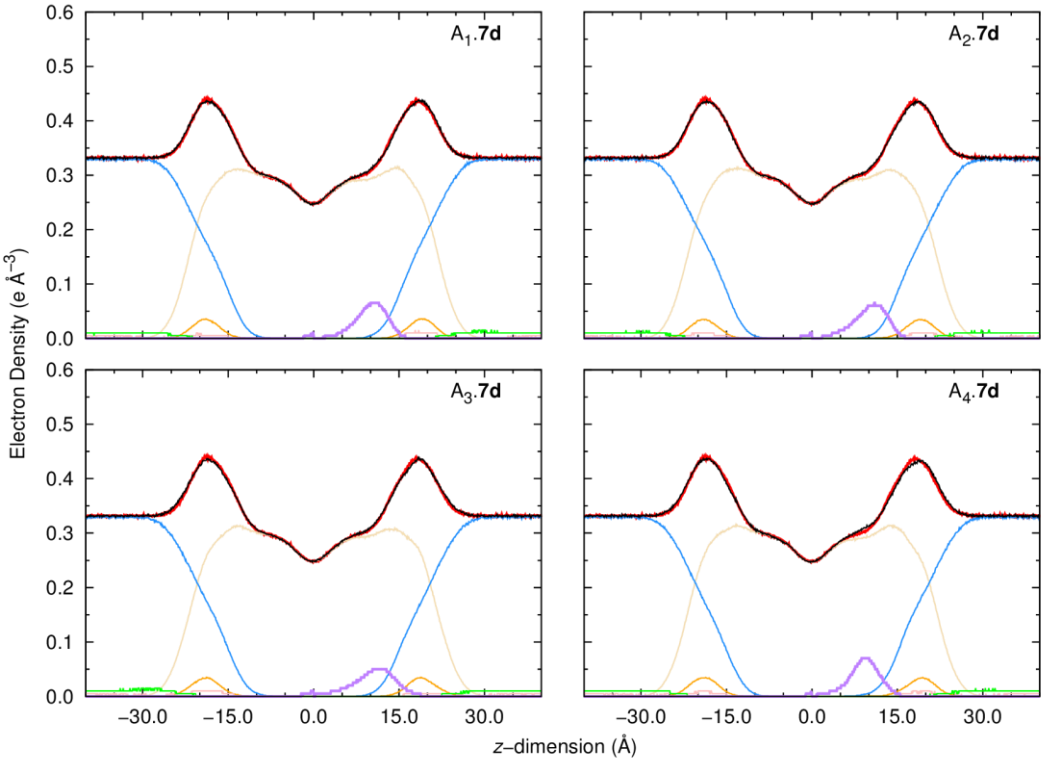


Figure B-4. Electron density profiles of simulations A1.**7d**–A4.**7d**, with **7d** plotted in purple. Remaining details as given in Figure B-1.

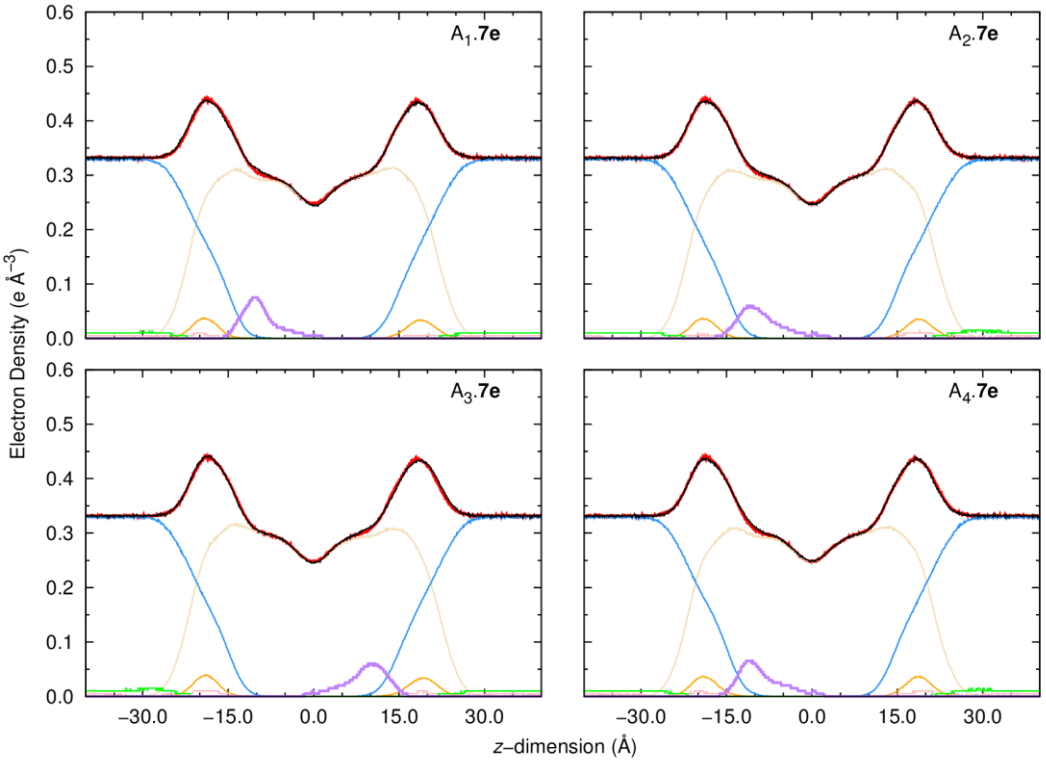


Figure B-5. Electron density profiles of simulations A1.**7e**–A4.**7e**, with **7e** plotted in purple. Remaining details as given in Figure B-1.

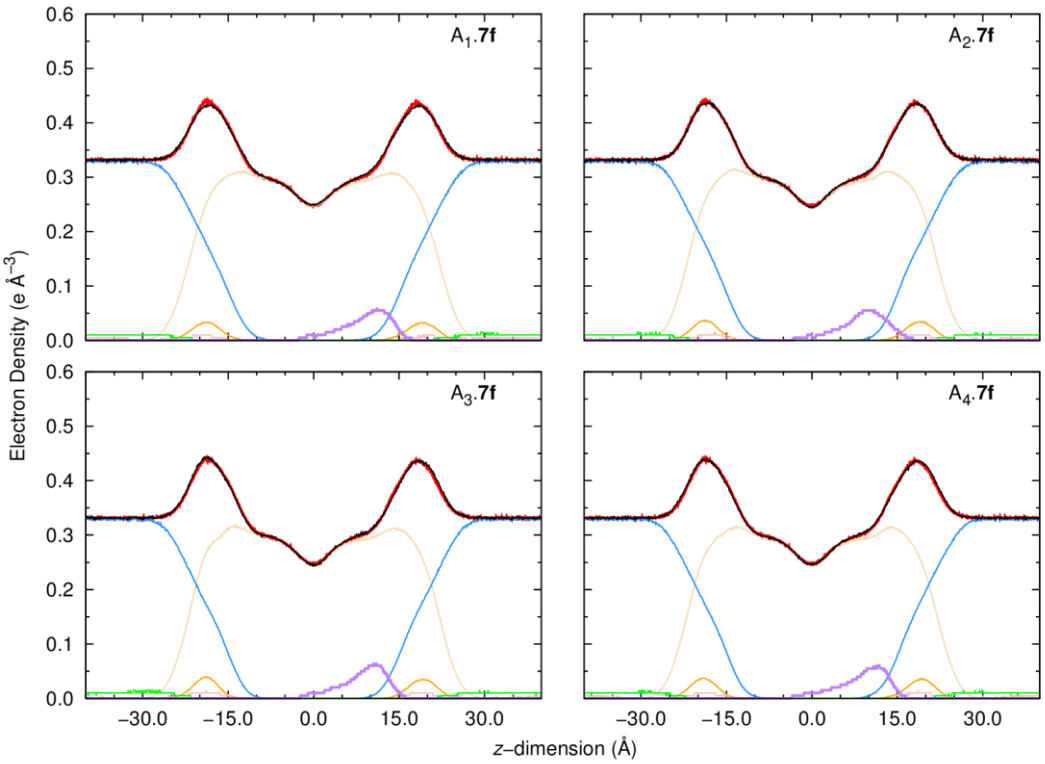


Figure B-6. Electron density profiles of simulations A1.**7f**–A4.**7f**, with **7f** plotted in purple. Remaining details as given in Figure B-1.

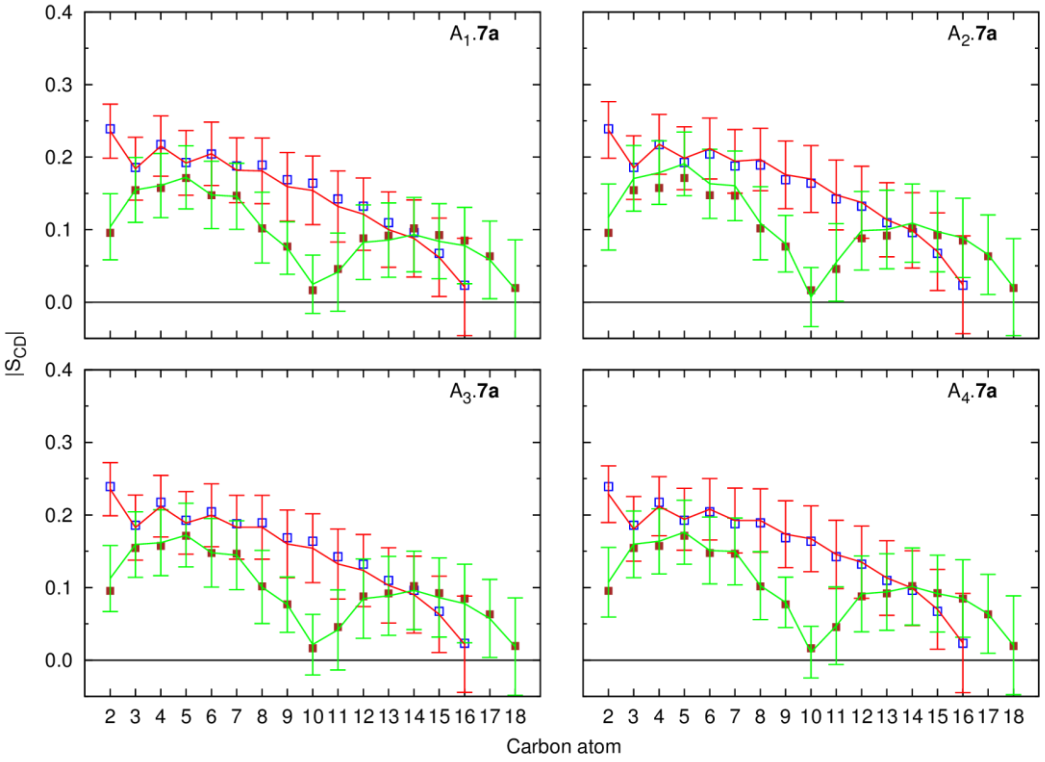


Figure B-7. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7a**–A4.**7a**. The $|S_{CD}|$ values calculated for the *sn*-1 chain are shown in red, while the values for the *sn*-2 chain are shown in green. The error bars associated with these results correspond to the SD. The computed $|S_{CD}|$ values from the free membrane are presented as blue ▫ (*sn*-1 chain), and brown ▪ (*sn*-2 chain).

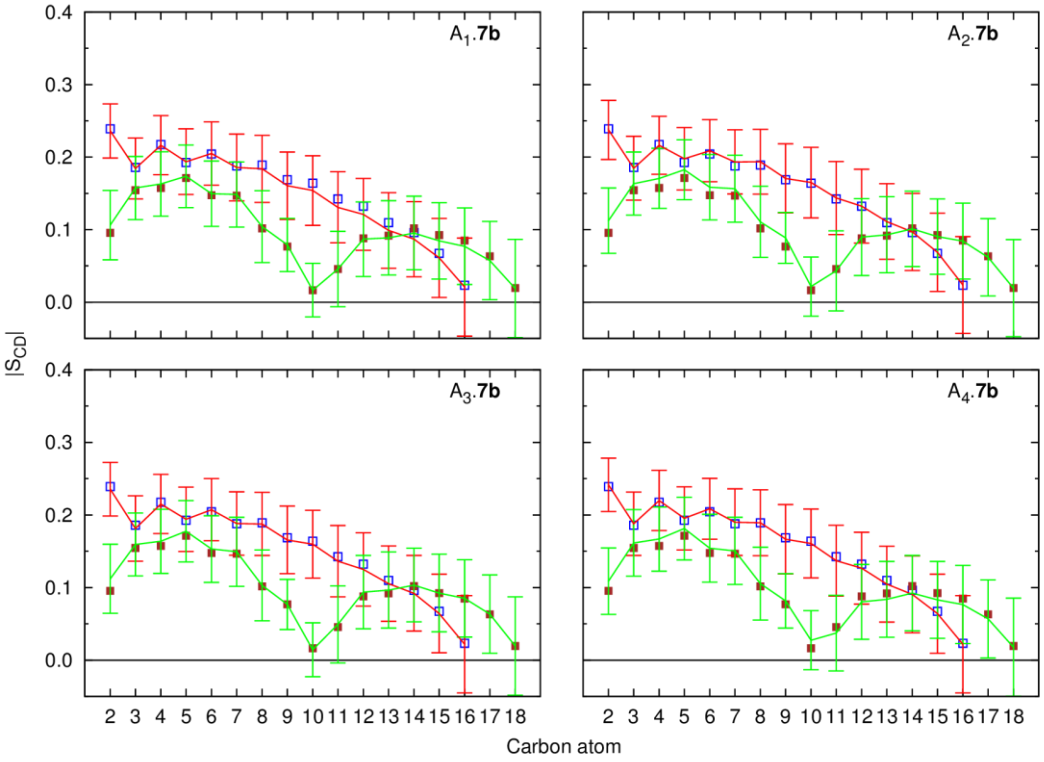


Figure B-8. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7b**–A4.**7b**. Remaining details as given in Figure B-7.

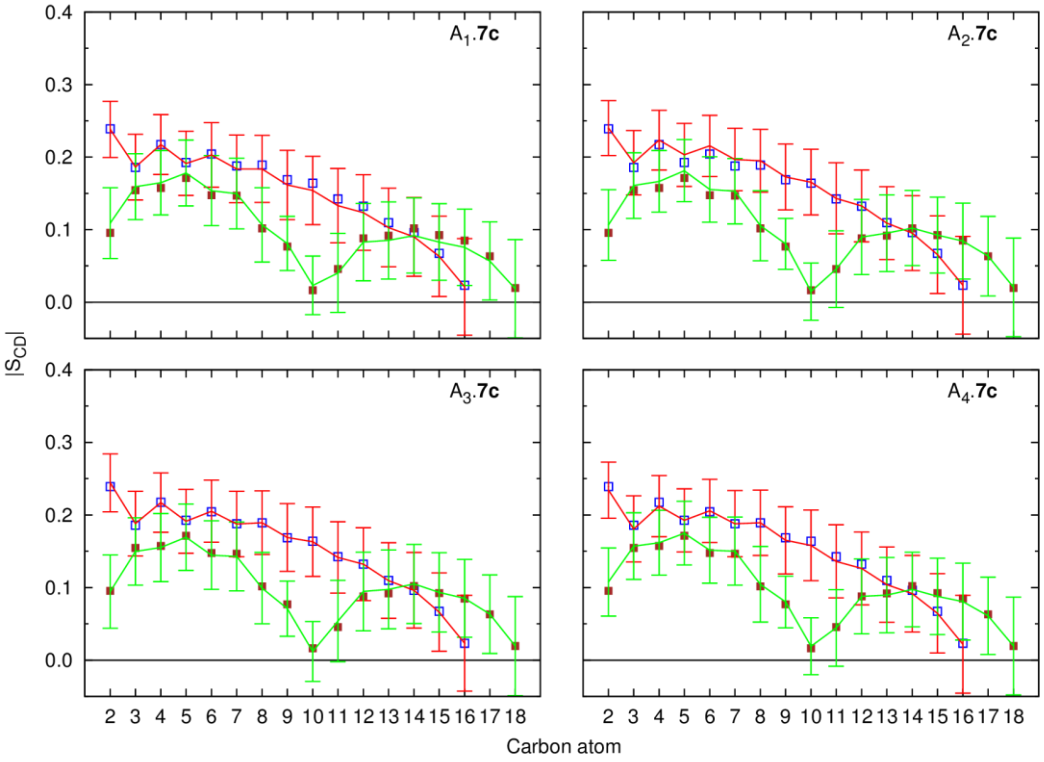


Figure B-9. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7c**–A4.**7c**. Remaining details as given in Figure B-7.

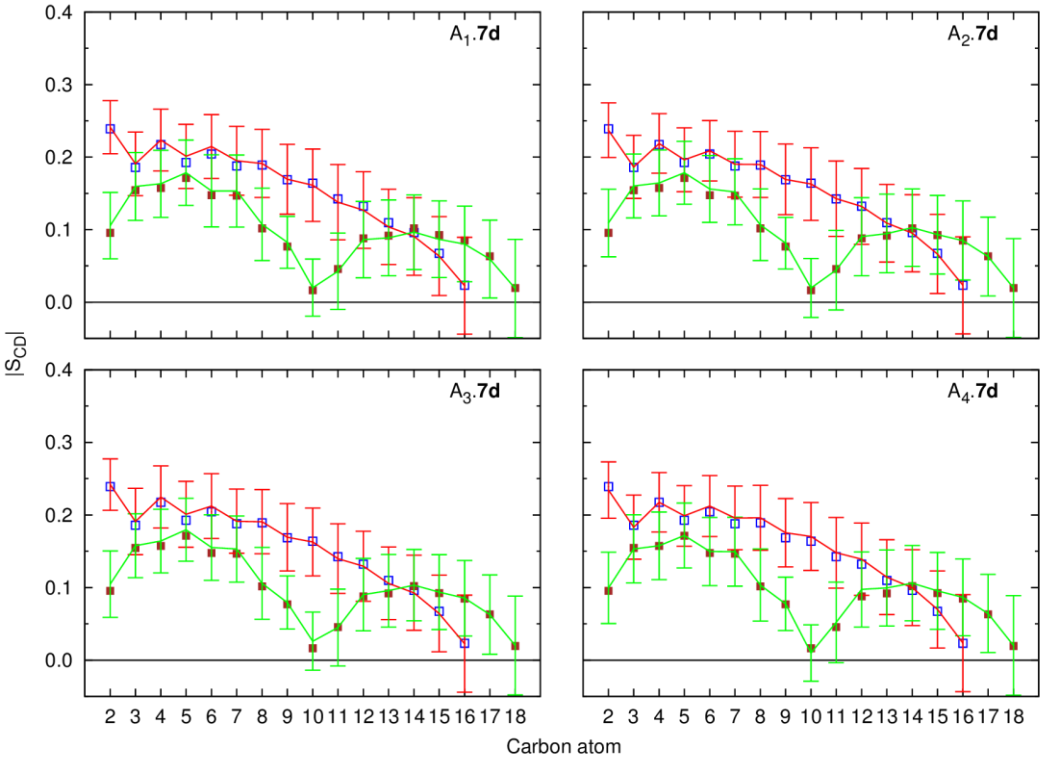


Figure B-10. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7d**–A4.**7d**. Remaining details as given in Figure B-7.

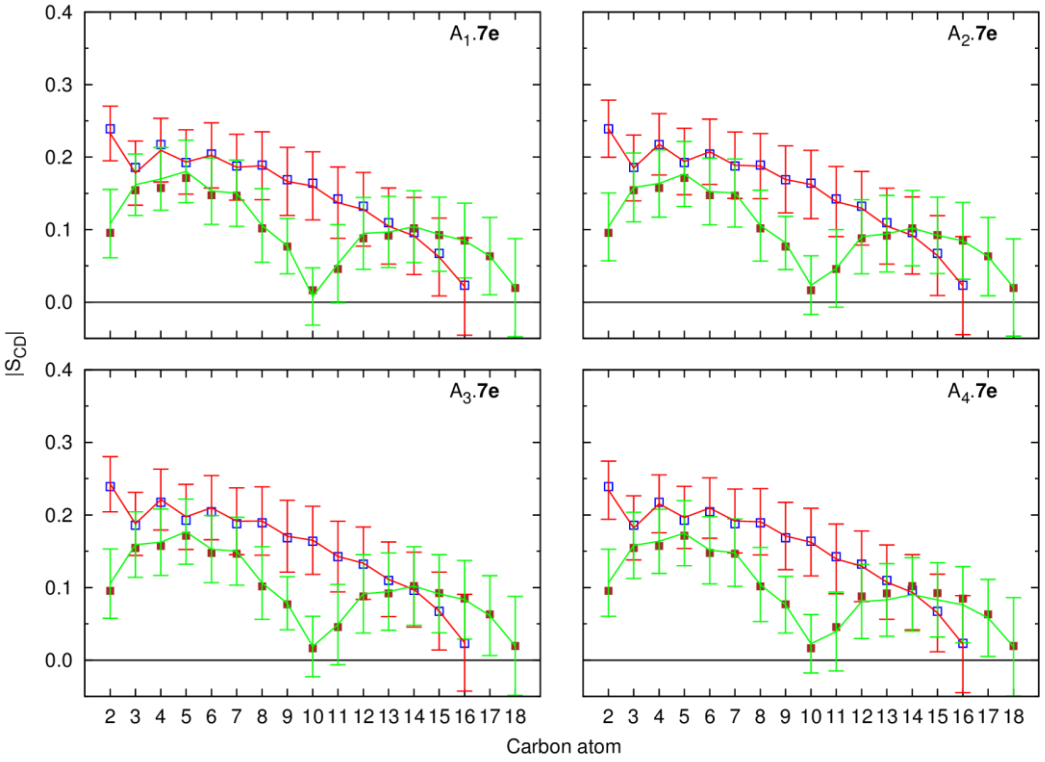


Figure B-11. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7e**–A4.**7e**. Remaining details as given in Figure B-7.

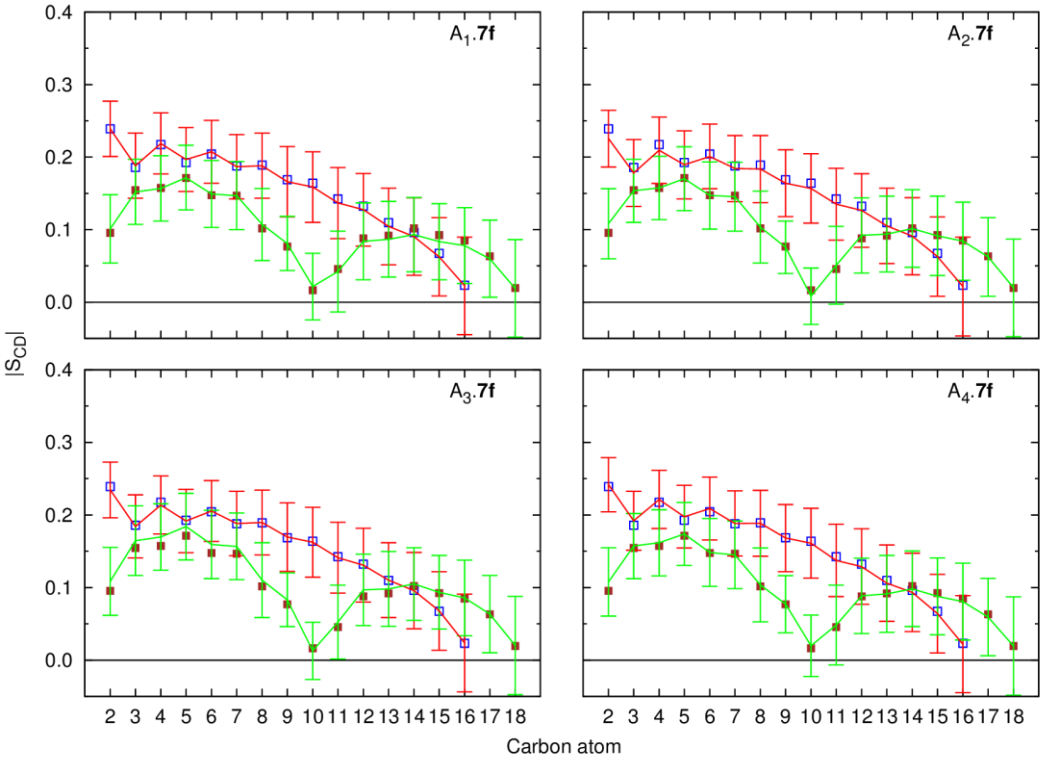


Figure B-12. Computed $|S_{CD}|$ for the palmitoyl and oleyl chains for the last 50 ns of MD runs A1.**7f**–A4.**7f**. Remaining details as given in Figure B-7.

B.3. Additional tables for simulations A1.7a to A4.7f

Table B-1. decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances (Å) averaged (Avg) for the last 50 ns of simulations A1.**7a**–A4.**7f** with the corresponding standard deviations (SD).[a]

Simulation ID	decalin$_{COM}$···P$_{int}$	N–H$_{COM}$···P$_{int}$	*p*–C$_{COM}$···P$_{int}$	tail$_{COM}$···P$_{int}$
	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD
A1.**7a**	8.11 ± 1.87	6.27 ± 2.01	3.32 ± 2.27	–
A2.**7a**	7.89 ± 2.21	6.77 ± 1.78	6.05 ± 1.87	–
A3.**7a**	5.35 ± 1.82	6.04 ± 1.75	6.79 ± 2.40	–
A4.**7a**	7.32 ± 2.39	6.02 ± 2.03	4.95 ± 2.24	–
A1.**7b**	5.73 ± 2.03	6.59 ± 1.73	8.06 ± 2.85	8.63 ± 3.46
A2.**7b**	9.21 ± 1.59	8.39 ± 1.31	8.23 ± 1.52	8.22 ± 1.80
A3.**7b**	7.71 ± 1.34	5.59 ± 1.28	2.45 ± 1.40	1.42 ± 1.52
A4.**7b**	6.02 ± 1.75	6.64 ± 1.55	8.49 ± 2.15	9.19 ± 2.50
A1.**7c**	8.82 ± 3.05	7.59 ± 2.36	6.75 ± 1.68	6.79 ± 1.99
A2.**7c**	6.66 ± 3.31	7.07 ± 2.96	8.75 ± 2.56	10.79 ± 2.14
A3.**7c**	9.75 ± 1.70	8.56 ± 1.58	8.30 ± 1.78	8.92 ± 2.35
A4.**7c**	8.80 ± 1.69	8.26 ± 1.51	8.29 ± 2.56	9.50 ± 3.94
A1.**7d**	8.42 ± 1.76	8.02 ± 1.60	8.68 ± 1.87	11.29 ± 2.69
A2.**7d**	7.23 ± 1.82	7.78 ± 1.67	9.53 ± 1.66	13.23 ± 2.01
A3.**7d**	6.55 ± 2.23	7.33 ± 2.25	9.47 ± 2.36	13.15 ± 2.45
A4.**7d**	8.74 ± 1.62	9.02 ± 1.49	10.10 ± 1.48	13.52 ± 1.65
A1.**7e**	8.08 ± 1.61	8.08 ± 1.34	8.78 ± 1.21	15.38 ± 1.43
A2.**7e**	7.70 ± 1.97	7.94 ± 1.80	9.62 ± 1.65	15.30 ± 2.00
A3.**7e**	7.40 ± 1.57	8.04 ± 1.33	10.09 ± 1.12	15.89 ± 1.43
A4.**7e**	7.03 ± 1.53	7.72 ± 1.41	9.80 ± 1.56	15.82 ± 1.51
A1.**7f**	6.77 ± 1.85	7.20 ± 1.68	9.13 ± 1.68	15.87 ± 2.20
A2.**7f**	7.43 ± 2.06	7.64 ± 1.76	9.15 ± 1.66	15.72 ± 2.37
A3.**7f**	8.03 ± 1.81	8.08 ± 1.56	9.42 ± 1.48	16.53 ± 1.68
A4.**7f**	7.23 ± 1.79	7.58 ± 1.50	9.59 ± 1.32	16.65 ± 2.36

[a] N = 5000.

Table B-2. decalin$_{COM}$···P$_{int}$, N–H$_{COM}$···P$_{int}$, *p*–C$_{COM}$···P$_{int}$ and tail$_{COM}$···P$_{int}$ distances (Å) averaged (Avg) for the last 50 ns of simulations B$_1$.**7a**–B$_2$.**7f** with the corresponding standard deviations (SD).[a]

Simulation ID	decalin$_{COM}$···P$_{int}$	N–H$_{COM}$···P$_{int}$	*p*–C$_{COM}$···P$_{int}$	tail$_{COM}$···P$_{int}$
	Avg ± SD	Avg ± SD	Avg ± SD	Avg ± SD
B$_1$.**7a**	7.07 ± 2.19	6.05 ± 1.81	5.03 ± 2.97	–
B$_2$.**7a**	7.02 ± 2.07	4.49 ± 2.03	0.93 ± 2.14	–
B$_1$.**7b**	9.11 ± 1.66	7.79 ± 1.43	7.14 ± 1.60	7.01 ± 1.85
B$_2$.**7b**	8.09 ± 1.60	8.63 ± 1.45	9.18 ± 2.31	9.43 ± 2.75
B$_1$.**7c**	7.94 ± 2.06	9.31 ± 1.97	11.60 ± 2.08	13.20 ± 2.39
B$_2$.**7c**	8.36 ± 1.93	7.48 ± 1.57	7.33 ± 1.52	7.82 ± 1.93
B$_1$.**7d**	7.37 ± 1.68	7.91 ± 1.78	9.82 ± 2.06	13.69 ± 2.10
B$_2$.**7d**	7.69 ± 1.64	7.66 ± 1.47	8.66 ± 1.43	11.85 ± 1.81
B$_1$.**7e**	7.23 ± 2.11	8.01 ± 1.91	10.31 ± 1.87	15.25 ± 2.14
B$_2$.**7e**	6.40 ± 2.52	7.00 ± 1.96	8.14 ± 1.53	13.37 ± 1.81
B$_1$.**7f**	8.02 ± 1.62	8.04 ± 1.52	9.13 ± 1.56	16.53 ± 1.93
B$_2$.**7f**	10.41 ± 2.14	9.14 ± 1.87	8.57 ± 1.72	14.03 ± 2.80

[a] N = 5000.

Table B-3. Average molecular mechanics energy terms (kcal mol^{-1}) with the corresponding standard deviations for each MD simulation.[a]

Simulation ID	ΔE_{TOTAL}	ΔE_{ELEC}	ΔE_{vdW}
	Avg ± SD	Avg ± SD	Avg ± SD
A_1.**7a**	−108.02 ± 10.65	−44.65 ± 9.20	−63.38 ± 4.87
A_2.**7a**	−84.79 ± 17.83	−27.82 ± 14.48	−56.97 ± 5.97
A_3.**7a**	−101.42 ± 7.41	−34.83 ± 6.73	−66.59 ± 4.31
A_4.**7a**	−88.35 ± 12.98	−28.52 ± 9.94	−59.83 ± 6.00
A_1.**7b**	−104.01 ± 14.15	−32.61 ± 13.56	−71.40 ± 4.36
A_2.**7b**	−115.38 ± 14.14	−40.19 ± 11.28	−75.19 ± 5.31
A_3.**7b**	−113.62 ± 16.87	−50.82 ± 15.04	−62.81 ± 5.50
A_4.**7b**	−99.82 ± 15.67	−31.73 ± 14.47	−68.09 ± 4.60
A_1.**7c**	−107.70 ± 14.26	−32.90 ± 11.43	−74.80 ± 5.13
A_2.**7c**	−118.93 ± 8.38	−34.81 ± 6.23	−84.12 ± 5.25
A_3.**7c**	−111.07 ± 13.60	−29.66 ± 9.13	−81.41 ± 6.20
A_4.**7c**	−114.07 ± 18.07	−36.28 ± 15.42	−77.79 ± 5.24
A_1.**7d**	−80.51 ± 7.60	−13.85 ± 5.85	−66.67 ± 5.49
A_2.**7d**	−124.48 ± 11.32	−32.30 ± 8.41	−92.18 ± 5.49
A_3.**7d**	−109.17 ± 8.84	−22.78 ± 6.92	−86.40 ± 4.90
A_4.**7d**	−138.89 ± 8.03	−47.32 ± 6.65	−91.57 ± 4.76
A_1.**7e**	−149.52 ± 8.35	−47.84 ± 6.72	−101.68 ± 4.79
A_2.**7e**	−130.64 ± 10.62	−32.55 ± 8.89	−98.09 ± 5.30
A_3.**7e**	−128.90 ± 17.21	−36.53 ± 12.62	−92.37 ± 6.86
A_4.**7e**	−127.90 ± 9.01	−30.56 ± 6.87	−97.34 ± 5.85
A_1.**7f**	−129.55 ± 14.37	−29.23 ± 12.08	−100.33 ± 5.54
A_2.**7f**	−119.45 ± 10.41	−18.52 ± 6.20	−100.93 ± 7.86
A_3.**7f**	−131.60 ± 14.20	−24.40 ± 12.84	−107.20 ± 5.09
A_4.**7f**	−130.03 ± 13.36	−28.27 ± 9.35	−101.76 ± 7.62

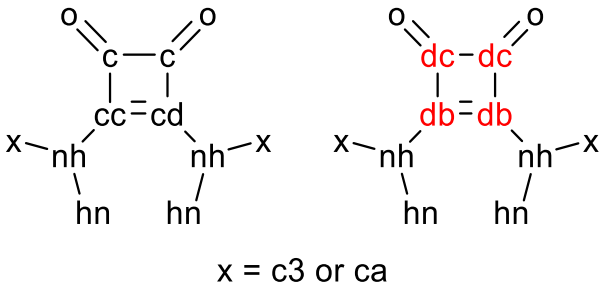
[a] N = 500 (100 ps interval between snapshots).

Appendix C.
Additional data and extended discussion of Chapter IV

C.1. Force field parameterisation of the squaramide core

The preliminary gas-phase MM optimisation of the twelve squaramide derivatives with the General AMBER Force Field (GAFF) default force field parameters and atom types, as given in Scheme C-1 and Table C-1, led to shorter cc-cd distances than in the crystal structures, while the c -c distances were longer. Furthermore, these structures were also optimised at the B3LYP/6-31G* level (in agreement with GAFF's current development), leading to bond lengths like those obtained *via* MM optimisations with default GAFF parameters. On the other hand, the analysis of the crystal structures revealed that most of these molecules are self-assembled, typically *via* N–H···O=C hydrogen bonding interactions, or interacting with solvent molecules or anion guests. These intermolecular interactions have a structural impact on the cc-cd and c -c distances discussed above. In these circumstances, given that the MD simulations take place in condensed phase (a membrane system composed of POPC lipids and water molecules), we decided to parameterise the squaramide core. The equilibrium bond lengths (r_{eq}) and equilibrium bond angles (θ_{eq}) were taken from the average distances and angles, respectively, assessed from 48 X-ray single crystal structures deposited with CCDC,[1] and three unpublished X-ray single crystal structures. In these 51 single crystal structures, there are 60 independent molecules, with 62 squaramide binding motifs that were used in this parameterisation effort.

New atom types, as identified in Scheme C-1, were created, with non-bonded and torsion angles parameters of GAFF's atom types cc/cd and c , but with bond lengths and bond angles force constants calculated in agreement with GAFF,[2,3] from the crystallographic data. The newly determined equilibrium bond lengths and bond angles are listed in Table C-1, along with the corresponding force constants (K_r and K_θ, respectively). Concerning the C-C bonds on the squaramide motif, the new bond length parameters assessed from the crystal structures are much closer to each other than the corresponding GAFF default parameters (1.408, 1.496 and 1.470 Å *vs.* 1.371, 1.550 and 1.462 Å). Moreover, the X-ray derived bond angles of the squaramide core are closer to the 90° of a right angle then GAFF's ones, which are much more obtuse (88.26 and 91.70° *vs.* 111.67 and 121.42°).



Scheme C-1. Schematic representation of the squaramide moiety with the atom types: left – default GAFF atom types; right – new atom types (in red) created for this specific parameterisation.

Table C-1. Force field parameters for the squaramide moiety: default GAFF parameters and X-ray derived parameters.

GAFF default parameters			X-ray derived parameters		
Bond lengths					
Parameter	r_{eq} (Å)	K_r (kcal mol^{-1} Å^{-2})	*Parameter*	r_{eq} (Å)	K_r (kcal mol^{-1} Å^{-2})
cc-cd	1.371	504.00	db-db	1.408	447.30
c -c	1.550	290.10	dc-dc	1.496	340.50
cc-c cd-c	1.462	377.40	db-dc	1.470	368.46
c -o	1.214	648.00	dc-o	1.224	624.81
cc-nh cd-nh	1.364	449.00	db-nh	1.334	496.30
nh-hn	1.014	401.20	nh-hn[a]	1.014	401.20
ca-nh	1.364	449.00	ca-nh[a]	1.364	449.00
c3-nh	1.458	332.70	c3-nh[a]	1.458	332.70
Bond angles					
Parameter	θ_{eq} (°)	K_θ (kcal mol^{-1} rad^{-2})	*Parameter*	θ_{eq} (°)	K_θ (kcal mol^{-1} rad^{-2})
c -c -cc c -c -cd	111.67	64.02	dc-dc-db	88.26	73.24
c -cc-cd c -cd-cc	121.42	65.25	dc-db-db	91.70	73.99
cd-cc-nh cc-cd-nh	123.89	68.69	db-db-nh	132.04	66.27
c -cc-nh c -cd-nh	118.57	67.80	dc-db-nh	136.21	63.60
o -c -c	120.99	67.16	o -dc-dc	135.36	65.15
o -c -cc o -c -cd	125.71	68.91	o -dc-db	136.35	65.75
cc-nh-hn cd-nh-hn	117.16	48.86	db-nh-hn[b]	117.16	49.79
cc-nh-ca cd-nh-ca	129.77	63.77	db-nh-ca[b]	129.77	64.46
cc-nh-c3 cd-nh-c3	119.23	64.17	db-nh-c3[b]	119.23	64.75

[a)] Parameters taken directly from GAFF; [b)] Bond angle and nh-hn, nh-ca and nh-c3 bond lengths taken directly from GAFF.

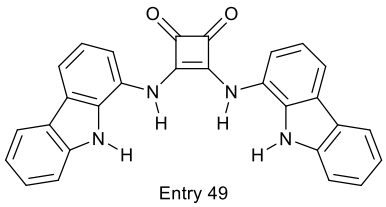
To further validate this parameterisation effort, both GAFF default and the newly developed parameters were used to undertake gas-phase MM optimisations on the 60 independent molecules obtained from the 51 single crystal structures used to derive the X-ray parameters. The Root-Mean-Square Deviation (RMSD) values listed in Table C-2 were calculated between the X-ray crystal structures and the MM minimised structures, with either set of parameters and standard RESP atomic charges calculated in agreement with GAFF development.[2, 3]

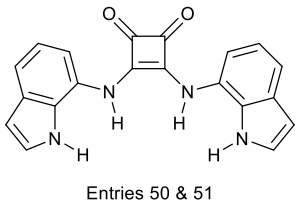
Table C-2. RMSD values calculated between X-ray crystal and MM optimised structures with RESP atomic charges and two force fields for the squaramide core.

CCDC RefCode	Reference	GAFF default parameters		X-ray derived parameters	
		Whole molecule[a]	Squaramide core[b]	Whole molecule[a]	Squaramide core[b]
AKOFIQ[c]	4	0.571	0.073	0.549	0.056
		0.588	0.047	0.568	0.036
AKOFOW	4	0.619	0.031	0.608	0.027
AKOGAJ	4	0.237	0.042	0.218	0.033
AKOGEN[c]	4	0.400	0.030	0.381	0.015
		1.209	0.039	1.196	0.044
CIBQIN	5	0.162	0.058	0.199	0.052
CIPYIL	6	1.306	0.050	0.467	0.037
CIPYOR	6	0.873	0.049	0.366	0.022
CIPYUX	6	0.846	0.062	0.468	0.041
CIPZAE	6	1.070	0.047	0.439	0.018
CIPZEI[c]	6	1.020	0.047	0.628	0.043
		1.027	0.043	0.668	0.031
DICQIQ	7	0.059	0.044	0.047	0.039
EWOCAV	8	0.236	0.066	0.353	0.062
EWOCAV01	8	0.238	0.069	0.355	0.064
FATSUP	9	0.307	0.038	0.171	0.063
FAWZAG	10	0.892	0.061	0.740	0.031
FAWZUA	10	0.366	0.067	0.294	0.041
FAXBEN	10	0.649	0.128	0.730	0.124
FAXBUD	10	0.309	0.062	0.245	0.029
FAXCAK	10	0.482	0.062	0.422	0.032
FAXCEO	10	0.328	0.061	0.283	0.031
FAXCIS[c]	10	0.812	0.060	1.012	0.073
		0.812	0.077	0.979	0.081
GAHMEH	11	0.177	0.060	0.272	0.055
GAQFAH	12	0.571	0.039	0.539	0.043
LANVOM	13	0.732	0.051	0.709	0.033
MUYYIP	14	0.470	0.066	0.386	0.038
MUYYOV[c]	14	0.429	0.104	0.351	0.085
		0.552	0.068	0.351	0.031
NANQUO	15	0.247	0.048	0.310	0.063
NANQUO02[c]	16	0.190	0.042	0.258	0.063
		0.244	0.050	0.311	0.078
NIZXIE	17	1.155	0.100	1.124	0.031
NIZXOK	17	0.140	0.048	0.187	0.045
NOLRIQ	18	1.930	0.075	1.942	0.067
OMUKAJ	19	0.656	0.072	0.592	0.047
OMUKEN	19	0.322	0.040	0.477	0.033
QORQIY	20	0.693	0.047	0.725	0.067
RAKZAG	21	0.622	0.043	0.602	0.028
REPGIE	22	0.474	0.088	0.840	0.069
SEWFAD	23	0.308	0.074	0.385	0.080
UBABAC	24	0.985	0.053	2.118	0.052
UBABAC01	24	0.675	0.040	0.823	0.057
UBADUY	25	0.321	0.034	0.336	0.047
VARDOJ	26	0.282	0.042	0.270	0.043
WECCAK	27	0.040	0.034	0.067	0.039
WECCEO	27	0.058	0.053	0.054	0.049
WECCIS	27	0.055	0.038	0.035	0.035
WELXUI[d]	28	0.281	0.052	0.452	0.048
			0.030		0.031
			0.035		0.032

CCDC RefCode	Reference	GAFF default parameters		X-ray derived parameters	
		Whole molecule[a]	Squaramide core[b]	Whole molecule[a]	Squaramide core[b]
XOPWEF	29	0.247	0.039	0.272	0.040
XOPWEF01	29	0.231	0.040	0.310	0.047
XUFMER[e]	30	0.581	0.065	0.636	0.060
		0.750	0.047	0.798	0.033
		0.770	0.044	0.736	0.028
Entry 49[f]	31	0.788	0.043	1.033	0.038
Entry 50[c,f]	31	0.151	0.051	0.136	0.021
		0.151	0.054	0.130	0.024
Entry 51[f]	31	0.292	0.049	0.277	0.032
Mean ± SD		0.533 ± 0.366	0.054 ± 0.018	0.520 ± 0.391	0.046 ± 0.019

[a] Only the non-hydrogen atoms were considered; [b] "Squaramide core" stands for the four carbon atoms, the two oxygen atoms and the two nitrogen atoms identified in Scheme C-1; [c] The asymmetric unit contains two independent molecules; [d] This molecule contains three squaramide binding units; [e] The asymmetric unit contains three independent molecules; [f] These molecules are sketched in Scheme C-2.





Scheme C-2. Structures of squaramide derivatives used in the force field parameterisation effort that are not yet available in the CCDC database.[31]

Table C-3. Average number of water molecules solvating the squaramide chloride complexes within the first solvation shell (cut-off = 3.4 Å).[a]

MD run	Squaramide derivative		Carbonyl groups[b]		Complexed chloride		*N*
	Avg ± SD	Range	Avg ± SD	Range	Avg ± SD	Range	
W'$_1$.**A1**	9.2 ± 3.6	[1.0 ; 23.0]	0.4 ± 0.8	[0.0 ; 8.0]	3.7 ± 1.0	[0.0 ; 8.0]	10000
W'$_1$.**A2**	7.2 ± 2.5	[1.0 ; 20.0]	0.2 ± 0.5	[0.0 ; 7.0]	3.9 ± 0.9	[1.0 ; 8.0]	10000
W'$_1$.**A3**	9.7 ± 10.4	[0.0 ; 56.0]	0.6 ± 1.9	[0.0 ; 14.0]	3.7 ± 1.1	[1.0 ; 8.0]	10000
W'$_1$.**B1**	8.3 ± 3.0	[1.0 ; 23.0]	0.5 ± 0.8	[0.0 ; 6.0]	3.5 ± 1.2	[1.0 ; 8.0]	10000
W'$_1$.**B2**	8.1 ± 3.1	[1.0 ; 29.0]	0.3 ± 0.7	[0.0 ; 6.0]	4.1 ± 1.1	[0.0 ; 9.0]	10000
W'$_1$.**B3**	10.1 ± 3.6	[1.0 ; 23.0]	0.4 ± 0.7	[0.0 ; 7.0]	3.8 ± 0.9	[1.0 ; 8.0]	10000
M'$_1$.**A1**	8.8 ± 3.0	[0.0 ; 22.0]	0.4 ± 0.8	[0.0 ; 6.0]	3.8 ± 1.0	[0.0 ; 7.0]	10000
M'$_1$.**A2**	6.6 ± 2.6	[0.0 ; 17.0]	0.1 ± 0.5	[0.0 ; 5.0]	3.5 ± 1.3	[0.0 ; 8.0]	10000
M'$_1$.**A3**	6.0 ± 2.7	[0.0 ; 19.0]	0.0 ± 0.3	[0.0 ; 5.0]	3.3 ± 1.0	[0.0 ; 7.0]	10000
M'$_1$.**B1**	10.0 ± 3.2	[0.0 ; 25.0]	0.5 ± 0.8	[0.0 ; 6.0]	4.2 ± 1.0	[1.0 ; 9.0]	10000
M'$_1$.**B2**	9.3 ± 2.7	[1.0 ; 21.0]	0.6 ± 0.8	[0.0 ; 6.0]	4.1 ± 0.9	[1.0 ; 8.0]	10000
M'$_1$.**B3**	8.9 ± 3.0	[1.0 ; 21.0]	0.4 ± 0.6	[0.0 ; 5.0]	3.9 ± 1.0	[1.0 ; 8.0]	10000
M$_1$.**C1**	11.0 ± 3.0	[1.0 ; 26.0]	1.6 ± 1.7	[0.0 ; 11.0]	2.7 ± 0.8	[0.0 ; 8.0]	10000
M$_2$.**C1**	10.3 ± 3.3	[1.0 ; 25.0]	1.4 ± 1.4	[0.0 ; 10.0]	2.5 ± 0.9	[0.0 ; 6.0]	10000
M$_1$.**C2**	9.5 ± 3.5	[0.0 ; 25.0]	2.5 ± 1.9	[0.0 ; 14.0]	2.2 ± 0.9	[0.0 ; 6.0]	10000
M$_2$.**C2**	10.0 ± 2.8	[2.0 ; 22.0]	2.0 ± 1.6	[0.0 ; 11.0]	2.5 ± 0.9	[0.0 ; 6.0]	10000
M$_1$.**C3**	8.3 ± 2.9	[0.0 ; 20.0]	1.7 ± 1.3	[0.0 ; 10.0]	2.1 ± 0.9	[0.0 ; 5.0]	10000
M$_2$.**C3**	10.3 ± 3.2	[1.0 ; 24.0]	2.6 ± 1.8	[0.0 ; 11.0]	2.1 ± 0.8	[0.0 ; 7.0]	10000
M$_1$.**D1**$_H$	8.6 ± 2.9	[1.0 ; 23.0]	4.4 ± 2.2	[0.0 ; 16.0]	0.1 ± 0.3	[0.0 ; 2.0]	40000
M$_2$.**D1**$_H$	8.8 ± 2.9	[1.0 ; 23.0]	4.2 ± 2.2	[0.0 ; 16.0]	0.0 ± 0.1	[0.0 ; 1.0]	40000
M$_1$.**D1**$_V$	10.0 ± 3.4	[2.0 ; 28.0]	4.6 ± 2.4	[0.0 ; 19.0]	0.2 ± 0.4	[0.0 ; 3.0]	40000
M$_2$.**D1**$_V$	10.7 ± 3.0	[2.0 ; 25.0]	4.9 ± 2.4	[0.0 ; 17.0]	0.1 ± 0.3	[0.0 ; 3.0]	40000
M$_1$.**D2**$_H$,	8.8 ± 3.0	[1.0 ; 26.0]	4.0 ± 2.1	[0.0 ; 15.0]	0.0 ± 0.1	[0.0 ; 2.0]	40000
M$_2$.**D2**$_H$	9.9 ± 2.8	[1.0 ; 23.0]	4.3 ± 2.1	[0.0 ; 17.0]	0.0 ± 0.2	[0.0 ; 2.0]	40000
M$_1$.**D2**$_V$	10.1 ± 2.7	[2.0 ; 24.0]	4.4 ± 2.1	[0.0 ; 16.0]	0.1 ± 0.3	[0.0 ; 2.0]	40000
M$_2$.**D2**$_V$	9.2 ± 2.8	[1.0 ; 25.0]	3.9 ± 2.1	[0.0 ; 19.0]	0.0 ± 0.2	[0.0 ; 2.0]	40000
M$_1$.**D3**$_H$	9.1 ± 3.4	[0.0 ; 27.0]	4.1 ± 2.2	[0.0 ; 17.0]	0.2 ± 0.4	[0.0 ; 2.0]	40000
M$_2$.**D3**$_H$	10.1 ± 3.1	[1.0 ; 28.0]	4.6 ± 2.4	[0.0 ; 17.0]	0.0 ± 0.1	[0.0 ; 2.0]	40000
M$_1$.**D3**$_V$	8.6 ± 3.7	[0.0 ; 25.0]	4.2 ± 2.3	[0.0 ; 17.0]	0.1 ± 0.3	[0.0 ; 2.0]	40000
M$_2$.**D3**$_V$	8.5 ± 3.1	[1.0 ; 26.0]	3.7 ± 2.1	[0.0 ; 16.0]	0.0 ± 0.1	[0.0 ; 1.0]	40000

[a] These parameters were evaluated throughout the last 100 ns of simulation time for the **A**, **B** and **C** series, and throughout the last 400 ns for the **D** series; [b] Sum of the number of water molecules around the independent C=O groups.

C.2. References

1. F. H. Allen, *Acta crystallographica. Section B, Structural science*, 2002, **58**, 380-388.
2. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
3. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2005, **26**, 114-114.
4. X. Q. Jin, Q. Q. Min, Y. F. Zheng, P. C. Wang, J. Zhu and H. B. Zhou, *Arkivoc*, 2010, 322-335.
5. A. P. Davis, S. M. Draper, G. Dunne and P. Ashton, *Chem. Commun.*, 1999, 2265-2266.
6. R. B. Elmes, P. Turner and K. A. Jolliffe, *Org. Lett.*, 2013, **15**, 5638-5641.
7. N. Fu, A. D. Allen, S. Kobayashi, T. T. Tidwell, S. Vukovic, S. Arumugam, V. V. Popik and M. Mishima, *J. Org. Chem.*, 2007, **72**, 1951-1956.
8. C. E. Silva, H. F. Dos Santos, N. L. Speziali, R. Diniz and L. F. de Oliveira, *J. Phys. Chem. A*, 2010, **114**, 10097-10109.
9. R. S. Muthyala, G. Subramaniam and L. Todaro, *Org. Lett.*, 2004, **6**, 4663-4665.

10. N. Busschaert, I. L. Kirby, S. Young, S. J. Coles, P. N. Horton, M. E. Light and P. A. Gale, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 4426-4430.
11. R. Mattes, G. Johann and C. Pieper, *Monatshefte für Chemie Chemical Monthly*, 1987, **118**, 105-113.
12. A. Lemmerer and S. A. Bourne, *Acta crystallographica. Section E, Structure reports online*, 2012, **68**, o384.
13. A. Frontera, M. Orell, C. Garau, D. Quinonero, E. Molins, I. Mata and J. Morey, *Org. Lett.*, 2005, **7**, 1437-1440.
14. V. Amendola, G. Bergamaschi, M. Boiocchi, L. Fabbrizzi and M. Milani, *Chem. Eur. J.*, 2010, **16**, 4368-4380.
15. B. Lunelli, P. Roversi, E. Ortoleva and R. Destro, *J Chem Soc Faraday T*, 1996, **92**, 3611-3623.
16. C. Gatti, E. May, R. Destro and F. Cargnoni, *J. Phys. Chem. A*, 2002, **106**, 2707-2720.
17. C. Rotger, B. Soberats, D. Quinonero, A. Frontera, P. Ballester, J. Benet-Buchholz, P. M. Deya and A. Costa, *Eur. J. Org. Chem.*, 2008, **2008**, 1864-1868.
18. J. P. Malerich, K. Hagihara and V. H. Rawal, *J. Am. Chem. Soc.*, 2008, **130**, 14416-14417.
19. A. Rostami, A. Colin, X. Y. Li, M. G. Chudzinski, A. J. Lough and M. S. Taylor, *J. Org. Chem.*, 2010, **75**, 3983-3992.
20. T. Kolev, R. W. Seidel, H. Mayer-Figge, M. Spiteller, W. S. Sheldrick and B. B. Koleva, *Spectrochim. Acta. A Mol. Biomol. Spectrosc.*, 2009, **72**, 502-509.
21. J. S. Oh, K. I. Kim and C. E. Song, *Org Biomol Chem*, 2011, **9**, 7983-7985.
22. M. Porel, V. Ramalingam, M. E. Domaradzki, V. G. Young, V. Ramamurthy and R. S. Muthyala, *Chem. Commun.*, 2013, **49**, 1633-1635.
23. Z. Dong, G. F. Qiu, H. B. Zhou and C. N. Dong, *Tetrahedron-Asymmetr*, 2012, **23**, 1550-1556.
24. R. Prohens, A. Portell, C. Puigjaner, S. Tomas, K. Fujii, K. D. M. Harris, X. Alcobe, M. Font-Bardia and R. Barbas, *Crystal Growth & Design*, 2011, **11**, 3725-3730.
25. L. Dai, H. Yang and F. Chen, *Eur. J. Org. Chem.*, 2011, **2011**, 5071-5076.
26. N. Suleymanoglu, R. Ustabas, Y. B. Alpaslan, F. Eyduran, C. Ozyurek and N. O. Iskeleli, *Spectrochim. Acta. A Mol. Biomol. Spectrosc.*, 2011, **83**, 472-477.
27. R. Prohens, A. Portell, C. Puigjaner, R. Barbas, X. Alcobe, M. Font-Bardia and S. Tomas, *Crystengcomm*, 2012, **14**, 5745-5748.
28. C. Jin, M. Zhang, L. Wu, Y. Guan, Y. Pan, J. Jiang, C. Lin and L. Wang, *Chem. Commun.*, 2013, **49**, 2025-2027.
29. A. Portell, R. Barbas, D. Braga, M. Polito, C. Puigjaner and R. Prohens, *Crystengcomm*, 2009, **11**, 52-54.
30. C. Estarellas, M. C. Rotger, M. Capo, D. Quinonero, A. Frontera, A. Costa and P. M. Deya, *Org. Lett.*, 2009, **11**, 1987-1990.
31. N. Busschaert, Private communication, 2013.

Appendix D.

Additional data for Chapter VI

D.1. Additional data for the MD simulations with Mac and Rot in solution

Table D-1. Statistics for the chalcogen bonding interactions evaluated throughout 50 ns in the individual MD simulations of **Rot** halide complexes.[a]

Anion	Run	Se···A distances (Å)[b]		C–Se···A angles (°)[b]	
		Avg ± SD	Range	Avg ± SD	Range
Cl^-	1	3.683 ± 0.196	[3.103 ; 5.215]	170.5 ± 3.8	[151.3 ; 180.0]
		3.692 ± 0.197	[3.122 ; 5.134]	171.1 ± 3.7	[150.6 ; 180.0]
	2	3.723 ± 0.226	[3.107 ; 13.138]	170.2 ± 4.1	[107.8 ; 180.0]
		3.690 ± 0.215	[3.134 ; 15.129]	170.4 ± 4.0	[146.0 ; 179.9]
	3	3.731 ± 0.217	[3.109 ; 6.034]	170.1 ± 4.1	[131.9 ; 179.9]
		3.719 ± 0.226	[3.148 ; 7.246]	170.6 ± 4.0	[149.3 ; 179.9]
Br^-	1	4.057 ± 0.280	[3.402 ; 7.186]	170.0 ± 4.8	[138.8 ; 180.0]
		4.035 ± 0.274	[3.357 ; 6.673]	170.0 ± 4.7	[131.4 ; 179.9]
	2	4.009 ± 0.257	[3.351 ; 8.966]	170.2 ± 4.6	[112.0 ; 179.9]
		4.067 ± 0.286	[3.373 ; 11.487]	170.1 ± 4.6	[135.7 ; 180.0]
	3	3.986 ± 0.237	[3.349 ; 7.633]	170.7 ± 4.3	[133.5 ; 180.0]
		3.978 ± 0.233	[3.329 ; 6.527]	170.5 ± 4.5	[129.1 ; 179.9]
I^-	1	4.328 ± 0.290	[3.623 ; 9.958]	169.4 ± 5.3	[132.4 ; 179.9]
		4.326 ± 0.294	[3.649 ; 9.534]	169.9 ± 5.0	[133.0 ; 179.9]
	2	4.284 ± 0.250	[3.645 ; 7.337]	169.9 ± 5.0	[134.8 ; 179.9]
		4.267 ± 0.245	[3.609 ; 7.851]	170.3 ± 4.8	[133.2 ; 180.0]
	3	4.341 ± 0.277	[3.653 ; 7.138]	169.6 ± 5.0	[125.3 ; 180.0]
		4.356 ± 0.290	[3.619 ; 7.154]	169.1 ± 5.3	[134.1 ; 180.0]

[a] N = 50000 for each independent MD run; [b] A stands for Cl^-, Br^- or I^-.

Table D-2. Statistics for the hydrogen bonding interactions evaluated throughout 50 ns in the individual MD simulations of **Rot** halide complexes.[a]

Anion	Run	C···A distances (Å)[b]		C–H···A angles (°)[b]	
		Avg ± SD	Range	Avg ± SD	Range
Cl^-	1	3.771 ± 0.182	[3.162 ; 4.601]	136.1 ± 14.2	[54.1 ; 179.3]
	2	3.747 ± 0.193	[3.094 ; 9.835]	132.4 ± 16.1	[49.0 ; 179.7]
	3	3.777 ± 0.186	[3.128 ; 4.540]	135.1 ± 15.1	[69.6 ; 179.3]
Br^-	1	3.936 ± 0.193	[3.266 ; 4.753]	135.9 ± 14.4	[63.9 ; 179.0]
	2	3.939 ± 0.192	[3.282 ; 6.810]	136.4 ± 14.5	[68.1 ; 178.6]
	3	3.949 ± 0.189	[3.242 ; 4.681]	135.2 ± 13.3	[52.4 ; 178.4]
I^-	1	4.144 ± 0.190	[3.488 ; 4.927]	136.5 ± 15.8	[51.9 ; 179.1]
	2	4.134 ± 0.199	[3.444 ; 5.043]	134.1 ± 15.1	[60.4 ; 179.1]
	3	4.144 ± 0.191	[3.471 ; 4.897]	137.0 ± 16.7	[42.4 ; 179.7]

[a] N = 50000 for each independent MD run; [b] A stands for Cl^-, Br^- or I^-.

Table D-3. Statistics for the chalcogen bonding interactions evaluated throughout 50 ns in the individual MD simulations of **Mac** halide complexes.[a]

Anion	Run	Se···A distances (Å)[b]		C–Se···A angles (°)[b]	
		Avg ± SD	Range	Avg ± SD	Range
Cl^-	1	3.758 ± 0.239	[3.134 ; 7.975]	168.9 ± 4.6	[126.1 ; 180.0]
		3.747 ± 0.236	[3.120 ; 7.831]	169.3 ± 4.5	[118.5 ; 179.9]
	2	3.762 ± 0.245	[3.148 ; 8.685]	168.9 ± 4.7	[115.1 ; 179.8]
		3.741 ± 0.234	[3.108 ; 8.680]	169.4 ± 4.5	[131.6 ; 179.9]
	3	3.757 ± 0.236	[3.141 ; 8.999]	169.2 ± 4.5	[120.9 ; 179.9]
		3.731 ± 0.226	[3.130 ; 7.546]	169.3 ± 4.5	[124.4 ; 179.9]
Br^-	1	4.083 ± 0.323	[3.380 ; 10.936]	167.9 ± 5.7	[111.2 ; 180.0]
		4.083 ± 0.321	[3.347 ; 10.131]	168.2 ± 5.7	[111.6 ; 180.0]
	2	4.056 ± 0.296	[3.318 ; 7.149]	168.6 ± 5.4	[114.7 ; 179.9]
		4.098 ± 0.312	[3.356 ; 8.006]	167.7 ± 5.6	[118.6 ; 180.0]
	3	4.070 ± 0.319	[3.320 ; 10.243]	168.2 ± 5.7	[95.4 ; 179.9]
		4.081 ± 0.320	[3.354 ; 10.147]	167.9 ± 5.7	[112.1 ; 179.9]
I^-	1	4.451 ± 0.350	[3.613 ; 7.679]	166.4 ± 6.8	[108.1 ; 179.9]
		4.395 ± 0.319	[3.627 ; 7.846]	167.7 ± 6.0	[119.8 ; 179.8]
	2	4.394 ± 0.316	[3.602 ; 9.272]	167.9 ± 5.7	[106.0 ; 179.9]
		4.426 ± 0.333	[3.612 ; 9.153]	166.9 ± 6.4	[94.9 ; 179.8]
	3	4.421 ± 0.326	[3.613 ; 7.159]	167.1 ± 6.2	[119.9 ; 180.0]
		4.388 ± 0.314	[3.603 ; 7.207]	167.8 ± 5.8	[106.4 ; 179.9]

[a] N = 50000 for each independent MD run; [b] A stands for Cl^-, Br^- or I^-.

Table D-4. Average A···MAC_{cent} (Å) distances,[a] between the anion guests and the centre of mass of the macrocycle of **Rot** or **Mac**, estimated for the individual MD simulations.[b]

Anion	Run	**Rot**		**Mac**	
		Avg ± SD	Range	Avg ± SD	Range
Cl^-	1	4.386 ± 0.443	[2.988 ; 6.799]	5.673 ± 0.605	[3.109 ; 10.141]
	2	4.362 ± 0.488	[3.002 ; 14.183]	5.778 ± 0.560	[3.142 ; 11.474]
	3	4.637 ± 0.574	[3.156 ; 7.300]	5.690 ± 0.576	[3.195 ; 10.377]
Br^-	1	4.759 ± 0.448	[3.268 ; 7.541]	5.952 ± 0.743	[2.990 ; 13.116]
	2	4.805 ± 0.471	[3.424 ; 9.736]	6.112 ± 0.631	[3.490 ; 9.346]
	3	4.850 ± 0.659	[3.228 ; 8.250]	6.056 ± 0.627	[3.746 ; 12.724]
I^-	1	5.242 ± 0.639	[3.369 ; 9.373]	6.385 ± 0.665	[3.768 ; 9.856]
	2	4.745 ± 0.372	[3.513 ; 7.747]	6.404 ± 0.668	[3.530 ; 10.799]
	3	5.504 ± 0.632	[3.851 ; 8.157]	6.352 ± 0.693	[3.333 ; 9.443]

[a] A stands for Cl^-, Br^- or I^-; [b] N = 50000 for each independent MD run.

Table D-5. Solvation of the chalcogen bonded halides in independent MD simulations of **Rot** or **Mac**.[a]

Anion	Run	**Rot**				**Mac**			
		No of water molecules		No of acetone molecules		No of water molecules		No of acetone molecules	
		Avg ± SD	Range	Avg ± SD	Range	Avg ± SD	Range	Avg ± SD	Range
Cl^-	1	4.0 ± 0.9	[1 ; 8]	0.5 ± 0.7	[0 ; 4]	3.1 ± 1.0	[1 ; 7]	2.1 ± 1.1	[0 ; 6]
	2	3.9 ± 0.9	[1 ; 8]	0.6 ± 0.7	[0 ; 4]	3.4 ± 1.0	[1 ; 8]	2.0 ± 1.0	[0 ; 6]
	3	4.1 ± 0.9	[1 ; 8]	0.5 ± 0.6	[0 ; 4]	3.3 ± 1.0	[1 ; 8]	2.1 ± 1.1	[0 ; 7]
Br^-	1	3.6 ± 1.0	[0 ; 8]	0.9 ± 0.8	[0 ; 5]	2.3 ± 1.0	[0 ; 7]	2.6 ± 1.2	[0 ; 7]
	2	3.6 ± 1.0	[1 ; 8]	0.9 ± 0.8	[0 ; 5]	2.5 ± 1.0	[0 ; 7]	2.5 ± 1.1	[0 ; 7]
	3	3.5 ± 1.0	[0 ; 7]	0.9 ± 0.8	[0 ; 5]	2.3 ± 1.1	[0 ; 7]	2.6 ± 1.2	[0 ; 7]
I^-	1	2.5 ± 1.2	[0 ; 7]	1.0 ± 0.9	[0 ; 6]	1.1 ± 0.9	[0 ; 5]	2.2 ± 1.2	[0 ; 7]
	2	2.3 ± 1.1	[0 ; 7]	1.0 ± 0.9	[0 ; 5]	1.0 ± 0.9	[0 ; 6]	2.3 ± 1.2	[0 ; 7]
	3	2.5 ± 1.1	[0 ; 8]	1.0 ± 0.9	[0 ; 6]	0.9 ± 0.9	[0 ; 6]	2.3 ± 1.2	[0 ; 8]

[a] N = 50000 for each independent MD run.

Table D-6. Solvation of the SeMe binding units in independent MD simulations of **Rot** or **Mac**,[a] either free or chalcogen bonded to anion guests.

Anion	Run	**Rot**				**Mac**			
		No of water molecules		No of acetone molecules		No of water molecules		No of acetone molecules	
		Avg ± SD	Range	Avg ± SD	Range	Avg ± SD	Range	Avg ± SD	Range
Unbound	1	0.3 ± 0.7	[0 ; 6]	1.3 ± 1.0	[0 ; 7]	0.1 ± 0.2	[0 ; 3]	2.4 ± 1.1	[0 ; 7]
	2	0.1 ± 0.4	[0 ; 5]	0.9 ± 0.8	[0 ; 6]	0.0 ± 0.2	[0 ; 3]	2.4 ± 1.1	[0 ; 8]
	3	0.4 ± 0.7	[0 ; 6]	1.2 ± 0.9	[0 ; 5]	0.1 ± 0.2	[0 ; 3]	2.4 ± 1.1	[0 ; 8]
Cl^-	1	0.8 ± 0.8	[0 ; 6]	0.3 ± 0.5	[0 ; 3]	0.8 ± 0.8	[0 ; 5]	1.5 ± 0.9	[0 ; 5]
	2	0.8 ± 0.8	[0 ; 5]	0.3 ± 0.5	[0 ; 3]	0.9 ± 0.8	[0 ; 5]	1.4 ± 0.9	[0 ; 5]
	3	0.9 ± 0.8	[0 ; 5]	0.3 ± 0.5	[0 ; 3]	0.8 ± 0.8	[0 ; 5]	1.5 ± 0.9	[0 ; 6]
Br^-	1	0.4 ± 0.6	[0 ; 5]	0.5 ± 0.6	[0 ; 4]	0.5 ± 0.6	[0 ; 4]	1.6 ± 0.9	[0 ; 5]
	2	0.5 ± 0.6	[0 ; 5]	0.6 ± 0.6	[0 ; 4]	0.5 ± 0.7	[0 ; 4]	1.5 ± 0.9	[0 ; 6]
	3	0.5 ± 0.7	[0 ; 4]	0.4 ± 0.6	[0 ; 3]	0.5 ± 0.6	[0 ; 4]	1.6 ± 0.9	[0 ; 5]
I^-	1	0.2 ± 0.5	[0 ; 4]	0.7 ± 0.7	[0 ; 4]	0.2 ± 0.4	[0 ; 3]	1.8 ± 0.9	[0 ; 6]
	2	0.2 ± 0.4	[0 ; 4]	0.6 ± 0.6	[0 ; 4]	0.2 ± 0.4	[0 ; 3]	1.8 ± 0.9	[0 ; 6]
	3	0.3 ± 0.5	[0 ; 4]	0.7 ± 0.7	[0 ; 4]	0.2 ± 0.4	[0 ; 4]	1.8 ± 0.9	[0 ; 6]

[a] N = 50000 for each independent MD run.