Functional Annotation using Molecular Dynamics: A Journey into Ion Channel Function



David Seiferth
Oriel College

A thesis submitted for the degree of
Doctor of Philosophy

Supervised by
Professor S. J. Tucker and Professor P. C. Biggin

Department of Condensed Matter Physics
University of Oxford

Trinity Term, 2024

Acta est fabula, plaudite.

Declaration

I declare that this thesis is entirely my own work and, except where stated, describes my own research.

David Seiferth
Oriel College

Acknowledgements

I would first like to express my gratitude to Professor Philip Biggin and Professor Stephen Tucker for their support and guidance as my DPhil supervisors. They provided a good balance of direction and independence, allowing me to steer the project in its finer details. I am also grateful to Dr Dheeraj Prakaash and Dr Zhiyi Wu for their assistance with technical challenges, and to all members of the structural bioinformatics and computational biochemistry (SBCB) group for their insightful discussions, with special thanks to Dr Linda Phan and Dr Simon Lichtinger. I would also like to extend my appreciation to the Biochemistry Department's support staff for maintaining the research infrastructure, particularly Borut Rozman and John Elder from the IT office. For their help in navigating the administrative aspects of my DPhil, I thank the Interdisciplinary Bioscience Doctoral Training Programme and Oriel College. I am also grateful to the Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) for funding this DPhil, as well as to Oriel College for supporting my conference travel through special grants. I would also like to extend my sincere thanks to Dr Shayan Hundrieser and Dr Linda Phan for their assistance in proofreading parts of my thesis. The SBCB community thrives not only on academic excellence but also on its unique culture. I am especially thankful to Dr Simon Lichtinger for leading the SBCB Christmas band and for his infectious enthusiasm during our gym sessions, which I enjoyed alongside Cyril Schroeder, Mingyu Qin, Sacha Salphati, Nitishwer Anand, Dr Iain Smith, and Dr Dheeraj Prakaash. Beyond my research group, my DPhil experience has been deeply enriched by extracurricular activities. While it is impossible to acknowledge everyone, I would like to highlight my gratitude for the time spent with the Hertford and Trinity College choirs, the Oriel College Boat Club, and the Oxford University Amateur Boxing Club. As the Romans famously said, *mens sana in corpore sano* — a healthy mind in a healthy body. Staying active helped me maintain balance, and I am grateful for the companionship of Sacha Salphati and Camila on countless bike tours and other adventures. Lastly, I am profoundly grateful to my family, my parents and my siblings, for their love and encouragement throughout my life.

Abstract

Ion channels are integral to the healthy function of all living cells. They are involved in maintaining cellular homeostasis by regulating the concentration of ions across the cell membrane. This regulation is vital for a wide range of cellular processes, including volume control, pH balance, and the regulation of metabolic activities. Ion channels embedded in the cell membrane are responsible for the controlled flow of ions such as sodium, potassium, calcium, and chloride into and out of cells. In the nervous system, ion channels enable the rapid transmission of signals across neurons, allowing for the coordination of complex processes such as muscle contraction, sensory perception, and cognitive functions.

In the past two decades, advances in structural biology, coupled with recent AI-driven structure prediction algorithms like AlphaFold, have unveiled numerous 3D ion channel and nanopore structures in various conformational states. However, these structures often require functional annotation to determine whether they are in an open (conductive) or closed state. This can be achieved by analysing the radius profile of the permeation pathway, as ion flux through a channel is typically regulated by steric occlusion of the pore. Additionally, ion permeation can be hindered by the formation of a desolvation barrier caused by hydrophobic residues lining the pore. Minor structural changes can induce transitions in confined hydrophobic regions between wet and dry states, effectively closing the pore without physically constricting the pathway — a phenomenon known as hydrophobic gating, which is reviewed in this thesis. Best practices for functional annotation of ion channels are described, with glycine receptors (GlyRs) serving as model systems in multiple chapters.

GlyRs are members of the Pentameric Ligand-Gated Ion Channel (pLGIC) family and play a key role in modulating synaptic activity. The heteromeric GlyR is responsible for inhibitory action in synapses and is composed of different subunits. Its asymmetric pore raises questions about how asymmetry influences conductance and barriers to permeation. I introduce a new software tool, PoreAnalyser, which quantifies asymmetry in ion channel pores and provides heuristic estimates of channel conductance based on pore profiles. Furthermore, I investigate GlyR and various other ion channels using molecular dynamics simulations. Water and ions behave differently when they permeate the pore of an ion channel than when they are in bulk.

Modeling ions and water in confined environments with molecular mechanics methods is known to be challenging. Consequently, polarisable models have been suggested to improve the accuracy of such calculations. A small model system of an ion carrier is used to assess the level of error in both fixed-charge and polarisable models. Then, conductive GlyR structures are studied using both polarisable and non-polarisable models.

Overall, this thesis reviews techniques for functional annotation of ion channel structures and the phenomenon of hydrophobic gating. It introduces the PoreAnalyser software to quantify pore asymmetry and estimate conductance, and compares the accuracy of fixed-charge and polarisable models in molecular dynamics simulations of ion permeation.

List of publications

The following is a list of articles, that were published as a result of the research carried out for this thesis.

- **Seiferth, D.** and Biggin, P.C., 2024. Exploring the influence of pore shape on conductance and permeation. Biophysical Journal. Biophysical Journal.
- Gibbs, E., Klemm, E., **Seiferth, D.**, Kumar, A., Ilca, S.L., Biggin, P.C., Chakrapani, S., 2023. Conformational transitions and allosteric modulation in a heteromeric glycine receptor. Nat Commun 14, 1363.
- Bharambe, N., Li, Z., **Seiferth, D.**, Balakrishna, A.M., Biggin, P.C., Basak, S., 2024. Cryo-EM structures of prokaryotic ligand-gated ion channel GLIC provide insights into gating in a lipid environment. Nat Commun 15, 2967.
- **Seiferth, D.**, Biggin, P.C., Tucker, S.J., 2022. When is a hydrophobic gate not a hydrophobic gate? Journal of General Physiology 154, e202213210.
- **Seiferth, D.**, Tucker, S.J., Biggin, P.C., 2023. Limitations of non-polarizable force fields in describing anion binding poses in non-polar synthetic hosts. Phys. Chem. Chem. Phys. 25, 17596–17608.

Table of contents

Table of contents xiii

List of figures xvi

List of tables xviii

Glossary xix

1 Introduction 1
- 1.1 Ion Transport ... 1
- 1.2 Ion Channels ... 2
  - 1.2.1 Features and Functions ... 2
  - 1.2.2 Ligand Gated Ion Channels ... 4
- 1.3 Properties of Water and Halide Ions ... 5
- 1.4 Mechanistic Understanding of Ion Channels from Simulations ... 7
- 1.5 Thesis Aims and Outline ... 9

2 Theoretical Background and Computational Methods 11
- 2.1 Theory ... 11
  - 2.1.1 Classical Mechanics ... 11
  - 2.1.2 Thermodynamic Ensembles ... 12
  - 2.1.3 Quantum Mechanics ... 13
- 2.2 Molecular Dynamic Simulations ... 16
  - 2.2.1 Force Fields ... 16
  - 2.2.2 Molecular Dynamics Algorithms ... 26
  - 2.2.3 Boltzmann Inversion of Water Number Density ... 33
  - 2.2.4 Potential of Mean Force Calculations ... 33

3 Simulation Protocols 35
- 3.1 Force Field Parametrisation ... 35
  - 3.1.1 GAFF and Parsley ... 35
  - 3.1.2 AMOEBA Parametrisation ... 35
- 3.2 System Set-up ... 36
  - 3.2.1 Solvent Boxes ... 36
  - 3.2.2 Membrane Systems ... 36
- 3.3 Equilibration Protocols ... 37
  - 3.3.1 Membrane System Equilibration Protocol ... 37
  - 3.3.2 Extended Equilibration Protocol for Pore Lining Residues ... 38
- 3.4 Simulation Settings ... 38
  - 3.4.1 Non-Polarisable Force Field Simulations ... 39
  - 3.4.2 Polarisable Force Field Simulations ... 39
- 3.5 Umbrella Sampling Protocol ... 40
  - 3.5.1 Ion Carriers ... 40
  - 3.5.2 Carbon Nanotubes ... 41
  - 3.5.3 Ion Channels ... 41
- 3.6 Hydration Shell Analysis ... 42
- 3.7 Energy Decomposition Analysis ... 42
  - 3.7.1 Symmetry-Adapted Perturbation Theory (SAPT) ... 43
  - 3.7.2 MD Interaction Energy ... 43
- 3.8 External Potential ... 43

4 Functional Annotation of Ion Channels 45
- 4.1 When is a Hydrophobic Gate not a Hydrophobic Gate? ... 45

4.1.1 Examples of Hydrophobic Gates in Ion Channels 48
4.1.2 Other Mechanisms of Gating Involving Nonpolar Residues 49
4.1.3 Functional Annotation of a Hydrophobic Gate 51
4.2 Examples of Functional Annotation 52
4.2.1 Functional Annotation of Heteromeric Glycine Receptors 52
4.2.2 Functional Annotation of Prokaryotic Ligand-Gated Ion Channel GLIC Structures 55
4.3 Conclusion 56

5 Exploring the Influence of Pore Shape on Conductance and Permeation 57
5.1 Introduction 57
5.2 Methods 60
5.2.1 Pore Geometry Calculations 60
5.2.2 Model Systems 61
5.3 Results and Discussion 62
5.3.1 Case Studies with Biological Ion Channels 62
5.3.2 Comparison with HOLE Capsule Option 64
5.3.3 Carbon Nanotubes to Investigate Conductance 64
5.3.4 Physical Model to Predict Conductance 66
5.4 Conclusion 71

6 Limitations of Non-polarisable Force Fields in Describing Anion Binding Poses in Non-polar Synthetic Hosts 73
6.1 Introduction 73
6.2 Methods 74
6.3 Results and Discussions 75
6.3.1 Ion Interactions with Methylated Biotin[6]uril in Acetonitrile 75
6.3.2 Ion Interactions with Unmethylated Biotin[6]uril in Water 77
6.3.3 Ion Interactions with Unmethylated Biotin-sulfoxide Macrocycle in Water 78
6.3.4 Comparison of the Three Macrocycles 79
6.3.5 Comparison of the Parsley and AMOEBA Force Fields 81
6.3.6 Interaction Energies Between Hosts and Ions 82
6.3.7 Role of Water in Interactions of Anions with Hosts 84
6.4 Conclusion 86

7 Conductive States of the GlyR 88
7.1 Introduction 88
7.2 Structure Selection 88
7.3 Pore Analysis 89
7.4 Heuristic for Conductance 90
7.5 Conductance Estimates with an External Potential 91
7.5.1 Force Field Sensitivity 92
7.5.2 Orientation of Side Chains 93
7.5.3 External Potential and Drude Particles 93
7.6 Potential of Mean Force and Hydration Profile 93
7.7 Conclusion 95

8 Conclusion and Outlook 96
8.1 Summary of Results 96
8.2 Future Directions 96
8.2.1 Conductance Estimates 96
8.2.2 Force Fields 97

Bibliography 99

Appendix A Carbon Nanotubes 114
A.1 Li *et al.* Ion Parameters 114
A.2 Histogram Overlap 114
A.3 PMF Convergence 115

Appendix B Simulations with polarisable and non-polarisable force fields 117
B.1 Ion Carriers 117
B.1.1 PMF Convergence and Histogram Overlap 117
B.1.2 Comparing RMSD and Dihedrals 118
B.1.3 Effects of Refitting Ion Parameters 118
B.1.4 Interaction Energy 119
B.2 GlyR 129
B.2.1 PMF 129
B.2.2 Convergence of PMFs 129
B.2.3 Position of Side Chains 132
B.2.4 Correlation between Conductance and Potential of Mean Force 132

List of figures

1.1 Molecular transport through a biological membrane. 1
1.2 Molecular structures of KcsA, Fluc and GlyR. 2
1.3 Cumulative number of ion channel structures deposited in the Protein Data Bank. 3
1.4 Pore lining residues of GlyR and GLIC. 5
1.5 Simulation box of GlyR. 7
1.6 Time scales associated with ion channels. 8

2.1 Schematic illustration of bonded interactions terms in molecular force fields. 18
2.2 Drude oscillator and multipole expansion. 22
2.3 Van der Waals interactions. 25

3.1 Extended Equilibration Protocol for Pore Lining Residues. 38

4.1 Different gating mechanisms. 46
4.2 The role of pore radius and hydrophobicity. 47
4.3 Best practices for functional annotation of ion channels. 52
4.4 Pore lining residues of the GlyR. 53
4.5 Functional annotation of three GlyR structures. 54
4.6 Functional annotation of two GLIC structures. 56

5.1 Pore geometry calculations with spherical probe particles. 60
5.2 Workflow diagram of the PoreAnalyser software. 61
5.3 Cylindrical carbon nanotubes with varying degrees of asymmetry. 62
5.4 Case Studies of the α1βB GlyR and NaK Channel. 63
5.5 Comparison of pathways found with elliptic probe particles of PoreAnalyser with the CAPSULE option from HOLE. 65
5.6 Evaluating the accuracy in calculating elliptic radii in 18 different carbon nanotube. 65
5.7 Barrier for ion permeation and ion conductance for carbon nanotubes as a function of the ratio between of the elliptic radii. 67
5.8 Potential of mean force (PMF) and ion hydration shell analysis. 68
5.9 Conductivity model for predicting channel conductance. 70
5.10 Conductivity model for Innexin. 71

6.1 Macrocycle made out of six biotin (sulfoxide) units. 74
6.2 Potential of mean force of anions and a biotin macrocycle in acetonitrile and water. 76
6.3 Different binding poses of ions within the biotin macrocycle. 77
6.4 Different binding poses of chloride in the biotin sulfoxide macrocycles. 78
6.5 Potential of mean force of ions and a deprotonated biotin sulfoxide macrocycle in two different conformations. 79
6.6 Potential of mean force of ions and a protonated biotin sulfoxide macrocycle in two different conformations. 80
6.7 Electrostatic surface potentials of different macrocycles. 81
6.8 Force field interaction energies along the reaction coordinate for three ion carriers and chloride. 82
6.9 Comparison of interaction energies between chloride and three hosts with DFT and AMOEBA. 83
6.10 Hydration shell of halide anions along the reaction coordinate. 85
6.11 Wetted and dewetted cavities. 86

7.1 Radius profiles and solvent number density of five glycine receptor (GlyR) models. 89
7.2 Conductance heuristic with PoreAnalyser software. 90
7.3 Conductance estimates. 91
7.4 Conductance comparison for CHARMM36m with and without NBFIX. 92

A.1 Barrier for ion permeation and ion conductance for carbon nanotubes as a function of the ratio between of the elliptic radii, calculated with ion parameters from Li *et al.* 114
A.2 Histogram overlap for carbon nanotubes. 115
A.3 PMF convergence for a 8x50 carbon nanotube. 116
A.4 PMF convergence for a 14x50 carbon nanotube. 116

B.1 Histogram overlap for methylated biotin macrocycle in acetonitrile with Parsley and with AMOEBA. 117
B.2 Convergence analysis for methylated biotin[6]uril in acetonitrile with AMOEBA. 118
B.3 Histogram overlap for unmethylated biotin macrocycle in water with Parsley, GAFF and AMOEBA. 119
B.4 Convergence analysis for methylated biotin[6]uril in water with AMOEBA. 120
B.5 Potential of mean force of ions and a biotin macrocycle in acetonitrile and in water calculated with GAFF and Li *et al.* ion parameters. 120
B.6 Time series of RMSD values comparing the AMOEBA trajectories to the last Parsley frame for three different biotin macrocycles in water and vice versa. 121
B.7 C-N-C-N Dihedral distributions for three different biotin macrocycles in water with two different forcefields and different binding states. 121
B.8 Effects of Refitting Ion Parameters. 122
B.9 Interaction energies along the reaction coordinate for three ion carriers and chloride in water evaluated from AMOEBA conformations with and without water with the AMOEBA force field. 123
B.10 Interaction energies along the reaction coordinate for three ion carriers and chloride in water evaluated from AMOEBA conformations with the Parsley force field for the host and two different chloride parameter sets. 124
B.11 Interaction energies along the reaction coordinate for three ion carriers and chloride evaluated from Parsley conformations with the Parsley force field for the host and the Li *et al.* chloride parameter set. 125
B.12 Interaction energies between chloride and three hosts for the non-central binding pose 125
B.13 Multipole interactions between chloride and the oxygens in the sulfoxide groups. 126
B.14 Radial distribution functions for halide anions and acetonitrile. 126
B.15 Interaction energies between chloride and the biotin macrocyle and a water molecule bound in the cavity and the macrocycle. 127
B.16 Time series of number of solvent molecules within 3 Å of the cavity center for three different biotin macrocycles in water. 128
B.17 Potential of mean force and hydration profile of a chloride ion permeating pores of five GlyR structures. 129
B.18 Histogram overlap for different GlyR models with CHARMM36m and with CHARMM Drude. 130
B.19 Convergence of CHARMM36m PMFs. 130
B.20 Convergence of CHARMM Drude PMFs. 131
B.21 Number of water hydrogens and protein atoms coordinating the chloride anion at the proline -2' and the leucine 9' constriction. 133
B.22 Orientation of leucine at 9' position during 250 ns simulations with 500 mV external potential with the CHARMM36m force field. 133
B.23 Orientation of leucine at 9' position during 250 ns simulations with 500 mV external potential with the CHARMM Drude force field. 134
B.24 Distribution of the distance between chloride nucleus and Drude particle in simulations with no external potential and with external potential. 134
B.25 Distance between chloride nucleus and Drude particle as a proxy for polarisation of chloride. 135
B.26 Correlation between conductance and potential of mean force. 135

List of tables

1.1 Gibbs Free Energy of Hydration and van der Waals Radii. 6

3.1 Membrane System Equilibration Protocol. 37
3.2 Hydration shell analysis. 42

5.1 Carbon nanotubes (CNTs) of length 50 Å and varying HOLE radii 62
5.2 Innexin and LRRC8 structures used to test the conductance model. 70

6.1 Estimates for the binding free energies for biotin-6-uril hexamethyl ester and different anions in acetonitrile. 76
6.2 Estimates for the binding free energies in kcal/mol for biotin-6-uril hexaacid ester and different anions in water. 77
6.3 Absolute Single-Ion Solvation Free Energies in water and in acetonitrile. 78
6.4 Estimates for the binding free energies for biotin-sulfoxide hexaacid ester and different anions in water. 79
6.5 Partial charge comparison 84

7.1 Summary of simulated GlyR structures with their respective pore constrictions and conductance estimates. 93

B.1 Comparison of *gromacs* and *tinker* simulation performance. 117
B.2 Interaction energies in kcal/mol for chloride bound to different biotin macrocycles in water. 122
B.3 AMOEBA interaction energies in kcal/mol for chloride bound to different biotin macrocycles in water. 123
B.4 Summary of pore constrictions and their associated potential of mean force (PMF) barrier. . 131

Glossary

5-HT$_3$R Serotonin Receptor Channel

AMBER Assisted Model Building with Energy Refinement

AMOEBA Atomic Multipole Optimized Energetics for Biomolecular Applications

Asp Aspartic Acid

ATP Adenosine Triphosphate

BCC Bond Charge Correction

BEST1 Bestrophin Chloride Channel

BK Large Conductance Ca^{2+}-activated Channels

CHAP Channel Annotation Package

CHARMM Chemistry at Harvard Macromolecular Mechanics

CNTs Carbon Nanotubes

COM Center of Mass

CRAC Ca^{2+} Release-activated Ca^{2+}

cryo-EM Cryogenic Electron Microscopy

DFT Density Functional Theory

DMA Distributed Multipole Analysis

ECC Electronic Continuum Correction

ECD Extracellular Domain

GAFF General AMBER Force Field

GLIC Gloeobacter Ligand-gated Ion Channel

GltPh Glutamate Transporter

Glu Glutamic Acid

GlyR Glycine Receptor

GPUs Graphics Processing Units

GTOs Gaussian-Type Orbitals

His Histidine

ICD Intracellular Domain

ITC Isothermal Titration Calorimetry

KcsA K Channel of Streptomyces A

LRRC8 Leucine-rich Repeat-containing Protein 8

MD Molecular Dynamics

ML Machine Learning

MM Molecular Mechanics

MscL Large Conductance Mechanosensitive Channel

MscS Small Conductance Mechanosensitive Channel

nAChR Nicotinic Acetylcholine Receptor

NPA Natural Population Analysis

PA PoreAnalyser

PBCs Periodic Boundary Conditions

PDB Protein Data Bank

pLGICs Pentameric Ligand-gated Ion Channels

PME Particle Mesh Ewald

PMF Potential of Mean Force

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

QM Quantum Mechanics

RDF Radial Distribution Function

RESP Restrained Electrostatic Potential Fit

RESPA Reversible Reference System Propagator Algorithm

RMSD Root Mean Square Deviation

SAPT Symmetry-Adapted Perturbation Theory

SK Small Conductance Ca^{2+}-activated Channels

SMA Stryrene Maleic Acid Polymer

STOs Slater-Type Orbitals

TMD Transmembrane Domain

TRP Transient Receptor Potential

US Umbrella Sampling

VDW van der Waals

WHAM Weighted Histogram Analysis Method

Chapter 1
Introduction

1.1 Ion Transport

Biological cell membranes are composed of amphipathic lipid molecules that assemble into a bilayer [1]. These membranes act as physical barriers, separating the intracellular environment from the extracellular space or dividing internal regions of the cell into distinct aqueous compartments [2]. Furthermore, membranes function as boundaries between two salt reservoirs with different concentrations, creating an electric potential difference across the membrane. The central region of these bilayers is hydrophobic, making it relatively impermeable to ions and other polar molecules [2]. To facilitate ion transport in and out of the cell or between intracellular compartments, transmembrane proteins, such as ion channels and transporters, are essential, see Fig. 1.1 [1].

Ion channels are membrane proteins that allow ions to pass through the membrane and are fundamental to the regulation of solute import and export across cellular membranes [3]. They play a key role in maintaining osmotic balance and the overall integrity of the cell. Many ion channels exhibit selectivity for specific ions such as potassium, sodium or chloride [3]. This selectivity is crucial for maintaining osmotic balance within cells [2]. For instance, potassium channels are vital for maintaining the cell's resting membrane potential, while chloride channels help in regulating cell volume and electrical excitability [4]. These channels ensure that the cell remains functional by selectively allowing certain ions to pass through the membrane while blocking others [5]. The generation of electrical signals in neurons depends on the selective permeability of the membrane and the specific ion concentration gradients across it [3]. Ion channels enable this selective permeability along a concentration gradient, while active transporters, another group of membrane proteins, are responsible for establishing and maintaining these ion gradients [1]. Over the course of evolution, diverse families of membrane transport proteins have evolved, each with unique structures and functions [6]. These families include various types of channels and transporters, each adapted to specific physiological roles [2]. Ion channels differ from transporters in their mechanisms of action and energy requirements:

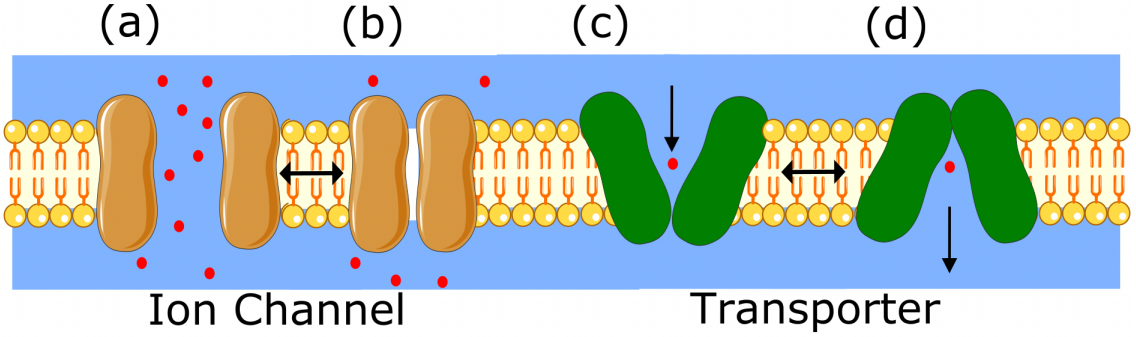


Figure 1.1: Membrane proteins facilitate the transport of solutes (red) across a membrane (yellow). Ion channels (brown) enable the diffusion of ionic species down their electrochemical gradients. Transporters (green) utilise energy to transport ions against their electrochemical gradient. A conductive (a) and a non-conductive (b) conformation of an ion channel, and an outward facing (c) and an inward facing (d) conformation of a transporter are shown.

Ion channels are water-filled, ion-conducting membrane proteins that allow ions to pass through without larger structural changes to the protein itself [3]. Fig. 1.1 depicts a conductive ion channel facilitating ion permeation through its pore and a non-conductive ion channel. This passive diffusion is driven by the transmembrane electrochemical potential difference of each permeant species, resulting in a high flux rate that can reach up to 10^6 ions per second, approaching diffusion-limited values in some channels [7, 8]. In contrast, transporters undergo conformational transitions during the transport process that are in general larger than from what could be expected by thermal motion alone. In Fig. 1.1, the substrate binds to an outward facing transporter and is released to the cell after a structural transition to an inward facing conformation. This structural rearrangement is associated with a higher energetic barrier, making the process slower compared to ion channels. Transporters can facilitate active transport, moving substrates against their concentration gradient using energy from adenosine triphosphate (ATP) hydrolysis or another solute gradient. This efficient transport mechanism is critical for various cellular processes such as regulating cell volume and enabling electrical excitability in neurons [2].

1.2 Ion Channels

1.2.1 Features and Functions

Channel proteins undergo conformational changes in response to external or internal stimuli, allowing or preventing the flow of ions across the membrane as depicted in Fig. 1.1. The process of opening and closing is known as gating: Opening is activation and closing is deactivation or desensitization. Ion channels can be activated by a variety of biochemical and physical stimuli, such as the binding of a ligand to an allosteric site,

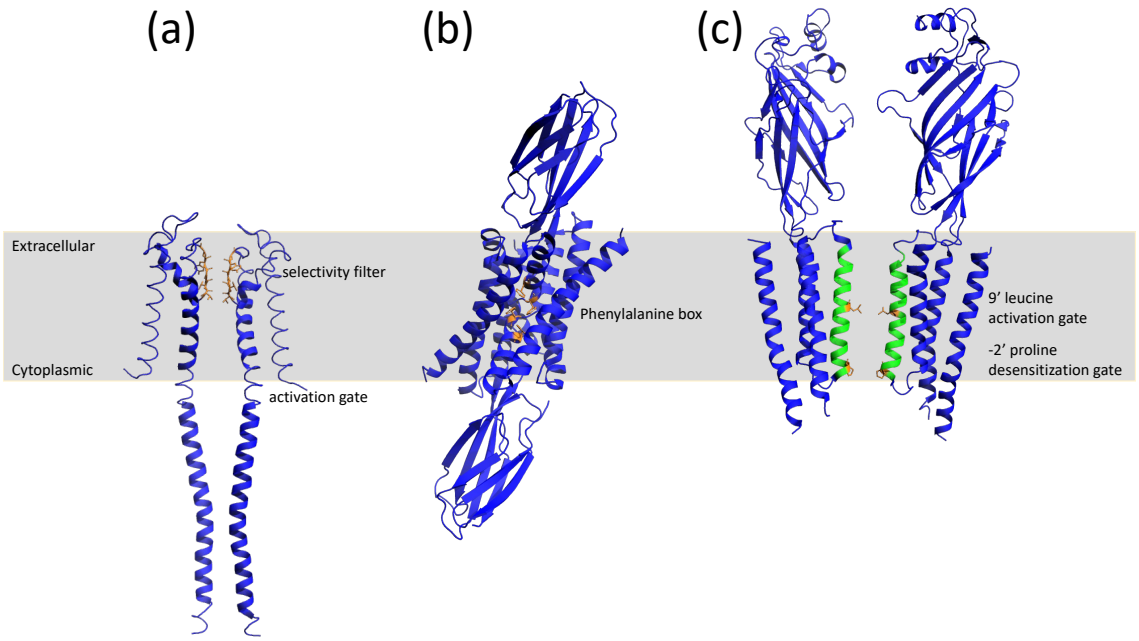


Figure 1.2: Molecular structures of (a) K channel of streptomyces A (KcsA), (b) Fluc and (c) GlyR are shown in a side view of two non-adjacent subunits in blue cartoon representation. Key functional regions are highlighted: The selectivity filter of KcsA, a tetrameric pore-loop potassium channel, is highlighted in orange in (a). Phenylalanine rings a fluoride selective Fluc channel are highlighted in orange in (b). The pore lining M2 helix of a chloride selective GlyR is shown in green in (c) and the residues at the activation and desensitization gate are shown in orange. The respective PDB codes are 3pjs [9], 5a43 [10] and 7tu9 [11]. The approximate position of the membrane is represented by a grey box.

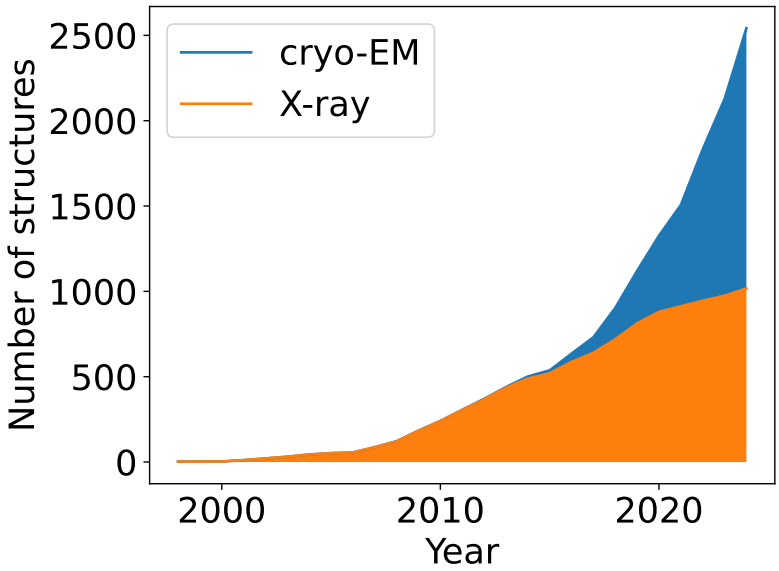


Figure 1.3: Cumulative number of ion channel structures deposited in the Protein Data Bank, as of beginning of September 2024, and its annual growth since the publication of the first ion channel structure [21].

changes in membrane potential, or mechanical tension [2]. When ion channels are activated and open, they typically exhibit selective permeability, allowing only specific classes of ions to pass through while excluding others. This selectivity is achieved through specific binding sites within the pore lining of the ion conduction pathway, which enables the protein to detect particular ions [5]. Most ion channels can distinguish between anions and cations, and some have developed the ability to differentiate between ions with the same charge, such as sodium and potassium ions. For example, potassium channels contain a narrow constriction known as the selectivity filter, which is composed of highly conserved amino acids. Backbone carbonyl groups at several sites within this filter play a role in selectively permitting the passage of potassium ions [9, 12]. The selectivity filter of K channel of streptomyces A (KcsA) is highlighted in orange in Fig. 1.2(a). Selectivity is not limited to cations; for instance, the double-barrelled fluoride channel exhibits a high specificity for fluoride ions over chloride ions. This specificity is conjectured to arise from the coordination of anions with the quadrupolar edges of conserved phenylalanine rings within the channel [10, 13, 14, 15]. A bacterial Fluc homologue and phenylalanine rings are depicted in Fig. 1.2(b). The activation and desensitization gate of a glycine receptor (GlyR) channel are shown in Fig. 1.2(c). There are increasing numbers of ion channel structures that have been discovered and structurally solved [16]. In Fig. 1.3, the number of channel structures solved with X-ray crystallography and cryogenic electron microscopy (cryo-EM) are plotted.

Ion channels regulate ion flux primarily through steric occlusion of their pores. In this thesis, another gating mechanism called hydrophobic gating is reviewed where hydrophobic residues within the pore lining create a barrier to permeation without physically blocking the pathway because the pore transitions between wet and dry states [17, 18, 19]. Ions passing a dehydrated pore would have to fully or partially loose their hydration shell which is associated with an energy penalty. In many recently published cryo-EM studies, ion channel structures show heteromeric subunit assembly, leading to pore asymmetry [11]. Many newly solved structure feature heteromeric subunit assembly, exemplified by synaptic α1βB Glycine and α4β2 Nicotinic receptors [11, 20]. Chloride is the most abundant anion in living organisms and is involved in processes such as controlling the membrane potential, the cell volume and pH balance [22]. Dysregulation of chloride ions can lead to diseases like cystic fibrosis [23]. Synthetic ionophores and artificial chloride channels have therapeutic potential [24], including anticancer activity [25]. The study of ion transport is

essential for understanding ion concentration gradients, but interactions with anions are more complex due to their high solvation energy and the presence of water, which can disrupt stabilizing interactions [26]. The mechanisms underlying selectivity in anion channels are not as well understood as those in their cation counterparts. While significant progress has been made in elucidating how cation channels achieve their high degree of selectivity [12], the processes by which anion channels discriminate between different anions remain less clear. In this thesis, these mechanisms are explored by studying the binding of anions to a narrow cavity within a synthetic ion carrier, as well as investigating the permeation of chloride ions through the anion-selective glycine receptor. By examining the interactions between anions and the synthetic ion carrier, I aim to shed light on the specific factors that contribute to selective anion binding. This includes identifying the structural features of the carrier that facilitate the recognition and binding of specific anions while excluding others.

1.2.2 Ligand Gated Ion Channels

Pentameric ligand-gated ion channels (pLGICs) are a family of ion channels crucial for rapid synaptic transmission. These channels are characterised by their pentameric structure, consisting of five homomeric or heteromeric subunits. Each subunit contributes to the formation of the ion-conducting pore, allowing for the passage of ions across the cell membrane in response to ligand binding [20]. Cys-loop receptors, a subclass of pLGICs, are named for a conserved loop formed by a disulfide bond between two cysteine residues. Among these, cationic Cys-loop receptors are permeable to positively charged ions. The serotonin receptor channel (5-HT_3R) and the nicotinic acetylcholine receptor (nAChR) are cation-selective [3]. 5-HT_3Rs are serotonin-gated ion channels, playing roles in both the central and peripheral nervous systems [27]. They are involved in processes such as pain perception, nausea, and anxiety [28]. Nicotinic acetylcholine receptors are found in both the nervous system and at neuromuscular junctions, they mediate the effects of the neurotransmitter acetylcholine and are crucial for muscle contraction, cognitive functions, and autonomic nervous system responses [29, 30]. In contrast, anionic Cys-loop receptors are permeable to negatively charged ions. GABA receptors are activated by the neurotransmitter gamma-aminobutyric acid (GABA). They are the primary mediators of inhibitory neurotransmission in the central nervous system, playing roles in reducing neuronal excitability and maintaining neural network stability. A GlyR is a chloride ion channel activated by the amino acid glycine. GlyRs contribute to inhibitory neurotransmission in the spinal cord and brainstem, modulating motor and sensory functions. Synaptic GlyRs are asymmetric, consisting of a heteromeric assembly of α and β subunits [11]. The glycine receptor serves as a central case study throughout this thesis. In Chapter 4, various structures of zebrafish α1βGlyR are used to illustrate best practices in functional annotation. In Chapter 5, the asymmetry of the GlyR pore is analysed. Finally, in Chapter 7, conductance values are estimated.

Prokaryotic variants of nAChRs, such as the Gloeobacter Ligand-gated Ion Channel (GLIC) receptor, can provide insights into the fundamental mechanisms of pLGICs. The GLIC receptor, identified in the bacterium *Gloeobacter violaceus*, serves as a model system for studying the structure and function of these channels. Despite differences in sequence and ligand specificity, GLIC shares the characteristic pentameric arrangement and overall architecture of eukaryotic pLGICs [31]. In section 4.2.2, two GLIC structures are used as a case study for functional annotation. The conserved pore architecture of pLGICs has two channel gates as depicted in Fig. 1.2(c): The conserved hydrophobic amino acid at the 9' position within the M2 domain is typically leucine in GlyRs [32] and isoleucine in GLIC [33]. These hydrophobic residues are located in the middle of the pore (see Fig. 1.4) and form the so-called activation gate [11]. In the closed state, the hydrophobic nature of these residues helps to occlude the pore, preventing ion permeation. Upon ligand binding, conformational changes in the receptor lead to the reorientation of these residues, allowing the

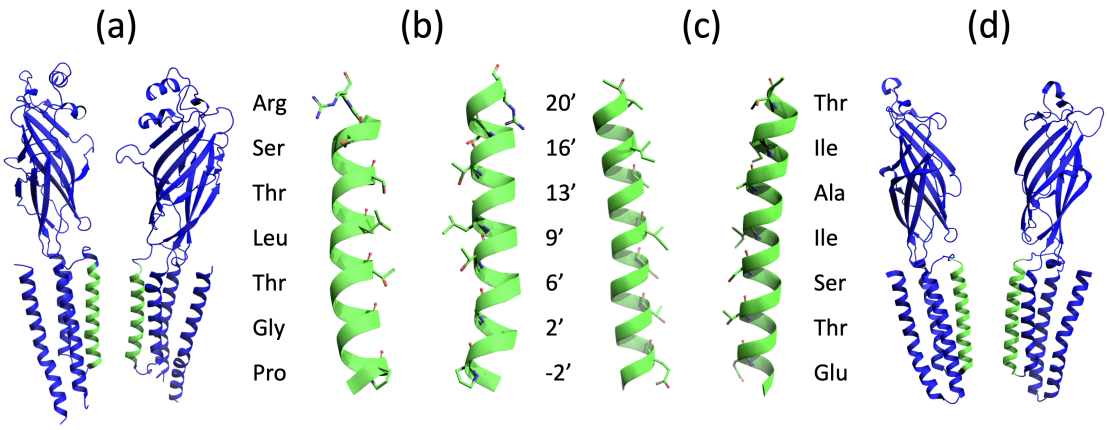


Figure 1.4: Two non-adjacent chains of GlyR, PDB code 7tu9, (a) and GLIC, PDB code 8wcr, (d). Two non-adjacent pore lining M2 helices with seven pore lining residues of GlyR (b) and GLIC (c). The 9' position corresponds to the activation gate and the -2' position to the desensitization gate.

pore to open and ions to pass through. The second gate is the desensitization gate at the intracellular end of the pore: The −2' position in αGlyR is a proline [34] and a glutamic acid in GLIC [33]. The pore-lining residues are depicted in Fig. 1.4. The canonical gating mechanism of pLGICs involves transitions from closed/resting, open/activated, and closed/desensitized states [35]. Different structures in presence of multiple agonist-bound closed states have been resolved and are accessible in the PDB [11, 34, 36, 37]. The glycine-bound receptor populates both open and desensitized states [34]. The family of pentameric ligand-gated ion channels is diverse, with each receptor type playing specific roles in various physiological processes. This diversity is reflected not only in the range of ligands that activate these receptors but also in their ion selectivity and the distinct cellular responses they mediate. Understanding the structural and functional nuances of these channels is essential for uncovering their roles in health and disease, as well as for developing therapeutic interventions targeting these proteins.

1.3 Properties of Water and Halide Ions

Water exhibits anomalous behaviour at the macroscopic level, such as the formation of droplets on a hydrophobic surface due to surface tension. Similar effects also manifest in nanoscale environments, where water molecules experience a strong repelling force from hydrophobic surfaces [38, 39]. When water is confined within a narrow pore that has a hydrophobic interior, it can undergo stochastic transitions between liquid and vapour states, a process referred to as pore dewetting [40]. In such a setting, so-called hydrophobic gates form when water molecules are expelled from the pore, creating an energetically unfavourable barrier for ions, which cannot pass through a vapour-filled pore [4]. This gating mechanism does not rely on physical blockage but rather on the altered energetic landscape caused by water's phase transitions [41]. In contrast, pores lined with polar surfaces retain water in its liquid phase, ensuring continuous ion permeability [17, 42]. Various external factors can influence the behaviour of these hydrophobic gates, including changes in the pore's radius, shifts in surface hydrophobicity (which may occur through allosteric transitions), or the application of external forces such as electric fields [43]. These modifications can induce pore wetting, effectively opening the hydrophobic gate. Temperature [42] and transmembrane voltage [44, 45, 46, 47] have also been shown to play a role, but the pore's radius and hydrophobicity remain the primary factors controlling dewetting and gating [43, 48].

Anions can be ranked according to their ability to stabilize proteins in aqueous solutions, following

Table 1.1: Gibbs Free Energy of Hydration at 298.1 K (kcal/mol) [49] and van der Waals Radii (Å) [50].

Ion	van der Waals (VDW) radius	Gibbs free energy of hydration
Fluoride	1.65	-111.1
Chloride	2.06	-81.3
Bromide	2.27	-75.3
Iodide	2.34	-65.7
Sodium	1.27	-87.2
Potassium	1.67	-70.5

the order established in the Hofmeister series: $F^- < Cl^- < Br^- < I^-$ [51]. This classification is based on the size, polarisability, and solvation characteristics of the anions. Smaller and harder halides, such as fluoride (F^-) and chloride (Cl^-), are less polarisable and form strong, well-defined solvation shells due to their high charge density [52]. These properties make them less likely to interact with hydrophobic interfaces, and they tend to decrease the solubility of nonpolar molecules—a phenomenon known as *salting out* [53]. In contrast, larger and softer anions like bromide (Br^-) and iodide (I^-) are more polarisable and possess weaker solvation shells, allowing them to increase the solubility of nonpolar molecules (*salting in*). They exhibit a greater propensity to migrate towards hydrophobic regions [52]. In Table 1.1, the VDW radii of halide ions, potassium and sodium are summarised. Atoms can be described by different radii: covalent radii of atoms, ionic radii of ions, and VDW radii of neutral atoms are typically derived from crystal structures. Determining the VDW radii of ions poses a challenge because the atomic distances in crystal structures are influenced by both VDW and electrostatic interactions. This overlap makes it difficult to clearly define the VDW sphere of ions [50]. Smith *et al.* discovered that VDW radii, derived from wavefunction analysis using the 0.0015 au electron density contour around atoms, showed excellent agreement with the experimentally determined VDW radii of noble gas atoms and calculated the VDW radii depicted in Table 1.1 based on this criterion [50]. Anion interactions with proteins and other solutes in aqueous environments are further complicated by the high solvation energy of these ions and the disruptive effects of water on stabilizing interactions compared to similarly sized cations [26], see Table 1.1 for free energy of hydration of halide ions compared to potassium or sodium. For instance, water molecules can interfere with direct ion-protein binding by forming extensive hydration shells, especially around smaller anions with stronger solvation energies. The hydration free energies of these halide anions vary systematically with their size and polarisability, as illustrated in Table 1.1. Thermodynamic measurements of salt and electrolyte solvation are relatively simple to obtain; however, separating the total solvation free energy into individual contributions from cations and anions requires reliance on an extrathermodynamic assumption, as it cannot be done directly through measurement alone [49, 54].

Extensive research has also focused on the behaviour of electrolytes at planar hydrophobic interfaces, where intermolecular interactions govern the adsorption of ions [55, 56, 57]. Adsorption leads to a concentration gradient of ionic species near the interface compared to the bulk solution [55]. This ion adsorption is driven primarily by cavitational energy, which arises from the disruption of water's hydrogen-bonding network as ions approach the interface, incurring an energetic cost [56]. Recent simulations have provided insights into the behaviour of halide ions at aqueous/hydrophobic interfaces, following the Hofmeister series [52, 57]. Smaller, hard ions like fluoride (F^-) retain their tightly bound hydration shells and are effectively excluded from the interface. In contrast, larger, softer ions, such as bromide (Br^-) and iodide (I^-), with more flexible and less defined hydration shells, show a greater affinity for the interface [57]. The anisotropic nature of these interfaces induces a dipole in polarisable ions, which is not present in bulk solution. The induced dipole interacts with surrounding water molecules, partially compensating for the ion's loss of hydration as

it adsorbs to the interface [58]. This stabilizing effect enables the ions to be energetically favourable at the interface despite their partial dehydration [59]. While direct experimental observation of ion behaviour at these interfacial regions remains challenging, spectroscopic techniques have provided indirect evidence [60, 61, 62, 63]. These experiments, capable of probing the surface compositions of electrolyte solutions, support the notion of ion-specific effects that align with the Hofmeister series. Additionally, simulations of ion solvation in small aqueous droplets and clusters using polarisable models have shown that halide movement towards the surface is primarily enthalpically driven [59]. The reorganization of water around partially dehydrated ions allows for more favourable water-water interactions, which, in turn, stabilizes the ions at the surface relative to their behaviour in bulk water.

1.4 Mechanistic Understanding of Ion Channels from Simulations

X-ray crystallography and cryo-EM allow for the visualization of ion channels at or near atomic resolution, providing high-resolution snapshots of these proteins [16]. The structural models capture the precise arrangement of atoms within the ion channel. While static pictures provide details about protein structures, they fall short in depicting the dynamic nature of these systems. Ion channels undergo conformational changes, transitioning between non-conductive and conductive states [64]. Understanding these dynamics is essential for comprehending their functional mechanisms. Ion channels are integral membrane proteins that interact with lipids and ligands, which bind to them and influence their activity. Conductive ion channels are wetted, meaning they facilitate the movement of ions and water across the cell membrane [3]. In some cryo-EM structures, some non-protein observed density can be interpreted as lipids, water, or ions. Many ion channels are solved within lipid nanodiscs, which helps preserve their native-like environment during structural determination. However, the nanodisc scaffold can influence the membrane protein structure by altering the properties of the lipid bilayer or directly interacting with the protein of interest [65]. For instance, pLGICs are sensitive to their membrane environment [66]. Dalal *et al.* suggest that larger nanodiscs may approximate a native-like membrane environment better than smaller ones [65]. CryoEM structures

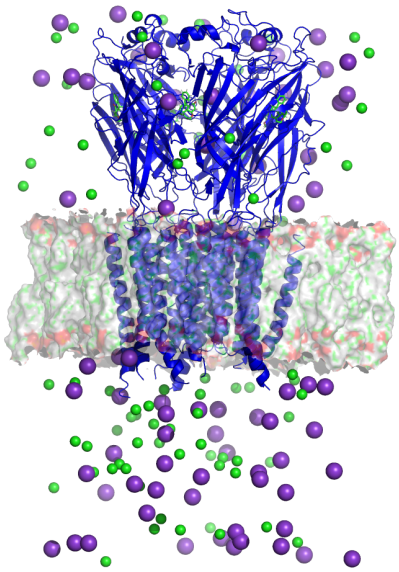


Figure 1.5: Simulation box: GlyR (blue cartoon) with strychnine bound (stick representation) embedded in a POPC bilayer (transparent surface) in water (omitted) with potassium (purple sphere) and chloride (green sphere), PDB code 7tu9.

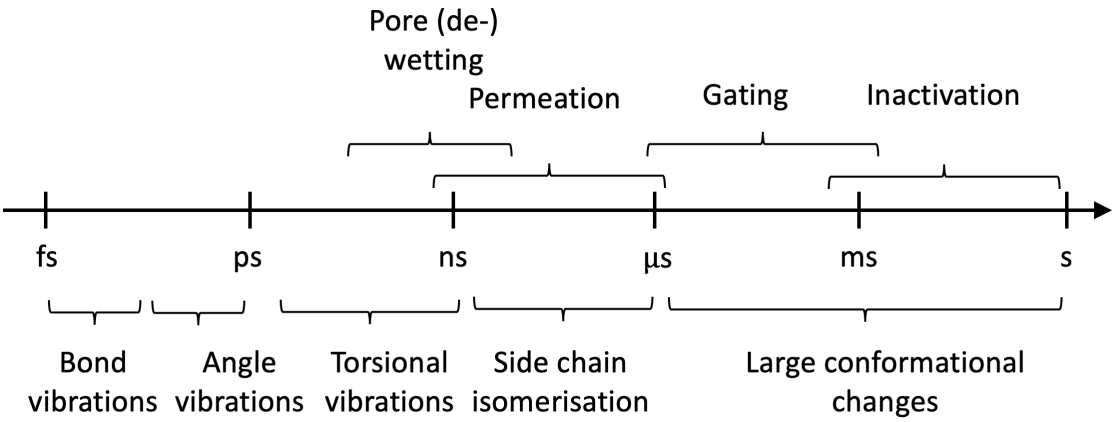


Figure 1.6: Time scales associated with ion channel activity (top) relating to structural fluctuations (bottom). Adapted from Flood *et al.* [67].

often, if at all, show only parts of the ion channel's environment, making it challenging to fully understand the protein's interactions with its surroundings. This limitation highlights the need for complementary techniques to study ion channels in their native states.

Molecular dynamics simulations aim to bridge the gap between static snapshots and dynamic behaviour by studying ion channels in a simulation box resembling its native environment as depicted in Fig. 1.5. These simulations provide a way to observe the temporal evolution of ion channels, offering insights into their functional states. A simulation box containing a GlyR structure embedded in a bilayer surrounded by water and ions is shown in Fig. 1.5. If one were to explain it to someone who is unfamiliar with scientific jargon, one might explain it like this: Computer simulations generate a "movie" from static pictures, revealing the dynamic behaviour of ion channels over time. An analogy can be drawn to the iPhone's "Live Photos" function, which captures a short video before and after a picture is taken, giving a sense of movement. It also allows the user to change the "conformation" of the depicted person from, for instance, a status of closed eyes to open eyes. Likewise, molecular dynamics (MD) simulations can reveal an ensemble of slightly different protein structures. A person might blink and hence their eyes are closed in a picture but the person opens their eyes shortly after. The analogous protein "blinking" could be an ion channel pore that undergoes a transition between a water-filled pore and a dewetting process on a nano second time scale. MD simulations are computationally expensive and typically produce short movies that span up to microseconds. This limitation is analogous to the brief videos captured by the iPhone's "Live Photos" function as opposed to a longer video. They both are characterised by short time scales. While an iPhone actually takes a short movie when the user takes a picture, the MD trajectories are typically generated by describing atoms as particles governed by classical physics with a set of force field parameters, describing interactions between bonded and non-bonded particles.

With increasing computational power and graphics processing units (GPUs) [68], MD simulations are becoming longer and more complex. Initially, simulations involved small systems and ran for picoseconds, starting with simple water boxes. Now, they encompass larger systems and longer durations, providing more detailed insights. Different time scales governing processes in ion channels are depicted in Fig. 1.6. Conformational changes that gate ion channels occur on a μs to ms time scale [67]. The time step of a simulation or the time difference between two frames in a movie is typically determined by the shortest time scale governing the system which are usually bond vibrations with 1-2 fs. Observing larger conformational transitions in simple MD trajectories is challenging because of the wide range of time scales involved.

Wetting and dewetting of ion channel pores can be observed in relatively short (ns) simulations [40] and permeation events can be observed on a ns-μs time scale [69].

Using a Harry Potter analogy, non-wizards (like the author of this thesis) can only create static paintings, while wizards can transform these into dynamic pictures that interact with observers. Similarly, researchers use MD simulations to turn static ion channel structures into dynamic models. Researchers continually ask questions and use MD simulations to seek answers about ion channel states, permeability, and ion passage rates. Unlike magic paintings that converse with wizards, MD simulations provide data through physics-based models, offering indirect yet valuable insights. The Harry Potter analogy is also limited: MD trajectories are not generated by moving a magic want, they are physics based models. The accessible time scales are short and an observer of a MD trajectory cannot interact with the moving protein in real time.

In summary, starting from a high-resolution snapshot of an ion channel obtained through cryo-EM or X-ray crystallography, researchers embed the protein in a native-like environment with a membrane, water, and ions, as depicted in Fig. 1.5. They then generate short movies (MD trajectories) based on initial atomic positions. These movies enable a subsequent analysis of protein dynamics, and thus provide a deeper understanding of ion channel functions.

1.5 Thesis Aims and Outline

In this thesis, MD simulations, supported by computational quantum mechanical calculations and machine learning (ML) techniques were leveraged to address several questions in the field of ion channel research. Specifically, the following questions are discussed:

- When is a hydrophobic gate not a hydrophobic gate?
- What are the best practices for functionally annotating ion channel structures?
- How does the shape of the pore influence conductance and permeation?
- How can we predict channel conductance?
- What are the limitations of non-polarisable force fields?
- How does anion binding occur in non-polar synthetic hosts?
- How can we predict channel conductance of GlyR with computational models?

Each of the following chapters contains their own detailed introduction. Chapter 2 provides the theoretical background and Chapter 3 the simulation protocols that underpin the research presented in this thesis.

In Chapter 4, the concept of hydrophobic gates is reviewed and an overview of functional annotation techniques for ion channel structures is presented. The first part of Chapter 4 is based on the viewpoint *When is a hydrophobic gate not a hydrophobic gate?* from Seiferth *et al.* [19]. The second part of Chapter 4 exemplifies best practices for functional annotating newly resolved GlyR and GLIC structures based on articles from Gibbs *et al.* [11] and Bharambe *et al.* [33] respectively.

Chapter 5 introduces an easy-to-use software tool for analysing ion channel pores using ellipsoidal probe particles, allowing the user to investigate the impact of pore shape and asymmetry on conductance and permeation. Additionally, a heuristic predictive model for ion channel conductance is presented. Chapter 5 is based on an article by Seiferth and Biggin [69] in which the PoreAnalyser tool is presented and ion permeation through carbon nanotube model systems is studied.

As described earlier, ion transport is usually highly selective for specific ions and is essential to maintain ion concentration gradients across biological membranes. The interaction of both biological and synthetic carriers with ions has been investigated for a long time, especially now that so many structures are available

(Fig. 1.3). The majority of studies, however, have focused on interactions with cations [12]. The interaction of ion carriers, synthetic or otherwise, with anions is more challenging because it must solve the problem of the relatively high solvation energy of anions compared to similar-sized cations [70]. Furthermore, the presence of water can negate the stabilizing intermolecular interactions that enable anion binding [26]. Chapter 6 and Chapter 7 focus on the limitations of non-polarisable force fields in describing anions in narrow and non-polar environments. In Chapter 6, anion binding in non-polar synthetic hosts is explored, while Chapter 7 examines chloride permeation through conductive GlyR structures, providing insights into the energetic barriers and structural features that influence ion permeation. Chapter 6 is based on an article by Seiferth *et al.* [71] in which different approaches to simulating anions in non-polar cavities are compared.

This thesis aims to enhance our understanding of ion channel function and presents best practices in the field of functional annotation.

Chapter 2
Theoretical Background and Computational Methods

When studying molecules through computer simulations, two primary approaches are utilised: quantum mechanics (QM) and classical mechanics. Quantum mechanics is essential for accurately describing the behaviour of electrons and nuclei. It provides detailed insights into electronic structures, chemical bonding, and molecular interactions. However, quantum mechanical calculations can become computationally prohibitive for larger systems. In such cases, classical mechanics, specifically molecular mechanics, is employed. This approach treats atoms as point masses and bonds as springs, simplifying the calculations by using empirical parameters to model molecular behaviour. While classical mechanics lacks the detailed electronic information provided by quantum mechanics, it allows for the simulation of much larger molecular systems and longer timescales, making it useful for studying structures, dynamics, and interactions in biochemical systems that are governed by time scales depicted in Fig. 1.6. In this chapter, an overview of the theoretical frameworks is given that underpin computational methods used in this thesis to study ion channels.

2.1 Theory

The choice of the appropriate theoretical framework is important because each physical theory excels at describing specific phenomena depending on the involved time and length scales. Classical mechanics is highly effective for studying macroscopic objects that are neither extremely massive nor moving at velocities close to the speed of light. In this regime, Newton's laws of motion provide accurate predictions of an object's behaviour. However, when one delves into the realm of atomic or subatomic particles, classical mechanics falls short. At scales comparable to an atom's diameter, typically around a few angstroms, quantum mechanics becomes relevant. Quantum mechanics accounts for the wave-like properties of particles and their probabilistic nature, offering a more accurate description of phenomena such as induced polarisation at these small scales. Quantum mechanical principles are essential for understanding the behaviour of electrons within atoms and the interactions that form chemical bonds. In this thesis, the primary focus is on describing ion channels using Newton's equations of motion, which are grounded in classical mechanics. However, one must be aware of the limitations inherent in this approach. Classical physics cannot capture the quantum mechanical interactions between atoms within ion channels such as induced polarisation. Therefore, to achieve a more detailed and accurate description of molecules on an atomistic level, quantum mechanical methods are employed as a bench mark and to parameterise classical force fields.

2.1.1 Classical Mechanics

When describing ion channels on an atomistic level, one can interpret atoms as point particles with a certain mass. This allows us to apply Newton's second law to describe their equations of motion. Newton's second law, formulated in 1687, states that the force $\vec{F}$ acting on an object is equal to the mass m of the object

multiplied by its acceleration $\vec{a}$:

$$\vec{F} = m\vec{a} = m\frac{d^2\vec{r}}{dt^2} = \frac{d\vec{p}}{dt}, \tag{2.1}$$

where $\vec{p}$ denotes the momentum and the acceleration $\vec{a}$ is the second derivative of the object's position $\vec{r}$. In a system containing N particles, each particle is characterised by its position and velocity. To fully describe the state of such a system, one must account for $3N$ position coordinates and $3N$ momentum components, collectively creating a $6N$-dimensional phase space. Each unique combination of these positions and momenta defines a distinct point within this phase space, and the evolution of the system can be visualized as a trajectory through this multidimensional space. Classical mechanics is often formulated in terms of the Hamiltonian H, for a conservative system representing the total energy of the system as the sum of kinetic and potential energies. The kinetic energy is a function of the velocities (or momenta) of the particles, while the potential energy depends on their positions. In the context of molecular dynamics, force fields can be interpreted as the Hamiltonian of the system. Force fields provide the mathematical description of the potential energy as a function of atomic positions. They include various terms representing bonded interactions (such as bonds, angles, and dihedrals) and non-bonded interactions (such as van der Waals forces and electrostatic interactions). By integrating Newton's equations of motion or using Hamilton's equations, one can simulate the time evolution of a system of particles. Hamilton's equations are given by:

$$\begin{aligned} \frac{d\vec{r}_i}{dt} &= \frac{\partial H}{\partial \vec{p}_i} \\ \frac{d\vec{p}_i}{dt} &= -\frac{\partial H}{\partial \vec{r}_i} \end{aligned} \tag{2.2}$$

These equations describe how the positions $\vec{r}_i$ and momenta $\vec{p}_i$ of particles change over time, allowing us to predict the future behaviour of the system based on its current state. By employing classical mechanics, one can model the dynamics of ion channels at an atomistic level. However, it is important to note that while classical mechanics offers a powerful framework, certain phenomena such as the effect of induced polarisation require additional quantum mechanical methods to achieve a more accurate description.

2.1.2 Thermodynamic Ensembles

In the study of molecular systems, macroscopic quantities such as system temperature are connected to the microscopic behaviour of particles. The temperature is related to the average kinetic energy of the particles in a system [72, 73]. This connection is realised through a large number of microstates, which are specific configurations of particles represented as points in phase space. An ensemble of microstates is defined as the collection of all possible microstates that correspond to a common value of a macroscopic variable. These ensembles provide a statistical framework for relating microscopic properties to macroscopic observables. Several types of ensembles are used in statistical mechanics: The microcanonical ensemble (NVE) represents a system with a fixed number of particles N, volume V, and energy E. It is suitable for isolated systems where energy is conserved, and no heat or particles are exchanged with the surroundings. In the canonical ensemble (NVT), the number of particles N, the volume V, and temperature T are fixed. The system can exchange energy with a heat bath, making it useful for studying systems in thermal equilibrium at a constant temperature. The isothermal-isobaric ensemble (NPT) allows for a constant number of particles N, pressure P, and temperature T. It is particularly useful for simulating systems under conditions where both temperature and pressure are controlled, such as biological membranes and macromolecules [72]. In the grand canonical ensemble, the system is characterised by a fixed temperature T, volume V, and chemical potential μ. Unlike the previous ensembles, the number of particles is not fixed, allowing for the

exchange of particles with a reservoir, making it ideal for studying open systems where the particle number can fluctuate. In this thesis, I mainly use the isothermal-isobaric ensemble.

A fundamental concept in statistical mechanics is the ergodic hypothesis, which posits that the ensemble average is equal to the time average for a system in equilibrium. This hypothesis implies that the behaviour of a single system over a long period can be represented by a large number of identical systems considered simultaneously. Thus, the time average of a single system's properties can be replaced by the ensemble average of many systems. If it were possible to visit all points in phase space, one could calculate the partition function by summing the values of $\exp(-E/k_BT)$, where E is the energy of a microstate, k_B is the Boltzmann constant, and T is the temperature. A phase space trajectory that visits all these points is called ergodic, meaning that the results would be independent of the initial configurations of the system [74]. However, for complex systems like proteins in membranes, the phase space is extraordinarily large, and achieving an ergodic trajectory is not feasible with current computational capabilities. As a result, molecular dynamics simulations can only provide an estimate of true energies and other thermodynamic properties.

2.1.3 Quantum Mechanics

The time-dependent Schrödinger equation describes how the quantum state of a physical system changes over time. It is given by:

$$\hat{H}|\Psi(\vec{r},t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(\vec{r},t)\rangle \tag{2.3}$$

where $\hat{H}$ is the Hamiltonian operator, which represents the total energy of the system, and $|\Psi(\vec{r},t)\rangle$ is the wavefunction, written in Dirac bra-ket notation [75]. The Hamilton operator can be expressed as

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V, \tag{2.4}$$

where ∇^2 is the Laplacian operator, representing the kinetic energy component, and V is the potential energy function [75]. This can be seen as analogous to the classical mechanics Hamiltonian, which is also the sum of kinetic and potential energy. When the external potential V in the Hamiltonian operator is independent of time, the wavefunction $|\Psi(\vec{r},t)\rangle$ can be separated into a product of a spatial part $|\Psi(\vec{r})\rangle$ and a time part $T(t)$:

$$|\Psi(\vec{r},t)\rangle = |\psi(\vec{r})\rangle T(t) \tag{2.5}$$

In this scenario, the Schrödinger equation can be simplified to a time-independent form:

$$\hat{H}|\Psi(\vec{r})\rangle = i\hbar E|\Psi(\vec{r})\rangle, \tag{2.6}$$

where E is the energy eigenvalue. This is known as the time-independent Schrödinger equation, an eigenvalue problem where the Hamiltonian operator $\hat{H}$ acts on the spatial part of the wavefunction [72].

Born-Oppenheimer Approximation

The Schrödinger equation cannot be solved exactly for any molecular system due to the complexity of interactions between particles. However, the Born-Oppenheimer approximation allows for the separation of electronic and nuclear motion by exploiting the mass difference between electrons and nuclei. Since nuclei are much heavier, they move much more slowly compared to electrons. Thus, electrons can be assumed to adjust instantaneously to changes in nuclear positions. Under this approximation, the total

wavefunction $\Psi(\text{nuclei, electrons})$ can be written as a product of nuclear and electronic wavefunctions [75]:

$$\Psi(\text{nuclei, electrons}) = \Psi(\text{nuclei})\Psi(\text{electrons}) \tag{2.7}$$

The total energy of the system is the sum of the nuclear and electronic energies. Solving the Schrödinger equation for atoms with more than one electron is more complex due to electron-electron interactions. For instance, the Helium atom, with two electrons and one nucleus, represents a three-body problem for which no exact solution exists. Thus, one must rely on approximation methods to find solutions. Electrons are indistinguishable particles, and the exchange of any pair of electrons must leave the electron density unchanged. According to the Born interpretation, electron density is equal to the square of the wavefunction. Therefore, the wavefunction must either remain unchanged or change sign upon the exchange of two electrons. The antisymmetry principle states that for fermions (such as electrons), the wavefunction changes sign when two electrons are exchanged [72]. The details of how to approximate solutions for systems with more than one electron involve advanced techniques such as Hartree-Fock, and density functional theory (DFT). A detailed description of those methods is beyond the scope of this discussion but a short overview is given to give some context of *ab initio* calculations later used to parameterise molecular dynamics force fields.

The Hartree-Fock (HF) method is an approach in quantum chemistry used to approximate the wavefunction and energy of a quantum many-body system in a stationary state. The method relies on the variation theorem, which asserts that the energy calculated from any approximate wavefunction is always an upper bound to the true ground state energy [72]. The closer the approximate wavefunction is to the true wavefunction, the lower the calculated energy will be. To describe the wavefunction of a multi-electron system, the Hartree-Fock method employs a Slater determinant, which ensures that the wavefunction satisfies the Pauli exclusion principle by being antisymmetric with respect to the exchange of any two electrons [72]. The Schrödinger equation and its approximations are computationally expensive and hence have limitations related to the time scales and system sizes they can effectively describe.

Basis Sets

A basis set in quantum chemistry is a set of functions used to describe the wavefunctions of electrons in a molecule. While Slater-Type Orbitals (STOs) closely resemble the actual shape of atomic orbitals, they are difficult to evaluate, especially when atomic orbitals are centred on different nuclei [72]. Gaussian-Type Orbitals (GTOs) are often used instead of STOs due to their computational advantages: The product of two Gaussians is another Gaussian, which simplifies the calculations. However, GTOs do not have a cusp at the origin and decay to zero more quickly than STOs. Therefore, instead of a single Gaussian, a linear combination of multiple Gaussians is used to approximate STOs more accurately. A minimal Basis Set contains the minimum number of functions required to represent all occupied orbitals of an atom. This set is limited in describing the electron density's non-spherical aspects. A double zeta basis set doubles the number of functions in the minimal basis set, allowing for a more accurate representation of the electron density. Instead of doubling the number of all functions, split-valence basis sets use more functions only for valence electrons, which are crucial for chemical bonding, and not for core electrons, which do not strongly affect chemical properties. An example is the 6-31G basis set. Polarisation functions are added to basis sets to allow for the perturbation of the electron cloud due to interactions with other atoms. These are indicated by an asterisk (*). Diffuse functions are used to describe anions and molecules with lone pairs where electron density extends far from the nuclei. These are indicated by a plus sign (+) [72].

Density Functional Theory

Hohenberg and Kohn established that the ground-state energy of a many-electron system is uniquely determined by its electron density [76]. Therefore, the energy is a unique functional of the electron density. In DFT, the key quantities are functionals, which are functions of functions. Specifically, DFT functionals map the electron density $\rho(\vec{r})$ to the total energy of the system [72]. DFT is generally more computationally efficient than methods that explicitly calculate the many-body wavefunction, such as Hartree-Fock and its post-Hartree-Fock descendants. A popular hybrid functional in DFT is B3LYP, combining the Becke three-parameter exchange functional with the Lee-Yang-Parr correlation functional [77, 78].

Intermolecular Interaction Energy - Symmetry-Adapted Perturbation Theory

The energy of interaction E_{int} between molecules A and B can be defined as the difference between the energy of dimer E_{AB} and the energies of the monomers E_A and E_B [79]:

$$E_{int} = E_{AB} - E_A - E_B \tag{2.8}$$

When calculating the monomer energies E_A and E_B, the internal coordinates used are assumed to be the same as those in the dimer. Consequently, the interaction energy depends on the distance between the centers of mass of the monomers and their mutual orientation [79]. At large intermolecular separations, both overlap and exchange effects diminish exponentially.

A simple perturbation approach to intermolecular interaction is the Rayleigh-Schrödinger perturbation method, applied to the eigenvalue problem for the electronic (Born-Oppenheimer) Hamiltonian of the dimer AB [79]:

$$(H_0 + \zeta V)\Phi_{AB} = E_{AB}\Phi_{AB}, \tag{2.9}$$

where $\zeta \in [0, 1]$ and $H_0 = H_A + H_B$ denotes the unperturbed operator, chosen as the sum of the monomer Hamiltonians H_A and H_B. The perturbation operator V represents the difference between the total Hamiltonian H and H_0. When $\zeta = 0$, the intermolecular interaction terms in the Hamiltonian are switched off and the dimer energy is equal to the sum of the energies of the unperturbed monomers $E_{AB}(0) = E_A + E_B$ and the dimer wave functions becomes the product of the monomer wave functions $\Phi_{AB}(0) = \Phi_A\Phi_B$. When $\zeta = 1$, the intermolecular interaction terms in the Hamiltonian given by V are completely included and the corresponding wave function $\Phi_{AB}(1)$ becomes the exact physical wave function and the eigenvalue $E_{AB}(1)$ becomes the exact energy of the dimer. One can expand $\Phi_{AB}(\zeta)$ and $E_{AB}(\zeta)$ as a power series in ζ. For a detailed perturbation theory, see Jeziorski *et al.* [79]. Symmetry-adapted perturbation theory (SAPT) directly computes the non-covalent interaction energy between two molecules, without requiring the computation of the total energies of the monomers or the dimer. Interaction energy corrections computed via the SAPT approach can be classified into terms with well-defined physical interpretations: electrostatic, induction, dispersion, and exchange [79, 80]. In SAPT, the Hamiltonian of the dimer is partitioned into contributions from each monomer described by the following equation:

$$H = K_A + W_A + K_B + W_B + V \tag{2.10}$$

In this equation, the Hamiltonian is expressed as the sum of the monomer Kohn-Sham operators $K_A + K_B$, the fluctuation potentials of each monomer, W_A and W_B, and the interaction potential V. The monomer Kohn-Sham operators K_i are treated as the zeroth-order Hamiltonian, and the interaction energy is evaluated through a perturbative expansion of V, W_A, and W_B. The first-order terms in V

include electrostatic and exchange interactions, while induction and dispersion interactions first appear at second-order in V[79, 80]. The simplest truncation, known as SAPT0, is defined by the following equation:

$$E_{SAPT0} = E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp,resp}^{(20)} + E_{exch-disp,resp}^{(20)}(est.) + \delta_{HF}^{(2)} \tag{2.11}$$

In this notation, $E^{(vw)}$ indicates the order in V and in $W_A + W_B$; the subscript $resp$ denotes that orbital relaxation effects are included [80]. The $\delta_{HF}^{(2)}$ term takes higher-order induction effects into account [80]. SAPT analysis can be used to study the balance of electrostatic and dispersion contributions. For instance, Kumar *et al*. used SAPT to study cation–π interactions of neutral aromatic ligands with the cationic amino acid residues arginine, histidine and lysine [81].

2.2 Molecular Dynamic Simulations

The previously described quantum mechanical frameworks are powerful to describe the behaviour of electrons within atoms and molecules. However, many biochemical systems of interests involve such a large number of particles that employing QM becomes impractical. The computational resources required to solve QM equations for systems with numerous particles grow enormously with the number of electrons, making it extremely time-consuming and resource-intensive [72]. Even in semi-empirical schemes, where some electrons might be ignored or approximations are made, the sheer number of particles still poses a major computational challenge. For instance, calculating the electronic structure of large biomolecules such as ion channels can be prohibitively expensive using purely quantum mechanical approaches. To address the limitations of QM methods, force field methods, also known as molecular mechanics (MM), are employed. These methods simplify the problem by ignoring the motions of electrons and instead calculate the energy of the system as a function of the positions of the nuclei alone. This approach reduces the computational load, enabling the study of much larger systems over longer timescales. Molecular Dynamics (MD) is a specific application of force field methods where the positions and velocities of atoms are tracked over time, using classical equations of Newtonian mechanics to describe their movements. While MD is much faster than QM methods, it falls short in accurately describing electronic distributions and electronic effects within molecules. This limitation arises because MD relies on several key assumptions and approximations. The Born-Oppenheimer Approximation was already described in section 2.1.3 and allows to separate nuclear and electronic motion, treating nuclei as stationary relative to the much faster-moving electrons. Nuclear positions can be calculated without explicitly solving for electronic motions. The electronic degrees of freedom are usually described with partial charges [72]. Force fields typically model the stretching of bonds, bending of angles, and rotation about single bonds using spring-like potentials. These potentials approximate the energetic penalties associated with deviations from equilibrium bond lengths and angles. One of the most important attributes of a force field used to model biochemical molecules is its transferability. This means that a set of parameters developed and tested on a relatively small number of cases can be applied to a much wider range of problems. For example, the same parameters used to model amino acids can be used to model different proteins. This transferability is particularly beneficial because it eliminates the need to define a new set of parameters for each individual molecule, saving both time and computational resources [72].

2.2.1 Force Fields

Force fields represent the interactions between atoms. More specifically, a force field is an interatomic empirical potential and different functional forms and parameter sets are employed to compute the

potential energy of a system at the atomic level [72]. The atom types within a force field are closely linked to the chemical environment of the atoms, including the atomic numbers of the nuclei, the system's geometry, overall charge, and details such as the atom's hybridization state. The simulation relies on a model Hamiltonian to compute energies and forces, ensuring that all configurations are sampled with the correct relative probability [74]. Force fields are empirical, meaning they are derived from experimental data and theoretical calculations. The functional forms used in force fields represent a compromise between accuracy and computational efficiency. For tasks such as energy minimization and time integration, it is desirable to calculate the first and second derivatives of the energy with respect to atomic coordinates. The forces on all atoms can be calculated by taking the gradient of the potential U:

$$\vec{F} = -\nabla U \tag{2.12}$$

Force fields describe both inter- and intramolecular forces and the potential U can be expressed as

$$U = U_{\text{bonded}} + U_{\text{nonbonded}}. \tag{2.13}$$

Intramolecular forces include energetic penalties for deviations of bonds and angles from their equilibrium values, modeled using spring-like interactions. Intermolecular forces, on the other hand, encompass interactions between non-bonded parts of the system. These interactions are often described using van der Waals forces, commonly represented by the Lennard-Jones potential, and electrostatic interactions computed based on Coulomb's law [72]. In this thesis, two different types of classical Hamiltonians, or force fields, are utilised. On the one hand, pairwise additive force fields, such as Assisted Model Building with Energy Refinement (AMBER) [82] and Chemistry at Harvard Macromolecular Mechanics (CHARMM), and on the other hand polarisable force fields like atomic multipole optimized energetics for biomolecular applications (AMOEBA) [83] and CHARMM Drude [84]. These force fields mainly differ in their treatment of non-bonded interactions, particularly van der Waals and electrostatic interactions. The pairwise additive force fields use fixed partial charges to model electrostatics, whereas polarisable force fields incorporate the dynamic response of the electron cloud to the surrounding environment, offering a more nuanced representation of molecular interactions. The energy due to interaction of two non-bonded atoms is independent of the position of all other atoms in non-polarisable or pairwise-additive force fields. In polarisable force fields nonbonded interactions facilitate many-body interactions. While polarisable force fields address certain limitations of classical pairwise additive force fields, they still fall short in physical accuracy and are for instance unable to describe chemical reactivity. Hybrid QM/MM approaches have gained popularity as computational tools for studying chemical reactivity within biomolecular systems. However, the high computational cost associated with performing high-level QM calculations has constrained the use of these methods in calculating free energies for chemical processes. To overcome these challenges, reference potentials and mean field approximations have been developed to accelerate high-level QM/MM calculations [85]. In this section, I first give a general overview over the bonded and nonbonded interactions in pairwise additive force fields and then focus on polarisable force fields.

Bonded Interactions

The first term in the potential energy function in eq. (2.13) describes the energy contributions that are associated with interactions between covalently linked atoms. The bonded potential U_{bonded}

$$U_{\text{bonded}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{prop. torsion}} + U_{\text{improp. torsion}} \tag{2.14}$$

can be decomposed into contributions from bond stretching, angle bending, torsional rotations and improper dihedral angles. Fig. 2.1 illustrates the different contributions to U_{bonded}. The bond stretching term U_{bond} is

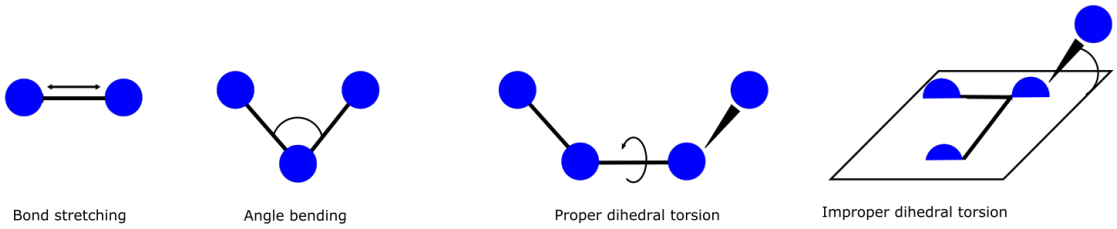


Figure 2.1: Schematic illustration of bonded interactions terms in molecular force fields.

usually modelled using a harmonic term

$$U_{\text{bond}} = \frac{k_b}{2}(x - x_0)^2 \tag{2.15}$$

where k_b is a spring constant associated with the bond between two atoms and x_0 is reference bond length x_0. The latter is a reference in the sense that the bond has length x_0 when all other terms in force field are set to zero. Therefore, x_0 does not necessarily equals the equilibrium bond length [72]. The bending of angles involving three atoms is usually described with a harmonic potential too:

$$U_{\text{angle}} = \frac{k}{2}(\theta - \theta_0)^2. \tag{2.16}$$

The bond stretching term is a 2-body and the angle bending term is a 3-body interaction. The *pairwise* in pairwise-additive force fields only refers to nonbonded interactions. Bonded interactions in non-polarisable force fields typically include even 4-body interactions such as dihedral angle interactions: A dihedral angle describes the angle between two planes formed by four atoms forming a linear bonded chain as illustrated in Fig. 2.1. The potential surface describing the rotation around the central bond can be complex and a common way to model it is a cosine expansion:

$$U_{\text{prop. torsion}} = \sum_n \frac{k_n}{2}(1 + \cos(n\omega - \gamma_n)), \tag{2.17}$$

where the relative contributions k_n and phases γ_n are usually fitted to optimally match a quantum-mechanical torsion scan. An improper dihedral is a bending mode where three atoms, each connected to a central atom, move along the normal vector of their arrangement. It can be used to enforce planar geometries, such as in sp² hybridized carbon systems and be modelled using the following functional form:

$$U_{\text{improp. torsion}} = \frac{1}{2}k_\psi(\psi - \psi_0)^2, \tag{2.18}$$

where changes to the flat geometry of, for instance an aromatic ring, or changes to tetrahedral geometries from their reference plane ψ_0 are associated with an energy penalty. Some force fields include further bonded coupling interactions, cross terms such as the Urey-Bradley term. The Urey-Bradley term describes changes in energy with angle bending as a functions of the distance between two non-covalently bonded atoms that are separated by two bonds forming the angle. These terms are typically included to improve agreement with experimental vibrational spectra.

Non-bonded Interactions

Non-bonded interactions refer to forces that occur between atoms that are not directly connected by chemical bonds. Electrostatic interactions arise from the forces between charged particles. In molecular systems, these interactions occur due to the presence of partial charges on atoms, which result from

differences in electronegativity. Electronegative elements, such as oxygen or nitrogen, attract electrons more strongly than less electronegative elements, like hydrogen or carbon. This unequal attraction leads to an uneven distribution of charge within a molecule, creating regions of partial positive and negative charges. To model these interactions, the charge distribution in a molecule is often represented by an arrangement of fractional point charges assigned to atoms. These partial atomic charges are designed to reproduce the electrostatic properties of the molecule. In pairwise additive force fields, the electrostatic contribution to the total energy of the system is calculated as the sum of interactions between all pairs of point charges, using Coulomb's law:

$$U_{\text{Coulomb}} = \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \tag{2.19}$$

where ϵ_0 is the dielectric permittivity, q_i the charge of particle i and r_{ij} is the distance between particles i and j.

Electric multipole moments are mathematical representations that describe the distribution of electric charge within a molecule. These moments provide insights into how molecules interact with external electric fields and with each other. The multipole expansion starts with the simplest moment, the monopole (total net charge), and proceeds to more complex distributions such as dipoles and quadrupoles. The simplest electric moment, apart from total net charge on molecule, is the dipole $\vec{\mu}$. The dipole moment of a distribution of charges q_i located at positions $\vec{r}_i$ is given by

$$\vec{\mu} = \sum q_i \vec{r}_i. \tag{2.20}$$

The dipole is a vector quantity. The quadrupole moment θ is the next level of complexity in the multipole expansion and represents the second-order distribution of charge. A quadrupole arises from an arrangement of four charges that sum to zero and are positioned in such a way that they do not produce a net dipole moment. The quadrupole moment is a tensor quantity with nine components, representing all pairwise combinations of the Cartesian coordinates. Mathematically, the quadrupole moment tensor θ can be represented as a 3x3 matrix:

$$\theta = \begin{pmatrix} \sum q_i x_i^2 & \sum q_i x_i y_i & \sum q_i x_i z_i \\ \sum q_i y_i x_i & \sum q_i y_i^2 & \sum q_i y_i z_i \\ \sum q_i z_i x_i & \sum q_i z_i y_i & \sum q_i z_i^2 \end{pmatrix} \tag{2.21}$$

Each element θ_{ij} of the matrix corresponds to a specific combination of the coordinates, capturing the extent to which the charge distribution deviates from spherical symmetry. Higher moments such as octupole and hexadecapole are often truncated in the series expansion [72]. The AMOEBA force field uses a multipole expansion for each atom [83].

Various methods have been developed to calculate partial charges. One such method is population analysis, which involves partitioning the electron density among the atoms in a molecule. During a Mulliken population analysis, partial charges are derived from a partitioning scheme of the molecular wavefunction, specifically from the molecular orbitals [86]. The Mulliken charges depend primarily on the internal bonding structure. This means that they reflect the electron density distribution based on atomic contributions to molecular orbitals rather than being explicitly designed to reproduce the properties that determine molecular interactions [72]. Cox and Williams introduced a method where the electrostatic potential at various points around the molecule is calculated from the molecular wavefunction [87]. A least-squares fitting procedure is then employed to determine a set of partial atomic charges that best reproduce the electrostatic potential at these points. The Restrained Electrostatic Potential Fit (RESP) method, developed by Bayly *et al.*, builds on the electrostatic potential fitting approach [88]. RESP uses hyperbolic restraints on

non-hydrogen atoms to ensure that the fitted charges are chemically reasonable and consistent. One of the key advantages of RESP charges is their relative insensitivity to molecular conformation, making them more robust for use in molecular dynamics simulations and other computational studies [72, 88]. For instance, the AMBER force field employs RESP charges [82]. To capture the anisotropy of the molecular charge distribution, more sophisticated models than simple point charges are sometimes used. These models distribute multipoles throughout the molecule. The Distributed Multipole Analysis (DMA) method developed by Stone provides a detailed way to calculate multipoles from a quantum mechanical wavefunction defined in terms of Gaussian basis functions [89]. As described in section 2.1.3, the overlap between two Gaussian functions can be represented by another Gaussian at a point along the line connecting them. Each product of basis functions corresponds to a charge density at a point, and this density can be expressed as a multipole expansion around that point. The highest multipole moment in a local expansion depends on the basis set used. For instance, using a basis set containing only s and p functions limits the local multipoles to no higher than quadrupoles. This ensures that the multipole expansion is appropriate for the level of detail in the basis set [72].

Electrostatic interactions do not account for all non-bonded interactions within a system. This is particularly evident when considering noble gas atoms, such as helium, neon, and argon, which lack permanent multipole moments yet still exhibit interactions that influence their behaviour. The deviations from ideal gas behaviour observed in noble gas atoms were quantified by VDW interactions. Named after Johannes Diderik van der Waals, these interactions describe how the actual behaviour of gases deviates from that predicted by the ideal gas law due to intermolecular forces. VDW interactions describe the interaction energy between atoms or molecules as a function of their separation. This energy varies with distance: At infinite distance, the interaction energy is zero and interactions become negligible at relatively short distances. However, at very close distances, a strong repulsive force dominates, preventing atoms from coming too close to each other. Dispersive interactions, also known as London dispersion forces, are a type of van der Waals force that contribute to the attractive interaction between atoms or molecules. These forces were first described by Fritz London in 1930 [90]. Dispersive forces arise due to the fluctuations in electron clouds within atoms or molecules. These fluctuations can create instantaneous dipoles, which are temporary and result from the momentary uneven distribution of electrons. An instantaneous dipole in one atom can induce a dipole in a neighbouring atom, leading to an attractive inductive effect between the two atoms. In addition to attractive dispersive forces, there are repulsive forces that act at very short distances. These repulsive interactions have a quantum mechanical origin and can be understood through the Pauli exclusion principle: No two electrons in a system can have the same set of quantum numbers. When atoms come very close to each other, their electron clouds overlap, and the Pauli exclusion principle leads to a repulsion between the overlapping electron clouds, preventing the atoms from collapsing into each other. Within molecular mechanics force-fields, the 12-6 Lennard-Jones potential is the most commonly used potential to model van der Waals interactions. The repulsive term is proportional to r^{-12} and the attractive term to r^{-6}:

$$U_{VDW} = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{2.22}$$

where σ refers to the zero of the potential and ϵ to the well depth of its minimum. Each atom i is characterised by a parameter pair σ_i, ϵ_i. The pairwise interaction between atom i and j can be modeled with Lorentz-Berthelot combination rules for radius and energy:

$$\begin{aligned} \sigma_{ij} &= \frac{\sigma_i + \sigma_j}{2} \\ \epsilon_{ij} &= \sqrt{\epsilon_i \epsilon_j} \end{aligned} \tag{2.23}$$

The 12-6 Lennard Jones potential is plotted in Fig. 2.3 for halide anions.

Polarisable Force Fields

The previously described electronic interactions are limited to the permanent features of charge distribution. However, electrostatic interactions also arise from changes in a molecule's or atom's charge distribution in response to an external field. For example, neighbouring molecules can induce a dipole within a molecule. In biomolecular MD simulations, the previously described pairwise additive force fields are the most commonly used, often relying on assigning fixed partial charges before the simulation begins. But, this approach, which ignores polarisation effects, limits physical accuracy [71, 91]. The charge density surrounding an atom is constantly changing in response to its environment, suggesting that partial charges would be better represented by dynamic and responsive models [92, 93]. Various frameworks have been developed to model polarisation in classical force fields, including transferable or fluctuating charges [94, 95], Drude oscillators [96], and polarisable atomic multipole models [83]. Fluctuating charge models allow partial charges on atoms to vary over time, while Drude oscillator force fields introduce massless Drude particles to model electronic degrees of freedom by displacing charge relative to the nucleus. Polarisable atomic multipole models, such as AMOEBA, incorporate inducible point dipoles and higher-order static multipoles. In this thesis, the atomic multipole framework is applied to describe ion binding to narrow and hydrophobic cavities of ion carriers in Chapter 6, while the Drude oscillator framework is used to explore the GlyR ion channel in Chapter 7.

Classical Drude Oscillator

In the classical Drude oscillator approach, an atom is described by a massive nucleus and a massless Drude oscillator which represents the electronic degrees of freedom. The net charge of an atom is divided between the nucleus and the Drude oscillator. A Drude particle is attached to the water oxygen in the SWM4-DP water model [96] that is depicted in Fig. 2.2(a). When the Drude oscillator is displaced $\vec{d}$ relative to the nucleus, it generates an atomic dipole moment. The interaction between the Drude oscillator and the nucleus is modeled as a harmonic spring, and a self-polarisation term is added to the potential energy:

$$U_{\text{self}} = \frac{1}{2} k_d \vec{d}^2, \tag{2.24}$$

where k_d is the force constant that, together with the charge of the Drude oscillator q_D, determines the polarisability α of the atom:

$$\alpha = \frac{q_D^2}{k_D}. \tag{2.25}$$

Drude particles are not influenced by van der Waals interactions and do not participate in any bonded interactions; they only interact through Coulomb forces. The CHARMM Drude model employs lone pair virtual sites on certain heavy atoms to more accurately represent the anisotropic distribution of electronic charges [84]. A Thole screening factor [97] is applied to dipole-dipole interactions, which is based on the concept of smearing out discrete point charges into a continuous charge density. To determine the Drude displacements that minimise the potential energy for fixed nucleus positions, a self-consistent field approach can be used solving the following equation

$$\nabla_{\vec{d}_i} U = \nabla_{\vec{d}_i} (U_{\text{self}} + U_{\text{Coulomb}}) = 0, \tag{2.26}$$

where i refers to polarisable atoms, U_{self} is the self-polarisation term defined in eq. (2.24) and U_{Coulomb} is the Coulomb potential defined in eq. (2.19). This approach directly reflects the Born Oppenheimer

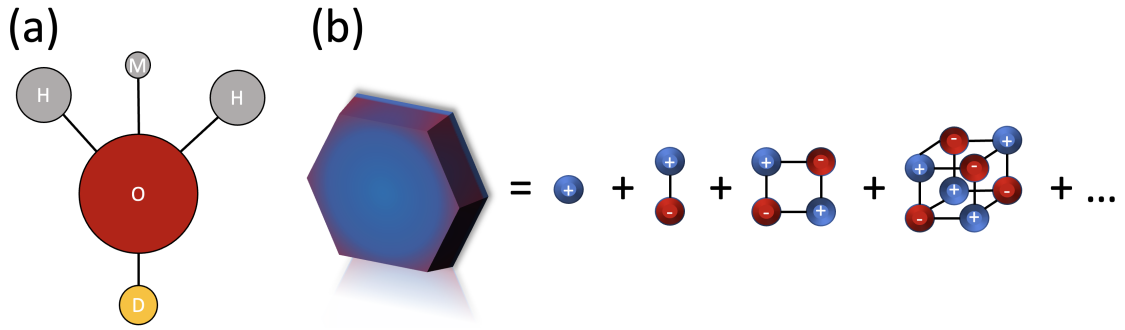


Figure 2.2: SWM4-DP water model for CHARMM Drude force field (a) and multipole expansion until octupole (b). The SWM4-DP water model shares the geometry of the nonpolarisable TIP4P water model with a single VDW interaction site on the red oxygen atom [96]. Charges are located on the hydrogens and a virtual "M"-site. A yellow Drude oscillator with opposite charge is attached to the oxygen atom via a harmonic spring.

approximation: The electronic degrees of freedom almost instantaneously relax and the potential solving eq. (2.26) can be used in equation of motion for the nuclei which are assumed to be fixed to calculate the self-consistent field.

AMOEBA

The AMOEBA force field uses different potentials for non-bonded interactions than pairwise additive force fields. Van der Waals interactions are described by a buffered 14-7 potential of the form

$$U_{\text{VDW}} = \epsilon_{ij} \left(\frac{1+\delta}{\rho_{ij}+\delta} \right)^7 \left(\frac{1+\gamma}{\rho_{ij}^7+\gamma} - 2 \right) \tag{2.27}$$

with $\delta = 0.07$ and $\gamma = 0.12$ [83]. The repulsion in the buffered 14-7 potential is softer in the Lennard Jones potential. Here $\rho_{ij} = r_{ij}/R_{ij}$ denotes the distance r_{ij} between two atoms i and j in units of the the distance R_{ij} for which the potential is minimal. The depth of the potential well ϵ_{ij} and R_{ij} for atoms of different types are defined as

$$\begin{aligned} R_{ij} &= \frac{R_i^3 + R_j^3}{R_i^2 + R_j^2} \\ \epsilon_{ij} &= \frac{4\epsilon_i \epsilon_j}{(\sqrt{\epsilon_i} + \sqrt{\epsilon_j})^2}. \end{aligned} \tag{2.28}$$

The buffered 14-7 potential is plotted in Fig. 2.3 for halide anions. In pairwise additive force fields, a Coulomb potential is used to describe electrostatic interactions, see eq. (2.19). A fixed partial charge model corresponds to a multipole expansion where dipole, quadrupole and higher-order terms are truncated. The AMOEBA force field takes multipoles up to quadrupoles at each atomic centre into account and truncates octupole and higher moments. A graphical representation of a multipole expansion is plotted in Fig. 2.2(b). The electrostatic energy in AMOEBA includes contributions from both permanent and induced multipoles. Only the dipole moment is treated as inducible. Monopole (i.e. charge) and quadrupole moments are invariable. Instead of the Coulomb potential for charges q_i, the electrostatic interaction U_{ele} in the AMOEBA force field is calculated as

$$U_{\text{ele}} = M_i^T T_{ij} M_j, \tag{2.29}$$

where each atom is associated with a multipole moment vector M with

$$M = \begin{pmatrix} q & \mu_x & \mu_y & \mu_z & \theta_{xx} & \theta_{xy} & \theta_{xz} & \dots & \theta_{zz} \end{pmatrix}^T. \tag{2.30}$$

Here, q, μ and θ are the monopole, dipole and quadruple moments respectively. Ren and Ponder follow Stone's notation [98] to express the multipole interaction T matrix elements as

$$\begin{aligned} T &= \frac{1}{R} \\ T_\alpha &= \nabla_\alpha T = -\frac{R_\alpha}{R^3} \\ T_{\alpha\beta} &= \nabla_\alpha T_\beta \\ T_{\alpha\beta\gamma} &= \nabla_\alpha T_{\beta\gamma}, \end{aligned} \tag{2.31}$$

where $4\pi\varepsilon_0$ is omitted [92]. The inclusion of explicit dipole polarisation allows the AMOEBA model to respond to changing or heterogeneous molecular environments [83]. An external field distorts the atomic electron density. A classical point dipole moment is induced at each polarisable atomic site according to the electric field felt by that site in AMOEBA. The induced dipole for each atom i is computed as

$$\mu_{i,\alpha}^{\text{ind}} = \alpha_i E_{i,\alpha} = \alpha_i \left(\sum_j T_\alpha^{ij} M_j^{\text{perm}} + \sum_{k \neq i} T_{\alpha\beta}^{ik} \mu_{k,\beta}^{\text{ind}} \right), \tag{2.32}$$

where α_i is the atomic polarisability and $E_{i,\alpha}$ is the sum of the fields generated by both permanent multipoles and induced dipoles [92]. The electronic degrees of freedom are again assumed to relax instantaneously. With the self-consistency condition, eq. (2.32) can be written as

$$\mu_{i,\alpha}^{\text{ind}}(n+1) = \mu_{i,\alpha}^{\text{ind}}(n) - \alpha_i \sum_j T_{\alpha\beta}^{ij} \mu_{j,\beta}^{\text{ind}}(n), \tag{2.33}$$

which can solved iteratively through successive overrelaxation [99]. In order to prevent the divergence of induced dipole moments at short interatomic distances, the so-called polarisation catastrophe, the AMOEBA force field incorporates Thole's damped interaction method [97]. Polarisation interaction at very short range is damped by smearing out the atomic multipole moments. The charge density in Thole's screening model is defined as

$$\begin{aligned} \rho &= \frac{3a}{4\pi} \exp(-a\tilde{r}_{ij}) \\ \tilde{r}_{ij} &= \frac{r_{ij}}{(\alpha_i \alpha_j)^{1/6}}, \end{aligned} \tag{2.34}$$

where $\tilde{r}_{ij}$ denotes the reduced interatomic distance depending on the atomic polarisabilities α_i. In the AMOEBA force field, all atoms have the same dimensionless width parameter $a = 0.39$ of the smeared charge distribution [83].

Electric Continuum Correction

The explicit treatment of electrostatics is computationally expensive, making it desirable to adjust existing parameters in pairwise additive force fields to enhance physical accuracy and offer a more efficient alternative to polarisable force fields. Electronic polarisation can be understood as the deformation of electronic shells, comprising two components: nuclear or inertial polarisation, which is dependent on the position of nuclei, and electronic or inertialess polarisation, which corresponds to the pure electronic response to an external field. Nuclear polarisation primarily influences low-frequency responses, while high-frequency responses are mainly due to electronic polarisation. In non-polarisable force fields, the pure electronic component of polarisation is not included in the electronic description, while the remaining nuclear component is effectively captured through the definition of atomic partial charges and other non-electrostatic parameters [100]. As a result, ions and other highly charged atoms in non-polarisable force fields tend to overbind to proteins. This overbinding occurs because the pure electronic polarisation, which leads to effects like

charge screening, is not modeled [101]. In electronic continuum correction (ECC), electronic polarisation is implicitly introduced into pairwise additive force fields by embedding the entire system within a dielectric continuum that has a dielectric constant corresponding to the electronic component of the polarisation response [102]. As a result, the charges are scaled [103]:

$$q_i^{\text{eff}} = \frac{q_i}{\sqrt{\epsilon_{el}}}, \tag{2.35}$$

where $\epsilon_{el} \approx 1.78$ for water such that the charge scaling factor $1/\sqrt{\epsilon_{el}} \approx 0.75$ [104]. ECC is restricted to electronically homogeneous systems and hence cannot be used for systems with a discontinuous dielectric constant such as the water air interface [104]. However, this limitation does not hold for protein-water interfaces. The ECC is not used in this thesis but it has been used in recent studies to investigate ion behaviour of ions in nanopore, ion pumps and ion channels [105, 106, 107].

Water Models

Modelling water accurately is a challenging task [96, 108, 109]. Many properties of water can be determined through computer simulations and compared with experimental data. Key characteristics such as radial pair distribution functions and the self-diffusion coefficient are used to assess the structure and dynamics of water at different temperatures [110]. The quality and transferability of a water model can be evaluated by its ability to replicate experimental properties of bulk water, including dielectric constant, surface tension, and bulk density [39]. Simple water models consist of three to five interaction sites with a rigid geometry, such as the Simple Point Charge (SPC) and the Transferable Intermolecular Potential (TIP) models. The TIP family includes models like TIP3P [108], a 3-point model with three interaction sites: a negative charge on oxygen and positive charges on each hydrogen atom. Other TIP models have additional interaction sites. In this thesis, I use the TIP3P model from the TIP family and the SPC/E (Extended Simple Point Charge) model from the SPC family [109]. The SPC/E water model also employs three sites for electrostatic interactions, with the partial positive charges on the hydrogen atoms balanced by a corresponding negative charge on the oxygen atom. The van der Waals interactions between two water molecules are computed using Lennard-Jones functions, with a single interaction point per molecule centered on the oxygen atom [110]. The CHARMM Drude polarisable SWM4-NDP water model features four regular interaction sites, similar to the TIP4P water model, with an additional Drude oscillator attached to the oxygen atom [96]. The SWM4-DP water model is depicted in Fig. 2.2(a). The AMOEBA03 water model is fully flexible, with its potential energy described by a harmonic Urey-Bradley term [111].

Ion Parameters

Many non-polarisable force fields are fitted to reproduce gas-phase conformational energies and geometries. Non-bonded interactions are particularly important in condensed phase simulations. Of the non-bonded interactions, electrostatics models are often polarised beyond what would be expected in the gas phase [112]. The transferability of parameters for pairwise additive force fields can be limited due to their fixed charges and their inability to model the effect of induced polarisation. Within molecular mechanics force fields, the 12-6 Lennard-Jones potential in eq. (2.22) is the most commonly used potential to model van der Waals interactions. There are different methods for improving the accuracy of anion MD properties by changing non-bonded interactions in a molecular mechanics force field. Li *et al*. re-optimized the Lennard-Jones parameters for van der Waals interactions to reproduce hydration free energies within the AMBER force field [113]. Parameters in the non-polarisable CHARMM force field systematically underestimate the interaction of chloride with proteins and lipids [114]. Hence, Orabi *et al*. formulated a series of pair-specific Lennard-

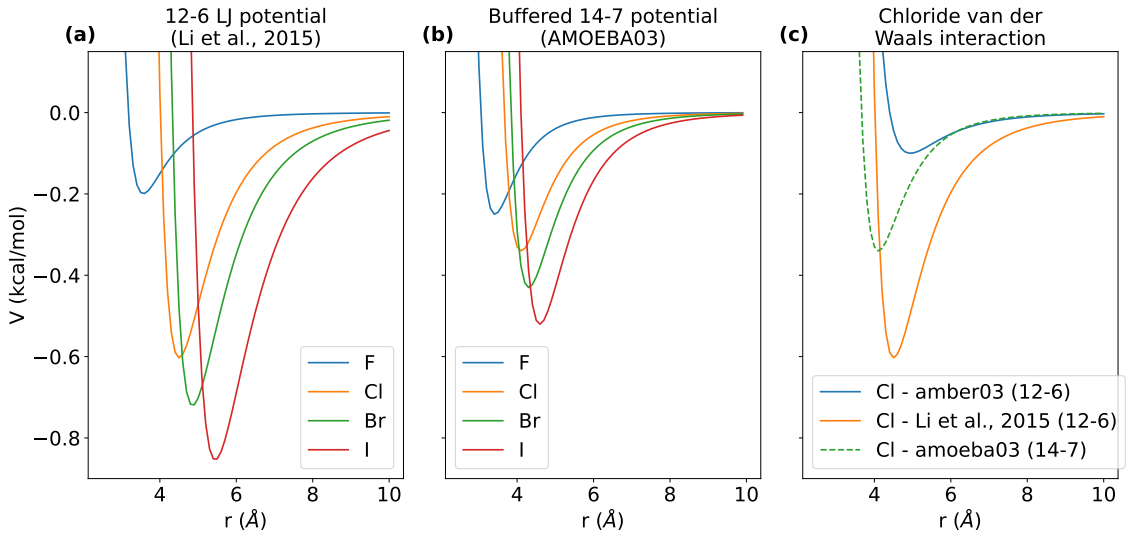


Figure 2.3: Van der Waals interaction modelled with a 12-6 Lennard-Jones potential (a) used for non-polarisable simulations and with a buffered 14-7 potential (b) used in AMOEBA. Different van der Waals models are compared in (c). In all plots, the van der Waals interactions are plotted for a pair of ions of the same type (fluoride – fluoride, chloride – chloride, etc).

Jones (NBFIX) parameters that were systematically calibrated against available experimental data as well as *ab initio* geometry optimizations and energy evaluations. Instead of tuning the van der Waals interactions, one can also change electrostatic interactions: As described before, standard additive force fields can be modified to include polarisation in a mean-field fashion. The ECC method models the polarisability of ions in an averaged static way by scaling the ionic charges by a factor of $1/\sqrt{\epsilon_{el}}$ as described in eq. (2.35), where ϵ_{el} is the dielectric constant of the solvent [102] and has been used to examine halide-nanotube interactions recently [105]. However, the ECC method cannot model how the strength of dipoles changes in response to their local environment over time [115]. Conventional pairwise additive force fields use the two-body Coulomb potential to describe charge–charge interactions and cannot take into account the effect of induced polarisation. A polarisable atomic multipole model such as AMOEBA describes the charge distribution of atoms not only with partial charges but also includes higher multipole moments and can model the effect of induced polarisation. The Drude oscillator model implemented in the CHARMM Drude force field is another approach that incorporates polarisation effects into classical MD simulations. Indeed, Drude oscillators have previously been used to investigate the effects of polarisation on the hydration of chloride [116].

Fig. 2.3 shows the 12-6 Lennard-Jones potential for the four halide anions. The chloride parameters from Li *et al*. in Fig. 2.3(a) relate both the radius and the well depth via a so-called “noble-gas” curve, which is a fit based on the VDW parameters of noble gas atoms to correlate the radius and the well depth in the Lennard-Jones potential in eq. 2.22 [113]. With increasing ion size, both σ and ϵ increase. As described earlier, the AMOEBA force field uses the buffered 14-7 potential in eq. 2.27 to describe VDW interactions. The 14-7 potential is plotted for the four halide ions in Fig. 2.3(b). Chloride described by the AMOEBA 14-7 potential has a slightly shallower well than the Li *et al*. chloride as depicted in Fig. 2.3(c).

2.2.2 Molecular Dynamics Algorithms

Energy Minimsation

On the potential energy surface, minimum points are of particular interest because they correspond to stable states of the system. A minimum point is a configuration where any small displacement of the atoms leads to an increase in the system's potential energy. An effective analogy for visualizing the potential energy surface is to think of it as a landscape with mountains and valleys [72]. In this analogy, minima are like valleys in the landscape, representing stable configurations where the energy is at a local minimum. Local minima represent a stable state, but not necessarily the most stable. A global minimum represent the deepest valley, corresponding to the most stable configuration with the lowest possible energy. Mathematically, a minimum on the potential energy surface is characterised by the first derivative of the energy function U with respect to each coordinate being zero and the second derivatives (the Hessian matrix of second partial derivatives) being positive, indicating that the curvature of the energy surface is upwards in all directions at the minimum point:

$$\begin{aligned} \frac{\partial U}{\partial x_i} &= 0 \\ \frac{\partial^2 U}{\partial x_i^2} &> 0 \end{aligned} \tag{2.36}$$

In a system with high dimensionality, such as one with $3N$ coordinates, finding analytical solutions for the minima is not feasible. Therefore, numerical methods are employed. These methods can be broadly classified into two categories: Methods using derivatives and methods not using derivatives. The former, such as gradient descent and conjugate gradient, utilise the first and sometimes second derivatives of the energy with respect to the coordinates to iteratively find the minima. The latter do not rely on derivatives and are often used when the derivative information is difficult to obtain or the energy surface is highly irregular.

The simplex or Nelder-Mead method is a non-derivative minimization technique often used in various optimization problems. While the simplex method is not utilised for minimizing the simulation box in the MD studies due to its computational intensity, it is employed for identifying a pathway through an ion channel pore, as detailed in Chapter 5. In an n-dimensional problem, the simplex algorithm operates on a geometric figure known as a simplex, which has $n + 1$ vertices. The algorithm navigates the potential energy surface akin to an amoeba, adapting its shape and position to locate the minimum energy point. The method involves three primary types of moves: Reflection, expansion and contraction [117]. In a reflection move, the vertex with the highest energy value is reflected through the opposite face of the simplex to generate a new point. This reflection aims to find a lower energy configuration. If the new point resulting from reflection has a lower energy than any other point in the simplex, an additional expansion move is applied to explore further in that direction [117]. When the simplex reaches a valley and reflection fails to yield a better point, the simplex contracts along one dimension from the highest energy point, reducing its size and honing in on the local minimum, a contraction move [72]. These moves are iteratively applied to dynamically adjust the simplex and drive it towards the minimum energy configuration. An initial simplex, formed by $3N + 1$ energy evaluations, is essential to begin the optimization process [72]. Given the high dimensionality and the need for numerous energy calculations, the Simplex method is deemed computationally expensive for MD systems [72].

Typically, initial protein structures derived from experimental methods such as X-ray crystallography or cryo-EM do not include hydrogen atoms. Before commencing atomistic MD simulations, usually hydrogens are added along with other components such as water, ions, and lipids to create a realistic simulation environment. Although a protein from a cryo-EM structure is likely in a local minimum, it requires fur-

ther minimization after adding hydrogens and placing it in the simulation box to ensure the stability of the simulated system. The potential energy in MD simulations is a differentiable function of the atomic coordinates and hence allows the use of derivative-based minimization methods. The steepest descent method is a first-order minimization technique used to find the local minimum of the potential energy surface. It iteratively adjusts the coordinates of the atoms to reduce the system's energy. In each iteration of the algorithm, the system moves in the direction parallel to the net force acting on it, which corresponds to the negative gradient of the potential energy. This is analogous to walking straight downhill on a landscape, following the steepest path downward. Once the direction of movement is determined, the algorithm needs to decide how far to move along this direction: Initially, a default step size is used for the movement. If the step results in a reduction in energy, the step size is increased in subsequent iterations to accelerate convergence. If the step increases the energy, the step size is reduced. This reduction is based on the assumption that the algorithm might have overshot the minimum and leapt across a valley containing the minimum energy point.

Integration Algorithm

In MD simulations, the set of atomic positions evolves over time by numerically integrating Newton's equations of motion, as described in section 2.1.1:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \nabla U, \tag{2.37}$$

where m_i refers to the mass of atom i, $\vec{r}_i$ to its position, and U is the potential energy defined by the force field. Molecular dynamics is a deterministic method, meaning that the future state of the system can be precisely predicted from its current state. Given the positions and velocities of all atoms at a particular time, the subsequent positions and velocities can be determined by solving Newton's equations step by step. Integrators are algorithms used to numerically solve Newton's equations of motion. They compute the trajectories of atoms by advancing their positions and velocities over small discrete time steps. Two commonly used integrators in MD simulations are the leap-frog algorithm and the velocity Verlet algorithm. The default integrator in *Gromacs* is the leap-frog algorithm. In this method, the positions and velocities of atoms "leap-frog" over each other, being calculated at staggered time points. Specifically, the velocities $v(t + \Delta t/2)$ are calculated at half-integer time steps, while the positions $r(t + \Delta t)$ are updated at integer time steps using these velocities. This results in kinetic energy and other velocity-dependent properties being evaluated half a time step out of sync with the atomic positions.

$$\begin{aligned} \vec{v}(t + \frac{\Delta t}{2}) &= \vec{v}(t - \frac{\Delta t}{2}) + \Delta t \frac{\vec{F}(y)}{m} \\ \vec{r}(t + \Delta t) &= \vec{r}(t) + \Delta t \vec{v}(t + \frac{\Delta t}{2}) \end{aligned} \tag{2.38}$$

Another popular integrator is the velocity Verlet algorithm, which updates positions and velocities at integer time steps, making it convenient for calculating properties that depend on both [118].

When using a polarisable force field such as AMOEBA, time integration is typically performed using the reversible reference system propagator algorithm (RESPA) [119] with an outer time step and an inner time step. Van der Waals and electrostatic multipole forces are updated in intervals of the outer time step and all other forces are evaluated at each inner time step. In Chapter 6, the self-consistent induced dipole moments between iterations were converged to $\epsilon < 0.000001$ D in eq. (2.33).

The self-consistent field method directly embodies the physical assumption underlying the Born-Oppenheimer approximation: that electronic degrees of freedom relax much faster than any nuclear motion, ensuring that the electrons in an MD simulation remain in the lowest energy state. However, in

practice, solving the corresponding equation requires an iterative process that demands repeated evaluation of electrostatic interactions, leading to increased computational expense. To address this, a more efficient dual-thermostat extended Lagrangian scheme is employed for simulations with the CHARMM Drude force field. In this method, each Drude oscillator is assigned a small mass m_D usually about 0.4 u, which is deducted from the mass of the nucleus to maintain the total atomic mass m_i. The reduced mass

$$m_i' = m_D \left(1 - \frac{m_D}{m_i}\right) \tag{2.39}$$

can then be viewed as a moment of inertia, allowing the Drude displacements to be integrated as additional dynamic variables [120]. It has been demonstrated that, to accurately replicate the dynamics observed in the self-consistent field regime, the Drude oscillators should be kept at a fictitious temperature T_D lower than T [96]. This is achieved by coupling the nuclear positions and Drude displacements to two separate thermostats. Instead of solving the classical equation of motions in eq. (2.37), dual Langevin dynamics of the form

$$\begin{aligned} m_i \frac{d^2\vec{r}_i}{dt^2} &= \nabla_{\vec{r}_i} U - \gamma \frac{d\vec{r}_i}{dt} + \sqrt{2\gamma' k_B T}\vec{R}(t) \\ m_i' \frac{d^2\vec{d}_i}{dt^2} &= \nabla_{\vec{d}_i} U - \gamma_* \frac{d\vec{d}_i}{dt} + \sqrt{2\gamma_* k_B T_D}\vec{R}_*(t) \end{aligned} \tag{2.40}$$

can be solved, where m_i' is the reduced mass defined in eq. (2.39), γ and γ_* are friction coefficients and $\vec{R}(t)$ and $\vec{R}_*(t)$ are zero mean delta-correlated Gaussian processes, which couple the dynamic variables to a heat batch of temperature T and T_D respectively [121]. Eq. (2.40) can be solved through an extension of the Verlet method described earlier [122]. When the Drude particles move far away from the nucleus, simulations become unstable and the polarisation catastrophe may occur. Therefore, a hard wall constraints on the length of Drude-nuclei bonds is introduced to avoid such instability. If the Drude particle moves further away from the nucleus than the hard wall threshold, it will be reflected back towards the nucleus [123].

Constraints

For stable integration using the leapfrog integrator, the time step must be sufficiently small to capture the fastest dynamical processes in the simulation system, typically the vibrations of covalent bonds, especially those involving hydrogen atoms. These vibrations restrict the time step to a maximum of 1 fs [72]. However, since bond vibrations are often of little interest, it is common practice to constrain these bonds to their equilibrium length, allowing for a larger time step, usually 2 fs [118]. The LINCS algorithm imposes these constraints, ensuring that the integrator updates all atom coordinates according to the forces provided while eliminating explicit high-frequency motion [124]. This is particularly advantageous for constraining bonds to hydrogen atoms, which have fast vibrational frequencies due to their small mass. The LINCS algorithm operates in two main steps: first, it updates the bond length, and then it corrects for bond rotations [125].

Temperature and Pressure Coupling

Direct integration of Newton's equations of motion produces the microcanonical or NVE ensemble. However, for biomolecular simulations that typically employ the isothermal-isobaric or NPT ensemble, thermostat and barostat algorithms are necessary to maintain the temperature T and pressure P at constant physiological values. In most simulations in this thesis, the velocity rescaling thermostat and the Parrinello-Rahman

barostat are used. The temperature T is related to kinetic energy K of N particles by the equation

$$K = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2 = \frac{3}{2} N k_B T. \tag{2.41}$$

Several algorithms exist for sampling at constant temperature; here, I primarily use the velocity rescaling thermostat with a stochastic term by Bussi *et al.*, which couples the system to an external heat bath fixed at a specified temperature T_0 [126]. The Bussi thermostat is based on the Berendsen thermostat but correctly samples the Boltzmann energy distribution. In Berendsen temperature coupling, the kinetic energy of all particles is proportionally scaled to correct any deviation of the system temperature from T_0, with deviations decaying exponentially with a time constant τ [73]:

$$\frac{dT(t)}{dt} = \frac{T_0 - T(t)}{\tau} \tag{2.42}$$

This method suppresses thermal fluctuations, and while the simulations remain near the target temperature T_0, they do not sample the canonical ensemble. The Berendsen thermostat is only used for equilibrations as described in Table 3.1. Bussi *et al.* improved this by including a stochastic term, propagating kinetic energy as an auxiliary dynamical variable according to the stochastic differential equation:

$$dK = (K_0 - K)\frac{dt}{\tau} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau}}, \tag{2.43}$$

where N_f is the number of degrees of freedom and dW is a Wiener process [126]. Switching off the noise dW reproduces the Berendsen algorithm.

An alternative approach to temperature coupling in molecular dynamics simulations is stochastic Langevin dynamics [121], which modifies the equation of motion for each particle by including friction and noise terms

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F}(\vec{r}) - m\gamma\frac{d\vec{r}}{dt} + \eta, \tag{2.44}$$

where γ is the friction constant and η is a noise process with an autocorrelation function

$$< \eta(t)\eta(t+\tau) >= 2\frac{k_B T}{m}\gamma\delta(\tau). \tag{2.45}$$

This modification is applied separately to all degrees of freedom without correlation between them. When $1/\gamma$ is large compared to the timescales of the simulated processes, this method effectively provides stochastic temperature coupling while dampening longer time-scale processes [121]. In the simulations, the Langevin thermostat is used when employing the CHARMM Drude force field.

To replicate experimental conditions that are often at atmospheric pressure rather than at a fixed volume, a pressure coupling algorithm is used. The instantaneous internal pressure within a simulation box can be calculated using the kinetic energy K and the forces on the particles in the virial form [118]:

$$P = \frac{2}{3V}K + \frac{1}{3V}\sum_{i=1}^{N} \vec{r}_i \cdot \vec{F}_i. \tag{2.46}$$

To achieve pressure coupling, the simulation box is scaled to match the desired external pressure. The Berendsen barostat accomplishes this by treating the system similarly to the Berendsen thermostat [73], using a first-order decay approach:

$$\frac{dP(t)}{dt} = \frac{P_0 - P(t)}{\tau_p}, \tag{2.47}$$

Here, τ_p represents the pressure coupling constant. The simulation system is weakly coupled to a pressure bath by scaling the volume by a certain scaling factor λ

$$\lambda = 1 - \kappa \frac{\Delta t}{\tau_p}(P - P_0), \tag{2.48}$$

where κ is the isothermal compressibility of the system. Instead of scaling the volume, one can scale the positions of atoms by $\lambda^{1/3}$. As the Berendsen thermostat, the Berendsen barosatat produces the correct average pressure but does not reproduce the correct ensemble and is only used for equilibration simulations. Furthermore, in systems that involve a membrane, semi-isotropic coupling should be employed. Parrinello and Rahman presented a barostat for anisotropic pressures where the system Hamiltonian is modified and the volume is made a dynamic variable [127]. A Parrinello-Rahman barostat is used in most simulations described in this thesis. This method allows for pressure changes in the xy-plane to be treated independently from those in the z direction, providing more accurate simulation conditions for membrane systems [127].

An alternative method for pressure coupling in molecular dynamics simulations is the Monte Carlo barostat, which employs a rejection sampling procedure to generate stochastic box volume changes [128]. I use the Monte Carlo barostat for simulations with the AMOEBA force field. In this method, a trial volume change $\Delta V$is chosen from a uniform distribution within the interval $[-\Delta V_{max}, \Delta V_{max}]$. The length of the simulation box is then scaled by the relative volume change according to the equation:

$$L_i' = \frac{V + \Delta V}{V} L_i, \tag{2.49}$$

where $i \in \{x, y, z\}$ denotes the spatial dimension. Similarly, particle positions are rescaled according to

$$r_i' = \frac{V + \Delta V}{V}(r_i - c_i) + c_i, \tag{2.50}$$

where $\vec{c}$ denotes the box centre. The thermodynamic cost of the proposed volume change is then calculated as

$$\Delta W = \Delta U_{pot} + p\Delta V - N k_B T \ln\left(\frac{V + \Delta V}{V}\right), \tag{2.51}$$

where U_{pot} is the change in the potential energy resulting from the change of volume. The volume move is accepted with a probability given by

$$P(\Delta W) = \begin{cases} 1 & \text{if } \Delta W > 0 \\ \exp\left(-\dfrac{\Delta W}{k_B T}\right) & \text{if } \Delta W \leq 0 \end{cases} \tag{2.52}$$

The parameter V_{max} is dynamically adjusted to achieve an acceptance probability of approximately 50% [128].

Periodic Boundary Conditions

The number of atoms in a simulation box is a critical factor that can limit the feasibility and accuracy of MD simulations. To ensure efficient computation, the smallest possible box that adequately represents the system of interest is used. However, boundary effects between the explicit solvent and the surrounding vacuum necessitate very large boxes to obtain bulk solution properties for a centrally solvated protein. To circumvent the issue of large simulation boxes, periodic boundary conditions (PBCs) are employed. PBCs eliminate the solution-vacuum interface by allowing atoms that leave through one face of the box to re-enter through the opposite face, thereby mimicking an infinite system. This technique makes the

environment within the box more bulk-like, which is essential for accurately simulating properties that depend on bulk behaviour [72]. Under PBCs, particle interactions are calculated using the minimum image convention. Each particle only interacts with the closest periodic image of another particle [129]. This ensures that the computational effort is minimized while maintaining realistic interactions. To avoid interactions between a particle and its own image, the box size must be at least twice as large as the non-bonded cutoff distance [118]. The computational complexity of MD simulations is heavily influenced by the calculation of non-bonded interactions, which scale with the square of the number of atoms, N^2. The short-range Lennard-Jones Potential potential in eq. (2.22) falls off rapidly with distance, following an r^{-6} dependence for the attractive term. The long-range potential Coulomb potential in eq. (2.19) decreases more slowly, with an r^{-1} dependence. To optimize calculations, a cutoff distance is introduced, beyond which interactions are not computed [118]. For the Lennard-Jones potential, this is feasible due to its rapid decay. The cutoff must be less than half the length of the cubic cell to ensure that a particle does not interact with its own image. Despite the cutoff, the initial calculation still requires determining the distance between every pair of atoms, making it computationally intensive [72]. To reduce the computational load, a non-bonded neighbour list is maintained, storing atoms within the cutoff distance and slightly beyond. This list allows efficient updates and calculations over multiple MD time steps, as atoms do not move significantly within this short period. The neighbour list is updated when the sum of maximum displacements of any two atoms exceeds the difference between the cutoff distance and the neighbour list distance [130]. For larger systems, the cell index method is used, dividing the simulation box into smaller cells with lengths longer than the non-bonded cutoff distance. This method helps manage and update neighbour lists more efficiently [118].

The introduction of a cutoff distance introduces discontinuities in potential energy and force near the cutoff value, which can affect energy conservation in MD simulations. To address this, one can use a shifted potential where a constant term is subtracted from the potential at all values, though this introduces a discontinuity in the force. Alternatively, one can use a polynomial function to multiply the potential energy, transitioning its value smoothly from 1 at $r = 0$ to 0 at the cutoff distance. This method ensures that both the potential energy and its first derivative, the force, smoothly approach zero at the cutoff, enhancing stability and energy conservation [72].

Particle Mesh Wald Method

The electrostatic potential decays slower with distance than VDW interactions. This slower decay necessitates a different approach for handling these interactions in computational simulations. Unlike short-ranged forces, where interactions quickly diminish and can be truncated beyond a certain distance, long-ranged electrostatic interactions extend much further. Charge-charge interactions decay with distance following a r^{-1} relationship. This extensive range poses a challenge because particles interact not only with their immediate neighbours but also with particles across the periodic boundaries of the simulation box [72]. One effective method to handle long-range electrostatic interactions in a periodic system is the Ewald summation method. This technique considers that each particle interacts with all other particles in the simulation box as well as with all of their periodic images in an infinite array of periodic cells. This inclusion ensures that the long-range nature of electrostatic interactions is properly accounted for. The Particle Mesh Ewald method does not adhere strictly to the minimum image convention as it is includes self interactions across periodic boundaries. The electrostatic potential of a box with N particles that interact with all their

periodic images can be written as a nested sum:

$$U_{\text{Coulomb}} = \frac{1}{8\pi\epsilon_0} \sum_{\vec{n}} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{|\vec{r}_{ij} + \vec{n}|}, \tag{2.53}$$

where the inner sums exclude $i = j$ if $|\vec{n}| = 0$. Here, $\vec{n}$ is a vector indicating the position of the periodic image [131]. However, the direct summation of these interactions converges slowly and is conditionally convergent. This means that if one sums only the positive terms or only the negative terms, each series will diverge. The trick in the Ewald summation method is to transform the problematic summation into two separate series that each converge more rapidly. The Ewald summation method achieves this by choosing an appropriate auxiliary function, $f(r)$, to handle the rapid variation of the $1/r$ term at small distances and the slow decay at large distances [132]. By doing so, the original interaction term is split into a short-ranged term and a long-ranged term, both of which are easier to sum. In practice, the Ewald method involves considering each charge to be surrounded by a neutralizing charge distribution of equal magnitude but opposite sign. A Gaussian charge distribution is commonly used for this purpose because it has desirable mathematical properties that facilitate the separation of the potential into rapidly and slowly converging series. The interaction in real space is short-ranged due to the neutralizing Gaussian charge [133]. This part is computationally manageable because it rapidly decays and can be truncated beyond a certain distance. The long-range interactions are handled in reciprocal (Fourier) space. In this domain, the interactions decay rapidly, making the summation more efficient. Since each particle interacts with its own neutralizing distribution, a correction term is included to subtract this self-interaction, ensuring the accuracy of the method. One has to choose an appropriate function $f(r)$ which deals with rapid variation of $1/r$ at small r and slow decay at long r.

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \tag{2.54}$$

In the Ewald method, each charge is considered to be surrounded by a neutralising charge distribution of equal magnitude but of opposite sign [132]. A Gaussian charge distribution is commonly used. The complementary error function

$$f(x) = erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \tag{2.55}$$

fulfills the desired properties and the overall Ewald sum becomes

$$U_{\text{Coulomb}} = \frac{1}{8\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{N} \begin{cases} \sum_{\vec{n}} \dfrac{q_i q_j erfc(\alpha|\vec{r}_{ij} + \vec{n}|)}{|\vec{r}_{ij} + \vec{n}|} \\ + \sum_{k \neq 0} \dfrac{q_i q_j}{\pi L^3} \dfrac{4\pi^2}{k^2} \exp\left(-\dfrac{k^2}{4\alpha^2}\right) \cos(\vec{k} \cdot \vec{r}_{ij}) \\ - \dfrac{\alpha}{\sqrt{\pi}} \sum_{k=1}^{N} q_k^2 + \dfrac{1}{6L^3\epsilon_0} |\sum_{k=1}^{N} q_k r_k| \end{cases}, \tag{2.56}$$

where α is a parameter to pick such that reasonable cutoffs can be used for both the real-space and the reciprocal sums. The first term in equation (2.56) is the real-space sum in terms of box translation vectors $\vec{n}$, the second term is the reciprocal-space sum in terms of reciprocal vectors $\vec{k}$, the third term is a correction to remove the the self-interaction of the added Gaussians and the fourth term is a correction to account for the surrounding medium [131]. Evaluating the sum is computationally still expensive and scales with N^2. Instead of direct summation of the wave vectors in the reciprocal part, the atomic charges are interpolated onto a grid in the particle mesh ewald (PME) method, which allows usage of the fast-fourier algorithm and results in computational costs scaling with $N \log N$ [72, 131].

2.2.3 Boltzmann Inversion of Water Number Density

When studying ion channel structures, pore hydration observed through equilibrium MD simulations can serve as a proxy for assessing ion permeability [27]. A dry or poorly hydrated pore imposes an energetic barrier on the ions, as they must shed their hydration shell to traverse the channel, leading to decreased ion permeability. Thus, I examine pore hydration with MD simulations to get a proxy for potential barriers to permeation when I functionally annotate ion channel structures in Chapter 4. The time-dependent number density $n(z)$ of water molecules at a given point z along the channel pore axis is related through the free energy $G(z)$ of water through

$$n(z) = C \exp\left(-\frac{G(z)}{k_B T}\right), \tag{2.57}$$

where C is a normalisation constant that is related to the partition function of the system. The free energy $G(z)$ of water along the channel axis can be obtained by inverting eq. (2.57):

$$G(z) = -k_B T(\ln n(x) - \ln C), \tag{2.58}$$

where one can determine constant C by the convention of setting the free energy in the bulk water region on both sides of the pore to zero. The number density along the channel axis z can be defined as the number of water particles per unit volume. Pore radius profiles calculated with HOLE [134], CHAP [40] or PoreAnalyser [69] can be used to estimate the pore volume. In MD simulations, every water molecule is characterised with a discrete position. Using kernel density estimation, where each water molecule position is associated with a Gaussian probability distribution, as implemented in CHAP [40], the overall denisty can be obtained via a summation of all Gaussians. Each Gaussian is characterised by a smoothing bandwidth, for which I choose 0.14 nm, the radius of a water molecules, in this thesis [33].

2.2.4 Potential of Mean Force Calculations

When one studies an ion channel, one is sometimes interested in energetic barriers for ion permeation along the channel pore axis. Or when studying an ion carrier, one might be interested in the binding affinity of a particular ion to the cavity of the ion carrier. The free energy change can be measured as a function of a so-called reaction coordinate which could be the centre of mass (COM) distance of the ion channel and the permeating ion. The free energy surface along the reaction coordinate is referred to as the potential of mean force (PMF). In principle, the free energy can be obtained directly by counting bound and unbound configurations in long unbiased simulations. When studying the energetic barrier for permeation of a certain ion type through a pore, one would have to sample enough permeating events. When calculating the binding affinity of an ion to an ion transporter, one would have to sample binding and unbinding events of the ion to the cavity of interest. In practice, this is rarely feasible due to energy barriers between phase space regions that prevent exhaustive sampling of all relevant phase space configurations in finite time as described in section 2.1.2. This is generally referred to as the problem of quasi-nonergodicity [74]. In silico calculations of free energy rely on three components: a model Hamiltonian, a sampling protocol, and a method to estimate the free energy difference [74]. A model Hamiltonian is used to compute energies and forces and to ensure that all configurations are sampled with the correct relative probability. In this thesis, two different types of classical Hamiltonians are used, generally referred to as force fields as discussed in section 2.2.1. They mainly differ in the way they describe non-bonded interactions; namely van der Waals and electrostatic interactions. In section 2.2.1 and Fig. 2.3, it is discussed how different force fields describe VDW interactions of chloride ions differently. The second component of free energy calculations is the sampling protocol. The (Gibbs) free energy difference between two states A and B, that represent the

bound and the unbound state of an ion with an ion carrier in an isothermal–isobaric (NPT) ensemble, can be written as

$$\Delta G = G_B - G_A = -RT \ln \frac{Z_B(NPT)}{Z_A(NPT)}. \tag{2.59}$$

The free energy difference depends on the ratio of the partition functions $Z_i(NPT)$ of the systems A and B. In Chapter 6, a physical pathway method is used to measure the reversible work, the potential of mean force, to transfer the ion from the bulk into the cavity of the host molecule. The relationship between the standard free energy of binding and the change in the PMF on binding is often assumed to be equivalent [135]. In this thesis, I calculate the binding free energy as the difference between the bulk and the minimum at the binding site as computed by the PMF. In Chapter 7, the free energy surface along the channel axis is calculated for chloride ions in GlyR to investigate the energy barriers to ion permeation.

In case of the ion carriers studied in Chapter 6, the PMF describes the reversible work to transfer an ion from the bulk to the cavity inside the host molecule [136]. The PMF is related to the probability $p(r)$ of finding the ion at a given distance r from the centre of the cavity through the Boltzmann equation

$$\mathrm{PMF}(r) = -RT \ln p(r) + C, \tag{2.60}$$

where C is an offset that does not influence the dynamics of the system. The offset is usually chosen with respect to some reference state. In principle, the probability $p(r)$ could be estimated directly from an MD trajectory. However, if the PMF contains high barriers, the simulation is unlikely to fully sample the relevant range of distances r. To circumvent the problem of quasi-nonergodicity [74], one can carry out N simulations, during each of which the force field potential is modified by a harmonic biasing potential

$$w_i(r) = k_r(r - r_i)^2, \tag{2.61}$$

where k_r denotes a force constant which restrains the distance r to a values near r_i. During N simulations, r_i is varied to ensure that the relevant range of distances r is sampled. This approach is called umbrella sampling (US). Each umbrella simulation gives a biased probability $p_i(r)$. The biased probabilities $p_i(r)$ can be combined to an unbiased probability $p(r)$ by employing the so-called weighted histogram analysis method (WHAM) [137, 138] which is based on iteratively solving the following coupled equations

$$\begin{aligned} p(r) &= \frac{\sum_{i=1}^{N} n_i p_i(r)}{\sum_{i=1}^{N} n_i \exp(-\beta(w_i(r) - F_i))} \\ F_i &= -\beta \ln \left(\int dr \frac{p(r)}{\exp(\beta w_i(r)} \right), \end{aligned} \tag{2.62}$$

where n_i denotes the total number of samples generated from the i-th simulation. Once can interpret F_i as the free energy required to introduce the i-th biasing potential into the system.

Chapter 3
Simulation Protocols

Disclaimer
Part of this chapter on computational methods are based on the method sections in my published work in the Journal of General Physiology, Seiferth *et al*. (2022) [19], in Nature Communications, Gibbs *et al*. (2023) [11], in Physical Chemistry Chemical Physics, Seiferth *et al*. (2023) [71], and in the Biophysical Journal, Seiferth and Biggin (2024) [69]. All investigation, writing and figures in these papers were produced by myself with feedback from the co-authors, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

This chapter gives a summary about simulation protocols that have been used in Chapter 4 to Chapter 7. It ranges from force field parametrisation in section 3.1, how to set up simulation boxes in section 3.2, equilibration protocols in section 3.3, simulation settings in section 3.4, umbrella sampling protocols in 3.5, hydration shell analysis protocols in 3.6 to energy decomposition analysis in section 3.7 and simulations with external potentials in section 3.8.

3.1 Force Field Parametrisation

When ion channels or carbon nanotubes are simulated in lipid bilayers, one can use already established force field parameters and does not have to parameterise parameters oneself. In Chapter 6, binding affinities are calculated of anions to the cavity of different ion carriers which consist of multiple biotin subunits forming a macrocycle. To parameterise the host molecules, QM methods were used as described in section 2.1.3.

3.1.1 GAFF and Parsley

For the simulations with non-polarisable, pairwise additive force fields in Chapter 6, parameters for the biotin macrocycles were generated with the second generation of the General AMBER force field (GAFF) [139] and the Parsley force field from the Open Force Field Initiative [140]. *Antechamber* and *tleap* were used for the GAFF2 force field [141]. For the non-bonded interactions, van der Waals parameters were inherited from the AMBER force fields [82] and partial charges were assigned according to the so-called Austin Model 1 bond charge corrections (AM1-BCC) model. Mulliken charges, see section 2.1.3, were obtained by a semi-empirical AM1 calculation, followed by a bond charge correction (BCC) scheme [139]. The Parsley force field relies on an AM1-BCC model for charge assignment as well [140].

3.1.2 AMOEBA Parametrisation

As described in section 2.2.1, the AMOEBA force fields incorporates a polarisable multipole framework up to quadrupole moments [142]. The parametrisation of the biotin macrocycles for the AMOEBA force field followed the POLTYPE protocol [143]. Density functional theory calculations were used to derive electrostatic parameters. The defaults in the POLTYPE protocol are *ab initio* calculations which were also used by Laury *et al*. for the parametrisation of the cucurbit[8]uril molecule [144]. The initial structure was optimized using the semi-empirical PM3 method and then optimized with B3LYP level of theory with

a 6-311G(1d,1p) basis set and empirical dispersion gdBJ using the Gaussian 16 software package. The optimized structure was used as an input for a distributed multipole analysis using Stone's *GDMA* [89] that determined initial atomic multipole estimates such as the charge, the components of the dipole vector and the quadrupole tensor. The *Tinker* program *POLEDIT* was used to choose local multipole frames and atomic polarisabilities and to define polarisation groups. A polarisation group in AMOEBA defines a group of atoms whose permanent multipoles do not polarise one another. In AMOEBA, the multipole moment vector is decomposed into a permanent contribution, independent of the environment and an induced component that can change according to the local electric field of other atoms. Within a polarisation group, the permanent multipoles are excluded from contributing to the field at an atomic site, and hence only the other induced dipoles polarise the atom in question [145]. A single point calculation at B3LYP level of theory with cc-ptvz basis and empirical dispersion gd3BJ was used as a reference point for electrostatic potential fitting. A Cartesian grid of points, at which to calculate the electrostatic potential, was created with the *Tinker POTENTIAL* program [146]. The QM electrostatic potential was evaluated at each of these grid points with the Gaussian *CUBEGEN* program. Then, the AMOEBA electrostatic potential was evaluated at each grid point and the dipole and quadrupole components were fitted to minimize the root mean square error between the QM and AMOEBA electrostatic potential with the *Tinker POTENTIAL* program while keeping the partial charges (the monopoles) fixed [146]. Finally, parameters for bonds, angles, stretch-bends, out-of-plane bends, torsions and van der Waals interactions were assigned by similarity to the AMOEBA09 parameter set with the *Tinker VALENCE* program [142, 146]. The AMOEBA09 parameter file contains atom classes for many functional groups such as saturated carbon and hydrogen, amine nitrogen, sulfides and carboxylic acids.

3.2 System Set-up

In Chapter 6, biotin macrocycles are simulated in solvent boxes with water or acetonitrile. In Chapter 4, Chapter 5 and Chapter 7, ion channel proteins and carbon nanotubes are simulated in lipid bilayers. In this section, I summarise how to set up the respective simulation boxes.

3.2.1 Solvent Boxes

To prepare simulation boxes of different ion carriers solvated either in water or in acetonitirle, *Gromacs* programs [129] were used for simulations with non-polarisable and *Tinker* programs [146] for simulations with the AMOEBA force field. For the AMOEBA simulations, a cubic, periodic boundary box was defined with a side length of 46 Å with the *Tinker xyzedit* programme which was also used to soak the host molecule in a box of solvent molecules of either water or acetonitrile. The setup included one host molecule of the biotin macrocycle, one anion (either fluoride, chloride, bromide, or iodide), and one potassium cation. In water, I ran simulations with the carboxyl groups of the biotin side chains protonated or deprotonated. To neutralise the charge of the box, six potassium ions were added accordingly. For the simulations with the GAFF or parsley force field, the simulation boxes were defined to be slightly larger (60 Å) with the *Gromacs editconf* programme, the simulation box was filled with the *solvent* programme and solvent molecules were replaced with ions with the *genion* programme.

3.2.2 Membrane Systems

In Chapter 5, carbon nanotubes were inserted into a bilayer composed of 128 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipids with the *InflateGro* method [147]. Fig. 5.3 shows an example of a

simulation box. Then, the simulation box was solvated with water and ions (KCl) at 150 mM concentration were added.

When studying ion channels such as GlyR in Chapter 4 or Chapter 7, the CHARMM-GUI Membrane Builder [148, 149] was used to prepare the simulation box. Unless specified otherwise, the ion channels were embedded in a POPC bilayer. The box was solvated with water and charge neutralised with ions. A physiological ion concentration of 150 mM (either KCl or NaCl) was maintained by adding the appropriate number of ions to the system.

3.3 Equilibration Protocols

3.3.1 Membrane System Equilibration Protocol

Energy minimization of the simulation box consisting of an ion channel in a membrane was performed using the steepest descent integrator with strong positional restraints on the backbone (4000 kJ/mol/nm²), side chains (2000 kJ/mol/nm²), and lipid phosphate groups (1000 kJ/mol/nm²). The minimization was set to converge when the maximum force on any atom dropped below 1000.0 kJ/mol/nm or after a maximum of 5000 steps. The neighbour list was updated every 10 steps using the Verlet cutoff scheme with a neighbour list cutoff of 1.2 nm. Van der Waals interactions were treated with a cut-off modified by a force-switch starting at 1.0 nm and switching off at 1.2 nm, while electrostatic interactions were calculated using the PME method with a cutoff of 1.2 nm. All hydrogen bonds were constrained using the LINCS algorithm. This energy minimization ensured that the system was brought to a stable low-energy configuration, reducing any unfavourable contacts and preparing it for subsequent equilibration and production simulations.

The membrane system equilibration was conducted using a multi-step protocol as detailed in Table 3.1. The protocol involved a series of simulations designed to gradually relax the system while maintaining stability and proper configuration of the membrane and associated molecules. The equilibration process was divided into five stages, each characterised by different ensembles, time steps, simulation durations, and force constants applied to various parts of the system, including the backbone, side chains, and lipid phosphate groups. As a first step after energy minimisation, the system was simulated in the NVT ensemble with a time step of 0.001 ps for 125 ps, applying strong positional restraints to the backbone (4000 kJ/mol/nm²), side chains (2000 kJ/mol/nm²), and lipid phosphate groups (1000 kJ/mol/nm²). In the second stage, the NVT ensemble was maintained for another 125 ps, but the force constants were reduced to 2000 kJ/mol/nm² for the backbone, 1000 kJ/mol/nm² for the side chains, and 400 kJ/mol/nm² for the lipid phosphate groups, allowing further relaxation. The third stage transitioned to the NPT ensemble with the same time step and simulation duration, further reducing the positional restraints to 1000 kJ/mol/nm² for the backbone, 500 kJ/mol/nm² for the side chains, and 400 kJ/mol/nm² for the lipid phosphate groups, enabling the system to equilibrate under constant pressure conditions. In the fourth stage, the time step was increased to 0.002 ps, and the simulation was extended to 500 ps with further reduced force

Table 3.1: Membrane System Equilibration Protocol.

	Ensemble	Time Step [ps]	Simulation Time [ps]	Backbone [kJ/mol/nm^2]	Side Chain [kJ/mol/nm^2]	Lipid Phosphate [kJ/mol/nm^2]
1	NVT	0.001	125	4000	2000	1000
2	NVT	0.001	125	2000	1000	400
3	NPT	0.001	125	1000	500	400
4	NPT	0.002	500	500	200	200
5	NPT	0.002	50,000	500	0	0

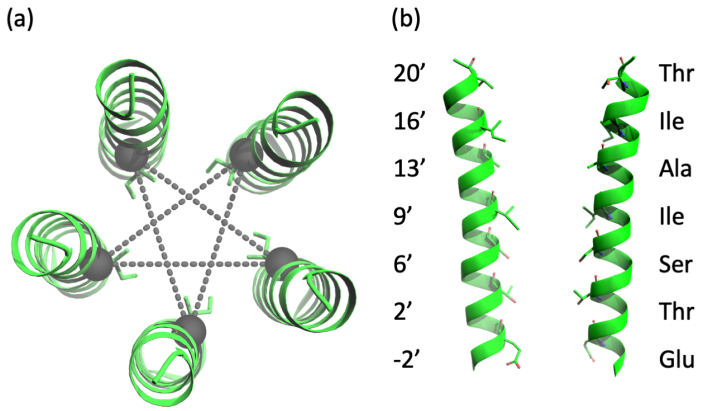


Figure 3.1: Pore lining M2 helices from the five GLIC subunits with the isoleucine residues at the 9' position in stick representation (a). The flat bottom restraints act on the C_α atoms and are indicated by dashed lines. Two non-adjacent pore lining M2 helices with seven pore lining residues (b).

constants of 500 kJ/mol/nm² for the backbone, 200 kJ/mol/nm² for the side chains, and 200 kJ/mol/nm² for the lipid phosphate groups, allowing more freedom for the side chains and lipid phosphate groups. The final stage involved a long equilibration period of 50 ns in the NPT ensemble with a time step of 0.002 ps, applying minimal restraints of 500 kJ/mol/nm² to the backbone and no restraints to the side chains and lipid phosphate groups, ensuring the system reached a fully relaxed state. Each stage built upon the previous one, and was a direct continuation of the previous step using the velocities from the last simulation frame. For the first step, the velocities were randomly generated. Temperature in step 1-4 was maintained with a Berendsen thermostat and pressure in step 4 was maintained by a Berendsen barostat [73]. In the final step 5 of the equilibration protocol, temperature and pressure were maintained, using the velocity-rescaling thermostat in combination with a semi-isotropic Parrinello and Rahman barostat. In all steps, temperature was set to 310 K and pressure to 1 bar with coupling constants of 1 and 5 ps, respectively.

3.3.2 Extended Equilibration Protocol for Pore Lining Residues

Dämgen and Biggin introduced an extended equilibration protocol targeting pore-lining residues in the GlyR, which was applied to relax a GLIC structure in section 4.2.2. In their work, Dämgen and Biggin noted that open-state GlyR structures tend to rapidly collapse during MD simulations. They demonstrated that a stable open state can be maintained when the leucine residues at the 9' position shift into well-conserved hydrophobic pockets located at the subunit interfaces, see Fig. 1.4(a). For the GLIC channel discussed in section 4.2.2, the pore was held open using flat-bottom harmonic cross-distance position restraints applied between selected C_α atoms of the pore-lining residues (Thr at 20', Ile at 16', Ala at 13', Ile at 9', Ser at 6', Thr at 2', and Glu at −2'). Fig. 3.1 illustrates the restraints to five C_α-pairs of the isoleucine at the 9' position. The flat-bottom potential only exerts a force when the C_α atoms of non-adjacent pore-lining residues approach each other more closely than their distance in the open-state cryo-EM structure, with the force being proportional to the distance difference (5000 kJ $mol^{-1}nm^{-2}$). This configuration permits the relaxation of the pore-lining side chains and is utilised for GLIC in section 4.2.2.

3.4 Simulation Settings

3.4.1 Non-Polarisable Force Field Simulations

Parsley Simulation of Macrocycles in Solvent Boxes

For the simulations in Chapter 6 with the Parsley force field, *Gromacs* [129] was used as the simulation engine using the leap-frog algorithm with a 2 fs time step. Electrostatics were computed with the PME method [131, 133] with a 14 Å cut-off. The thermostat employing velocity rescaling with a stochastic term maintained the temperature for simulations carried out with *Gromacs* with 0.1 ps coupling time. The corresponding gromcas keyword is v-rescale. The pressure was maintained by the Parrinello–Rahman barostat [127] with 2 ps coupling time for production simulations. The isothermal compressibility for water was set to $4.5 \cdot 10^{-5}$ bar^{-1} and for acetonitrile was set to $10.8 \cdot 10^{-5}$ bar^{-1} [150].

CHARMM36m Simulations of Ion Channels in Membranes

Ion channels in membranes were simulated using the TIP3P water model [108] and the CHARMM36m forcefield [151] with pair-specific Lennard-Jones parameters to correct the chloride interactions with proteins, lipids, and alkali cations [114] unless specified otherwise. The *Gromacs* simulation engine [129, 68] was used to integrate Newton's equations of motions employing a leap-frog algorithm with a 2 fs time step. The neighbour list was updated every 20 steps using the Verlet cutoff scheme, with a list cutoff radius of 1.2 nm. Van der Waals interactions were treated with a cutoff scheme modified by a force-switch, switching on at 1.0 nm and off at 1.2 nm. Electrostatic interactions were calculated using the PME method with a cutoff of 1.2 nm [133]. Temperature coupling was managed using the v-rescale algorithm with separate coupling groups for the solute, membrane, and solvent, each with a time constant of 1.0 ps and a reference temperature of 310 K. Pressure coupling was applied using the Parrinello-Rahman method in a semi-isotropic manner, with a time constant of 5.0 ps, isothermal compressibility for water $4.5 \cdot 10^{-5}$ bar^{-1}, and a reference pressure of 1.0 bar. Reference coordinate scaling was set to the center of mass when restraints were applied on the protein to maintain a certain conformation. All bonds involving hydrogen atoms were constrained using the LINCS algorithm, allowing for a larger time step by maintaining the geometric constraints of these bonds [124]. To maintain overall system stability, the center of mass motion was removed every 100 steps for the groups solute, membrane, and solvent, using a linear center of mass removal algorithm.

3.4.2 Polarisable Force Field Simulations

AMOEBA Simulations of Macrocycles in Solvent Boxes

The AMOEBA simulation in Chapter 6 were carried out with *Tinker9* [146]. Similar parameters were used as in a recent study using AMOEBA for calculating PMFs of an ion channel [152]. Time integration was performed using the reversible reference system propagator algorithm (RESPA) [119] with an outer time step of 2 fs and an inner time step of 0.25 fs. The van der Waals and electrostatic multipole forces were updated in intervals of the outer time step and all other forces were evaluated at each inner time step. The self-consistent induced dipole moments between iterations were converged to ε < 0.000001 D. The Ewald cut-off was set to 7.0 Å and the van der Waals cut-off to 10 Å. For polarisable atomic multipole interactions, *Tinker* provides regular Ewald summation method to include long range electrostatic interactions in a system with periodic boundary conditions. The temperature of the simulation system was maintained at a temperature of 298.15 K or 300 K using a Bussi thermostat [126] with 0.1 ps coupling time. The *Tinker* Bussi and the gromcas v-rescale thermostat are the same. 300 K was used for simulations where the system was solvated in water and 298.15 K was used for simulations in acetonitrile to match the temperature at which

isothermal titration calorimetry was performed by Lisbjerg *et al.* [153]. A Monte Carlo barostat was used to maintain a pressure of 1 atm with 2 ps coupling time.

CHARMM Drude Simulations of GlyR

The CHARMM Drude simulations of the GlyR in Chapter 7 were conducted with *openMM* [154]. Drude particles were added to the equilibrated non-polarisable system (see Table 3.1) with the Drude Prepper in the CHARMM-GUI [148, 155]. For the Drude simulations, a smaller time step of 0.001 ps as opposed to the 0.002 ps production time step for CHARMM36m simulations was chosen. Electrostatic interactions were calculated using the PME method with an Ewald error tolerance of 0.0001. Van der Waals interactions were treated with the Switch method, employing a switch-on distance of 1.0 nm and a switch-off distance of 1.2 nm. Simulations were maintained at a constant temperature of 310 K using Langevin dynamics with a friction coefficient of 5, see eq. (2.40). The Drude particles, representing polarisable elements, were controlled at a much lower temperature of 1.0 K with a higher friction coefficient of 20 and a hardwall constraint of 0.02 nm. Pressure coupling was enabled, with the reference pressure set to 1.0 bar by a Monte Carlo membrane barostat. The pressure coupling frequency was 100 steps. Hydrogen bonds were constrained during the simulations.

3.5 Umbrella Sampling Protocol

As described in section 2.2.4, the free energy along a reaction pathway can be calculated using different biased simulation windows. In this thesis, the free energy is calculated for ions permeating ion channels or ions binding to the cavity of an ion carrier. In this section, the set-up is described for umbrella sampling windows to obtain the potential of mean force. Eq. (2.62) can be solved by Grossfield's implementation of WHAM [137] or by the *Gromacs* WHAM version [138].

3.5.1 Ion Carriers

In Chapter 6, the PMF was calculated with umbrella sampling for two different force fields (Parsley, AMOEBA), two different solvents (water and acetonitrile) and three halide anions (chloride, bromide, iodide). The distance between the z-coordinates of the centre of mass (COM) of the cavity (defined by twelve carbon atoms) and the z-coordinate of the ion was set as the reaction coordinate as depicted in Fig. 6.1(c). Position restraints to three carbon atoms of the cavity were applied to retain the orientation of the carrier during the simulations while still allowing the cavity full flexibility. The z-coordinate of the ion was harmonically restrained to a specific distance with respect to the COM of the cavity in each umbrella window to ensure sampling along the reaction coordinate, see Fig. 6.1(d). For Parsley simulations, 25 × 25 ns windows were simulated with 1 Å spacing for the distance bias of 500 kJ/mol/nm^2. The reaction coordinate was sampled between 0 Å < z < 25 Å. For some systems, additional windows with higher bias were needed to ensure sufficient sampling in energy barrier regions along the reaction coordinate. In addition to the distance restraint of the z-coordinate, flat bottom restraints were introduced in the xy-plane to force the ion to stay within a cylinder of radius $R_{cyl} = 2.5$ Å or 3.5 Å. Within the cylinder, the ion could move freely and explore the phase space. Outside of the cylinder, there were high harmonic restraints of 5000 kJ/mol/nm^2, forcing the ion back to the cylinder. This cylinder ensured that the ion could not move laterally away from the carrier when its z-coordinate was restrained to a value which corresponds to a region where binding and unbinding events were very frequent. Unbinding events in the xy-plane are irreversible on the time scale of the window (25 ns) and affect convergence of the potential of mean force calculations. In long unbiased simulations (500 ns), one can observe a handful of binding and unbinding events. Simulating each

window for 500 ns was not feasible and hence I studied only the binding and unbinding dynamics along the z-coordinate and restrained the movement of the ion in the xy-plane to an area of the size of the cavity. This allowed me to calculate the area sampled by the ion orthogonal to the reaction coordinate and limits the sampling required in each window [135]. The binding free energy

$$\Delta G = -RT\ln(C^0 K_b) \tag{3.1}$$

with equilibrium binding constant K_b and standard concentration $C^0 = 1/1661\,\text{Å}^{-3}$ can be corrected by the term

$$\Delta G_{\text{corr}} = -RT\ln\frac{\pi R_{\text{cyl}}^2}{A^0}, \tag{3.2}$$

where πR_{cyl}^2 is the surface of the cylinder [135, 156]. When the surface of the cylinder equals the so-called standard surface A_0, the correction is zero. I define the standard surface as $A_0 = \pi r_0^2 = 170\,\text{Å}^2$ if r_0 denotes the radius of a sphere with standard volume 1661 Å^3.

The first nanosecond of each window was discarded as equilibration time. The remaining 24 ns were divided into three sub-windows of 8 ns length defining three data sets. From each of the three sets the PMF was calculated. The bulk region was defined as the zero point. The minimum of the PMF averaged over three data sets is interpreted as binding free energy and the standard deviation of the minimum as the error. For AMOEBA simulations, the set-up is similar except for that the z-coordinate of the ion was biased with 2 kcal/mol/Å^2 and 0.5 Å spacing in 60 windows between −15 Å < z < 15 Å distance to the COM of the cavity. Each window was simulated for 13 ns of which the first nanosecond was discarded as equilibration time. The remaining simulation time was divided into three parts to define three data sets. The histogram overlap for each PMF was also checked to ensure good overlap between windows. Example overlap data for this is shown in Fig. B.1 and Fig. B.3 in the appendix.

3.5.2 Carbon Nanotubes

In Chapter 5, the distance between the respective ion and the centre of mass of the CNT structure along the z-axis of the simulation box was chosen as the collective variable. Starting configurations were generated by swapping positions of the respective ion with water molecules along the permeation pathway. The minimum of the biasing potential was shifted by 0.2 nm between each umbrella sampling window. A force constant of 500 kJ/mol/nm^2 restrained the z-coordinate of potassium or chloride ion to the respective distance to the centre of mass of the CNT. In addition to the distance restraint of the reaction coordinate, flat bottom restraints were introduced in the xy-plane to force the ion to stay within a cylinder with a sufficiently large radius $r_{cyl} > a \geq b$, where a and b denote the radii of the elliptical cross-sectional area. Within the cylinder, the ion could move freely while outside of the cylinder, there were high harmonic restraints of 5000 kJ/mol/nm^2, forcing the ion back to the cylinder. This cylinder ensured that the ion could not move laterally away from the CNT when its z-coordinate was restrained to the entry or exit of the CNT. The energetic barrier was corrected as described in eq. (3.2) [71].

3.5.3 Ion Channels

In Chapter 7, the umbrella sampling simulations of GlyR were conducted using two different engines: the *openMM* engine [154] for CHARMM Drude simulations and the *Gromacs* engine [129, 68] for CHARMM36m simulations with pair specific Lennard Jones (NBFIX) parameters [151, 114]. The starting configurations were prepared by swapping positions of the respective chloride ion with water molecules along the permeation pathway. The minimum of the biasing potential was incrementally shifted by 0.1 nm between each umbrella

sampling window, creating a total of 91 windows spanning from -45 Å to +45 Å around the COM of the protein. A force constant of 1000 kJ/mol/nm² was used to restrain the z-coordinate of the chloride ion relative to the COM of the protein. In *Gromacs*, this was achieved using the pull code, while in *openMM*, a custom external force was defined for this purpose. Each window underwent an initial equilibration period of 1 ns. For the non-polarisable CHARMM36m simulations, the subsequent production run for each window lasted 21 ns, whereas for the polarisable CHARMM Drude simulations, it lasted 9 ns. The total window time was further divided into three sub-windows of 7 ns or 3 ns each, resulting in three datasets per window. The PMF was calculated from each of these three datasets. The bulk region served as the reference zero point for the PMF calculations. To estimate the error, the PMF was averaged across the three datasets, providing a measure of the variability and reliability of the results.

3.6 Hydration Shell Analysis

In Chapter 6, the hydration shell of anions binding ion carriers is plotted as a function of the distance to the cavity. Chapter 7 investigates how the hydration shell of chloride is affected when chloride permeates GlyR. For each umbrella sampling window, a python analysis script noted which and how many atoms were within a cutoff range of the respective ion. The cut-off distances were derived by analysing the RDF of each respective ion with the corresponding solvent, as depicted in Fig. B.14. For interactions of ions with the host molecules, biotin macrocycles in Chapter 6 and GlyR in Chapter 7, the cut-off distances were set at 3.0 Å for chloride, 3.25 Å for bromide, and 3.5 Å for iodide as depicted in Table 3.2. For acetonitrile, the end of the first solvation shell was defined for different atomic interactions: the C-H_3 interaction cut-off distances were 5.5 Å for chloride, 5.9 Å for bromide, and 6.2 Å for iodide; for the hydrogen interactions, the cut-off distances were 4.4 Å for chloride, 4.5 Å for bromide, and 4.7 Å for iodide. For water, the end of the first solvation shell was defined separately for oxygen and hydrogen atoms. The oxygen interaction cut-off distances were 3.75 Å for chloride, 4.0 Å for bromide, and 4.25 Å for iodide, while the hydrogen interaction cut-off distances were 2.75 Å for chloride, 3.0 Å for bromide, and 3.25 Å for iodide.

3.7 Energy Decomposition Analysis

In Chapter 6, the interaction energy decomposition is calculated between ion and host as a function of distance of the ion to the cavity. Even though one cannot experimentally verify an interaction energy decomposition, it can shed light on the relative importance of different contributions such as electrostatic and van der Waals interactions.

Table 3.2: Hydration shell analysis. Cut-off distances for counting ion interactions with biotin, acetonitrile and water. The cut-offs are determined by analysing the RDF of the respective ion with the respective solvent as depicted in Fig. B.14.

	Distances in Å	Chloride	Bromide	Iodide
Biotin or Protein	Cutoff distance for interactions	3.0	3.25	3.5
Acetonitrile	End of 1st solvation shell (C-H_3)	5.5	5.9	6.2
	End of 1st solvation shell (hydrogen)	4.4	4.5	4.7
Water	End of 1st solvation shell (oxygen)	3.75	4.0	4.25
	End of 1st solvation shell (hydrogen)	2.75	3.0	3.25

3.7.1 Symmetry-Adapted Perturbation Theory (SAPT)

To perform symmetry-adapted perturbation theory (SAPT) with DFT, the B3LYP level of theory was utilised to determine the interaction energy between the hosts and chloride ions [79]. The jun-cc-pVDZ basis set, which includes chloride, was employed for these calculations. Initially, ten representative frames were extracted from the AMOEBA simulations. These frames were used to compare the energy decomposition obtained from DFT with that from the AMOEBA force field. Due to the computational expense of DFT calculations, SAPT0 was only performed for the bound poses and not along the entire reaction coordinate.

The calculations were conducted using the *Psi4* software package [80] with the following configuration: a memory allocation of 250 GB, the jun-cc-pVDZ basis set, density fitting (DF) for the self-consistent field (SCF) method, and the core electrons were frozen during the computations to reduce computational cost. The SAPT0 energy calculations were initiated with the command *energy('sapt0')*. This approach provided a "ground truth" for the interaction energies, allowing a detailed comparison of the energy components derived from DFT and the AMOEBA force field in Chapter 6.

3.7.2 MD Interaction Energy

To quantify the interaction strength between anions and host molecules in MD simulations, the nonbonded interaction energy between these species was computed. Although this quantity is not a free energy or binding energy and is not typically parameterised to be physically meaningful in most force fields, it provides insight into the interaction dynamics along the reaction coordinate. For *Gromacs* simulations, the energy decomposition capability was utilised to calculate short-range nonbonded energies between defined groups using the *energygrps* keyword in the molecular dynamics parameter file. This allows to monitor how interaction energies varied along the reaction coordinate. For AMOEBA simulations conducted with the *Tinker9* engine [146], the *analyze* module with the D option (Details for All Individual Interactions) was employed to obtain Van der Waals, atomic multipoles, and polarisation contributions. From each umbrella sampling window, 13 frames were selected from the AMOEBA simulations spanning 13 ns, with each frame being 1 ns apart. This allowed for the calculation of interaction energy profiles along the reaction coordinates and the decomposition of the total energy into individual components. The interaction energy profiles were referenced to bulk conditions, setting the interaction energy of the host and guest to zero at a distance of 15 Å from the center of the cavity. This approach enabled a detailed analysis of the interaction energy variations along the reaction coordinate, providing insights into the dynamics of host-anion interactions in the simulated systems.

3.8 External Potential

In some molecular dynamics simulations, a constant electric field is applied to model the influence of an external voltage on the membrane system. Each partial charge q_i experiences an additional force due to the electric field $\vec{E}$, expressed as $q_i\vec{E}$. The electric field is oriented in the z-direction, which is perpendicular to the plane of the membrane. The voltage V applied across the membrane is related to the electric field and the length of the simulation box in the z-direction L_z, given by the relation

$$V = EL_z. \tag{3.3}$$

Here, L_z represents the length of the unit cell in the z-direction and not the physical thickness of the membrane. In Chapter 4 and Chapter 5, three independent simulation runs were conducted for each condition, each lasting 250 ns. During these simulations, a potential of 500 mV was applied and the number of permeation events was monitored. Then the mean and standard deviation was calculated

of the permeation events observed across the three repeats, providing a measure of the variability and consistency of the results. In Chapter 7, I try to relate permeation events observed during MD simulations to experimental single channel conductance. The current I can be estimated using

$$I = \frac{Nq}{T}, \tag{3.4}$$

where N is the number of permeation events of charges q observed in a time interval T. Following Sotomayor *et al.*, the errors σ_I on ionic currents I can be estimated by assuming a Poisson distribution for permeation events [157]:

$$\sigma_I = \frac{I}{\sqrt{N}} \tag{3.5}$$

With longer simulation time T and more observed permeation events N, the uncertainty increases. Conductances g can be estimated using

$$g = \frac{I}{V}, \tag{3.6}$$

where V is the electrostatic potential drop across the membrane.

With periodic boundary conditions and a single-bilayer setup, the simulation box is treated as if it repeats infinitely in all directions. This means that the two sides of the membrane are effectively part of the same continuous compartment. When an external electric field is applied across the simulation box, it creates a potential difference across the membrane, but due to the periodic boundary conditions, the physical environment on either side of the membrane remains identical in terms of ion concentration and overall composition [158, 159]. The ionic concentrations on both sides of the membrane are the same.

An alternative method is computational electrophysiology [160, 161]. Unlike the method of applying an external electric field, the voltage in computational electrophysiology simulations is internal in the sense that it arises from an ion concentration difference between two aqueous compartments. This setup mirrors the physiological conditions in cells, where Nernst concentration gradients exist across the cell membrane. To achieve the separation of two compartments in a periodic simulation system, the membrane patch is duplicated in the direction of the membrane normal. This arrangement creates an inner and outer compartment, with the latter connected across the periodic boundaries. Instead of applying an external potential across the simulation box, computational electrophysiology establishes a maintained ion gradient across the membrane, resulting in sustained ion flux through the channel. To preserve the chosen ionic gradient, ion/water pairs are exchanged between both sides of the membrane [161]. However, the focus is on simulations employing an external field in this thesis. In Chapter 7, simulations with an external potential field are compared for the pairwise additive CHARMM36m force field [151] and the polarisable CHARMM Drude force field. At the time of writing this thesis, computational electrophysiology is only implemented in *gromcas*p. The *openMM* engine is used for CHARMM Drude simulations and cannot use computational electrophysiology when using a polarisable force field.

Chapter 4
Functional Annotation of Ion Channels

Disclaimer
Parts of this chapter on hydrophobic gating are based on my published work in the Journal of General Physiology, Seiferth *et al.* (2022) [19], see also in the list of publications. The example of functional annotations in section 4.2.1 and 4.2.2 are based on my simulations that I performed for the published work in Nature Communications, Gibbs *et al.* (2023) [11] and Bharambe *et al.* (2024) [33]. All investigation, writing and figures in that paper were produced by myself with feedback from the co-authors, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

The flux of ions through a channel is usually regulated by changes that result in steric occlusion of its pore as depicted in Fig. 4.1. But ion permeation can also be prevented by formation of a desolvation barrier created by hydrophobic residues that line the pore. Relatively minor structural changes can result in confined hydrophobic regions in channels undergoing transitions between wet and dry states to gate the pore closed without physical constriction of the permeation pathway. This concept is referred to as hydrophobic gating, and many examples of this process have been demonstrated. However, the term is also now being used in a much broader context that often deviates from its original meaning. In this chapter, the hydrophobic gating concept is recapitulated, and examples of this process are compared with other gating mechanisms that simply exploit hydrophobic residues and/or lipids in steric closure of the pore. Furthermore, best practices for identification of a hydrophobic gate are described and exemplified by annotating newly resolved GlyR and GLIC structures.

4.1 When is a Hydrophobic Gate not a Hydrophobic Gate?

In the 1990s, Nigel Unwin solved the structure of the torpedo nicotinic acetylcholine receptor structure and observed a wide, but very hydrophobic, central pore [162, 163]. Previously published studies indicated that the hydrophobic nature of the pore in this new channel structure could act as a potential barrier to permeation and that it might not be conductive even though it appeared wide enough to accommodate a hydrated or partially hydrated ion: Lum *et al.*, studied the thermodynamics of liquid-to-vapour transitions in water confined between two hydrophobic plates [164], and Brovchenko *et al.* performed simulations of confined water [165]. These studies laid the foundations for to the concept of hydrophobic gating, which is now shown to occur in several different classes of ion channels. The gating mechanism of the nicotinic receptor is depicted in Fig. 4.1(b).

The phenomenon of hydrophobic gating was initially demonstrated using molecular dynamics simulations in model nanopores [17, 42, 166] and in carbon nanotubes designed to mimic ion channels [167, 168]. It was shown that, if hydrophobic enough, narrow pores wider than a hydrated ion could form a hydrophobic gate that prevents permeation due to liquid–vapour transitions of water within the pore. A functionally closed, dewetted pore could then be opened in two ways: firstly, by increasing the diameter of the narrowest region of the pore; or secondly, by increasing the hydrophilicity of that region allowing it to become wetted and thus permeable as depicted in Fig. 4.2 (a) [42]. This phenomenon is now referred to as “hydrophobic gating” [17, 169], thereby reflecting the behaviour of water within these confined spaces, but has also sometimes been called a vapour lock [41] or bubble gating [170].

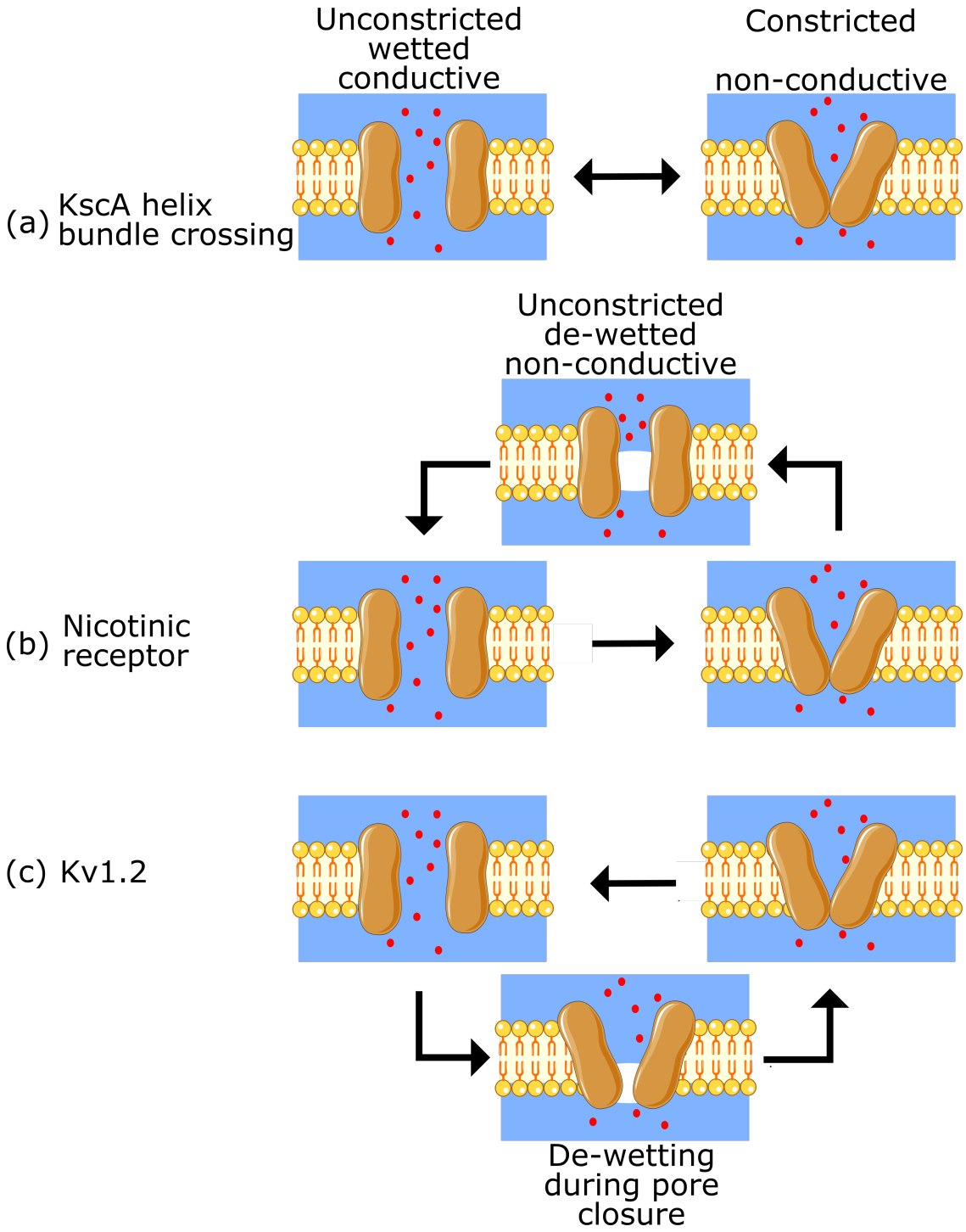


Figure 4.1: Different gating mechanisms. Water (blue), ions (red), ion channel (brown), and lipid bilayer (yellow). (a) Gating via helix bundle crossing, where ion movement is prevented by steric occlusion of the helix bundle crossing in the non-conductive state. An example is the KcsA channel. (b) In the resting state of the nicotinic receptor, ion movement is prevented by a **hydrophobic gate** where the pore pathway is not sterically occluded (see middle top panel). Further changes then open the hydrophobic gate, and the channel becomes conductive. (c) Similar to KcsA, the closure of the bundle crossing gate in Kv1.2 produces a physically occluded closed state. The hydrophobicity of the bundle-crossing gate permits tight packaging of the helices in the closed state and also produces dewetting during closure. The formal definition of hydrophobic gating would exclude such cases where pore dewetting immediately precedes steric closure of the pore.

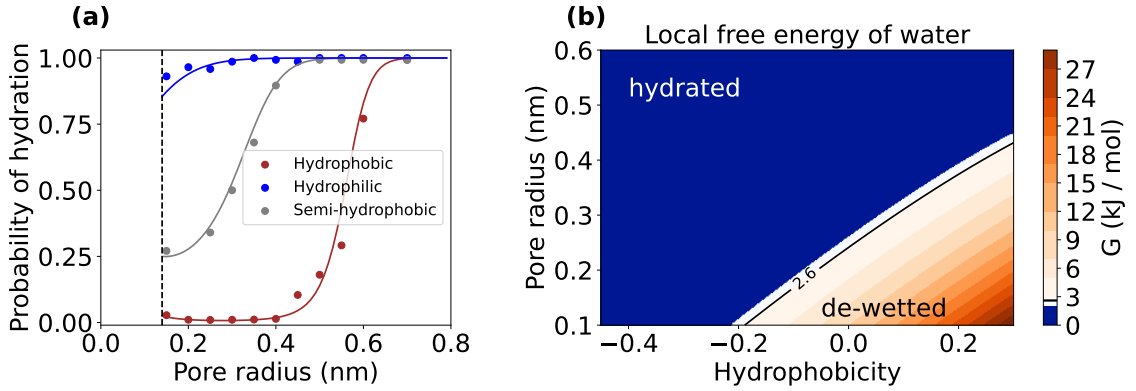


Figure 4.2: The role of pore radius and hydrophobicity. (a) The pore hydration probability depends on its radius and hydrophobicity. Based on studies of a model nanopore, a hydrophilic pore remains hydrated if its radius is larger than that of a water molecule (0.13 nm, indicated by the dotted line). By contrast, a hydrophobic pore dewets at radii below 0.6 nm (original data from [42]). (b) Using a simple heuristic based on analysing the hydration of many different pore structures [43], the radius and hydrophobicity of a pore can be used to predict its hydration state. The hydrophobicity employs the Wimley-White scale [171], with more positive values indicating greater hydrophobicity. The water-free energy is based on mean measurements from pore radius, hydrophobicity, and water-free energy profiles obtained from number density calculations.

Water can exhibit anomalous behaviour at the macroscale, for example, where surface tension promotes the formation of droplets on a hydrophobic surface. However, these properties can also manifest when water is confined within hydrophobic environments on the nanoscale, such as those found in some ion channel pores, where it promotes alternating evaporation and condensation events. These spontaneous dewetting transitions present a major energetic barrier to a hydrated or partially hydrated ion without the usual requirement for steric occlusion of the pathway. In contrast, water confined by polar surfaces in pores with similar size exists exclusively in the condensed phase, and these pores are therefore permeable. Other factors such as temperature [42] and transmembrane voltage have also been shown to affect this process [44, 45, 46, 47], but the primary determinants of such hydrophobic gates appear to be the radius of the pore and its local hydrophobicity.

Many ion channel and nanopore structures in different conformational states have been revealed over the last two decades, thanks to advances in structural biology along with recent artificial intelligence–driven structure prediction algorithms, such as AlphaFold [172]. In Fig. 1.3, the number of available channel structures in the PDB is plotted over time. However, in nearly every case, these structures require functional annotation [40]. In some cases, it is relatively easy to define whether a given structure corresponds to an open or a closed state by determining the physical dimensions of the pore with applications such as HOLE [134], CHAP [40] or PoreAnalyser [69]. Likewise, sterically occluded pores can quickly be annotated as functionally closed. However, a simple determination of pore radius is often not enough due to the process of hydrophobic gating because hydrophobicity also plays a major role in defining pore permeability.

There is considerable experimental evidence to support the anomalous behaviour of water at the nanoscale [173, 174]. But, it is not currently possible to directly observe water dynamics within a channel pore, and evidence for hydrophobic gating relies mainly on molecular simulation studies supported by mutagenesis studies [38]. Molecular dynamics simulations therefore help bridge the gap between structural and functional observations, and more detailed computational approaches such as free energy profiles for ion permeation can also be determined, as well as using the behaviour of water in a channel pore as a proxy for ionic permeability [27]. Molecular simulations can aid the functional annotation of channel pore

structures with software such as CHAP, which was specifically developed for this purpose [40]. Parameters such as pore radius and hydrophobicity seemed so influential that Rao *et al*. developed a simple heuristic that is capable of predicting the permeability of any given pore structure [43]. The free energy barrier for water is plotted in Fig. 4.2(b) as a heuristic function of hydrophobicity and pore radius. However, when simulating water in confined environments, the water models and force fields used must also be chosen carefully [39], and in some cases, the use of polarisable force fields may be required to capture effects where the electronic structure of an atom is altered by other charges in its immediate environment [14, 152].

The concept of hydrophobic gating in channels and nanopores has gained considerable traction over the past few years with many different examples now being observed. However, the increased visibility of this phenomenon has also resulted in the term hydrophobic gate being used in a much broader context that often deviates from its original definition. For instance, nonpolar residues have been proposed to form a hydrophobic gate in several cases where wet–dry transitions have not been observed [175, 176], whilst in other cases, the direct occlusion of pores by lipids has been referred to as hydrophobic gating [177].

The hydrophobic gating mechanism depicted in Fig. 4.1(b) involves a change in radius (without steric closure) or smaller scale changes such as rotation of side chains that can change both radius and hydrophobicity or sometimes a combination of both. A change in radius or a change in hydrophobicity of the pore can promote dewetting. However, based upon its original meaning, the expression hydrophobic gating should only be used to describe the mechanism by which movement of a hydrated ion through the dewetted sections of a channel pore is prevented by the free energy barrier created by liquid–vapour transitions within the pore, rather than by van der Waals clashes with tightly packed pore residues, even if these residues are hydrophobic. It is important to note that this definition of hydrophobic gating also formally excludes cases such as in Fig. 4.1(c) where dewetting immediately precedes the subsequent steric closure of a pore, e.g., the dewetting events observed prior to closure of the helix-bundle crossing gate in simulations of the Kv1.2 pore domain [178].

4.1.1 Examples of Hydrophobic Gates in Ion Channels

According to this original definition of a hydrophobic gate, genuine examples have been demonstrated in several distinct classes of ion channels, in particular, within the family of pentameric ligand-gated ion channels (pLGICs) or Cys-loop receptors [20]. Hydrophobic residues in the pore-lining M2 helix of the Torpedo nAChR act as a hydrophobic gate that creates a large desolvation barrier for permeating ions [29, 30, 48]. Likewise, in GLIC, a bacterial homolog of nAChR, tilting of the pore-lining helices induces cooperative drying and closing of the pore by dewetting and not by constriction [31]. Similarly, in the serotonin 5-HT_3R, strong dewetting within the central hydrophobic gate region can be observed [27, 28, 152], whilst the GlyR also appears to show the same features [36, 37, 179]. That being said, recent structures of Cys-loop receptors, where both resting and desensitized structures have been solved, have shown that the minimum radius in both states can be very similar; this, therefore, implies a potential role for steric occlusion in the resting state gate, even if simulations show substantial dewetting in this region.

The TWIK-1 K2P potassium channel has also been shown to possess a hydrophobic barrier deep within its inner pore, and stochastic dewetting of this hydrophobic constriction acts as a major barrier to potassium permeation. The hydrophobic nature of the TWIK-1 pore restricts full hydration of the inner cavity, and polar substitutions at this hydrophobic cuff increase channel activity by directly disrupting this hydrophobic barrier to water and hydrated ions [18]. However, like many other aspects of TWIK-1 function, the mechanisms that regulate this barrier in the wild-type channel remain unclear.

Hydrophobic gates have also been found in the small conductance Ca^{2+}-activated channels (SK) and large conductance Ca^{2+}-activated channels (BK) [64, 180]. In the functional closed state of the BK channel,

the pore undergoes hydrophobic dewetting transitions and is not permeable to ions even though its pore remains physically open and accessible to relatively large molecules [180]. Simulations suggest that upon calcium-binding, the pore-lining helices bend and twist to move hydrophobic residues away from the pore and expose two glutamic acid residues to the pore, thus making it more hydrophilic [38, 180].

The heptameric small conductance mechanosensitive channel (MscS) has a hydrophobic pore with branched hydrophobic leucine side chains pointing into the pore. Anishkin and Sukharev (2004) demonstrated dewetting in MscS, which allowed for reinterpretation of the first MscS structure [41] that had previously been considered open [181]. Mutation of this leucine ring to polar residues caused a gain of function phenotype [182], whilst later work showed that the gate could also be opened by a rotation of these same leucine residues away from the pore, thus increasing its diameter at that point [183]. Simulation studies have also suggested a similar hydrophobic gate in the related large conductance mechanosensitive channel (MscL) [184], though it may be the precise dynamics of the dewetting process itself that are more important for gating [185].

A hydrophobic gate has also been proposed as the major determinant of inactivation in the mechanosensitive Piezo channels [186]. Members of the transient receptor potential (TRP) channel family have also been shown to exhibit hydrophobic gates that involve the motion of polar side chains to stabilize either the dry or wet state of the channel [187, 188].

Other examples include hydrophobic gates in the Ca^{2+} release-activated Ca^{2+} (CRAC) channel formed by two rings of pore-lining valine and phenylalanine residues that are rotated out of the pore upon channel opening [189], whilst in simulations of the magnesium channel CorA, spontaneous and reversible hydration events concurrent with small-amplitude fluctuations in pore diameter can be observed [190]. Likewise, in the Bestrophin chloride channel (BEST1), dewetting events that act as a hydrophobic gate have been observed in the narrow neck region of the pore that is lined by three rings of hydrophobic amino acids [191].

4.1.2 Other Mechanisms of Gating Involving Nonpolar Residues

Although nonpolar pore-lining residues are clearly important in the formation of hydrophobic gates, not every hydrophobic residue along the pore of a channel constitutes such a gate. As described above, the underlying principle of the hydrophobic gating mechanism is the creation of a free-energy barrier to permeation that is associated with stripping ions of their hydration shell. A genuine hydrophobic gate is not sterically occluded and must remain wide enough to allow a hydrated ion to pass if the pore was otherwise hydrophilic in that region. However, in some cases, hydrophobic pore segments and/or constrictions are sometimes referred to as hydrophobic gates in the sense they contain patches of hydrophobic residues that play an important role in gating and/or selectivity, but they do not meet the formal criteria originally described for this process.

A particularly interesting case is the early study of the Kv1.2 pore domain where a dewetting transition was observed immediately prior to steric occlusion of the pore by the closure of the bundle crossing gate [178]. Such bundle-crossing gates are clearly important in the gating of many members of this superfamily of tetrameric cation channels where the pore-lining helices intersect at the cytoplasmic entrance to seal the permeation pathway shut, see cartoon in Fig. 4.1 (a) and (c). In such cases, it is quite common to find patches of small hydrophobic residues at the point where these helices intersect as this permits tight packing of the helices in the closed state [192, 193]. Consequently, as these hydrophobic helices move toward the sterically occluded closed state, it is therefore likely that at some point in this transition a dewetting event will occur, and in some cases this is thought to assist the closure of the gate [192]. However, such dewetting

events immediately preceding such hydrophobic collapse and steric occlusion of a gate would not normally be classified as a classical hydrophobic gating mechanism.

In another, more recent case, two pore regions lined by hydrophobic residues were identified as hydrophobic gates within the transient chloride conductive channel-like state of the glutamate transporter (GltPh) [175]. These hydrophobic residues at the extracellular and intracellular openings clearly play an important role in the selectivity of the putative chloride permeation pathway in GltPh, and a similar role has also been proposed in its human homolog (EAAT1) [176]. However, no dewetting transitions were demonstrated, and computational studies suggest uninterrupted aqueous pathways in both cases. Thus, despite their location and obvious functional importance, these hydrophobic residues may not meet the criteria required of a genuine hydrophobic gate.

Another area of possible confusion arises from the fact that bound lipids and/or detergents are now being identified in many experimentally determined ion-channel structures [4, 194, 195, 196, 197, 198, 199]. Such lipids can influence channel function in various ways ranging from lipid-induced conformational changes [199] to direct blocking of a pore. For example, molecular dynamics simulations suggest that lipids can favourably assemble into a bilayer within the large hydrophobic pore of the calcium homeostasis modulator protein, CALHM2, but not within the smaller, more hydrophilic pore of CALHM1 [200]. There is also evidence that CALHM4 and CALHM6 may be gated by a lipid block in a similar manner to CALHM2 [201]. Hydrophobic residues within these pores thus clearly play a role in this “lipid plugging” process, but its mechanism and time scale are fundamentally different from that of hydrophobic gating. Dewetting occurs on a nanosecond timescale, whereas lipid plugging functionally closes the channel on a much longer time scale. Steric block by lipids may therefore be a possible mechanism for regulating large pores, but such plugging events are unlikely to be used for the dynamic regulation of a typical ion channel gate.

Direct pore occlusion by lipids has also been suggested in mechanosensitive channels such as TRAAK [196] and MscS [197]. Many K2P potassium channels such as TRAAK and the related TWIK-1 have large side portals or fenestrations at the interface between their transmembrane helices that potentially expose the ion conduction pathway to the lipid core of the bilayer [4, 196], though to what extent such lipid intrusion contributes to gating remains to be seen. In the TWIK-1 channel, simulations demonstrated that lipid tails could enter the fenestrations, but not far enough to sterically occlude the pore [4]. Interestingly, it has recently been argued that lipids seen within the pore of the mechanosensitive Piezo channel may result from a computational artifact because absolute binding free energy calculations indicate that the presence of these lipids within the pore is thermodynamically unfavourable and most likely results from being kinetically trapped in the pore during simulations [202]. Three pore-wetting mutations targeting the S6 helix of the human NaX channel have also recently been shown to unlock a voltage-insensitive leak conductance [203] and are referred to as a possible hydrophobic gate. However, further work is required to investigate whether this S6 gate is indeed a genuine hydrophobic gate or whether lipids identified in the structure occlude the pore.

Dynamic conformational changes in the amino terminus of the human Pannexin channel (PANX1) may also be associated with lipid movement in and out of its pore [204], and similar mechanisms have also been proposed for some of the other large-pore channels such as innexin [205] and connexin [206]. Such lipids blocking the permeation pathway in these channels have also been referred to as a hydrophobic gate [177], but these lipids sterically occlude the pore without evidence of a vapour lock being formed. Such references have therefore also added further confusion to the original definition of a hydrophobic gate, which is based on a fundamentally different principle of gating.

4.1.3 Functional Annotation of a Hydrophobic Gate

To help avoid such confusion, a series of best practices are summarised as nearly every newly solved ion channel structure requires some form of functional annotation to address whether it is open i.e., conductive or not. Fig. 4.3 illustrates best practices in functional annotation of ion channels. As an initial analysis step, the radius profile of the permeation pathway can be calculated with tools such as HOLE [134], CHAP [40] or PoreAnalyser [69]. The radius and hydrophobicity profile can be directly determined with CHAP from the PDB structure with no need for MD simulations. If the narrowest constriction exceeds the radius of the respective hydrated ion and is not hydrophobic, then the structure is considered to be open, and ion permeation can be further studied using more demanding approaches such as computational electrophysiology [160] or simulations with an external potential as described in section 3.8. However, if the narrowest constriction suggests a hydrated ion might pass, but appears to be hydrophobic, then further analysis is needed. Based on radius and hydrophobicity, a simple heuristic can also be used to predict the energetic barrier for water as illustrated in Fig. 4.2(b) [43], and if this score indicates a major barrier, then the channel structure should be embedded in a lipid bilayer and simulated with an atomistic force field. Appropriate backbone or C_{α} atom restraints should also be used to ensure that the functional annotation refers to the input structure and not some subsequent conformational change. If dewetting of the pore or a discontinuous water pathway is seen that creates an associated free energy barrier [40], then the structure can be considered to be closed and non-conductive due to a hydrophobic gate, but only if there is no evidence that this dewetted region subsequently constricts to sterically block the permeation pathway. In summary, attempts to functionally annotate the ever-increasing number of ion channels and pore structures have led to confusion about the formal definition of a hydrophobic gate. Not every hydrophobic cluster within a channel pore necessarily generates pore dewetting, and a genuine hydrophobic gate only functions as a gate if that region of the pore would otherwise remain permeable if it were more hydrophilic. Furthermore, dewetted regions within a pore do not function as hydrophobic gates if that region subsequently becomes sterically occluded. The term hydrophobic gating should therefore only be used to describe the mechanism by which dewetting events generate a free energy barrier for hydrated ions to prevent permeation through an otherwise wide enough pathway.

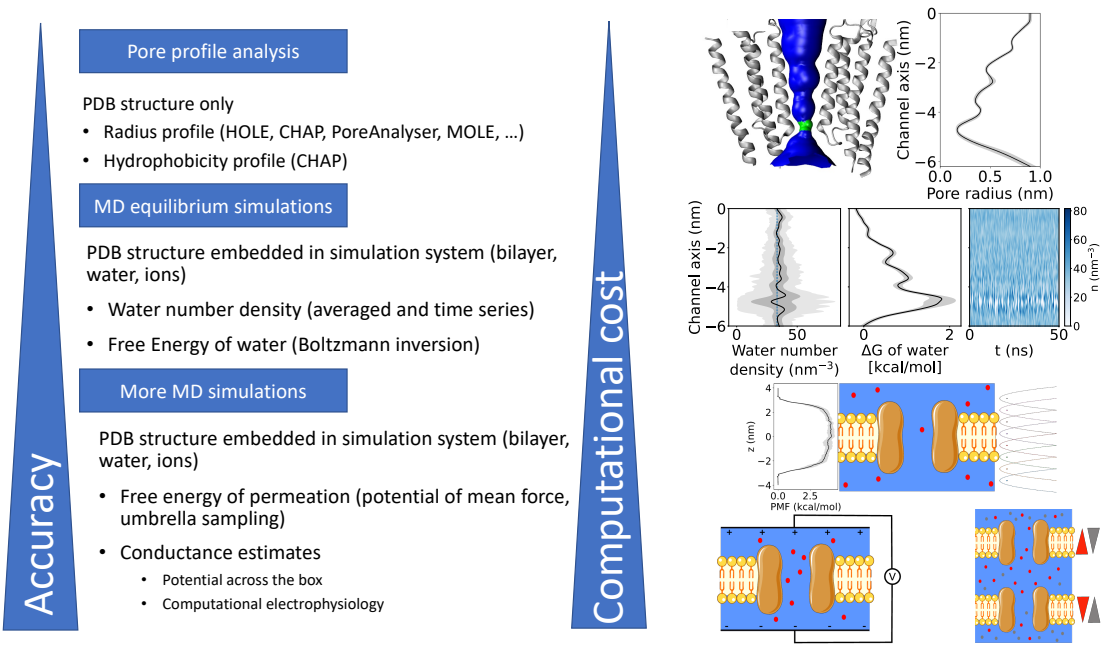


Figure 4.3: Summary of best practices for functional annotation of ion channels.

4.2 Examples of Functional Annotation

In the final section of this chapter, the best practices for functional annotation summarised in Fig. 4.3 will be applied to two distinct ion channels, specifically focusing on pLGICs. The classification of these receptors is typically based on the type of ions they conduct. For instance, cation-selective receptors such as the previously discussed 5-HT_3R and nAChR are categorized as excitatory receptors, meaning they facilitate the excitation of neurons. On the other hand, anion-selective receptors such as GlyR and γ-aminobutyric acid (GABA) receptors are classified as inhibitory receptors, which generally work to inhibit neural activity. The GlyR is one of the most widely distributed inhibitory receptors in the central nervous system. It plays an essential role in a variety of physiological processes, particularly in mediating inhibitory neurotransmission in the spinal cord and brainstem [207]. GlyR is activated by a range of agonists, including glycine, β-alanine, taurine, and ivermectin [208]. Additionally, it can be selectively blocked by the high-affinity competitive antagonist strychnine. Structurally, these ligand-gated ion channels are composed of five subunits that surround a central pore, with each subunit consisting of four α-helical transmembrane segments [32]. The second ion channel under consideration is the GLIC receptor, which is a bacterial ligand-gated ion channel derived from Gloeobacter. GLIC serves as a homolog to the nAChR and provides another example for the concept of hydrophobic gating in ion channels. The GLIC receptor is proton-gated, meaning it opens in response to the binding of a proton, and is cation-selective, allowing only positive ions to pass through. Structurally, GLIC, like the nAChRs, functions as a pentameric oligomer, meaning the channel operates as an assembly of five subunits. By applying functional annotation techniques described earlier to these two ion channels, one can gain insights into their physiological roles and mechanisms of action.

4.2.1 Functional Annotation of Heteromeric Glycine Receptors

Synaptic glycine receptors are composed of a heteromeric assembly of α and β subunits, which together form the functional receptor responsible for mediating inhibitory neurotransmission in the central nervous system [11]. Recent advances in cryo-EM have allowed for the detailed structural elucidation of the full-

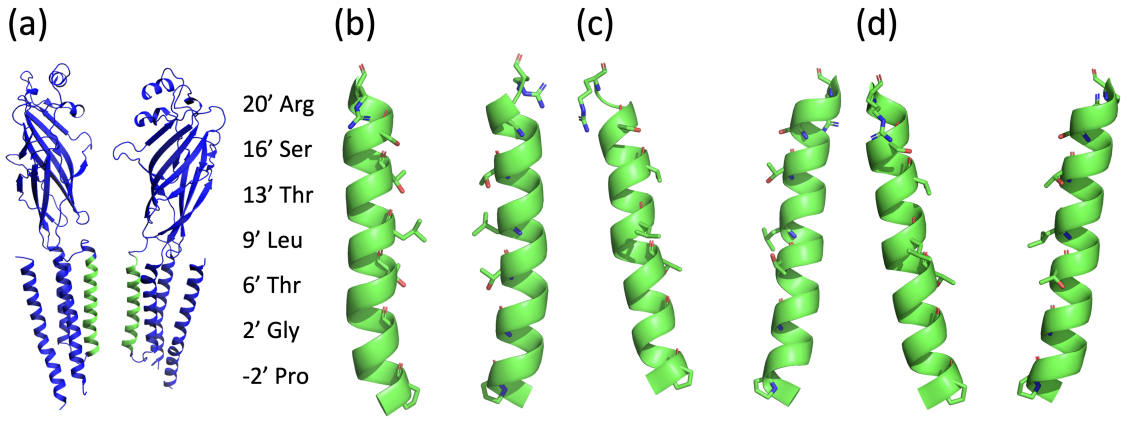


Figure 4.4: Two non-neighbouring α chains of GlyR with the TM4 region highlightes in green (a). TM4 region with pore-lining residues of αβGlyR-Stry with PDB code 7tu9 (b). TM4 region of αβGlyR-Gly with PDB code 7tvi (c) and TM4 region of αβGlyR-Gly-Ivm with PDB code 8fe1 (d).

length zebrafish α1βGlyR in various states, providing insights into how these structures behave under different conditions. Specifically, structures have been resolved in the presence of the antagonist strychnine, the agonist glycine, and a combination of glycine with the positive allosteric modulator ivermectin [11]. Each of these cryo-EM structures reveals a distinct conformation of the pore, with varying degrees of asymmetry, highlighting the dynamic nature of the receptor in response to different ligands. The heteromeric assembly of the GlyR, composed of different subunits, motivates an investigation into the role of pore asymmetry and how it influences ion conductance in the next chapter. In this section, the functional states of the receptor in its various forms — αβGlyR-Stry (antagonist-bound), αβGlyR-Gly (agonist-bound), and αβGlyR-Gly-Ivm (agonist with positive allosteric modulator) — were assessed using MD simulations. Pore lining residues of the TM4 region of the three structures are depicted in Fig. 4.4.

The MD simulations were performed with each receptor conformation embedded within a hydrated lipid bilayer composed of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine as described in section 3.2. Pore water free energy profiles were computed for each conformation using the Channel Annotation Package [40]. These computations were based on three replicates of 50 ns equilibrium simulations, carried out at physiological salt concentration (150 mM NaCl). The equilibration protocols are described in more detail in section 3.3. To maintain the structural integrity of each cryo-EM derived conformation during the simulations, harmonic restraints were applied to the protein backbone atoms with a force constant of 500 kJ/mol/nm^2 and side chains were flexible. Simulation trajectories were analysed at 100 ps intervals, with a bandwidth of 0.14 nm applied for estimating water density within the pore. To assess the chloride ion conduction properties of the receptor, additional simulations were conducted for each backbone-restrained receptor structure. These simulations extended to 200 ns and involved the application of a +500 mV transmembrane potential difference, with the positive potential on the cytoplasmic side. This potential was imposed by applying an external, uniform electric field in the membrane's normal direction, a technique described in section 3.8. During the 50 ns of simulation runs with backbone restraints, all three pore conformations remained stable with no major changes to the overall pore radii besides fluctuations arising from side-chain movements of residues lining the pore (Fig. 4.5). pLGIC pores do not have well-defined ion coordination sites but rather pass solvated ions. As such, the pore hydrophobicity, as well as the pore radius, are determinants of channel conductivity [43, 40].

The results from these simulations are plotted in Fig. 4.5 and reveal that the αβGlyR-Stry conformation exhibits clear channel dewetting near the leucine residue at the 9' position, with occasional dewetting

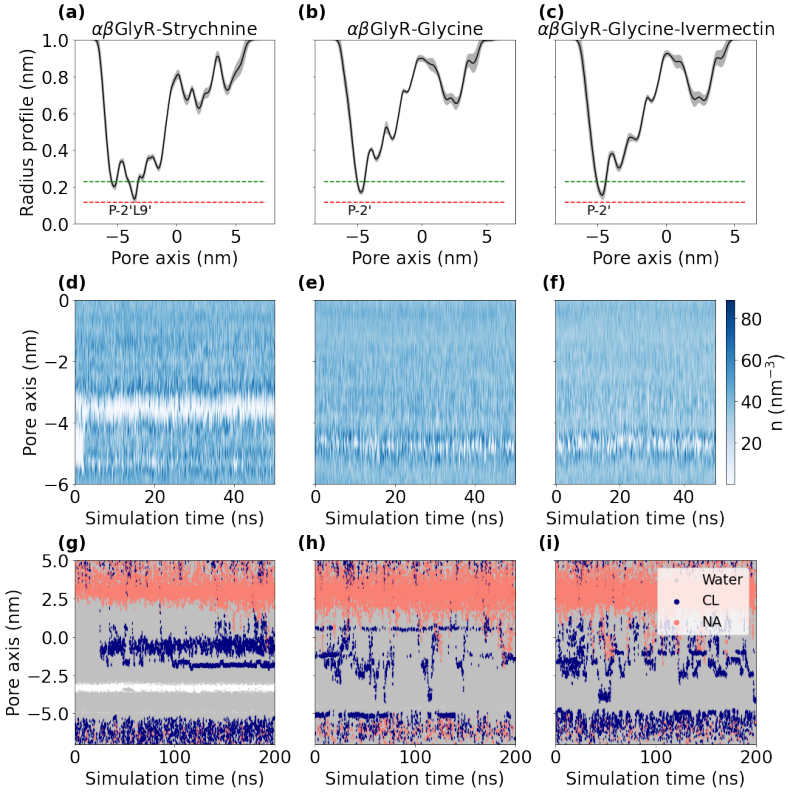


Figure 4.5: Mean pore radius profile and standard deviations averaged across 50 ns equilibrium simulations for (a) αβGlyR-Stry, PDB code 7tu9, (b) αβGlyR-Gly, PDB code 7tvi, and (c) αβGlyR-Gly-Ivm, PDB code 8fe1, along the central pore axis. The one-standard-deviation range is shown as a grey band, and the thick line corresponds to the mean of a 50 ns simulation. Water density within the channel pore (d)-(f). Each discrete water molecule position is associated with a Gaussian distribution, and the sum of all Gaussians yields the density function of water along the channel center line. Bulk water has a density of 33 nm^{-3}. The time series of water density was calculated for 50 ns equilibrium simulations. Trajectories along the pore axis of water molecules and ion coordinates within 10 Å of the channel axis inside the pore, in the presence of a +500 mV transmembrane potential difference (i.e., with the cytoplasmic side having a positive potential) in (g)-(i). No chloride permeation was observed past L9' or P-2' for αβGlyR-Stry, consistent with a closed channel state. No chloride permeation was observed past P-2' for αβGlyR-Gly and αβGlyR-Gly-Ivm, consistent with a desensitized channel state. During all simulations, positional restraints were placed on the protein backbone, in order to preserve the experimental conformational state while permitting rotameric flexibility in amino acid side chains.

observed at the αP-2' site. In contrast, the αβGlyR-Gly conformation showed a hydrated pore at the leucine residue at the 9' position but partial dewetting near the αP-2' (or βA-2') site. The αβGlyR-Gly-Ivm conformation, which involved the positive allosteric modulator ivermectin, demonstrated more pronounced dewetting at the αP-2' (or βA-2') site, which correlates with the smaller channel radius observed in this state. The pore lining residues of the α-subunit are depicted in Fig. 4.4. Pore lining residues are not well conserved between αGlyR and βGlyR and are more hydrophobic in βGlyR [37]. In particular, the 6' position is a threonine in αGlyR but leucine in βGlyR [11]. The desensitization gate is asymmetric at the P-2' position, with the αA interface radially outward from the pore center and slightly lower than the other −2' residues [11]. The radius profile will be studied in more detail in Chapter 5 with asymmetric (ellipsoidal) probe particles. These findings suggest that the receptor is likely in a non-conducting, desensitized state when bound to the agonist and positive allosteric modulator. The insights gained from this analysis will serve as a foundation for further investigations in the next chapters. In Chapter 5, the focus will be on understanding the relationship between pore asymmetry and ion conductance using GlyR as a model system. In Chapter 7,

conductive conformations of GlyR are studied.

4.2.2 Functional Annotation of Prokaryotic Ligand-Gated Ion Channel GLIC Structures

Bharambe *et al*. resolved structures of the GLIC receptor at two different pH levels: pH 7.5 and pH 4.0. The structural differences in the two structures provided an opportunity to explore how pH influences the functional states of GLIC. To define these functional states, MD simulations were performed on the GLIC structures obtained at pH 7.5 ($GLIC_{pH7.5}$) and on the structure derived from the pH 4.0 condition ($GLIC_{pH4.0}$). Each of these structures was embedded within a lipid bilayer composed of dioleoylphosphatidylethanolamine (DOPE), palmitoyloleoylphosphatidylserine (POPS), and palmitoyloleoylphosphatidylcholine (POPC) in a 2:1:1 ratio. The bilayer was positioned within simulation cells of dimensions 11.6×11.6×15.8 nm^3, following the system setup protocol detailed in section 3.2. In line with previous studies, several acidic residues, such as Aspartic Acid (Asp), Glutamic Acid (Glu) and Histidine (His), in the $GLIC_{pH4.0}$ structure were protonated to simulate the low pH environment [209, 210]: Glu26, Glu35, Glu67, Glu75, Glu82, Asp86, Asp88, Glu177, and Glu243. Additionally, the histidine residue His277 was doubly protonated to accurately reflect the acidic conditions. To ensure that the structural integrity of each cryo-EM derived conformation was maintained during the simulations, harmonic restraints were applied to the protein backbone atoms with a force constant of 1000 kJ/mol/nm^2. These MD simulations were run for 50 ns as described in section 3.3. For the $GLIC_{pH4.0}$ structure, the initial equilibration phase was extended using a specific simulation protocol designed to support the relaxation of stable open states in pLGICs. This protocol, described in detail in section 3.3.2, involved an additional restrained simulation step with pore restraints that lasted for 200 ns. The final frame from this extended equilibration was then used as the starting configuration for three independent simulation repeats, each initiated with different velocities. The pore water-free energy profiles for these alternative protein conformations were computed using the Channel Annotation Package [40], with each profile based on 200 ns equilibrium simulations conducted at a physiological salt concentration of 150 mM NaCl. Simulation trajectories were analysed at 100 ps intervals, with a bandwidth of 0.14 nm applied for estimating water density within the pore. Furthermore, ion conduction was assessed in 200 ns simulations for each backbone-restrained receptor structure. This was done at 150 mM NaCl concentration under the influence of a +500 mV and -500 mV transmembrane potential difference, with the positive potential applied on the cytoplasmic and extracellular sides, respectively. The transmembrane potential was generated by imposing an external, uniform electric field in the membrane's normal direction, as described in section 3.8.

The results depicted in Fig. 4.6 of these simulations revealed stable pore radii in both GLIC structures during the 200 ns equilibration period, with no major changes observed. For $GLIC_{pH7.5}$, the simulations indicated that the pore at the 9' position (with an isoleucine) remained dewetted throughout the entire simulation. This dewetted state was corroborated by the presence of an energetic barrier for water molecules at this region, as evidenced by the free energy profiles. Conversely, the simulations of $GLIC_{pH4.0}$ showed no signs of an energetic barrier for water at the 9' position, indicating a hydrated pore in this region. The pore lining residues are depicted in 1.4(b). When MD simulations were conducted under an external potential, the $GLIC_{pH7.5}$ structure continued to exhibit a dewetted environment around the isoleucine residue at the 9' position, with no Na^+ ion permeation events observed, suggesting that $GLIC_{pH7.5}$ represents a non-conductive state. In contrast, several Na^+ ion permeation events were observed during the same simulations with $GLIC_{pH4.0}$, indicating that this structure corresponds to an open, conductive state. In summary, the combination of cryo-EM structures and MD simulations has provided clear evidence that $GLIC_{pH7.5}$ is in a non-conductive state, while $GLIC_{pH4.0}$ represents an open, conductive state.

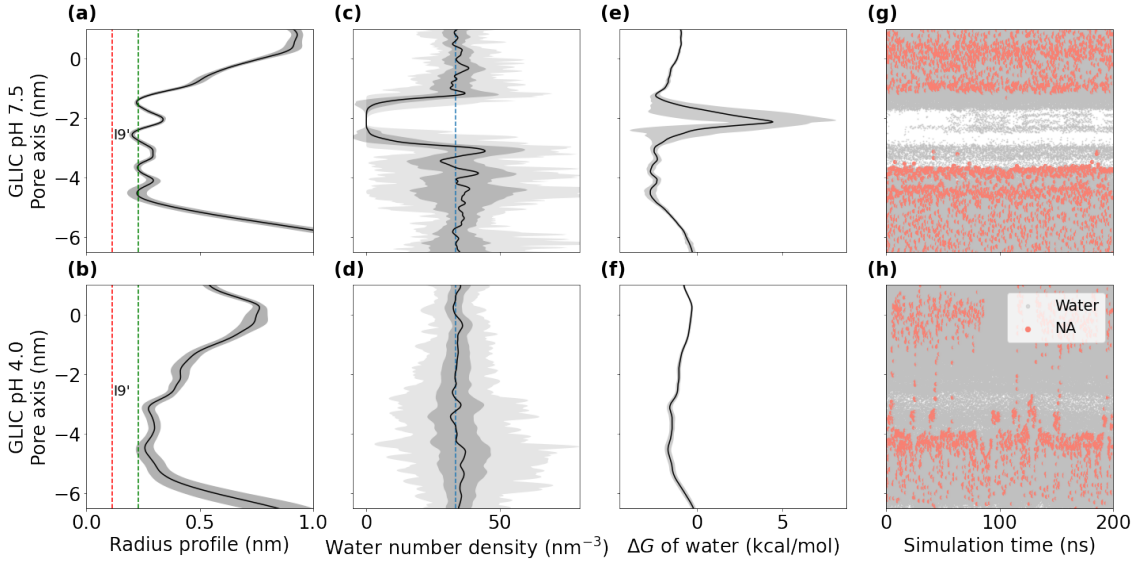


Figure 4.6: The mean pore radius of GLICpH7.5 (a) and GLICpH4.0-O (b) along with the corresponding standard deviation in grey (simulation average and standard deviation from one out of three independent equilibrium simulation runs of 200 ns with a sampling interval of 0.1 ns). The red and green dashed lines represent the diameter of one or two water molecule respectively. The mean water number density profile along the channel axis is plotted in (c) and (d) with one standard deviation coloured in dark grey and minimum and maximum in light grey. The dashed blue line corresponds to the bulk water density. Free-energy profiles of a water molecule along the central axis are shown for the corresponding structures with standard deviation in gray in (e) and (f). Time series of the water molecules (grey) and Na+ ions (salmon) are plotted for a simulation time of 200 ns for each corresponding structure in (g) and (h). White stretches reflect regions lacking water.

4.3 Conclusion

In sections 4.2.1 and 4.2.2, newly solved GlyR and GLIC structures were functionally annotated following the best practices presented in Fig. 4.3. After analysing an initial pore radius profile, equilibrium MD simulations of the ion channel embedded in a lipid bilayer were used to assess whether the pore is dewetted. The water number density averaged over short MD trajectories can give an initial estimate of energetic barriers to ion channel permeation. Simulations with an external potential can be used to obtain a crude conductance estimate. I will further characterise the pore profiles of heteromeric GlyR structures studied in section 4.2.1 in the next chapter to quantify the asymmetry of the pore. In Chapter 7, conductive GlyR structures with a wider pore are studied with polarisable and non-polarisable force fields.

Chapter 5

Exploring the Influence of Pore Shape on Conductance and Permeation

Disclaimer
This chapter is based on my published work in the Biophysical Journal, Seiferth and Biggin (2024) [69], see also in the list of publications. All investigation, writing and figures in that paper were produced by myself with feedback from the co-author, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

There is an increasing number of ion channel structures with heteromeric subunit assemblies, such as synaptic α1βB Glycine and α4β2 Nicotinic receptors. These structures show inherent pore asymmetry, though its functional relevance remains unclear. Molecular dynamics simulations of symmetrical homomeric channels often result in thermal distortion, producing asymmetrical conformations. As discussed in section 4.1.3, a first step in functional annotation is typically to annotate ion channels based on minimal constrictions determined by radius-profile calculations using software like HOLE or CHAP and pore hydrophobicity assessments. However, these tools use spherical probe particles, which limits their ability to accurately capture pore asymmetry. In this chapter, I introduce an algorithm that uses ellipsoidal probe particles, offering a more comprehensive representation of pore geometry. My analysis shows that non-spherical ellipsoids provide a more accurate and easily interpretable depiction of conductance. To quantify the impact of pore asymmetry on conductance, I systematically study carbon nanotubes (CNTs) with varying degrees of asymmetry as model systems. The conductance through these channels revealed surprising effects not predicted by spherical probes. These findings have broad implications for the functional annotation of biological ion channels and the design of synthetic channel systems for applications like water filtration. Additionally, this more accurate pore characterization is used to refine a physical conductance model, offering a heuristic estimate for single channel conductance. The PoreAnalyser software is freely available as a pip-installable Python package and provided as a web service.

5.1 Introduction

Advances in structural biology, combined with the development of artificial intelligence-driven structure prediction algorithms, have greatly expanded our understanding of three-dimensional ion channel and nanopore structures in various conformational states. The number of available channel structures in the PDB is plotted in Fig. 1.3. Notably, AlphaFold [172] and cryo-EM have been instrumental in uncovering the details of these molecular structures. Despite these advances, functionally annotating these structures remains a critical challenge. To bridge this gap, several tools have been developed, including HOLE [134], Caver [211], Moleonline [212], and the Channel Annotation Package (CHAP) [40]. These tools facilitate the analysis of the physical dimensions and characteristics of the pores in ion channels. One of the pioneering programs for computing the pore radius of an ion channel is HOLE [134]. The HOLE program uses a Monte Carlo simulated annealing procedure to find the optimal path for a sphere with varying radius to traverse the channel. Conversely, CHAP combines calculations of pore radius, hydrophobicity, and water density to predict hydrophobic gates in ion channels [40]. Written in C++, CHAP leverages the trajectory analysis

framework of the MD simulation software *Gromacs*. Other tools, such as MOLEonline [212] and CAVER [211], are also used for detecting and characterizing channels, pores, and tunnels within molecules, and they do not use a probe-based algorithm for pathfinding. MOLEonline provides information about channel-lining residues, physicochemical properties, and identifies cavities using Voronoi diagrams and molecular surfaces. CAVER employs a grid-based approach and convex approximation to determine the lowest-cost centerline path between a starting point and the molecule's surface [211]. In later versions, CAVER also incorporates Voronoi diagrams [213]. Despite not using a probe-based algorithm, the visualization of paths in CAVER or MOLEonline relies on probe particles. By using the maximum radii for each node along the path, the characteristics of the identified channels or pores can be effectively visualized and analysed. While some methods allow for detailed visualization of highly asymmetric surfaces, probe-based methods offer a more intuitive way to interpret the geometry using simple statistics like radii. The HOLE program has been widely used by structural biologists to quickly and reliably provide key information on pore geometry. Until recently, most transmembrane channel structures were highly symmetrical, typically due to homomeric subunits forming higher-order multimer structures combined with symmetrical averaging techniques in structural processing. Homomeric channels are not perfectly symmetric in a continuous rotational sense; they are only symmetric about an n-fold rotational axis. However, as structural biology techniques advance, more structures displaying inherent asymmetry are being identified, and understanding their potential role and relationship to function is gaining attention. Minniberger *et al*. explored the asymmetry and ion selectivity properties of bacterial channel NaK mutants derived from ionotropic glutamate receptors [214]. They observed structural asymmetry in the upper part of the selectivity filter, proposing that local asymmetrical conformational changes could efficiently alter ion channel function without requiring large-scale conformational transitions. Similarly, Roy *et al*. studied a nonselective sodium-potassium ion channel with structural asymmetry in the selectivity filter, suggesting that different degrees of conformational plasticity are required for conducting potassium and sodium [215]. Zhang *et al*. proposed a model for asymmetric activation of the 5-HT_3A serotonin receptor [216], while Shi *et al*. highlighted the permeation mechanisms of sodium and potassium in NaK channels [217]. Lewis *et al*. observed structural changes between sodium and potassium-bound states, accompanied by increased structural heterogeneity with Na^+. Chugunov *et al*. reported asymmetric gating at the selectivity filter in molecular dynamics simulations of the TRPV1 pore upper gate [218]. These studies collectively emphasize that conformational plasticity and structural asymmetry in the selectivity filter play crucial roles in regulating ion selectivity and conductance. Moreover, local asymmetrical conformational changes emerge as an energetically efficient means to modulate ion channel function without requiring global conformational transitions across all subunits. The presence of asymmetry in ion channel pores raises the question of whether using a spherical probe is appropriate. This issue was anticipated years ago, and the HOLE software includes a CAPSULE option to measure pore anisotropy with a spherocylinder [134]. However, an asymmetrical pore can be more accurately described with an ellipsoidal probe. Along the channel axis, two radii are reported, corresponding to the longer and shorter axes of the ellipsoid. The volume of the pore characterised with ellipsoidal probe particles will be larger than that found using spherical probes. Building on the well-established HOLE program and its MDAnalysis Python API [219], I have developed a program called PoreAnalyser. This tool, available as a pip-installable Python package (`https://github.com/DSeiferth/PoreAnalyser`) and as an interactive web service (`https://poreanalyser.bioch.ox.ac.uk`), allows users to capture asymmetry using an ellipsoidal probe. To demonstrate its utility, an assessment of the influence of asymmetry on conductance predictions is presented. Even with an ellipsoidal probe particle, the asymmetry of a channel cannot always be fully accounted for, as the ellipsoidal shape is often more symmetric than the channel structure. However, reporting two radii offers a good compromise, providing an easily interpretable representation of

the pore geometry. To systematically quantify the implications of pore asymmetry, CNTs are investigated as model systems, examining conductance through channels with varying degrees of pore asymmetry. This investigation builds upon the cylindrical approximation model, a physical model used to predict channel conductance. Hille characterised homogeneous conducting materials by a bulk property known as resistivity ρ_{bulk}, which is the inverse of conductivity κ_{bulk} [3]. The conductance model treats the channel as a cylinder and considers both the pore resistance R_{pore}, associated with the length of the pore, and the access resistance at the pore entrance (R_{access}), as shown in the equation below:

$$R = R_{pore} + R_{access} = \rho_{bulk}\frac{L}{A} + R_{access}, \tag{5.1}$$

where A denotes the cross-sectional area and L the length of the cylinder. Subsequent studies utilised Hille's cylindrical model but neglected access resistance, approximating an ion channel as a series of cylinders with different radii as determined by HOLE analysis [220]. Mejri *et al*. extended this work using CNTs, revealing that increasing the CNT diameter enhanced conductance values, while increasing tube length decreased conductance, suggesting a series-like action of access resistance and pore resistance [221]. Manghi *et al*. further investigated the relative contributions of channel and access resistances in experimentally measured currents [222]. CNTs are well-established model systems for ion channels in molecular dynamics studies due to their controllable dimensions, making them invaluable for investigating ion transport phenomena [223, 224]. CNTs have been explored in the context of dewetting processes within confined hydrophobic regions of channels. As discussed in chapter 4, these regions may undergo transitions between wet and dry states to gate the pore closed without physical constriction, a phenomenon known as hydrophobic gating [19]. Beckstein *et al*. proposed a hydrophobic gating mechanism for nanopores, emphasizing the role of geometry, diameter, and the hydrophobicity of pore lining residues [17]. In this chapter, the focus is on pores wide enough to rule out dewetting transitions, and only conductive CNTs with radii larger than 4 Å are considered, ensuring consistently wetted pores. Mendonça *et al*. examined water diffusion within deformed CNTs with varying degrees of eccentricity, providing insights into water molecule behaviour within these nanoscale conduits and emphasizing water freezing under certain conditions [225, 226]. Rezaee *et al*. studied monoatomic fluid flow through elliptical CNTs, systematically examining the influence of tube size, temperature, and pressure gradient on fluid dynamics [227]. Balme *et al*. explored ionic transport through hydrophobic nanopores, observing modifications in ion solvation structure with nanopore diameters less than 2 nm and varying NaCl concentrations [228]. Furthermore, Mejri *et al*. investigated on the role of water models (SPC/E, TIP3P, TIP4P/2005) in studying ionic conductance using MD simulations [221]. They found that while the choice of water model did not strongly impact conductance trends with increasing tube diameter, it did influence conductance values. Hydrogen bond analyses revealed similar water arrangements within the pore for both the SPC/E and TIP4P/2005 models, highlighting the robustness of certain water models in capturing essential aspects of ion transport through CNTs. Mejri *et al*. also studied the impact of different chemical functionalisation and geometric parameters, such as radius and chirality, on conductance through MD simulations with an external potential [229]. Together, these studies contribute to understanding the complex interplay between CNT characteristics, water behaviour, and ionic conductance. Therefore, I selected CNTs as model systems because they allow us to easily control the extent of pore asymmetry. Building on the cylindrical approximation, an improved physical model to predict conductance is presented. These results have broad implications for the functional annotation of biological ion channels and the design of synthetic channel systems for applications such as water filtration. I anticipate that the PoreAnalyser software package will be more effective in accounting for the asymmetrical properties of ion channel structures.

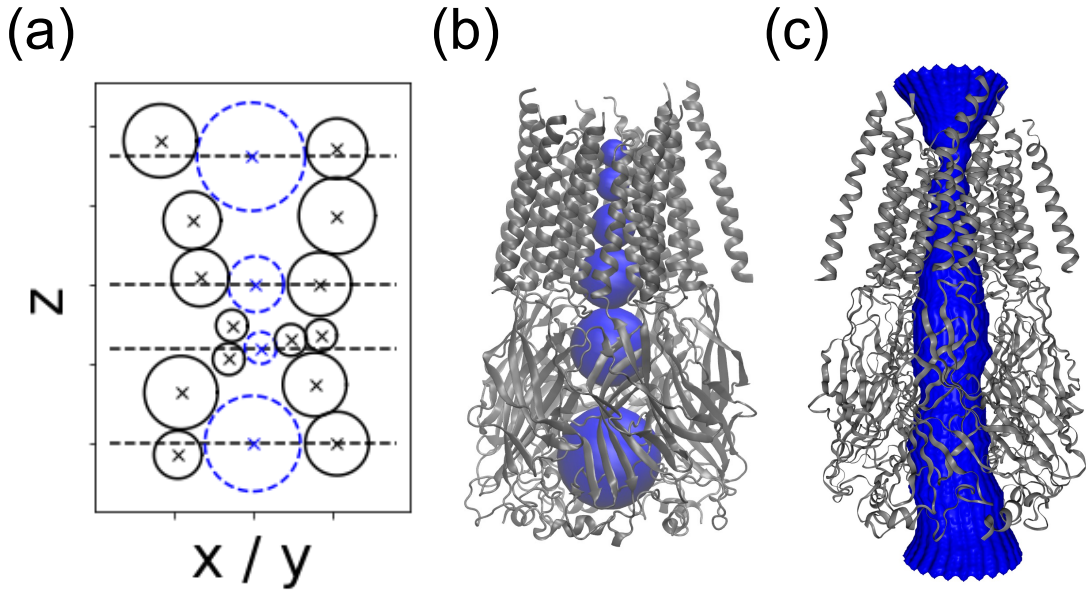


Figure 5.1: Pore geometry calculations with spherical probe particles used in HOLE [134] or CHAP [40]. (a) The position of a spherical probe (blue and dashed) is iteratively adjusted in parallel planes to achieve the maximum radius while ensuring that there is no overlap with the van der Waals sphere of any atom (black circles) within the ion channel. (b) Positions of a few probe particles (blue) along the pathway of the pore within the α1βB Glycine receptor (PDB 7tvi) [11]. (c) Surface representation of the pore of the α1βB Glycine receptor.

5.2 Methods

5.2.1 Pore Geometry Calculations

In the PoreAnalyser workflow, the principal axis of the input structure is aligned to the z-axis by default, although users can modify this setting according to their preferences. Users of the web interface can simply upload their own coordinates. To calculate the pore geometry, probe-based pathway finding with spherical probe particles is used as a first step, then each sphere is transformed into an ellipsoid. Initially, HOLE is executed using a spherical probe particle, with the COM of the input structure as the default starting point. The probe's position is then optimized in subsequent parallel x-y planes to maximize its radius without overlapping the van der Waals sphere of any ion channel atom. Fig. 5.1 provides a schematic representation of van der Waals spheres, the positions of their centers in the xz-plane of a channel aligned to the z-axis, and a surface representation of the pore. HOLE is integrated into the Python workflow using its MDAnalysis Python API [219]. To incorporate an ellipsoidal probe particle, the spherical probe particles are iteratively transformed into ellipsoids. To initiate the calculation, the output file from HOLE with the positions and radii of the probes is loaded. Then, a loop is performed through all the spherical probe particles and an ellipsoid is initialized using the parameters obtained from the HOLE output. The parameters of the ellipsoid include the position of its center $[x, y, z]$, orientation θ, and radii $[a, b]$, where $a \geq b$. The smaller radius b remains constant along the z-coordinate. To optimize the parameters, a Nelder-Mead (or downhill simplex) four-dimensional optimization algorithm [117] is employed, first using smaller bounds for the parameters $[x, y, a, \theta]$, and then with larger boundaries to further increase the ellipsoid (Fig. 5.2). See section 2.2.2 for a description of the downhill simplex algorithm. Finally, the volumes of the pores based on the use of spherical and ellipsoidal probe particles are compared. This comparison provides insights into the impact of

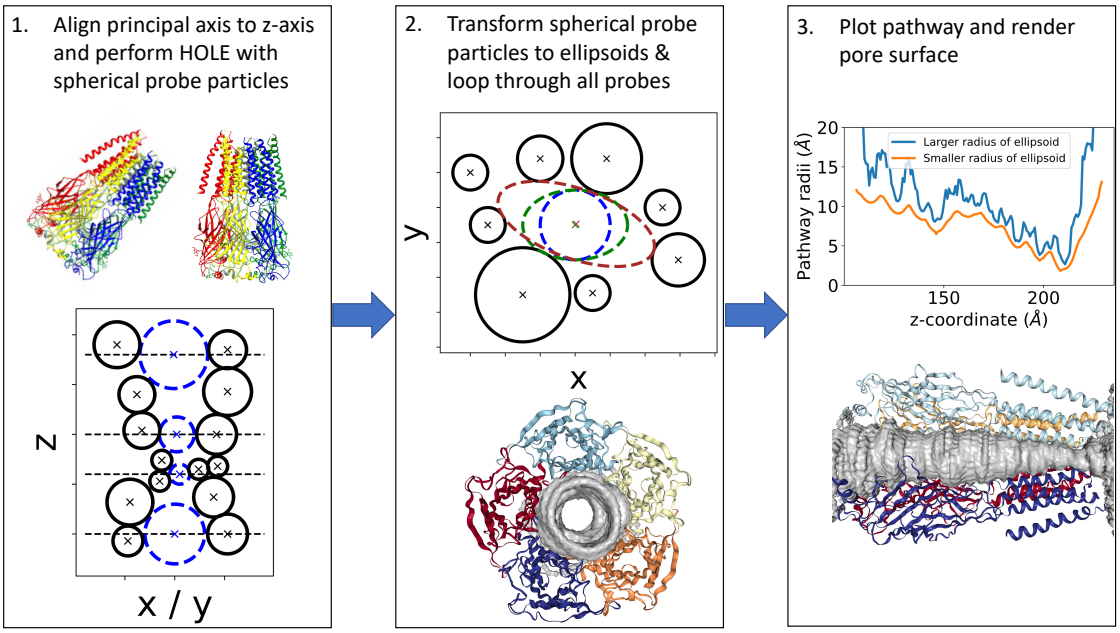


Figure 5.2: Workflow diagram of the PoreAnalyser software. The input protein is aligned to the z-axis and then a HOLE analysis is performed (1). The HOLE spheres are then iteratively transformed into ellipsoid (2). As a final step, the pathway is plotted in terms of both the larger and smaller ellipsoid radii and the pore surface visualized (3).

probe shape on pore geometry calculations.

5.2.2 Model Systems

A pristine armchair CNT capped with hydrogen atoms is used as a model system, featuring a length of 50 Å, a dimension that enables them to span a lipid bilayer completely. The key characteristic of the CNTs lies in the variation of their radius, carefully selected to be wide enough to prevent dewetting of the pore, ensuring stable hydration throughout the simulation duration. All CNT atoms have zero partial charge. The simulation setup involves the insertion of the CNT into a bilayer (see Fig. 5.3) composed of 128 POPC lipids, facilitated by the InflateGro method [147] as described in section 3.2.2. The system is solvated with a 0.15 M KCl solution. The CNT is modelled using the OPLS all-atom force field [230], united-atom lipids, and the SPC/E water model [109]. The temperature is maintained at 300 K. To preserve orientation, the z-coordinates of the CNT are restrained, with harmonic restraints of 500 kJ/mol/nm^2. Introducing varying degrees of asymmetry, the area in the xy-plane $A_{xy} = \pi r^2 = \pi ab$ is manipulated, where r represents the radius of the CNT circular plane, a denotes the larger, and b the smaller radius of the ellipsoid. The asymmetry is quantified by the ratio of the smaller to the larger elliptic radius (1/2,3/4, and 1), corresponding to different degrees of eccentricity $e = \sqrt{1 - b^2/a^2}$, depicted in Fig. 5.3 (a)-(c). The cross-sectional area A_{xy} is preserved for varying degrees of eccentricity. To retain a non-circular shape, harmonic restraints are defined such that a major axis of $2a$ and a minor axis of $2b$ are enforced. Notably, the pores remain stably hydrated throughout the entirety of the simulations, providing a robust foundation for investigating the influence of varying degrees of asymmetry on the behaviour of CNTs within lipid bilayers. (8,8), (10,10), (12,12), (14,14), (16,16) and (18,18) pristine armchair CNTs with different degrees of eccentricity are studied, where the native (8,8) and the (14,14) have HOLE radii of 3.4 Å and 7.0 Å respectively (see Table 5.1). A (14,14) armchair CNT with length 50 Å is denoted as 14 x 50 in this article. The CNT type, denoted by the pair of integers (n, m), signifies that a specific CNT can be derived by rolling a graphene sheet characterised by lattice vectors $\vec{a}_1$

Armchair CNT of type (n,m)	(8,8)	(10,10)	(12,12)	(14,14)	(16,16)	(18,18)
HOLE radius (Å)	3.4	4.7	5.9	7.0	8.3	9.8

Table 5.1: Carbon nanotubes (CNTs) of length 50 Å and varying HOLE radii

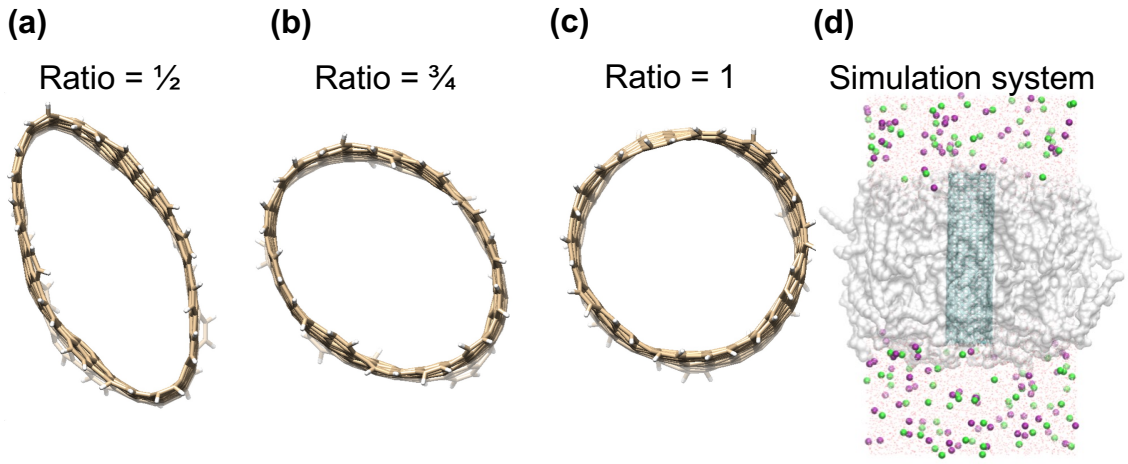


Figure 5.3: Cylindrical carbon nanotubes with varying degrees of asymmetry. View of xy-plane. (a) Elliptical xy-plane with ratio of ½ between smaller and larger radius. (b) elliptical xy-plane with ratio of ¾ between smaller and larger radius. (c) Circular xy-plane (smaller and larger radius of ellipse are identical). The area $A_{xy} = \pi ab$ in the xy-plane is conserved and virtual bonds are defined to maintain the degree of asymmetry. (d) Cartoon of the simulation box with the carbon nanotube in cyan, the lipid in grey translucent spheres, water in red and white wire and K^+ and Cl^- ions in purple and green respectively.

and $\vec{a}_2$. In this configuration, atoms align theoretically atop one another when positioned at a distance of $n\vec{a}_1 + m\vec{a}_2$ on the graphene sheet. Ion conduction was measured in 250 ns simulations at 150 mM KCl concentration and in the presence of a +500mV and -500mV transmembrane potential difference. This was applied by imposing an external, uniform electric field in the membrane normal direction as described in section 3.8. Umbrella sampling was performed to obtain one-dimensional single-ion potential of mean force profiles for potassium and chloride passing through the pore of different CNTs with varying degree of asymmetry as described in section 3.5. Unbiasing was performed with the weighted histogram analysis method (WHAM) using the *Gromacs* implementation [138] as described in section 2.2.4.

5.3 Results and Discussion

5.3.1 Case Studies with Biological Ion Channels

To illustrate how the geometrical profile of a pore is influenced by using an ellipsoidal probe, the pore geometry of the full-length zebrafish α1β GlyR is examined that was already studied in section 4.2.1. A cryo-EM structure of the full-length zebrafish α1β GlyR is analysed in the glycine-bound state (PDB code 7tvi). In Fig. 5.4, the local minima of the pathway radii obtained using a spherical probe particle, representing the smaller elliptical radius, coincided with the local minima obtained from the larger radius of the ellipsoid. Gibbs *et al*. found that the desensitization gate is asymmetric at the P-2' position, with the αA interface

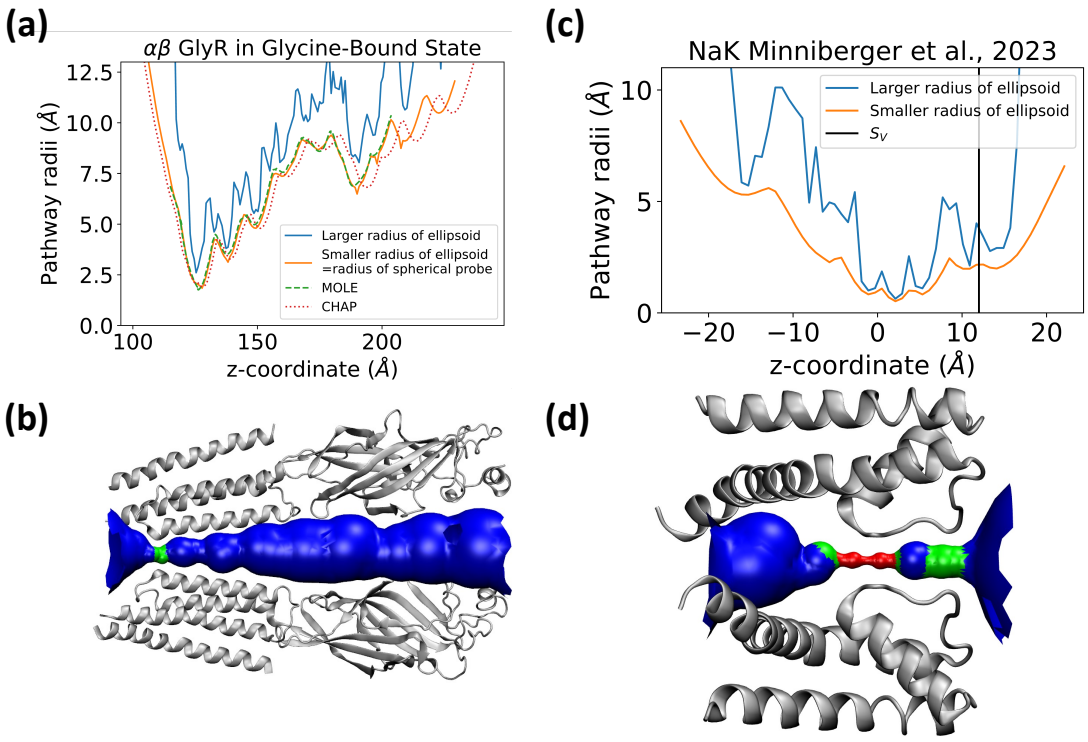


Figure 5.4: Case studies of the α1βB GlyR and NaK Channel. (a) Pathway analysis for a cryo-EM structure of full-length zebrafish α1βB GlyR with heteromeric subunit assembly in the presence of glycine [11] with the larger and smaller ellipsoid radii shown as blue and orange lines respectively. The radii profiles for MOLE [212] and CHAP [40] are plotted as green dashed and dotted pink lines respectively. (b) Cartoon showing the pore profile for α1βB (PDB code 7tvi). (c) Pathway profile of a of bacterial channel NaK mutant derived from ionotropic glutamate receptors (PDB code 7oph) by Minniberger *et al.* [214]. The black line labelled with S_V denotes the position of the selectivity filter. (d) cartoon showing the pore profile (pdb code 7oph).

radially outward from the pore center and slightly lower than the other −2' residues [11]. The radius profiles from MOLE [212] and CHAP [40] are also plotted in Fig. 5.4(a). These profiles match the profiles obtained using a spherical probe particle. CHAP uses a custom pore coordinate s instead of the z-coordinate, which explains the slight differences from the MOLE and HOLE profiles. The ratio of pore volume based on ellipsoidal and spherical probe particles is approximately 1.6. This indicates that using a spherical probe particle alone underestimates the actual pore volume. Incorporating an ellipsoidal probe particle provides a more accurate representation of the true pore volume, offering valuable insights into the size and dimensions of the GlyR pore. Minniberger *et al.* observed structural asymmetry in the selectivity filter of bacterial NaK channel mutants designed to mimic ionotropic glutamate receptors, which are non-selective tetrameric cation channels [214]. As depicted in Fig. 5.4(c), the radius profile reveals a high degree of asymmetry in the upper part of the selectivity filter, where the larger elliptical axis is twice as long as the smaller one. These findings highlight the importance of considering different probe shapes when calculating pore geometry and emphasize the limitations of relying solely on spherical probe particles for accurate measurements.

5.3.2 Comparison with HOLE Capsule Option

As mentioned earlier, Smart *et al*. introduced the CAPSULE option within the HOLE algorithm to allow the user to capture pore asymmetry [220], a feature that has yet to gain widespread adoption within the scientific community. Departing from the traditional approach of optimizing a single sphere, HOLE with the CAPSULE option extends its capability to a spherocylinder, essentially an extension of a sphere. The spherocylinder is uniquely characterised by two independently moving centres, described by three properties: its radius r and the positions of these two centres with distance L between them [220]. Unlike the original routine, which optimized the sphere's radius, the CAPSULE option focuses on maximizing the area of the capsule on a given plane. Subsequently, this area is converted into an effective radius $r_{\text{eff}} = \sqrt{r^2 + Lr/\pi}$, providing a parameter for comparisons. In Fig. 5.5, the performance of the HOLE CAPSULE algorithm is compared with the PoreAnalyser package, where ellipsoidal probe particles are employed to explore the intricacies of pore geometry. The smaller elliptical radius b of the PoreAnalyser algorithm and the spherical radius r of HOLE are identical by construction and are plotted in blue in Fig. 5.5. The larger elliptical radius a of the PoreAnalyser algorithm is consistently larger than the effective radius r_{eff} of the capsule both for an elliptical CNT 10 x 50 model system with ratio of elliptical radii $b/a = 1/2$ depicted in Fig. 5.5(a)-(b) and for a GlyR ion channel depicted in Fig. 5.5(c)-(d). The PoreAnalyser algorithm can accurately capture both axes of ellipsoidal probe particles, whereas the HOLE capsule option underestimates the cross-sectional area. The elliptical CNT model system has a volume of $V_{ana} = L\pi ab = 3100\,\text{Å}^3$, where $L = 50$ Ådenotes the length of the carbon nanotube and $2a$ and $2b$ denote the minor and major axis of the elliptical cross-sectional area. The volumes of the model system based on a PoreAnalyser and HOLE profile are $V_{PA} = 2500\,\text{Å}^3$ and $V_{HOLE} = 1900\,\text{Å}^3$ respectively. For the GlyR ion channel, the minima of larger and smaller radii of the PoreAnalyser algorithm coincide. When moving away from the narrowest constrictions, the differences between the major and the minor elliptical axis become more pronounced, possibly reflecting increased deviation away from spherical geometry in the the less restricted region. The pore volume $V_{PA} = 32400\,\text{Å}^3$ obtained from PoreAnalyser is larger than the one $V_{HOLE} = 21700\,\text{Å}^3$ obtained from HOLE capsule. For both systems, the PoreAnalyser algorithm describes the respective pore more accurately.

5.3.3 Carbon Nanotubes to Investigate Conductance

To assess the accuracy of the PoreAnalyser algorithm when using ellipsoidal probe particles, various Carbon Nanotube (CNT) model systems with varying degree of asymmetry were analysed. The ellipticity of each CNT is defined by two radii: the longer radius a and the shorter radius b. The ellipticity was enforced using harmonic restraints to maintain the elliptical shape with specified ratios of b/a (1, 3/4, and 1/2), as shown in Fig. 5.3 and described in section 5.2.2. First, I checked whether the PoreAnalyser software could accurately capture the radii maintained by the harmonic restraints. In Fig. 5.6, there are higher standard deviations for CNTs with larger elliptical radii, indicating greater variability when measuring the major axis. The pores identified using spherical probe particles effectively characterised the minor axis, providing a useful benchmark for comparison. 18 distinct CNTs were examined, each with six different areas in the xy-plane and three varying ratios between the longer and shorter radii (1, 3/4, 1/2). The HOLE radii of the circular CNTs (with a radius ratio of one) are listed in Table 5.1. As expected, the HOLE radius, representing the shorter radius, aligned well with the minor elliptical axis. However, the larger radius was less accurately maintained, as might be anticipated from using harmonic restraints. Nonetheless, this level of accuracy was considered acceptable. Having checked whether the model systems are behaving as expected, these CNTs with varying degrees of pore asymmetry are used as model systems to investigate how pore asymmetry affects ionic conductance. Smart *et al*. predict the conductance g of a channel based on the pore geometry

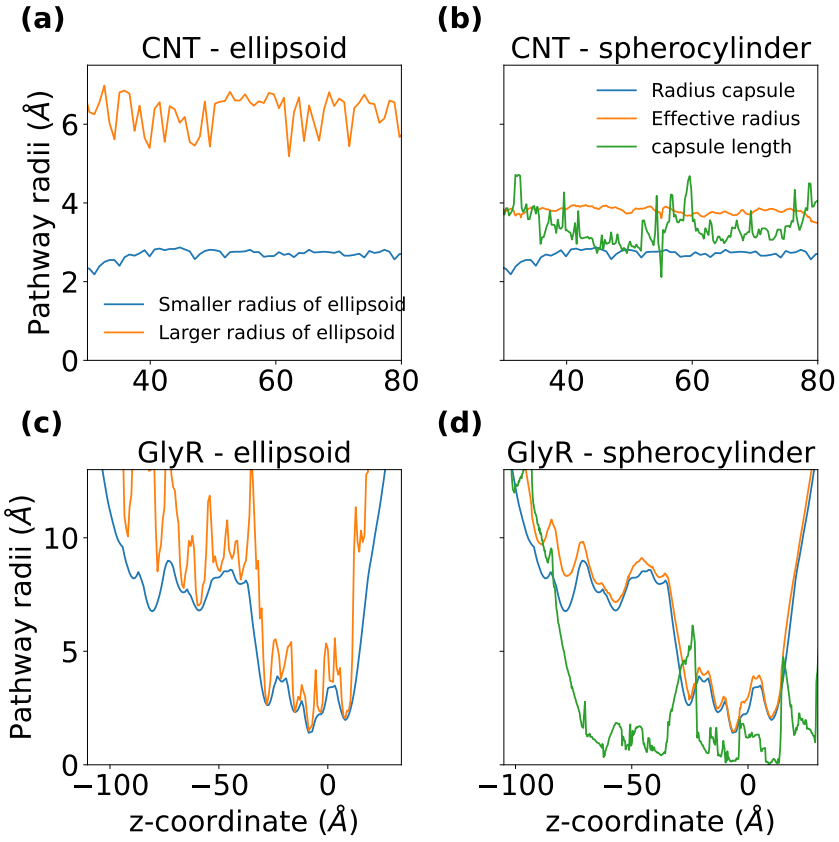


Figure 5.5: Comparison of pathways found with elliptic probe particles of PoreAnalyser with the CAPSULE option from HOLE. A carbon nanotube profile of a 10 x 50 system with ratio 1/2 obtained from PoreAnalyser (a) and HOLE CAPSULE option (b). Profiles obtained from PoreAnalyser (c) and HOLE CAPSULE (d) for the Glycine receptor (PDB code 7TU9).

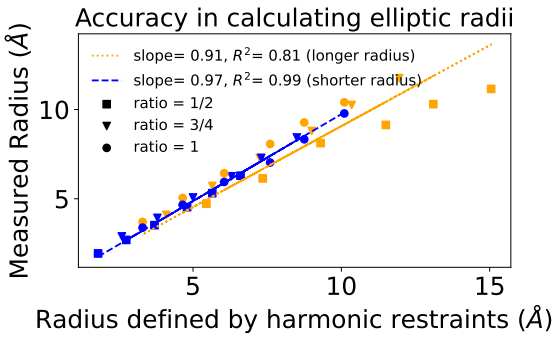


Figure 5.6: Evaluating the accuracy in calculating elliptic radii in 18 different carbon nanotubes (see Table 5.1) with varying degree of asymmetry. Points are represent as squares, triangles and circles for elliptical ratios ½, ¾ and 1 respectively. It can be seen that the longer radius is less well maintained than the shorter radius, particularly at radii above 9 Å.

found by HOLE as

$$R_{HOLE} = \frac{1}{g} = \rho_{bulk} \sum_{z} \frac{dz}{A_{xy}(z)}, \tag{5.2}$$

where the resistivity ρ_{bulk} of the permeant ions is regarded as a function of the molar ion concentration c [220]. The pore is approximated as a sequence of cylinders with area $A_{xy}(z)$ in the xy-plane and length dz. This model predicts higher resistance R (and lower conductance g) for longer and narrower pores. However, the model would predict the same conductance for a channel with circular $A_{xy} = \pi r^2$ and for a channel with elliptical area $A_{xy} = \pi ab$ if the cross-sectional area A_{xy} is the same. The conductance and the barrier to permeation was measured for potassium and chloride in four different CNT systems with varying degree of asymmetry. As one would expect, a higher barrier to permeation corresponds to a lower conductance (Fig. 5.7). Here, the conductance for potassium is lower than for chloride and consistent with this trend, the energetic barrier is higher for potassium than for chloride. Four different CNT systems were studied with increasing cross-sectional area (CNT 8 x 50 has the lowest cross-sectional area and CNT 14 x 50 the highest), and the conductance increases with increasing cross-sectional area A_{xy} in line with eq. (5.2). When the ratio of the elliptical axes is varied while preserving the cross-sectional area A_{xy}, one finds that a system with ratio 1, corresponding to a spherical cross-sectional area, has consistently higher conductance values than the same system with ratio 1/2, corresponding to major axis twice as long as the minor axis. The conductance of chloride with a high standard deviation in the CNT 14 x 50 is an outlier of this trend, see Fig. 5.7(b). The energetic barrier decreases for less elliptical shapes. Histogram overlaps are plotted in Fig. A.2 and convergence is assessed in Fig. A.3 and Fig. A.4 in the appendix. Neither potassium nor chloride conductance events were observed in the repeats of 250 ns in the narrowest CNT system (CNT 8 x 50). The energetic barrier for potassium and chloride to permeate the pore increases with higher eccentricity of the cross-sectional area. The differences are more pronounced for narrower CNTs with smaller cross-sectional areas (Fig. 5.7). The maximum PMF barrier is plotted in Fig. 5.7(c), (d) and the PMF profile along the z-axis is plotted in Fig. 5.8 (a) and (b). Note that one only observes conduction when the PMF barrier is below 5 kcal/mol. The PMFs were calculated and the conductance were estimated also for the ion parameters obtained from Li *et al.* [113]. The barrier for ion permeation and the ion conductance are depicted in Fig. A.1. Chapter 6 illustrates that ion parameters can influence free energy barriers strongly [71] but when comparing Fig. 5.7 and Fig. A.1, one can see a similar trend for energetic barriers and conductance. As described in section 3.6, one can obtain a measure of the hydration shell along the reaction coordinate, by simply reporting the mean and standard deviation of how many water molecules are within a certain distance of the respective ion. In Fig. 5.8, the hydration shell of potassium is distorted in the CNT with ratio 1/2 between minor and major axis of the elliptical cross-sectional area. The RDF for potassium, Fig. 5.8(e), and chloride, Fig. 5.8(f), in bulk and in the narrow CNT system reveals that for potassium, the second peak, corresponding to the second hydration shell, is reduced compared to bulk. Potassium inside a CNT with ratio 1/2 between minor and major axis has a smaller RDF compared to potassium inside a spherical CNT (ratio 1).

5.3.4 Physical Model to Predict Conductance

In the development of a physical model to predict conductance through ion channels, Hille's approach [3] initially considered a cylindrical approximation with length L and cross-sectional area A, allowing the resistance R to be expressed as:

$$R = \frac{1}{g} = \frac{\rho L}{A} \tag{5.3}$$

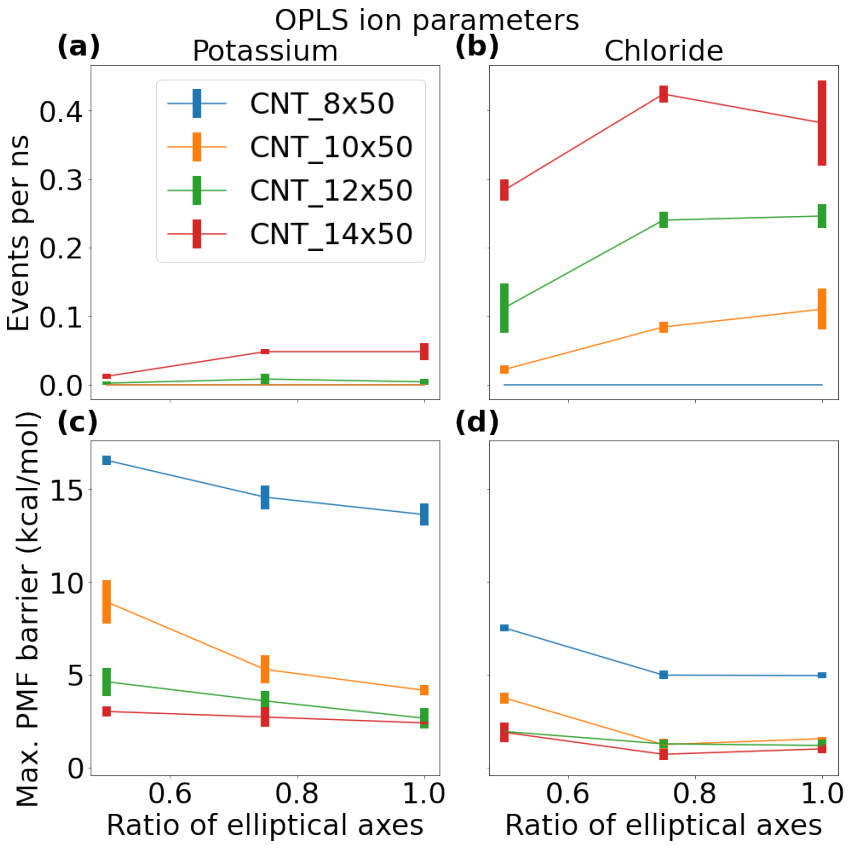


Figure 5.7: Barrier for ion permeation and ion conductance for carbon nanotubes (CNTs) as a function of the ratio between of the elliptic radii. The conductance is measured as events per ns observed in multiple 250 ns repeats with external potential of 500 mV for (a) potassium ions and (b) chloride ions. The energetic barrier is the maximum of the Potential of Mean Force (PMF) obtained from umbrella sampling for (c) potassium ions and (d) chloride ions. The different CNT systems 8x50, 10x50, 12x50 and 14x50 (see Table 5.1 for corresponding HOLE radii) are all 50 Å in length and shown as blue, orange, green and red lines respectively. Histogram overlaps are plotted in Fig. A.2 and convergence is assessed in Fig. A.3 and Fig. A.4 in the appendix.

where ρ represents bulk resistivity. This simplistic model was subsequently refined [134, 231] suggesting a more accurate representation of ion channels as a series of stacked cylinders, where the resistance accumulates. Considering ohmic principles and utilizing the HOLE software, which measures cross-sectional areas $A(z)$ along the channel axis z, the refined resistance model becomes:

$$R_{HOLE} = \rho_{bulk} \sum_{z} \frac{z_i - z_{i-1}}{\pi r_i^2} \tag{5.4}$$

where the resistivity ρ is approximated with bulk property ρ_{bulk} and r_i denotes the radius of a spherical probe particle. However, relying on bulk property resistivity $\rho = 1 \,/\, \kappa$ becomes problematic, as the conductivity κ depends on the diffusion coefficients of ions. The bulk conductivity κ_{bulk} of a KCl solution with concentration c is defined as:

$$\kappa_{bulk} = q_e^2 c \frac{D_K + D_{CL}}{k_B T} \tag{5.5}$$

where q_e is the elementary charge, D_K and D_{CL} are diffusion coefficients of potassium and chloride ions, k_B is the Boltzmann constant, and T is temperature [228]. To refine the model for ion channel conductance further, a conductivity model is introduced, expressing the conductivity $\kappa(a_i, b_i)$ as a function of the radii a_i and b_i of ellipsoidal probe particles. For larger radii, the ion movement is relatively unconstrained,

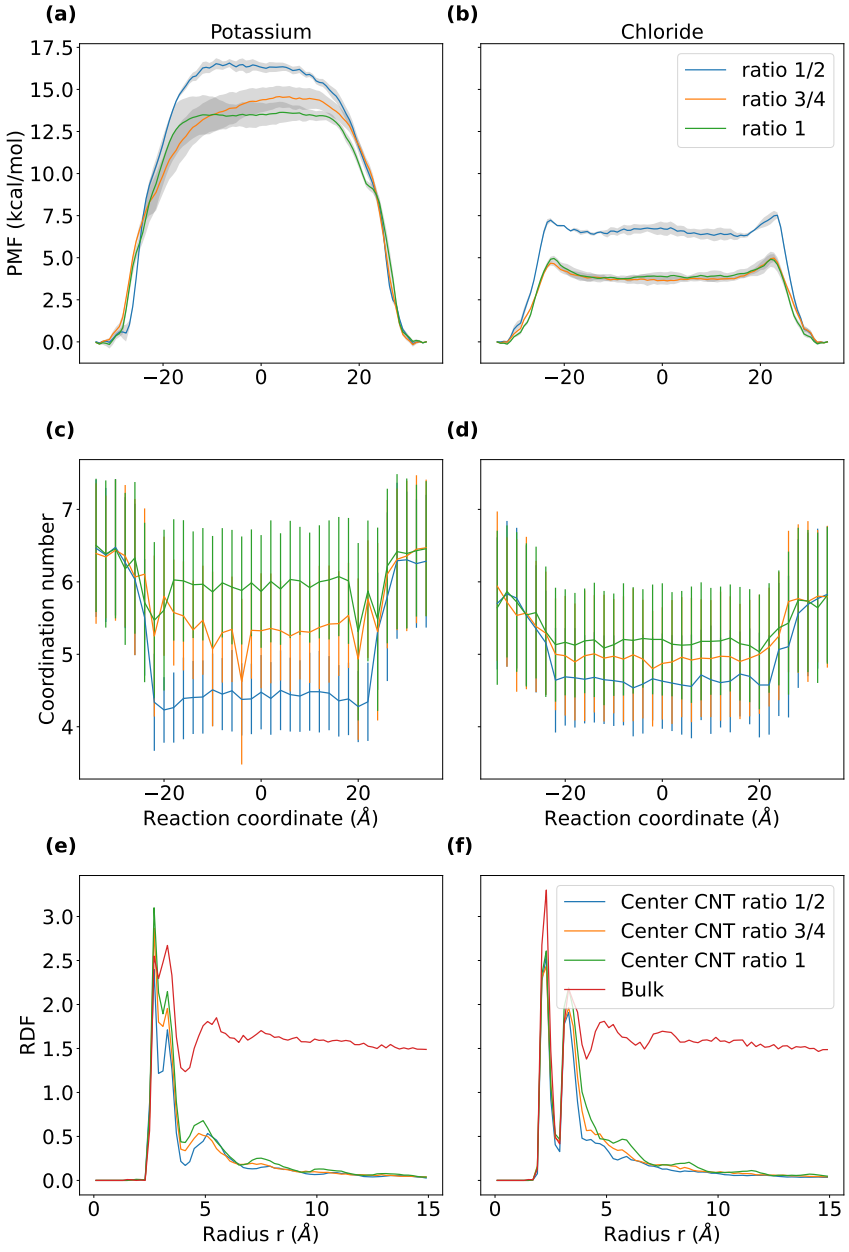


Figure 5.8: Potential of mean force (PMF) and ion hydration shell analysis. PMFs for Potassium (a) and chloride (b) to permeate 8 x 50 carbon nanotubes (CNTs) with different degrees of asymmetry. The reaction coordinate is parallel to the z-axis of the simulation box and the distance to the z-coordinate of the center of mass of the CNT. The hydration shell is plotted as a function of the reaction coordinate for potassium (c) and chloride (d) for a narrow CNT. The radial distribution function (RDF) of potassium (e) and chloride (f) in bulk and in the center of a 8 x 50 CNT with varying degree of asymmetry. CNT ellipsoid ratios of ½, ¾ and 1 are shown as blue, orange and green lines respectively. Bulk values are shown as red lines.

resulting in $\kappa(a_i, b_i) \approx \kappa_{bulk}$, while narrower constrictions with smaller radii lead to reduced conductivity $\kappa(a_i, b_i) < \kappa_{bulk}$. Hence, one can further adapt the model for channel resistance / conductance based on the PoreAnalyser (PA) profile to:

$$R_{PA} = \frac{1}{g_{PA}} = \sum_i \frac{z_i - z_{i-1}}{\kappa(a_i, b_i)\pi a_i b_i} \tag{5.6}$$

The conductivity function is represented by a double-sigmoid function with parameters c_1 and c_2, given by:

$$\kappa(a, b) = \frac{1}{1 + \exp\left(-\frac{a}{c_1} + c_2\right)} \cdot \frac{1}{1 + \exp\left(-\frac{b}{c_1} + c_2\right)} \kappa_{bulk} \tag{5.7}$$

To fit parameters c_1 and c_2 of the model, the MD conductance of 18 CNT systems are utilised, encompassing 6 CNTs with distinct cross-sectional areas A_{xy} and three different elliptic radii ratios (a and b), as training data. Every model system of the training data corresponds to a point in Fig. 5.9(a). In Fig. 5.9(b), the ratio $\kappa(a_i, b_i)/\kappa_{bulk}$ is illustrated as a function of varying radii, keeping one radius constant while the other varies. The conductance of these model systems is calculated based on their mean radii and compared to MD conductance. The conductivity model

$$\kappa(a, b) = \frac{g_{MD}}{\pi ab/L} \tag{5.8}$$

is employed for fitting, and Fig. 5.9(a) depicts the ratio $\kappa(a_i, b_i)/\kappa_{bulk}$ in a two-dimensional plot as a function of both radii, with training data points plotted in the plane spanned by the radii a and b. When utilizing the bulk conductivity κ_{bulk} to predict the conductance of CNTs with eq. (5.5), an overestimation of conductance is evident, as depicted in Fig. 5.9(c). However, employing the conductivity model from eq. (5.7) yields an improved correlation between MD conductance g_{MD} and the heuristic conductance g_{PA}, resulting in an R^2 value of 0.916.

To further validate the enhanced physical conductance model, the comparison to biological ion channels is extended by embedding various Innexin and Leucine-rich repeat-containing protein 8 (LRRC8) structures (see Table 5.2) in a POPC membrane and measuring MD conductance through simulations with an external potential. In Fig. 5.9(d), MD conductance and the heuristic based on HOLE or PoreAnalyser pathways are compared, both with and without the conductivity model, for the Innexin [232, 233] and LRRC8 [234, 235] test systems. The use of PoreAnalyser or HOLE pathway profiles in conjunction with bulk conductivity consistently overestimates conductance. Conversely, employing the conductivity model from (5.7) with a HOLE pore profile often results in underestimation. The optimal agreement between MD conductance and the heuristic is achieved by using the conductivity model $\kappa(a, b)$ in combination with the PoreAnalyser profile based on ellipsoidal probe particles. The Pearson correlation coefficients of the conductance based on HOLE and PoreAnalyser profiles with the conductivity model are 0.916 and 0.918 respectively. For the INX-6 hemichannel structures (6kfg and 6kfh in Table 5.2) from Burendei *et al.* [232], the conductance is overestimated and the model based on a HOLE profile overestimates the conductance least. For the other four test systems, the model based on the PoreAnalyser profile performs best. The errors of the MD conductance correspond to one standard deviation of three 250 ns simulations with external potential, as described in section 3.8. To estimate the errors of the conductance model, the pore pathway for ten frames from the simulations is calculated to calculate the standard deviation. The final data point in in Fig. 5.9(d) is an outlier and represents the 5h1q system. Conductivity models based on both HOLE and PoreAnalyser profiles underestimate the 5h1q MD conductance of 5 nS, with the PoreAnalyser profile yielding the smallest error. The LRRC8 structures 6m04, 8b42, and 8b41 exhibit relatively low MD conductance values of 0.7, 0.8, and 1.0 nS, respectively. The conductivity models also slightly underestimate their MD conductance values. Once again, the PoreAnalyser profile outperforms HOLE in accurately predicting the MD conductance.

PDB code	Channel	Authors
5h1q	INX-6 gap junction hemichannel	Oshima *et al.* [233]
6kfg	INX-6 hemichannel in detergent	Burendei *et al.* [232]
6kfh	INX-6 hemichannel in a nanodisc	Burendei *et al.* [232]
6m04	homo-hexameric LRRC8D	Nakumara *et al.* [234]
8b40	homomeric LRRC8C	Rutz *et al.* (44)
8b41	heteromeric LRRC8A/C	Rutz *et al.* [235]
8b42	heteromeric LRRC8A/C	*Rutz et al.* [235]

Table 5.2: Innexin and LRRC8 structures used to test the conductance model.

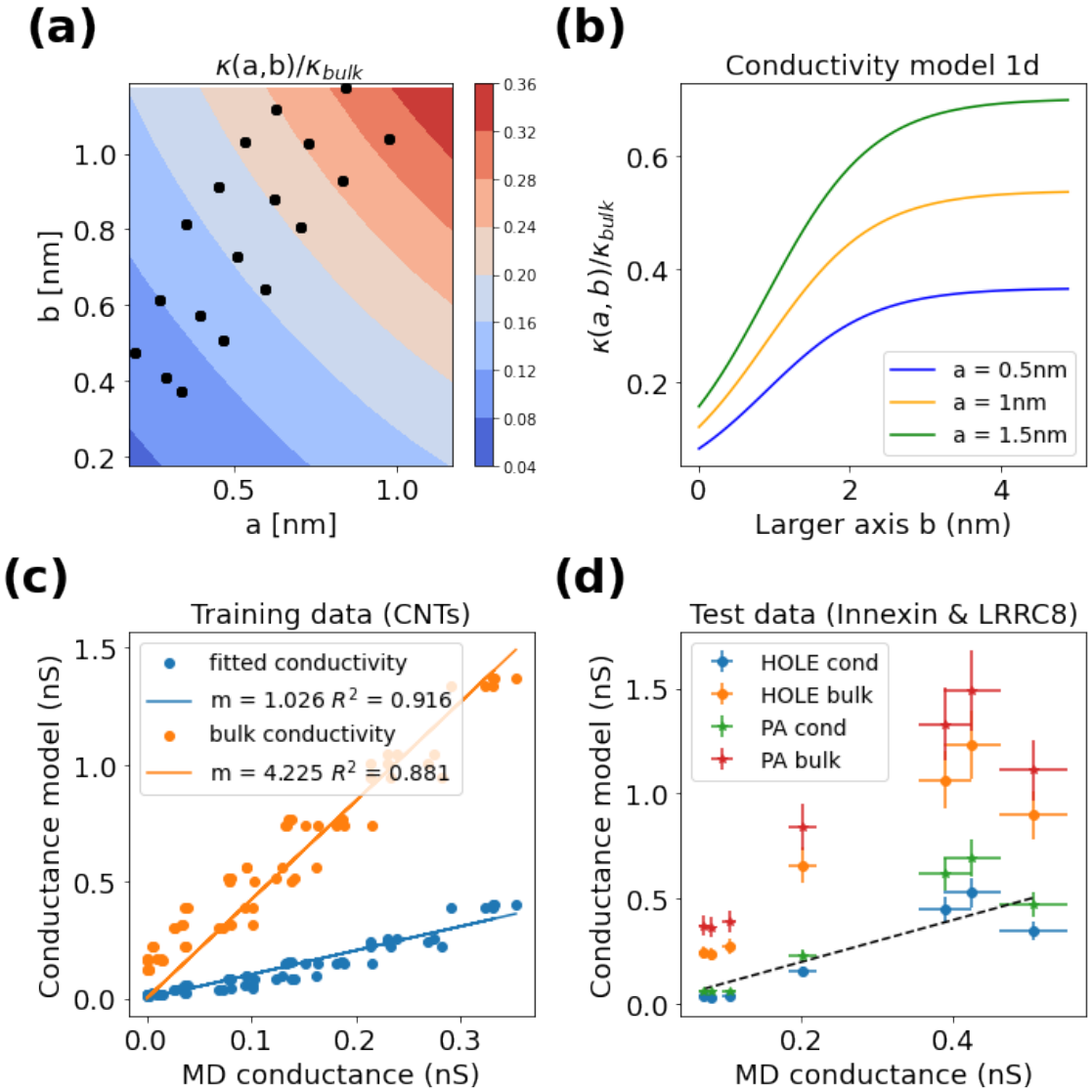


Figure 5.9: (a) Conductivity model for KCl in confined environments characterised by radii a and b. The color surface of the ratio $\kappa(a,b)/\kappa_{bulk}$ characterised by the two radii a and b. The black dots represent the CNT systems summarized in Table 5.1 with varying degree of asymmetry. (b) One-dimensional sigmoid function with one radius fixed. (c) Comparison of MD conductance (considered to be ground truth) and conductance model using bulk conductivity (orange) and conductivity model (blue) for the CNT systems summarized in Table 5.1. (d) Predicting conductance values for biological ion channels (see Table 5.2) using the bulk conductivity or the conductivity model based on a HOLE (HOLE bulk or HOLE cond) or a PoreAnalyser (PA bulk or PA cond) profile. The closer the data points are to the dashed line, the better the model predicts the MD conductance.

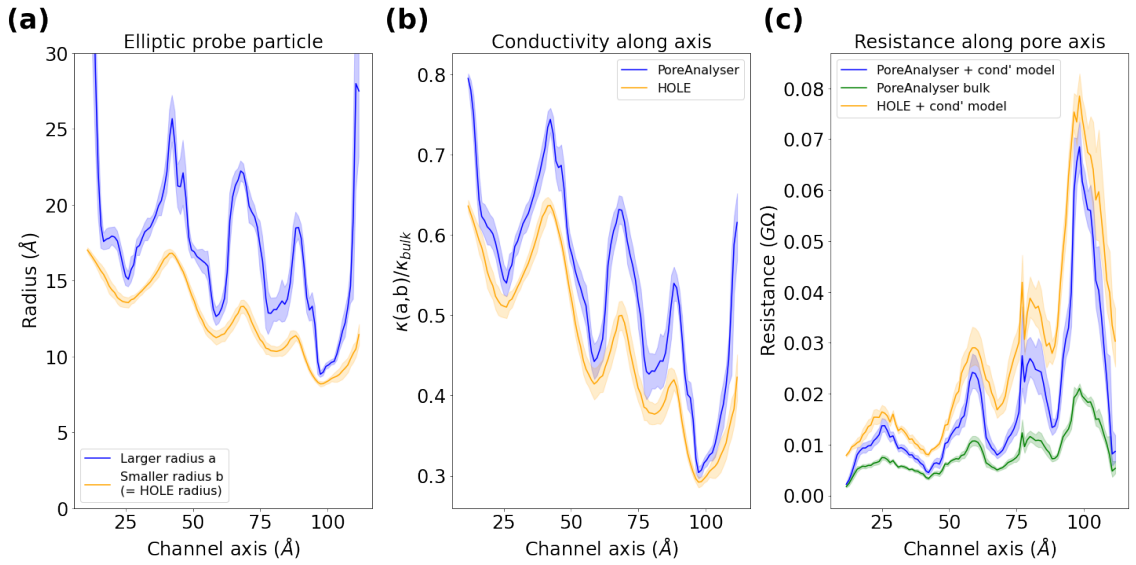


Figure 5.10: Conductivity model for Innexin (PDB code 6kfg). (a) Pore profile described by radii a (blue line) and b (yellow line) along the channel axis. (b) Conductivity ratio $\kappa(a,b)/\kappa_{bulk}$ along the channel axis as computed by HOLE (yellow) and PoreAnalyser (blue). (c) Resistance along the channel axis employing pore profile based on HOLE (yellow line) and PoreAnalyser with (blue line) and without (green line) the conductivity model, see eq. (5.2) and (5.1). Errors are one standard deviation derived from an analysis of an ensemble of structures generated from a single MD trajectory.

To further enhance the understanding of the conductance model in eq. (5.2) and highlight the contributions of different terms, the pore profile of an innexin system is presented in Fig. 5.10(a). The conductivity along the channel axis is influenced by the radii a and b of the elliptical probe particle, as illustrated by the minima in Fig. 5.10(b). The conductivity $\kappa(r_{HOLE})$ of a HOLE profile is smaller than the conductivity $\kappa(a,b)$ based on the PoreAnalyser profile. Ions, confined in a narrow environment, have a decreased mobility than ions in bulk solution. Calculating resistance based on the cylindrical approximation for each probe particle and plotting it along the channel axis in Fig. 5.10(c), resistance maxima correspond to the narrow constrictions in Fig. 5.10(a). Substituting the bulk conductivity κ_{bulk} instead of the conductivity model $\kappa(a,b)$ reduces channel resistance, resulting in an overestimation of channel conductance, as demonstrated in Fig. 5.10(c).

5.4 Conclusion

Th focus of this chapter was on capturing pore asymmetry and its influence on conductance and barriers to permeation. Asymmetry can be observed in crystallographic or cryo-EM structures, which can result from the heterogeneous composition of subunits within the channel complex. While crystal structures of channels with homogeneous subunit decomposition exhibit a symmetrical arrangement, the symmetry is broken when these proteins are simulated in more realistic conditions. By incorporating new features into the PoreAnalyser methodology, I was able to accurately capture and analyse the asymmetry of the pore, shedding light on its functional implications and providing a more comprehensive understanding of ion channel dynamics. The additional complexity of the pore computation adds a slight overhead to speed of the pathfinding process, which can take approximately 1 minute per frame when using 12 CPUs in parallel, but for most use-cases this is a reasonable timing. Consideration of asymmetry allows the user to refine a physical conductance model to obtain a heuristic estimate for single channel conductance. This model allows users to predict single channel conductance without running costly simulations and can be used as

an easy and cheap method for quickly predicting the functional state of new channel structures.

To make the package as usable as possible, an interactive web-service is provided that allows users to calculate the pore profile of any input structure without the need for installation or downloads. Additionally, the PoreAnalyser package is pip installable and the source code is available on GitHub, which includes scripts for visualizing results in popular molecular graphics programs such as Chimera, VMD [236], and PyMOL [237].

Chapter 6

Limitations of Non-polarisable Force Fields in Describing Anion Binding Poses in Non-polar Synthetic Hosts

Disclaimer

This chapter is based on my published work in Physical Chemistry Chemical Physics, Seiferth *et al.* (2023) [71], see also in the list of publications. All investigation, writing and figures in that paper were produced by myself with feedback from the co-authors, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

6.1 Introduction

Biological ion channels and transporters are crucial for the proper function of all living cells. They facilitate the selective passage of ions and water across cell membranes, regulating cellular electrical activity and maintaining cellular and ionic homeostasis. Chloride, the most prevalent anion in living organisms, plays a vital role in controlling membrane potential, cell volume, and pH balance [22]. Imbalances in chloride ion regulation can lead to various diseases, such as cystic fibrosis [23]. Synthetic ionophores that transport anions across membranes have potential therapeutic applications to compensate for defective ion transport [24], and some artificial chloride channels have shown anticancer properties [25]. Ion transport is typically highly specific to certain ions, essential for maintaining ion concentration gradients across biological membranes. The interactions of both biological and synthetic carriers with anions and cations have been studied extensively, especially now that so many structures are available. However, most research has focused on cation interactions. Anion interactions pose a greater challenge due to their relatively high solvation energy compared to similarly sized cations. Additionally, the presence of water can disrupt the stabilizing intermolecular interactions necessary for effective anion binding [26]. Selective binding and discrimination of anions requires an understanding of how anions interact with synthetic receptors. For instance, Hewitt *et al*. studied anion interaction with synthetic receptors with structural data and DFT [238]. In recent studies, Lisbjerg *et al*. introduced and investigated a biotin based macrocyclic anion receptor [153, 239, 240, 241]. This carrier comprises six biotin units forming a ring through their nitrogen atoms, with twelve CH-groups from the convex side of each biotin unit directed towards the center, creating a cavity as shown in Fig. 6.1. The CH-bonds are slightly polarized due to the inductive electron withdrawal by the urea moieties, making the cavity somewhat hydrophobic [240]. The biotin monomer itself does not recognize anions [241]. However, it was discovered that the carrier with methylated side chains preferentially binds smaller halides, such as chloride and bromide, over the larger halide iodide in the nonpolar solvent acetonitrile [240]. In water, this trend reverses, with the biotin macrocycle exhibiting a higher binding affinity for larger anions like iodide and bromide over chloride. Solvent effects influence anion selectivity in receptor molecules, with water being a strongly competitive solvent that can negate most stabilizing intermolecular forces necessary for anion binding in host-guest systems [26]. Additionally, the sulfide moiety in each biotin monomer of the macrocycle can be oxidized, see Fig. 6.1(a), resulting in a sulfoxide moiety that is more electron-withdrawing

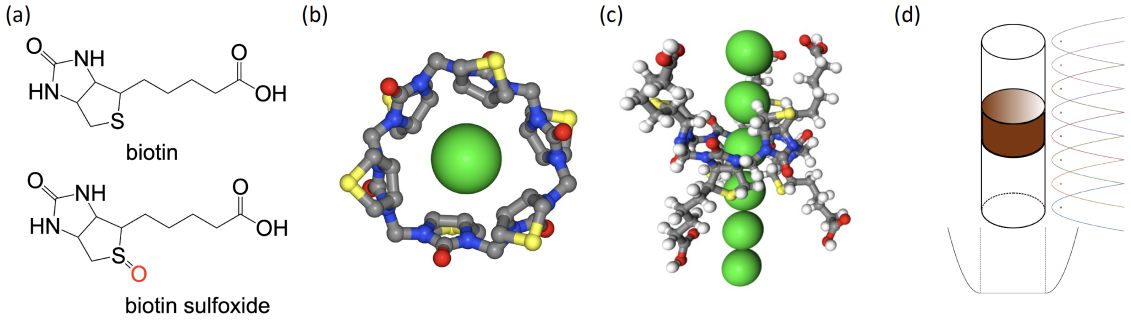


Figure 6.1: (a) Biotin and biotin sulfoxide unit. (b) Chloride (green sphere) inside the cavity of biotin[6]uril macrocycle. Side chains and hydrogen atoms are not shown. The centre of the cavity is defined as the centre of geometry of the 12 inside-pointing carbon atoms (grey spheres) from the convex side of the biotin units. (c) Visualisation of the reaction coordinate for PMF calculations. The distances of neighbouring chloride ions to the centre of the cavity differ by 4 Å as illustrated by the green spheres along the reaction coordinate. (d) System set-up. The position and orientation of the ion carrier (brown cylinder) is preserved by restraining three carbon atoms. Umbrella potentials are applied along the z-axis to ensure sufficient sampling along the reaction coordinate (distance of the ion to the cavity centre). A flat-bottom potential in the xy-plane prevents the ion from moving far away from the carrier.

than sulfide, thereby altering the cavity potential of the macrocycle [241]. These characteristics make this carrier an interesting case study for different computational approaches that appear to influence anion interactions.

Anions can be ranked based on their ability to stabilize proteins in water according to the Hofmeister series: F < Cl < Br < I [51]. Smaller, harder halides like fluoride and chloride are less polarisable and possess a stronger, more well-defined solvation shell. These anions decrease the solubility of nonpolar molecules (”salting out”) and are less likely to migrate toward hydrophobic interfaces. In contrast, larger, softer, and more polarisable anions such as bromide and iodide have a weaker, less well-defined solvation shell. They increase the solubility of nonpolar molecules (”salting in”) and are more inclined to migrate toward hydrophobic interfaces.

In this chapter, electrostatic effects, including induced polarisation, are investigated by calculating the free energy of binding of different anions to biotin macrocycles solvated in water and in acetonitrile using MD simulation techniques.

6.2 Methods

Here, two types of classical Hamiltonians, or force fields, are used which primarily differ in their descriptions of non-bonded interactions, specifically van der Waals and electrostatic interactions as described in section 2.2.1. First, the expected error levels in both fixed charge and polarisable models is determined. The accuracy of molecular mechanics methods in capturing solvation and binding properties of ions is known to be problematic. For instance, explaining the high selectivity for fluoride over chloride, despite their identical charge and similar size, in fluoride-selective ion channels [15], or the potassium/sodium selectivity in potassium channels [242], is challenging. Ion parameters are typically calibrated against reference data on ion–water interactions, but their transferability to interactions with other functional groups is limited [243].

There are different methods for improving the accuracy of anion MD properties by changing non-bonded interactions in a molecular mechanics force field, as discussed in section 2.2.1. In this chapter, the Lennard-Jones parameters for van der Waals interactions reoptimised by Li *et al.* to reproduce hydration free energies within the AMBER force field [113] are used in combination with the Parsley force field [140].

Many non-polarisable force fields are calibrated to reproduce gas-phase conformational energies and geometries. However, non-bonded interactions, especially electrostatics, play a crucial role in condensed phase simulations. Electrostatics models in non-polarisable force fields often exhibit polarisation effects that exceed gas-phase expectations [112]. The transferability of parameters in pairwise additive force fields is limited due to their fixed charges and inability to model induced polarisation effects. The Parsley force field employs a fixed partial charge model and a Coulomb potential for electrostatic interactions. This model is akin to a truncated multipole expansion, including only monopole (charge) terms while omitting dipole, quadrupole, and higher-order terms. In contrast, the AMOEBA force field [142, 83] uses different potentials for non-bonded interactions, describing van der Waals forces with a buffered 14-7 potential. AMOEBA accounts for multipoles up to quadrupoles at each atomic center, truncating octupole and higher-order moments as described in section 2.2.1. Its electrostatic energy calculations include contributions from both permanent and induced multipoles, with only the dipole moment being inducible, while monopole (charge) and quadrupole moments remain fixed. This explicit dipole polarisation allows the AMOEBA model to adapt to varying or heterogeneous molecular environments [83]. When an external field distorts atomic electron density, AMOEBA induces a classical point dipole moment at each polarisable atomic site based on the electric field experienced by that site. To prevent the polarisation catastrophe, where induced dipole moments diverge at short interatomic distances, AMOEBA employs Thole's damped interaction method [97], which dampens polarisation interactions at very short ranges by smearing out atomic multipole moments. For parametrisation with the AMOEBA force field, the POLTYPE protocol [143] was used, with a detailed description of the process provided in section 3.1.2.

The PMF, see section 2.2.4, was calculated using umbrella sampling for two different force fields (Parsley and AMOEBA), two solvents (water and acetonitrile), and three halide anions (chloride, bromide, and iodide) as described in section 3.5. The reaction coordinate was defined as the distance between the z-coordinate of the ion and the z-coordinate of the COM of the cavity, which is defined by twelve carbon atoms, as shown in Fig. 6.1(c). Position restraints were applied to three carbon atoms of the cavity to maintain the orientation of the carrier during simulations while allowing full flexibility of the cavity. The ion's z-coordinate was harmonically restrained to specific distances from the COM of the cavity in each umbrella window to ensure proper sampling along the reaction coordinate, as depicted in Fig. 6.1(d). The first nanosecond of each window was discarded as equilibration time, and the remaining 24 ns were divided into three sub-windows of 8 ns each, creating three data sets. The PMF was calculated from each of these sets. The bulk region was defined as the zero point, and the PMF averaged over the three data sets was interpreted as the binding free energy, with the standard deviation representing the error. For AMOEBA simulations, a similar setup was used, but the z-coordinate of the ion was biased with 2 kcal/mol/Å^2 and 0.5 Å spacing across 60 windows between -15 Å and 15 Å relative to the COM of the cavity. Each window was simulated for 13 ns, with the first nanosecond discarded as equilibration time. The remaining simulation time was divided into three parts to create three data sets. Histogram overlap for each PMF was checked to ensure good overlap between windows, with example overlap data provided in the Fig. B.1 in the appendix.

6.3 Results and Discussions

6.3.1 Ion Interactions with Methylated Biotin[6]uril in Acetonitrile

To calculate the free energy of binding between various ions and hosts, molecular dynamics simulations were conducted using both a non-polarisable additive force field (Parsley) and the polarisable AMOEBA force field. Given that calculations with GAFF revealed similar PMF profiles (Fig. B.5), the focus here lies on the comparison between Parsley (representative of non-polarisable force fields) and AMOEBA. In Fig. 6.2,

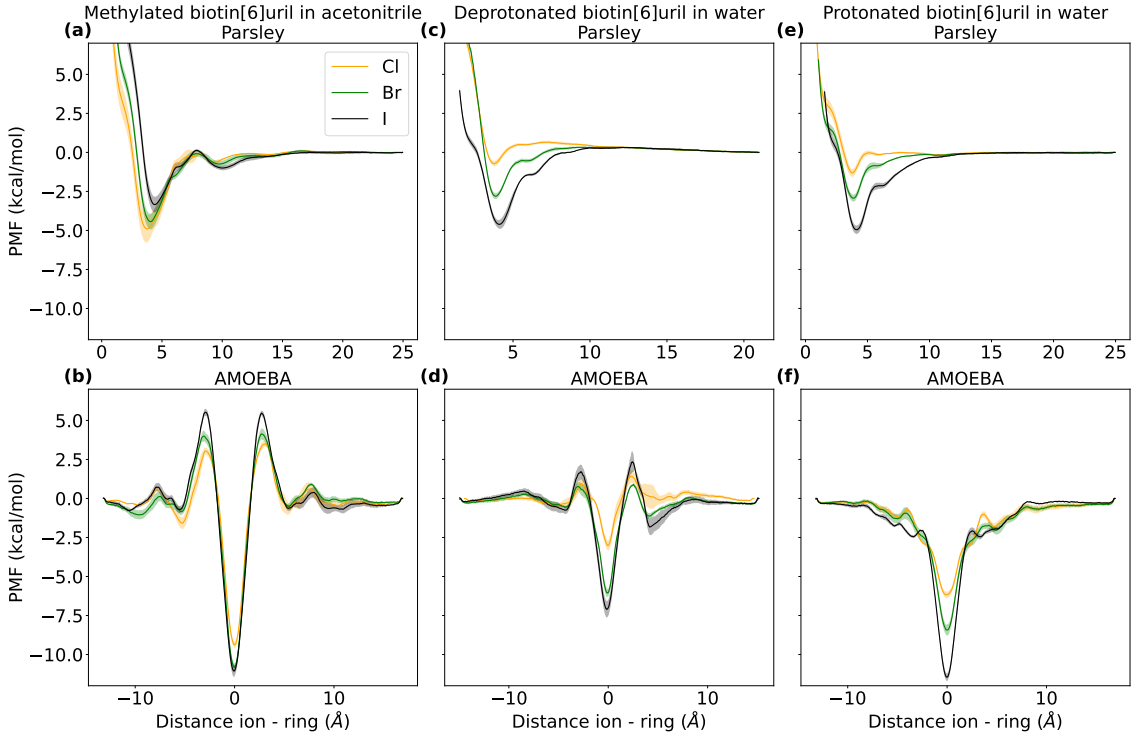


Figure 6.2: Potential of mean force (PMF) of ions and a biotin macrocycle with methylated side chains in (a and b) acetonitrile and an unmethylated macrocycle in water in deprotonated (c and d) and protonated (e and f) forms for two different force fields. Each PMF here corresponds to the average of three data sets. Error bars represent one standard deviation. Two representative binding poses are depicted in Fig. 6.3. The histogram overlap for these PMFs can be found in the Fig. B.1. Chloride is depicted in yellow, bromide in green and iodide in black.

the PMF of the interaction between the biotin macrocycle and different anions in acetonitrile and water is shown. The estimated binding free energies are summarized in Tables 6.1 and 6.2, and are compared to experimental free energy values obtained through Isothermal Titration Calorimetry (ITC) [153, 240]. The solvent strongly influences the binding affinities.

Table 6.1: Estimates for the binding free energies for biotin[6]uril hexamethyl ester and different anions in acetonitrile. Binding free energies are estimated by the difference between minimum and bulk PMF values. The corresponding PMFs are depicted in Fig. 6.2 Note that Lisbjerg *et al.* do not report error estimates [240].

kcal/mol	Chloride	Bromide	Iodide
ITC in acetonitrile [153]	-6.3	-6.3	-4.2
PMF (Parsley)	-4.0$\pm$0.9	-3.5$\pm$0.5	-2.4$\pm$0.5
PMF (AMOEBA)	-8.1$\pm$0.2	-9.5$\pm$0.2	-9.8$\pm$0.4

According to ITC data from Lisbjerg *et al.* [153], in the nonpolar solvent acetonitrile, chloride and bromide exhibit higher binding affinities than iodide. The simulations using the non-polarisable Parsley force field show a similar trend, with smaller and harder anions binding more effectively than larger, softer anions, as shown in Table 6.1. In contrast, the binding affinities estimated with the polarisable AMOEBA force field are higher, but no differences in anion selectivity are observed, with all three halides appearing stably bound. The inclusion of multipoles and polarisation therefore does not appear to improve the correlation with the experimental ITC data. However, a closer examination reveals substantial differences between the PMFs obtained from Parsley and AMOEBA (Fig. 6.2). Contrary to the experimentally derived structure

Table 6.2: Estimates for the binding free energies in kcal/mol for biotin[6]uril hexaacid ester and different anions in water. Binding free energies are estimated by the difference between minimum and bulk PMF. The corresponding PMFs are depicted in Fig. 6.2. In water, the acid could exist as protonated or deprotonated and thus both forms were considered. The PMF minima are corrected by eq. (3.2) to take the orthogonal degrees of freedom into account (note that the binding energy is positive for chloride and the deprotonated host when using the Parsley force field).

kcal/mol	Chloride	Bromide	Iodide
ITC in water [239, 241]	-1.5	-3.8	-4.8
Deprotonated (Parsley)	+0.6±0.2	-1.5±0.2	-3.3±0.3
Deprotonated (AMOEBA)	-1.8±0.3	-4.8±0.2	-5.8±0.5
Protonated (Parsley)	-0.4±0.3	-2.0±0.2	-4.1±0.3
Protonated (AMOEBA)	-4.9±0.2	-7.1±0.4	-10.2±0.2

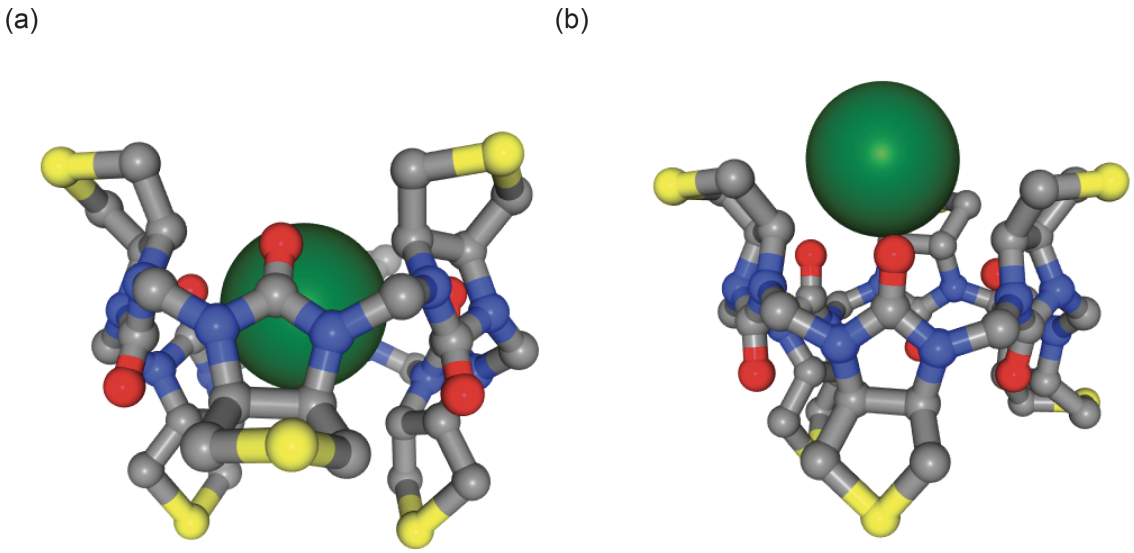


Figure 6.3: Different binding poses of ions (white spheres) within the biotin macrocycle. Side chains and hydrogen atoms are omitted for clarity. The distance between the centre of mass of the cavity (defined by 12 carbon atoms) and the ion is zero and 3.5 Å respectively in (a and b). Simulations with the AMOEBA force field predict binding within the cavity (a) whereas non-polarisable simulations predict the favourable binding site to be slightly outside the cavity (b).

[239], Parsley suggests that ion binding within the cavity is not favourable. In Parsley simulations, the PMF minima appear between 3–4 Å on the reaction coordinate, indicating that an ion within the cavity (zero distance in the reaction coordinate) must overcome a steep energy barrier, see Fig. 6.2(a) and Fig. 6.3(b). In stark contrast, the AMOEBA force field simulations show that ions within the cavity are in a favourable position, with the origin of the reaction coordinate corresponding to the energetic minimum, see Fig. 6.2(b) and Fig. 6.3(a). Ions coming from the bulk face an energetic barrier of 3 kcal/mol to bind symmetrically within the cavity.

6.3.2 Ion Interactions with Unmethylated Biotin[6]uril in Water

The selectivity trends observed in acetonitrile are reversed in water. According to ITC experiments, the larger and softer iodide binds more strongly to the unmethylated biotin macrocycle in water than the smaller and harder bromide, and MD simulations replicate this trend. Larger, softer, and more polarisable anions like bromide and iodide have weaker and less well-defined solvation shells compared to the smaller

Table 6.3: Absolute Single-Ion Solvation Free Energies (kcal/mol) in water and in acetonitrile [70].

Ion	Water	Acetonitrile
Chloride	-74.5	-62.4
Bromide	-68.3	-59.3
Iodide	-59.9	-53.0

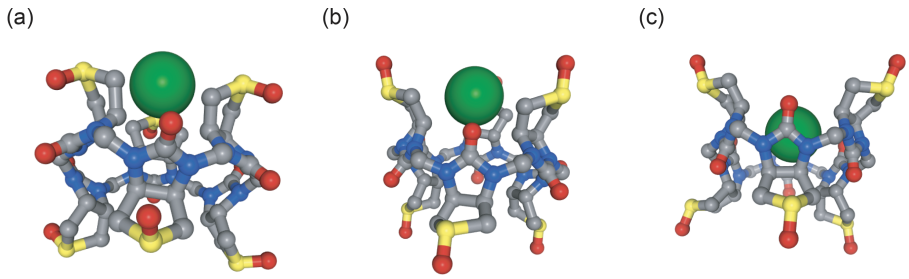


Figure 6.4: Non-central binding pose of chloride (green sphere) within the biotin-D-sulfoxide (a) and biotin-L-sulfoxide (b). Central binding pose of iodide within the biotin-L-sulfoxide macrocycle (c). Side chains and hydrogen atoms are omitted for clarity. The distance between the centre of mass of the cavity (defined by 12 carbon atoms) and the ion is 4 Å and corresponds to a free energy minimum.

chloride ion. The hydration free energies of these halide anions are summarized in Table 6.3. Although the hydration free energies of halides are of smaller magnitude in acetonitrile than in water, the overall trend remains consistent across both solvents. The PMFs shown in Fig. 6.2 indicate that the energetic landscape for iodide has forms deeper wells compared to bromide or chloride across all three force fields. Table 6.2 summarizes the binding free energies of ions with the carrier in water, demonstrating that the larger and more polarisable the ion, the higher its binding affinity, thus confirming a preference for larger anions. Both force fields capture this overall trend, but the AMOEBA force field once again predicts larger binding free energies than the non-polarisable Parsley. Both force fields again differ in where they predict the ions to bind. In simulations with AMOEBA, ions bind centrally in the cavity as depicted in Fig. 6.3(a), while Parsley simulations suggest non-central binding, see Fig. 6.3(b).

6.3.3 Ion Interactions with Unmethylated Biotin-sulfoxide Macrocycle in Water

The sulfide moiety of each biotin monomer of the macrocycle can be oxidized. There are two possible conformations as a result: the oxygen can point away from the cavity (D-isomer) or perpendicular to the cavity (L-isomer) or as depicted in Fig. 6.4.

Furthermore, the PMF for both sulfoxide isomers solvated in water were calculated. Andersen *et al.* could not unambiguously determine the stereochemistry of the two sulfoxide macrocycles, but suggested that the hexasulfoxide used in their ITC study was the L-isomer [241]. The oxidized macrocycle exhibits the same selectivity trend in water as the methylated biotin[6]uril macrocycle, preferring larger ions over smaller ones. This selectivity trend was successfully reproduced with the Parsley force field. The binding free energies corresponding to the PMFs in Fig. 6.5 and 6.6 are presented in Table 6.4. Halide anions bind more effectively to biotin-D-sulfoxide than to L-sulfoxide, likely due to the negatively charged oxygen in the L-isomer pointing slightly towards the cavity, creating an unfavourable electrostatic interaction between the anion and the macrocycle.

The selectivity trend in water could not be reproduced with AMOEBA and the binding free energies for the halide anions are within error (Table 6.4). There are two stable binding poses for anions with biotin-L-sulfoxide; the central one in the cavity and the second 3.5–4 Å from the centre of the cavity. For the

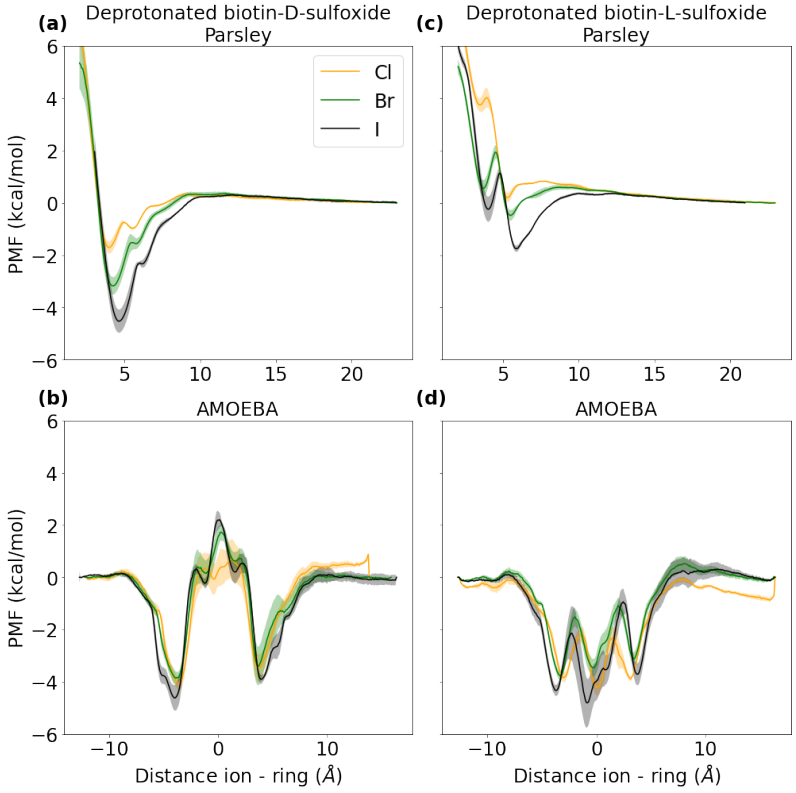


Figure 6.5: Potential of mean force (PMF) of ions and a deprotonated biotin sulfoxide macrocycle in two different conformations (D-conformer in (a) and (b), L-conformer in (c and d)) with unmethylated side chains in water for two different force fields (Parsley in the first row, AMOEBA in the second row). Each PMF here corresponds to the average of three data sets. The standard deviation of the three data sets is plotted as error bar. Chloride is depicted in yellow, bromide in green and iodide in black.

Table 6.4: Estimates for the binding free energies for biotin-sulfoxide[6]uril hexaacid ester and different anions in water. Binding free energies are estimated by the difference between minimum and bulk PMF in kcal/mol.

Conformation		Chloride	Bromide	Iodide
L-conformer (deprotonated)	ITC [241]	/	-2.2	-3.8
	PMF (Parsley)	unstable	+0.8±0.2	-0.5±0.1
	PMF (AMOEBA)	-2.9±0.3	-2.5±0.2	-3.5±0.9
D-conformer (deprotonated)	PMF (Parsley)	-0.4±0.1	-1.9±0.1	-3.2±0.2
	PMF (AMOEBA)	-3.0±0.4	-2.6±0.3	-3.3±0.5
L-conformer (protonated)	PMF (Parsley)	unstable	+0.7±0.1	-1.1±0.1
	PMF (AMOEBA)	-2.7±0.7	-4.0±0.4	-3.9±0.5
D-conformer (protonated)	PMF (Parsley)	-2.3±0.1	-3.2±0.1	-4.6±0.2
	PMF (AMOEBA)	-4.3±0.5	-3.9±0.2	-3.7±0.3

D-isomer, only a non-central binding pose is stable and the halide anions cannot stably bind to the cavity centre.

6.3.4 Comparison of the Three Macrocycles

Overall, the sulfide macrocycle serves as a ”better host” than the sulfoxide macrocycle, a trend confirmed by both ITC and molecular dynamics simulations, with and without polarisable force fields. The sulfoxide

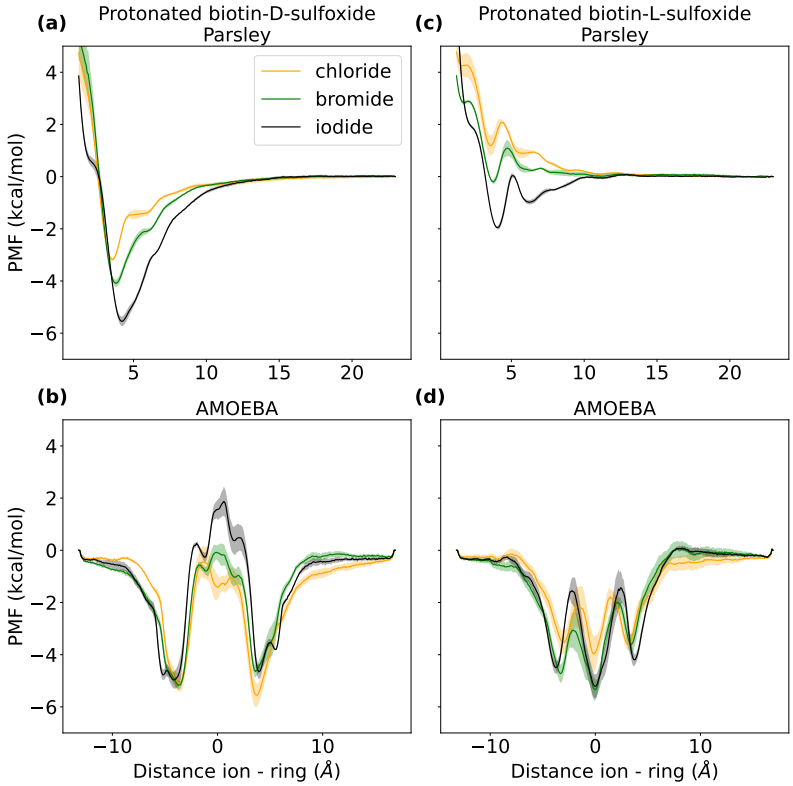


Figure 6.6: Potential of mean force (PMF) of ions and a protonated biotin sulfoxide macrocycle in two different conformations (D-conformer in (a and b), L-conformer in (c) and (d)) with unmethylated side chains in water for two different force fields (Parsley in the first row, AMOEBA in the second row). Each PMF here corresponds to the average of three data sets. The standard deviation of the three data sets is plotted as error bar. Chloride is depicted in yellow, bromide in green and iodide in black.

moiety is more electron-withdrawing than sulfide, resulting in different electrostatic surface potentials for the sulfoxide macrocycles. Andersen *et al.* calculated these potentials at the CAM-B3LYP/6-311G(d) level of theory [241]. The cavity potential for biotin[6]uril and biotin-D-sulfoxide[6]uril is approximately +130 kJ/mol, while the L-sulfoxide conformer has a more positive surface potential of around +200 kJ/mol (see Fig. 6.7 for calculations at the B3LYP/cc-pvtz level of theory, used for electrostatic fitting of monopole, dipole, and quadrupole AMOEBA parameters as described in section 3.1.2). For all three systems, the cavity itself is neutral (green). The oxidized carriers are more polar on either side of the cavity due to the additional negatively charged oxygen (red).

Polarisable and non-polarisable force fields predict different binding poses for the biotin[6]uril macrocycle and ions. In non-polarisable force fields, halide anions cannot stably bind in the center of the hydrophobic cavity. Instead, they bind at positions 3.5–4 Å from the center, which do not correspond to a free energy minimum in simulations with the polarisable AMOEBA force fields (as shown in Fig. 6.2). Oxidizing the sulfide moiety increases the system's polarity. One can observe the same stable binding poses in simulations with both non-polarisable and polarisable force fields. For both the L- and D-isomers, a free energy minimum is found around 3.5–4 Å from the center of the cavity with both force field types, and this non-central binding pose (Fig. 6.5) is stable in both AMOEBA and Parsley simulations. For the biotin-L-sulfoxide macrocycle, the polarisable AMOEBA force field reveals an additional stable binding pose not seen in non-polarisable simulations. The central binding pose is stable only for the biotin-L-sulfoxide (and biotin[6]uril) macrocycle within the AMOEBA framework.

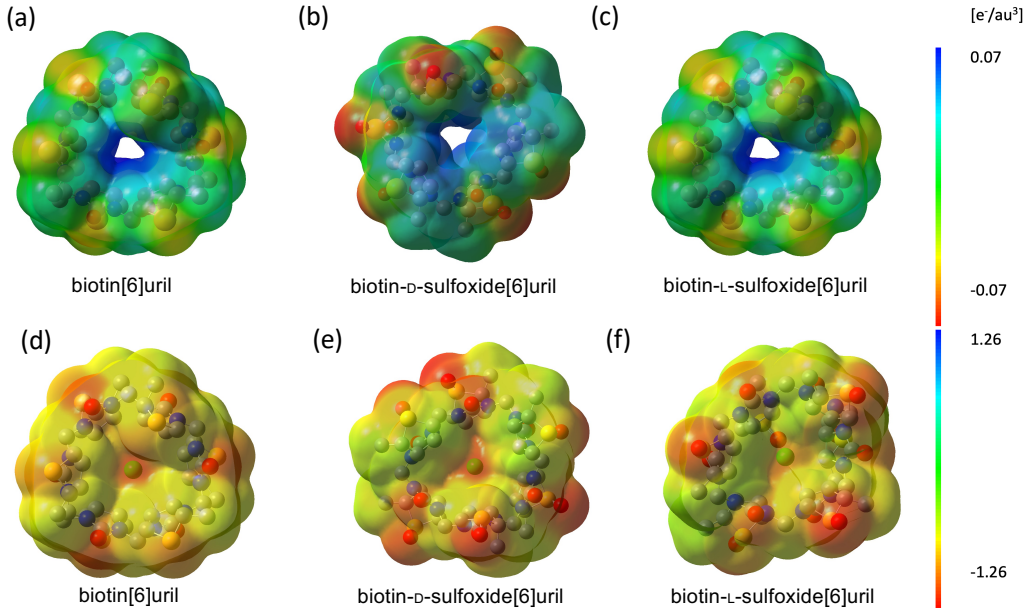


Figure 6.7: Electrostatic surface potentials of (a) biotin[6]uril, (b) biotin-D-sulfoxide[6]uril and (c) biotin-L-sulfoxide[6]uril in the apo state and (d–f) with chloride bound in the central cavity modelled at B3LYP/cc-pvtz level of theory. The macrocycles are shown top-down (as drawn in Fig. 6.1) looking into the cavities. The side chains have been truncated. Red, green and blue correspond to negative, neutral and positive potentials respectively. The cavity potential for the sulfide and D-sulfoxide macrocycle corresponds to approximately +130 kJ/mol whereas the L-sulfoxide isomer has a more positive cavity potential.

6.3.5 Comparison of the Parsley and AMOEBA Force Fields

The conformational ensembles and the flexibilities and rigidities of biotin sulfide and biotin sulfoxide using both the AMOEBA and Parsley force fields were investigated. By comparing the root mean square deviation (RMSD) values of the twelve cavity carbon atoms between the two force fields (Fig. B.6 in the appendix), I found that the average RMSD was less than 1 Å. To further understand the conformational behaviour, the C–N–C–N dihedral distributions was analysed, which describe the orientation of two subunits relative to each other. These distributions were similar across all systems and force fields (Fig. B.7 in the appendix). My findings suggest that the host behaviour is consistent in both force fields, indicating that the biotin[6]uril and biotin-sulfoxide macrocycles exhibit similar flexibilities and rigidities across different force fields. It is important to note that different van der Waals interaction potentials can result in different radii (Fig. 2.3). To determine whether the difference in radius contributed to the inability of the chloride anion to bind centrally in the cavity, I fitted the Lennard-Jones potential to the buffered 14-7 potential (see Fig. B.8). The chloride parameters from Zhang *et al.* [244] were also used. With both of these additional parameterisations, I found that chloride did not stably bind in the center of the cavity. Moreover, fluoride, which has a smaller radius than chloride, was also not stable in the cavity, see Fig. B.8(c) for a PMF for fluoride. These results suggest that smaller ions are not stable in the center of the cavity and that the instability of chloride is not solely due to the use of the 12-6 potential from Li *et al.* [113]. Thus, while Parsley captures the trends observed in ITC, it does not accurately represent the observed binding mode. On the other hand, AMOEBA

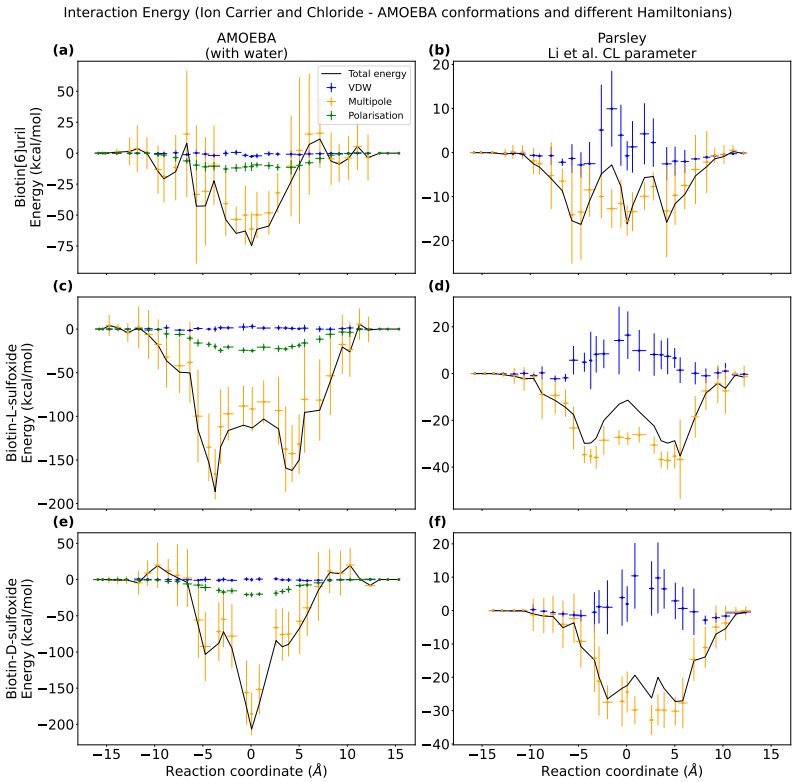


Figure 6.8: Interaction energies along the reaction coordinate for three ion carriers and chloride in water evaluated from the AMOEBA conformations but calculated with either the AMOEBA (a, c and e) or the Parsley (b, d and f) force field. For the Parsley simulations, interactions are computed for VDW (blue) and electrostatics (yellow). For each 13 ns long umbrella window, 13 conformations were selected to evaluate the interaction energy. Note the different energy scales.

captures the binding mode but is less effective in capturing the observed trends. Therefore, the host-ion energetics were explored in greater detail.

6.3.6 Interaction Energies Between Hosts and Ions

Given that AMOEBA captures the binding mode accurately, conformations from these simulations were utilised to compute the interaction energies between the chloride guest and the three different hosts using both force fields. As described in section 3.7, 13 frames were selected from AMOEBA simulations spanning 13 ns, with each frame being 1 ns apart. By calculating the interaction energy profiles along the reaction coordinates, the total energy was decomposed into its individual components. These profiles were computed with bulk as the reference point: the interaction energy of the host and guest is set to zero at a distance of 15 Å from the center of the cavity. To assess the pairwise contributions to the interaction energies using the AMOEBA parameters, the *Tinker analyze* program was employed. Given that polarisable force fields are nonlinear, accounting for multi-body contributions was necessary. Therefore, the interaction energies were evaluated both in the presence and absence of solvent to capture these effects. The total interaction energy between the chloride guest and the host differed by a few kcal/mol when water was present or not during the calculations (Fig. B.9), but the overall trend along the reaction coordinate remained consistent. The minima in the PMF corresponded to the minima of the interaction energy (Fig. 6.8).

To evaluate the interaction energies using the Parsley force field in *Gromacs*, the frames obtained from the AMOEBA simulation were converted to *Gromacs* format using MDAnalysis [219]. Then, these

AMOEBA frames were assessed with the Parsley force field using two different chloride parameter sets: (1) the Li *et al.* parameter set, which has a larger VDW radius than that in AMOEBA (see Fig. 2.3) [113], and (2) re-fitted VDW parameters that resemble the VDW chloride radius in AMOEBA (Fig. B.8). When evaluating the interaction energies of the conformations sampled in AMOEBA simulations with the Parsley force field (Fig. B.11), one can observe that the Lennard-Jones potential is repulsive near the cavity when using the Li *et al.* parameter set for chloride [113]. However, this repulsive VDW contribution vanishes when one uses the re-fitted parameters (Fig. B.10). The interaction energies were also assessed along the reaction coordinate for the Parsley umbrella simulations with the Parsley force field (Fig. B.11). Although the central binding pose is not stable and the potential of mean force shows no minimum (Fig. 6.4 and 6.6), the interaction energy is minimal near the cavity center. The VDW contributions are not repulsive near the cavity when evaluating the Parsley-derived conformations. Therefore, I conclude that the difference in binding behaviour is not due to the difference in VDW parameters for the ions. Ions with smaller VDW radii than the Li *et al.* chloride [113] are not stably bound in the cavity at all (Fig. B.8).

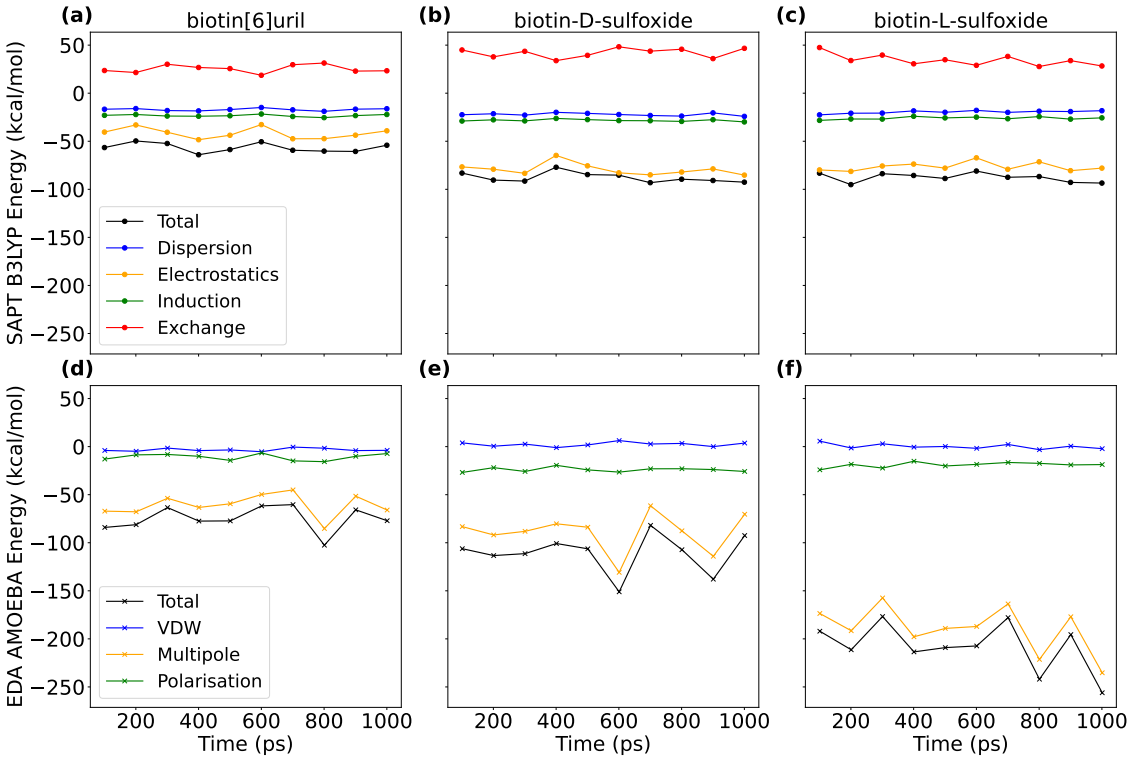


Figure 6.9: Interaction energies between chloride and three hosts (biotin[6]uril, biotin-D-sulfoxide[6]uril and biotin-L-sulfoxide[6]uril). For ten AMOEBA frames, the interaction energy was evaluated with symmetry-adapted perturbation theory (SAPT) on B3LYP level of theory with a jun-cc-pVDZ basis set (a–c) and performed interaction energy decomposition analysis (EDA) with the AMOEBA force field (d–f). Mean and standard deviation of the energy components are summarized in Tables B.2 and B.3. Interaction energies for the non-central binding pose are plotted in Fig. B.12.

After comparing interaction energies for non-polarisable and polarisable force fields, symmetry-adapted perturbation theory (SAPT) was performed at the B3LYP level of theory to obtain a *ground truth* for the interaction energy between the hosts and chloride. As described in section 3.7, the jun-cc-pVDZ basis set was used, which includes chloride. Ten AMOEBA frames were extracted and then the energy decomposition from DFT was compared with an equivalent energy decomposition from the AMOEBA force field (Fig. 6.9 and Tables B.2 and B.3). In the AMOEBA force field, there is no exchange term (red in Fig. 6.9), and the van der Waals term is less favourable than the dispersion term.

Comparing the interaction energies between the Parsley and AMOEBA force fields with the DFT energy decomposition reveals that host systems with narrow and hydrophobic cavities cannot be accurately described using conventional additive force fields like Parsley, which lack higher multipoles and polarisation. While AMOEBA does not include the exchange term, which is non-negligible, it offers a good balance in terms of accuracy, sampling, and performance. The AMOEBA multipole interaction energy for the L-sulfoxide macrocycle is higher than that for the D-sulfoxide macrocycle (Fig. 6.9 and Table B.3). In the D-sulfoxide macrocycle, the oxygen in the sulfoxide group is closer to the centrally bound chloride, resulting in repulsive multipole interactions between these oxygens and chloride. Conversely, in the L-sulfoxide system, the non-favourable Cl–O interaction is weaker, making the overall interaction energy larger in the central binding pose (see Fig. B.13).

Finally, we compare charges derived from quantum mechanical methods and polarisable models to assess the role of charge transfer. In MD simulations, chloride consistently exhibits a charge of -1. In contrast, QM calculations with B3LYP/cc-pvtz level of theory reveal a Mulliken charge of -0.758970 and a Natural Population Analysis (NPA) charge of -0.94392 for chloride bound in the cavity of biotin[6]uril as depicted in Table 6.5. These discrepancies highlight the limitations of fixed-charge MD frameworks, which cannot account for charge transfer effects. Fluctuating charge frameworks such as the CHARMM version do not explicitly model the movement of electrons between molecules [94, 95]. Instead, they redistribute charges within the boundaries of predefined atoms or fragments. Hence, charge transfer from one molecule to another such as from the chloride to the cavity of the macrocycle is not possible either in the CHARMM fluctuating charge force field [95]. To properly describe charge transfer between molecule, one could for instance combine QM for the cavity and chloride with classical MD for the rest of the system or use reactive force fields such as ReaxFF [245].

Table 6.5: Partial charge of chloride bound in the cavity of three macrocycles.

bound chloride in	MD charge	Mulliken charge	NPO charge
biotin[6]uril	-1	-0.758970	-0.94392
biotin-D-sulfoxide	-1	-0.771887	-0.95157
biotin-L-sulfoxide	-1	-0.727659	-0.87970

6.3.7 Role of Water in Interactions of Anions with Hosts

Given AMOEBA's performance so far, the role of water molecules was examined in the interactions of anions with biotin[6]uril, biotin-D-sulfoxide, and biotin-L-sulfoxide. The behaviour of the hydration shells, relative to the distance between the ion and the cavity center using the AMOEBA force field, is illustrated in Fig. 6.10, Table 3.2, and Fig. B.14. There are more interactions between ions and cavity atoms of the sulfoxide macrocycles compared to the sulfide moiety. When an ion binds centrally to the cavity, the hydration shell is partially replaced by cavity interactions in all three systems, but this replacement is more extensive in the sulfoxide macrocycles. Additionally, ions begin replacing their hydration shell earlier when interacting with the sulfoxide host: the hydration shell remains almost intact for an ion 4 Å away from the cavity center of the sulfide macrocycle. This non-central binding pose is unstable and does not correspond to a PMF minimum. However, at the same distance in the sulfoxide macrocycle, the hydration shell is already partially replaced, and this non-central binding pose is stable, corresponding to a PMF minimum for both sulfoxide isomers.

Furthermore, the interaction energy was compared between chloride ions and the biotin[6]uril macrocycle with that of a water molecule bound within the cavity of the macrocycle. As anticipated, the polarisation

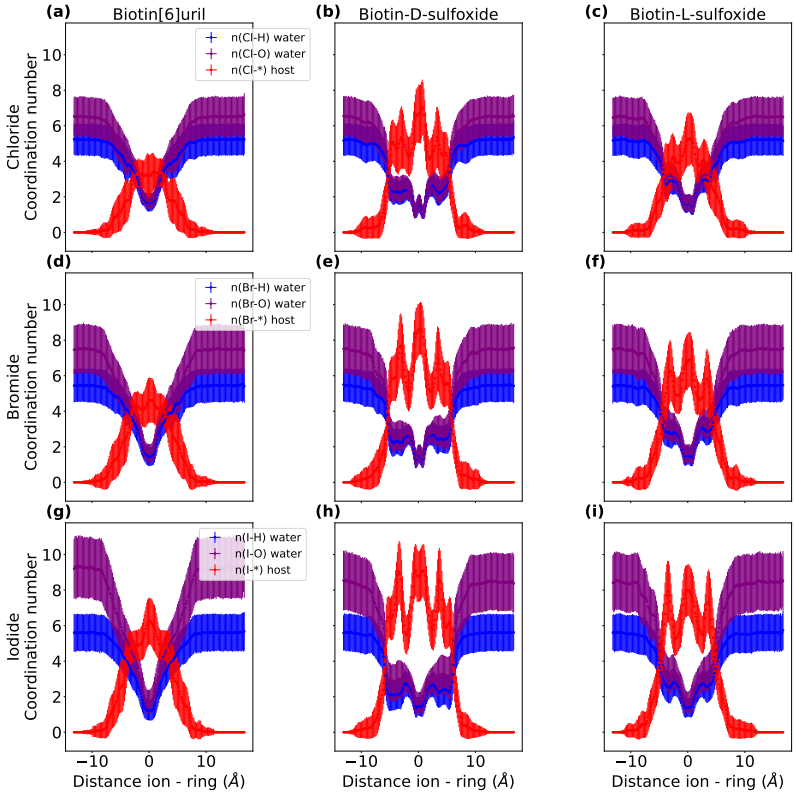


Figure 6.10: Hydration shell of halide anions in biotin[6]uril (a–c), biotin-D-sulfoxide (d–f) and biotin-L-sulfoxide (g–i) in water (AMOEBA). For each time-step in each umbrella window, the number of interactions within a cut-off are counted. The interactions and the distance of the ion to the biotin ring are stored for each time-step and binned. The error bars correspond to the standard deviation of each bin.

interaction energy between chloride and the host is stronger than between water and the host (Fig. B.15). Both guests are centrally bound in the cavity. The electrostatic surface potential of the cavities in the different macrocycles varies (see Fig. 6.7), influencing water behaviour within the confined environment of the cavity. One can observe two distinct states: either the cavity is hydrated and there are water molecules within 3 Å of the cavity center, see Fig. 6.11(a) or dewetted and there are no water molecule within the cavity, see Fig. 6.11(b). In macrocycles with less polar cavities, the cavity is dewetted 10–15% of the simulation time. In the sulfide macrocycle, these transitions occur more frequently and are shorter-lived compared to D-sulfoxide. In the more polar L-sulfoxide macrocycle, the cavity is dewetted only 1% of the simulation time. When an ion enters the L-sulfoxide cavity, it competes with a relatively stable water molecule, explaining the weaker binding affinity of the L-sulfoxide host compared to D-sulfoxide and sulfide hosts. In simulations using the Parsley force field, the cavities of all three hosts are dewetted less than 0.1% of the simulation time. Thus, accurately simulating the behaviour of water in confined environments with varying polarity requires a polarisable force field.

Dewetting transitions can also be observed in narrow and hydrophobic ion channels as demonstrated for GlyR in Fig. 4.5 or for GLIC in Fig. 4.6, but they occur on a nanosecond timescale [19]. In contrast, dewetting transitions in smaller anion carriers occur on a picosecond timescale and are more dynamic than the dewetted state in larger ion channels with a similar radius of about 3 Å. In ion channels, changing the polarity of the pore lining can destabilize the dewetted state, as discussed in Chapter 4. Consequently, the cavity of the macrocycle is more likely to be hydrated in the polar L-sulfoxide version, while in the sulfide macrocycle with a nonpolar cavity, dewetting transitions occur more frequently (Fig. B.16).

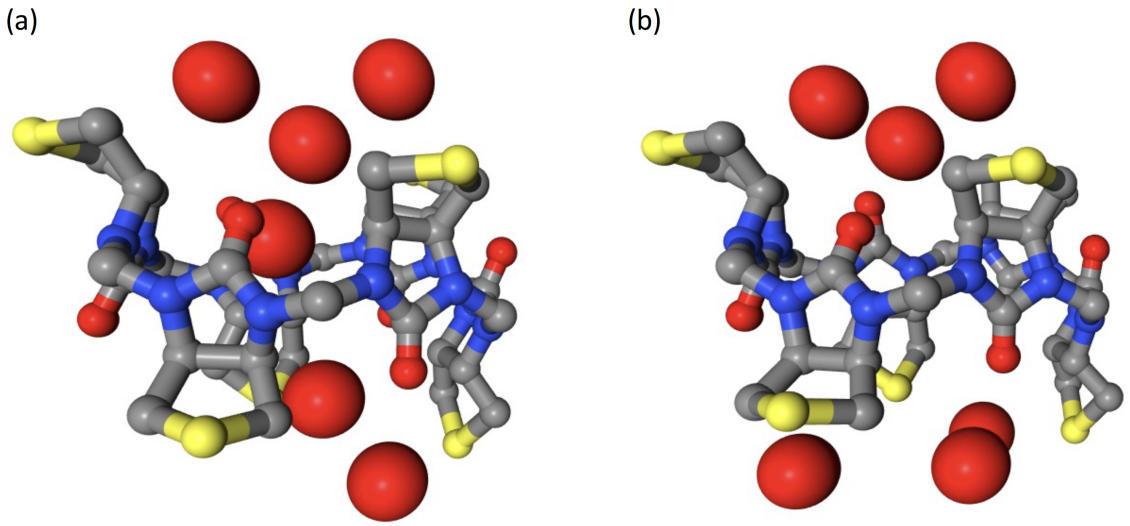


Figure 6.11: Observed hydration patterns: (a) wetted cavity and (b) dewetted cavity of biotin-sulfide macrocycle. The side chains of the biotin units and hydrogens are not shown.

6.4 Conclusion

In this chapter, binding free energies were calculated for biotin macrocycles and different anions in water and acetonitrile with *in silico* methods employing both non-polarisable and polarisable force fields. The anion binding showed strong solvent dependency, which is consistent with results from experimental studies [240, 241]. However, the free energy profiles obtained by PMF and the preferred binding positions of anions were found to depend on whether the force field can model the effect of multipoles and induced polarisation. Lisbjerg *et al*. reported that anion binding in the biotin[6]uril family is solvent-dependent [153]. Biotin[6]uril hexa-acid preferentially binds larger anions in water, while biotin[6]uril hexamethyl ester binds smaller anions in organic solvents such as acetonitrile and can transport anions across cell membranes [241]. The solvent's influence has direct functional implications: Biotin[6]uril hexamethyl ester will bind anions strongly inside a low dielectric lipid membrane environment for transport but binds weakly at a high dielectric lipid–water interface, facilitating spontaneous ion uptake and release. It also binds chloride more strongly in a nonpolar solvent than in water. Confirming the selectivity trend in two different solvents with different force fields is therefore encouraging. However, the preferred binding pose and the magnitude of the binding affinity are highly dependent on the force field type. Simulations with the polarisable force field AMOEBA overestimate the binding free energies, as summarized in Tables 6.1 and 6.2. The anionophore cavity is lined with 12 nonpolar CH-groups and is reportedly hydrophobic [240]. The interactions of anions with soft hydrogen bond donors in the cavity depend on the force field, and the partial charges of the CH-groups are fixed when using Parsley. In simulations with non-polarisable force fields, ions are not stably bound in the nonpolar cavity. They prefer configurations where the ion is 3.5–4 Å from the center of the ring, interacting with a more polar environment, as depicted in Fig. 6.2. The inclusion of multipoles and polarisation allows the AMOEBA force field to respond to changing heterogeneous molecular environments. The binding pose from AMOEBA simulations in Fig. 6.3(a) resembles the one reported in the crystal structure [239], with the ions encapsulated in the slightly hydrophobic cavity. Parametrisation processes can sometimes result in different effective radii. When comparing the VDW potential for chloride used in the Parsley and AMOEBA simulations in Fig. 2.3(c), it is clear that the Li *et al*. VDW radius is larger than the AMOEBA VDW radius. This difference could explain why chloride in non-polarisable force fields may prefer not to sit at the center of

the cavity. However, even after decreasing the VDW radius of chloride in the Parsley simulations, one still does not observe a stable binding pose in the cavity. Interestingly, fluoride, which has a smaller radius than chloride, also fails to form a stable binding pose in the cavity, suggesting that the instability is not due to the size of the ion. The cavity of both isomers of the biotin-sulfoxide macrocycle is more polar than that of the biotin[6]uril cavity. The differences in the free energy landscapes for the more polar macrocycles are less pronounced between non-polarisable and polarisable force fields compared to the more hydrophobic biotin[6]uril. The different force field frameworks predict entirely different binding poses for the more hydrophobic biotin[6]uril macrocycle. In contrast, the polarisable AMOEBA force field predicts the same binding pose for biotin-sulfoxide as the non-polarisable Parsley force field. For the L-isomer, the polarisable simulations indicate an additional binding pose centrally in the cavity that is unstable in simulations with the Parsley force field. Polarisable force fields are computationally more expensive than pairwise additive force fields (see Table B.1), but non-polarisable force fields cannot capture the effects of induced polarisation, where the electronic structure of an atom is altered by the varying distribution of charges in its environment. This chapter underscores the sensitivity of binding free energies and binding positions to the choice of force fields. While AMOEBA accurately captures binding poses, reproducing energetic trends in these systems requires further refinement. Ion binding in a nonpolar environment with soft hydrogen-bond donors serves as a useful case study, demonstrating that the higher computational costs of a polarisable framework are justified and provide a more accurate description than pairwise additive force fields. Additionally, the ease of performing these calculations will increase as more parts of the process become automated [246]. Overall, these results have strong implications not only for our understanding of anion-host interactions in synthetic ionophores but also in the narrow cavities of biological ion channels.

Chapter 7
Conductive States of the GlyR

7.1 Introduction

In the previous chapter, I discussed the computational trade-offs between polarisable and additive force fields, highlighting that while polarisable force fields are more computationally expensive, they are essential for accurately capturing the effects of induced polarisation. The effect of polarisation becomes particularly important in environments that are narrow and/or hydrophobic, such as the passage of chloride ions through channels. In these scenarios, the local electronic environment can vary strongly, making it an interesting use case for polarisable force fields to accurately model these interactions. In this chapter, I extend my analysis to a specific biological system by calculating the conductance and the free energy barriers associated with chloride ion permeation through the glycine receptor (GlyR). The energetic landscape is studied that governs chloride transport, providing insights into the molecular mechanisms underlying GlyR function. This chapter highlights the sensitivity of energetic barriers and of conductance estimates to force fields. In Chapter 5, I presented the PoreAnalyser software that provides a conductance estimate based on a pore profile. Here, the conductance model from eq. (5.6) is used to obtain an initial estimate which will be compared to conductance estimates from MD simulations with an external potential. The interactions of chloride with proteins and lipids is underestimated in many non-polarisable force fields as described in section 2.2.1 [114]. In the previous chapter, van der Waals parameters from Li *et al*. were used to describe chloride binding to the hydrophobic and narrow cavity of different macrocycles. Li *et al*. re-optimised the van der Waals parameters for the AMBER force field [113]. For the CHARMM force field [151], Orabi *et al*. formulated a series of pair-specific Lennard-Jones (NBFIX) parameters to describe the interaction of chloride with other residues properly [114]. In this chapter, the CHARMM36m force field [151] is used with NBFIX parameters for chloride [114] for non-polarisable force field simulations and the CHARMM Drude force field for simulations with a polarisable force field. In section 4.2.1, I delved into the functional annotation of the GlyR, focusing on its structural and functional characteristics. A key aspect of this annotation involves understanding the restrictions at specific positions within the receptor, notably the -2' and 9' positions, which are occupied by leucine and proline, respectively. These residues play crucial roles in the gating of the GlyR. The GlyR structures from Fig. 4.5 are in a non-conductive state. The minimal constriction at the P-2' position is smaller than twice the radius of a water molecule (dashed green line in Fig. 4.5). In this chapter, I want to investigate conductive conformations of GlyR.

7.2 Structure Selection

Available GlyR structures from the PDB [16] were screened and their pore radius profiles were examined, ensuring that the minimal constriction is larger than the radius of two water molecules. After the initial screen of GlyR structures, the structures solved by Kumar *et al.* [247] and Yu *et al.* [34] exhibited a large enough minimal constriction to theoretically allow chloride permeation. The pore profiles depicted in Fig. 7.1 identify the narrowest constriction at the P-2' position, with a secondary barrier at the L9' position. Kumar *et al.* provide models for GlyR in the presence of 1mM glycine and 32µM tetrahydrocannabinol, represented by PDB entries 7m6q and 7m6s respectively [247]. Additionally, I study models from Yu *et al.*, which includes GlyR bound to taurine (PDB entry 6pm2) and GlyR bound to glycine (PDB entry 6pm6)

[34]. Yu *et al.* isolated the GlyR using stryrene maleic acid polymer (SMA) and determined high-resolution structures of the receptor bound to glycine, taurine, or GABA in the SMA environment [34]. The analysis is focused on the 6pm2 and 6pm6 structures, and excludes the expanded-open state because the latter has a minimum diameter of 7 Å at the -2' position, and thus should be large enough to conduct isethionate. Isethionate is an organic anion with an unhydrated diameter of 6.2 Å [248], and cannot permeate pLGIC anionic channels [249]. Yu *et al.* therefore suggested that the expanded-open state was too large to be the physiologically relevant open state [34]. The PDB codes of the structures and their minimal constrictions are listed in Table 7.1. Dämgen and Biggin found that the open state structure solved by Du *et al.* with PDB code 3jae [250] rapidly collapsed in MD simulation [36]. Therefore, they developed an extended equilibration protocol to allow pore lining side chains to relax. I summarised the simulation protocol in section 3.3.2 and used it to stabilize an open structure of GLIC in section 4.2.2. The physiological relevance of the open state structure has sparked lively discussions [179, 251, 252]. This chapter will shed light on this debate by using more accurate force fields to describe chloride-protein interactions. To keep the simulation box small, I focus exclusively on the transmembrane domain (TMD) of GlyR. Omitting the extracellular domain (ECD) allows me to simulate a smaller box and hence reduces the complexity and computational expense of the simulations. The box vector in the z-direction could be reduced from 170 Å when simulating the whole GlyR structures to 120 Å when omitting the ECD. Simulating not the whole receptor has been recently criticised [252] as the channel conductance is determined by the rate-limiting step of the permeation process. In the open state of the GlyR, the rate-limiting step could be in the intracellular domain (ICD) or the ECD. For the specific case of the α1 homomeric GlyR, the rate-limiting step seems to lie within the TMD, because Moroni *et al.* found that the removal of the ICD does not seem to affect single-channel conductance [253]. Moraga-Cid *et al.* found that replacing the ECD with the ECD from GLIC displays the same conductance as the full-length wild-type GlyR α1, including a number of sub-conductive states [254].

7.3 Pore Analysis

The GlyR channels were embedded in a lipid bilayer, following the simulation protocols previously described in section 3.3. 50 ns of equilibrium simulations were conducted with the CHARMM36m force field [151] and

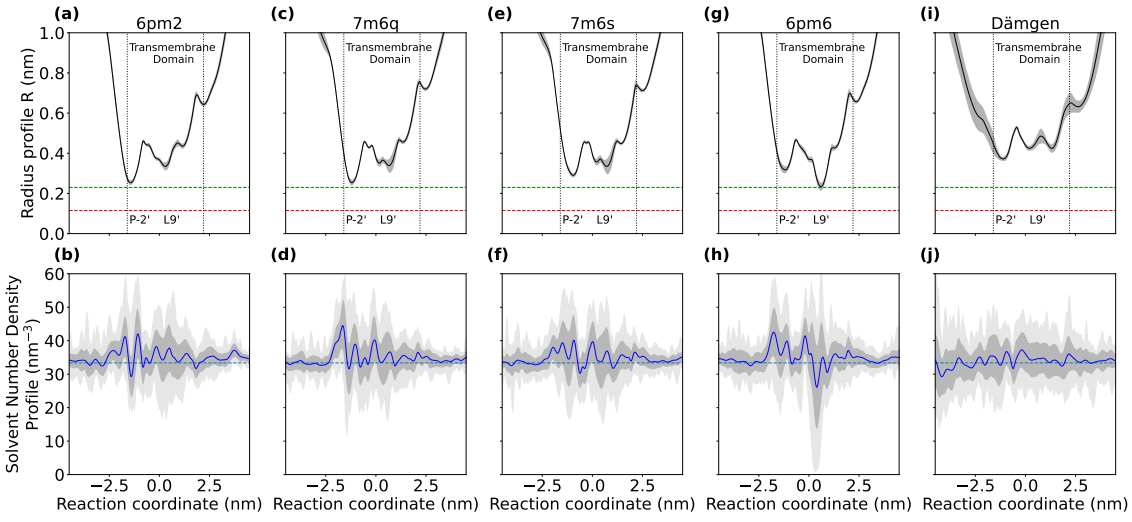


Figure 7.1: Radius profiles (a, c, e, g, i) and solvent number density (b, d, f, h, j) for 50 ns simulations of five GlyR models. The red and the green dashed horizontal lines indicate the radius of one or two water molecules respectively. The blue dashed lines mark the number density of bulk water (33 nm^{-3}).

pair specific Lennard-Jones parameters for chloride [114] with small backbone restraints to assess whether the pores were wetted. As described in section 4.1.3, wetting of the pores is an important prerequisite for subsequent analyses as it indicates that the ion channel is in a state suitable for conducting further detailed studies, such as free energy calculations or simulations with an external electric field. The results of these simulations revealed that all examined GlyR pores were wetted, confirming their potential for ion conduction. Specifically, the minimal pore radius in each structure were larger than twice the radius of a water molecule - the green dashed line in Fig. 7.1. This indicates that the channels are sufficiently wide to support ion flow. Moreover, the water number density was evaluated along the channel axis over the 50 ns simulation period. The average water density in the channel is found to be close to the bulk water number density, which is 33 nm^{-3}. Notably, the density exhibits higher variation at narrower regions of the channel, such as in the L9' position in the 6pm6 structure, reflecting the dynamic behaviour of water molecules in these constricted areas. Detailed pore profiles and wetting characteristics are illustrated in Fig. 7.1.

7.4 Heuristic for Conductance

After equilibrium MD simulations suggested that the chosen GlyR models were indeed wetted, their conductance was estimated with the PoreAnalyser software before running computationally expensive MD simulations with an external potential. The PoreAnalyser software estimates the single channel conductance based on a radius profile with the cylindrical approximation, see eq. (5.6). Fig. 7.2 depicts the PoreAnalyser radius profile obtained with ellipsoidal probe particles. One can relate the radius profile with a resistance and conductance using the model in eq. (5.6). When using the conductivity model for ions within confined environments, the conductance model in eq. (5.6) predicts conductance ranging from 38 to 88 pS. If one uses the bulk conductivity, eq. (5.5), in the cylindrical approximation, the conductance estimates range from 240 to 411 pS. Single-channel studies by Borman *et al.* have shown that GlyRα1 exhibits five conductance states ranging from 18 pS to 90 pS [255]. Using the bulk conductivity in the cylindrical approximation results

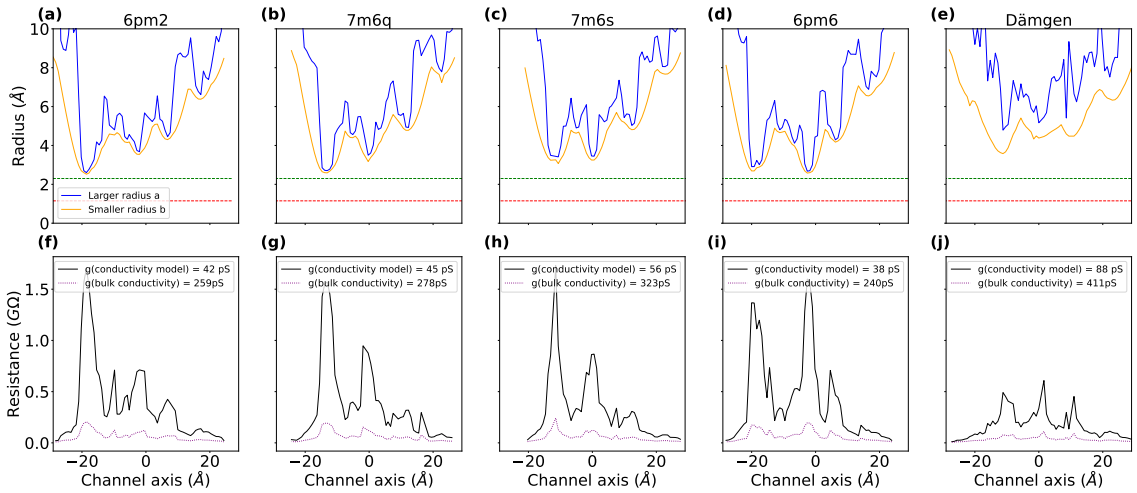


Figure 7.2: Conductance heuristic. The PoreAnalyser radius profiles in (a)-(e) can be used to obtain an estimate for the resistance or conductance of the channel as described in Chapter 5. The red and the green dashed horizontal lines indicate the radius of one or two water molecules respectively. The blue line marks the radius of the larger axis of an ellipsoidal probe particle, the orange line marks the radius of the smaller axis. The narrowest constriction at the P-2' and L9' position have the largest contribution to the resistance across the channel axis (bottom row). To relate radius profile and conductance in (f)-(j), one can either use the bulk conductivity (dotted purple line) or the conductivity model for ions in confined environments (solid black line).

in overestimation of the conductance as already described in Chapter 5. The conductivity model was trained on CNT model systems with potassium and chloride ions. Here, sodium was used instead of potassium. Therefore, the diffusion coefficient D_K in eq. (5.5) was replaced with D_{NA}. Given that GlyR channels are chloride selective, I do not expect this to be a severe limitation of the heuristic. The PoreAnalyser conductance prediction for the 6pm6 structure solved by Yu *et al.* [34] was the lowest among the five structures in Fig. 7.2 and 6pm6 is characterised by a constriction at the L9' position that is equal to the radius of two water molecules and by a secondary constriction at the P-2' position. The equilibrated structure by Dämgen and Biggin [36] had the widest pore and hence the highest PoreAnalyser conductance estimate.

7.5 Conductance Estimates with an External Potential

After having established that the GlyR pores are wetted, a +500 mV transmembrane potential difference was applied with the positive potential on the cytoplasmic side to observe chloride influx. I used the CHARMM36m force field with pair specific Lennard Jones parameters and the CHARMM Drude force field and counted the number of permeation events in 250 ns long simulations. The Drude Prepper [155] within CHARMM-GUI was used to convert the simulation boxes with CHARMM36m force field [151] parameters to those compatible with the CHARMM Drude polarisable force field [84]. Fig. 7.3 compares the conductance estimates based on simulations with the two force fields and the PoreAnalyser conductance. The errors on the ionic currents were estimated by assuming a Poisson distribution for permeation events, as described in section 3.8 [157]. The variance equals the mean in the Poisson distribution and therefore, the uncertainty increases at high voltages such as 500 mV, when more permeation events are expected to be observed. This is reflected in the large error bars in Fig. 7.3. All conductance estimates are within the experimentally determined range from 18 pS to 90 pS [255]. There is no observable general trend between polarisable and non-polarisable force fields. The Drude conductance is higher than the CHARMM36m conductance for the 7m6s structure and lower for the 6pm2 structure. For the other three structures, the Drude and CHARMM36m conductance are within error. For the 7m6q structure and the structure refined by Dämgen and Biggin, all three estimates lie within error of each other. Consistent with the PoreAnalyser conductance

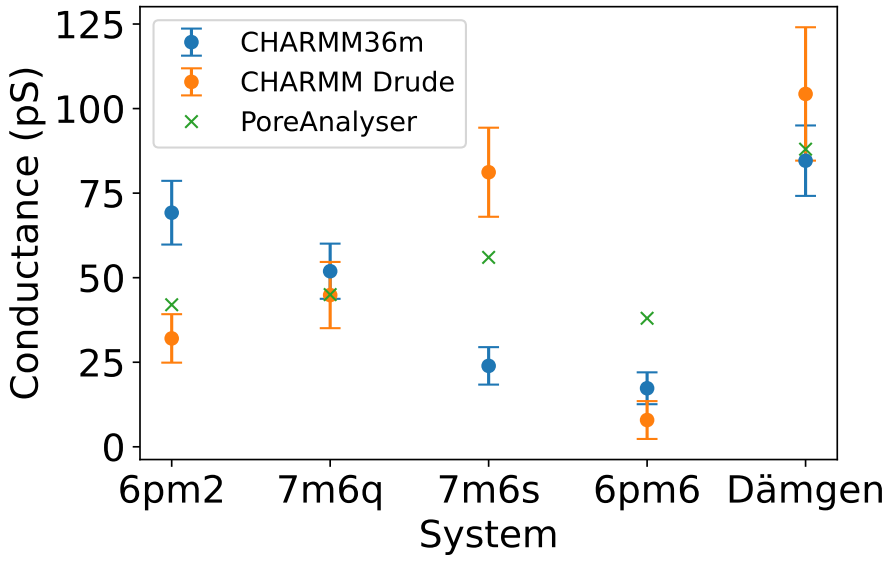


Figure 7.3: Conductance estimates based on PoreAnalyser profiles and MD simulations with two force fields for the five GlyR structures in Table 7.1.

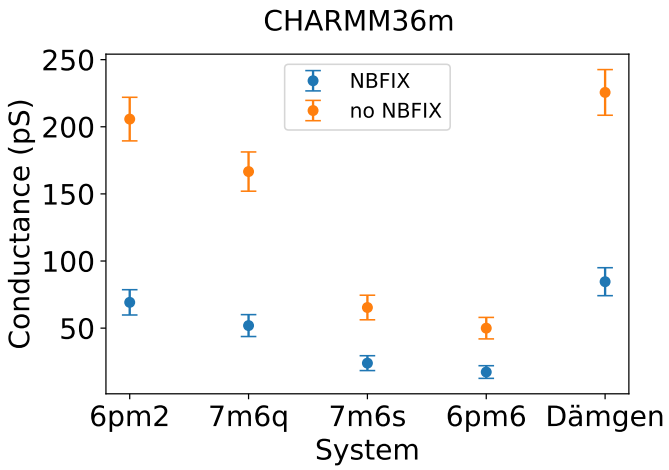


Figure 7.4: Conductance comparison: Conductance estimates for five GlyR structure based on simulations with an external electric field with the pairwise additive CHARMM36m force field [151] with and without NBFIX parameters [114].

estimates, the MD conductance is largest for the wide pore of the Dämgen structure and smallest for the narrow 6pm6 pore. Rao *et al.* suggested that MD simulations of GlyRs were able to indicate relative but not absolute conductance values [256]. Backbone restraints of 500 kJ/nm^2/mol were used to maintain the conformation of the structures in the bilayer during the 250 ns simulation time. The five structures only represent a subset of the whole ensemble of conductive GlyR structures. Therefore, the conductance estimates cannot reproduce the correct thermodynamic properties of the physiological open state [252].

7.5.1 Force Field Sensitivity

Conductance estimates of the GlyR are sensitive to the water model used [252]. Cerdan *et al.* found that TIP4P predicts a much lower conductance for the wide-open structure (PDB code 3jae) of 228 pS compared with 413 pS when using TIP3P [179, 252]. Here TIP3P was used for the simulations with CHARMM36m and SWM4-NDP for the CHARMM Drude force field. In previous GlyR studies, Cerdan *et al.* used the CHARMM27 [257] force field [179] and the CHARMM36m force field (without NBFIX parameters) [251] while Dämgen and Biggin used the AMBER99SB-ILDN [82] force-field [36]. Cerdan and Cecchini used the CHARMM36m force field [151] to evaluate the Dämgen and Biggin structure [251]. They employed computational electrophysiology [160] on a reduced molecular system composed of the transmembrane domain only in a lipid bilayer and harmonic restraints on the heavy atoms [251]. Cerdan and Cecchini computed a conductance of 382–480 pS, which they argue is too large to correspond to the physiologically relevant open state [251]. However, conductance estimates are very force field sensitive and Cerdan and Cecchini did not use pair-specific Lennard-Jones parameters for chloride. As discussed before, the native chloride parameters in the CHARMM36m force field underestimate the interactions of chloride with non-polar residues [114] and chloride exhibits more "friction" when simulated with NBFIX parameters or with a polarisable force field. Fig. 7.4 compares conductance estimates based on simulations with the CHARMM36m force field with and without NBFIX parameters. Not using NBFIX parameters increases the conductance by more than a factor of two to three. In Table 7.1, all conductance estimates are summarised.

Table 7.1: Summary of simulated GlyR structures with their respective pore constrictions and conductance estimates.

Structure	6pm2	6pm6	7m6q	7m6s	Dämgen
authors	Yu *et al.* [34]		Kumar *et al.* [247]		Dämgen and Biggin [36]
PDB constriction [Å]	2.5	2.6	2.6	3.2	3.6
g(PoreAnalyser) [pS]	42	38	45	56	88
g(CHARMM36m with NBFIX) [pS]	69±9	17±5	52±8	24±6	85±10
g(CHARMM36m) [pS]	206±16	50±8	167±15	65±9	226±17
g(Drude) [pS]	32±7	8±6	45±10	81±13	104±20

7.5.2 Orientation of Side Chains

Dämgen and Biggin found that the -2' Pro gate region is highly dynamic in their MD simulations and exhibited higher than usual root-mean-square fluctuation values for the C_α atoms in that region [36]. The cryo-EM map from Du *et al.* exhibits poor density in that region [250, 36]. The extended equilibration protocol from section 3.3.2 was developed by Dämgen and Biggin to allow pore facing side chains such as proline and leucine at the -2' and 9' position to move to hydrophobic pockets by introducing flat bottom restraints on C_α atoms to keep the pore open during equilibration. In the simulations, the movement of the backbone atoms was restrained and only side chains were allowed to move freely. In Fig. B.22 in the appendix, I check whether leucine sits in the hydrophobic pocket or is pore facing. In simulations with the CHARMM Drude force field, the side chains of the leucine residues at the 9' position were oriented more towards their hydrophobic pockets than to the pore as depicted in Fig. B.23. This might have contributed to the higher conductance observed for simulations with the CHARMM Drude force field for the 7m6s structure in Fig. 7.3.

7.5.3 External Potential and Drude Particles

In Fig. B.24 in the appendix, the distribution of the distance between chloride nucleus and Drude particle is plotted for simulations with and without an external potential and I find that there is no difference in the distributions between these simulations. This serves as a sanity check that the external potential does not overly polarise the chloride ions. The final bin at 0.2 Å of the distribution is more populated because of the *drude hardwall* parameter that is set to the default 0.2 Å. When the Drude particles move far away from the nucleus, simulations become unstable and the so-called polarisation catastrophe may occur, as described in section 2.2.2. Therefore, a hard wall constraint on the length of Drude-nuclei bonds is introduced to avoid such instability. As previously described in section 2.2.1, the length of Drude bonds are checked and compared with *drude hardwall* in each MD step. When the length of a Drude bond is larger than *drude hardwall*, the relative velocities along bond vector are flipped and scaled down according to the temperature of the Drude bond [123].

7.6 Potential of Mean Force and Hydration Profile

In Chapter 5, the barrier of permeation was calculated for potassium and chloride and the conductance of carbon nanotubes and observed the expected negative correlation between these two quantities. To understand the energetics of ion permeation through the GlyR pore, umbrella sampling simulations were employed to calculate the potential of mean force (PMF) as described in section 3.5. This should allow us to compare the free energy profile of ion movement along the pore axis for simulations employing different force fields. Good overlap was observed between neighbouring umbrella sampling windows as

depicted in Fig. B.18 in the appendix. Convergence is analysed in Fig. B.19 in the appendix. In the PMFs, one can observes different energy profiles for the Drude polarisable force field and the non-polarisable CHARMM36m force field with pair-specific Lennard-Jones parameters (NBFIX) [151, 114]. Chloride anions partially lose their hydration shell when passing through the GlyR pore, see Fig. B.17 in the appendix. The PMF profiles in Fig. B.17 show distinct barrier characteristics at critical points along the pore. The proline (P) at the -2' position marks the narrowest constriction along the permeating path and the leucine (L) at the 9' position marks a secondary barrier. The energetic barrier associated with the constrictions at P-2' and L9' are within error of each other for PMFs based on the CHARMM36m force field for the 6pm2, 7m6q and 7m6s structures. Chloride ions described by the CHARMM36m force field have to overcome two barriers of similar height when they permeate the GlyR pore.

The barrier associated with the constriction P-2' is smaller for the Drude PMF compared to the CHARMM36m PMF for structures 6pm2 and 7m6q. For 7m6s, depicted in Fig. B.17, the barrier associated with P-2' is within error for both force fields. The Drude force field is computationally more expensive and each umbrella sampling window was run for only 9 ns instead of 21 ns. Convergence is analysed in Fig. B.20. To unbias the windows with the umbrella bias with WHAM, a cyclic reaction coordinate was used. The first and the last bin of the reaction coordinate, both in the bulk region, were assumed to be neighbours. However, the Drude PMFs in Fig. B.17 are not flat in the bulk-water regions. When I calculate a non-periodic PMF, I observe an offset between the two bulk water regions. The offset is equally distributed along the reaction coordinate to enforce a periodic PMF resulting in the non-flat PMF in the bulk region. The non-flat PMFs in Fig. B.17 indicate that the Drude PMFs are not converged and that 9 ns of simulation time is not enough sampling time per umbrella window. Hub *et al.* suggest to re-weigh the terms in eq. (2.62) with a term proportional to the integrated autocorrelation time of each umbrella window [138]. In a protein pore, the integrated autocorrelation times are higher than the ones in bulk [138]. Increasing the simulation time per window for the Drude systems was computationally not feasible at the time of writing and hence preliminary PMF results can be found in the appendix.

The distribution of the number of water-hydrogens and protein atoms coordinating the chloride ion at the proline -2' and the leucine 9' constriction is depicted in Fig. B.21 in the appendix. Chloride is heavily coordinated by proline at the -2' position in simulations with CHARMM36m and moderately coordinated in simulations with the CHARMM Drude force field for structures 6pm2 and 7m6q, see Fig. B.21(a) and (c). Apart from the behaviour at the -2' position for structures 6pms and 7m6q, there does not seem to be a large difference between the CHARMM36m and Drude force field in terms of chloride coordination at the narrowest constriction. The 6pm2 and 7m6q structures exhibit more chloride friction than the wider 7m6s structure. The PoreAnalyser radius profile in Fig. 7.2 shows that the elliptical radii coincide at the constriction points and that the -2' and 9' position is symmetric. No large difference between CHARMM36m and CHARMM Drude is observe because the 7m6s pore is sufficiently wide to accommodate a hydrated chloride ion and not much dehydration is needed for permeation.

The differences in barrier height predicted by the different force fields highlight the sensitivity of predicting free energy barriers to the choice of the Hamiltonian. If one compares the interactions of chloride described with NBFIX parameters in the CHARMM36m and the polarisable Drude force field in Fig. B.21 in the appendix, one does not see increased friction or increased interaction of chloride at the proline gate at the -2' position. Chloride ions exhibit high affinity to uncharged residues and both force fields seem to capture this characteristic well.

The discrepancy between chloride conductance in simulations with and without NBFIX parameters depicted in Fig. 7.4 indicate that a polarisable force field might not be necessary to observe increased friction between chloride and pore lining residues. Using NBFIX parameters in a non-polarisable force field

or employing a polarisable force field both result in conductance estimates within the expected range. The CHARMM Drude force field is computationally more expensive because of a smaller time step (0.001 ps) and a higher number of particles in the same simulation box. This limited the sampling time for the umbrella windows and resulted in convergence problems. The barrier associated with the proline gate is lower in the CHARMM Drude force field than in the non-polarisable force field, see Fig. B.17. However, the Drude PMFs are not converged yet and further investigation and more simulation time is needed.

7.7 Conclusion

In this chapter, GlyR structures were analysed where the minimal constriction of the pore was larger than the radius of two water molecules. The simulation protocol was employed for functional annotation of ion channels, as previously outlined in Chapter 4. The analysis began with an initial radius profile assessment, followed by equilibrium MD simulations, which indicated that the pores remained wetted throughout the simulations. This observation suggests that the structures most likely represent conductive states of the channels. Furthermore, the radius profile generated by the PoreAnalyser software was utilised to make a heuristic estimate of the channel conductance as described in Chapter 5. The estimated conductance values were found to fall within the range of conductance values determined through experimental methods by Borman *et al.* [255], reinforcing the validity of my approach. Backbone restraints were used to maintain the conformation of the five selected structures. Therefore, the conductance estimates cannot represent the whole ensemble of conductive GlyR structures and cannot reproduce the correct thermodynamic properties of the physiological open state [252].

The effects of different force fields on the simulations was also explored. These simulations revealed that the energetic barrier to permeation and conductance estimates were sensitive to the choice of force field. As discussed before, some non-polarisable force fields systematically underestimate the interaction of chloride with proteins and lipids [114]. For the AMBER force field family, Li *et al*. re-fitted the Lennard-Jones parameters of anions [113] and for CHARMM36m Orabi *et al*. developed pair-specific Lennard-Jones parameters. Cerdan *et al*. found that GlyR conductance estimates were sensitive to the water model too: They observed a much lower conductance for the wide-open structure (PDB code 3jae) of 228 pS with simulations employing TIP4P compared with 413 pS when using TIP3P [179]. Their conductance estimates are unphysiologically large and obtained with MD simulations employing the CHARMM force field family without correcting for the underestimation of chloride interactions with proteins [179, 251]. As depicted in Fig. 7.4, using the NBFIX parameters for chloride increases the friction or interaction of the permeating chloride with the pore and decreases MD conductance estimates to physiological values. Overall, these findings underscore the importance of carefully selecting the appropriate force field for MD simulations in the study of ion channel conductance and highlights the need for further analysis to clarify the discrepancies between barriers to permeation and conductance observed with the Drude force field.

Chapter 8
Conclusion and Outlook

8.1 Summary of Results

In this thesis, I explored the functional annotation of ion channels through simulation-free methods and molecular dynamics simulations. Using software tools such as CHAP and PoreAnalyser, I characterised pore profiles and estimated conductance based on a cylindrical approximation. This approach was applied to newly solved ion channel structures, such as GlyR and GLIC in Chapter 4, demonstrating their utility in functional annotation. MD simulations provided additional insights, particularly for narrow constrictions and hydrophobic environments, to check whether or not the pores were wetted. The concept of hydrophobic gating was discussed extensively in Chapter 4; even in the absence of a steric occlusion, dewetted regions can present major energetic barriers to ion passage.

Furthermore, the limitations of non-polarisable force fields in accurately describing ions in narrow and hydrophobic environments were highlighted. This was particularly evident in the study of macrocycles and anions, where different force fields predicted different binding poses in Chapter 6. Limitations of commonly used force fields to accurately describe ions in narrow and hydrophobic environments have implications for MD simulations of narrow ion channels. Polarisable force fields are computationally more expensive than pairwise additive force fields. In Chapter 7, I demonstrated that a non-polarisable force field with improved ion parameters can yield conductance estimates in the physiologically expected range. Pair-specific Lennard-Jones parameters [114] were used to circumvent the problem that many non-polarisable force fields underestimate interactions between anions and protein residues.

In most chapters, different GlyR structures were used as model systems. The asymmetric GlyR structures in Chapter 4 motivated the investigation of the influence of pore shape and asymmetry on conductance and permeation barriers in Chapter 5. For future research, simulations with polarisable force fields are promising for accurately describing the energetic barriers faced by chloride ions permeating a GlyR pore as suggested in Chapter 7.

The methodologies and best practices described in this thesis are intended to guide those unfamiliar with MD simulations, enabling them to perform functional annotation of ion channels effectively. To this end, I developed the user-friendly PoreAnalyser software, which can be easily accessed via a pip-installable Python package or a graphical user interface on a web service. I hope that these tools will facilitate broader adoption and application of MD simulations in ion channel research and beyond.

8.2 Future Directions

8.2.1 Conductance Estimates

To observe permeation events in MD simulations of ion channels, I applied external electric fields of 500 mV that are much higher than physiological relevant potential differences. These artificially high fields enhanced ion permeation rates within the limited timescales of simulations. However, this approach introduces inaccuracies, as it does not reflect the true membrane potential found in biological systems. The simulation times of 250 ns are generally too short to produce realistic permeation events. Consequently, relying on artificially elevated electric fields to speed up simulations can lead to misleading conclusions about

the behavior of ion channels. Instead of relying on the potential across the simulation box, computational electrophysiology techniques could be employed. These techniques, as developed by Kutzner and colleagues, provide a more accurate representation of the electrostatic environment within the simulation [160, 161].

An alternative to using external potentials in simulations is to estimate channel conductance by calculating the PMF and diffusion coefficients along the reaction coordinate. Instead of directly simulating ion flux under an external field, one could solve the Langevin equation, which accounts for the spatially varying free energy, represented by the PMF, and the spatially varying diffusion coefficient D [67]. In the simplest scenario of ion diffusion within a bulk medium, the diffusion coefficient D can be derived from the mean square displacement (MSD) using the Einstein relation in multiple dimensions. Alternatively, D can be determined through the autocorrelation of ion velocity [258]. Both of these methods assume that the PMF is flat, meaning that the ion is undergoing Brownian motion without major energy barriers. However, when the PMF is not flat, ions can become trapped in local energy minima, limiting the MSD and causing the velocity autocorrelation to vanish [67]. In such cases, obtaining an accurate diffusion coefficient requires separating the random, thermal forces from the systematic forces generated by the PMF. This separation is achieved by calculating a position-dependent diffusion coefficient $D(r)$ that is unaffected by the underlying energy landscape. In this thesis, I focused on the diffusion coefficient of ions permeating through carbon nanotubes. Studies by Mendonça and colleagues provided insights into the behaviour of water in confined environments, which is relevant to understanding ion diffusion in similar conditions [225, 226]. Furthermore, Zhu and Hummer have applied a Smoluchowski diffusion model to calculate single-channel conductance at low ion concentration using the one-dimensional PMF for ion passage through the GLIC ion channel [259].

In Chapter 5, a heuristic conductivity model was fitted to obtain an estimate of resistance and conductance of channels based on their pore profile. The bulk conductivity in eq. (5.5) is proportional to the bulk diffusion coefficient of the permeating ions. The conductivity of ions in constrained environments was fitted to match MD conductance estimates. Instead of fitting the conductivity model, one could calculate a diffusion coefficient $D(a, b)$ that depends on the constriction defined by radii a and b based on the PMF data. The diffusion coefficient $D(a, b)$ could be converted to a conductivity with eq. (5.5). By comparing the conductivities obtained from a fitted model to those derived from the PMF, one could assess the validity of the model.

8.2.2 Force Fields

The accuracy and reliability of MD simulations heavily depend on the force fields used, which are empirical models describing the potential energy of a system of particles. There is a general trade-off between accuracy and computational efficiency. For instance, some non-polarisable force fields systematically underestimate the interaction of chloride with proteins and lipids [114]. Therefore, Li *et al*. re-fitted the Lennard-Jones parameters of anions [113] for the AMBER force field family which was used in Chapter 6. Orabi *et al*. developed pair-specific Lennard-Jones parameters for CHARMM36m [114] which were used in Chapter 7. Instead of changing van der Waals interactions, one can also employ polarisable models. In Chapter 6, only simulations with the polarisable AMOEBA force field reproduced the experimentally determined binding pose of anions in a narrow and hydrophobic cavity. In Chapter 7, the polarisable CHARMM Drude and the non-polarisable CHARMM36m with NBFIX parameters both predicted conductance values within the physiological range. Polarisable force fields attempt to account for the response of the electron cloud to changes in the environment, offering improvements over classical pairwise additive force fields. However, polarisable force fields still do not fully capture the complexity of chemical interactions, such as electron transfer, highlighting the limitations of purely classical approaches in modeling certain types of molecular behaviour [260]. One approach to improve physical accuracy and applicability is the

use of hybrid methods. Hybrid QM/MM methods have become increasingly popular for studying chemical reactivity within biomolecular systems [85]. The chemically reactive part of the system is treated with high-level quantum mechanics, while the rest of the system is described using classical molecular mechanics. While these methods enhance the physical accuracy of simulations, they are computationally expensive. This cost has limited their application, especially in calculating free energies, for which extensive sampling is necessary. To address these challenges, researchers have developed reference potentials and mean-field approximations to accelerate QM/MM calculations, thereby making these high-accuracy methods more feasible for practical use [85]. In the specific case of modeling water, which is a crucial component in many biological and chemical systems, all commonly used water models exhibit limitations. For instance, models such as the widely used SPC and TIP3P have been shown to have issues such as elevated self-diffusion rates, incorrect static dielectric constants, and overestimation of bulk potential energy [67]. Higher-order models, such as TIP4P, offer improved accuracy in representing water's physical properties but come with increased computational costs and may require shorter time steps. Furthermore, the consistency between these water models and the protein force fields, that are used in combination with them, can be problematic. For example, CHARMM and AMBER force fields are parameterised to be used with the 3-site TIP3P water model [151, 82], while GROMOS force fields are designed for use with SPC, and the TIP4P model is frequently paired with the OPLS-AA force field [230]. These combinations must be carefully chosen to ensure accurate and consistent results across different simulations. Overall, the choice of force field is a delicate balance between accuracy, computational efficiency, and the specific requirements of the system that is being studied.

Bibliography

[1] R. Phillips, J. Kondev, and J. Theriot, *Physical Biology of the Cell*. New York: Garland Science, Taylor & Francis Group, Nov. 2008.

[2] J. D. Firth, C. P. Conlon, T. M. Cox, F. Ashcroft, and P. Tammaro, "Ion channels and disease," in *Oxford Textbook of Medicine*, Oxford University Press, Mar. 2020.

[3] B. Hille, "Ionic channels in excitable membranes. Current problems and biophysical approaches," *Biophysical Journal*, vol. 22, pp. 283–294, May 1978.

[4] P. Aryal, F. Abd-Wahab, G. Bucci, M. S. Sansom, and S. J. Tucker, "Influence of lipids on the hydrophobic barrier within the pore of the TWIK-1 K2P channel," *Channels*, vol. 9, pp. 44–49, Jan. 2015.

[5] B. Corry, "Understanding ion channel selectivity and gating and their role in cellular signalling," *Mol. BioSyst.*, vol. 2, pp. 527–535, 2006.

[6] M. H. Saier, "Transport protein evolution deduced from analysis of sequence, topology and structure," *Current Opinion in Structural Biology*, vol. 38, pp. 9–17, 2016.

[7] J. H. Morais-Cabral, Y. Zhou, and R. MacKinnon, "Energetic optimization of ion conduction rate by the k+ selectivity filter," *Nature*, vol. 414, pp. 37–42, Nov. 2001.

[8] D. Kozono, M. Yasui, L. S. King, and P. Agre, "Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine," *The Journal of Clinical Investigation*, vol. 109, pp. 1395–1399, June 2002.

[9] S. Uysal, L. G. Cuello, D. M. Cortes, S. Koide, A. A. Kossiakoff, and E. Perozo, "Mechanism of activation gating in the full-length KcsA K+ channel," *Proceedings of the National Academy of Sciences*, vol. 108, no. 29, pp. 11896–11899, 2011.

[10] R. B. Stockbridge, L. Kolmakova-Partensky, T. Shane, A. Koide, S. Koide, C. Miller, and S. Newstead, "Crystal structures of a double-barrelled fluoride ion channel," *Nature*, vol. 525, pp. 548–551, Sept. 2015.

[11] E. Gibbs, E. Klemm, D. Seiferth, A. Kumar, S. L. Ilca, P. C. Biggin, and S. Chakrapani, "Conformational transitions and allosteric modulation in a heteromeric glycine receptor," *Nature Communications*, vol. 14, p. 1363, Mar. 2023.

[12] W. Kopec, B. S. Rothberg, and B. L. de Groot, "Molecular mechanism of a potassium channel gating through activation gate-selectivity filter coupling," *Nature Communications*, vol. 10, p. 5366, Nov. 2019.

[13] R. K. Skitchenko, D. Usoltsev, M. Uspenskaya, A. V. Kajava, and A. Guskov, "Census of halide-binding sites in protein structures," *Bioinformatics*, vol. 36, pp. 3064–3071, 02 2020.

[14] Z. Yue, Z. Wang, and G. A. Voth, "Ion permeation, selectivity, and electronic polarization in fluoride channels," *Biophysical Journal*, vol. 121, pp. 1336–1347, Apr. 2022.

[15] D. T. Gomez, L. R. Pratt, D. N. Asthagiri, and S. B. Rempe, "Hydrated Anions: From Clusters to Bulk Solution with Quasi-Chemical Theory," *Accounts of Chemical Research*, vol. 55, pp. 2201–2212, Aug. 2022.

[16] S. K. Burley, H. M. Berman, C. Bhikadiya, C. Bi, L. Chen, L. Di□Costanzo, C. Christie, K. Dalenberg, J. M. Duarte, S. Dutta, Z. Feng, S. Ghosh, D. S. Goodsell, R. K. Green, V. Guranović, D. Guzenko, B. P. Hudson, T. Kalro, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Periskova, A. Prlić, C. Randle, A. Rose, P. Rose, R. Sala, M. Sekharan, C. Shao, L. Tan, Y.-P. Tao, Y. Valasatava, M. Voigt, J. Westbrook, J. Woo, H. Yang, J. Young, M. Zhuravleva, and C. Zardecki, "RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy," *Nucleic Acids Research*, vol. 47, pp. D464–D474, Oct. 2018.

[17] O. Beckstein, P. C. Biggin, and M. S. P. Sansom, "A Hydrophobic Gating Mechanism for Nanopores," *The Journal of Physical Chemistry B*, vol. 105, pp. 12902–12905, Dec. 2001.

[18] P. Aryal, F. Abd-Wahab, G. Bucci, M. S. P. Sansom, and S. J. Tucker, "A hydrophobic barrier deep within the inner pore of the TWIK-1 K2P potassium channel," *Nature Communications*, vol. 5, p. 4377, July 2014.

[19] D. Seiferth, P. C. Biggin, and S. J. Tucker, "When is a hydrophobic gate not a hydrophobic gate?," *Journal of General Physiology*, vol. 154, p. e202213210, Nov. 2022.

[20] A. Gharpure, C. M. Noviello, and R. E. Hibbs, "Progress in nicotinic receptor structural biology," *Neuropharmacology*, vol. 171, p. 108086, July 2020.

[21] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon, "The structure of the potassium channel: Molecular basis of k+ conduction and selectivity," *Science*, vol. 280, no. 5360, pp. 69–77, 1998.

[22] Y. de Koninck, "Altered chloride homeostasis in neurological disorders: a new target," *Current Opinion in Pharmacology*, vol. 7, pp. 93–99, Feb. 2007.

[23] C. Farinha and I. Callebaut, "Molecular mechanisms of cystic fibrosis – how mutations lead to misfunction and guide therapy," *Bioscience Reports*, vol. 42, p. BSR20212006, July 2022.

[24] N. Busschaert and P. A. Gale, "Small□molecule lipid□bilayer anion transporters for biological applications," *Angewandte Chemie International Edition*, vol. 52, pp. 1374–1382, Jan. 2013.

[25] C. Ren, X. Ding, A. Roy, J. Shen, S. Zhou, F. Chen, S. F. Yau Li, H. Ren, Y. Y. Yang, and H. Zeng, "A halogen bond-mediated highly active artificial chloride channel with high anticancer activity," *Chemical Science*, vol. 9, no. 17, pp. 4044–4051, 2018.

[26] M. A. Yawer, V. Havel, and V. Sindelar, "A Bambusuril Macrocycle that Binds Anions in Water with High Affinity and Selectivity," *Angewandte Chemie*, vol. 127, pp. 278–281, Jan. 2015.

[27] J. L. Trick, S. Chelvaniththilan, G. Klesse, P. Aryal, E. J. Wallace, S. J. Tucker, and M. S. P. Sansom, "Functional Annotation of Ion Channel Structures by Molecular Simulation," *Structure*, vol. 24, pp. 2207–2216, Dec. 2016.

[28] S. Yuan, S. Filipek, and H. Vogel, "A Gating Mechanism of the Serotonin 5-HT3 Receptor," *Structure*, vol. 24, pp. 816–825, May 2016.

[29] P. V. Plazas, M. J. De Rosa, M. E. Gomez-Casati, M. Verbitsky, N. Weisstaub, E. Katz, C. Bouzat, and A. B. Elgoyhen, "Key roles of hydrophobic rings of TM2 in gating of the α9α10 nicotinic cholinergic receptor," *British Journal of Pharmacology*, vol. 145, no. 7, pp. 963–974, 2005.

[30] E. Zarkadas, E. Pebay-Peyroula, M. J. Thompson, G. Schoehn, T. Uchański, J. Steyaert, C. Chipot, F. Dehez, J. E. Baenziger, and H. Nury, "Conformational transitions and ligand-binding to a muscle-type nicotinic acetylcholine receptor," *Neuron*, vol. 110, pp. 1358–1370.e5, Apr. 2022.

[31] F. Zhu and G. Hummer, "Pore opening and closing of a pentameric ligand-gated ion channel," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 19814–19819, Nov. 2010.

[32] A. Miyazawa, Y. Fujiyoshi, and N. Unwin, "Structure and gating mechanism of the acetylcholine receptor pore," *Nature*, vol. 423, pp. 949–955, June 2003.

[33] N. Bharambe, Z. Li, D. Seiferth, A. M. Balakrishna, P. C. Biggin, and S. Basak, "Cryo-em structures of prokaryotic ligand-gated ion channel glic provide insights into gating in a lipid environment," *Nature Communications*, vol. 15, p. 2967, Apr. 2024.

[34] J. Yu, H. Zhu, R. Lape, T. Greiner, J. Du, W. Lü, L. Sivilotti, and E. Gouaux, "Mechanism of gating and partial agonist action in the glycine receptor," *Cell*, vol. 184, pp. 957–968.e21, Feb. 2021.

[35] B. Katz and S. Thesleff, "A study of the 'desensitization' produced by acetylcholine at the motor end-plate," *The Journal of Physiology*, vol. 138, no. 1, pp. 63–80, 1957.

[36] M. A. Dämgen and P. C. Biggin, "A Refined Open State of the Glycine Receptor Obtained via Molecular Dynamics Simulations," *Structure(London, England:1993)*, vol. 28, pp. 130–139.e2, Jan. 2020.

[37] A. Kumar, S. Basak, S. Rao, Y. Gicheru, M. L. Mayer, M. S. P. Sansom, and S. Chakrapani, "Mechanisms of activation and desensitization of full-length glycine receptor in lipid nanodiscs," *Nature Communications*, vol. 11, p. 3752, July 2020.

[38] M. Yazdani, Z. Jia, and J. Chen, "Hydrophobic dewetting in gating and regulation of transmembrane protein ion channels," *The Journal of Chemical Physics*, vol. 153, p. 110901, Sept. 2020.

[39] C. I. Lynch, G. Klesse, S. Rao, S. J. Tucker, and M. S. P. Sansom, "Water Nanoconfined in a Hydrophobic Pore: Molecular Dynamics Simulations of Transmembrane Protein 175 and the Influence of Water Models," *ACS Nano*, vol. 15, pp. 19098–19108, Dec. 2021.

[40] G. Klesse, S. Rao, M. S. P. Sansom, and S. J. Tucker, "CHAP: A Versatile Tool for the Structural and Functional Annotation of Ion Channel Pores," *Journal of Molecular Biology*, vol. 431, pp. 3353–3365, Aug. 2019.

[41] A. Anishkin and S. Sukharev, "Water Dynamics and Dewetting Transitions in the Small Mechanosensitive Channel MscS," *Biophysical Journal*, vol. 86, pp. 2883–2895, May 2004.

[42] O. Beckstein and M. S. P. Sansom, "The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores," *Physical Biology*, vol. 1, pp. 42–52, Mar. 2004.

[43] S. Rao, G. Klesse, P. J. Stansfeld, S. J. Tucker, and M. S. P. Sansom, "A heuristic derived from analysis of the ion channel structural proteome permits the rapid identification of hydrophobic gates," *Proceedings of the National Academy of Sciences*, vol. 116, pp. 13989–13995, July 2019.

[44] J. Dzubiella and J.-P. Hansen, "Electric-field-controlled water and ion permeation of a hydrophobic nanopore," *The Journal of Chemical Physics*, vol. 122, p. 234706, June 2005.

[45] J. Li, X. Gong, H. Lu, D. Li, H. Fang, and R. Zhou, "Electrostatic gating of a nanometer water channel," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 3687–3692, Mar. 2007.

[46] J. L. Trick, C. Song, E. J. Wallace, and M. S. P. Sansom, "Voltage Gating of a Biomimetic Nanopore: Electrowetting of a Hydrophobic Barrier," *ACS Nano*, vol. 11, pp. 1840–1847, Feb. 2017.

[47] G. Klesse, S. J. Tucker, and M. S. Sansom, "Electric field induced wetting of a hydrophobic gate in a model nanopore based on the 5-HT3 receptor channel," *ACS nano*, vol. 14, no. 8, pp. 10480–10491, 2020.

[48] O. Beckstein and M. S. P. Sansom, "A hydrophobic gate in an ion channel: the closed state of the nicotinic acetylcholine receptor," *Physical Biology*, vol. 3, pp. 147–159, July 2006.

[49] Y. Marcus, "Thermodynamics of solvation of ions. part 5.—Gibbs free energy of hydration at 298.15 K," *J. Chem. Soc., Faraday Trans.*, vol. 87, pp. 2995–2999, 1991.

[50] M. Smith, Z. Li, L. Landry, K. M. Merz, and P. Li, "Consequences of Overfitting the van der Waals Radii of Ions," *Journal of Chemical Theory and Computation*, vol. 19, pp. 2064–2074, Apr. 2023.

[51] F. Hofmeister, "Zur Lehre von der Wirkung der Salze: Zweite Mittheilung," *Archiv für Experimentelle Pathologie und Pharmakologie*, vol. 24, pp. 247–260, Feb. 1888.

[52] N. Schwierz, D. Horinek, U. Sivan, and R. R. Netz, "Reversed hofmeister series—the rule rather than the exception," *Current Opinion in Colloid Interface Science*, vol. 23, pp. 10–18, 2016.

[53] B. Kang, H. Tang, Z. Zhao, and S. Song, "Hofmeister series: Insights of ion specificity from amphiphilic assembly and interface property," *ACS Omega*, vol. 5, pp. 6229–6239, Mar. 2020.

[54] A. Grossfield, P. Ren, and J. W. Ponder, "Ion solvation thermodynamics from simulation with a polarizable force field," *Journal of the American Chemical Society*, vol. 125, pp. 15671–15682, Dec. 2003.

[55] Y. Levin, "Polarizable ions at interfaces," *Phys. Rev. Lett.*, vol. 102, p. 147803, Apr. 2009.

[56] M. Peng and A. V. Nguyen, "Adsorption of ionic surfactants at the air-water interface: The gap between theory and experiment," *Advances in Colloid and Interface Science*, vol. 275, p. 102052, 2020.

[57] D. J. Tobias, A. C. Stern, M. D. Baer, Y. Levin, and C. J. Mundy, "Simulation and theory of ions at atmospherically relevant aqueous liquid-air interfaces," *Annual Review of Physical Chemistry*, vol. 64, no. Volume 64, 2013, pp. 339–359, 2013.

[58] P. B. Petersen and R. J. Saykally, "On the nature of ions at the liquid water surface," *Annual Review of Physical Chemistry*, vol. 57, no. Volume 57, 2006, pp. 333–364, 2006.

[59] C. Caleman, J. S. Hub, P. J. van Maaren, and D. van der Spoel, "Atomistic simulation of ion solvation in water explains surface preference of halides," *Proceedings of the National Academy of Sciences*, vol. 108, no. 17, pp. 6838–6842, 2011.

[60] D. Liu, G. Ma, L. M. Levering, and H. C. Allen, "Vibrational spectroscopy of aqueous sodium halide solutions and air–liquid interfaces: observation of increased interfacial depth," *The Journal of Physical Chemistry B*, vol. 108, pp. 2252–2260, Feb. 2004.

[61] B. C. Garrett, "Ions at the air/water interface," *Science*, vol. 303, no. 5661, pp. 1146–1147, 2004.

[62] L. Piatkowski, Z. Zhang, E. H. G. Backus, H. J. Bakker, and M. Bonn, "Extreme surface propensity of halide ions in water," *Nature Communications*, vol. 5, p. 4083, June 2014.

[63] Y. Tong, I. Y. Zhang, and R. K. Campen, "Experimentally quantifying anion polarizability at the air/water interface," *Nature Communications*, vol. 9, p. 1313, Apr. 2018.

[64] P. Aryal, M. S. Sansom, and S. J. Tucker, "Hydrophobic Gating in Ion Channels," *Journal of Molecular Biology*, vol. 427, pp. 121–130, Jan. 2015.

[65] V. Dalal, M. J. Arcario, J. T. Petroff, B. K. Tan, N. M. Dietzen, M. J. Rau, J. A. J. Fitzpatrick, G. Brannigan, and W. W. L. Cheng, "Lipid nanodisc scaffold and size alter the structure of a pentameric ligand-gated ion channel," *Nature Communications*, vol. 15, p. 25, Jan. 2024.

[66] M. J. Thompson and J. E. Baenziger, "Structural basis for the modulation of pentameric ligand-gated ion channel function by lipids," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1862, no. 9, p. 183304, 2020.

[67] E. Flood, C. Boiteux, B. Lev, I. Vorobyov, and T. W. Allen, "Atomistic Simulations of Membrane Ion Channel Conduction, Gating, and Modulation," *Chemical Reviews*, vol. 119, pp. 7737–7832, July 2019.

[68] C. Kutzner, S. Páll, M. Fechner, A. Esztermann, B. L. de Groot, and H. Grubmüller, "Best bang for your buck: Gpu nodes for gromacs biomolecular simulations," *Journal of Computational Chemistry*, vol. 36, no. 26, pp. 1990–2008, 2015.

[69] D. Seiferth and P. C. Biggin, "Exploring the influence of pore shape on conductance and permeation," *Biophysical Journal*, 2024.

[70] C. P. Kelly, C. J. Cramer, and D. G. Truhlar, "Single-ion solvation free energies and the normal hydrogen electrode potential in methanol, acetonitrile, and dimethyl sulfoxide," *The Journal of Physical Chemistry B*, vol. 111, pp. 408–422, Jan. 2007.

[71] D. Seiferth, S. J. Tucker, and P. C. Biggin, "Limitations of non-polarizable force fields in describing anion binding poses in non-polar synthetic hosts," *Physical Chemistry Chemical Physics*, vol. 25, no. 26, pp. 17596–17608, 2023.

[72] A. R. Leach, *Molecular modelling : principles and applications*. Harlow, England: Prentice Hall, 2nd ed ed., 2001.

[73] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, vol. 81, pp. 3684–3690, Oct. 1984.

[74] C. D. Christ, A. E. Mark, and W. F. van Gunsteren, "Basic ingredients of free energy calculations: A review," *Journal of Computational Chemistry*, vol. 31, no. 8, pp. 1569–1582, 2010.

[75] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics*. Cambridge ; New York, NY: Cambridge University Press, third edition ed., 2018.

[76] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov. 1964.

[77] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A*, vol. 38, pp. 3098–3100, Sept. 1988.

[78] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, pp. 785–789, Jan. 1988.

[79] B. Jeziorski, R. Moszynski, and K. Szalewicz, "Perturbation theory approach to intermolecular potential energy surfaces of van der waals complexes," *Chemical Reviews*, vol. 94, pp. 1887–1930, Nov. 1994.

[80] D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, I. Schaefer, Henry F., A. Y. Sokolov, K. Patkowski, I. DePrince, A. Eugene, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, "PSI4 1.4: Open-source software for high-throughput quantum chemistry," *The Journal of Chemical Physics*, vol. 152, p. 184108, May 2020.

[81] K. Kumar, S. M. Woo, T. Siu, W. A. Cortopassi, F. Duarte, and R. S. Paton, "Cation–π interactions in protein–ligand binding: theory and data-mining reveal different roles for lysine and arginine," *Chem. Sci.*, vol. 9, pp. 2655–2665, 2018.

[82] C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, and C. Simmerling, "ff19sb: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution," *Journal of Chemical Theory and Computation*, vol. 16, pp. 528–552, Jan. 2020.

[83] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, "Current Status of the AMOEBA Polarizable Force Field," *The Journal of Physical Chemistry B*, vol. 114, pp. 2549–2564, Mar. 2010.

[84] P. E. M. Lopes, J. Huang, J. Shim, Y. Luo, H. Li, B. Roux, and A. D. MacKerell Jr., "Polarizable force field for peptides and proteins based on the classical drude oscillator," *Journal of Chemical Theory and Computation*, vol. 9, pp. 5430–5449, Dec. 2013.

[85] F. Duarte, B. A. Amrein, D. Blaha-Nelson, and S. C. Kamerlin, "Recent advances in qm/mm free energy calculations using reference potentials," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1850, no. 5, pp. 954–965, 2015.

[86] R. S. Mulliken, "Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I," *The Journal of Chemical Physics*, vol. 23, pp. 1833–1840, Oct. 1955.

[87] S. R. Cox and D. E. Williams, "Representation of the molecular electrostatic potential by a net atomic charge model," *Journal of Computational Chemistry*, vol. 2, no. 3, pp. 304–323, 1981.

[88] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model," *The Journal of Physical Chemistry*, vol. 97, pp. 10269–10280, Oct. 1993.

[89] A. J. Stone, "Distributed multipole analysis: stability for large basis sets," *Journal of Chemical Theory and Computation*, vol. 1, pp. 1128–1132, Nov. 2005.

[90] F. London, "Zur Theorie und Systematik der Molekularkräfte," *Zeitschrift für Physik*, vol. 63, pp. 245–279, Mar. 1930.

[91] P. S. Nerenberg and T. Head-Gordon, "New developments in force fields for biomolecular simulations," *Current Opinion in Structural Biology*, vol. 49, pp. 129–138, 2018.

[92] P. Ren and J. W. Ponder, "Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation," *The Journal of Physical Chemistry B*, vol. 107, pp. 5933–5947, June 2003.

[93] P. Cieplak, F.-Y. Dupradeau, Y. Duan, and J. Wang, "Polarization effects in molecular mechanical force fields," *Journal of Physics: Condensed Matter*, vol. 21, p. 333102, July 2009.

[94] F.-Y. Lin and A. D. MacKerell, Jr, "Force fields for small molecules," in *Methods in Molecular Biology*, Methods in molecular biology (Clifton, N.J.), pp. 21–54, Springer New York, 2019.

[95] S. Patel and C. L. Brooks III, "Charmm fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations," *Journal of Computational Chemistry*, vol. 25, no. 1, pp. 1–16, 2004.

[96] G. Lamoureux, J. MacKerell, Alexander D., and B. Roux, "A simple polarizable model of water based on classical Drude oscillators," *The Journal of Chemical Physics*, vol. 119, pp. 5185–5197, Sept. 2003.

[97] B. Thole, "Molecular polarizabilities calculated with a modified dipole interaction," *Chemical Physics*, vol. 59, pp. 341–350, Aug. 1981.

[98] A. J. Stone, *The Theory of Intermolecular Forces*. Oxford University Press, July 1996.

[99] D. M. Young and W. Rheinboldt, *Iterative Solution of Large Linear Systems*. Saint Louis: Elsevier Science, 2014.

[100] I. V. Leontyev and A. A. Stuchebrukhov, "Electronic continuum model for molecular dynamics simulations," *The Journal of Chemical Physics*, vol. 130, p. 085102, Feb. 2009.

[101] E. Duboué-Dijon, P. Delcroix, H. Martinez-Seara, J. Hladílková, P. Coufal, T. Křížek, and P. Jungwirth, "Binding of divalent cations to insulin: Capillary electrophoresis and molecular simulations," *The Journal of Physical Chemistry B*, vol. 122, pp. 5640–5648, May 2018.

[102] I. V. Leontyev, M. V. Vener, I. V. Rostov, M. V. Basilevsky, and M. D. Newton, "Continuum level treatment of electronic polarization in the framework of molecular simulations of solvation effects," *The Journal of Chemical Physics*, vol. 119, pp. 8024–8037, Oct. 2003.

[103] B. J. Kirby and P. Jungwirth, "Charge scaling manifesto: A way of reconciling the inherently macroscopic and microscopic natures of molecular simulations," *The Journal of Physical Chemistry Letters*, vol. 10, pp. 7531–7536, Dec. 2019.

[104] E. Duboué-Dijon, M. Javanainen, P. Delcroix, P. Jungwirth, and H. Martinez-Seara, "A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization," *The Journal of Chemical Physics*, vol. 153, p. 050901, Aug. 2020.

[105] L. X. Phan, C. I. Lynch, J. Crain, M. S. Sansom, and S. J. Tucker, "Influence of effective polarization on ion and water interactions within a biomimetic nanopore," *Biophysical Journal*, vol. 121, no. 11, pp. 2014–2026, 2022.

[106] L. X. Phan, V. C. Chamorro, H. Martinez-Seara, J. Crain, M. S. Sansom, and S. J. Tucker, "Influence of electronic polarization on the binding of anions to a chloride-pumping rhodopsin," *Biophysical Journal*, vol. 122, no. 8, pp. 1548–1556, 2023.

[107] L. X. Phan, A. P. Owji, T. Yang, J. Crain, M. S. Sansom, and S. J. Tucker, "Electronic polarizability tunes the function of the human bestrophin 1 cl^- channel," *bioRxiv*, 2024.

[108] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, pp. 926–935, July 1983.

[109] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, "The missing term in effective pair potentials," *The Journal of Physical Chemistry*, vol. 91, pp. 6269–6271, Nov. 1987.

[110] P. Mark and L. Nilsson, "Structure and dynamics of the tip3p, spc, and spc/e water models at 298 k," *The Journal of Physical Chemistry A*, vol. 105, pp. 9954–9960, Nov. 2001.

[111] P. Ren and J. W. Ponder, "Temperature and pressure dependence of the amoeba water model," *The Journal of Physical Chemistry B*, vol. 108, pp. 13427–13437, Sept. 2004.

[112] V. Lim, D. Hahn, G. Tresadern, C. Bayly, and D. Mobley, "Benchmark assessment of molecular geometries and energies from small molecule force fields [version 1; peer review: 2 approved]," *F1000Research*, vol. 9, no. 1390, 2020.

[113] P. Li, L. F. Song, and K. M. Merz, "Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model," *Journal of Chemical Theory and Computation*, vol. 11, pp. 1645–1657, Apr. 2015.

[114] E. A. Orabi, T. N. Öztürk, N. Bernhardt, and J. D. Faraldo-Gómez, "Corrections in the CHARMM36 Parametrization of Chloride Interactions with Proteins, Lipids, and Alkali Cations, and Extension to Other Halide Anions," *Journal of Chemical Theory and Computation*, vol. 17, pp. 6240–6261, Oct. 2021.

[115] E. E. Bruce and N. F. A. Van Der Vegt, "Does an electronic continuum correction improve effective short-range ion-ion interactions in aqueous solution?," *The Journal of Chemical Physics*, vol. 148, p. 222816, June 2018.

[116] S. J. Stuart and B. J. Berne, "Effects of Polarizability on the Hydration of the Chloride Ion," *The Journal of Physical Chemistry*, vol. 100, pp. 11934–11943, Jan. 1996.

[117] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, pp. 308–313, Jan. 1965.

[118] S. Hug, *Classical Molecular Dynamics in a Nutshell*, pp. 127–152. Totowa, NJ: Humana Press, 2013.

[119] M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics," *The Journal of Chemical Physics*, vol. 97, pp. 1990–2001, Aug. 1992.

[120] M. Allen and D. Tildesley, *Computer Simulation of Liquids*. Oxford science publications, Oxford University Press, 2017.

[121] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, "Efficient algorithms for langevin and dpd dynamics," *Journal of Chemical Theory and Computation*, vol. 8, pp. 3637–3649, Oct. 2012.

[122] A. T. Brunger, C. L. Brooks, and M. Karplus, "Stochastic boundary conditions for molecular dynamics simulations of st2 water," *Chemical Physics Letters*, vol. 105, pp. 495–500, 1984.

[123] J. Chowdhary, E. Harder, P. E. M. Lopes, L. Huang, A. D. MacKerell Jr., and B. Roux, "A polarizable force field of dipalmitoylphosphatidylcholine based on the classical drude model for molecular dynamics simulations of lipids," *The Journal of Physical Chemistry B*, vol. 117, pp. 9142–9160, Aug. 2013.

[124] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[125] B. Hess, "P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation," *Journal of Chemical Theory and Computation*, vol. 4, pp. 116–122, Jan. 2008.

[126] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Chemical Physics*, vol. 126, p. 014101, Jan. 2007.

[127] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied Physics*, vol. 52, pp. 7182–7190, Dec. 1981.

[128] J. Åqvist, P. Wennerström, M. Nervall, S. Bjelic, and B. O. Brandsdal, "Molecular dynamics simulations of water and biomolecules with a monte carlo constant pressure algorithm," *Chemical Physics Letters*, vol. 384, no. 4, pp. 288–294, 2004.

[129] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1-2, pp. 19–25, Sept. 2015.

[130] L. Verlet, "Computer "experiments" on classical fluids. I. thermodynamical properties of lennard-jones molecules," *Phys. Rev.*, vol. 159, pp. 98–103, July 1967.

[131] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, pp. 10089–10092, June 1993.

[132] P. P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921.

[133] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method," *The Journal of Chemical Physics*, vol. 103, pp. 8577–8593, Nov. 1995.

[134] O. S. Smart, J. G. Neduvelil, X. Wang, B. Wallace, and M. S. Sansom, "HOLE: A program for the analysis of the pore dimensions of ion channel structural models," *Journal of Molecular Graphics*, vol. 14, pp. 354–360, Dec. 1996.

[135] S. Doudou, N. A. Burton, and R. H. Henchman, "Standard Free Energy of Binding from a One-Dimensional Potential of Mean Force," *Journal of Chemical Theory and Computation*, vol. 5, pp. 909–918, Apr. 2009.

[136] D. Markthaler, S. Jakobtorweihen, and N. Hansen, "Lessons learned from the calculation of one-dimensional potentials of mean force [article v1.0]," *Living Journal of Computational Molecular Science*, vol. 1, p. 11073, Nov. 2019.

[137] A. Grossfield, "Wham: the weighted histogram analysis method." version 2.0.9.

[138] J. S. Hub, B. L. De Groot, and D. Van Der Spoel, "g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates," *Journal of Chemical Theory and Computation*, vol. 6, pp. 3713–3720, Dec. 2010.

[139] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *Journal of Computational Chemistry*, vol. 25, pp. 1157–1174, July 2004.

[140] J. Wagner, M. Thompson, D. Dotson, Hyejang, SimonBoothroyd, and J. Rodríguez-Guerra, "openforcefield/openff-forcefields: Version 1.3.1 "Parsley" Update," June 2021.

[141] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," *Journal of Molecular Graphics and Modelling*, vol. 25, pp. 247–260, Oct. 2006.

[142] P. Ren, C. Wu, and J. W. Ponder, "Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules," *Journal of Chemical Theory and Computation*, vol. 7, pp. 3143–3161, Oct. 2011.

[143] J. C. Wu, G. Chattree, and P. Ren, "Automation of AMOEBA polarizable force field parameterization for small molecules," *Theoretical Chemistry Accounts*, vol. 131, p. 1138, Mar. 2012.

[144] M. L. Laury, Z. Wang, A. S. Gordon, and J. W. Ponder, "Absolute binding free energies for the sampl6 cucurbit[8]uril host–guest challenge via the amoeba polarizable force field," *Journal of Computer-Aided Molecular Design*, vol. 32, pp. 1087–1095, Oct. 2018.

[145] P. Ren and J. W. Ponder, "Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations," *Journal of Computational Chemistry*, vol. 23, no. 16, pp. 1497–1506, 2002.

[146] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, "Tinker 8: Software Tools for Molecular Design," *Journal of Chemical Theory and Computation*, vol. 14, pp. 5273–5289, Oct. 2018.

[147] C. Kandt, W. L. Ash, and D. Peter Tieleman, "Setting up and running molecular dynamics simulations of membrane proteins," *Methods*, vol. 41, pp. 475–488, Apr. 2007.

[148] S. Jo, T. Kim, V. G. Iyer, and W. Im, "Charmm-gui: A web-based graphical user interface for CHARMM," *Journal of Computational Chemistry*, vol. 29, no. 11, pp. 1859–1865, 2008.

[149] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im, "Charmm-gui membrane builder toward realistic biological membrane simulations," *Journal of Computational Chemistry*, vol. 35, no. 27, pp. 1997–2004, 2014.

[150] H. Piekarski, K. Kubalczyk, and M. Wasiak, "Volumes, Heat Capacities, and Compressibilities of the Mixtures of Acetonitrile with *N* , *N* -Dimethylacetamide and Propylene Carbonate," *Journal of Chemical & Engineering Data*, vol. 55, pp. 5435–5440, Dec. 2010.

[151] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell, "Charmm36m: an improved force field for folded and intrinsically disordered proteins," *Nature Methods*, vol. 14, pp. 71–73, Jan. 2017.

[152] G. Klesse, S. Rao, S. J. Tucker, and M. S. Sansom, "Induced Polarization in Molecular Dynamics Simulations of the 5-HT $_3$ Receptor Channel," *Journal of the American Chemical Society*, vol. 142, pp. 9415–9427, May 2020.

[153] M. Lisbjerg, B. E. Nielsen, B. O. Milhøj, S. P. A. Sauer, and M. Pittelkow, "Anion binding by biotin[6]uril in water," *Organic & Biomolecular Chemistry*, vol. 13, no. 2, pp. 369–373, 2015.

[154] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, "Openmm 7: Rapid development of high performance algorithms for molecular dynamics," *PLOS Computational Biology*, vol. 13, pp. 1–17, July 2017.

[155] A. A. Kognole, J. Lee, S.-J. Park, S. Jo, P. Chatterjee, J. A. Lemkul, J. Huang, A. D. MacKerell Jr., and W. Im, "Charmm-gui drude prepper for molecular dynamics simulation using the classical drude polarizable force field," *Journal of Computational Chemistry*, vol. 43, no. 5, pp. 359–375, 2022.

[156] V. Limongelli, M. Bonomi, and M. Parrinello, "Funnel metadynamics as accurate binding free-energy method," *Proceedings of the National Academy of Sciences*, vol. 110, pp. 6358–6363, Apr. 2013.

[157] M. Sotomayor, V. Vásquez, E. Perozo, and K. Schulten, "Ion conduction through mscs as determined by electrophysiology and simulation," *Biophysical Journal*, vol. 92, no. 3, pp. 886–902, 2007.

[158] B. Roux, "The membrane potential and its representation by a constant electric field in computer simulations," *Biophysical Journal*, vol. 95, no. 9, pp. 4205–4216, 2008.

[159] J. Gumbart, F. Khalili-Araghi, M. Sotomayor, and B. Roux, "Constant electric field simulations of the membrane potential illustrated with simple systems," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1818, no. 2, pp. 294–302, 2012.

[160] C. Kutzner, H. Grubmüller, B. de□Groot, and U. Zachariae, "Computational Electrophysiology: The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail," *Biophysical Journal*, vol. 101, pp. 809–817, Aug. 2011.

[161] C. Kutzner, D. A. Köpfer, J.-P. Machtens, B. L. de Groot, C. Song, and U. Zachariae, "Insights into the function of ion channels by computational electrophysiology simulations," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1858, no. 7, Part B, pp. 1741–1752, 2016.

[162] N. Unwin, "Acetylcholine receptor channel imaged in the open state," *Nature*, vol. 373, pp. 37–43, Jan. 1995.

[163] N. Unwin, "Projection structure of the nicotinic acetylcholine receptor: Distinct conformations of the α subunits," *Journal of Molecular Biology*, vol. 257, no. 3, pp. 586–596, 1996.

[164] K. Lum, D. Chandler, and J. D. Weeks, "Hydrophobicity at Small and Large Length Scales," *The Journal of Physical Chemistry B*, vol. 103, pp. 4570–4577, June 1999.

[165] I. Brovchenko, D. Paschek, and A. Geiger, "Gibbs ensemble simulation of water in spherical cavities," *The Journal of Chemical Physics*, vol. 113, pp. 5026–5036, Sept. 2000.

[166] R. Allen, J.-P. Hansen, and S. Melchionna, "Molecular dynamics investigation of water permeation through nanopores," *The Journal of Chemical Physics*, vol. 119, pp. 3905–3919, Aug. 2003.

[167] G. Hummer, J. C. Rasaiah, and J. P. Noworyta, "Water conduction through the hydrophobic channel of a carbon nanotube," *Nature*, vol. 414, pp. 188–190, Nov. 2001.

[168] M. S. P. Sansom and P. C. Biggin, "Water at the nanoscale," *Nature*, vol. 414, pp. 157–159, Nov. 2001.

[169] O. Beckstein and M. S. P. Sansom, "Liquid–vapor oscillations of water in hydrophobic nanopores," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 7063–7068, June 2003.

[170] R. Roth, D. Gillespie, W. Nonner, and R. E. Eisenberg, "Bubbles, Gating, and Anesthetics in Ion Channels," *Biophysical Journal*, vol. 94, pp. 4282–4298, June 2008.

[171] W. C. Wimley and S. H. White, "Experimentally determined hydrophobicity scale for proteins at membrane interfaces," *Nature Structural Biology*, vol. 3, pp. 842–848, Oct. 1996.

[172] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, pp. 583–589, Aug. 2021.

[173] M. Eriksson, M. Tuominen, M. Järn, P. M. Claesson, V. Wallqvist, H.-J. Butt, D. Vollmer, M. Kappl, J. Schoelkopf, P. A. C. Gane, H. Teisala, and A. Swerin, "Direct Observation of Gas Meniscus Formation on a Superhydrophobic Surface," *ACS Nano*, vol. 13, pp. 2246–2252, Feb. 2019.

[174] B. D. Brandner, P. M. Hansson, A. Swerin, P. M. Claesson, M. Wåhlander, J. Schoelkopf, and P. A. C. Gane, "Solvent segregation and capillary evaporation at a superhydrophobic surface investigated by confocal Raman microscopy and force measurements," *Soft Matter*, vol. 7, pp. 1045–1052, Jan. 2011.

[175] I. Chen, S. Pant, Q. Wu, R. J. Cater, M. Sobti, R. J. Vandenberg, A. G. Stewart, E. Tajkhorshid, J. Font, and R. M. Ryan, "Glutamate transporters have a chloride channel with two hydrophobic gates," *Nature*, vol. 591, pp. 327–331, Mar. 2021.

[176] S. Pant, Q. Wu, R. Ryan, and E. Tajkhorshid, "Microscopic Characterization of the Chloride Permeation Pathway in the Human Excitatory Amino Acid Transporter 1 (EAAT1)," *ACS Chemical Neuroscience*, vol. 13, pp. 776–785, Mar. 2022.

[177] C. L. Anderson and R. J. Thompson, "Intrapore lipids hydrophobically gate pannexin-1 channels," *Science Signaling*, vol. 15, p. eabn2081, Feb. 2022.

[178] M. Jensen, D. W. Borhani, K. Lindorff-Larsen, P. Maragakis, V. Jogini, M. P. Eastwood, R. O. Dror, and D. E. Shaw, "Principles of conduction and hydrophobic gating in K+ channels," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 5833–5838, Mar. 2010.

[179] A. H. Cerdan, N. E. Martin, and M. Cecchini, "An Ion-Permeable State of the Glycine Receptor Captured by Molecular Dynamics," *Structure*, vol. 26, pp. 1555–1562.e4, Nov. 2018.

[180] Z. Jia, M. Yazdani, G. Zhang, J. Cui, and J. Chen, "Hydrophobic gating in BK channels," *Nature Communications*, vol. 9, p. 3408, Aug. 2018.

[181] R. B. Bass, P. Strop, M. Barclay, and D. C. Rees, "Crystal Structure of *Escherichia coli* MscS, a Voltage-Modulated and Mechanosensitive Channel," *Science*, vol. 298, pp. 1582–1587, Nov. 2002.

[182] S. Miller, "Domain organization of the MscS mechanosensitive channel of Escherichia coli," *The EMBO Journal*, vol. 22, pp. 36–46, Jan. 2003.

[183] W. Wang, S. S. Black, M. D. Edwards, S. Miller, E. L. Morrison, W. Bartlett, C. Dong, J. H. Naismith, and I. R. Booth, "The Structure of an Open Form of an E. coli Mechanosensitive Channel at 3.45 Å Resolution," *Science*, vol. 321, pp. 1179–1183, Aug. 2008.

[184] A. Anishkin, B. Akitake, K. Kamaraju, C.-S. Chiang, and S. Sukharev, "Hydration properties of mechanosensitive channel pores define the energetics of gating," *Journal of Physics: Condensed Matter*, vol. 22, p. 454120, Oct. 2010.

[185] J. S. Najem, I. Rowe, A. Anishkin, D. J. Leo, and S. Sukharev, "The voltage-dependence of MscL has dipolar and dielectric contributions and is governed by local intramembrane electric field," *Scientific Reports*, vol. 8, p. 13607, Dec. 2018.

[186] W. Zheng, E. O. Gracheva, and S. N. Bagriantsev, "A hydrophobic gate in the inner pore helix is the major determinant of inactivation in mechanosensitive Piezo channels," *eLife*, vol. 8, p. e44003, Jan. 2019.

[187] M. A. Kasimova, A. Yazici, Y. Yudin, D. Granata, M. L. Klein, T. Rohacs, and V. Carnevale, "Ion Channel Sensing: Are Fluctuations the Crux of the Matter?," *The Journal of Physical Chemistry Letters*, vol. 9, pp. 1260–1264, Mar. 2018.

[188] W. Zheng, R. Hu, R. Cai, L. Hofmann, Q. Hu, M. Fatehi, W. Long, T. Kong, J. Tang, P. Light, V. Flockerzi, Y. Cao, and X.-Z. Chen, "Identification and characterization of hydrophobic gate residues in TRP channels," *The FASEB Journal*, vol. 32, no. 2, pp. 639–653, 2018.

[189] M. Yamashita, P. S.-W. Yeung, C. E. Ing, B. A. McNally, R. Pomès, and M. Prakriya, "STIM1 activates CRAC channels through rotation of the pore helix to open a hydrophobic gate," *Nature Communications*, vol. 8, p. 14512, Feb. 2017.

[190] C. Neale, N. Chakrabarti, P. Pomorski, E. F. Pai, and R. Pomès, "Hydrophobic Gating of Ion Permeation in Magnesium Channel CorA," *PLOS Computational Biology*, vol. 11, p. e1004303, July 2015.

[191] S. Rao, G. Klesse, P. J. Stansfeld, S. J. Tucker, and M. S. P. Sansom, "A BEST example of channel structure annotation by molecular simulation," *Channels*, vol. 11, pp. 347–353, July 2017.

[192] M. Yonkunas and M. Kurnikova, "The Hydrophobic Effect Contributes to the Closed State of a Simplified Ion Channel through a Conserved Hydrophobic Patch at the Pore-Helix Crossing," *Frontiers in Pharmacology*, vol. 6, Nov. 2015.

[193] M. V. Yelshanskaya, D. S. Patel, C. M. Kottke, M. G. Kurnikova, and A. I. Sobolevsky, "Opening of glutamate receptor channel to subconductance levels," *Nature*, vol. 605, pp. 172–178, May 2022.

[194] N. Bocquet, H. Nury, M. Baaden, C. Le Poupon, J.-P. Changeux, M. Delarue, and P.-J. Corringer, "X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation," *Nature*, vol. 457, pp. 111–114, Jan. 2009.

[195] A. N. Miller and S. B. Long, "Crystal Structure of the Human Two–Pore Domain Potassium Channel K2P1," *Science*, vol. 335, pp. 432–436, Jan. 2012.

[196] S. G. Brohawn, E. B. Campbell, and R. MacKinnon, "Physical mechanism for gating and mechanosensitivity of the human TRAAK K+ channel," *Nature*, vol. 516, pp. 126–130, Dec. 2014.

[197] B. Reddy, N. Bavi, A. Lu, Y. Park, and E. Perozo, "Molecular basis of force-from-lipids gating in the mechanosensitive channel MscS," *eLife*, vol. 8, p. e50486, Dec. 2019.

[198] R. Jin, S. He, K. A. Black, O. B. Clarke, D. Wu, J. R. Bolla, P. Johnson, A. Periasamy, A. Wardak, P. Czabotar, P. M. Colman, C. V. Robinson, D. Laver, B. J. Smith, and J. M. Gulbis, "Ion currents through Kir potassium channels are gated by anionic lipids," *Nature Communications*, vol. 13, p. 490, Jan. 2022.

[199] T. S. Turney, V. Li, and S. G. Brohawn, "Structural Basis for pH-gating of the K+ channel TWIK1 at the selectivity filter," *Nature Communications*, vol. 13, p. 3232, Dec. 2022.

[200] J. L. Syrjanen, K. Michalski, T.-H. Chou, T. Grant, S. Rao, N. Simorowski, S. J. Tucker, N. Grigorieff, and H. Furukawa, "Structure and assembly of calcium homeostasis modulator proteins," *Nature Structural & Molecular Biology*, vol. 27, pp. 150–159, Feb. 2020.

[201] K. Drożdżyk, M. Sawicka, M.-I. Bahamonde-Santos, Z. Jonas, D. Deneka, C. Albrecht, and R. Dutzler, "Cryo-EM structures and functional properties of CALHM channels of the human placenta," *eLife*, vol. 9, p. e55853, May 2020.

[202] W. Jiang, J. Lacroix, and Y. L. Luo, "Lipids in the Piezo1 channel pore: gating mechanism, or simulation artifact?," tech. rep., bioRxiv, Nov. 2021.

[203] C. L. Noland, H. C. Chua, M. Kschonsak, S. A. Heusser, N. Braun, T. Chang, C. Tam, J. Tang, C. P. Arthur, C. Ciferri, S. A. Pless, and J. Payandeh, "Structure-guided unlocking of NaX reveals a non-selective tetrodotoxin-sensitive cation channel," *Nature Communications*, vol. 13, p. 1416, Mar. 2022.

[204] M. Kuzuya, H. Hirano, K. Hayashida, M. Watanabe, K. Kobayashi, T. Terada, M. I. Mahmood, F. Tama, K. Tani, Y. Fujiyoshi, and A. Oshima, "Structures of human pannexin-1 in nanodiscs reveal gating mediated by dynamic movement of the N terminus and phospholipids," *Science Signaling*, vol. 15, no. 720, p. eabg6941.

[205] B. Burendei, R. Shinozaki, M. Watanabe, T. Terada, K. Tani, Y. Fujiyoshi, and A. Oshima, "Cryo-EM structures of undocked innexin-6 hemichannels in phospholipids," *Science Advances*, vol. 6, no. 7, p. eaax3157.

[206] H.-J. Lee, H. Jeong, J. Hyun, B. Ryu, K. Park, H.-H. Lim, J. Yoo, and J.-S. Woo, "Cryo-EM structure of human Cx31.3/GJC3 connexin hemichannel," *Science Advances*, vol. 6, no. 35, p. eaba4996.

[207] S. Rajendra, J. W. Lynch, and P. R. Schofield, "The glycine receptor," *Pharmacology & Therapeutics*, vol. 73, no. 2, pp. 121–146, 1997.

[208] Q. Shan, J. L. Haddrill, and J. W. Lynch, "Ivermectin, an unconventional agonist of the glycine receptor chloride channel*," *Journal of Biological Chemistry*, vol. 276, no. 16, pp. 12556–12564, 2001.

[209] C. Bergh, S. A. Heusser, R. Howard, and E. Lindahl, "Markov state models of proton- and pore-dependent activation in a pentameric ligand-gated ion channel," *eLife*, vol. 10, p. e68369, Oct. 2021.

[210] C. Bergh, U. Rovšnik, R. Howard, and E. Lindahl, "Discovery of lipid binding sites in a ligand-gated ion channel by integrating simulations and cryo-em," *eLife*, vol. 12, p. RP86016, Jan. 2024.

[211] M. Petřek, M. Otyepka, P. Banáš, P. Košinová, J. Koča, and J. Damborský, "CAVER: a new tool to explore routes from protein clefts, pockets and cavities," *BMC Bioinformatics*, vol. 7, p. 316, Dec. 2006.

[212] L. Pravda, D. Sehnal, D. Toušek, V. Navrátilová, V. Bazgier, K. Berka, R. Svobodová□Vařeková, J. Koča, and M. Otyepka, "MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update)," *Nucleic Acids Research*, vol. 46, pp. W368–W373, July 2018.

[213] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, and J. Damborsky, "CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures," *PLoS Computational Biology*, vol. 8, p. e1002708, Oct. 2012.

[214] S. Minniberger, S. Abdolvand, S. Braunbeck, H. Sun, and A. J. Plested, "Asymmetry and Ion Selectivity Properties of Bacterial Channel NaK Mutants Derived from Ionotropic Glutamate Receptors," *Journal of Molecular Biology*, vol. 435, p. 167970, Mar. 2023.

[215] R. N. Roy, K. Hendriks, W. Kopec, S. Abdolvand, K. L. Weiss, B. L. De Groot, A. Lange, H. Sun, and L. Coates, "Structural plasticity of the selectivity filter in a nonselective ion channel," *IUCrJ*, vol. 8, pp. 421–430, May 2021.

[216] Y. Zhang, P. M. Dijkman, R. Zou, M. Zandl-Lang, R. M. Sanchez, L. Eckhardt-Strelau, H. Köfeler, H. Vogel, S. Yuan, and M. Kudryashev, "Asymmetric opening of the homopentameric 5-HT3A serotonin receptor in lipid bilayers," *Nature Communications*, vol. 12, p. 1074, Feb. 2021.

[217] C. Shi, Y. He, K. Hendriks, B. L. De Groot, X. Cai, C. Tian, A. Lange, and H. Sun, "A single NaK channel conformation is not enough for non-selective ion conduction," *Nature Communications*, vol. 9, p. 717, Feb. 2018.

[218] A. O. Chugunov, P. E. Volynsky, N. A. Krylov, D. E. Nolde, and R. G. Efremov, "Temperature-sensitive gating of TRPV1 channel as probed by atomistic simulations of its trans- and juxtamembrane domains," *Scientific Reports*, vol. 6, p. 33112, Sept. 2016.

[219] R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, and O. Beckstein, "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," (Austin, Texas), pp. 98–105, 2016.

[220] O. Smart, J. Breed, G. Smith, and M. Sansom, "A novel method for structure-based prediction of ion channel conductance properties," *Biophysical Journal*, vol. 72, pp. 1109–1126, Mar. 1997.

[221] A. Mejri, K. Mazouzi, G. Herlem, F. Picaud, T. Hennequin, J. Palmeri, and M. Manghi, "Molecular dynamics investigations of ionic conductance at the nanoscale: Role of the water model and geometric parameters," *Journal of Molecular Liquids*, vol. 351, p. 118575, Apr. 2022.

[222] M. Manghi, J. Palmeri, F. Henn, A. Noury, F. Picaud, G. Herlem, and V. Jourdain, "Ionic Conductance of Carbon Nanotubes: Confronting Literature Data with Nanofluidic Theory," *The Journal of Physical Chemistry C*, vol. 125, pp. 22943–22950, Oct. 2021.

[223] C. Xing, J. Sheng, L. Wang, and W. Fei, "Research progress in molecular dynamics simulation of CNT and graphene reinforced metal matrix composites," *Oxford Open Materials Science*, vol. 1, p. itab008, Nov. 2020.

[224] R. Rafiee and R. M. Moghadam, "On the modeling of carbon nanotubes: A critical review," *Composites Part B: Engineering*, vol. 56, pp. 435–449, Jan. 2014.

[225] B. H. Mendonça, D. N. De Freitas, M. H. Köhler, R. J. Batista, M. C. Barbosa, and A. B. De Oliveira, "Diffusion behaviour of water confined in deformed carbon nanotubes," *Physica A: Statistical Mechanics and its Applications*, vol. 517, pp. 491–498, Mar. 2019.

[226] B. H. S. Mendonça, P. Ternes, E. Salcedo, A. B. De Oliveira, and M. C. Barbosa, "Water diffusion in rough carbon nanotubes," *The Journal of Chemical Physics*, vol. 152, p. 024708, Jan. 2020.

[227] M. Rezaee, M. Namvarpour, A. Yeganegi, and H. Ghassemi, "Comprehensive study of monatomic fluid flow through elliptical carbon nanotubes," *Physics of Fluids*, vol. 32, p. 092006, Sept. 2020.

[228] S. Balme, F. Picaud, M. Manghi, J. Palmeri, M. Bechelany, S. Cabello-Aguilar, A. Abou-Chaaya, P. Miele, E. Balanzat, and J. M. Janot, "Ionic transport through sub-10 nm diameter hydrophobic high-aspect ratio nanopores: experiment, theory and simulation," *Scientific Reports*, vol. 5, p. 10135, June 2015.

[229] A. Mejri, N. Arroyo, G. Herlem, J. Palmeri, M. Manghi, F. Henn, and F. Picaud, "Impact of Single-Walled Carbon Nanotube Functionalization on Ion and Water Molecule Transport at the Nanoscale," *Nanomaterials*, vol. 14, p. 117, Jan. 2024.

[230] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," *Journal of the American Chemical Society*, vol. 118, pp. 11225–11236, Nov. 1996.

[231] M. Sansom and I. Kerr, "Transbilayer pores formed by beta-barrels: molecular modeling of pore structures and properties," *Biophysical Journal*, vol. 69, pp. 1334–1343, Oct. 1995.

[232] B. Burendei, R. Shinozaki, M. Watanabe, T. Terada, K. Tani, Y. Fujiyoshi, and A. Oshima, "Cryo-EM structures of undocked innexin-6 hemichannels in phospholipids," *Science Advances*, vol. 6, p. eaax3157, Feb. 2020.

[233] A. Oshima, K. Tani, and Y. Fujiyoshi, "Atomic structure of the innexin-6 gap junction channel determined by cryo-EM," *Nature Communications*, vol. 7, p. 13681, Dec. 2016.

[234] R. Nakamura, T. Numata, G. Kasuya, T. Yokoyama, T. Nishizawa, T. Kusakizako, T. Kato, T. Hagino, N. Dohmae, M. Inoue, K. Watanabe, H. Ichijo, M. Kikkawa, M. Shirouzu, T. J. Jentsch, R. Ishitani, Y. Okada, and O. Nureki, "Cryo-EM structure of the volume-regulated anion channel LRRC8D isoform identifies features important for substrate permeation," *Communications Biology*, vol. 3, p. 240, May 2020.

[235] S. Rutz, D. Deneka, A. Dittmann, M. Sawicka, and R. Dutzler, “Structure of a volume-regulated heteromeric LRRC8A/C channel,” *Nature Structural & Molecular Biology*, vol. 30, pp. 52–61, Jan. 2023.

[236] W. Humphrey, A. Dalke, and K. Schulten, “VMD: Visual molecular dynamics,” *Journal of Molecular Graphics*, vol. 14, pp. 33–38, Feb. 1996.

[237] DeLano, Warren L., “PyMOL: An Open-Source Molecular Graphics Tool,” *CCP4 Newsletter On Protein Crystallography*, vol. 40, pp. 82–92, 2002.

[238] S. H. Hewitt, G. Macey, R. Mailhot, M. R. J. Elsegood, F. Duarte, A. M. Kenwright, and S. J. Butler, “Tuning the anion binding properties of lanthanide receptors to discriminate nucleoside phosphates in a sensing array,” *Chem. Sci.*, vol. 11, pp. 3619–3628, 2020.

[239] M. Lisbjerg, B. M. Jessen, B. Rasmussen, B. E. Nielsen, A. Madsen, and M. Pittelkow, “Discovery of a cyclic 6 + 6 hexamer of d-biotin and formaldehyde,” *Chem. Sci.*, vol. 5, pp. 2647–2650, 2014.

[240] M. Lisbjerg, H. Valkenier, B. M. Jessen, H. Al-Kerdi, A. P. Davis, and M. Pittelkow, “Biotin[6]uril Esters: Chloride-Selective Transmembrane Anion Carriers Employing C—H···Anion Interactions,” *Journal of the American Chemical Society*, vol. 137, pp. 4948–4951, Apr. 2015.

[241] N. N. Andersen, K. Eriksen, M. Lisbjerg, M. E. Ottesen, B. O. Milhøj, S. P. A. Sauer, and M. Pittelkow, “Entropy/Enthalpy Compensation in Anion Binding: Biotin[6]uril and Biotin- I -sulfoxide[6]uril Reveal Strong Solvent Dependency,” *The Journal of Organic Chemistry*, vol. 84, pp. 2577–2584, Mar. 2019.

[242] D. Asthagiri, P. Dixit, S. Merchant, M. Paulaitis, L. Pratt, S. Rempe, and S. Varma, “Ion selectivity from local configurations of ligands in solutions and ion channels,” *Chemical Physics Letters*, vol. 485, pp. 1–7, Jan. 2010.

[243] V. Wineman-Fisher, Y. Al-Hamdani, I. Addou, A. Tkatchenko, and S. Varma, “Ion-Hydroxyl Interactions: From High-Level Quantum Benchmarks to Transferable Polarizable Force Fields,” *Journal of Chemical Theory and Computation*, vol. 15, pp. 2444–2453, Apr. 2019.

[244] Y. Zhang, Y. Jiang, Y. Qiu, and H. Zhang, “Rational Design of Nonbonded Point Charge Models for Highly Charged Metal Cations with Lennard-Jones 12-6 Potential,” *Journal of Chemical Information and Modeling*, vol. 61, pp. 4613–4629, Sept. 2021.

[245] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, “The reaxff reactive force-field: development, applications and future directions,” *npj Computational Materials*, vol. 2, p. 15011, Mar. 2016.

[246] B. Walker, C. Liu, E. Wait, and P. Ren, “Automation of AMOEBA polarizable force field for small molecules: Poltype 2,” *Journal of Computational Chemistry*, vol. 43, pp. 1530–1542, Sept. 2022.

[247] A. Kumar, K. Kindig, S. Rao, A.-M. Zaki, S. Basak, M. S. P. Sansom, P. C. Biggin, and S. Chakrapani, “Structural basis for cannabinoid-induced potentiation of alpha1-glycine receptors in lipid nanodiscs,” *Nature Communications*, vol. 13, p. 4862, Aug. 2022.

[248] M. Sunesen, L. P. de Carvalho, V. Dufresne, R. Grailhe, N. Savatier-Duclert, G. Gibor, A. Peretz, B. Attali, J.-P. Changeux, and Y. Paas, “Mechanism of cl- selection by a glutamate-gated chloride (glucl) receptor revealed through mutations in the selectivity filter*,” *Journal of Biological Chemistry*, vol. 281, no. 21, pp. 14875–14881, 2006.

[249] R. D. Schwartz and X. Yu, “Optical imaging of intracellular chloride in living brain slices,” *Journal of Neuroscience Methods*, vol. 62, no. 1, pp. 185–192, 1995.

[250] J. Du, W. Lü, S. Wu, Y. Cheng, and E. Gouaux, “Glycine receptor mechanism elucidated by electron cryo-microscopy,” *Nature*, vol. 526, pp. 224–229, Oct. 2015.

[251] A. H. Cerdan and M. Cecchini, “On the functional annotation of open-channel structures in the glycine receptor,” *Structure*, vol. 28, no. 6, pp. 690–693.e3, 2020.

[252] M. A. Dämgen, A. M. Zaki, and P. C. Biggin, “Comment on “on the functional annotation of open-channel structures in the glycine receptor”,” *Structure*, vol. 28, no. 6, pp. 601–603, 2020.

[253] M. Moroni, I. Biro, M. Giugliano, R. Vijayan, P. C. Biggin, M. Beato, and L. G. Sivilotti, “Chloride ions in the pore of glycine and gaba channels shape the time course and voltage dependence of agonist currents,” *Journal of Neuroscience*, vol. 31, no. 40, p. 14095 – 14106, 2011.

[254] G. Moraga-Cid, L. Sauguet, C. Huon, L. Malherbe, C. Girard-Blanc, S. Petres, S. Murail, A. Taly, M. Baaden, M. Delarue, and P.-J. Corringer, “Allosteric and hyperekplexic mutant phenotypes investigated on an α_1 glycine receptor transmembrane structure,” *Proceedings of the National Academy of Sciences*, vol. 112, pp. 2865–2870, Mar. 2015.

[255] J. Bormann, N. Rundström, H. Betz, and D. Langosch, “Residues within transmembrane segment m2 determine chloride conductance of glycine receptor homo- and hetero-oligomers.,” *The EMBO Journal*, vol. 12, no. 10, pp. 3729–3737, 1993.

[256] S. Rao, G. Klesse, C. I. Lynch, S. J. Tucker, and M. S. P. Sansom, “Molecular simulations of hydrophobic gating of pentameric ligand gated ion channels: Insights into water and ions,” *The Journal of Physical Chemistry B*, vol. 125, pp. 981–994, Feb. 2021.

[257] A. D. Mackerell Jr., “Empirical force fields for biological macromolecules: Overview and issues,” *Journal of Computational Chemistry*, vol. 25, no. 13, p. 1584 – 1604, 2004.

[258] H. Flyvbjerg and H. G. Petersen, “Error estimates on averages of correlated data,” *The Journal of Chemical Physics*, vol. 91, pp. 461–466, July 1989.

[259] F. Zhu and G. Hummer, “Theory and Simulation of Ion Conduction in the Pentameric GLIC Channel,” *Journal of Chemical Theory and Computation*, vol. 8, pp. 3759–3768, Oct. 2012.

[260] R. L. Spicer, A. D. Stergiou, T. A. Young, F. Duarte, M. D. Symes, and P. J. Lusby, “Host–guest-induced electron transfer triggers radical-cation catalysis,” *Journal of the American Chemical Society*, vol. 142, pp. 2134–2139, Feb. 2020.

[261] P. Li, L. F. Song, and K. M. Merz, “Parameterization of Highly Charged Metal Ions Using the 12-6-4 LJ-Type Nonbonded Model in Explicit Water,” *The Journal of Physical Chemistry B*, vol. 119, pp. 883–895, Jan. 2015.

Appendix A

Carbon Nanotubes

A.1 Li *et al.* Ion Parameters

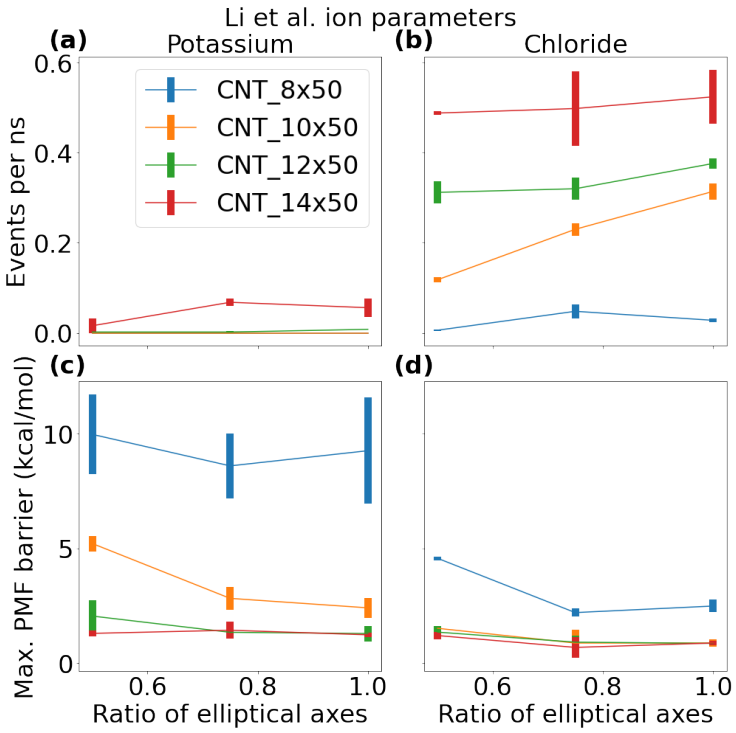


Figure A.1: Barrier for ion permeation and ion conductance for carbon nanotubes (CNTs) as a function of the ratio between of the elliptic radii. The conductance is measured as events per ns observed in multiple 250 ns repeats with external potential of 500 mV for (a) potassium ions and (b) chloride ions. The energetic barrier is the maximum of the Potential of Mean Force (PMF) obtained from umbrella sampling for (c) potassium ions and (d) chloride ions. The different CNT systems 8x50, 10x50, 12x50 and 14x50 (see Table 5.1 for corresponding HOLE radii) are all 50 Å in length and shown as blue, orange, green and red lines respectively). Chloride and potassium are described by force field parameter from Li *et al.* [261].

In Fig. 5.8, I plotted the conductance and barrier to permeation for potassium and chloride when I use the OPLS ion parameters. In chapter 5, only anions described by the Li *et al.* parameter set [113] bound to macro-cyclic hosts when using a pairwise additive force field. I replaced the OPLS ion parameters with the Li *et al.* ones to check the sensitivity of the results from chapter 5 to ion parameters. Fig. A.1 confirms the overall trend observed in Fig. 5.8 between conductance and permeation barrier as a function of the ratio between elliptic radii. The Lennard Jones potential for chloride described with ion parameters from Li *et al.* is plotted in Fig. 2.3.

A.2 Histogram Overlap

The windows for the umbrella sampling simulations in Chapter 5 exhibit good overlap, see Fig. A.2.

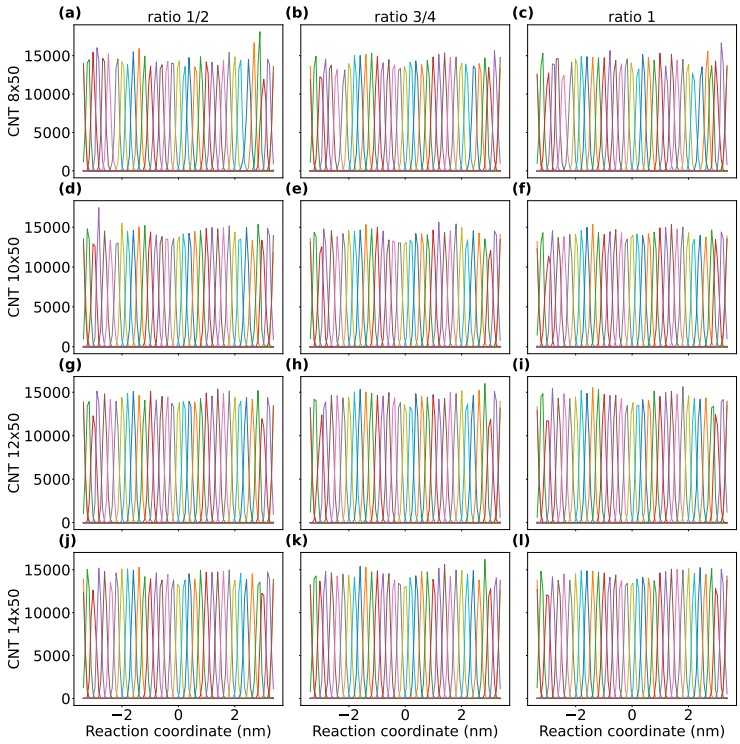


Figure A.2: Histogram overlap for 8x50, 10x50, 12x50 and 14x50 CNTs model systems (see Table 5.1 for corresponding HOLE radii) with varying ratio of the larger and smaller radius.

A.3 PMF Convergence

PMF convergence is analysed as described in section 3.5 in Fig. A.3 and Fig. A.4.

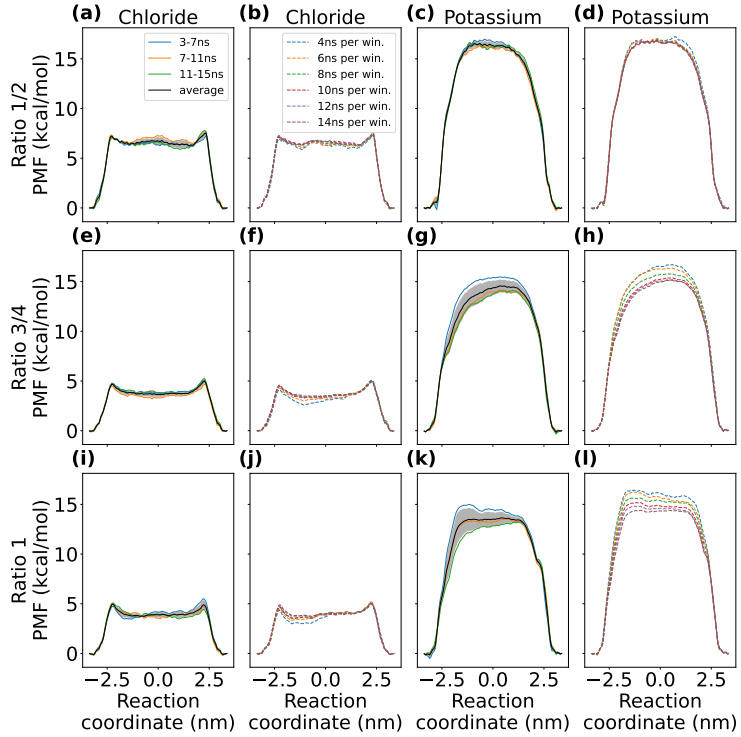


Figure A.3: PMF convergence of chloride and potassium permeating a 8x50 carbon nanotube with three different ratios of the longer and smaller radius.

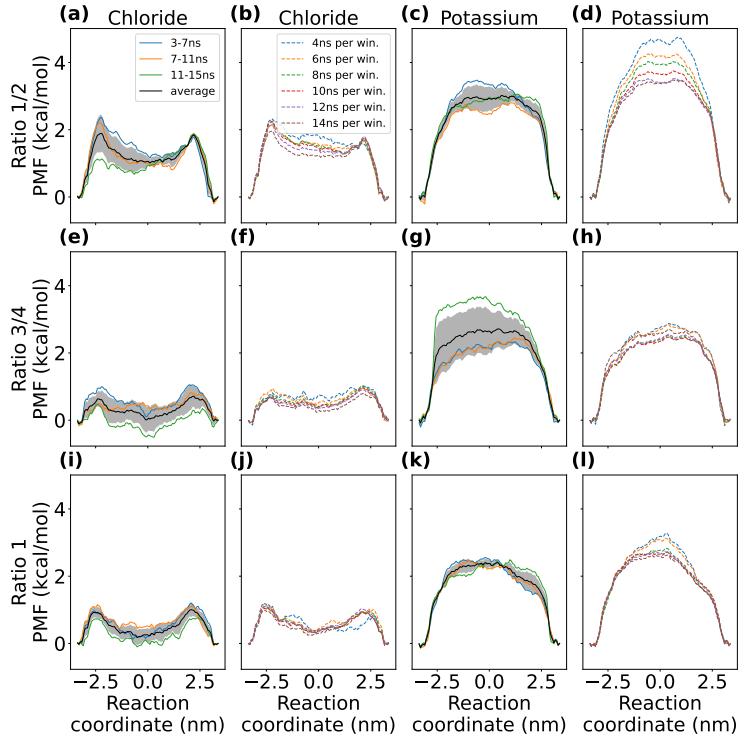


Figure A.4: PMF convergence of chloride and potassium permeating a 14x50 carbon nanotube with three different ratios of the longer and smaller radius.

Appendix B

Simulations with polarisable and non-polarisable force fields

B.1 Ion Carriers

Table B.1: Comparison of *gromacs* and *tinker* simulation performance. Using a single A40 GPU and 6 cores of an AMD 7402 2.8GHz node.

	Box size	Performance
Gromacs 2020.3	18029 atoms	308 ns/day
Tinker9	9089 atoms	38.7 ns/day

B.1.1 PMF Convergence and Histogram Overlap

The windows for the umbrella sampling simulations in Chapter 6 exhibit good overlap, see Fig. B.1 for acetonitrile and Fig. B.3 for water. PMF convergence is analysed as described in section 3.5 in Fig. B.2 for

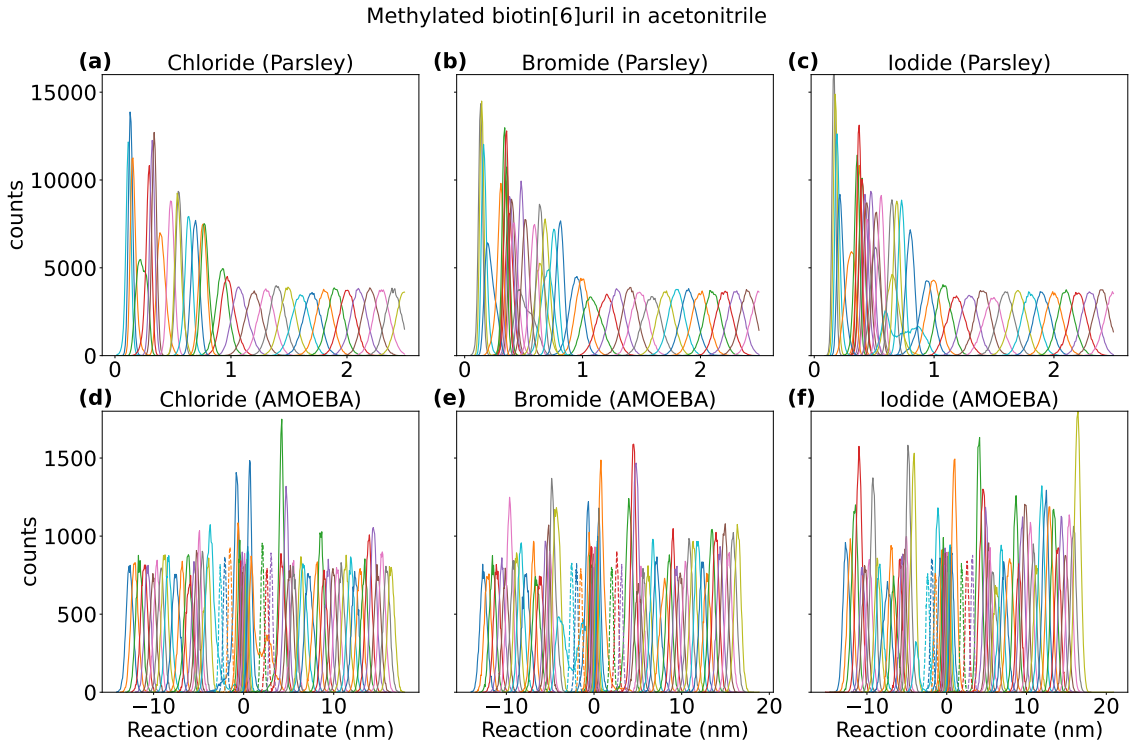


Figure B.1: Histogram overlap for methylated biotin[6]uril in acetonitrile with Parsley (a, b and c) and with AMOEBA (d, e, f). The dashed lines indicate a higher force constant (10 kcal/mol/Å2) rather than 2 kcal/mol/Å^2.

acetonitrile and Fig. B.4 for water.

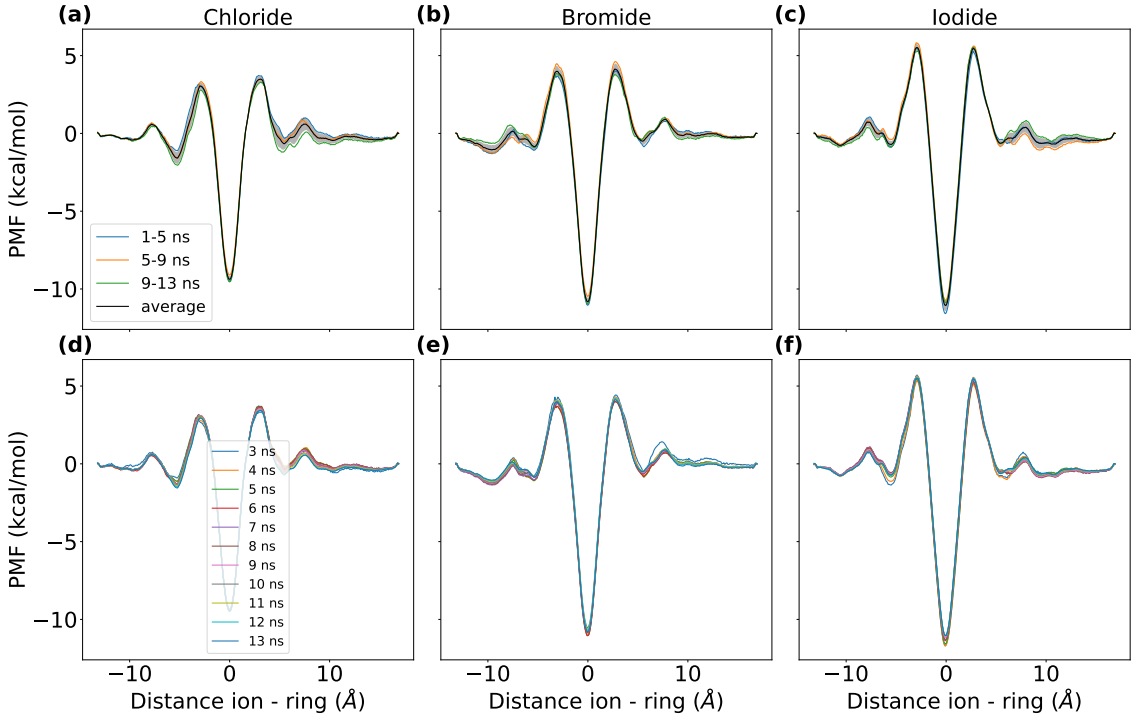


Figure B.2: Convergence analysis for methylated biotin[6]uril in acetonitrile with AMOEBA.

PMF calculations with GAFF in Fig. B.5 revealed similar PMF profiles to the PMF profiles calculated with Parsley in Fig. 6.2.

B.1.2 Comparing RMSD and Dihedrals

The conformational ensembles and the flexibilities and rigidities of biotin sulfide and biotin sulfoxide using both the AMOEBA and Parsley force fields were investigated. By comparing the RMSD values of the twelve cavity carbon atoms between the two force fields in Fig. B.6, I found that the average RMSD was less than 1 Å. To further understand the conformational behaviour, the C–N–C–N dihedral distributions was analysed, which describe the orientation of two subunits relative to each other. These distributions were similar across all systems and force fields, see Fig. B.7.

B.1.3 Effects of Refitting Ion Parameters

The chloride with fitted Lennard Jones potential was placed centrally in the ring, energy minimized, and carefully equilibrated. 1 ns NVT, 1 ns NPT, 1 ns NPT with restraints of 1500 and 1000 and 500 kJ/mol/nm^2 restraints respectively to ensure that the ion is centrally bound in the cavity. To allow flexibility in the xy-plane of the ring, I applied flat-bottom restraints similar in set-up as in the umbrella simulations. All restraints were lifted for production runs of 100 ns. I set up three repeats with the set-up described above and in all three repeats, the chloride with modified VDW parameters left the cavity almost immediately. In the combined simulation time of 300 ns, I did not observe any re-binding event in Fig. B.8(b). I also calculated the potential of mean force profile along the reaction coordinate for the chloride with modified VDW parameters and for fluoride, which has an even smaller VDW radius, see Fig. B.8(c). Both ions cannot be stably bound by the biotin[6]uril host. Even though the interaction energies between host and ions with smaller VDW radius are favorable in Fig. B.8(d) and (e), the ions do not bind centrally to the cavity.

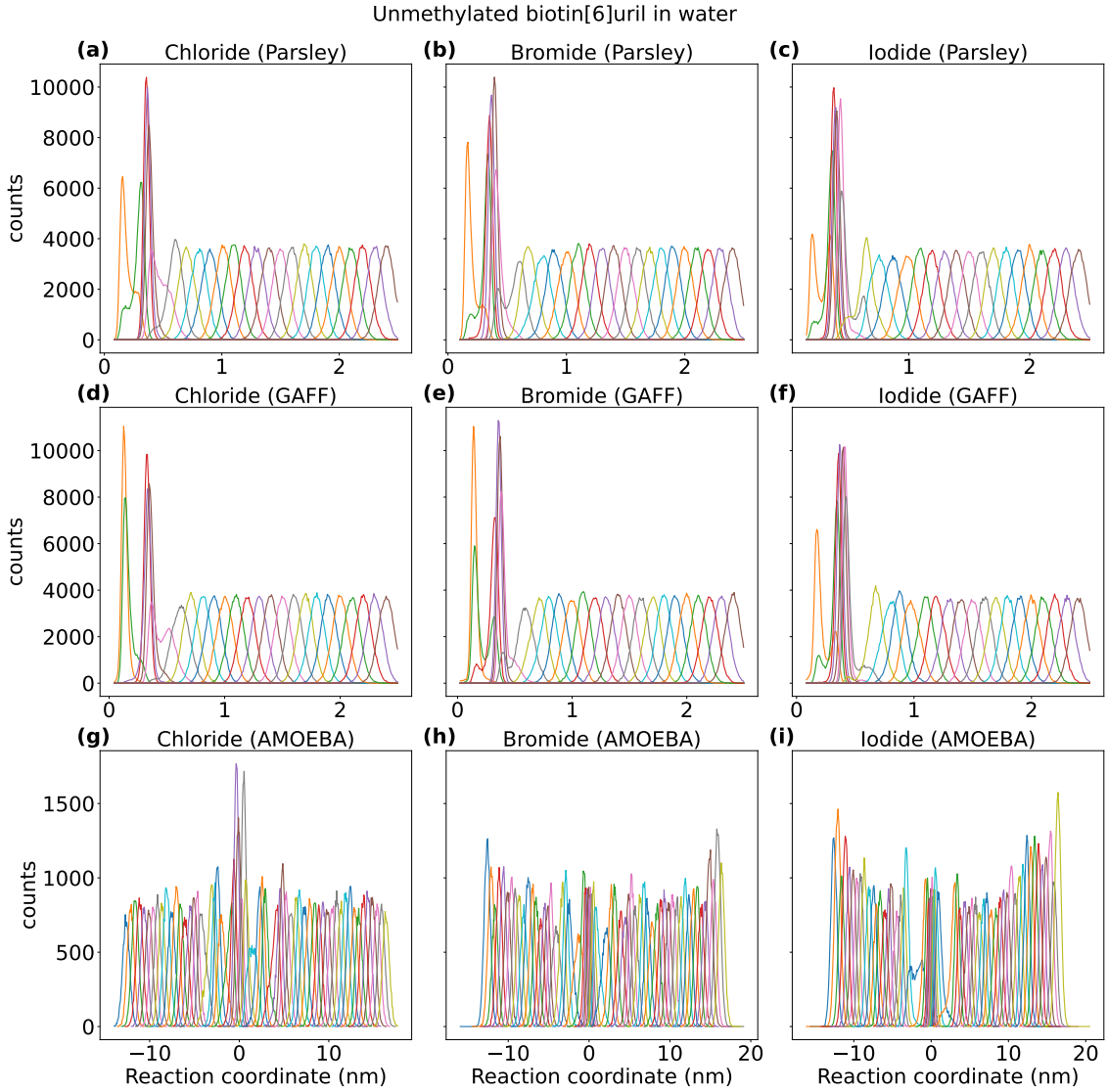


Figure B.3: Histogram overlap for unmethylated biotin[6]uril in water with Parsley (a, b and c), GAFF (d, e, f) and AMOEBA (g, h, i).

B.1.4 Interaction Energy

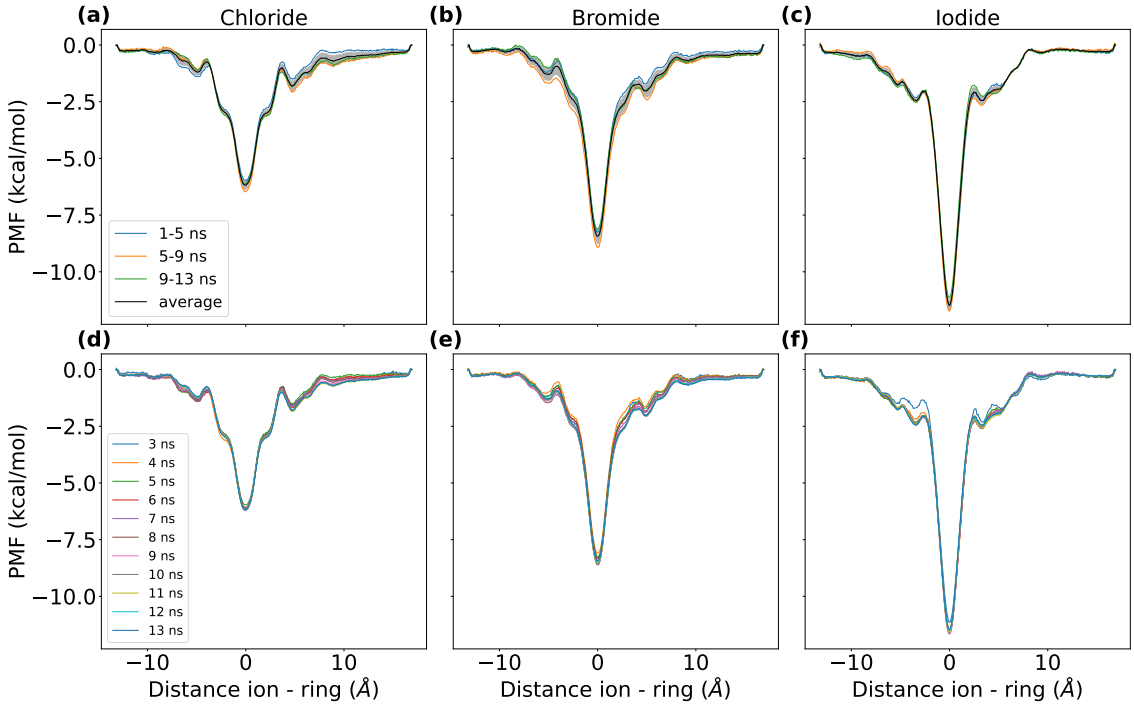


Figure B.4: Convergence analysis for methylated biotin[6]uril in water with AMOEBA.

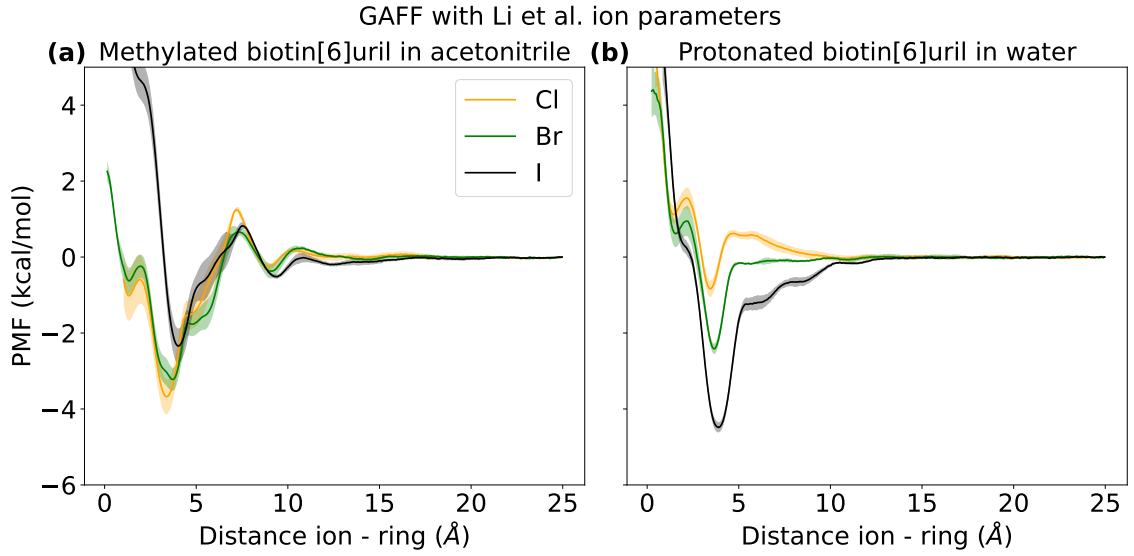


Figure B.5: Potential of mean force (PMF) of ions and a biotin macrocycle with methylated side chains in acetonitrile (a) and an unmethylated macrocycle in water (b) in protonated forms calculated with GAFF and Li*et al.* ion parameters. Each PMF here corresponds to the average of three data sets. Error bars represent one standard deviation. The ions are not centrally bound in the cavity and the PMF profile look similar to those obtained with Parsley, see Fig. 6.2.

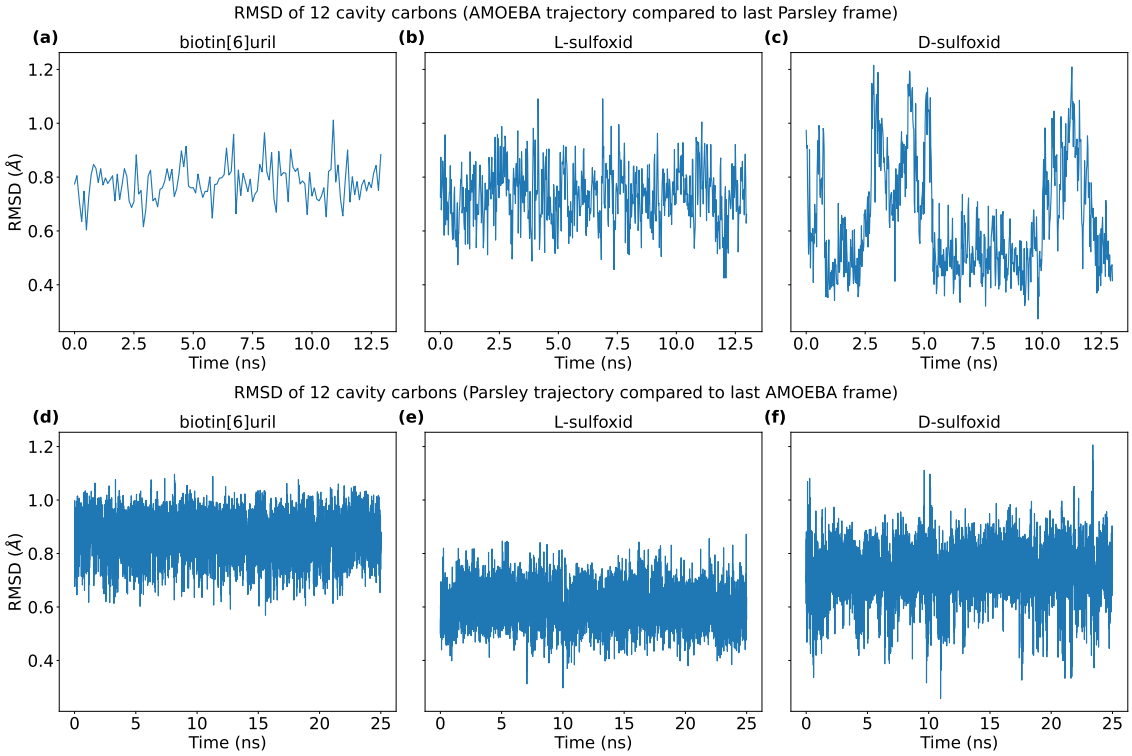


Figure B.6: (a)-(c) Time series of RMSD values comparing the AMOEBA trajectories to the last Parsley frame for three different biotin macrocycles in water. (b) Time series of RMSD values comparing the Parsley trajectories to the last AMOEBA frame for three different biotin macrocycles in water.

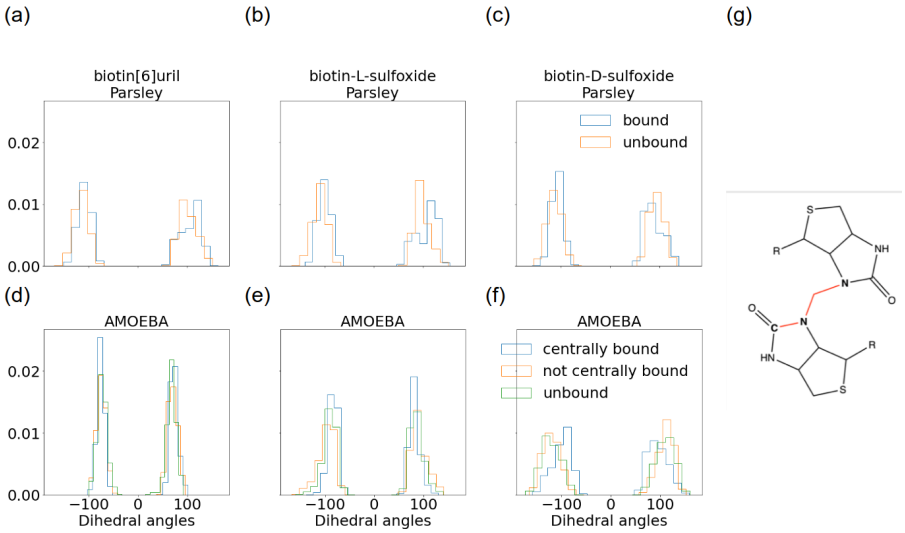


Figure B.7: C-N-C-N Dihedral distributions for three different biotin macrocycles in water with two different forcefields and different binding states. The C-N-C-N dihedral (highlighted in red) in (g) spans atoms from different sub-units and the second carbon atom is the carbon linking two biotin subunits.

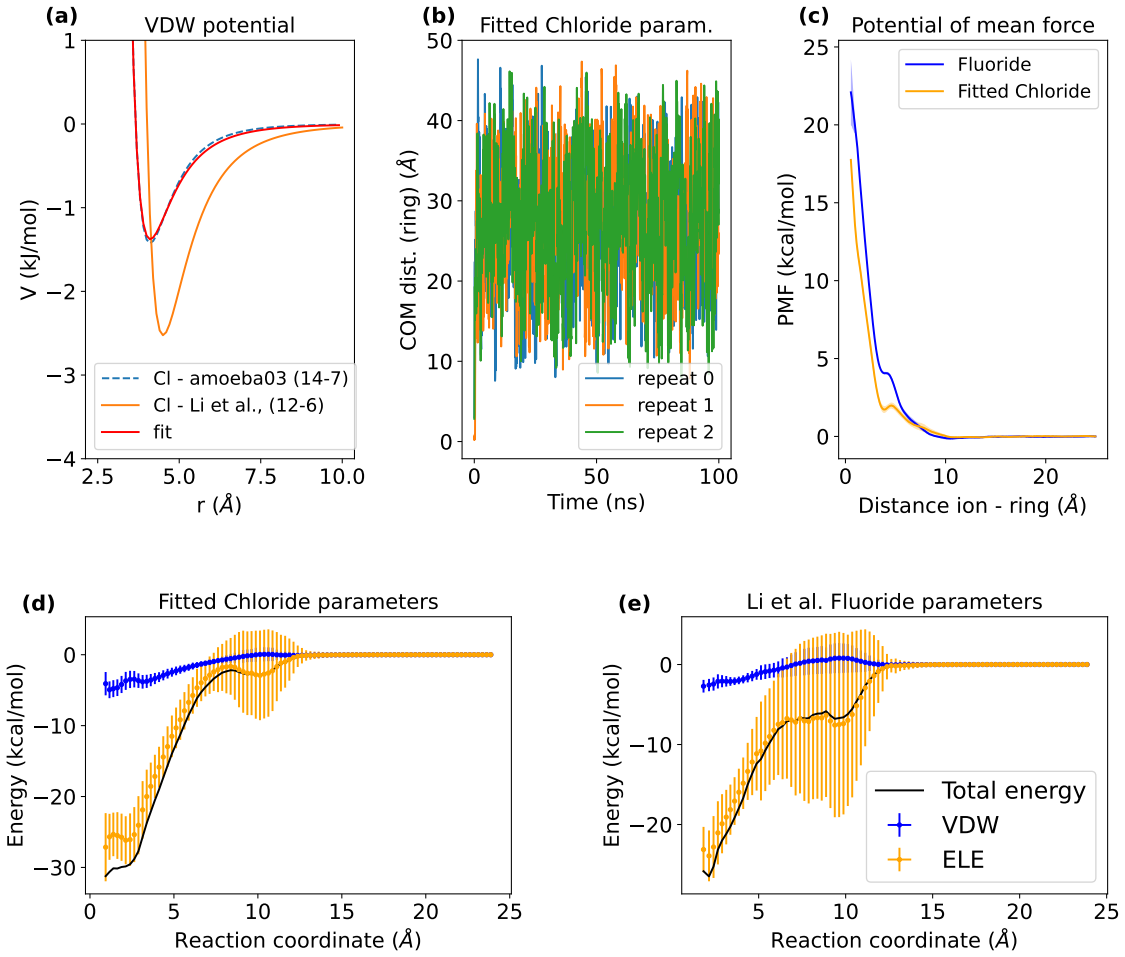


Figure B.8: Van der Waals potential for chloride. (a) Modifying the chloride parameters by Lennard- Jones fitting to resemble more closely the AMOEBA buffered 14-7 potential does not preserve the hydration free energy of chloride in water. The potential is depicted for a chloride-chloride pair. (b) The chloride with modified parameters is not stably bound in the cavities of the macrocycles. (c) The potential of mean force profile for Fluoride (Li *et al.* parameters) and for the modified chloride indicate that neither ion can stably bind to the cavity despite their smaller VDW radius. (d) Interaction energies between the biotin[6]uril host and the Li *et al.* fluoride ion. (e) Interaction energies between the biotin[6]uril host and fluoride with the Li *et al.* parameters.

Table B.2: Interaction energies in kcal/mol for chloride bound to different biotin macrocycles in water. For ten AMOEBA frames, we evaluated the interaction energy with symmetry-adapted perturbation theory (SAPT) on B3LYP level of theory with a jun-cc-pVDZ basis set.

Macrocycle	Interaction energ	Dispersion	Electrostatics	Induction	Exchange
Biotin[6]uril	-56.6 ± 4.5	-17.0 ± 1.1	-41.6 ± 5.3	-23.3 ± 1.1	25.3 ± 3.9
Biotin-D-sulfoxide	-87.7 ± 4.9	-22.1 ± 1.3	-79.3 ± 5.8	-28.3 ± 1.0	42.0 ± 4.7
Biotin-L-sulfoxide	-87.7 ± 4.5	-19.6 ± 1.4	-76.5 ± 4.3	-26.0 ± 1.3	34.4 ± 5.8

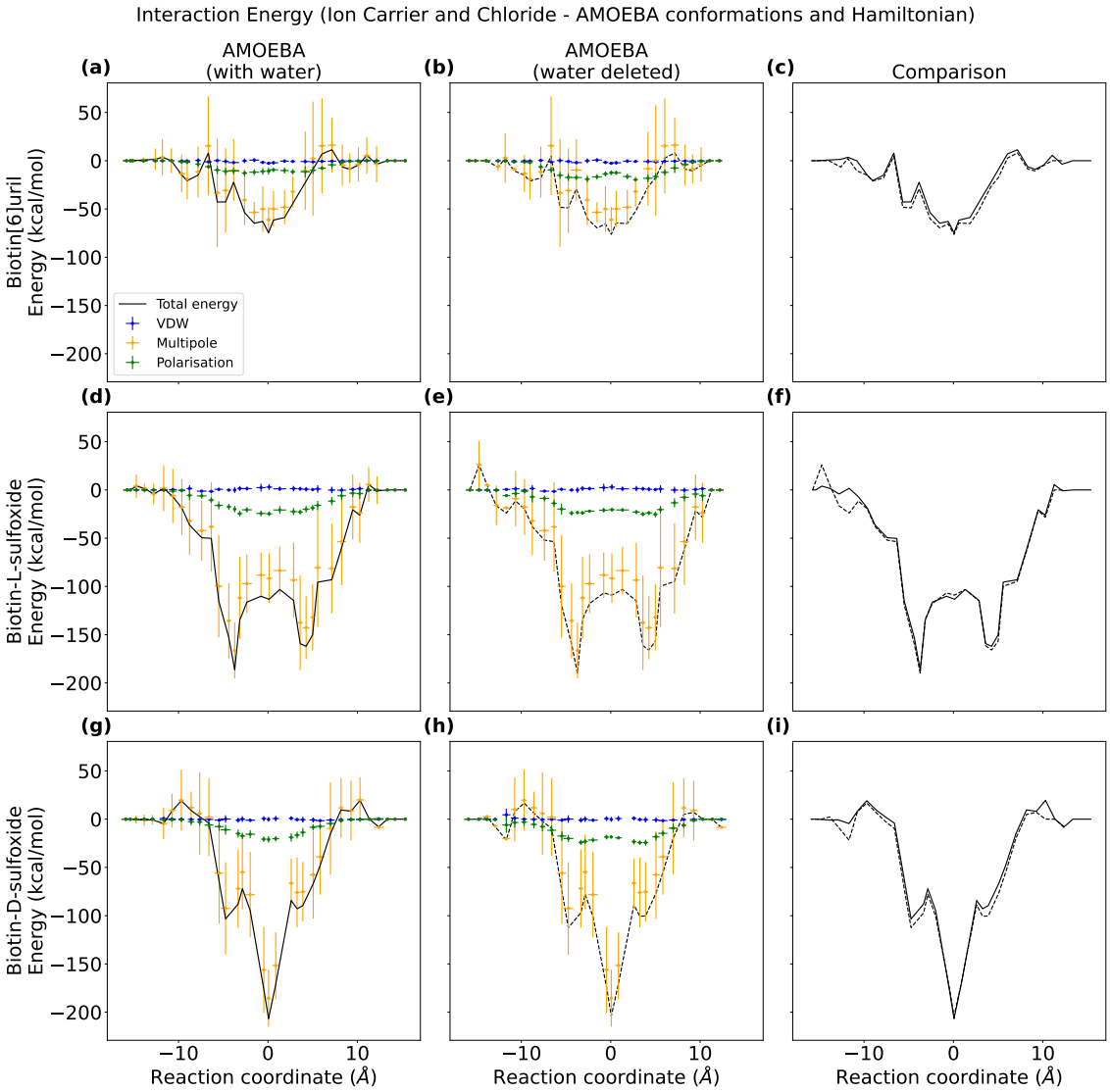


Figure B.9: Interaction energies along the reaction coordinate for three ion carriers and chloride in water evaluated from AMOEBA conformations with (a, d and g) and without water (b, e and h) with the AMOEBA force field. A direct comparison of the total energy is also shown (c, f and i). For each 13 ns long umbrella window, 13 conformations were selected to evaluate the interaction energy.

Table B.3: AMOEBA interaction energies in kcal/mol for chloride bound to different biotin macrocycles in water. For ten AMOEBA frames, we performed energy decomposition analysis with the AMOEBA force field.

Macrocycle	Interaction energy	VDW	Electrostatics	Polarisation
Biotin[6]uril	-75.1 ± 12.3	-3.3 ± 1.5	-60.9 ± 11.1	-10.8 ± 3.2
Biotin-D-sulfoxide	-110.9 ± 19.2	2.4 ± 2.0	-89.2 ± 19.0	-24.1 ± 2.2
Biotin-L-sulfoxide	-208.2 ± 24.1	0.2 ± 2.6	-189.4 ± 23.0	-19.0 ± 2.5

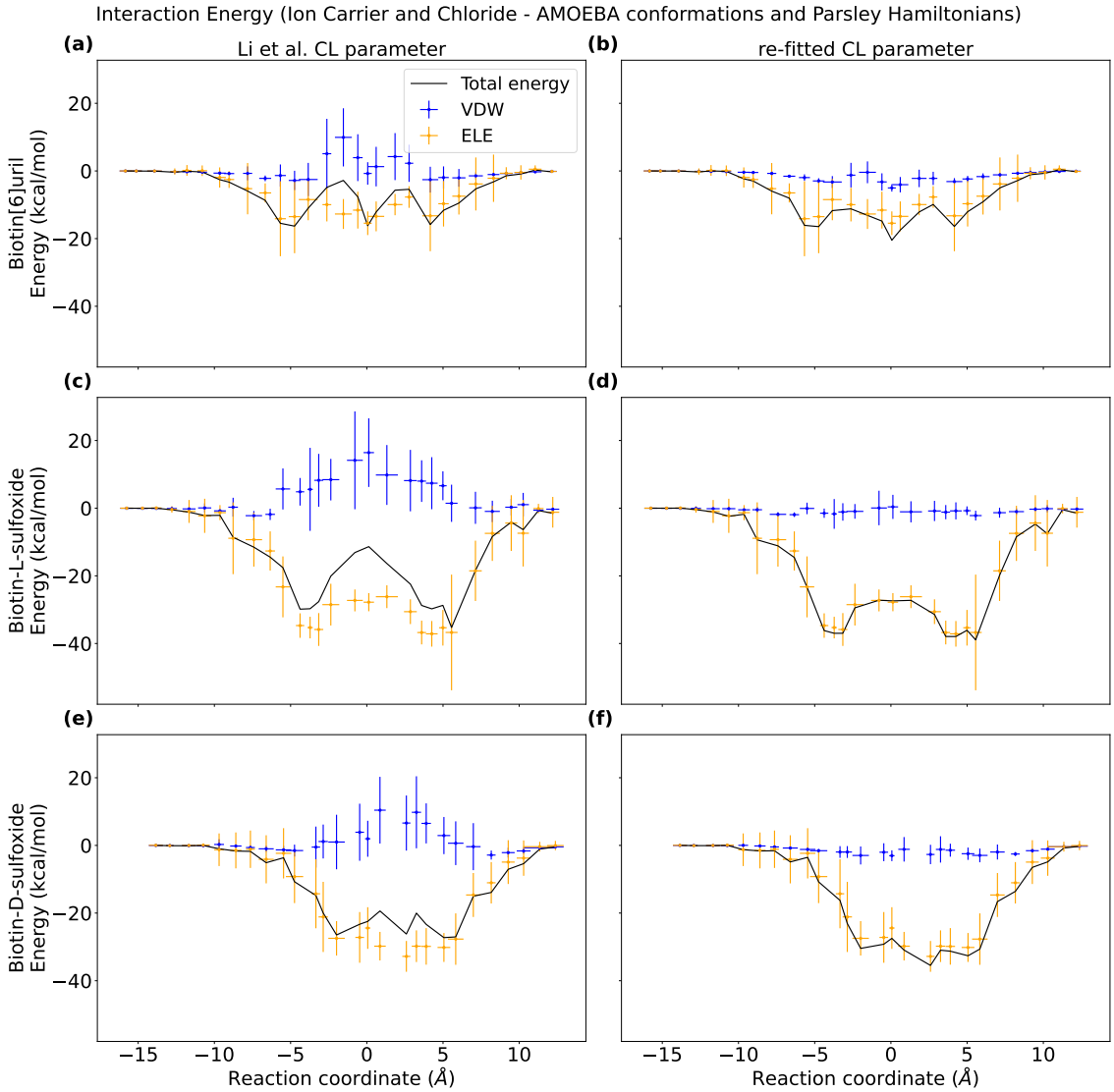


Figure B.10: Interaction energies along the reaction coordinate for three ion carriers and chloride in water evaluated from AMOEBA conformations with the Parsley force field for the host and two different chloride parameter sets. For each 13 ns long umbrella window, 13 conformations were selected to evaluate the interaction energy. The Li *et al.* parameter set and the modified/re-fitted chloride parameter set are compared in Figure S4.

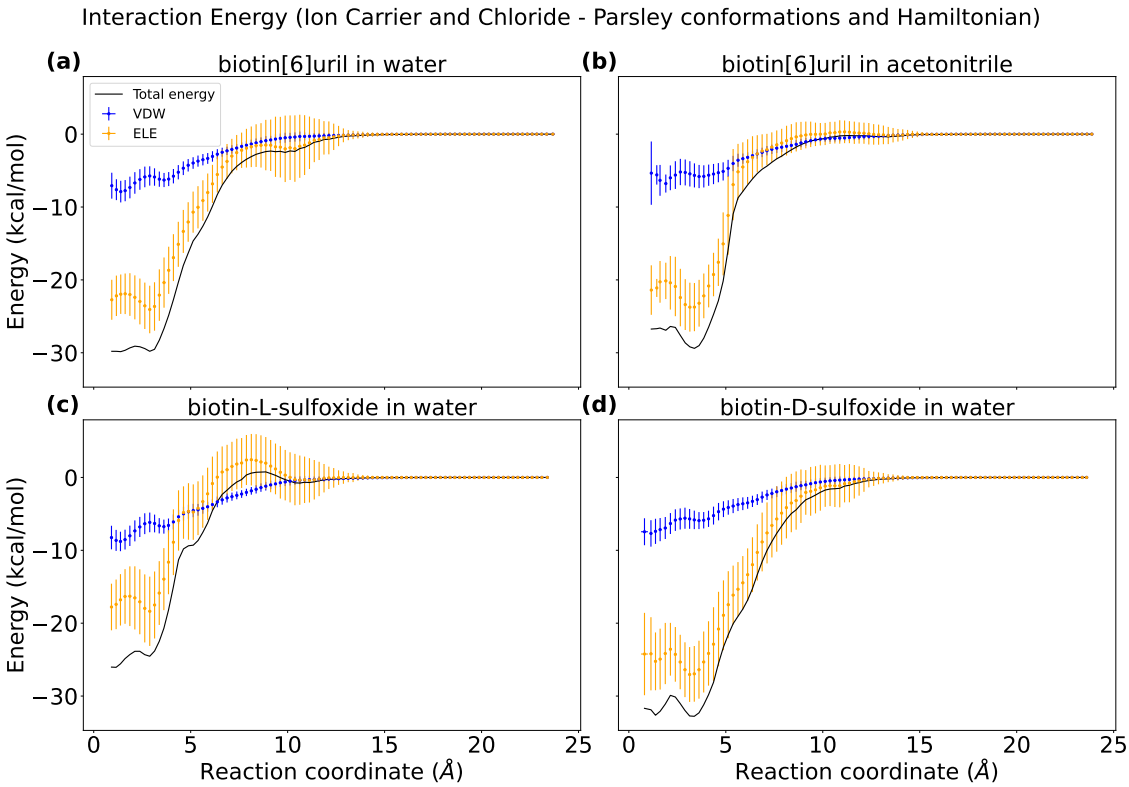


Figure B.11: Interaction energies along the reaction coordinate for three ion carriers and chloride evaluated from Parsley conformations with the Parsley force field for the host and the Li *et al.* chloride parameter set. For each 25ns long umbrella window, 1250 conformations were selected to evaluate the interaction energy.

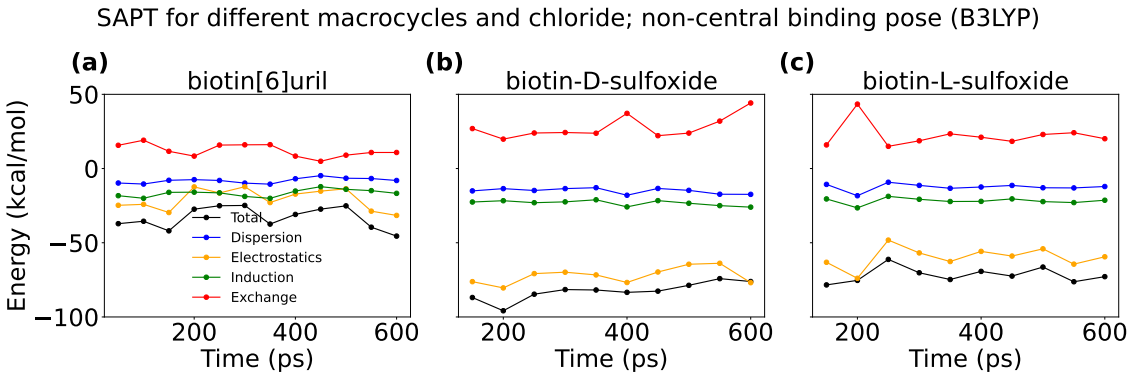


Figure B.12: Interaction energies between chloride and three hosts (biotin[6]uril, biotin-D-sulfoxide[6]uril and biotin-L-sulfoxide[6]uril). For ten AMOEBA frames and the non-central binding pose, we evaluated the interaction energy with symmetry-adapted perturbation theory (SAPT) on B3LYP level of theory with a jun-cc-pVDZ basis set.

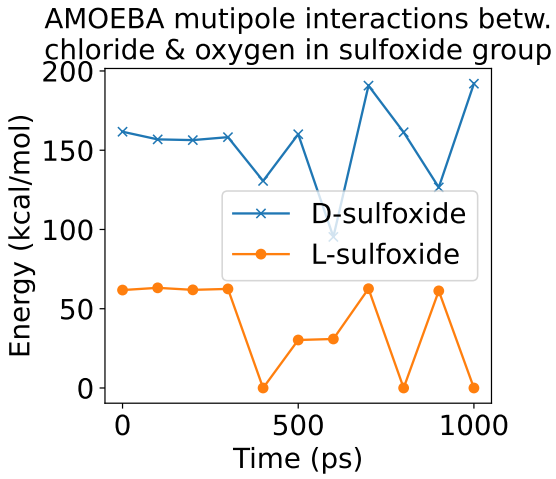


Figure B.13: Multipole interactions between chloride and the oxygens in the sulfoxide groups.

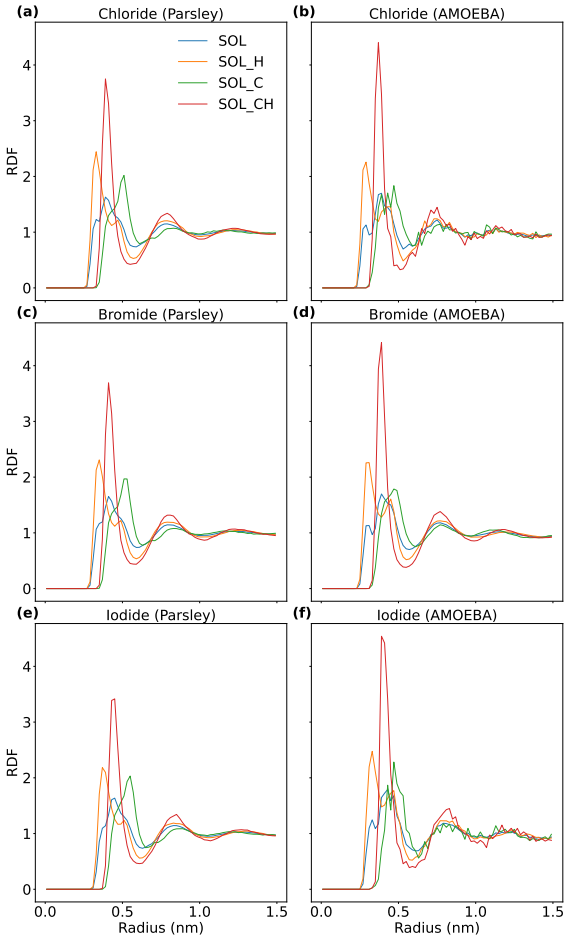


Figure B.14: Radial distribution functions (RDF) for halide anions and acetonitrile. SOL refers to the whole acetonitrile molecules, H refers to hydrogen, C to the carbon bound to nitrogen and CH to the carbon bound to three hydrogens. The RDFs are used to determine the cut-offs used in Table. S3 to analyse the hydration shell in umbrella windows.

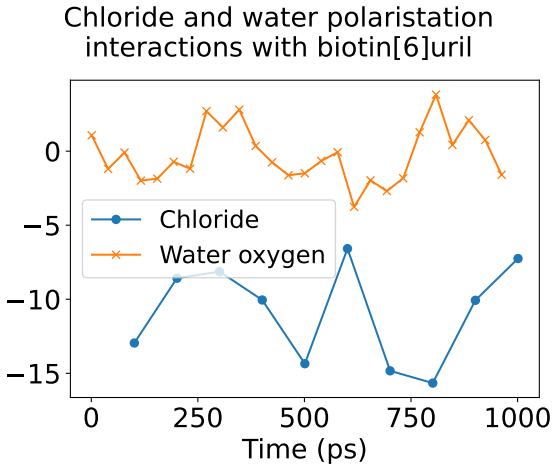


Figure B.15: Interaction energies between chloride and the biotin[6]uril macrocyle and a water molecule bound in the cavity and the macrocycle. The guests are both centrally bound in the cavity. As expected, the chloride has a stronger polarization interaction with the host than a water molecule.

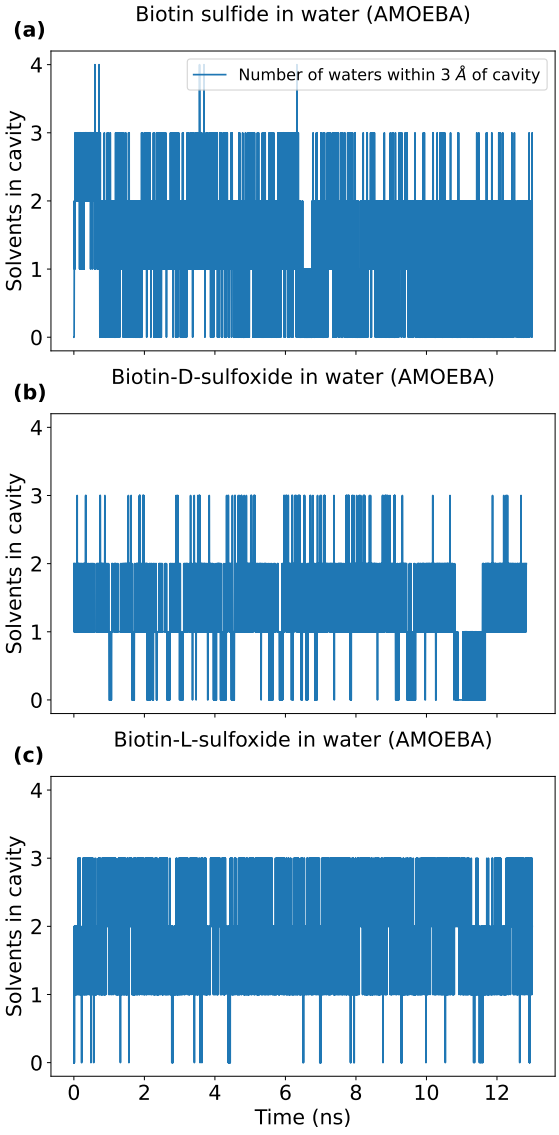


Figure B.16: Time series of number of solvent molecules within 3 Å of the cavity center for three different biotin macrocycles in water. A representative snap-shot of a hydrated and a dewetted cavity are shown in Fig. 6.11.

B.2 GlyR

B.2.1 PMF

To understand the energetics of ion permeation through the GlyR pore, I employed umbrella sampling simulations to calculate the potential of mean force (PMF) as described in section 3.5. This allowed me to compare the free energy profile of ion movement along the pore axis for simulations employing different force fields. I used the Drude Prepper [155] within CHARMM-GUI to convert the simulation boxes with CHARMM36m force field [151] parameters to those compatible with the CHARMM Drude polarisable force field [84]. Good overlap between neighbouring umbrella sampling windows was observed in Fig. B.18. In my PMF calculations, I observed different energy profiles for the Drude polarisable force field and the non-polarisable CHARMM36m force field with pair-specific Lennard Jones parmeters (NBFIX) [151, 114].

B.2.2 Convergence of PMFs

The Drude PMFs for the 6pm2 and 7m6q structures exhibit a lower energy barrier compared to CHARMM36m, indicating more favourable conditions for chloride ion passage. Chloride anions partially lose their hydration shell when passing through the GlyR pore, see Fig. B.17. The PMF profiles in Fig. B.17 show distinct barrier characteristics at critical points along the pore. The proline (P) at the -2' position marks the narrowest constriction along the permeating path and the leucine (L) at the 9' position marks a secondary barrier. The energetic barrier associated with the constrictions at P-2' and L9' are within error of each other for PMFs based on the CHARMM36m force field. Chloride ions described by the CHARMM36m force field have to overcome two barriers of similar height when they permeate the GlyR pore. However, the barrier associated with the narrowest constriction at P-2' is significantly smaller than the barrier at the slightly wider

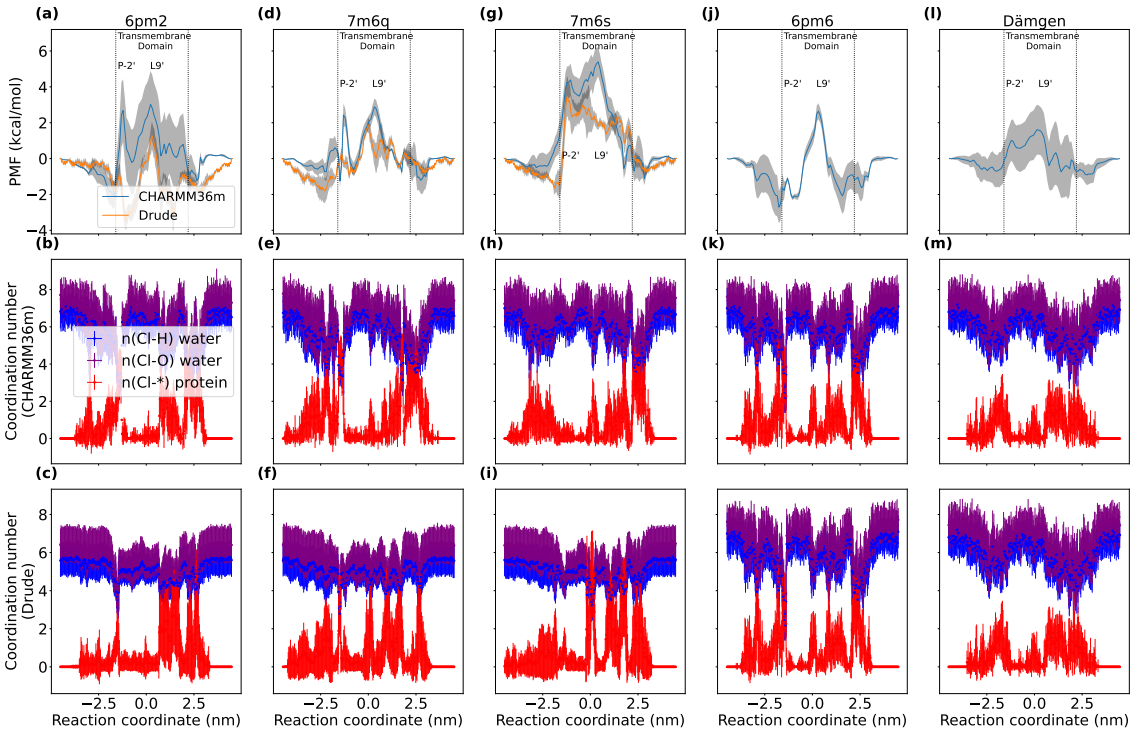


Figure B.17: Potential of mean force of a chloride ion permeating pores of five GlyR structures (a,d,g,j,l). Number of water oxygen, hydrogen and protein atoms coordinating the chloride anion in simulations with the CHARMM36m force field (b,e,h,k,m) and the CHARMM Drude force field (c,f,i).

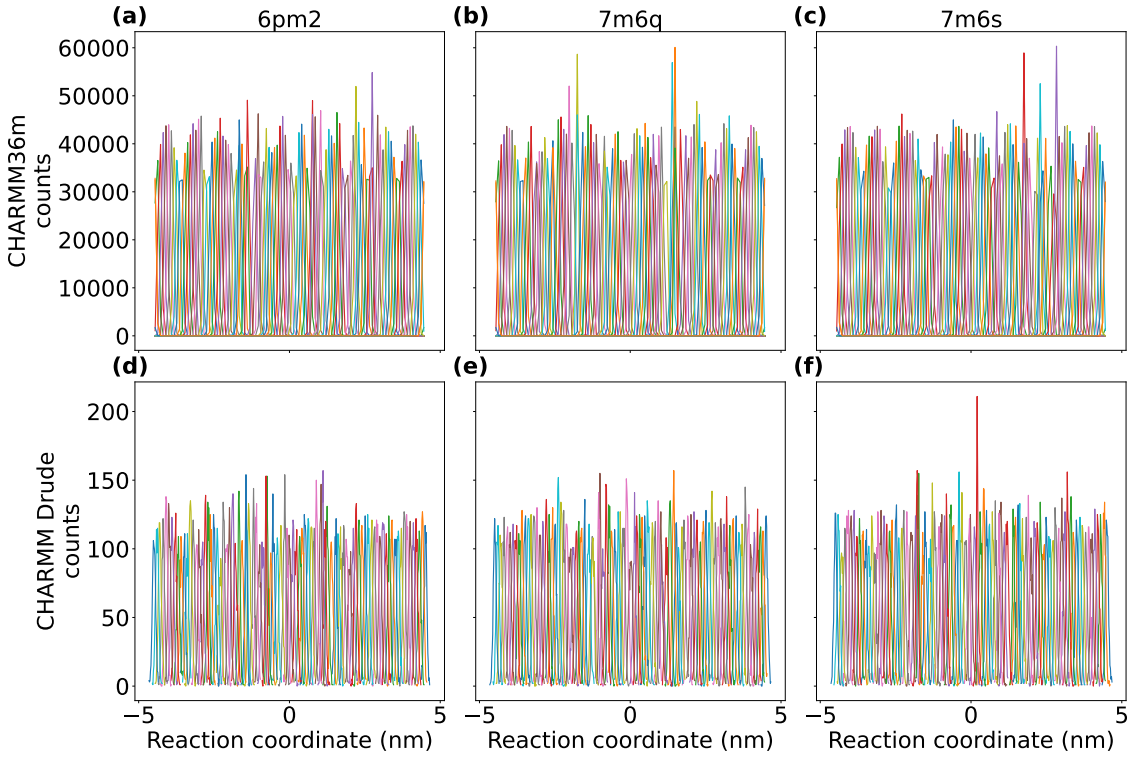


Figure B.18: Histogram overlap for different GlyR models with CHARMM36m (a, b and c) and with CHARMM Drude (d, e, f).

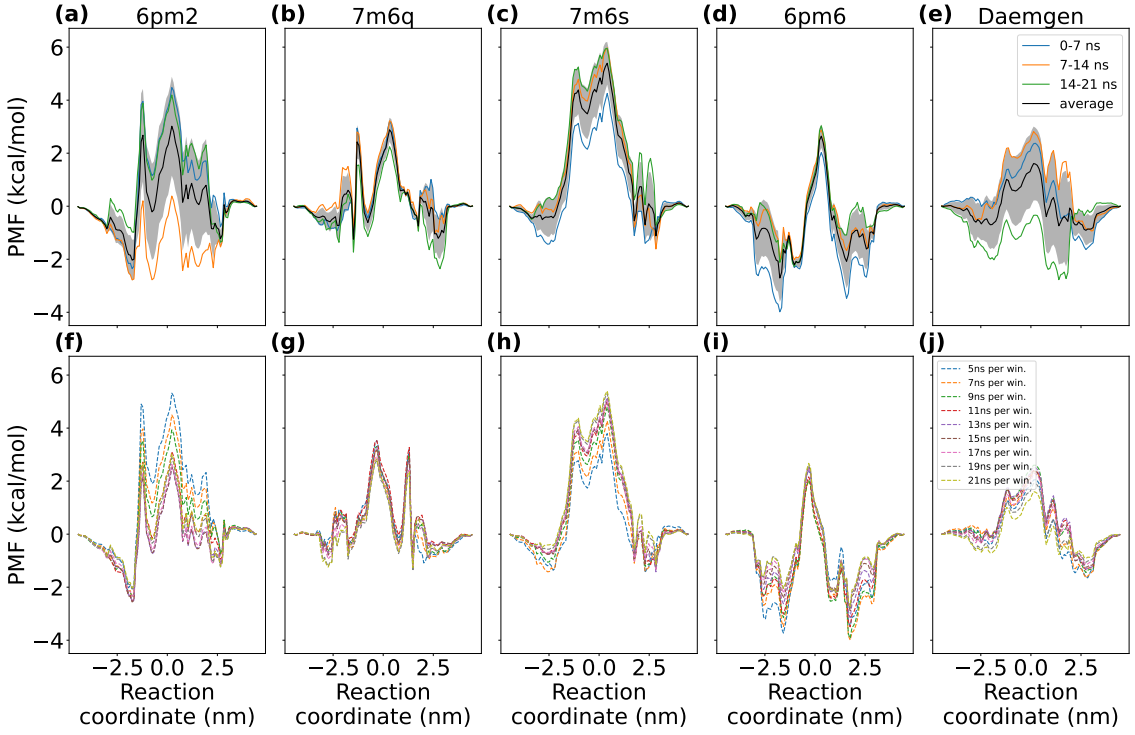


Figure B.19: Convergence of CHARMM36m PMFs.

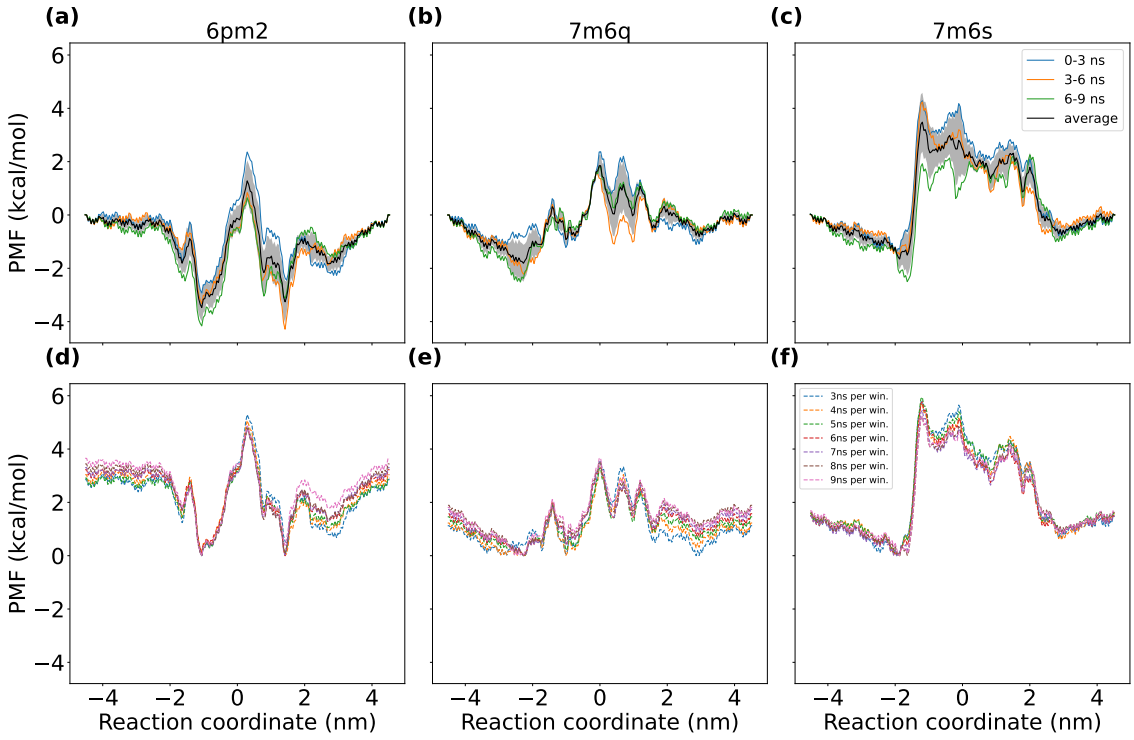


Figure B.20: Convergence of CHARMM Drude PMFs.

Table B.4: Summary of pore constrictions and their associated PMF barrier. The 6pm2 structure was solved in a stryrene maleic acid polymer (SMA) environment [34]. Kumar *et al.* reconstituted GlyR into a lipid nanodisc in complex with Δ^9-Tetrahydrocannabinol (THC) under varying concentrations of glycine (gly).

	PDB code	6pm2	7m6q	7m6s
	state	GlyR-taurine-open-SMA	GlyR-1gly-THC-state1	GlyR-1gly-THC-state3
	authors	Yu *et al.* [34]	Kumar *et al.* [247]	
CHARMM36m	PDB constriction [Å] at Pro -2'	2.5	2.6	3.2
	Constriction (sim. av) [Å]	2.5	2.5	2.9
	Barrier[kcal/mol]	3.0 ± 1.9	2.9 ± 0.5	5.4 ± 0.8
	Position	P-2' and L9' gate within error	P-2' and L9' gate within error	L9' gate
Drude	Barrier[kcal/mol]	1.3 ± 0.8	1.8 ± 0.4	3.5 ± 1.1
	Position	L9' gate	L9' gate	P-2' gate

L9' constriction in the Drude PMF for structures 6pms and 7m6q. The highest barrier for permeation for structures 6pm2 and 7m6q is not associated with the narrowest constriction at P-2' but with the secondary constriction at L9'. The difference between the barrier associated with P-2' and L9' is larger for Drude PMFs than for CHARMM36m PMFs. While the CHARMM36m force field predicts barriers of roughly equal height at both constriction sites, the Drude force field shows almost no barrier at P-2', while the only significant barrier is observed at the L9' position. The 7m6s structure represents a corner case, with the P-2' gate originally having a 3.2 Å restriction in the PDB entry, which reduces to 2.9 Å in the simulation, see Tab. 7.1. For this structure, I observe a higher barrier to permeation for both force fields: The CHARMM36m force field shows the highest barrier at the L9' position, whereas the Drude force field shows the highest barrier at the P-2' position. In both force fields, the maximum barrier for chloride permeating the 7m6s structure is larger than for the other two structures. I note that the barrier associated with the constriction P-2' is significantly smaller for the Drude PMF compared to the CHARMM36m PMF for structures 6pm2 and 7m6q. For 7m6s, depicted in Fig. B.17(g), the barrier associated with P-2' is within error for both force fields. The mean distance between the nucleus of the chloride restrained in the umbrella sampling windows and its Drude particle is plotted in Fig. B.25. The standard deviation drops at the -2' position.

I depict the distribution of the number of water hydrogens and protein atoms coordinating the chloride anion at the P-2' and the L9' constriction in Fig. B.21. Chloride is heavily coordinated by proline at the -2' position in simulations with CHARMM36m and moderately coordinated in simulations with the CHARMM Drude force field for structures 6pm2 and 7m6q, see Fig. B.21(a) and (c). Apart from the behaviour at the -2' position for structures 6pms and 7m6q, there does not seem to be a significant difference between the CHARMM36m and Drude force field in terms of chloride coordination at the narrowest constriction.

B.2.3 Position of Side Chains

In Fig. B.22, I check whether leucine sits in the hydroppobic pocket or is pore facing. I calculate the COM of the leucine backbone and the COM of its side chain. Then I calculate the distance d in the xy-plane of both COMs to the protein COM and the orientation o of respective leucine residue:

$$\begin{aligned} d_i &= \sqrt{(\mathrm{COM}_i(x) - \mathrm{COM}_{\text{protein}}(x))^2 + (\mathrm{COM}_i(y) - \mathrm{COM}_{\text{protein}}(y))^2} \\ o &= d_{\text{side chain}} - d_{\text{back bone}} \end{aligned} \tag{B.1}$$

If the orientation o is positive, then the COM of the leucine side chain is further away from the channel axis then the leucine backbone and it is completely in the hydrophobic pocket. The more negative the orientation score o, the closer is the side chain to the pore. The leucine side chains are not completely in the hydrophobic pockets but also do not block the ion permeation pathway. Note that back bone restraints were applied during the simulations with the external potential.

B.2.4 Correlation between Conductance and Potential of Mean Force

In Chapter 5, I observed the expected negative correlation between conductance and PMF barrier height for the carbon nanotube model systems. In Fig. B.26, I plot the conductance against the PMF barrier for the GlyR systems. For the CHARMM36m force field with NBFIX parameters, I observe negatvie correlation while for the (unconverged) CHARMM Drude force field, I observe surprsingly positive correlation.

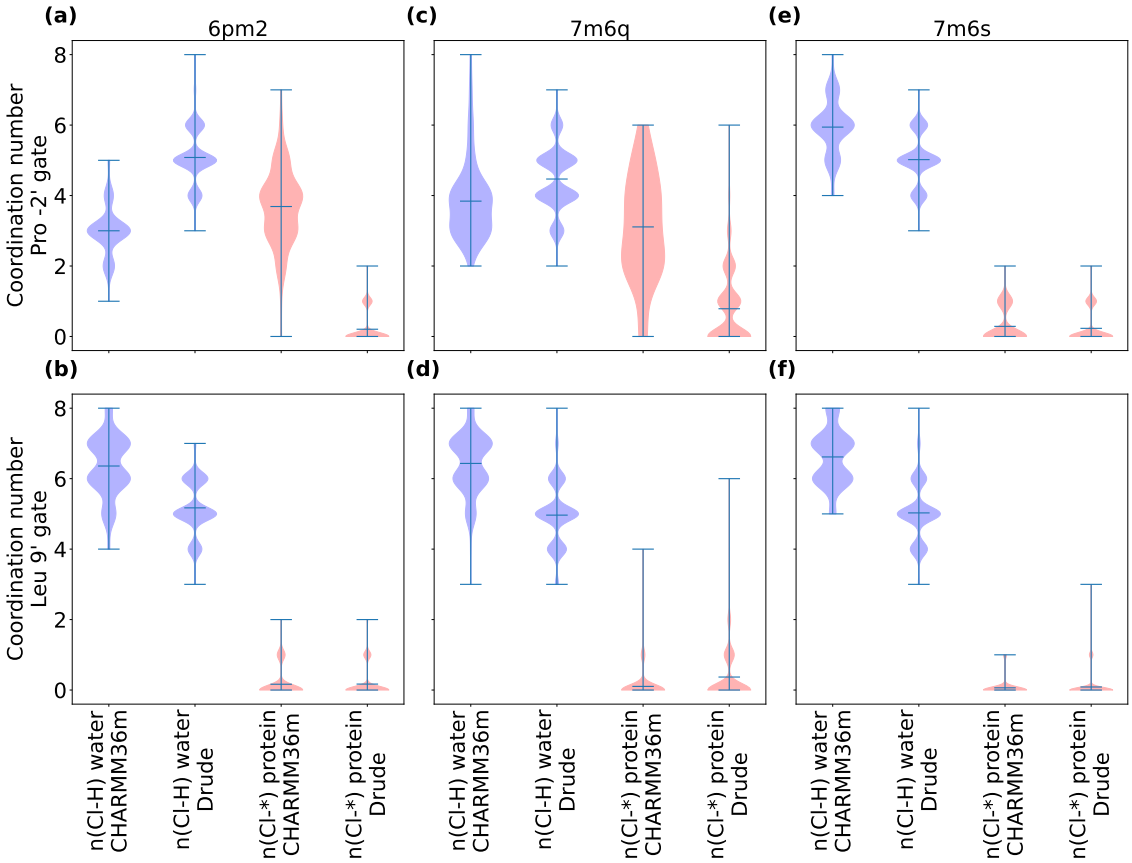


Figure B.21: Number of water hydrogens and protein atoms coordinating the chloride anion at the proline -2' (a,c,e) and the leucine 9' constriction in simulations with the CHARMM36m force field and the CHARMM Drude force field.

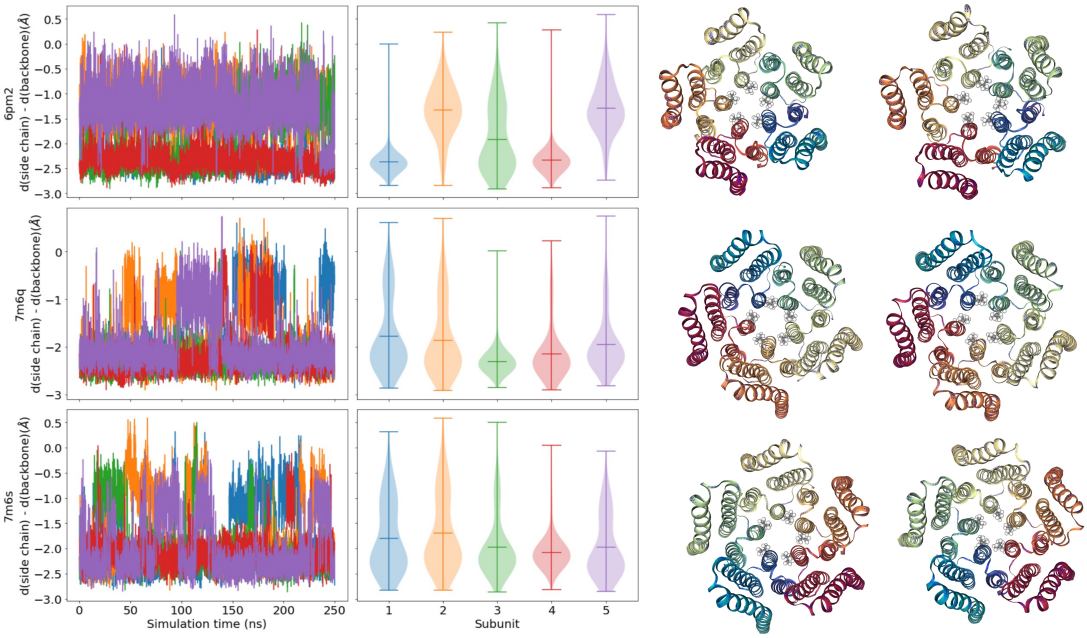


Figure B.22: Orientation o of leucine at 9' position during 250 ns simulations with 500 mV external potential as calculated in eq. (B.1). For each structure, we present the conformation after minimisation and the final frame after 250 ns simulations with an external potential and backbone restraints. The more negative the orientation o, the more pore facing is the respective leucine side chain at 9' position.

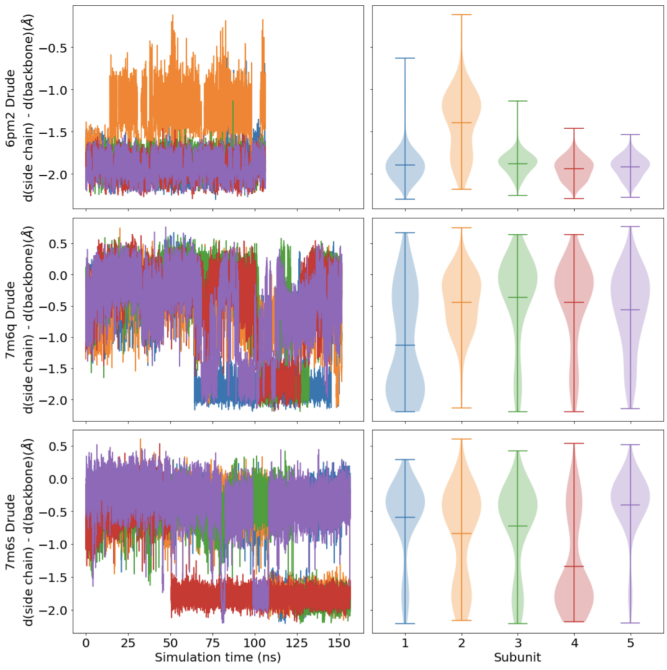


Figure B.23: Orientation o of leucine at 9' position during CHARMM Drude simulations with 500 mV external potential as calculated in eq. (B.1). The more negative the orientation o, the more pore facing is the respective leucine sidechain at 9' position.

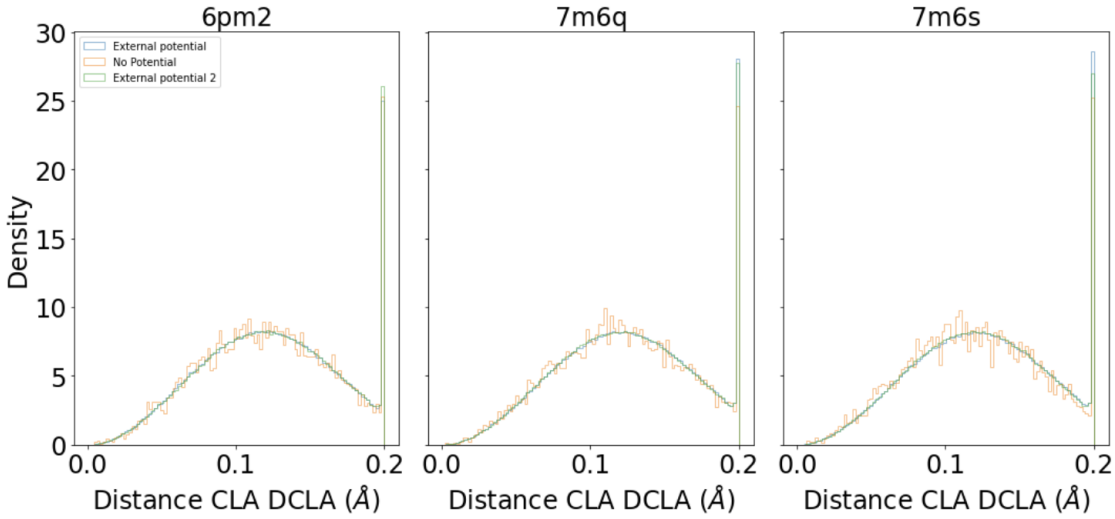


Figure B.24: Distribution of the distance between chloride nucleus and Drude particle in simulations with no external potential and with external potential.

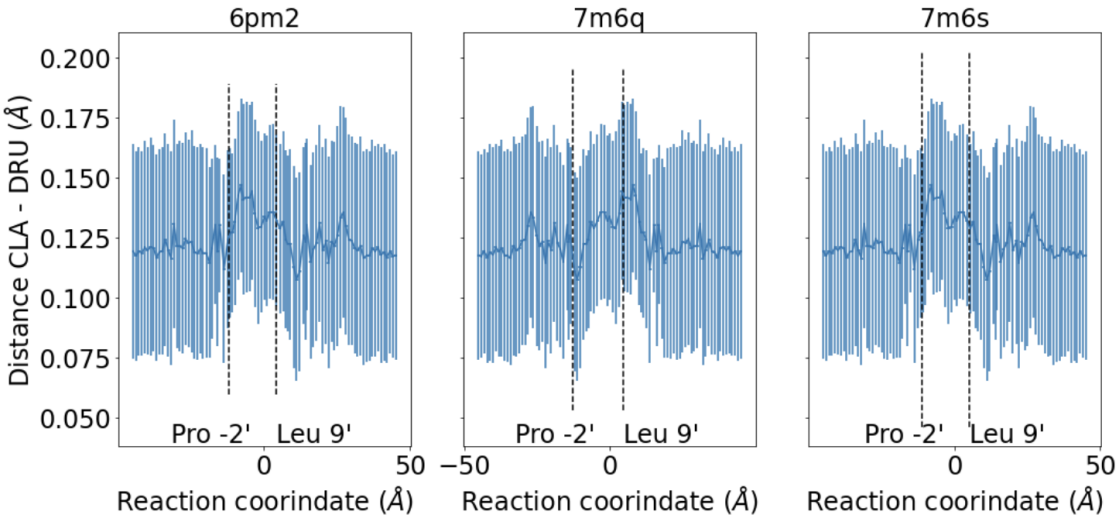


Figure B.25: Distance between chloride nucleus and Drude particle as a proxy for polarisation of chloride.

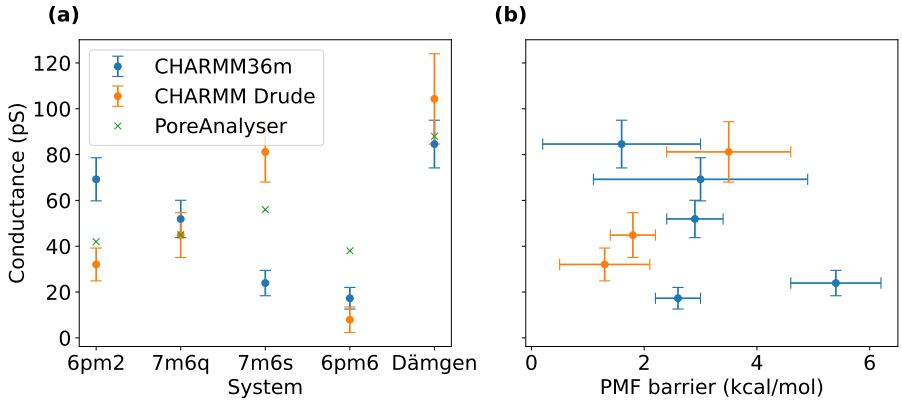


Figure B.26: Conductance comparison: (a) Conductance estimates based on PoreAnalyer profiles and MD simulations with two force fields for the GlyR structures in Table 7.1. (b) The conductance estimates based on MD simulations with external potential against the barrier to permeation based on potential of mean force from Fig. B.17.