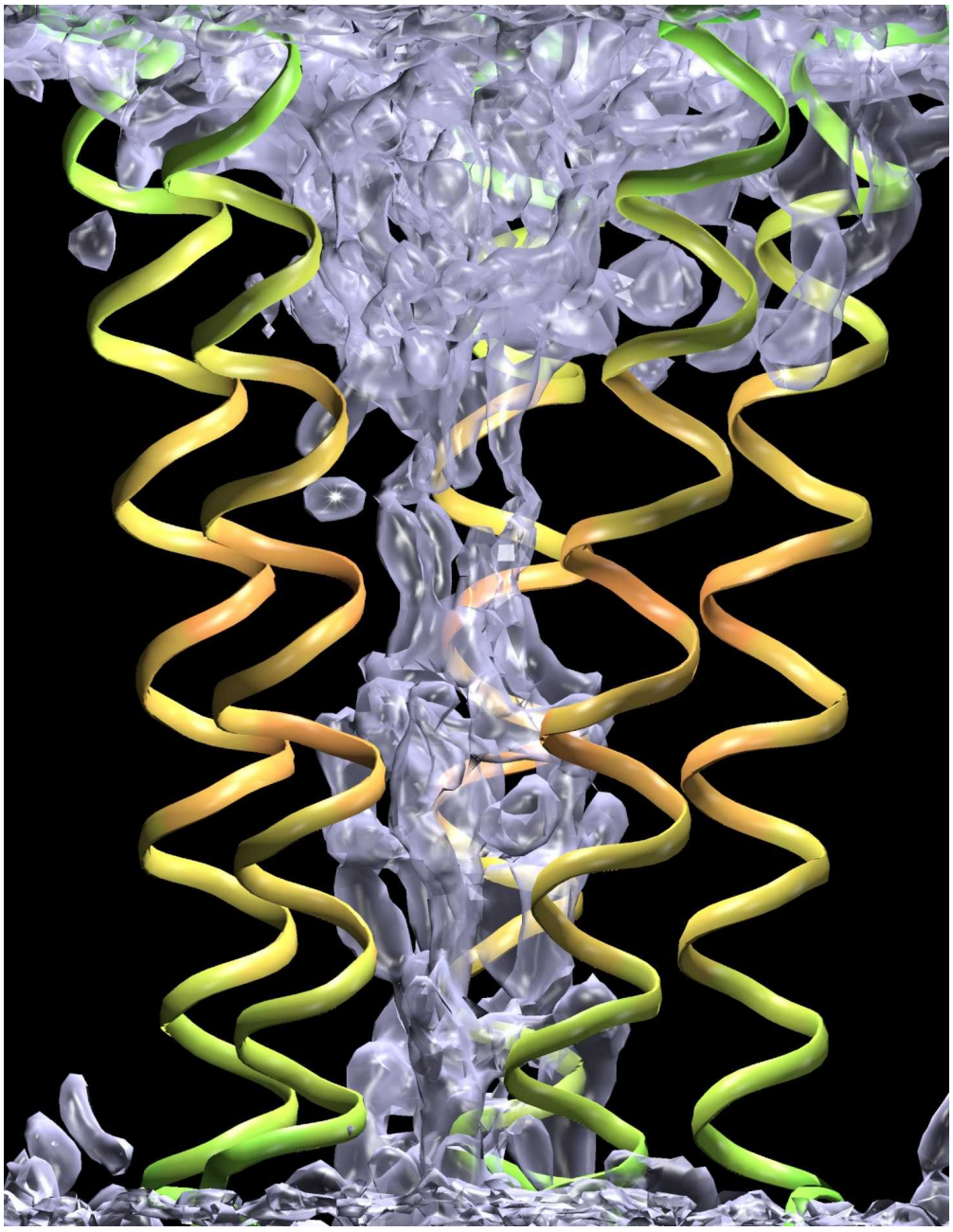
Principles of Gating Mechanisms of Ion Channels



Oliver Beckstein

Principles of Gating Mechanisms of Ion Channels

Oliver Beckstein

Laboratory of Molecular Biophysics and
Merton College, Oxford

Michaelmas 2004

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

First version 2005-05-29.
Second version 2007-08-13 with minor corrections.
Third version 2008-08-09 with minor corrections.
Fourth version 2012-01-08 with minor corrections.

Viva voce examination held on 7 March 2005 by Dame Prof. Louise N. Johnson (University of Oxford) and Prof. Jean-Pierre Hansen (University of Cambridge).

This document was set in 12pt Palatino with LaTeX, using KOMA-Script. It was generated with the command `make thesis`. See the author's homepage `http://sbcb.bioch.ox.ac.uk/users/oliver/` for electronic copies, movies, errata, and software described in Appendix D.
The title page shows the water density in the transmembrane region of the nicotinic acetylcholine receptor, which is formed by the five M2 helices. The density is contoured at the bulk density. The region of broken density marks the hydrophobic gate.

Principles of Gating Mechanisms of Ion Channels

Oliver Beckstein
Laboratory of Molecular Biophysics and Merton College, Oxford
Submitted for the degree of Doctor of Philosophy, Michaelmas 2004

Abstract

Ion channels such as the nicotinic acetylcholine receptor (nAChR) fulfil essential roles in fast nerve transmission and cell signalling by converting an external signal into an ionic current, which in turn triggers further down-stream signalling events in the cell. Increasing structural evidence suggests that the actual mechanisms by which channels *gate* (i.e. switch their ion permeability) are fairly universal: Conduction pathways are either physically occluded by localised sidechains or the pore is narrowed by large-scale protein motions so that a constriction lined by hydrophobic sidechains is formed. In this work the latter mechanism, termed *hydrophobic gating*, is investigated by atomistic computer simulations.

Simple hydrophobic model pores were constructed with dimensions estimated for the putative gate region of nAChR (length 0.8 nm, radius varied between 0.15 nm and 1.0 nm). In long classical molecular dynamics (MD) simulations, water confined in the pore was found to oscillate between a liquid and a vapour phase on a nano second time scale. Water would rarely permeate a pore less wide than three water molecules. A simple thermodynamic model based on surface energies was developed, which explains the observed *liquid-vapour oscillations* and their dependence on pore radius and surface hydrophobicity. Similarly, Na^+ ion flux is only appreciable for pore radii greater than 0.6 nm. Calculation of the free energy profile of translocating ions showed barriers to permeation of greater $10\,kT$ for pore radii less than 0.4 nm. Comparison to continuum-electrostatic Poisson-Boltzmann calculations indicates that the behaviour of the solvent, i.e. water, is crucial for a correct description of ions in apolar pores. Together, these results indicate that a hydrophobic constriction site can act as a *hydrophobic gate*.

An ongoing debate concerns the nature and position of the gate in nAChR. Based on the recent cryo-electron microscopy structure of the transmembrane domain at 4 Å resolution, and using techniques established for the model pores, equilibrium densities and free energy profiles were calculated for Na^+, Cl^-, and water. It was found that ions would have to overcome a sizable free energy barrier of about $10\,kT$ at a hydrophobic girdle between residues L9′ and V13′, previously implicated in gating. This suggests strongly that nAChR contains a hydrophobic gate. Furthermore, charged rings at both ends of the pore act as concentrators of ions up to about six times the bulk concentration; an effect which would increase the ion current in the open state. The robustness of the results is discussed with respect to different parameter sets (force fields) and the applied modelling procedure.

Acknowledgements

Had it not been for Mark Sansom (MSPS) and Nigel Unwin (NU) sitting on a bus from the airport to a conference location some time in 2000, you would read a completely different story. Something like the following conversation must have taken place:*

NU: "Mark, I wondered lately. The nicotinic receptor pore seems rather wide but probably very hydrophobic. Do you think that this would actually be enough to keep ions from going through?"

MSPS: "Hmm…I remember a recent paper[1] where they did some simulations of water in hydrophobic, spherical cavities and said it did not stay in there."

NU: "Interesting."

MSPS: "Well, might be a dumb idea, but we could set up a simple pore and do some MD on it. You know, I have some motivated students coming in for a short 5 month project; one of them could do it."

NU: "Good, let me know what comes out…We really need some more images to get the resolution down…"

The student was me, and the project proved to be so fruitful that I also spent the next three years on it—thanks to Mark's and Nigel's insightful conversation for which I am very grateful to both of them.

I would also like to thank Mark Sansom for being a very supportive supervisor, who allowed me enough leeway so that I could pursue questions in structural biology which were probably not strictly canonical.

José Faraldo-Gómez was incredibly helpful in my fledgling days as a computational biophysicist; thanks to him the model pores became a reality (as much as such a thing can become real). He masterly played the rôle of the *advocatus diaboli*, which ensured many interesting and enjoyable discussions. I also have to thank the members of the Sansom group, past and present, who all contribute to an environment that is fun to work in. In particular Paul Barrett, Kaihsu Tai, and Carmen Domene, and Jeff Campbell and Jen Johnston made the lab (or the climbing tower…) a place that was not only concerned with work but also a place to hang out with friends. Different people were involved at different stages of the project, and their contributions are listed on a per chapter basis:

*Note: This conversation is reconstructed from rumours and should not be construed as factually accurate in every detail.

Confined Water I would like to thank José Faraldo-Gómez, Andrew Horsfield, Joanne Bright, and Peter Pohl for very interesting discussions of various aspects of water in pores—be it imagined or real.

Hydrophobic Gating I am grateful to Paul Barrett, Kaihsu Tai, Campbell Millar, José Faraldo-Gómez, Benoît Roux, Nigel Unwin, and Bob Evans for their interest, encouragement, and readiness for critical discussions about ions and water (again!) in toy model pores.

Ion PMFs in Model Pores This work was carried out in a very stimulating collaboration with Kaihsu Tai, who carried out many of the Poisson-Boltzmann calculations. Furthermore, I would like to thank Nathan Baker for help with apbs and Born energy calculations, Graham Smith and Marc Baaden for useful discussions about umbrella sampling and WHAM.

A Gate in nAChR and the α7 receptor I gratefully acknowledge discussions in and comments from the nAChR subgroup, namely Shiva Amiri, Phil Biggin, Andy Hung, Jennifer Johnston, and Kaihsu Tai. Shiva Amiri provided the model of the α7 receptor. Nigel Unwin is thanked for his continuing interest and for making available their model of the transmembrane domain of *Torpedo* nAChR prior to publication.

A Consistent Definition of Volume Ideas presented in this appendix evolved through very enjoyable discussions with Andrew Horsfield during his time at the Laboratory of Molecular Biophysics.

Funding The Wellcome Trust's programme in Structural Biology "From Molecules to Cells" generously funded my DPhil. Merton College, Oxford and the British Biophysical Society contributed towards travels to conferences.

However, work is not the only important thing in life, and so I am also indebted to all the friends not mentioned yet who made these four years a great time for me: The First Cohort (Jonathan Malo, Frank Cordes, Maria Höllerer, John Briggs), Werner Bär, Volker Blum, Moritz und Tina Engl, Peter Fleck, Matt Harding, Markus Hormess, Tom Krieger, Carsten Küppersbusch, Uli Lang, Thomas Rechtenwald, and Zyfflich.

I'm lucky in not only having good friends but also a wonderful family. My brothers Nik und Chris are caring and daring company, better than anyone could wish for. My parents, Ursula und Ernst Beckstein, gave me their never-ceasing support and love. Dafür danke ich Euch von ganzem Herzen!

Oxford, November 2004

Contents

List of Figures

List of Tables

1 What is it about?

This work is concerned with *ion channels*, membrane-embedded proteins which act as facilitators of diffusive ion transport through the cell's lipid bilayer. A defining characteristic of ion channel function—*gating*—is examined closer. Physical *principles* are investigated that are employed by protein channels to control the flow of ions or water. A hypothesis is put forward to explain how some known channels are gated. It is tested with the help of computer simulations, both in simple models and a real protein structure, the nicotinic acetylcholine receptor.

1.1 Transport through the cell membrane

A defining step in the evolution of life on earth[2,3] was the appearance of membranes[4,5] and hence the distinction between the inside of a cell and the outside world.[6] Life relies on the efficient progression of chemical reactions.* Compartmentation through membranes greatly facilitates chemical reactions by increasing the concentration of reactive species in the enclosed volume, thus promoting higher reaction rates.[7] Furthermore, the interface between water and a non-polar phase, such as lipids or polycyclic aromatic hydrocarbons, provides a unique microenvironment which accumulates organic molecules and increases their chemical reactivity and also helps peptides to self assemble or to fold through the hydrophobic effect.[6] It also provides a place to anchor macromolecular complexes so to increase the probability of specific interactions between these structures and between the complexes and membrane-associated substrates.[4] Although only measuring about 3 nm in thickness, the phospholipid bilayer provides a formidable permeability barrier to ionic species such as protons or sodium ions. All organisms either use a proton or a Na^+ gradient across the membrane (or both) to store free energy and to drive the production of ATP.[8] But the tight boundary that a membrane provides has also got its drawbacks as it almost completely abolishes the influx of vital ionic nutrients (such as amino acids, nucleotides, and phosphate) or the efflux of toxic waste products. The permeability barrier is so high that only a few solute ions per minute will cross the lipid bilayer. On the other hand, a growing bacterial cell can and must

*Here *life* is viewed as a spatially confined network of chemical and physical processes which harnesses free energy from the surroundings to promote its replication and integrity.

take up millions of nutrient solute molecules per minute.[4] This task requires the help of specialised membrane proteins which facilitate transport across the membrane.[7] Broadly, they can be categorised in two groups: passive *channels* or *pores* and active *carriers*. Active transporters use a source of free energy to move a solute against the electro-chemical gradient; if they use another concentration gradient as their source they are termed co-transporters, if they use chemical energy they are called pumps (for instance, protons are actively pumped out of the cytosol* using either chemical energy from the breakdown of nutrients or the energy of sunlight). Channels and pores facilitate transport down the electro-chemical or osmotic gradient by providing a permeation pathway through the membrane. Lastly, some solutes such as hydrophobic or lipophilic molecules or water can also simply diffuse through the membrane.[7,9]

This is, however, not the whole story. For instance, if fast water transport is required in a cell (as for example in the kidney) specialised *water* channels are expressed in the tissue, which allow water to cross the cell membrane much faster than by diffusion.[10,11] Organisms have evolved to such a degree that any process that is crucial for the functioning of the cell is catalysed by a specific protein, including "simple" transport phenomena. Instead of relying on random diffusion, proteins evolved to speed up very specific processes such as the permeation of Na^+ ions or water through the membrane.

How do membrane proteins shape these transport processes? The question is a complicated one and here we will confine ourselves to passive transport (or assisted diffusion) of ions and water, i.e. the proteins of interest are ion channels[12] and—to a lesser degree—water pores.[11,13] Biochemical and biophysical methods have contributed greatly to our understanding in this area. But in principle much of the function of a protein is already implicitly laid out in its gene sequence. The nascent peptide chain folds into its three dimensional structure which determines its function. So it should be possible to understand the working of transport proteins from the knowledge of the positions of the atoms that make up the protein. The total number of protein structures in the protein data bank[14] has reached 27761 (October 2004) of which 86 are unique membrane proteins, though including only eight ion channel structures.[15] The first atomically-detailed protein structure of an ion channel was only published in 1998,[16] which for the first time enabled the explanation of the characteristics of ion transport in terms of the structural features.

*It is suggested that the direction of the proton gradient reflects the conditions faced on earth by the earliest organisms before ca. 3.9 Gyr.[6,8] The primordial oceans were probably acidic as are the environments near today's deep-ocean geothermal vents, which are considered as possible cradles of life.[2] Hence a proton gradient could be established by extruding protons and increasing the cytosol's pH to a neutral value. Modern eukaryotic cells still maintain a neutral cytosol but pump protons across the thylakoid membrane into chloroplasts (in plants) or mitochondria (in animals).[8]

1.2 Characteristics of ion channels and pores

It is now known that ion channels and water pores are proteins or protein assemblies which form an aqueous pore through the membrane.[12] These passive transport systems share three key properties.[12,17]

Conductivity They *conduct* the permeant species at very high rates ($> 10^6$ s^{-1} for ions and about 10^9 s^{-1} for water), typically comparable to the diffusion rate in free solution. The rate is not stoicheiometrically coupled to the consumption of energy in the form of e.g. ATP; the permeators follow the electro-chemical or osmotic gradient. Rapid transport of ions is required, for instance, in the conduction of nerve impulses, both in nerve and muscle cells.

Selectivity Channels typically do not conduct all solutes at the same rate, they are *selective* for one particular one or a certain class. This is necessary so that different metabolic or signalling pathways can be regulated independently. For example, potassium channels discriminate between K^+ and Na^+ ions with an accuracy of 1 000 : 1 or better[12] and most channels are permeable for either cations or anions. Some water pores (aquaporins) are only permeable for water molecules but not for protons or other ions.[11,13]

Gating Lastly, all known channels can be regulated. An external signal determines if the channel is open so that e.g. ions can pass through, or if it is closed, which prevents the flow of ions. This property is termed *gating*. The signal can be the binding of a ligand to the channel* as, for instance, in the family of the ligand gated ion channels,[12,18] a change in the transmembrane potential (the voltage gated ion channels[12,19]), a change in pressure (the mechanosensitive channels[20]), or temperature (members of the transient receptor potential (TRP) channel family[21]). The *gate* is the region of the channel which prevents the permeant species to diffuse through the pore in the closed state. Gate and sensor (the region of the channel where the signal is detected, e.g. the pocket that binds a ligand) do not have to be close to each other. It is rather typical that they lie in completely different domains, often many nanometres apart.† Gating is often meant to entail the whole sequence of actions, from sensing the signal

*Ligands can come in various shapes, e.g. neuro transmitters such as acetylcholine (ACh) or γ-aminobutyric acid (GABA) or general cell signalling molecules (cyclic adenosine monophosphate (cAMP), or calcium ions); in particular if the ligands are protons (H^+) then the channel is said to be proton- or pH-gated.

†A general organising principle of proteins, especially of eukaryotic ones, is a division in separate domains. Domains can evolve independently, being able to carry out very different tasks. When these domains (or rather, their sequences) are combined, a new protein can emerge which combines the functions of the separate building blocks.

(e.g. binding of a ligand), through the transduction of the signal across the protein, to the structural change which opens the gate. A typical gating process takes at least about 1 ms (in fast synaptic transmission).

We would like to understand what happens during gating, and how it depends on the structure of the channel. Gated ion channels are responsible for fast synaptic transmission in the central nervous system and the neuro-muscular junction, and sensory perception of taste, temperature, and pressure,[12] so knowledge about these "switches" will contribute to our understanding of phenomena as diverse as muscle control, consciousness,[22,23] hearing, tasting and feeling.[21,24] Furthermore, some hereditary diseases can be traced back to mutations that affect channel gating, for instance, slow or fast channel syndrome, a form of congenital myasthenia (hereditary muscle weakness), or nocturnal frontal lobe epilepsy.[25]

1.3 Gating mechanisms

The gate creates a barrier to permeation. Gating—the process of opening or closing the transmembrane pore—has at its the core the temporary removal or establishing of this permeation barrier. If we want to understand the very complicated process of gating then we first need to understand *where* the gate is located and *how* the permeators are prevented from passing through the gate. These two questions define what we mean by the term *gating mechanism*.

It would seem that the location of a gate is easily pinpointed, once the structure of a channel is solved at atomic resolution by X-ray protein crystallography (PX) or cryo electron microscopy (cryo-EM). Ideally, one would require two high resolution structures, one in the functionally open state, the other in the closed state. Then comparison might identify regions along the pore which changed, and by definition, the gate is a region that changes so that the channel conducts. But even this ideal case (which has not been realised although Jiang et al.[26] argue that the structures of KcsA[27] and MthK[28] are essentially the closed and open state of potassium channels) is not free from ambiguities. If the conformational changes are large then it might be difficult to identify a particular region as "the" gate—or could a gate possibly span the whole conduction pathway? Furthermore, a crystal structure is a static snapshot of a dynamical protein. It could be captured in a non-physiological state, or dynamical side-chain motions might be required during conduction, which cannot be seen in the structure. In the absence of two structures a closed state structure might still indicate a region of, for instance, most narrow constriction, which one might want to identify with the gate.

The last paragraph should have made clear that an idea of what constitutes a gate is required, even if all the structural evidence is available. For example,

Hille[12] considers twelve theoretical models for gates but only two seem to be realised in the channel structures available so far: steric occlusion through larger scale protein motions and blockage of a constriction site by a single sidechain. However, the case of steric occlusion is already ambiguous, as the following discussion of gates in ion channels will show.

1.3.1 Gates in ion channels

Since 1998 (Ref. 16) only a handful of channel structures have been published at atomic resolution, i.e. a resolution which allows identification of the sidechains (typically, better than 4 Å). All of the channels solved by X-ray crystallography are of prokaryotic or archaeal origin because their eukaryotic homologues are inherently difficult to crystallise. The structure of the electric ray *Torpedo marmorata* nicotinic acetylcholine receptor, solved by cryo-EM,[29] is the only eukaryotic membrane channel structure available. Generally, it is tacitly assumed that bacterial channel structures also illuminate the function of the homologous eukaryotic channels. The case of the ClC chloride channels, however, suggests some caution in this respect. The ClC structure from *Salmonella enterica serovar typhimurium* (StClC, 3.0 Å resolution[30]) and *Escherichia coli* (EcClC, 2.5 Å, Ref. 31) was later shown to belong to a Cl^-/H^+ exchange transporter[32] whereas the mammalian homologue is a true channel.

Local sidechain block ClC channels can be gated by pH and by voltage.[33] The EcClC structure suggests that a conserved glutamate residue competes with Cl^- ions at anion-binding site in the pore.[31,34] At low pH the glutamate carboxyl group is protonated so that the electrostatic interaction at the binding site favours Cl^- and hence the gate opens. Another example of local sidechain blocking is conjectured in the outer membrane[35–37] protein A (OmpA, PX at 2.5 Å, Ref. 38). An arginine residue can alternatively form a salt bridge across the pore in the closed state, or to a neighbouring glutamate, thus opening the pore.[39] A similar "electrostatic switch" mechanism might operate in the annexins,[40] a family of calcium-dependent phospholipid-binding proteins with ion channel activity.

Large scale steric occlusion For the channels to be discussed in the following paragraphs there is generally evidence that opening the putative gate requires some larger scale motion such as bending or twisting of pore lining α-helices[45] or large, iris-like reconfiguration of the whole channel.[46] Hence the discussion will focus on the pore radius in the closed state. The pore radius R is important because for a first approximation one can assume that an ion will not penetrate a pore whose radius is smaller than the ionic radius. Bare ionic

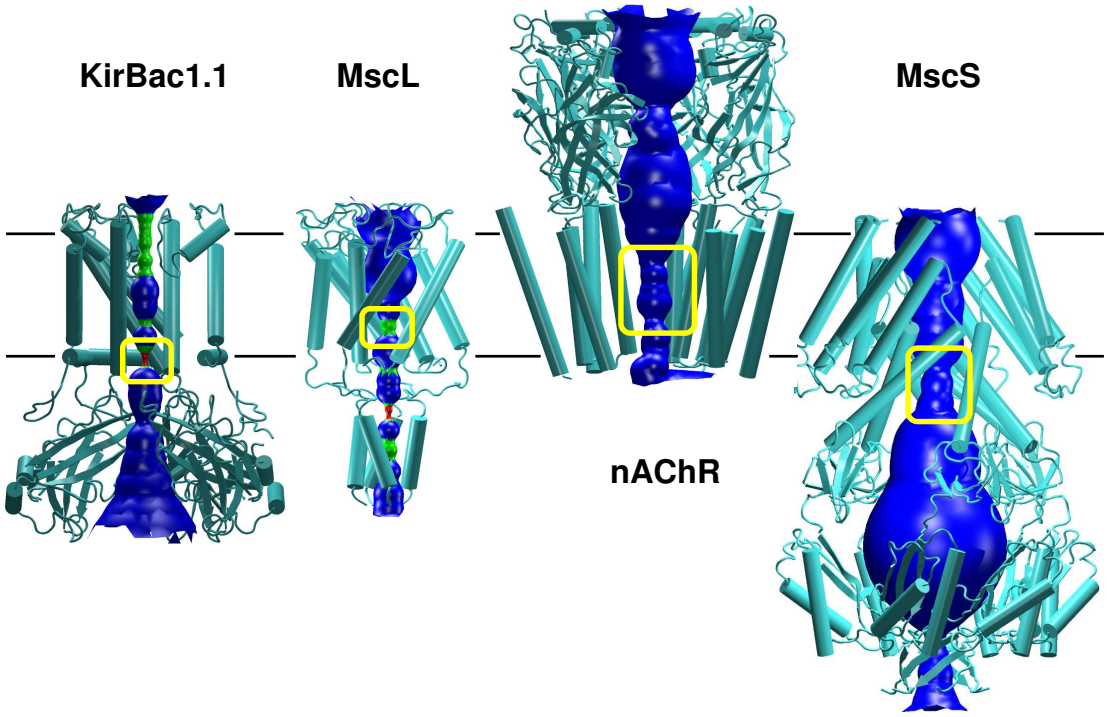


Figure 1.1: Putative occlusion gates in selected ion channels. The approximate position of the membrane is indicated by the black line. The cytosol is at the bottom of the figure. The surface of the pore is rendered within the secondary structure cartoon representation of each channel. One subunit was omitted to allow a view into the constriction site (boxed in yellow), which is formed by hydrophobic residues. Pore radius in the constriction: KirBac1.1 0.05 nm, MscL 0.16 nm, nAChR 0.31 nm (model of the $\alpha 7$ receptor from Amiri et al.[41]), MscS 0.32 nm. (Pore surface calculated with hole,[42] images created with vmd[43] and rendered with Raster3D[44])

Pauling radii range from about 0.1 nm for small cations (0.095 nm for Na^+ and 0.133 nm for K^+) to almost 0.2 nm for larger anions (0.181 nm for Cl^-)[12] with the typical radius of a water molecule, $r_w = 0.14$ nm, occupying an intermediate position.[47]

The bacterial potassium channel KirBac1.1 (3.65 Å, Ref. 48) was crystallised in the closed state and the gate was identified at a constriction (formed by phenylalanines) where the pore radius drops to 0.05 nm (see Fig. 1.1). The KcsA channel (2.0 Å, Ref. 27), which is generally believed to be a closed state structure (especially when compared to the clearly open structure of the related MthK channel[28]), displays a pore of minimum radius 0.12 nm.* The constriction site

*The selectivity filter region has a minimum radius of 0.05 nm but the filter is known not to participate in gating. Furthermore, the filter interacts with K^+ ions in a very specific manner,[27] which is the basis of its selectivity.[49] See also the discussion of the K-channel filter vs. the

is formed by one valine sidechain from each of the four monomers. The potassium channels are believed to open through a bending motion of the pore lining α-helices at a conserved glycine "hinge".[26,50]

A second class of gated channels comprises the mechanosensitive channels of large (MscL) and small (MscS) conductance. MscL (3.5 Å, Ref. 51) appears to have two gates: The narrow $R < 0.05$ nm S1 gate at the intracellular side and a second gate ($R = 0.16$ nm) at the centre of the membrane, formed by a stretch of hydrophobic residues such as valine and leucine.[46,52–54] It turns out that the S1 gate is not a proper gate as it leaks ions[46] so the real gate is formed by the hydrophobic constriction. The situation is less clear for MscS (3.9 Å, Ref. 55). Originally, it was suggested that the structure is captured in the open state[55] because the prominent constriction site, formed by two rings of leucine residues, opened to a radius of about 0.32 nm. The classification as open, however, was only based on a rough conductance estimate and not confirmed by any other method.* The mechanosensitive channels are believed to open by a large scale motion of the α-helical segments in an iris-like expansion.[20,46,54,57]

The structure of the transmembrane domain of the closed nicotinic acetylcholine receptor (nAChR) was solved by cryo-EM to about 4 Å resolution.[29] The open state structure is not known at atomic resolution but the 9 Å structure[58] clearly shows a wide ($R \approx 0.65$ nm) aqueous pore. The minimum constriction in the closed state is still rather wide with $R = 0.31$ nm. Based on the radius, the physiologically closed pore appears to belong to an open channel. Comparing the two structures reveals that the pores not only differ in radius but also in the residues which line the constriction. The constriction of the closed state is formed by a hydrophobic girdle of conserved valines or leucines. On opening, the α-helices, which line the pore, probably twist slightly so to move the valines out of the pore and expose the polar peptide backbone to the pore lumen.[59–61]

1.3.2 The hydrophobic gating hypothesis

A common theme in the putative gates discussed so far is the presence of *hydrophobic residues* in the constriction site. Furthermore, the nAChR gate and the MscS gate (if the structure is indeed captured in a closed state) can be hardly described as occluded. With $R = 0.3$ nm the pore is wide enough to allow the concurrent passage of three water molecules ($r = 0.14$ nm) or of a cation with half of its solvation shell intact. It seems likely that the *chemical character* of the pore wall contributes significantly to creating a barrier for ion permeation.

nicotinic receptor hydrophobic girdle in Section 6.1.2.

*Subsequently, Anishkin and Sukharev[56] argued for a closed state for reasons to be discussed in the following pages (see in particular Section 6.3.3 on page 136 and the footnote on page 137).

Motivated by the structural evidence we posed the hydrophobic gating hypothesis:[62]

> A gate in an ion channel can be formed by a constriction site, formed by hydrophobic sidechains such as valine, leucine, isoleucine or phenylalanine. The constriction need not be narrower than about 0.3 nm in radius because this is already sufficient to require *partial dehydration* of the permeant ion. The energetic cost for stripping even a few water molecules from the hydration shell of the ion will prevent ions from permeating the *hydrophobic gate*.

The effect of a hydrophobic transmembrane pore is twofold. Not only does it cost free energy to remove water molecules from the hydration shell but the hydrophobic sidechains also do not stabilize the ion at the centre of the membrane where it experiences a large electrostatic dielectric barrier (the dielectric or Born barrier).[63,64] In the case of nAChR[65] and MscL[51,66] it had already been suggested that the hydrophobic residues form a hydrophobic gate. Intuitively, the general mechanism appears to be simple and effective (and has been voiced in one form or another since the first half of the last century[12]) but without further evidence it largely remains a "just so story".[67] Large parts of this work investigate the hydrophobic gating hypothesis in atomic detail with the aim of providing quantitative evidence for or against it.

1.3.3 Water and gas channels

Ion channels are not the only passive transporters which seem to make use of hydrophobic pore lining. The aquaporins[11] are highly selective for water whilst the closely related glycerol facilitators[68] additionally conduct small neutral molecules with OH groups; together they form the large and ancient family of aquaglyceroporins. They are highly selective for their substrate and can exquisitely discriminate against protons (or hydronium ions H_3O^+) even though H^+ can travel through water chains by fast flipping of O-H bonds (the Grotthuss mechanism). Atomic resolution structures of Aqp1 (2.2 Å, Ref. 69, from *Bos taurus*), AqpZ (2.5 Å, Ref. 70) and GlpF (2.2 Å, Ref. 71, both from *E. coli*) show a 2 nm long and narrow (0.1 nm $\leq R \leq$ 0.2 nm) predominantly hydrophobic pore, which is punctuated by hydrogen bond donors or acceptors.[69,72] Thus, water is selectively stabilized through very specific interactions in an otherwise energetically unfavourable environment. There is also some indication that some aquaporins can be gated by pH.[73] A 3 Å resolution structure of a junction between two Aqp0 (solved by electron diffraction[74]) suggests that forming of the junction closes the pore (the average pore radius is about 0.1 nm compared to 0.2 nm in Aqp1) and converts Aqp0 into a structural protein. Some

histidine and tyrosine residues are also tentatively identified that could play a role in gating[74] through local sidechain block.

Only recently, the structure of an ammonia channel (AmtB from *E. coli*) has been published at an resolution of 1.35 Å,[75] unprecedented for transmembrane channels, both with and without bound Am (Am refers to either neutral ammonia NH_3 or the ammonium cation NH_4^+). It reveals a 2 nm long and narrow hydrophobic pore, dotted with two histidines. The proposed conduction mechanism progresses through the deprotonation of NH_4^+ at the extracellular vestibule, the diffusion of neutral NH_3 through the pore in an *unsolvated* state (stabilized by the hydrogen bond donating histidines), and re-protonation at the intracellular vestibule. AmtB (which has wrongly be named an active transporter) does not conduct water molecules or K^+ ions[75] even though the latter are similar in size to NH_4^+ ions ($r_{K^+} = 0.133$ nm vs. $r_{NH_4^+} = 0.148$ nm). The permeant species passes through the pore in the gas phase. By changing its chemical character transiently, the ammonium ion does not require stabilization by its hydration shell any more and as a neutral molecule it is not affected by the dielectric barrier posed by the membrane—unlike metal ions, which can not become neutral, or water molecules, which require hydrogen bonds for stabilization.

1.3.4 Enzyme tunnels

So far the impression was given that only membrane proteins contain pores. But other proteins also need to transport substrates between spatially separated regions. Some globular enzymes have evolved long (> 3 nm) tunnels between active sites.* In these cases, active sites in different domains catalyse different steps of an overall reaction. The tunnel facilitates diffusion of intermediate reaction products.[76] This has the advantage that the intermediate is not lost by diffusion into the cytosol, toxic intermediates are prevented from entering the cytosol, chemically labile species are protected from reactions with the solvent, transit time is decreased, and reaction rates are increased manyfold because a three-dimensional diffusion/binding process is converted to a one dimensional one.[77]

There are many examples for hydrophobic tunnels in enzymes, particularly in the family of glutamine amido transferases.[76] Tryptophan synthase[77] features a 3 nm hydrophobic tunnel for its non-polar intermediate, indole. The tunnel in carbamoyl phosphate synthase connects three active sites.[76] Ammonia can diffuse from site 1 to 2 over 4.5 nm, where it is processed into carbamate, which subsequently travels 3.5 nm to site 3. The ammonia tunnel is lined by conserved hydrophobic residues and some backbone atoms whereas the car-

*Globular proteins are normally tightly packed and the presence of continuous voids in their core, i.e. tunnels, signifies an important functional role for these.

bamate tunnel is somewhat more polar and less conserved. Overall, the average radius is about 0.34 nm. Not all ammonia tunnels are hydrophobic, though. In the case of glutamate synthase[78] the tunnel is lined by tyrosines, serines and glutamates. The longest hydrophobic tunnel was found in carbon monoxide dedydrogenase/acetyl-CoA synthase,[79,80] which forms a 310 kDa heterotetramer. The active sites are connected by a 13.8 nm long hydrophobic channel. Carbon monoxide, which is generated at the first site, can diffuse to the second site, where it is incorporated into acetyl-CoA.

The common theme in these enzymes is the transport of a gaseous intermediates. In many cases the solution is the same as the one adopted by the ammonia channel AmtB[75], discussed in the preceding section: A narrow and hydrophobic pore, rather unsuitable for ions or polar water but well adapted to harbour neutral molecules. Montet et al.[81] view hydrophobic tunnels as gas reservoirs and suggest that all gas-metabolizing enzymes will have hydrophobic channels or cavities for gas storage and active site access. There is some indication that these tunnels are also gated but the processes are probably tightly integrated with control of the enzymatic activity and not easily dissected.[76] In the current context of ion channel gating, enzyme tunnels serve as an example of the interplay between gaseous substrates and hydrophobic confinement—a theme that will be thoroughly pursued in Chapter 3.*

1.4 The hydrophobic effect

In the previous sections hydrophobicity and hydrophobic surfaces emerged as concepts central to ion channel gating. To put the work in subsequent chapters on firmer ground, we will briefly review the hydrophobic effect.[82–86]

Non-polar substances (denoted by S in the following) do not readily dissolve in water W but in oil O. That means, the change in free energy for the transfer of the solute from the oil phase (SO) to the water phase (SW) is unfavourable, i.e. positive. Formally, it is expressed as the change in the molar chemical potential at standard pressure[82]

$$\mu^\circ_{SW} - \mu^\circ_{SO} = \Delta\mu^\circ(T) = \Delta h^\circ(T) - T\Delta s^\circ(T) > 0,$$

where Δh° is the change in molar enthalpy, and Δs° the change in molar entropy. For simple (non-water) solvents the main contribution to $\Delta\mu^\circ$ is the Δh° term at all temperatures, i.e. the interaction between S and solvent W is much weaker than the S–O interaction—the low solubility is purely enthalpic. For water as a

*More correctly speaking, we will look at water *vapour* in hydrophobic channels; a vapour is a fluid that fills a volume like gas but which can be liquefied by pressure alone because it is below its critical temperature.

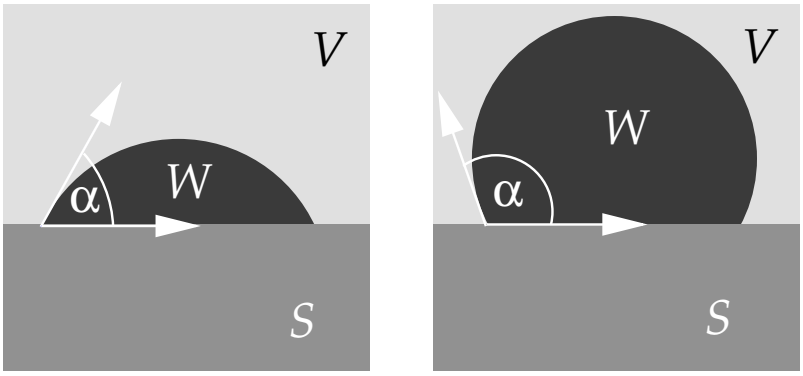
solvent, however, the enthalpic change is close to zero and the entropic term Δs° is very negative. A non-polar solute molecule decreases the entropy of water in its vicinity by imposing additional order onto the water molecules in its first hydration shell. A distinguishing feature of water is its capability to form extended hydrogen bond networks. Close to a solute to which no hydrogen bonds are possible the water molecules rearrange to maximise the number of hydrogen bonds despite the geometric frustration of the solute and thus decreases the number of favourable configurations, thus decreasing the entropy.[87] These hydrogen bonds close to a solute can be stronger than bulk water bonds, which explains the enthalpic signature. However, entropy only dominates for temperatures below ca. 350 K; for larger temperatures hydrogen bonds are less favourable and hence ordering effects decrease whilst the enthalpic costs increase so that the overall effect is dominated by the enthalpy. This is the signature of the hydrophobic effect (together with a strong increase in the isobaric heat capacity, which is again interpreted as a strengthening of hydrogen bonds which require more heat to be broken).

Hence, a hydrophobic substance is a non-polar substance that does not readily dissolve in water and whose solvation energy exhibits the particular temperature dependence described above. It is clear that *hydrophobicity* is a property that is only defined when the substance is brought in contact with *water*.

So far, small apolar molecules were discussed. For small solutes the hydrophobic effect is mainly entropic. It is equivalent to the process of creating a cavity in the solvent[86,88] (plus enthalpic interactions between water and solute). On the other hand, large solutes, which create extended planar interfaces, can not be straddled by water molecules to achieve maximum hydrogen bonding.[89,90] They always force a water molecule to forfeit one bond and disrupt the water structure. Consequently, water density is depleted near the substrate; this collective effect resembles a drying transition induced by the hydrophobic surface.[91] The solvation free energy $\Delta\mu^\circ > 0$ is always of enthalpic origin due to the reduction in water hydrogen bonds, not the substrate-water interaction. It can be related to the macroscopic surface tension. The crossover between these two different regimes of the hydrophobic effect is found at the nanometre length scale.[90]

A macroscopic definition of hydrophobicity comes from the observation that a droplet of water W may either completely wet a substrate surface S, form a droplet with a contact angle $0^\circ < \alpha \leq 90^\circ$, or a droplet with $90^\circ < \alpha \leq 180^\circ$ (Fig. 1.2 on the following page).[92] The surface is called hydrophilic in the former two cases and hydrophobic in the latter. The contact angle is related to the surface tension γ_{SW} (or free energy to create an interfacial area) by Young's equation

$$\gamma_{SV} - \gamma_{SW} - \gamma_{WV} \cos\alpha = 0$$



(a) Hydrophilic surface (b) Hydrophobic surface

Figure 1.2: Macroscopic definition of hydrophobicity. (a) If a water droplet (dark grey, W) on a surface (grey, S) forms a contact angle $\alpha \leq 90^\circ$ then the surface material is called *hydrophilic* ("water-liking"). (b) If $\alpha > 90^\circ$ then the material is termed *hydrophobic* or "water-fearing". Water is in equilibrium with its vapour V.

where γ_{SV} denotes the surface tension between water vapour and the surface and γ_{WV} the surface tension of the liquid water-water vapour interface.

Scaled particle theory[93] describes hydrophobic solvation of small solutes by ascribing the macroscopic surface tension to the interfacial area of the solute cavity and motivates the application of the macroscopic concepts to the microscopic domain, as will be done in subsequent chapters.

Although not subsequently pursued any further, it should be noted that two hydrophobic particles in water interact, mediated by the solvent. Depending on their distance they are attracted or repulsed.[86,88,90,94] Contact is always thermodynamically more favourable than separation because it decreases the solvent-exposed surface area. This effect is called hydrophobic interaction or hydrophobic attraction. It is an effective interaction, a collective effect of the solvent in the vicinity of the solutes. The effect is attractive because water molecules attract each other much stronger than either the solute attracts water molecules, or solute particles each other. Between extended hydrophobic surfaces (up to a separation of about 2 nm) the water film is unstable and evaporates[90,94] which leads to attraction between the surfaces.

1.5 Complementing structural data

The preceding discussion of the location of gates in protein channel structures showed that although the structure is crucial to understanding gating behaviour it does not contain all information necessary. Proteins are dynamic mol-

ecules, and a static structure can not quantitate protein movements. Furthermore, solvent and solute effects are only addressed insofar as sometimes water or ions are co-crystallised, thus indicating putative binding sites, or by using molecular graphics and displaying the protein surface characteristics, indicating polar and hydrophobic areas. Typically, other methods are also used to probe different functional aspects and to complement structural data. There is available a vast array of biochemical (mutagenesis, fluorescence probe labelling, spin labelling, cysteine scanning mutagenesis, to name a few) and biophysical (patch clamp single channel recordings, flux measurements, ...) methods.[12] In addition, theoretical and especially computational methods are increasingly used to model the dynamics of membrane-embedded proteins.[95]

In this work, computer "simulations" are carried out with the aim to understand the behaviour of pure water and of solvated ions in confined geometries, as presented by putative hydrophobic gates in ion channels, i.e. to investigate the hydrophobic gating hypothesis, and to explore the possibility of such a gate in nAChR. These *molecular dynamics* calculations treat water, ions, and protein with atomistic detail (albeit as a classical not a quantum mechanical system). They allow a dynamic and microscopic view on processes that are difficult or impossible to observe in experiment. Computer simulations pose a twofold challenge: to connect the computational results to experiments, and to understand the results in terms of physical models.

The following overview lays out how these aims were approached and met.

1.6 From toy models to the nicotinic receptor

A recurring theme is the question how the behaviour of water and ions depends on the environment that is provided by a (protein) pore. In **Chapter 2** the ground work is laid by recapitulating the thermodynamics and some of the statistical mechanics of confined fluids. In later chapters this will be used to explain the results of the computer simulations. The bulk of these calculations are fully atomistic molecular dynamics simulations (to be introduced here), both in equilibrium and biased towards unfavourable regions. The latter is performed by umbrella sampling, which is introduced together with the WHAM procedure (weighted histogram analysis method) to unbias the biased simulations. A central quantity in the discussion of permeation is the potential of mean force (PMF), the free energy required to position a test particle at a given position in space. The PMF can be obtained from equilibrium and biased simulations. A more coarse grained approach to the PMF is provided by continuum electrostatic Poisson-Boltzmann calculations which allow to calculate the electrostatic (or Born) contribution to the PMF, albeit in a fraction of the computing time required for the fully atomistic calculations.

Chapter 3 introduces a "toy model," a radically simplified hydrophobic pore with only a limited number of parameters. This provides greater control over the environment experienced by the permeators and helps to disentangle different effects. Although ultimately interested in ion permeation, first the behaviour of the solvent, i.e. *water*, is investigated. Despite being little more than a hole through a membrane the toy model system exhibits a wealth of sometimes surprising behaviour. In pores of radius of about 0.5 nm water is seen to oscillate between a liquid and a vapour phase on a time scale of nano seconds. Narrow pores of the dimension of the nAChR gate $R \leq 0.35$ nm are seen to be void of water. The absence of liquid water in narrow hydrophobic pores suggests that an ion will not be able to permeate the pore because its solvation shell will not be stable. This chapter is based on O. Beckstein and M. S. P. Sansom, *Proc. Natl. Acad. Sci. U.S.A.* **100** (2003), 7063–7068 (Ref. 96), but includes extensions not presented in the paper and new insights, based on subsequent calculations, which supersede some findings in the original paper.*

The hydrophobic gating hypothesis is tested in **Chapter 4** by extending the previous calculations to sodium and chloride ions. Equilibrium calculations demonstrate that hydrophobic pores of the dimensions of the nAChR gate are impermeable to ions on the 100 ns time scale. If a narrow hydrophobic pore is made more hydrophilic it can retain liquid water. Thus radius (geometry) and chemical character of the pore wall are identified as determinants of permeability. A thermodynamic model is developed which explains the liquid-vapour transitions in the pores in terms of surface tensions. So far the chapter closely follows O. Beckstein and M. S. P. Sansom, *Physical Biology* **1** (2004), 42–52 (Ref. 97). In addition, the effect of pore wall flexibility is investigated and the seemingly counter-intuitive result is found that a flexible pore is *less* permeable to water than a rigid one. A qualitative argument, which also applies to similar behaviour at higher temperatures, suggests why this is so.

A quantitative assessment of the impermeability of narrow hydrophobic pores is conducted in **Chapter 5** with the calculation of PMFs of ions and water molecules. The barriers to permeation are found to be substantial ($> 10\ kT$) in pores with $R < 0.4$ nm. This is the strongest evidence in favour of the hydrophobic gating hypothesis. These atomistic results are compared to continuum-electrostatics calculations which fail to reproduce the molecular dynamics PMF because they cannot properly account for the solvent behaviour in the pores. The comparison reinforces the notion that the water in the pore is key to understanding the behaviour of the ions [O. Beckstein, K. Tai, and M. S. P. Sansom, *J. Am. Chem. Soc.* **126** (2004), 14694–14695 (Ref. 98)].

After having gained some understanding of the principles of hydrophobic

*The interested reader will find explained in footnotes where later results improve upon the ones published in Ref. 96.

gating from the toy model pores the same reasoning is applied to the pore of the muscle-type nicotinic acetylcholine receptor in **Chapter 6**. The nAChR structure only became available in March 2003.[29] Historically, the model pores had been designed to mimic the dimensions of the nAChR pore, based on data available at the time[59] (autumn 2000). In 2003 the opportunity arose to compare the models to the real structure. Fully atomistic free energy calculations yield the PMF for Na^+, Cl^-, and water translocation through the nAChR transmembrane pore. The permeation barrier is located at the hydrophobic girdle, as predicted by hypothesis. The height and width of the barrier (about 8 kT over 2 nm) strongly supports the idea that the nAChR structure is indeed in the closed state and that nAChR contains a *hydrophobic gate*. The results are robust with respect to different force field parameters and simplifications to the protein setup, namely the reduction of the pore to only the five pore lining M2 helices.

The main part closes with a discussion which highlights how the idea of hydrophobic confinement has already been applied in various systems. Appendices list results that could not be incorporated in the main chapters, either because of length or of their more speculative nature. PMFs were calculated for a model of the neuronal $\alpha 7$ nicotinic receptor[41], which show some differences to the muscle-type nAChR. Other appendices deal with the problem of calculating volumes at the molecular scale, which could be used for thermodynamic models such as the one used in Chapter 4, a list of publications which arose from this work, and a short overview of software developed for analysis. A bibliography and an index complete the work.

2 Theory and Methods

Before reporting the results on water and electrolytes in a variety of model pores and in the "real" pore of the nicotinic receptor in the next chapters we will set the scene by recapitulating some of the physics of fluids in inhomogeneous systems, recognising that the theory is the same regardless of the complexity of the actual pore. We will then briefly describe the methods used for our machine calculations.* Primarily this means atomistic molecular dynamics, both in equilibrium and as biased simulations, but continuum electrostatics Poisson-Boltzmann theory is also discussed, which is used for Born energy and pK_a calculations.

2.1 Thermodynamics and statistical mechanics of interfaces

In this work we will be concerned with the behaviour of water or an electrolyte in a given environment. In order to interpret the computer simulations some basic concepts of thermodynamics are required, which are recapitulated in this chapter, following largely Chaikin and Lubensky[99], Rowlinson and Widom[100], and Hansen and McDonald[101].

2.1.1 The surface tension

The environment will take the form of a planar surface (an approximation to a biological membrane) or a pore through such a membrane (a highly simplified model for ion channels). The difference between a homogeneous bulk system ("water in a box") and the systems with a membrane-mimetic slab (with or without a hole) is the presence of the interface between the slab (the "wall") and the fluid phase.† The surface introduces a new field, the surface tension or

*In the early years of computer simulations (1954 to about 1970) theoreticians referred to what is nowadays known as *simulations* as *machine calculations*, as opposed to solving equations with pencil and paper. Perhaps in the future this particular field will rather be simply known as *computer-assisted biophysics*.

†The term "fluid" encompasses both the liquid and the gaseous state, which are distinguished from the solid state by their higher degree of symmetry. The solid and the liquid ("condensed matter") differ from the gaseous state in their higher density and degree of spatial correlations.[99]

surface free energy γ, into the total energy $U(S, V, N, A)$,[99,100]

$$dU = T\,dS - p\,dV + \mu\,dN + \gamma\,dA. \tag{2.1}$$

Here T is the temperature, S the entropy, p the pressure, V the volume, μ the chemical potential of the fluid, N the number of particles in the fluid, and A the area of the wall-fluid interface. (For an M-component system we would write $\sum_{i=1}^{M} \mu_i\,dN_i =: \boldsymbol{\mu} \cdot d\mathbf{N}$ with the chemical potentials of the M species μ_i and their particle numbers N_i.) More useful thermodynamic potentials are the Helmholtz free energy $F(T, V, N, A)$

$$dF = -S\,dT - p\,dV + \mu\,dN + \gamma\,dA, \tag{2.2}$$

which is minimal in equilibrium when a closed system* is in contact with a heat bath so that the average total energy is fixed, the Gibbs function $G(T, p, N, A)$ (also Gibbs free energy),

$$dG = -S\,dT + V\,dp + \mu\,dN + \gamma\,dA \tag{2.3}$$

(especially useful for biological processes, which typically occur at constant pressure and temperature) and the grand potential

$$d\Omega = -S\,dT - p\,dV - N\,d\mu + \gamma\,dA. \tag{2.4}$$

For an open system at equilibrium (which can exchange heat and particles) $\Omega(T, V, \mu, A)$ takes on a minimal value. The grand potential is best suited for a system like a pore connected to reservoirs where the chemical potential μ and the temperature T is fixed by the external reservoir, and the pore volume V is constant.

Integrating Eq. (2.4) yields

$$\Omega(T, V, \mu, A) = -pV + \gamma A, \tag{2.5}$$

and when this is compared to the grand potential of the homogeneous system $\Omega(T, V, \mu) = -pV$ at the *same* thermodynamic state (T, V, μ) then the excess free energy

$$\Omega^s(T, V, \mu, A) := \Omega(T, V, \mu, A) - \Omega(T, V, \mu) = \gamma A, \tag{2.6}$$

is obtained[100], i.e. the additional free energy required to create the interface. Hence, the surface tension

$$\gamma = \frac{\Omega^s}{A} = \frac{\Omega + pV}{A} = \left(\frac{\partial \Omega}{\partial A}\right)_{T,V,\mu} \tag{2.7}$$

*In a closed system the number of particles is fixed but heat and mechanical energy can be exchanged with the surroundings.

is the free energy required to create the interface per unit of area of the interface between the fluid phase and the wall.*

2.1.2 Microscopic description

A connection between the microscopic, atomistic parameters (e.g. the interaction between fluid molecules and the wall) and the macroscopic thermodynamic observables is provided by statistical mechanics. Although we will really be concerned with water in the following chapter, by no means a "simple liquid", we will restrict the rest of this section to simple liquids made from spherical "molecules" because the following equations will not be used in a quantitative manner but shall only be used to deepen our understanding of the behaviour of confined liquids in a very general manner.

The fluid consists of N particles with mass m and momentum $\mathbf{p}_i$, which interact through a potential $V(\{\mathbf{x}_j\})$ of the particle coordinates $\mathbf{x}_j$ alone.† Without the wall, the Hamiltonian $\mathcal{H}_0$ is simply

$$\mathcal{H}_0(\{\mathbf{p}_j, \mathbf{x}_j\}) = \mathcal{H}_{\text{kin}} + \mathcal{U} = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + V(\{\mathbf{x}_j\}). \tag{2.8}$$

$\mathcal{H}_{\text{kin}}$ is the kinetic part of the Hamiltonian (here it only contains a translational component but it can be generalised to non-spherical molecules[100]). The potential energy $\mathcal{U}$ is taken to be an interaction potential $V(\{\mathbf{x}_j\})$ that only depends on the positions of the N particles.

Canonical ensemble For a given Hamiltonian $\mathcal{H}$ the partition function of the canonical ensemble for N particles is given by

$$Z_N(T, V) = \operatorname{Tr} e^{-\beta\mathcal{H}}. \tag{2.9}$$

$\beta = (kT)^{-1}$ is the inverse temperature (with Boltzmann's constant $k = 1.3806503(24) \times 10^{-23}\ \text{J} \cdot \text{K}^{-1}$) and Tr denotes the trace over all states of the quantum mechanical probability density operator or in classical mechanics the phase-space integral

$$\operatorname{Tr} \equiv \frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^3p_i\, d^3x_i}{h^3}.$$

*For F a similar relationship, $\gamma = F^s/A$, holds, provided the Gibbs dividing surface is chosen as the surface of vanishing adsorption ($\boldsymbol{\mu} \cdot \mathbf{N}^s = 0$).[100]

†The notation $f(\{\mathbf{x}_j\}) \equiv f(\{\mathbf{x}_j\}_{1 \le j \le N})$ is a shorthand for $f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ and indicates that the function (or observable) f depends on *all* coordinates $\mathbf{x}_j$. Alternatively this could be written $f(\mathbf{x}^N)$ or $f(x^{3N})$.

The Helmholtz free energy is then

$$\beta F(T,V,N) = -\ln Z_N(T,V). \tag{2.10}$$

A thermodynamic average of an observable $\mathcal{A}(\{\mathbf{p}_j,\mathbf{x}_j\})$ is given by

$$\langle \mathcal{A} \rangle = Z_N^{-1}\ \mathrm{Tr}\, \mathcal{A}(\{\mathbf{p}_j,\mathbf{x}_j\})\, e^{-\beta\mathcal{H}} \tag{2.11}$$

in the canonical ensemble.

Grand canonical ensemble If the number of particles is allowed to vary (but the average $\langle N \rangle$ is fixed by an imposed chemical potential μ) then the partition function of the grand canonical ensemble is

$$\Xi(T,V,\mu) = \mathrm{Tr}\, e^{-\beta(\mathcal{H}-\mu N)}. \tag{2.12}$$

Now the trace is also summed over all possible numbers of particles $0 \leq N < \infty$; classically this means

$$\mathrm{Tr} \equiv \sum_{N=0}^{\infty} \frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^3p_i\, d^3x_i}{h^3};$$

hence Ξ can also be written $\Xi = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_N$. The grand potential

$$\beta\Omega(T,V,\mu) = -\ln \Xi(T,V,\mu) \tag{2.13}$$

is the thermodynamic potential appropriate for an open system. The thermodynamic average of any observable $\mathcal{A}$ is again given as the $\mathcal{A}$-weighted sum over all states, normalised by the partition function

$$\langle \mathcal{A} \rangle = \Xi^{-1}\ \mathrm{Tr}\, \mathcal{A}(\{\mathbf{p}_j,\mathbf{x}_j\})\, e^{-\beta(\mathcal{H}-\mu N)}, \tag{2.14}$$

where the grand canonical trace includes the sum over all possible N-states.

Inhomogeneous systems For simplicity, we do not consider the wall a part of the system but represent it through an external field $u(\mathbf{x})$. Then the potential energy $\mathcal{U}$ must be augmented by the contribution from the external potential

$$\mathcal{U}_{\mathrm{ext}} = \sum_{i=1}^{N} u(\mathbf{x}_i) = \int d^3x\, u(\mathbf{x}) n(\mathbf{x}) \tag{2.15}$$

with the density(operator)*

$$n(\mathbf{x}) := \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{x}_i). \tag{2.16}$$

The inhomogeneous Hamiltonian is thus

$$\mathcal{H} = \mathcal{H}_{\text{kin}} + \mathcal{U} + \mathcal{U}_{\text{ext}} = \mathcal{H}_0 + \int d^3x\, u(\mathbf{x}) n(\mathbf{x}). \tag{2.17}$$

The external potential couples linearly to the density, just as the chemical potential [$\mu N = \int d^3x\, \mu n(\mathbf{x})$]. In the grand canonical ensemble one can interpret the external potential as a shift in the chemical potential

$$\mu \to \mu(\mathbf{x}) = \mu - u(\mathbf{x}), \tag{2.18}$$

i.e. in the inhomogeneous system the chemical potential appears to depend on the position.† Hence the grand partition function is now a functional of $\mu(\mathbf{x})$,

$$\Xi[T, V, \mu(\mathbf{x})] = \operatorname{Tr} e^{-\beta(\mathcal{H} - \mu N)} = \operatorname{Tr} \exp\left[-\beta\left(\mathcal{H}_0 - \int d^3x\, \mu(\mathbf{x}) n(\mathbf{x})\right)\right]. \tag{2.19}$$

In the homogeneous system the average density $\langle n(\mathbf{x})\rangle = \langle N\rangle / V$ is constant (where $\langle N\rangle$ is fixed through the value of μ). In the system with a wall the addition of the inhomogeneity $u(\mathbf{x})$ will also induce inhomogeneity in the average density,‡ i.e. the variation in $\mu(\mathbf{x})$ (or equivalently $-u(\mathbf{x})$) results in

$$\langle n(\mathbf{x})\rangle = \frac{\delta \ln \Xi[\mu(\mathbf{x})]}{\delta \beta \mu(\mathbf{x})} = -\frac{\delta \Omega[\mu(\mathbf{x})]}{\delta \mu(\mathbf{x})}. \tag{2.20}$$

Not only is the equilibrium density $\langle n(\mathbf{x})\rangle$ of interest but also *correlations* in the density, which contain information about the local structure of the liquid.

*The equations from (2.20) onwards are only valid for classical fluids as in the derivation it is assumed that the Hamiltonian $\mathcal{H}$ commutes with the local density operator $n(\mathbf{x})$. For this work this is no limitation as only classical fluids are simulated by classical molecular dynamics simulations.

†Rowlinson and Widom[100] call $\mu(\mathbf{x})$ the *intrinsic chemical potential* $\mu_{\text{int}}(\mathbf{x})$ [their Eq. (4.132)] and we follow their definition; Chaikin and Lubensky[99] prefer to define $\mathcal{U}_{\text{ext}}$ in (2.15) with a minus sign so that $\mu(\mathbf{x}) = \mu + u(\mathbf{x})$. In either case it is still μ which is fixed in the grand canonical ensemble.

‡The addition of $u(\mathbf{x})$ breaks the symmetry of the homogeneous system $\mathcal{H}_0$ so that the inhomogeneous one has to adapt a lower symmetry; but this route of reasoning will not be further pursued in this work.

Scattering experiments effectively measure the two point density correlation function[99]

$$C_{nn}(\mathbf{x}_1, \mathbf{x}_2) := \left\langle \sum_{i,j} \delta(\mathbf{x}_1 - \mathbf{x}_i)\delta(\mathbf{x}_2 - \mathbf{x}_j) \right\rangle = \langle n(\mathbf{x}_1) n(\mathbf{x}_2) \rangle \,. \tag{2.21}$$

For large separations $|\mathbf{x}_1 - \mathbf{x}_2|$ the correlations vanish* so that $C_{nn}(\mathbf{x}_1, \mathbf{x}_2)$ tends to the product $\langle n(\mathbf{x}_1) \rangle \langle n(\mathbf{x}_2) \rangle$ (i.e. the densities at the two points are uncorrelated). Thus the correlated part of the two point density correlation function is the Ursell function[99]

$$S_{nn}(\mathbf{x}_1, \mathbf{x}_2) := C_{nn}(\mathbf{x}_1, \mathbf{x}_2) - \langle n(\mathbf{x}_1) \rangle \langle n(\mathbf{x}_2) \rangle \,. \tag{2.22}$$

The pair distribution function $g(\mathbf{x}_1, \mathbf{x}_2)$, defined by

$$\langle n(\mathbf{x}_1) \rangle \, g(\mathbf{x}_1, \mathbf{x}_2) \, \langle n(\mathbf{x}_2) \rangle := \left\langle \sum_{i \neq j} \delta(\mathbf{x}_1 - \mathbf{x}_i)\delta(\mathbf{x}_2 - \mathbf{x}_j) \right\rangle = \langle n(\mathbf{x}_1) n(\mathbf{x}_2) \rangle - \langle n(\mathbf{x}_1) \rangle \, \delta(\mathbf{x}_1 - \mathbf{x}_2), \tag{2.23}$$

is the probability to observe a particle in the volume element d^3x at $\mathbf{x}_2$, provided that there is another, different particle at $\mathbf{x}_1$. For $|\mathbf{x}_1 - \mathbf{x}_2| \to \infty$ it tends to 1; for any liquid whose molecules have a hard core, $g(\mathbf{x}_1, \mathbf{x}_2)$ approaches 0 when $\mathbf{x}_2 \to \mathbf{x}_1$. For a spatially homogeneous and translationally invariant fluid ($\langle n(\mathbf{x}) \rangle = \langle n(\mathbf{x} + \mathbf{R}) \rangle$, hence $\langle n(\mathbf{x}) \rangle = \langle n \rangle = \langle N \rangle / V$), the pair distribution function is only a function of the difference $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$,

$$g(\mathbf{x}) = \frac{1}{\langle n \rangle} \left\langle \sum_{i \neq 0} \delta(\mathbf{x} - \mathbf{x}_i + \mathbf{x}_0) \right\rangle . \tag{2.24}$$

For an arbitrary particle at $\mathbf{x}_0$, the number of other particles in the volume element at a distance $\mathbf{x}$ equals $\langle n \rangle \, g(\mathbf{x}) \, d^3x$. For isotropic fluids g only depends on the distance $r = |\mathbf{x}|$ from $\mathbf{x}_0$ and the number of particles in shells of thickness dr around any particle is given by $\langle n \rangle \, 4\pi r^2 g(r) \, dr$; $g(r)$ is called the radial distribution function. For non-isotropic fluids "a" $g(r)$ can still be computed as the spherical average

$$g(r) = (4\pi)^{-1} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, g(r, \phi, \theta). \tag{2.25}$$

Occasionally the pair correlation function

$$h(\mathbf{x}_1, \mathbf{x}_2) := g(\mathbf{x}_1, \mathbf{x}_2) - 1 \tag{2.26}$$

*... unless the system is close to a phase transition, which is heralded by a diverging correlation length.

is also used, which only contains the correlated part of $g(\mathbf{x}_1, \mathbf{x}_2)$. The correlation functions can also be obtained as variations $\delta\Omega[\mu(\mathbf{x})]$ with $\delta\mu(\mathbf{x}_1)$ and $\delta\mu(\mathbf{x}_2)$:

$$\begin{aligned}\beta S_{nn}(\mathbf{x}_1, \mathbf{x}_2) &= -\frac{\delta^2\Omega[\mu(\mathbf{x})]}{\delta\mu(\mathbf{x}_1)\,\delta\mu(\mathbf{x}_2)} = \frac{\delta\,\langle n(\mathbf{x}_1)\rangle}{\delta\mu(\mathbf{x}_2)} \\ &= \beta\big(\,\langle n(\mathbf{x}_1)\rangle\, h(\mathbf{x}_1, \mathbf{x}_2)\,\langle n(\mathbf{x}_2)\rangle + \langle n(\mathbf{x}_2)\rangle\,\delta(\mathbf{x}_1 - \mathbf{x}_2)\big) \quad (2.27)\end{aligned}$$

Another route to the pair distribution function (2.23) can be taken if the particle-particle interaction potential $V(\{\mathbf{x}_j\})$ in Eq. (2.8) only consists of pairwise interactions $v(\mathbf{x}_i, \mathbf{x}_j)$,

$$\begin{aligned}V(\{\mathbf{x}_j\}) &= \sum_{i<j} v(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{2}\sum_{i\neq j} v(\mathbf{x}_i, \mathbf{x}_j) \\ &= \frac{1}{2}\int d^3x_1\, d^3x_2\, v(\mathbf{x}_1, \mathbf{x}_2)\left[\sum_{i,j}\delta(\mathbf{x}_1 - \mathbf{x}_i)\delta(\mathbf{x}_2 - \mathbf{x}_j) - \sum_i \delta(\mathbf{x}_1 - \mathbf{x}_2)\delta(\mathbf{x}_1 - \mathbf{x}_i)\right].\end{aligned} \quad (2.28)$$

Then it follows immediately that $g(\mathbf{x}_1, \mathbf{x}_2)$ is the functional derivative of the grand potential with respect to the interaction potential,

$$\langle n(\mathbf{x}_1)\rangle\, g(\mathbf{x}_1, \mathbf{x}_2)\,\langle n(\mathbf{x}_2)\rangle = 2\frac{\delta\Omega[v(\mathbf{x}_1, \mathbf{x}_2)]}{\delta v(\mathbf{x}_1, \mathbf{x}_2)}. \quad (2.29)$$

This relationship also explicitly demonstrates that the structure of the liquid (as embodied by g) is rooted in the (pairwise) interaction between fluid particles. The influence of the wall, $u(\mathbf{x})$, is implicitly taken into account by the Boltzmann factor in the partition function $\Xi[u(\mathbf{x})] = \mathrm{Tr}\exp\big[-\beta\big(\mathcal{H}_0 - \sum_i[\mu - u(\mathbf{x}_i)]\big)\big]$. Regions "in" the wall carry a high potential energy for a particle there, which translates into a low probability by virtue of the Boltzmann factor. Hence this region of space contributes negligibly to the partition function or any thermodynamic average.

Surface tension and density as functionals of the wall potential Because the partition functions [Eq. (2.9) or Eq. (2.12)] contain the microscopic parameters in the interaction potential $V(\{\mathbf{x}_j\})$ and the wall potential $u(\mathbf{x})$ it is possible to derive expressions for the surface tension which show how it depends on these parameters. In particular we will be interested in how the surface tension γ depends on the parameters λ_j of the wall potential $u(\mathbf{x}; \{\lambda_j\})$. First we will need the dependence of the grand potential and the equilibrium density on the

wall parameters,

$$\frac{\partial\Omega}{\partial\lambda_i} = \int d^3x\ \langle n(\mathbf{x})\rangle\ \frac{\partial}{\partial\lambda_i} u(\mathbf{x};\{\lambda_j\}), \quad \text{and} \tag{2.30}$$

$$\frac{\partial\ \langle n(\mathbf{x}_1)\rangle}{\partial\lambda_i} = -\beta \int d^3x_2\, S_{nn}(\mathbf{x}_1,\mathbf{x}_2)\ \frac{\partial}{\partial\lambda_i} u(\mathbf{x}_2;\{\lambda_j\}), \tag{2.31}$$

which follows from straight differentiation of the partition function (2.12) and Eq. (2.13). The surface tension can be obtained as the change of Ω with the interfacial area A [Eq. (2.7)] and hence

$$\frac{\partial\gamma}{\partial\lambda_i} = \left(\frac{\partial^2\Omega}{\partial\lambda_j\,\partial A}\right)_{T,V,\mu} = -\frac{1}{\beta}\frac{\partial}{\partial\lambda_j}\frac{\partial}{\partial A}\ln\Xi[T,V,\mu(\mathbf{x}),A]. \tag{2.32}$$

The partial derivative in A is carried out by a dimensional argument.[100,101] For simplicity, we look at the special case of a flat interface perpendicular to the z direction. The Cartesian coordinates change by $x \to (1+\epsilon)x$, $y \to (1+\epsilon)y$, and $z \to (1+\epsilon)^{-2}z = (1-2\epsilon)z + O(\epsilon^2)$, where $|\epsilon| \ll 1$. This distortion conserves the volume (as required by the partial derivative) and leads to a change in area $\delta A = 2\epsilon A$. The only terms affected by the change in coordinates are the ones that depend on inter-particle distances, i.e. the interaction potential $v(\mathbf{x}_i,\mathbf{x}_j)$ in (2.28). The form

$$\gamma = \frac{1}{2}\int_{-\infty}^{\infty} dz_1 \int d^3x_{12}\ \frac{x_{12}^2 - 3z_{12}^2}{r_{12}^2}\frac{\partial v(r_{12})}{\partial r_{12}}\ \langle n(\mathbf{x}_1)\rangle\, g(\mathbf{x}_1,\mathbf{x}_2)\ \langle n(\mathbf{x}_2)\rangle$$
$$\text{with } r_{12} := |\mathbf{x}_{12}| = |\mathbf{x}_1 - \mathbf{x}_2| \tag{2.33}$$

obtained by this approach is called the Kirkwood-Buff formula (where the integration is over all difference vectors $\mathbf{x}_{12} := \mathbf{x}_1 - \mathbf{x}_2$). Consequently, the change in the interaction parameters

$$\frac{\partial\gamma}{\partial\lambda_i} = \frac{1}{2}\int_{-\infty}^{\infty} dz_1 \int d^3x_{12}\ \frac{x_{12}^2 - 3z_{12}^2}{r_{12}^2}\frac{\partial v(r_{12})}{\partial r_{12}}\frac{\partial}{\partial\lambda_j}\ \langle n(\mathbf{x}_1)\rangle\, g(\mathbf{x}_1,\mathbf{x}_2)\ \langle n(\mathbf{x}_2)\rangle\,, \tag{2.34}$$

affects γ only through a change in the pair distribution function

$$\frac{\partial}{\partial\lambda_i}\ \langle n(\mathbf{x}_1)\rangle\, g(\mathbf{x}_1,\mathbf{x}_2)\ \langle n(\mathbf{x}_2)\rangle =$$
$$-\beta\int d^3x_3 \Big\{\langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3)\rangle - \langle n(\mathbf{x}_1)n(\mathbf{x}_2)\rangle\ \langle n(\mathbf{x}_3)\rangle$$
$$+\big[\langle n(\mathbf{x}_1)n(\mathbf{x}_2)\rangle - \langle n(\mathbf{x}_1)\rangle\ \langle n(\mathbf{x}_2)\rangle\big]\delta(\mathbf{x}_1-\mathbf{x}_2)\Big\}\frac{\partial u(\mathbf{x}_3;\{\lambda_j\})}{\partial\lambda_i} \tag{2.35}$$

Inserting (2.35) into (2.34) shows that the change in surface tension depends on a third order correlation function (the term in curly braces in (2.35)), weighted by the change of the wall potential $\partial u(\mathbf{x}; \{\lambda_j\})/\partial\lambda_i$. Thus, if $u(\mathbf{x})$ only changes appreciably in a narrow region, say, close to the surface, then it will be molecules in the surface layer that contribute predominantly to the change in surface tension.* Eq. (2.34) or (2.33) are not especially useful in their current form because the correlation functions and equilibrium densities are not known. One could use approximations for $\langle n \rangle$ and g (see Rowlinson and Widom[100], pp91) but this is not the subject of this work. These equations are nevertheless useful as an interpretative guide, which describes where and how changes in the wall influence the equilibrium density and the surface tension, the two quantities that are studied by computer simulations in subsequent chapters. Furthermore, exact relations like Eq. (2.34) have been used to accurately calculate surface tensions from the density $\langle n(\mathbf{x}) \rangle$ obtained from computer simulations.[102–104†] A more exhaustive list of sum rules like Eq. (2.34) and (2.31) can be found in van Swol and Henderson[102].

2.1.3 Application to computer simulations

In this work, theoretical descriptions as presented in the preceding sections are predominantly sought to understand and analyse computer simulations. Molecular dynamics simulations (see Section 2.2) yield trajectories of the simulated system, i.e. the position (and if desired, also the velocities) of all particles at each recorded time step. From a statistical mechanics point of view this is the complete though not very enlightening description of the system. Statistical mechanics and thermodynamic concepts are then applied to extract "interesting properties."[105] Which observables these correspond to depends on the questions asked and it also depends on the simulation being conducted in such a way that the questions *can* be answered meaningfully.

*This is rather obvious for a square-well potential[102] where the integral in Eq. (2.35) will only have contributions from within the square well.

†Nijmeijer et al.[104] obtain a much simpler formula for the change of the surface tension with the well depth ϵ of their square well potential, namely

$$\frac{\partial\gamma}{\partial\epsilon} = L^{-2} \int_{-L/2}^{L/2} dx\, dy \int_0^{z_c} dz\, \langle n(\mathbf{x}) \rangle \frac{u(\mathbf{x}; \{\epsilon, a\})}{\epsilon},$$

z_c being the cut-off for the inter-particle interactions, a the range of the wall potential, and L the simulation box size. It is not quite clear how they arrived there and how general it is.

Averages

Many equilibrium properties of interest can be expressed as thermodynamic averages of an observable $\mathcal{A}$, Eq. (2.11) or Eq. (2.14). These averages are *ensemble averages*, i.e. the value of the observable for a given configuration $\mathcal{A}(\{\mathbf{x}_i\})$ times the probability for this configuration $\varrho(\{\mathbf{x}_i\}) = Z^{-1} \exp(-\beta\mathcal{H})$ or $\varrho(\{\mathbf{x}_i\}) = \Xi^{-1} \exp\big(-\beta(\mathcal{H} - \mu N)\big)$:

$$\langle \mathcal{A} \rangle = \operatorname{Tr} \varrho \mathcal{A} \tag{2.36}$$

From molecular dynamics trajectories one can also compute averages, but these are *time averages* of the instantaneous value $A(t) := \mathcal{A}(\{\mathbf{x}_i(t)\})$ of the observable $\mathcal{A}$,

$$\overline{\mathcal{A}} = \lim_{\mathfrak{T}\to\infty} \frac{1}{\mathfrak{T}} \int_0^{\mathfrak{T}} dt\, A(t), \tag{2.37}$$

where $\mathfrak{T}$ is the time during which the system evolves. The ergodic theorem states the equivalence of the ensemble average (2.36) and the time average (2.37) in equilibrium.[106] In computer simulations, however, finite resources do not allow us to obtain the limit $\mathfrak{T} \to \infty$.* In fact, the ergodic theorem is largely irrelevant for our purposes because it requires time scales for $\mathfrak{T}$ which scale roughly with e^N, where N is the number of particles.[108] Even for the small systems used in simulations ($N = 10^2 \ldots 10^6$) the length of these *Poincaré cycles* is beyond all computational capabilities.† In practice, the time average Eq. (2.37) can still be a good approximation to Eq. (2.36) for $\mathfrak{T} < \infty$, provided that the trajectory samples a large enough volume of the relevant phase space[110] (where "relevant" is an attribute of configurations $\{\mathbf{x}_i\}$ which contribute strongly to the average (2.36) due to their large Boltzmann factor).[105]

Measures of spatial variation A special group of averages measures mobility or flexibility of particles. A particle in a bound state performs thermal vibrations about its equilibrium position $\mathbf{x}_0 = \langle \mathbf{x} \rangle$. The range of movement of the particle (or in a certain sense, its flexibility) is characterised by the root mean square deviation

$$\rho := \sqrt{\langle (\mathbf{x} - \mathbf{x}_0)^2 \rangle} = \sqrt{\overline{(\mathbf{x}(t) - \overline{\mathbf{x}(t)})^2}}. \tag{2.38}$$

As shown in Section 4.3.1 on page 95, ρ can simply be related to a harmonic force constant, which in turn is a direct measure of how strongly the particle is

*…unless one considers rather non-standard scenarios for the nature of reality, namely a "omega-point" at the end of the universe in a Big Crunch.[107]

†For $N = 100$ particles and a time step of $\Delta t = 1$ fs $= 10^{-12}$ s the *simulated* time $\mathfrak{T}$ would already have to be $\Delta t\, e^N = 2.7 \times 10^{31}$ s $= 6 \times 10^{13}\, t_0$ where $t_0 = 14.1 \pm 0.9 \times 10^9$ yr is the age of the universe.[109]

bound. An ion or water molecule in solution, on the other hand is not bound but diffuses freely. In equilibrium, its movement is described by a random walk.[111] Thus, the average position of the particle does not change over time ($\langle \mathbf{x}(t) \rangle = 0$) but the root mean square deviation

$$\rho(t) := \sqrt{\langle (\mathbf{x}(t) - \mathbf{x}(t_0))^2 \rangle} \tag{2.39}$$

does. It measures how far on average a particle has drifted. Einstein[112] showed that for a random walk and times t sufficiently larger than the molecular collision time, $\rho(t) \propto \sqrt{t}$ (see also Eq. (3.12) on page 66 in Section 3.2.3).

Constrained free energies

A general concept in statistical mechanics is the idea of a *constrained free energy* (for instance, the phenomenological treatment of phase transitions through Landau free energies).[99] We ask what value a free energy $\mathcal{W}$ takes on when an observable $\mathcal{A}(\{\mathbf{x}_i\})$ has a particular prescribed value a. (Depending on the situation, $\mathcal{W}$ could be the Helmholtz free energy F, the Gibbs function G, or the grand potential Ω.) Calculating the value of $\mathcal{W}$ for all a defines the constrained free energy $\mathcal{W}(a)$. The constraint is expressed in the partition function by performing a *partial* trace over only those configurations $\{\mathbf{x}_i\}$ where $\mathcal{A}$ has the value a,

$$Z(a) := \operatorname*{Tr}_{\mathcal{A}(\{\mathbf{x}_i\})=a} e^{-\beta \mathcal{H}} \quad \text{and} \tag{2.40}$$

$$\Xi(a) := \operatorname*{Tr}_{\mathcal{A}(\{\mathbf{x}_i\})=a} e^{-\beta(\mathcal{H}-\mu N)}. \tag{2.41}$$

The corresponding constrained free energy is

$$\beta F(a) = -\ln Z(a) \quad \text{and} \tag{2.42}$$

$$\beta \Omega(a) = -\ln \Xi(a). \tag{2.43}$$

The connection to simulations is established through the probability to observe the system in a state where $\mathcal{A}$ takes the value a,

$$p(\mathcal{A} = a) = \frac{\mathcal{Z}(a)}{\mathcal{Z}} \tag{2.44}$$

where $\mathcal{Z}$ stands for the partition function appropriate for the free energy $\beta\mathcal{W} = -\ln \mathcal{Z}$. $\mathcal{Z}$ measures the total volume of phase space accessible to the system whereas $\mathcal{Z}(a)$ describes the (smaller) volume of phase space that is accessible under the constraint $\mathcal{A}(\{\mathbf{x}_i\}) = a$. Because all points in phase space have the same probability weight, the probability to find the system in state a is simply

the quotient of the two phase space volumes. $p(a)$ can be obtained from experiments or simulations, and hence the constrained free energy is

$$\beta\mathcal{W}(a) = -\ln p(a) + \beta\mathcal{W}. \tag{2.45}$$

Because the absolute free energy $\mathcal{W}$ is difficult to obtain in most cases, constrained free energies are often only known up to a constant, i.e. $\mathcal{W}$. It should be noted that $\mathcal{W}$ is only a constant with respect to $\mathcal{A}$; if other thermodynamic potentials like temperature or volume change, then $\mathcal{W}$ changes, too. If the absolute value of $\mathcal{W}(a)$ is required then one possibility is to determine $\mathcal{W}$ for a reference system analytically and try to bridge between the simulated system and the reference system. In many cases, however, one is more interested in the relative shape of the energy landscape as a function of a and in these cases the constant term does not matter.

2.1.4 Volume and other fuzzy concepts

For thermodynamic analysis we would like to be able to compute densities from the simulations. It is easy to count particles in a given region of space but in order to compute a density one needs to know the volume that is available to the particle. Often thermodynamics yields surprisingly good results even when applied to microscopic systems though it is by no means required to do so.* An example is classical nucleation theory which seems to be valid down to clusters of a few atoms.[113] When trying to compute a density we are really trying to describe a discrete system (made from particles) by a continuum description (the density, defined in the whole of space). In macroscopic and mesoscopic systems one is used to apply concepts like volume, area, or pressure without further thoughts. Rowlinson[114] pointed out that these quantities are not well defined at the molecular scale. A good example is the interface between liquid and vapour. At the length scale of the fluid molecules the interfacial density profile varies over a few molecular diameters.[100,101,115] Similarly, the surface between a solid wall, modelled as Lennard-Jones spheres (and not hard spheres), and a fluid is complicated because the particles behave like compressible or "squidgy" balls and not hard billiard balls. Given enough energy, a fluid molecule can approach the centre of a wall atom arbitrarily close. The problem is hardly a new one, and for interfaces the concept of the Gibbs dividing surface is well-known,[100] where the position of the surface is recognised to be arbitrary in the sense that no thermodynamic function may depend on its position.

Still, it seems desirable to be able to use concepts like "volume"[116] or "surface area",[117] at least in a consistent fashion (unless one is prepared to completely do without them as suggested by Rowlinson[114]). Here we describe two

*Equally, there is no obvious reason why it should *not* give meaningful results for molecular systems (barring quantum effects).

methods to compute volumes and, more importantly, densities as "number of particles divided by some suitable volume". It should be noted, that the *microscopic density* $n(\mathbf{x})$ is rather unambiguous when interpreted as the average number of particles in a small volume element d^3x as in Eq. (2.16). Problems arise at boundaries between different phases when macroscopic densities for a whole system are computed from them. For a fluid in the pore the question really is what is meant by "in the pore."

The molecular surface

Qualitatively, a simple van der Waals surface of overlapping "hard sphere" atoms can describe the region of e.g. the pore lumen of an ion channel that is accessible to solvent. At a higher level of sophistication only the truly solvent accessible surface[118] and volume[119] are calculated by accounting for re-entrant surfaces.* In a similar manner, the program hole[42] traces out the solvent accessible surface of a pore. All these methods treat atoms as hard spheres with a well defined radius which is typically taken to be the van der Waals radius. In Appendix B a method to calculate consistent volumes is proposed which explicitly takes into account the softness of atoms; however, it is not pursued any further in the main body of this work.

The basic problem that we face is to calculate the volume of a cylindrical open pore. The two approaches that are used in the following are simple and try to define an effective radius that can be used in the same way as hard-sphere surface radius.

Shifted van der Waals surface

Atoms in the pore wall are treated as hard van der Waals spheres, and in fact, they are assumed to form a smooth, unrugged surface, which can be described by a single radius R. Here R describes the *van der Waals surface* of the pore, i.e. the interface where the atomic surfaces of a water molecule and a wall atom would touch.† The neglect of the re-entrant surfaces and of the ruggedness is justified as calculations of the exact rugged volume (neglecting re-entrancy) of a pore in Appendix C show.

The effective boundaries between, e.g. fluid and wall, are chosen so that for the largest system quantities like the density are consistent when calculated either from the volume or from an average over small sub volumes. For the

*The re-entrant surface bounds the regions between overlapping atomic spheres inaccessible to the van der Waals surface of a solvent molecule.

†The formal definition of the van der Waals surface is the union of the van der Waals spheres of all atoms in the molecule. Note that this can include internal cavities which are not solvent accessible.

volume of a cylindrical pore this amounts to calculating a correction δR to the van der Waals surface radius of the interior R by requiring that

$$\pi(R+\delta R)^2 L = \int_0^L dz \int_0^{2\pi} d\phi \int_0^{R_0} dr\, r n(r,\phi,z), \tag{2.46}$$

where R_0 is the radius where the density vanishes.* (2.46) evaluated for the largest pore yields the value of $\delta R = -0.03$ nm which is used throughout all calculations. The effective volume from a given pore radius R is hence given

$$V_{\text{eff}} := \pi(R+\delta R)^2 L = \pi R_{\text{eff}}^2 L. \tag{2.47}$$

($R_{\text{eff}} := R + \delta R$ is the effective radius.)

Fluid "radius of gyration"

Alternatively, we can estimate the effective radius of a cylindrical cavity by the average radius of gyration of the fluid molecules in a disk of thickness dz.† In fact we are interested in R as a function of z, i.e. the *pore profile* $R(z)$, as it fully characterises a cylindrically symmetric pore and can also be used for simple estimates of the ionic conductance of the pore.[12,42] The definition

$$R_{\text{G}}^2(z) := \frac{2\int_0^{2\pi} d\phi \int_0^{R'} r\, dr\, r^2\, n(r,\phi,z)}{\int_0^{2\pi} d\phi \int_0^{R'} r\, dr\, n(r,\phi,z)} \tag{2.48}$$

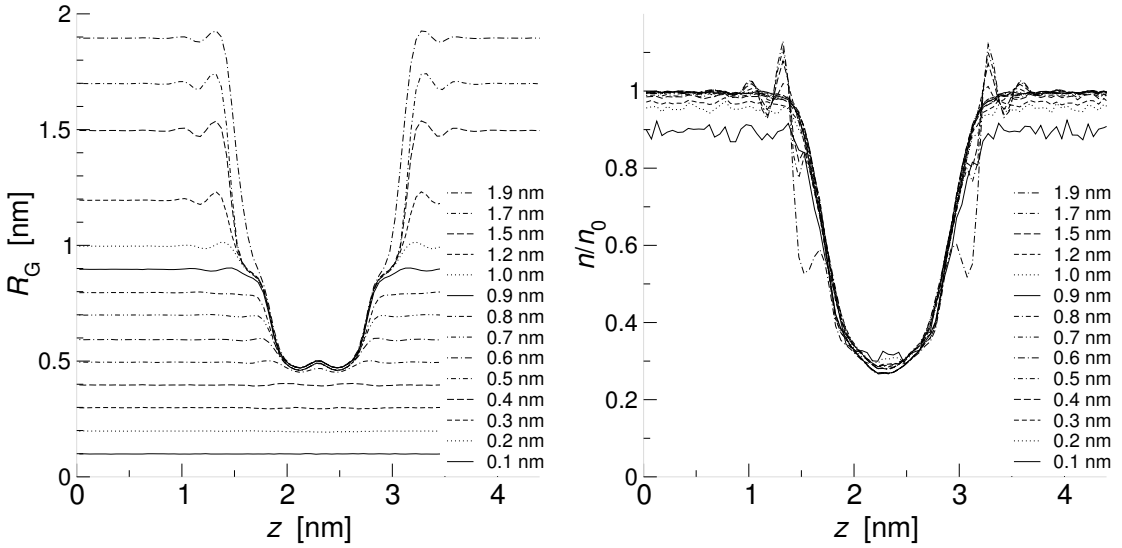
has the correct normalisation so that (2.48) yields the correct result $R_{\text{G}}(z) = R(z)$ for a constant, cylindrically symmetric density $n(r,z)$ with a sharp interface at R, i.e. $n(r,z) = n_0(z)\,\Theta\big(R(z)-r\big)$, when the integration range $R' \geq R$ is chosen sufficiently large (Θ is the Heaviside step function). $R_{\text{G}}(z)$ can be used as a pore profile; it could be described as an "ensemble pore profile" because it

*More precisely, R_0 should be defined as $R_0(\phi,z)$ but in practical applications (cf. Section 3.2.3 on page 62) we computed a radial density first and found R_0 as the value where the radial density vanished.

†The term "radius of gyration" is used in a rather loose sense for an ensemble of fluid molecules; it has a much more precise meaning when applied to a single (macro) molecule such as a protein or a simple polymer, which consists of atoms connected by bonds:

$$R_G = \sqrt{\frac{\sum_{i=1}^N m_i\,(\mathbf{x}_i - \mathbf{x}_{\text{com}})^2}{\sum_{i=1}^N m_i}},$$

where $\mathbf{x}_{\text{com}}$ is the centre of mass of the polymer consisting of N atoms of mass m_i each.



(a) Pore profile $R(z)$ (b) Axial average density $n(z)$

Figure 2.1: Density $n(z)$ and pore profile as calculated from the fluid radius of gyration R_G, Eq. (2.48). The fluid is SPC water in a model pore of nominal radius $R = 0.55$ nm in a simulation box of dimensions $4 \times 4 \times 4.6$ nm^3. The integrals (2.48) and (2.49) were calculated for increasing radial limits R', as indicated by the different line styles. Note that the radius of the mouth region of the pore is 1 nm.

is built from the spatial movements of an ensemble of particles in the pore. The average density $n(z)$ in a disk centred at z is then calculated as

$$n(z) = \frac{\int_0^{2\pi} d\phi \int_0^{R'} r\, dr\, n(r, \phi, z)}{\pi R_G^2(z)}. \tag{2.49}$$

Test calculations showed that both Eq. (2.48) and (2.49) are rather insensitive to the value of the integration boundary R'; typical data for water in a model pore are shown in Fig. 2.1. Of course, only surfaces are detected which are included in the integration. Eq. (2.49) will be referred to as calculating the average density profile from the "local density" (i.e. the grid cells, in which the density is well defined). This method was implemented as a function in the grid analysis program a_ri3Dc (Appendix D.2.3) and density profiles along the pore axis were calculated according to Eq. (2.49).

For simple model pores the density was calculated in a simpler fashion by averaging the density in disks of constant radius (and this proved to be sufficient). But for the pore of the nicotinic receptor the method based on the fluid radius of gyration had to be employed because of its irregular shape. It should also be noted that the pore profile $R_G(z)$ essentially describes the *solvent accessible*

surface[*] [120] (SAS) of the pore because the calculation is based on the density or distribution of the centres of mass of the probe particles. Calculating the *solvent excluded surface*[†] [122] (SES) is not difficult in principle either. One would have to smear out the density $n(\mathbf{x})$ with the volume of the probe molecule, modelled by an appropriate shape function $\chi(r)$.[‡] The new "volumetric" density $n_V(\mathbf{x})$ would be the convolution

$$n_V(\mathbf{x}) = \chi * n = \int d^3x' \, \chi(\mathbf{x}')n(\mathbf{x} - \mathbf{x}').$$

2.2 Molecular Dynamics Simulations

For comparatively simple systems, analytical methods can already give profound insights into the physics of the problem. For more complex systems one turns to computational methods that integrate the fundamental equations of motions of all particles (molecular dynamics) or generate system configurations according to a given thermodynamic ensemble (Monte Carlo methods). These methods are often referred to as "simulations" because they show how a system behaves which is governed by an approximation to the laws of the real system. It does *not* imply that these simulations or "computer experiments" can be carried out instead of a scientific experiment in the real world. Ideally, a combination of experiment and computer modelling/theory helps us to better understand the problem in question.

2.2.1 The Born-Oppenheimer approximation

A full quantum mechanical description of a molecule is provided by the Hamiltonian

$$\mathcal{H} = \hat{T}_\mathrm{e} + \hat{T}_\mathrm{N} + \hat{V}_\mathrm{ee} + \hat{V}_\mathrm{eN} + \hat{V}_\mathrm{NN} \quad \text{with} \tag{2.50}$$
$$\hat{T}_\mathrm{e} = \sum_i \frac{\mathbf{p}_i^2}{2m}, \quad \hat{T}_\mathrm{N} = \sum_N \frac{\mathbf{P}_N^2}{2M_N},$$

[*] The solvent accessible surface is the surface traced out by the *centre* of the (spherical) probe while it rolls over the van der Waals surface of the protein.

[†] The solvent excluded surface is the surface traced out by the probe surface when it rolls over the van der Waals surface. It may contain re-entrant surfaces which are the concave patches "between" atomic spheres. The interested reader is referred to more specialised articles [118,119,121] for a discussion of molecular surfaces.

[‡] Possible shape functions: $\chi(\mathbf{x}) := 1 - \exp[-\beta V^\mathrm{rep}(|\mathbf{x}|)]$, where the repulsive part of the Weeks-Chandler-Anderson decomposition [123] [Eq. (4.10)] of the Lennard-Jones potential is used; or a Gaussian-like approximation $\chi(\mathbf{x}) := \exp[-(|\mathbf{x}|/\sigma)^{2n}]$, which has the advantage that the normalisation $4\pi \int_0^\infty dr \, r^2 \chi(r)$ can be determined analytically.

where $\mathcal{H}$ contains the kinetic energy operators for the electrons 'e' (mass m) and the nuclei 'N' (individual masses M_N), the electron-electron interaction, the electron-nuclei interaction, and the interaction between the nuclei. The position and momenta of the electrons are denoted $\mathbf{x}_i$ and $\mathbf{p}_i$ with $x \equiv (\mathbf{x}_1, \ldots \mathbf{x}_{N_\mathrm{e}})$, and $\mathbf{X}_N$ and $\mathbf{P}_N$ with $X \equiv (\mathbf{X}_1, \ldots, \mathbf{X}_{N_\mathrm{N}})$.*

Born and Oppenheimer[124] realised that the large difference in mass between electrons and nuclei leads to an effective separation of the motion of the nuclei from the electronic motion; the electronic degrees of freedom x adjust instantaneously to a varying nuclear configuration X.

Following Ref. 125, pp257 we determine the wave function of the electrons $\psi(x|X)$ while keeping the nuclear positions X fixed,

$$(\hat{T}_\mathrm{e} + \hat{V}_\mathrm{ee} + \hat{V}_\mathrm{eN} + \hat{V}_\mathrm{NN})\psi(x|X) = E_\mathrm{e}(X)\psi(x|X) \tag{2.51}$$

where the nuclear positions play the role of parameters in the energy eigenvalue

$$\epsilon(X) = V_\mathrm{NN}(X) + E_\mathrm{e}(X). \tag{2.52}$$

$\epsilon(X)$ is the sum of the nuclear interactions and the energy eigenvalue of the electrons moving in the potential provided by the frozen nuclei. As an *ansatz* for the wave function of the whole molecule

$$\Psi(x, X) = \psi(x|X)\,\Phi(X)$$

we use a product of the electron wave function at frozen nuclear coordinates times the wave function $\Phi(X)$ of the nuclei themselves. From the stationary Schrödinger equation $\mathcal{H}\Psi = E\Psi$ follows

$$\begin{aligned}\psi(x|X)\left(\hat{T}_\mathrm{N} + \epsilon(X)\right)\Phi(X) = \psi(x|X)\,E\,\Phi(X) \\ - \sum_N \frac{-\hbar^2}{2M_N}\left[\Phi(X)\,\nabla_X^2\psi(x|X) + 2\nabla_X\Phi(X)\nabla_X\psi(x|X)\right].\end{aligned} \tag{2.53}$$

Multiplying by $\psi^*(x|X)$, integrating over x, and making the Born-Oppenheimer approximation, i.e. neglecting contributions from the sum over the nuclei (which can be shown to be smaller by a factor $m/M \leq 1/1836$ (Ref. 125)), yields the *Born-Oppenheimer* equation

$$\left(\hat{T}_\mathrm{N} + \epsilon(X)\right)\Phi(X) = E\Phi(X). \tag{2.54}$$

(2.54) is an effective Schrödinger equation for the nuclei. They move in the effective potential $\epsilon(X)$, Eq. (2.52), and arrange in such a way as to minimise the

*In this section only we use the notation more common in quantum mechanics, i.e. $\mathbf{x}$ for electrons and $\mathbf{X}$ for nuclei. In later sections we will again focus exclusively on the atoms (i.e. essentially the nuclei) and denote their positions with $\mathbf{x}$ as was done before.

sum of inter-nuclear Coulomb repulsion V_{NN} and electronic energy E_{e}, i.e. the equilibrium coordinates of the nuclei are those X which minimise $\epsilon(X)$. Hence the force on a nucleus K can be calculated as $-\boldsymbol{\nabla}_K \epsilon(\mathbf{X}_1, \ldots, \mathbf{X}_K, \ldots, \mathbf{X}_N)$, and because almost all of an atom's mass is concentrated in the nucleus, this determines the position of the whole atom.

2.2.2 Classical Molecular Dynamics

The Born-Oppenheimer equation Eq. (2.54) implies that the electronic degrees of freedom can be subsumed in an effective potential, and the dynamics of the atoms can be reduced to the classical dynamics of the nuclei in this potential. Then the system of N particles is described by the classical Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{x}_1, \ldots, \mathbf{x}_N), \tag{2.55}$$

with masses m_i, positions $\mathbf{x}_i(t)$, and momenta $\mathbf{p}_i$. U is chosen as to approximate the quantum-mechanical Born-Oppenheimer energy surface Eq. (2.52), i.e. $U(\mathbf{x}_1, \ldots, \mathbf{x}_N) \approx \epsilon(X)$. A Newtonian formulation of Eq. (2.55) leads to Newton's equations of motion[126]

$$m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i \tag{2.56}$$

where $\ddot{\mathbf{x}}_i(t)$ is the acceleration. The force $\mathbf{F}_i$ on the i-th particle is derived from the potential U,

$$\mathbf{F}_i = -\boldsymbol{\nabla}_i U(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N), \tag{2.57}$$

which depends on the positions of all N particles. A classical molecular dynamics trajectory $\mathbf{x}^N(t) := \big(\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t)\big)$ is generated from integrating Eq. (2.56) with (2.57) for all particles.

The parameterization of U is called a *force field* and is discussed in the next section.

2.2.3 Force Fields

A classical force field consists of the functional form of $U(\{\mathbf{x}_j\})$ (the potential function) and a set of parameters for different atom types. For MD simulations the gromacs software package[127,128] was used, and so this section follows the package's implementation notes[129] and other standard texts on the subject.[105,130]

Potential functions

The potential function $U(\{\mathbf{x}_j\})$ is a true many body potential which in principle contains all two-, three- up to N-body interactions. It is approximated

by a simpler form containing pairwise-additive two-body potentials for non-bonded interactions and selected two-, three-, and four-body interactions to describe bonded interactions. The classical approximation does not calculate the quantum mechanical wave functions but describes each atom as a mass point, carrying (fractional) point charges, which is bonded to other atoms by harmonic (or other simple) forces—generally referred to as a "balls and springs" model. The functional form of the potential function U is chosen as to give a convenient frame work for parameterization of the Born-Oppenheimer energy surface.

A form typically found in MD force fields is

$$U = U_{\mathrm{NB}} + U_{\mathrm{B}} \tag{2.58}$$

where the first term contains the non-bonded contributions,

$$U_{\mathrm{NB}}(\mathbf{x}_1, \dots \mathbf{x}_N) = \sum_{i<j} V_{\mathrm{LJ}}(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i<j} V_{\mathrm{Coul}}(\mathbf{x}_i, \mathbf{x}_j) \tag{2.59}$$

(neutral atom interactions and charged atoms interactions). The second one approximates bonded interactions with

$$U_{\mathrm{B}}(\mathbf{x}_1, \dots \mathbf{x}_N) = \sum_{(i,j)} V_{\mathrm{bond}}(\mathbf{x}_i, \mathbf{x}_j) + \sum_{(i,j,k)} V_{\mathrm{angle}}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \sum_{(i,j,k,l)} V_{\mathrm{torsions}}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l). \tag{2.60}$$

Non-bonded interactions At the quantum mechanical level the only interaction that plays a role is the Coulomb interaction between two charged particles.* At the classical level the effect of this interaction (in conjunction with the rules of quantum mechanics) is split up into different separate interactions.

The effect of the Pauli exclusion principle between the valence electrons of two atoms at distance $r \equiv x_{ij} = |\mathbf{x}_j - \mathbf{x}_i|$ is modelled as a steeply repulsive potential $\propto r^{-n}, n > 6$ whereas the dipole-dipole attraction between two neutral atoms is based on a $-r^{-6}$ term. Taken together, this defines the **Lennard-Jones potential**,

$$V_{\mathrm{LJ}}(x_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{x_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{x_{ij}}\right)^{6}\right], \tag{2.61}$$

where σ_{ij} is the distance at which the potential energy vanishes and the energy ϵ_{ij} is the depth of the attractive well at the equilibrium separation $2^{1/6}\sigma_{ij}$ An alternative equivalent form is

$$V_{\mathrm{LJ}}(x_{ij}) = \frac{C_{ij}^{(12)}}{x_{ij}^{12}} - \frac{C_{ij}^{(6)}}{x_{ij}^{6}}. \tag{2.62}$$

*Gravitation is much weaker than the electrostatic force and neither strong nor weak force have a direct influence on the behaviour of atoms due to their short range.

The parameters $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ (or ϵ_{ij} and σ_{ij}) depend on the atom types of the pair (i, j).* If a particular combination has not been explicitly parameterized the cross terms can be calculated by the empirical combination rules

$$C_{ij}^{(6)} = \sqrt{C_{ii}^{(6)}\, C_{jj}^{(6)}} \quad \text{and} \quad C_{ij}^{(12)} = \sqrt{C_{ii}^{(12)}\, C_{jj}^{(12)}} \tag{2.63}$$

or

$$\sigma_{ij} = \tfrac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad \text{and} \quad \epsilon_{ij} = \sqrt{\epsilon_{ii}\, \epsilon_{jj}}. \tag{2.64}$$

Electrostatic (Coulomb) interactions are calculated from Coulomb's law between two charged particles as

$$V_{\text{Coul}}(x_{ij}) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{x_{ij}}. \tag{2.65}$$

An important step in the parameterization of a force field is the assignment of partial charges q_i to atoms of a molecule; these partial charges interact through Eq. (2.65).

Bonded interactions Two atoms connected by a chemical bond behave like a classical harmonic oscillator for small deviations from their equilibrium bond length. Hence bonds are typically represented by a harmonic **bond-stretching potential**

$$V_{\text{bonds}}(x_{ij}) = \frac{k_{ij}^{\text{b}}}{2}\left(x_{ij} - b_{ij}\right)^2 \tag{2.66}$$

around the atom-type dependent equilibrium bond length b_{ij} with a force constant k_{ij}^{b}. An equally valid approximation is a fourth-power potential†

$$V_{\text{bonds}}(x_{ij}) = \frac{\tilde{k}_{ij}^{\text{b}}}{4}\left(x_{ij}^2 - b_{ij}^2\right)^2. \tag{2.67}$$

Directionality of bonds is incorporated through **angle potentials** which embody harmonic angular vibrations between three atoms (i, j, k),

$$V_{\text{angle}}(\theta_{ijk}) = \frac{k_{ijk}^{\theta}}{2}\left(\theta_{ijk} - \alpha_{ijk}\right)^2 \qquad \text{with} \quad \theta_{ijk} := \arccos\frac{\mathbf{x}_{ji}\cdot\mathbf{x}_{jk}}{x_{ji}x_{jk}}, \tag{2.68}$$

*The relation between the two parameter sets is $\sigma = (C^{(12)}/C^{(6)})^{1/6}$ and $\epsilon = \frac{1}{4}(C^{(6)})^2/C^{(12)}$, or $C^{(6)} = 4\epsilon\sigma^6$ and $C^{(12)} = 4\epsilon\sigma^{12}$.

†The force constants in Eq. (2.66) and (2.67) are related by $2\,\tilde{k}_{ij}^{\text{b}}\, b_{ij} = k_{ij}^{\text{b}}$.

where α_{ijk} is the equilibrium angle between atoms of the types represented by i, j, and k.* Sometimes a simplified version of Eq. (2.68) is used instead,†

$$V_{\text{angle}}(\theta_{ijk}) = \frac{\tilde{k}^{\theta}_{ijk}}{2}\left(\cos\theta_{ijk} - \cos\alpha_{ijk}\right)^2 \qquad \text{with} \quad \cos\theta_{ijk} := \frac{\mathbf{x}_{ji}\cdot\mathbf{x}_{jk}}{x_{ji}x_{jk}}. \tag{2.69}$$

Torsional vibrations are described by dihedral potentials. The dihedral angle between the planes (i,j,k) and (j,k,l) is defined by

$$\phi_{ijkl} = \arccos\frac{(\mathbf{x}_{ji}\times\mathbf{x}_{jk})\cdot(\mathbf{x}_{kl}\times\mathbf{x}_{jk})}{|\mathbf{x}_{ji}\times\mathbf{x}_{jk}|\cdot|\mathbf{x}_{kl}\times\mathbf{x}_{jk}|} \tag{2.70}$$

so that $\phi_{ijkl} = 0$ when i and l are in the *cis* conformation (the IUPAC/IUB convention). One choice for a periodic dihedral potential is

$$V_{\text{dih}}(\phi_{ijkl}) = k^{\phi}_{ijkl}\left(1 + \cos(n\phi_{ijkl} - \phi_{0,ijkl})\right) \tag{2.71}$$

with dihedral force constant k^{ϕ}_{ijkl} (typically depends only on atom types of j and k), multiplicity n and phase shift $\phi_{0,ijkl}$. Another choice is the Ryckaert-Bellemans form

$$V_{\text{dih}}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n \cos^n(\phi_{ijkl} - \pi). \tag{2.72}$$

A special form of (2.72) is used in the OPLS-AA force field,

$$V_{\text{dih}}(\phi_{ijkl}) = V_0 + \frac{1}{2}\sum_{n=1}^{3} V_n(1 + \cos n\phi_{ijkl}), \tag{2.73}$$

which are the first four terms of a Fourier series in the dihedral potential. An improper dihedral angle ξ_{ijkl} is defined in the same way as ϕ_{ijkl} in Eq. (2.70) but the four atoms can also be part of a tetrahedral or a ring structure. The **improper dihedral** potential

$$V_{\text{impr}}(\xi_{ijkl}) = k^{\xi}_{ijkl}(\xi_{ijkl} - \xi_{0,ijkl})^2 \tag{2.74}$$

is used to keep rings planar or prevent molecules from flipping over to their mirror image.

In the following sections the values that were chosen for the parameters in Eqs. (2.61) to (2.74) are briefly discussed.

*The definition used for connecting vectors $i \to j$ is $\mathbf{x}_{ij} := \mathbf{x}_j - \mathbf{x}_i$.

†The force constants in Eq. (2.68) and (2.69) are related by $\tilde{k}^{\theta}_{ijk}\sin^2\alpha_{ijk} = k^{\theta}_{ijk}$.

Protein force field

For a simulation, parameters have to be chosen for the protein, ligands, water, and ions. All these parameter sets have to be compatible and some care needs to be taken when combining different parameter sets.

Model pores Many simulations in this work were carried out on systems containing primarily water and model pores, which are much simpler than real proteins. These pores were constructed from methane molecules (see in particular Section 3.2.1 and also Appendix D.1.1) to make them hydrophobic. A methane molecule is modelled as a unified atom, i.e. the effect of the hydrogen atoms is incorporated into the LJ potential of the unified CH_4 "atom". The LJ parameters for the interaction between a methane molecule and the water oxygen are $\epsilon_{CO} = 0.906493$ kJ mol^{-1} and $\sigma_{CO} = 0.342692$ nm and were taken directly from the GROMOS96 force field.

Proteins Today there exists a range of well-tested force fields suitable for molecular simulations of proteins.[131] In Chapter 6 calculations are carried out for two different force fields to assess the reliability of the results.

The **GROMOS96**[132,133] force field is a united atom force field, i.e. it treats non-polar hydrogen atoms as part of a larger united atom; hydrogens on aromatic rings and polar hydrogens are treated explicitly.* The united atom scheme reduces the typical system size from about 15 000 atoms to about 10 000 (cf. Tab. 6.1 on page 132) with a corresponding speed up in simulation execution. GROMOS96 uses Eq. (2.67) for bonds, Eq. (2.69) for angles, Eq. (2.71) for proper dihedrals, and all other interactions as described in Section 2.2.3.

The second force field, **OPLS-AA**[134–140], is an all atom force field, which treats all hydrogens explicitly.† An all atom force field is arguably a better choice when it comes to investigating hydrophobic gating. The important hydrophobic side chains are modelled as unified atoms in GROMOS96 but simulations of model pores showed that the wall-water interaction is important and conceivably a better description of this interaction is provided if all hydrogen atoms are present as is the case in OPLS-AA. The functional form of OPLS-AA differs from GROMOS96 in that bonds are calculated by Eq. (2.66), angles by Eq. (2.68), and proper dihedrals by Eq. (2.72).

GROMOS96 is optimised with the SPC water model[132] whereas TIP4P[141] is the native water model of OPLS-AA. For long (50 ns) simulations of the 36-residue Villin headpiece van der Spoel and Lindahl[142] evaluated combinations

*All hydrogens were "dummified", i.e. their position was calculated from the position of their connecting heavy atoms,[129] saving the calculation of forces and speeding up the calculation.

†Hydrogens were dummified[129] as for GROMOS96.

Table 2.1: Water model parameters. Partial charges (in units of e) on the oxygen and hydrogens are given by $q(\mathrm{O})$ and $q(\mathrm{H})$. In the 4-point model TIP4P, $q(\mathrm{M})$ gives the total charge associated with the oxygen lone pairs, $r(\mathrm{OM})$ the distance of the fourth interaction site M from the oxygen, lying on the bisector of the HOH angle. ϵ and σ are the Lennard-Jones parameters of the oxygen atom; the hydrogen sites only contribute through their Coulomb interaction.

	Experiment	SPC	TIP4P
$r(\mathrm{OH})/\mathrm{nm}$	0.0957	0.10	0.0957
$\theta(\mathrm{HOH})/^\circ$	104.52	109.47	104.52
$q(\mathrm{O})$		-0.82	0.0
$q(\mathrm{H})$		0.41	0.52
$q(\mathrm{M})$			-1.04
$r(\mathrm{OM})/\mathrm{nm}$			0.015
σ/nm		0.316565	0.315365
$\epsilon/\mathrm{kJ}\cdot\mathrm{mol}^{-1}$		0.650167	0.648520
$C^{(6)}\times 10^3/\mathrm{kJ}\cdot\mathrm{mol}^{-1}\cdot\mathrm{nm}^6$		2.61735	2.55190
$C^{(12)}\times 10^6/\mathrm{kJ}\cdot\mathrm{mol}^{-1}\cdot\mathrm{nm}^{12}$		2.63413	2.51041
Reference	47	132, 143	144, 145

of water model with force field and compared to NMR data. They concluded that best results are obtained by the combination of GROMOS96 with SPC water or OPLS-AA with TIP4P, although the combination OPLS-AA with SPC does not score significantly worse.

Water and ion parameters

A range of water models exist that have been widely used and tested.[146–148] Each water molecule is explicitly simulated, using the single point charge (SPC) model[132,143] in most simulations.* It consists of a oxygen site with a partial negative charge and a LJ-potential, and two hydrogen sites, which only interact through their positive partial charge. The HOH angle and bond lengths are rigid; parameters are listed in Tab. 2.1. Like all point-charge water models it lacks polarization effects, which are compensated by a dipole moment that is larger than the real dipole moment of the water molecule. The TIP4P[141,144] water model was used in some simulations which included proteins. It differs from SPC in that the oxygen is split into a LJ-site, which carries the mass, and a charge site, which only carries the partial negative charge (see Tab. 2.1). TIP4P

*Short test simulations of water in a model pore (cf. Chapter 3) using the SPC/E, flexible SPC, and the TIP3P model did not display significant differences from the SPC simulations.

Table 2.2: Ion parameters. Given are the Lennard-Jones parameters used in Eq. (2.61) and Eq. (2.62).

	Na^+		Cl^-	
	GROMOS96	OPLS-AA	GROMOS96	OPLS-AA
q/e	$+1$	$+1$	-1	-1
m/u	22.98980	22.98977	35.45300	35.45300
σ/nm	0.257536	0.333045	0.444796	0.441724
ϵ/kJ · mol^{-1}	0.061749	0.0115980	0.445708	0.492833
$\frac{C^{(6)} \times 10^4}{\text{kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^6}$	0.72063121	0.633083	138.0625	146.442
$\frac{C^{(12)} \times 10^7}{\text{kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{12}}$	0.21025	0.863929	1069.156	1087.86
Reference	133	149	133	150

was used because it is the recommended water model with the OPLS-AA force field.

Ions are parameterized as LJ-particles with charge $+1e$ (Na^+) or $-1e$ (Cl^-), listed in Tab. 2.2. For calculations that only included the SPC water model and methane molecules (i.e. the model pore calculations), the ion parameters appropriate for GROMOS96 were chosen. Simulations of the nicotinic receptor, which were carried out with both GROMOS96 and OPLS-AA, used the ion parameters appropriate for the protein force field (see Tab. 2.2). Test calculations for a NaCl bulk electrolyte (GROMOS96 and SPC) yielded radial distribution functions in agreement with experimental ionic radii.[12,151]

2.2.4 Algorithms

In order to integrate the equations of motions Eq. (2.56) the forces need to be calculated as the gradient of the force field, Eq. (2.57). For the force field formulation (2.58)–(2.74) the derivatives are calculated analytically.[129]

Integration of the equations of motion

The Verlet algorithm[152] is a simple and robust scheme to integrate Newton's equations of motion,

$$\mathbf{x}(t + \Delta t) \approx 2\mathbf{x}(t) - \mathbf{x}(t - \Delta t) + \frac{\mathbf{F}}{m} \Delta t. \tag{2.75}$$

It can be derived from the Taylor expansion of $\mathbf{x}(t+\Delta t) + \mathbf{x}(t-\Delta t)$.[105,130,153] The error in the positions is of order Δt^4; velocities $\mathbf{v}(t) = \frac{1}{2\Delta t}\big(\mathbf{x}(t+\Delta t) - \mathbf{x}(t-\Delta t)\big) + \mathcal{O}(\Delta t^2)$ are less accurate but are only required to compute the kinetic energy and thus the temperature. The Verlet algorithm has good long time stability and energy conservation because it is a symplectic integrator. It is not just a scheme to solve a set of differential equations but it also conserves properties inherent in a Hamiltonian system, namely energy conservation, time reversal symmetry, volume of phase flow and the sum of areas of phase elements.[130,153]

Leap-frog integrator In gromacs a variant of the Verlet integrator is employed, the leap-frog algorithm,[154] which defines the velocities at half-integer time steps[105,129,130,153]

$$\mathbf{v}(t+\tfrac{1}{2}\Delta t) := \frac{\mathbf{x}(t+\Delta t) - \mathbf{x}(t)}{\Delta t} \quad \text{and} \quad \mathbf{v}(t-\tfrac{1}{2}\Delta t) := \frac{\mathbf{x}(t) - \mathbf{x}(t-\Delta t)}{\Delta t}.$$

This leads to the update for the positions

$$\mathbf{x}(t+\Delta t) \approx \mathbf{x}(t) + \mathbf{v}(t+\tfrac{1}{2}\Delta t)\,\Delta t \tag{2.76}$$

and the update of the velocities is taken from the Verlet scheme (2.75)

$$\mathbf{v}(t+\tfrac{1}{2}\Delta t) \approx \mathbf{v}(t-\tfrac{1}{2}\Delta t) + \frac{\mathbf{F}}{m}\,\Delta t. \tag{2.77}$$

It can be shown to be formally equivalent to the Verlet scheme and thus shares its symplectic properties, but has the advantages of a smaller error on the velocities and is less prone to the introduction of numerical imprecision.[105]

Bond constraining The time step Δt in the integration scheme depends on the fastest motions in the system. For liquids and biomolecules in the biological temperature range these are bond vibrations, which are about 10–50 times faster than translational motions.[153] By treating harmonic bonds as rigid, time steps can be increased by about a factor of four.[155] Rigid bonds are implemented as constraints, which are enforced after each time step by resetting the bonds to their fixed length. For protein bonds the LINCS[155] algorithm was employed (an improvement over traditional SHAKE[156] in terms of stability and speed), whereas for the rigid water models the optimised SETTLE[157] procedure was used.

Electrostatics under periodic boundary conditions

The size of the systems that can be currently simulated with MD is still so small that the surface region of the system is large compared to the volume. Because

one is interested in volume behaviour the surface is eliminated by periodically repeating the simulation cell, known as periodic boundary conditions. A particle exiting on one side of the cell enters from the opposite face with the same velocity.

Care must be taken when calculating non-bonded interactions. For short range interactions, which decay faster than r^{-3} a simple cut-off scheme suffices, where these interactions are only calculated between particles i and j within a cut-off radius R_{cut} when $|x_{ij}| \leq R_{\text{cut}}$. The cut-off radius cannot be larger than the shortest cell dimension.

The Coulomb interaction (2.65), however, has a long range and in a periodic, i.e. crystal-like, arrangement contributions from an infinite number of neighbouring image cells need to be considered.[158] Truncation (cut-off) schemes exist for long range forces with varying degrees of sophistication. In particular for water and electrolytes they have been shown to introduce serious artifacts due to ordering just outside the cut-off region.[158,159] In this work great emphasis is put on the behaviour of water and ions and hence an exact method was chosen to evaluate all contributions to the Coulomb energy. It is based on the Ewald summation method,[160] initially devised to calculate the electrostatic energy of ionic crystals.

Following Deserno and Holm[161] the total electrostatic energy of a periodic system (box length L, periodic in three dimensions $\mathbf{n} = (n_1, n_2, n_3)$) can be written as

$$U_{\text{Coul}} = \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{i,j=1}^{N}{\sum_{\mathbf{n}\in\mathbb{Z}^3}}' \frac{q_i q_j}{|\mathbf{x}_{ij} + \mathbf{n}L|}, \tag{2.78}$$

where the primed sum indicates that for $i = j$ the term $\mathbf{n} = 0$ has to be omitted. Eq. (2.78) is only conditionally convergent so its value actually depends on the limiting procedure employed.[158] The slowly converging sum can be split into two fast converging ones, based on the identity

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

one term collects the rapid variations for short distances, the other the slow decay at large separations. The Ewald choice $f(r) = \text{erfc}(r) := \frac{2}{\sqrt{\pi}}\int_r^{+\infty} dt\, \exp(-t^2)$ screens the point charges by Gaussian charge densities (which makes their interactions decay rapidly with distance) and then corrects for the lattice of screening charges (which can be computed in reciprocal space). The Ewald energy is

$$U_{\text{Coul}} = U^{(\text{r})} + U^{(\text{k})} + U^{(\text{s})} + U^{(\text{d})}, \tag{2.79}$$

where the contribution from real space is (now using Gaussian units for notational convenience)

$$U^{(\mathrm{r})} = \frac{1}{2}\sum_{i,j}\sum_{\mathbf{m}\in\mathbb{Z}^3}{}' q_i q_j \frac{\operatorname{erfc}(\alpha|\mathbf{x}_{ij}+\mathbf{m}L|)}{|\mathbf{x}_{ij}+\mathbf{m}L|}, \tag{2.80}$$

from reciprocal space

$$U^{(\mathrm{k})} = \frac{1}{2L^3}\sum_{\mathbf{k}\neq 0}\frac{4\pi}{k^2}e^{-\frac{k^2}{4\alpha^2}}|\tilde{\rho}(\mathbf{k})|^2, \tag{2.81}$$

the self-energy

$$U^{(\mathrm{s})} = -\frac{\alpha}{\sqrt{\pi}}\sum_i q_i^2, \tag{2.82}$$

and the dipole correction

$$U^{(\mathrm{d})} = \frac{2\pi}{(1+2\epsilon')L^3}\left(\sum_i q_i\mathbf{x}_i\right)^2. \tag{2.83}$$

The Fourier transform of the charge density $\rho(\mathbf{x})$ is

$$\tilde{\rho}(\mathbf{k}) = \int d^3x\,\rho(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}} = \sum_{j=1}^{N} q_j e^{-i\mathbf{k}\cdot\mathbf{x}_j}. \tag{2.84}$$

The Ewald tuning parameter α adjusts the relative contributions of real and reciprocal space terms. The advantage of splitting Eq. (2.78) in such a way is that the sums (2.80) and (2.81) converge exponentially with $\mathbf{m}$ and $\mathbf{k}$ and so a small cut-off can be chosen for them without much loss of accuracy.[153,158,161] CPU time for the Ewald summation scales with $N^{3/2}$ due to the Fourier sum Eq. (2.81) with (2.84), which is $\mathcal{O}(N^2)$.[153] For large systems a more efficient approach to computing $U^{(\mathrm{k})}$ is provided by the Fast Fourier Transform (FFT, Ref. 162), which is $\mathcal{O}(N\log N)$. However, FFT requires the charges to be located on a regular grid or mesh. Deserno and Holm[161] present a careful study of a number of algorithms (particle-particle/particle-mesh (PPPM, Ref. 163), particle-mesh (PME, Ref. 164), and smooth PME (SPME, Ref. 165) Ewald summation; the SPME method was used in this work), which all employ the same idea: First the particle charges are spread or interpolated over a regular mesh. Then the mesh density is fast-Fourier transformed (the equivalent of Eq. (2.84)) and the reciprocal space energy $U^{(\mathrm{k})}$ (analogously to Eq. (2.81)) is evaluated

(this amounts to solving Poisson's equation). Then forces on the interpolated charges are calculated as the derivative of the Coulomb energy (SPME calculates analytical derivatives, which is fast in terms of computational speed but not as exact as a Fourier space derivative, i.e. multiplication by ik, Ref. 161). Finally, the forces are back-interpolated on the original particle positions to yield the Coulomb contribution to the total force on the particle.

All calculations in the following chapters employed the SPME[165] method and used the same set of parameters: the real space cutoff in Eq. (2.80) was 1 nm, the Fourier spacing was 0.15 nm^{-1} (which fixes the maximum wave vector in the sum Eq. (2.81)), and $\alpha = 10^{-5}$. Fourth order cubic splines were used in the SPME interpolation procedure. These parameters should give electrostatic energies accurate to 5×10^{-3}, which is better than the typical accuracy of the LJ terms.[129]

Constant temperature, constant pressure simulations

Laboratory experiments are typically conducted under conditions of constant temperature and constant pressure. The trajectories obtained from the Hamiltonian (2.55) are total energy conserving and thus the system is described by the microcanonical ensemble. Different schemes have been proposed to generate other ensembles. The first class of schemes modifies Eq. (2.55) and are known as extended Hamiltonian methods;[130,153] they can be shown to generate well defined thermodynamic ensembles. Methods of a second class, named weak coupling schemes,[130,166] are conceptually simpler but do not lead to an exactly defined ensemble, although the practical deviations are believed to be negligible.[129] In this work, the weak coupling schemes were used to simulate at constant temperature and pressure.

Temperature coupling The temperature T is computed from the velocities through the kinetic energy of the whole system,

$$\tfrac{1}{2}\nu\, kT = E_{\mathrm{kin}} = \tfrac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \quad \text{with} \quad \nu := 3N - N_{\mathrm{constraints}} - \nu_{\mathrm{com}}, \tag{2.85}$$

where ν is the number of degrees of freedom, i.e. all translational degrees of spherical atoms minus the number of constraints applied to these atoms (e.g. fixed bond lengths) and $\nu_{\mathrm{com}} = 3$ for centre of mass momentum removal. If one of the constant temperature algorithms is employed then the centre of mass momentum must be periodically removed from the simulation to avoid a transfer of energy to the centre of mass, which leads to a cooling of the atoms while the total system accelerates—an effect known as the "flying ice cube";[167] this is

not a problem for constant energy simulations as the Verlet algorithm has very good momentum conservation properties.

As described in Ref. 129, the weak coupling algorithm[166] mimics the coupling to a heat bath with a given temperature T_0. The effect is a slow relaxation of the system temperature T to the target temperature T_0 according to

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}.$$

Initial deviations will exponentially decay with constant τ. The algorithm changes the particle velocities at each step by a factor λ, which is calculated as

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T}\left(\frac{T_0}{T(t - \frac{1}{2}\Delta t)} - 1\right)} \qquad \text{where} \quad \tau = \tau_T \frac{2C_V}{k\nu}, \tag{2.86}$$

C_V being the system's heat capacity at constant volume. The use of the leap-frog algorithm Eq. (2.77) limits the temperature to the one at the last half-step. In gromacs τ_T and T_0 can be prescribed. τ/τ_0 varies between 1 (ideal gas $C_V = \frac{3}{2}Nk$) and about 3 for water. In all simulations reported here, a long decay time of $\tau_T = 0.1$ ps was chosen as to only minimally influence the dynamics of the system (for a step size $\Delta t = 2\,\text{fs} = 2 \times 10^{-3}$ ps).

Pressure coupling The instantaneous pressure $p(t)$ is computed as the trace

$$p = \tfrac{1}{3}\,\mathrm{Tr}\,\mathsf{p} \tag{2.87}$$

of the pressure tensor

$$\mathsf{p} = \frac{2}{V}\left(\frac{1}{2}\sum_{i=1}^{N} m_i \mathbf{v}_i \otimes \mathbf{v}_i - \Xi\right) \tag{2.88}$$

from the virial tensor

$$\Xi = \frac{1}{2}\sum_{i<j} \mathbf{x}_{ij} \otimes \mathbf{F}_{ij}. \tag{2.89}$$

Simulations at constant pressure were also carried out with the help of a weak coupling scheme,[166] which rescales the particle coordinates $\mathbf{x}_i$ and simulation cell vectors at each time step with a scaling matrix

$$\mu_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau_p}\beta_{ij}\left(p_{0,ij} - p_{ij}(t)\right). \tag{2.90}$$

τ_p is the coupling time constant and β is an estimate for the isothermal compressibility of the system; it suffices to choose a value applicable to liquids like the one for water, $\beta = 4.6 \times 10^{-5}$ bar^{-1}. This scheme also leads to a exponentially damped relaxation

$$\frac{d\mathsf{p}}{dt} = \frac{\mathsf{p}_0 - \mathsf{p}}{\tau_p}.$$

All simulations were carried out at $p_0 = 1$ bar with a slow relaxation time of $\tau_p = 1$ ps. Pressure coupling was only applied in the z-direction, parallel to the pore axis by setting all $\beta_{ij} = 0$ except for $\beta_{zz} = 4.5 \times 10^{-5}$ bar^{-1}. This kept the "membrane" in the xy-plane of the simulation cell at constant size while the simulation cell could adjust in the direction perpendicular to the membrane.*

2.3 Free energy calculations

Free energies such as the Helmholtz free energy, the Gibbs function, or the grand potential are important quantities in the description of thermodynamic systems in equilibrium because—depending on the boundary conditions—one of them takes on a minimum value in equilibrium. A system displaced from equilibrium by a change of constraints will rearrange in such a way that its free energy becomes minimal again while satisfying the constraints. By comparing the constrained free energy $\mathcal{W}(\mathcal{A} = a)$ (introduced in Section 2.1.3) for different values a and a' of the constraint $\mathcal{A}$ one obtains the reversible work required for the transition from state a to state a'. For instance, a could be the state "a particular Na^+ ion is in the bulk" and a' "the ion is in the channel pore." If $\mathcal{W}(a') > \mathcal{W}(a)$ then the presence of ion in the pore would be disfavoured by the Boltzmann factor $\exp[-\beta(\mathcal{W}(a') - \mathcal{W}(a))]$, and the system is more likely to be found with no ions in the pore. The free energy is grounded in a statistical mechanics description of the system and contains both enthalpic and entropic effects, unlike simple calculations of interaction energies.

2.3.1 The potential of mean force

A particularly useful constrained free energy is obtained when the constraint is taken to be the position $\mathbf{x}_1$ of a tagged particle 1 such as a water molecule or an ion. The resulting *potential of mean force*[168] (PMF) $\mathcal{W}(\mathbf{x}_1)$ is the free energy required to position the particle at the position $\mathbf{x}_1$, relative to some reference

*The "membrane" was actually mimicked by an Einstein crystal of hydrophobic methane molecules (see Section 3.2.1 on page 60 for details), which lent itself naturally to a simulation at fixed area.

state. The PMF itself is the most fundamental quantity for the description of permeation and can be used as an input to more coarse grained methods to compute *I-V* curves or single channel conductances from first principles.[169] The PMF can be written as[169]

$$e^{-\beta \mathcal{W}(\mathbf{x}_1)} = C \int d^3x_2 \cdots d^3x_N \, e^{-\beta U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)}, \tag{2.91}$$

where all coordinates are integrated out except for the coordinate of the particle of interest; C is a constant fixed by the reference state. The name "potential of mean force" derives from the fact that the gradient $-\boldsymbol{\nabla}_1 \mathcal{W}(\mathbf{x}_1)$ is the effective or mean force that is exerted on particle 1 by all other particles 2 to N in the system.*

For transport processes it is convenient to define a reaction coordinate ξ which parameterizes the path $\mathbf{x}(\xi)$ that a particle of interest moves along. Then the PMF along ξ becomes a one dimensional (1D) free energy profile

$$e^{-\beta[\mathcal{W}(\xi) - \mathcal{W}(\xi_0)]} = C \frac{\int d^3x_1 \cdots d^3x_N \, \delta\big(\mathbf{x}(\xi) - \mathbf{x}_1\big) \, e^{-\beta U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)}}{\int d^3x_1 \cdots d^3x_N \, \delta\big(\mathbf{x}(\xi_0) - \mathbf{x}_1\big) \, e^{-\beta U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)}}, \tag{2.92}$$

where ξ_0 is a reference point, and C is again arbitrary. In the case of straight ion channels and pores with their pore axis parallel to the z-axis, ξ is generally taken to be the z coordinate of the particle.[12] Such a 1D free energy profile is used as an input to simple kinetic rate theory and 1D Nernst-Planck models.[12,95,169] The reduction of the dimensionality of the PMF is only justified if z is a good reaction coordinate, i.e. all other degrees of freedom reach equilibrium rapidly whilst z is only slowly varying;[95,169] in this case the PMF can be written as

$$e^{-\beta \mathcal{W}(z)} := e^{-\beta[\mathcal{W}(0,0,z_0) - \mu_0]} \, \frac{\int dx \int dy \, e^{-\beta \mathcal{W}(x,y,z)}}{\int dx \int dy \, e^{-\beta \mathcal{W}(x,y,z_0)}} \tag{2.93}$$

where the normalisation was chosen so that the PMF vanishes in the bulk far from the pore (μ_0 being the excess chemical potential of the sampled species in the bulk). Eq. (2.93) shows that $\mathcal{W}(z)$ is only meaningfully defined in the pore region where the integration over x and y is bounded. Far from the channel, the 1D PMF is ill defined.[169,170] The 3D PMF $\mathcal{W}(\mathbf{x}) \equiv \mathcal{W}(\mathbf{x}_1)$ (2.91) does not suffer from this shortcoming and can be used to relate the 1D PMF in the pore to the constrained free energy in the bulk. The constant C in Eq. (2.92) is arbitrary, and hence there is no unique definition of $\mathcal{W}(z)$ (which allows the choice made in (2.93)) and only the variations within the PMF along z are meaningful; the 3D PMF is required to make statements about all processes which involve a

*Ref. 169 defines the PMF in a more rigorous and general way but for the present purpose the definition of the one-particle PMF $\mathcal{W}(\mathbf{x})$ should suffice.

particle in the bulk and in the channel, e.g. binding energies in the pore or barrier heights.

The PMF is related to the equilibrium density* $n(\mathbf{x})$ through the density of the reference state in the bulk far from the pore n_0,

$$\beta\mathcal{W}(\mathbf{x}) = -\ln\frac{n(\mathbf{x})}{n_0} + \beta(\mathcal{W}_0 - \mu_0), \tag{2.94}$$

because $n(\mathbf{x})/n_0\, d^3x$ is the probability to find the particle at position $\mathbf{x}$. Similarly, the 1D PMF is

$$\beta\mathcal{W}(z) = -\ln\frac{n(z)}{n_0} + \beta(\mathcal{W}_0 - \mu_0) \quad \text{with} \tag{2.95}$$

$$n(z) = \frac{1}{A}\int dx \int dy\, n(\mathbf{x}), \tag{2.96}$$

where A is the area in the xy-plane which is integrated over.

2.3.2 Umbrella sampling

From equilibrium MD trajectories the density $n(\mathbf{x})$ can be determined, and with the help of Eq. (2.94) also the PMF (up to the constant term). However, in most circumstances it is not possible to reliably estimate the density from simulations because a region of large PMF corresponds to a region of low occupancy. Furthermore, particles might be unable to cross a barrier (a local maximum in $\mathcal{W}(\xi)$) to explore regions of space behind the barrier.† The *umbrella sampling* method of Valleau and Torrie[172] (see also Ref. 173) overcomes the problem of insufficient sampling for certain regions on the reaction coordinate ξ by introducing an additional biasing potential $w(\xi)$ which forces the sampled particle into the high energy regions.‡ The biased simulations are generated using the

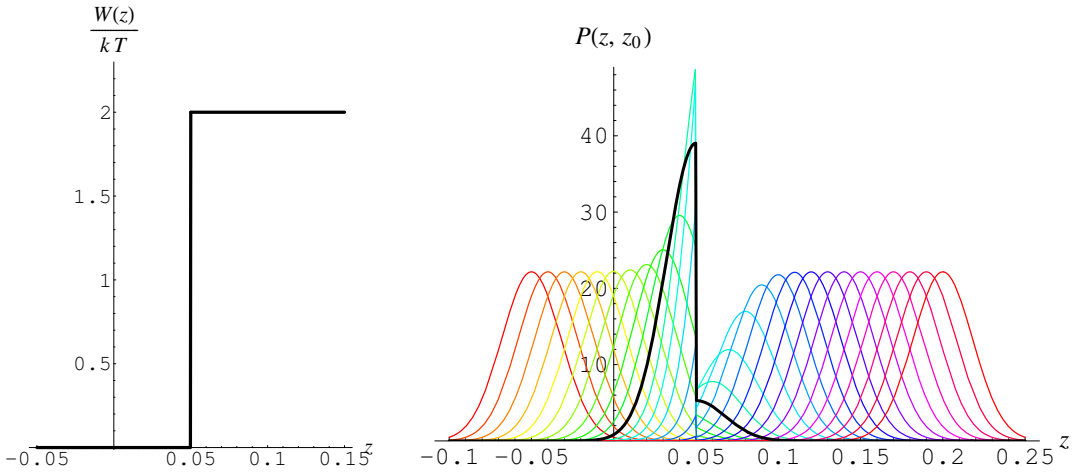
*From here onwards the symbol $n(\mathbf{x})$ should be understood as the expectation value of the density operator Eq. (2.16), i.e. as $\langle n(\mathbf{x})\rangle$.

†Example to illustrate the difficulty to sample a barrier region in equilibrium MD (following Smith and Sansom[171]): The rate of barrier crossing k is assumed to follow an Arrhenius law

$$k = v\exp(-\beta G^{\ddagger}).$$

The activation energy $G^{\ddagger}$ is taken as the barrier height, for instance $G^{\ddagger} = 40\ kT$ (see Chapter 5 for an example). With an estimate of the jump attempt frequency of a Na^+ ion across the barrier as $v = 1\ \mathrm{ps}^{-1}$, we arrive at a mean time between successful jumps of $\tau \approx 2.4\times10^5$ s $= 65$ h. A typical MD simulation of the system runs at 1 ns simulated time per 10 h real wall-clock time. To observe a *single* successful barrier crossing one would need to simulate on average for 2.4×10^{14} h $= 27\,\mathrm{Gyr} = 1.9\,t_0$ where $t_0 = 14.1\pm0.9\times10^9$ yr is the age of the universe.[109]

‡The further discussion will generally follow Roux[174], which is more accessible than the original publication.[172]



(a) Step potential (b) Probability densities

Figure 2.2: Probability density for a harmonically restrained particle at a step barrier. (a) The PMF has the shape of a step function of height 2 kT, centred on $z = 0.05$, $\beta\mathcal{W}(z) = 2\Theta(z - 0.05)$. (b) A test particle in this PMF is restrained to a window centre z_0 with a harmonic force. The probability of the particle to be in an interval $[z, z + dz]$ can be calculated analytically (see Appendix D.3.2) and plotted for different z_0. The probability distribution at the discontinuity $z_0 = 0.05$ is highlighted.

new potential energy $U(\{\mathbf{x}_j\}) + w(\xi)$. Typically, a number of simulations with different biasing potentials are carried out, each one confining the variations in the position of the particle to a particular region or "window" i on ξ (whence the method derived the name umbrella sampling). A common choice for the window potentials w_i is a harmonic one,

$$w_i(\xi) = \tfrac{1}{2} K (\xi - \xi_i)^2, \tag{2.97}$$

which restrains the particle to the vicinity of the window centre ξ_i. The w_i are chosen so to overlap and cover the whole range of interest on the reaction coordinate. In practice, an equilibrium simulation is run for each window i and the probability distribution of ξ is recorded during the run. The results from all windows are unbiased and combined to yield the PMF.

An example for the influence of umbrella windows on the distribution of a test particle is shown in Fig. 2.2. The PMF is a step barrier and in the regions of constant PMF the probability distribution is Gaussian, as expected for a harmonically bound particle. Close to the barrier, the Gaussian distribution is distorted by the influence of the PMF, and it is this difference from the known Gaussian distribution which allows the calculation of the underlying PMF.

The average distribution function (or density) is

$$\langle \rho(\xi) \rangle := \frac{\int d^3x_1 \cdots d^3x_N \, \delta\big(\xi(\{\mathbf{x}_j\}) - \xi\big) \, e^{-\beta U(\{\mathbf{x}_j\})}}{\int d^3x_1 \cdots d^3x_N \, e^{-\beta U(\{\mathbf{x}_j\})}}, \tag{2.98}$$

where the reaction coordinate typically only depends on a subset of all $\mathbf{x}_i$. In terms of the average distribution function, the PMF is (cf. Eq. (2.95))

$$\mathcal{W}(\xi) = -\beta^{-1} \ln \frac{\langle \rho(\xi) \rangle}{\langle \rho(\xi_0) \rangle} + \mathcal{W}(\xi_0). \tag{2.99}$$

The biased distribution around the i-th window is then (using the new potential energy of the system $U(\{\mathbf{x}_j\}) + w_i(\xi)$)

$$\langle \rho(\xi) \rangle_i = e^{-\beta w_i(\xi)} \, \langle \rho(\xi) \rangle \, \langle e^{-\beta w_i(\xi)} \rangle^{-1} \tag{2.100}$$

and hence the unbiased distribution is

$$\langle \rho(\xi) \rangle_i^{\text{unbiased}} = e^{+\beta w_i(\xi)} \, \langle \rho(\xi) \rangle_i \, \langle e^{-\beta w_i(\xi)} \rangle. \tag{2.101}$$

(where we explicitly named $\langle \rho(\xi) \rangle$ (from Eq. (2.100)) $\langle \rho(\xi) \rangle_i^{\text{unbiased}}$). The PMF for the i-th window from the unbiased distribution Eq. (2.101) becomes with the help of (2.99)

$$\mathcal{W}_i(\xi) = -\beta^{-1} \ln \frac{\langle \rho(\xi) \rangle_i}{\langle \rho(\xi_0) \rangle} - w_i(\xi) + F_i + \mathcal{W}(\xi_0), \tag{2.102}$$

$$\text{with} \quad e^{-\beta F_i} := \langle e^{-\beta w_i(\xi)} \rangle. \tag{2.103}$$

Apart from the F_i (and the essentially arbitrary $\mathcal{W}(\xi_0)$) all other terms in Eq. (2.102) can be obtained from an umbrella sampling simulation (see Fig. 2.3 for an example of $\langle \rho(\xi) \rangle_i$ from real data). The constants F_i represent the free energy required to introduce the biasing potential into the system. They are required to combine the $\mathcal{W}_i(\xi)$ into the total PMF $\mathcal{W}(\xi)$. The correct choice of F_i ensures a continuous PMF across the windows. A variety of methods can be employed to match up the $\mathcal{W}_i$ but the most versatile and robust one[174] is the weighted histogram analysis method.

2.3.3 The Weighted Histogram Analysis Method

The weighted histogram analysis method[175] (WHAM) provides an estimate for the F_i in Eq. (2.103) using *all* data in the N_w biased distribution functions

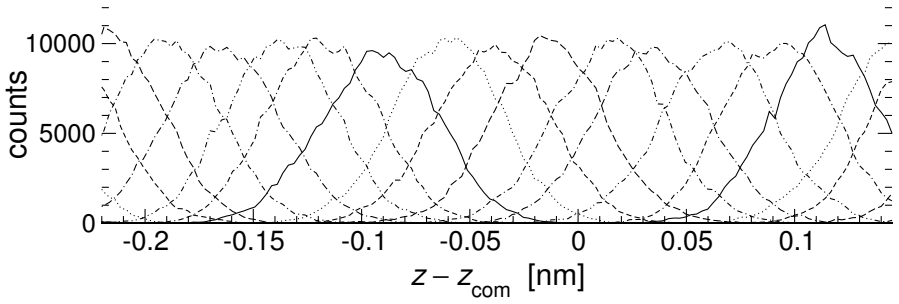


Figure 2.3: Raw umbrella sampling histograms. The example shows how often a Na^+ ion in the nAChR pore, harmonically restrained to z_0, visited a region between z and $z + \Delta z$ (see Chapter 6). For the figure, the bin width Δz was chosen as 2×10^{-3} nm but in general a resolution of $\Delta z = 8 \times 10^{-3}$ nm is sufficient in the WHAM procedure.

$\langle \rho(\xi) \rangle_i$.* Kumar et al.[175] derive an optimal estimate for the total unbiased distribution as a ξ-dependent weighted sum of the N_w unbiased distribution functions

$$\langle \rho(\xi) \rangle = \sum_{i=1}^{N_w} \langle \rho(\xi) \rangle_i^{\text{unbiased}} \; \frac{n_i \, e^{-\beta(w_i(\xi) - F_i)}}{\sum_{j=1}^{N_w} n_j \, e^{-\beta(w_j(\xi) - F_j)}} \tag{2.104}$$

by minimising the statistical error in an auxiliary function (the density of states); n_i is the number of data points used in constructing $\langle \rho(\xi) \rangle_i$. Using Eq. (2.101) and (2.103), the unbiased distribution is

$$\langle \rho(\xi) \rangle = \sum_{i=1}^{N_w} e^{+\beta w_i(\xi)} \; \langle \rho(\xi) \rangle_i \; e^{-\beta F_i} \; \frac{n_i \, e^{-\beta(w_i(\xi) - F_i)}}{\sum_{j=1}^{N_w} n_j \, e^{-\beta(w_j(\xi) - F_j)}},$$

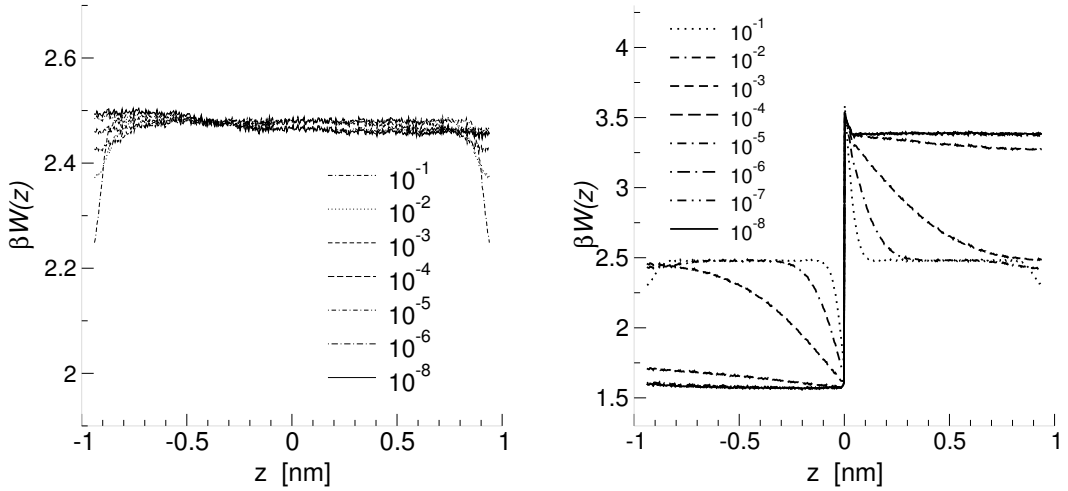
which implies the two WHAM equations

$$\langle \rho(\xi) \rangle = \frac{\sum_{i=1}^{N_w} n_i \, \langle \rho(\xi) \rangle_i}{\sum_{j=1}^{N_w} n_j \, e^{-\beta(w_j(\xi) - F_j)}} \quad \text{and} \tag{2.105}$$

$$e^{-\beta F_i} = \int d\xi \, e^{-\beta w_i(\xi)} \; \langle \rho(\xi) \rangle \, . \tag{2.106}$$

These two equations are solved self-consistently for the F_i. In practice this is done iteratively starting from a guess for the F_i until the maximum change $\max_i |F_i - F_i^{\text{last}}|$ between iterations is smaller than a preset tolerance.

*The description given here continues to follow Roux[174]. For a discussion of other methods than WHAM see also Roux[174].



(a) Flat constant PMF, generated from $\beta W(z) = 0$

(b) Step function of height 2 kT PMF, generated from $\beta W(z) = 2\Theta(z)$

Figure 2.4: Test cases for the WHAM procedure and check of convergence. Input for two underlying PMFs was simulated and WHAM was used to recover the PMFs. The recovered profiles are plotted for decreasing tolerances (in kT) of the self-consistent WHAM algorithm. The PMFs from WHAM are only determined up to a constant.

2.3.4 Umbrella sampling and WHAM in practice

The practical problems of applying umbrella sampling and WHAM to obtain a free energy profile in channels are briefly discussed. This includes tests of the WHAM procedure and a protocol to obtain $\mathcal{W}(z)$ from a combination of equilibrium and umbrella sampled PMFs.

Test input The WHAM procedure is implemented in the program g_wham (see Appendix D.2.5 for details). The implementation was tested by simulating input as if obtained by umbrella sampling a given PMF $\mathcal{W}(z)$ (using fakepmf, see Appendix D.3.2) and processing it with g_wham. Results are shown in Fig. 2.4 for increasing tolerances (in kT). A constant PMF poses no problem although a very small drift is visible. A step function of height 2 kT is resolved provided the tolerance is taken to be at least 10^{-5} kT. The step discontinuity itself has the correct height of 2 kT but the adjacent constant stretch of potential is about 0.2 kT to low. Given the extreme case of a discontinuity, WHAM was judged to work satisfactory; for processing of MD umbrella sampling data the tolerance was chosen to be 10^{-5} or smaller.*

*WHAM is carried out in double precision (`double` variables) so the required tolerance is not related to the loss of precision when only using `float` variables.

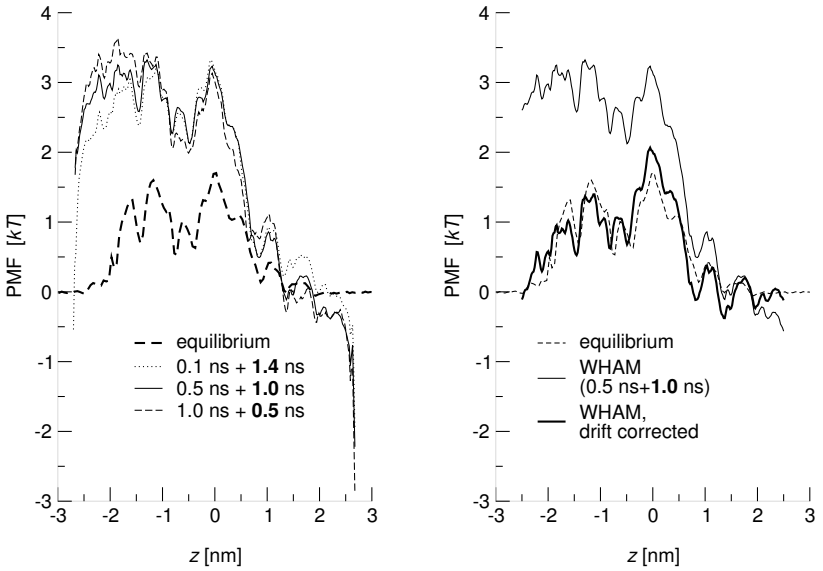
Matching umbrella and equilibrium PMF As pointed out before, the 1D PMF is not well defined outside the pore region. In particular, it is not possible to umbrella sample along z without confinement in x and y. One possibility is to add an artificial confining potential[170] (and see Appendix D.3.1) whose influence can later be removed from the PMF.* Here a less rigorous and conceptually simpler method was employed. The equilibrium density is always well sampled in the bulk region and in the entrance ("mouth") regions of the pore, which allows to define a PMF that is accurate in bulk and mouth but is typically inaccurate in the pore itself. Umbrella sampling, on the other hand, yields accurate results in the mouth and the pore region, where lateral particle movements are bounded, but not in the bulk. Both PMFs should be accurate in the mouth region and as they describe the same system one can match them to obtain a continuous PMF. This operation was carried out manually; errors associated with this procedure are estimated to be smaller than 0.5 kT, sometimes less than 0.1 kT if the match is very good. The matching of the equilibrium PMF with the umbrella PMF also ensures that absolute barrier heights (relative to the bulk) can be obtained.

Equilibration and drift correction Umbrella sampling in complex pores such as the nAChR pore (Chapter 6) is complicated by the fact that—unlike in simple model pores—it is by no means clear what the PMF should look like. However, if the density is sufficiently sampled during an equilibrium trajectory then Eq. (2.95) yields the PMF. A PMF obtained with umbrella sampling and WHAM can then compared to the "true" PMF. For ions the barriers are too high in the closed nAChR pore but water permeates rather freely, which is expressed in barriers less than 1.5 kT high (Fig. 2.5(a)). A water molecule was umbrella sampled in 101 windows for 1.5 ns in each window (see Section 6.2.4 and Tab. 6.2 on page 133 for details). An initial part of length $\mathfrak{T}_{\text{eq}}$ of the umbrella sampled time series $\big(t, z(t) - z_i(t)\big)$ is discarded as equilibration time, with $\mathfrak{T}$ remaining as production data. As Fig. 2.5(a) shows, the influence of the equilibration phase is moderate, and $\mathfrak{T}_{\text{eq}} = 0.5$ ns appears to suffice for a converged PMF.

A more serious problem is a systematic "drift" or linear shift in the umbrella PMF compared to the equilibrium PMF (which was observed in many but not all other PMFs that were analysed in this work). For symmetric problems such as the model pores, very good agreement between the PMFs was achieved by symmetrizing the PMF (see Fig. 5.4 on page 115 and Section 5.2.2). This is not possible for asymmetric pores so it was decided to apply a "drift correction". A function

$$\Delta_{\text{drift}}(z) = \lambda z \tag{2.107}$$

*Confining potentials were considered but tests did not show a significant advantage over the method that was finally used.



(a) Influence of equilibration time and sampling time, $\mathfrak{T}_{\text{eq}}+\mathfrak{T}$.

(b) PMF, drift corrected by $W(z) \to W(z) + \Delta_{\text{drift}}(z)$.

Figure 2.5: Equilibration time and drift correction of the PMF for water in nAChR. (a) The influence of the equilibration phase on the PMF is comparatively small. (b) Good agreement with the PMF from the equilibrium density is only obtained after adding a correction $\Delta_{\text{drift}}(z) = 0.614\ kT\ \text{nm}^{-1} \times z$ for "drift" in the WHAM procedure.

with the adjustable parameter λ is added to $\mathcal{W}(z)$ so that the corrected umbrella PMF $\mathcal{W}(z) + \Delta_{\text{drift}}(z)$ can be matched to the equilibrium PMF. Although similar "hysteresis" effects have been reported[176] and attributed to correlations in subsequent starting configurations, this is not likely to be the case here as the starting positions in adjacent windows tend to be from different times of the equilibrium trajectory and are therefore uncorrelated. It probably rather reflects the fact that the WHAM procedure integrates histogram data drom the whole range and can be somewhat sensitive to incomplete sampling in the outermost bins. In principle the PMF could be constrained to be the same at either end of the range because no potential difference occurs. In practice, this systematic error can be corrected for by comparison (or calibration) to the equilibrium PMF as was done here.

2.4 Poisson-Boltzmann electrostatics

Molecular dynamics simulations are very CPU intensive. An estimate for binding energies or the PMF for an ion in a pore can also be obtained from a static

structure. In this approach, interactions are reduced to electrostatic contributions, i.e. only the static or "fixed" charge distribution of a protein

$$\rho_{\mathrm{f}}(\mathbf{x}) = \sum_i q_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{2.108}$$

interacts with another protein or an ion. Dynamic and hence entropic effects are not included. The solvent is treated as a continuum so the size of the problem is reduced considerably, and because averaging effects can be incorporated, such as the rapid re-orientation of the strong water dipoles, which is responsible for the high dielectric constant of water. Water is treated as a continuous medium with dielectric constant $\epsilon_{\mathrm{w}} \approx 78.5$ whereas proteins are given $2 \leq \epsilon_{\mathrm{p}} \leq 16$. Mobile ions in solution are treated in Debye-Hückel theory[177] as continuous distributions

$$\begin{aligned} \rho_+(\mathbf{x}) &= +z_+ e\, n_+(\infty)\, e^{-\beta z_+ e \phi(\mathbf{x})} \\ \rho_-(\mathbf{x}) &= -z_- e\, n_-(\infty)\, e^{+\beta z_- e \phi(\mathbf{x})}, \end{aligned} \tag{2.109}$$

where $z_\pm$ is the valence and $n_\pm(\infty)$ the bulk concentration of the ions. The density depends on the electric potential $\phi(\mathbf{x})$ through the Boltzmann factor; $\phi(\mathbf{x})$ is generated by the fixed charges (2.108) and by the mobile charges (2.109) themselves. Inserting the total charge $\rho := \rho_{\mathrm{f}} + \rho_+ + \rho_-$ into Poisson's equation

$$-\boldsymbol{\nabla} \cdot [\epsilon(\mathbf{x}) \boldsymbol{\nabla} \phi(\mathbf{x})] = \frac{\rho(\mathbf{x})}{\epsilon_0} \tag{2.110}$$

yields the *Poisson-Boltzmann equation*.[82] For a salt with symmetric charges on the ions (valence $z := z_+ = z_-$, e.g. NaCl $z = 1$) and bulk concentration $n_0 = n_\pm(\infty)$ it can be written[178,179]

$$-\boldsymbol{\nabla} \cdot [\epsilon(\mathbf{x}) \boldsymbol{\nabla} \cdot \phi(\mathbf{x})] = -\frac{2ze n_0 \chi(\mathbf{x})}{\epsilon_0} \sinh[\beta z e \phi(\mathbf{x})] + \frac{\rho_{\mathrm{f}}}{\epsilon_0}. \tag{2.111}$$

The shape function $\chi(\mathbf{x})$ is 1 in the solvent and 0 in the protein; it ensures that the effect of the mobile ions only applies to the region accessible to the ions. Similarly, $\epsilon(\mathbf{x})$ equals ϵ_{p} in the protein interior and ϵ_w in the region accessible to solvent.

For $ze\phi(\mathbf{x}) \ll kT$ the non-linear term $\sinh[\beta z e \phi(\mathbf{x})]$ can be approximated by $\beta z e \phi(\mathbf{x})$ so that Eq. (2.111) becomes the linearised Poisson-Boltzmann equation,

$$-\boldsymbol{\nabla} \cdot [\epsilon(\mathbf{x}) \boldsymbol{\nabla} \cdot \phi(\mathbf{x})] = -\tilde{\kappa}^2(\mathbf{x}) \phi(\mathbf{x}) + \frac{\rho_{\mathrm{f}}}{\epsilon_0} \qquad \text{with} \tag{2.112}$$

$$\tilde{\kappa}^2(\mathbf{x}) := \frac{2\beta e^2 n_0 z^2 \chi(\mathbf{x})}{\epsilon_0} = \frac{2\beta e^2 I \chi(\mathbf{x})}{\epsilon_0}. \tag{2.113}$$

For electrolytes of differing valence z_i, the term $n_0 z^2$ can be replaced by the ionic strength $I := \frac{1}{2}\sum_i n_i z_i^2$ of all ionic species.* In the electrolyte region where $\chi(\mathbf{x}) = 1$ and $\epsilon(\mathbf{x}) = \epsilon_w$, $\tilde{\kappa}^2$ is related to the Debye screening length λ,

$$\epsilon_w \tilde{\kappa}^{-2} = \lambda^2 = \frac{\epsilon_w \epsilon_0}{2\beta e^2 I},$$

which is the thickness of electrolyte solution that reduces the potential of a point charge by a factor of $1/e$;[82] for $I = 0.1\ \mathrm{mol}\cdot\mathrm{l}^{-1}$ NaCl, $\lambda = 0.962$ nm.

Electrostatic free energy For the linear Poisson-Boltzmann equation (2.112) the total electrostatic energy is[180]

$$\Delta G^{\mathrm{el}} = \tfrac{1}{2}\int d^3x\, \rho_f(\mathbf{x})\, \phi(\mathbf{x}); \tag{2.114}$$

it is the free energy required to charge the system (the fixed charges) in the presence of the mobile ions. The electrostatic binding energy of an explicit ion in a protein channel at subsequent positions z along the pore axis (or electrostatic PMF) can be calculated as

$$\Delta G_B(z) = \Delta G_{PI}^{\mathrm{el}}(z) - \Delta G_I^{\mathrm{el}} - \Delta G_P^{\mathrm{el}}. \tag{2.115}$$

$\Delta G_{PI}^{\mathrm{el}}$ denotes the electrostatic energy of the system comprising the pore and the ion, located at position z. ΔG_I^{el} is the electrostatic energy of an ion in the solvent. In order to minimise artifacts due to the calculation algorithm, $\Delta G_I^{\mathrm{el}}(z)$ is calculated for every position z of the ion in the channel. ΔG_P^{el} is the total electrostatic energy of the pore/slab system immersed in high dielectric solvent. The explicit ion is modelled as a spherical cavity of the same dielectric as the protein† with a radius (Born radius) that reproduces experimental solvation energies.[181]

Poisson-Boltzmann solver In order to solve the Poisson-Boltzmann equation (2.111) or (2.112) for arbitrary shapes $\chi(\mathbf{x})$ and $\epsilon(\mathbf{x})$ the problem is discretized on a grid. For the Born profile calculations in Chapter 5 and 6, the program apbs[179] was used, which uses a finite elements algorithm for discretization.[182,183] For the calculation of the protonation states, i.e. the pK_a s,[184,185] of the residues in the nicotinic acetylcholine receptor transmembrane domain (Chapter 6), DelPhi[186] was used, which employs a finite difference scheme.

*If the ionic strength is given in $\mathrm{mol}\cdot\mathrm{l}^{-1}$ instead of m^{-3} then it needs to be converted to m^{-3}: $I/(\mathrm{mol}\cdot\mathrm{l}^{-1})\cdot 10^3\, N_A = I/\mathrm{m}^{-3}$ with Avogadro's constant $N_A = 6.02214179\times 10^{23}\ \mathrm{mol}^{-1}$.

†For the ion itself a dielectric of $\epsilon = 1$ would be more appropriate but the program apbs is not yet capable of calculating multiple low-dielectric regions.

3 The surprising behaviour of confined water

Water confined to short ($L = 0.8$ nm) hydrophobic pores (radius 0.35 nm $\leq R \leq$ 1.0 nm) is studied by atomistic molecular dynamics simulations. The water density in the pore fluctuates on a nanosecond time scale between a liquid and vapour filled pore. These liquid-vapour oscillations can be explained as capillary evaporation alternating with capillary condensation, driven by local pressure fluctuations in the water outside the pore. The free energy landscape shows how a metastable liquid state gradually develops with increasing radius. For radii larger than ca. 0.55 nm the liquid becomes the globally stable state and the vapour state vanishes. One dimensional confinement affects the dynamic behaviour of the water molecules and increases the self diffusion by a factor of two to three compared to bulk water. Permeabilities for the narrow pores are of the same order of magnitude as for biological water pores. Water flow is not continuous but occurs in bursts.

Channel and transporter proteins control flow of water, ions and other solutes across cell membranes. In recent years several channel and pore structures have been solved at near atomic resolution[16,30,51,55,69,71] which together with three decades of physiological data[12] and theoretical and simulation approaches[95] allow us to describe transport of ions, water or other small molecules at a molecular level. Water plays a special role here: it either solvates the inner surfaces of the pore and the permeators (for example, ions and small molecules like glycerol), or it is the permeant species itself as in the aquaporin family of water pores[13,187,188] or in the bacterial peptide channel gramicidin A (gA), whose water transport properties are well studied.[189–191] Thus, a better characterisation of the behaviour of water would improve our understanding of the biological function of a wide range of transporters. The remarkable water transport properties of aquaporins—water is conducted through a long (ca. 2 nm) and narrow (ca. 0.3 nm diameter) pore at bulk diffusion rates while at the same time protons are strongly selected against—are the topic of recent simulation studies.[72,192]

The shape and dimensions of biological pores and the nature of the pore lining atoms are recognised as major determinants of function. How the behaviour of water depends on these factors is far from understood.[9] Water is not a

simple liquid due to its strong hydrogen bond network. When confined to narrow geometries like slits or pores it displays an even more diverse behaviour than already shown in its bulk state.[193,194]

3.1 Water in pores

A biological channel can be crudely approximated as a "hole" through a membrane. Earlier molecular dynamics (MD) simulations showed pronounced layering effects and a marked decrease in water self diffusion in infinite hydrophobic pore models.[195,196] Hummer et al.[197], Beckstein et al.[62], and Allen et al.[198] observed that water in finite narrow hydrophobic pores exhibited a distinct two-state behaviour. The cavity was either filled with water at approximately bulk density (liquid-like) or it was almost completely empty (vapour-like). Similar behaviour was seen in Gibbs ensemble Monte Carlo simulations (GEMC) in spherical[1] and cylindrical pores.[199]

In our previous simulations[62*] we explored model pores of the dimensions of the gating region of the nicotinic acetylcholine receptor nAChR.[59†] Hydrophobic pores of radius $R \geq 0.7$ nm were filled during the whole simulation time (up to 6 ns) whereas narrow ones ($R \leq 0.4$ nm) were permanently empty. Changing the pore lining from a hydrophobic surface to a more hydrophilic (polar) one rendered even narrow pores water—and presumably ion—conducting. At intermediate radii (0.4 nm $< R <$ 0.7 nm) the pore-water system was very sensitive to small changes in radius or character of the pore lining. In a biological context, a structure close to the transition radius would confer the highest susceptibility to small conformational rearrangements (i.e. gating) of a channel.

The previous short-time scale simulations[62] gave a tantalising hint at interesting emerging phenomena at longer time scales. In this chapter considerably longer simulations are presented ($>$ 50 ns instead of 1 ns) in order to explore the longer timescale behaviour of the water-pore system. Starting from the observed oscillations in water density (Fig. 3.1 on the facing page) we analyse the kinetics and the free energy of the system. We compare the water transport properties of the pores to experimental and theoretical data. First, however, we look at the behaviour of water in the simplest inhomogeneous system, namely at a planar hydrophobic surface. So we are able to pinpoint the effect of the

*The simulations reported in Beckstein et al.[62] were comparatively short (typically, only 1 ns per simulation) and although the main conclusions are still valid, more detailed and extensive simulations are the subject of this and the following chapter. Therefore, the results from our previous paper are not explicitly shown here; the conclusions arising from the material presented here recapitulate and extend the results of the previous paper.

†Note that at the time (2001/2002) the atomic resolution structure of the nAChR transmembrane domain[29] was not known yet.

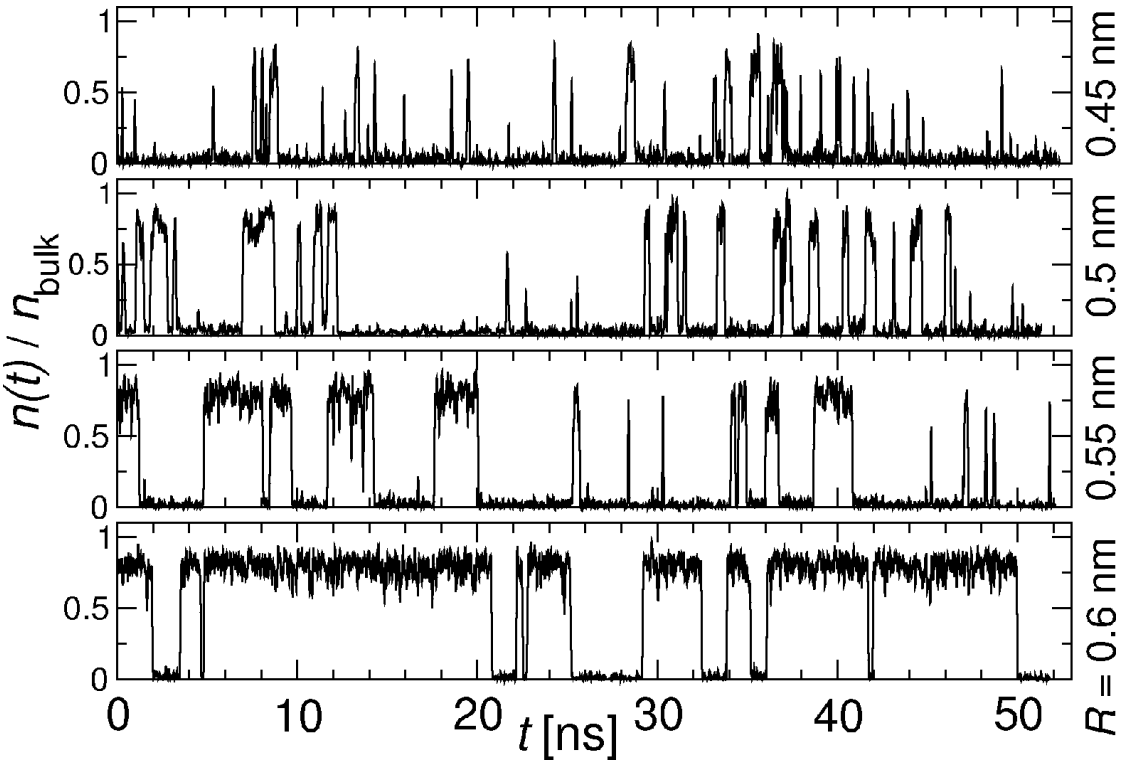


Figure 3.1: Oscillating water density in model pores of increasing pore radius R. The water density $n(t)$ (in units of the bulk water density n_0) over the simulation time shows strong fluctuations on a greater than ns time scale (50 ps moving average smoothing). Two distinctive states are visible: *open* at approximately n_0 (liquid water), and *closed* with very few or no water in the pore (water vapour).

higher degree of confinement conferred by the cylindrical pores compared to the less confining planar surface whilst using the bulk system ("neat water") as a reference for both. In this sense, a planar surface can be seen as the limiting case of a pore with an infinite radius.

3.2 Methods

For completeness, some of the methodological aspects, that were already mentioned in Chapter 2, are repeated here, together with the parameters of the MD simulations that were used for the calculations of this chapter.

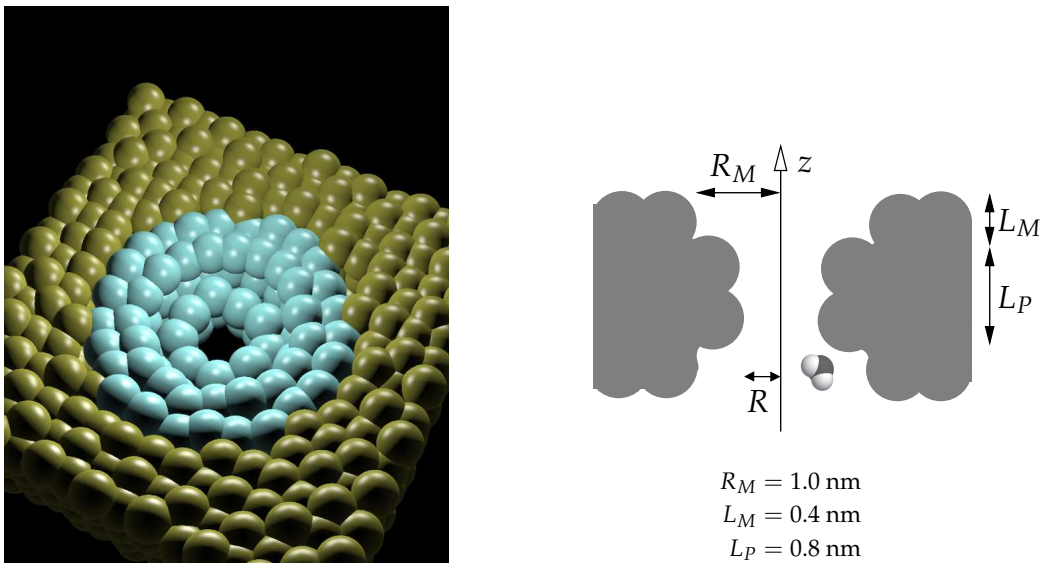
3.2.1 Model

The system is designed to capture the minimal characteristics of a protein channel embedded in a membrane: A planar slab of material (the membrane mimetic) introduces two surfaces, and a cylindrical conduit ("hole")—the pore—runs through the "membrane". Even though both are described separately, they very much form a unit, which provides the confining environment for water molecules to move in.

Membrane mimetic slab A membrane mimicking slab was constructed as a cubic crystal of methane molecules with lattice constant 0.39 nm. In fact, this construct is not a real crystal because it is not held together by intermolecular forces between the molecules. Instead, harmonic forces restrain the individual molecules to the lattice positions, forming an Einstein crystal.[200*] Its purpose is to create a defined simple surface with only a few parameters to vary as we assume that the major effect on the water behaviour is primarily due to confinement and only secondly due to the internal details of the surface representation. Methane (CH_4) is used as the surface material because it approximates the chemical characteristics of the methyl groups (–CH_3) of hydrophobic side chains in proteins (Val, Ile, Leu, Ala, Met) which are, as mentioned in the introduction, implicated in playing a role in the gating of some ion channels.

Hydrophobic pores The pore model was designed to mimic the dimensions of a biological pore (e.g. the gate region of nAChR[59]), whilst retaining the tractability of a simple model. Cylindrical pores of finite length were constructed from concentric rings of pseudo-atoms (Fig. 3.2) using the program pgeom (see Appendix D.1.1). As described for the membrane mimetic, these hydrophobic pseudo-atoms have the characteristics of methane molecules (uncharged Lennard-Jones spheres with a van der Waals radius of 0.195 nm). The pore consists of two mouth regions (radius $R_M = 1$ nm, length $L_M = 0.4$ nm) and an inner pore region ($L_P = 0.8$ nm) of radius 0.35 nm $\leq R \leq$ 1.0 nm, which is the minimum van der Waals radius of the cavity. The centres of the pore lining atoms are placed on circles of radius $R + 0.195$ nm. The model was embedded in a membrane mimetic of thickness ca. 1.5 nm or 1.9 nm. Pseudo-atoms were harmonically restrained to their initial position with a force constant of 1000 kJ mol^{-1} nm^{-2}, resulting in positional fluctuations of ca. 0.1 nm, comparable to those of pore lining atoms in membrane protein simulations.[201] Of course, this setup does not allow for global collective motions of the pore which

*In fact, this is not an ideal Einstein crystal as the lattice molecules can interact with each other; however the harmonic forces are much stronger than the inter-molecular Lennard-Jones interactions at this lattice spacing so that the molecules act effectively as uncoupled harmonic oscillators.



(a) Van der Waals surface (b) Schematic

Figure 3.2: Model pore system. (a) The pore model consists of methane pseudo-atoms of van der Waals radius 0.195 nm. The pore (cyan) is embedded in a membrane mimetic slab (gold). Pore and slab consist of the same kind of methane atoms. (b) The geometry of the mouth region (length L_M and radius R_M) and the length L_P of the pore have been kept fixed throughout the calculations in this chapter but the pore radius $R \equiv R_P$ was varied. A water molecule is drawn to scale.

are necessary in real proteins for gating to occur. However, here the aim is *not* to simulate these large scale motions but to investigate the behaviour of water in a environment characteristic of either the open or the closed state. Because the time scale of gating motions is on the millisecond time scale[12] but solvent rearranges on much shorter time scales (pico- to nanoseconds) our simplifications appear to be reasonable ones.

3.2.2 Simulation Details

MD simulations were performed with gromacs version 3.0.5 (Ref. 127) and the SPC water model.[132] Parameters for the unified methane atoms are described in Section 2.2.3 on page 38. The integration time step was 2 fs and coordinates were saved every 2 ps. With periodic boundary conditions, long range electrostatic interactions were computed with a particle mesh Ewald method (real space cutoff 1 nm, grid spacing 0.15 nm, 4th order interpolation[164,165]) while the short ranged van der Waals forces were calculated within a radius of 1 nm. The neighbour list (radius 1 nm) was updated every 10 steps. Weak coupling

algorithms[166] were used to simulate at constant temperature ($T = 300$ K, time constant 0.1 ps) and pressure ($p_z = 1$ bar, compressibility 4.5×10^{-5} bar^{-1}, time constant 1 ps) where the box dimensions in the plane of the membrane were held fixed.

Membrane mimetic slab A box (size 6 nm × 6 nm × 4.6 nm) with a membrane mimetic of thickness 1.2 nm and altogether 3 nm of SPC water was simulated for 20 ns. The system contained 675 methane molecules and 3910 water molecules.

Hydrophobic pores For the simulations of water in the hydrophobic pore models, the pore/slab system was solvated with SPC water, forming a simulation cell with lengths in x and y of 6 nm (or 3.9 nm for the 80 ns simulation of the $R = 0.35$ nm pore). The length in z was 4.6 nm in both cases (ensuring bulk-like water behaviour far from the membrane mimetic). The large (small) system contained about 700 (300) methane pseudo-atoms and 4000 (1500) SPC water molecules. Simulation times $\mathfrak{T}$ ranged from 52 ns to 80 ns (altogether 460 ns).

Bulk water Bulk properties of SPC water were obtained from simulations of neat water (i.e. just water) in a cubic cell of length 3 nm (895 molecules) with isotropic pressure coupling at 300 K and 1 bar for 5 ns.

3.2.3 Analysis

To a certain degree, analysis of simulations is similar to analysis of experimental data though it differs in that one actually *has* detailed knowledge of all particles involved because the positions (and sometimes velocities) are recorded in the trajectory over time. But this level of detail is overwhelming and does not lend itself to facilitating a deeper understanding of the simulated system. So the task is to define and "measure" (i.e. compute) meaningful observables in order to make sense of the intricate dance that is performed by the atoms.

Time courses and density For the density time courses (Fig. 3.1 on page 59) the pore occupancy $N(t)$, i.e. the number of water molecules in the pore cavity (a cylinder of height $L_P = 0.8$ nm containing the pore lining atoms) was counted with g_count (Appendix D.2.1). The density $n(t)$ is given by $N(t)$ divided by the water-accessible pore volume

$$V = L_P \, \pi R_{\text{eff}}^2 \tag{3.1}$$

and normalised to the bulk density of SPC water at 300 K and 1 bar ($n_0 = 53.67 \pm 0.03$ mol $\cdot$ l^{-1}).* The effective pore radius for all pores is $R_{\text{eff}} = R - \delta R$. Choosing $\delta R = 0.03$ nm fixes the density $\langle N \rangle / V$ in the (most bulk-like) $R = 1.0$ nm-pore at the value calculated from the radial density, $R_0^{-1} \int_0^{R_0} n(r)\, dr$, where $R_0 = 1.05$ nm is the radius at which $n(r)$ vanishes (see Section 2.1.4 on page 28).

The density $n(\mathbf{x})$ was determined on a grid of cubic cells with spacing 0.05 nm, using g_ri3Dc (Appendix D.2.3). Two- and one-dimensional densities were computed by integrating out the appropriate coordinate(s) with a_ri3Dc. Because an equilibrium MD simulation samples states according to the Boltzmann factor, the potential of mean force (PMF) of a water molecule is related to the water density $n(\mathbf{x})$ by

$$\beta \mathcal{W}(\mathbf{x}) = -\ln \frac{n(\mathbf{x})}{n_0} + \beta(\mathcal{W}_0 - \mu_0) \tag{2.94}$$

(see also Section 2.3.1 on page 46). $\mathcal{W}_0 - \mu_0$ is not determined in our simulations, and because it is generally different for different systems (e.g. pores of different radii) we normally cannot compare the position of different $\mathcal{W}(\mathbf{x}; R)$ relative to each other.† We can identify the PMF $\mathcal{W}(\mathbf{x})$ of Eq. (2.94) with the constrained grand potential $\Omega(\mathbf{x})$, as detailed in the next section.

Free energy density We consider the free energy of the fluid in the pore. To this end we imagine a subsystem of the full simulation box which contains the pore and a buffer region. Because the full system is simulated at constant pressure (and temperature), the water density and thus the chemical potential near the pore entrance is essentially fixed. Thus the fluid in the pore is most conveniently treated in the grand canonical ensemble[202] (fixed temperature, volume, and chemical potential, see Section 2.1 on page 17).‡ Hence the free energy is the grand potential. The number of particles N in the subsystem fluctuates, but the average is fixed by μ. In order to distinguish different phases of the fluid in the pore we introduce the pore occupancy N^*, the number of particles in the pore of volume $V = \pi R^2 L$. Then the density in the pore region $n^* := N^*/V$

*The symbol n is used as a density in a fairly wide sense, and interchangeably denotes number density or concentration, or mass density; for SPC water $n_0 = 0.997 \times 10^3$ kg $\cdot$ m^{-3}.

†The situation is analogous to the case of umbrella sampling (cf. Section 2.3.2) where a procedure has to be found to join neighbouring histograms. WHAM (Section 2.3.3) precisely determines the relative free energy offsets by a histogram reweighting technique.

‡In Ref. 96 we took the free energy to be the Helmholtz free energy but it was subsequently pointed out by Allen et al.[202] that the most appropriate thermodynamic potential is the grand potential where the number of water molecules in the pore takes the role of a reaction coordinate. Similarly, Hunter III and Reinhardt[203] use the average volume per particle $v := V/N$ as the order parameter in a constrained free energy of the isothermal-isobaric ensemble.

is a suitable *order parameter*. (n^* is the inverse of the average volume per water molecule $\overline{v} = V/N^*$ which could also be used as an order parameter.[203]) We are interested in the free energy as a function of the order parameter, i.e. the constrained free energy

$$\beta\Omega(n^*) = -\ln \Xi(n^*), \tag{3.2}$$

(cf. Section 2.1.3 on page 27) where the grand canonical partition function (2.12) is constrained to values which correspond to a pore filled with fluid at density n^*,

$$\Xi(n^*) = \operatorname*{Tr}_{n^*(\{\mathbf{x}_i\})=n^*} e^{-\beta(\mathcal{H}-\mu N)}. \tag{3.3}$$

$n^*(\{\mathbf{x}_i\})$ is the density in the pore for a particular configuration of molecules $\{\mathbf{x}_i\}$. The simulations yield the probability $p(n^*)$ to observe the subsystem in a state with fluid density n^*. Hence the constrained free energy can be calculated from the simulation data as

$$\beta\Omega(n^*) = -\ln p(n^*) + \beta\Omega. \tag{3.4}$$

and transformed into a free energy density

$$\omega(n^*) \equiv \omega(T, \mu; n^*) = \Omega(T, V, \mu; n^*)/V, \tag{3.5}$$

A sixth order polynomial in n^* was least-square fitted to $\beta\omega(n^*)$, except for the pores with radii 0.15 nm, 0.2 nm, and 1.0 nm for which a cubic regression sufficed. The $\beta\omega(n^*)$ curves obtained for different radii R from the simulations are only determined within the additive constant Ω, which is the free energy of the pore system $\beta\Omega = -\ln V \int_0^\infty dn^*\, \Xi(n^*)$.

For the free energy density Eq. (3.5) we can show that all curves coincide for $n^* = 0$, which effectively describes the vapour state. Let ω be the free energy density in the pore region and $\omega' = -p'$ the free energy density in the adjacent buffer and bulk regions. In equilibrium these densities have the same value $\omega = \omega'$, and, using the model Eq. (4.4), which will be introduced in detail in the next chapter (see page 90), this simply expresses mechanical equilibrium across the meniscus,

$$-p + V^{-1}\left(2\pi RL\gamma_{fw}(n^*) + 2\pi R^2\gamma_{fl}(n^*)\right) = -p'.$$

The surface tensions γ between the fluid f in the pore (vapour or liquid) and the wall w or the liquid l in the buffer/bulk regions vary with the density n^* of the fluid; the dependency on temperature is not explicitly shown. With the introduction of the aspect ratio of the pore $\eta = R/L$, the free energy density in the pore depends on n^* as

$$\omega(n^*) = -p' - 2L^{-1}\left(\eta^{-1}\gamma_{fw}(n^*) + \gamma_{fl}(n^*)\right). \tag{3.6}$$

We note that $\gamma_{fw}(n^* = 0)$ vanishes because without particles there is no interaction between the fluid and the wall.* (Similarly, $\gamma_{fl}(n^* = n_0)$ vanishes when there is no meniscus.) When the surface tensions are expanded in powers series in n^*,

$$\gamma_{fw}(n^*) = \sum_{k=0}^{\infty} a_k n^{*k} \quad \text{and} \quad \gamma_{fl}(n^*) = \sum_{k=0}^{\infty} b_k n^{*k},$$

then $\gamma_{fw}(n^* = 0) = 0$ implies $a_0 = 0$. Hence the power series of Eq. (3.6) reads

$$\omega(n^*) = -p' - 2L^{-1}\left[b_0 + \sum_{k=1}^{\infty}(\eta^{-1}a_k + b_k)n^{*k}\right], \tag{3.7}$$

which demonstrates that $\omega(n^* = 0) = -p' - 2L^{-1}b_0$ does not depend on the radius $R = \eta L$. Eq. (3.7) also indicates that it should be feasible to fit the simulation data by polynomials in n^*.

Kinetics and free energy difference between states The time series $n(t)/n_0$ of the water density in the pore was analysed in the spirit of single channel recordings[204] by detecting open (high-density; in the following denoted by a subscript 'o') and closed (approximately zero density, subscript 'c') pore states, using a Schmitt-trigger with an upper threshold of 0.65 and a lower threshold of 0.15. A characteristic measure for the behaviour of these pores is the *openness*

$$\langle\omega\rangle = \frac{\mathfrak{T}_o}{\mathfrak{T}}, \tag{3.8}$$

the ratio between the total time spent in the open state, $\mathfrak{T}_o$, and the total simulation time $\mathfrak{T} = \mathfrak{T}_o + \mathfrak{T}_c$. In other words, the openness is the probability for the pore being in the open state;[62] errors can be conveniently estimated from a block-averaging procedure.[205,206] Analysis of the simulation data shows that the distributions of the lifetimes t_α of state $\alpha = \{o, c\}$ are exponentials $\tau_\alpha^{-1} e^{-t_\alpha/\tau_\alpha}$. The maximum-likelihood estimator for the characteristic times τ_o and τ_c is the mean $\tau_\alpha = \langle t_\alpha \rangle$.

The free energy difference between closed and open state,

$$\Delta\Omega = \Omega_c - \Omega_o,$$

can be calculated in two ways. Firstly, we obtain it from the equilibrium constant

$$K = \frac{\mathfrak{T}_c}{\mathfrak{T}_o} = \frac{\mathfrak{T} - \mathfrak{T}_o}{\mathfrak{T}_o} = \langle\omega\rangle^{-1} - 1$$

*The wall itself is treated as an external potential and is not part of the thermodynamics; otherwise the value would be the one of the "dry surface" tension γ_{SO}, see Ref. 92.

of the two-state system *open*⇌*closed* as

$$\beta\,\Delta\Omega_{\mathrm{kin}} = -\ln K. \tag{3.9}$$

Secondly, $\Delta\Omega$ is determined from $p(N^*)$ as the ratio between the probability that the pore is in the closed state and the probability for the open state,

$$\beta\,\Delta\Omega_{\mathrm{eq}} = -\ln\frac{P_{\mathrm{c}}}{P_{\mathrm{o}}} = -\ln\frac{\sum_{N^*\le N^\ddagger} p(N^*)}{\sum_{N^*>N^\ddagger} p(N^*)}. \tag{3.10}$$

The definition of state used here is independent of the kinetic analysis. It only depends on $N^\ddagger$, the pore occupancy in the transition state, the state of lowest probability between the two probability maxima that define the closed and open state. Of course, Eq. (3.10) is only applicable if both a liquid and a vapour state exist; Eq. (3.9) can be formally applied even if there is only one stable state. The relationship involving K can be inverted to describe the openness in terms of $\Delta\Omega(R)$,

$$\langle\omega(R)\rangle = \frac{1}{1+\exp[-\beta\,\Delta\Omega(R)]}. \tag{3.11}$$

Self-diffusion coefficient The three Cartesian components of the self-diffusion coefficient of water are calculated from the Einstein relations[112]

$$\left\langle \big(x_i(t) - x_i(t_0)\big)^2 \right\rangle = 2D_i\,(t-t_0) \tag{3.12}$$

(which is understood to be exact in the limit $t\to\infty$). We can rewrite (3.12) for the diffusion in space as

$$\left\langle \big(\mathbf{x}(t) - \mathbf{x}(t_0)\big)^2 \right\rangle = 2\,\mathrm{Tr}\,\mathsf{D}\,(t-t_0), \quad \text{with } D_{ij} := D_i\delta_{ij}, \tag{3.13}$$

where the diffusion constant is now a diagonal Cartesian tensor D.* This formulation is motivated by a generalisation of Fick's law for the current

$$\mathbf{j}(\mathbf{x},t) = -\mathsf{D}\boldsymbol{\nabla} n(\mathbf{x},t) \tag{3.14}$$

and the continuity equation

$$\frac{\partial n(\mathbf{x},t)}{\partial t} - \boldsymbol{\nabla}\cdot\mathsf{D}\boldsymbol{\nabla} n(\mathbf{x},t) = 0 \tag{3.15}$$

*Tr is the standard trace operation on tensors, $\mathrm{Tr}\,\mathsf{D} = \sum_{i=1}^{3} D_{ii}$. The tensor transformation properties follow from Eq. (3.14).

(expressing the conservation of particle number in the volume element d^3x). For isotropic fluids all three components of the tensor are identical, $D_i \equiv D$, and (3.13) reduces to the well-known 3D random walk result[99,111]

$$\left\langle \left(\mathbf{r}(t) - \mathbf{r}(t_0)\right)^2 \right\rangle = 6D\,(t - t_0).$$

We further generalise to a position-dependent diffusion tensor $\mathsf{D}(\mathbf{x})$. By calculating the mean square displacement (MSD) $\langle \mathbf{x}^2 \rangle$ from (3.15) we arrive at*

$$\frac{\partial}{\partial t} \int\limits_V d^3x\, \mathbf{x}^2 n(\mathbf{x}, t) = \frac{\partial \langle \mathbf{x}^2 \rangle}{\partial t} = 2 \left\langle \operatorname{Tr} \mathsf{D} + (\boldsymbol{\nabla}^\dagger \mathsf{D}) \cdot \mathbf{x} \right\rangle \tag{3.16}$$

or alternatively,

$$\lim_{t\to\infty} \langle \mathbf{x}^2 \rangle \asymp 2 \left\langle \operatorname{Tr} \mathsf{D} + (\boldsymbol{\nabla}^\dagger \mathsf{D}) \cdot \mathbf{x} \right\rangle t. \tag{3.17}$$

expresses the fact that for large times the MSD increases linearly with t. For position independent D the gradient vanishes and (3.13) is recovered. In the case of a slab system with broken translational symmetry in the z direction we consider the volume integration in (3.16) to be carried out only in slices of finite thickness Δz at positions z so that both the mean square displacement $\lambda^2 := \langle \mathbf{x}^2 \rangle$ and the effective diffusion constant $D_{\text{eff}} := \frac{1}{3} \langle \operatorname{Tr} \mathsf{D} + \mathbf{x} \cdot \boldsymbol{\nabla} \mathsf{D} \rangle$ become functions of z,

$$\lambda^2(z) \approx 6 D_{\text{eff}}(z) t. \tag{3.18}$$

The equation only holds approximately because a particle will diffuse out of its slice of finite thickness in a finite time (on the order of $\tau = \Delta z^2 / 2 D_{\text{eff}}$) and hence it is no longer possible to obtain the limit $t \to \infty$. In other words there is a trade-off between obtaining good spatial resolution (small Δz) and good mean square deviations (long t) and hence more accurate diffusion constants. These considerations show that a local diffusion constant (or tensor) is, after all, not a rigorously defined quantity at a spatial resolution comparable to the molecular length scale. Instead it can be viewed as the effect of variations of the local free energy barriers lumped into an effective diffusion coefficient. Eq. (3.18) also shows that in regions where the diffusion constant does not vary strongly the spherical average over the diffusion tensor in a Δz slice should equal the isotropic bulk value

$$\frac{1}{3} \langle \operatorname{Tr} \mathsf{D}(z) \rangle_{\Delta z} = D, \tag{3.19}$$

which can be used as a check on the simulation results. In order to obtain the components of the diffusion tensor from the MD trajectories the simulation box

*The notation $(\boldsymbol{\nabla}^\dagger \mathsf{D}) \cdot \mathbf{x}$ is perhaps clearer written in components as $\sum_{j,k} x_j \partial_k D_{kj}(\{x_i\})$.

was stratified perpendicular to the pore axis or slab normal; the central layer of thickness $\Delta z = L$ contained the whole pore region. During the simulation time $\mathfrak{T}$ the MSD of water molecules in a given layer was accumulated for 10 ps and after discarding the first 2 ps, a straight line was fit to the MSD to obtain D_i. These diffusion coefficients were averaged in each layer for the final result.

Water current density The current density (flux per area) was calculated as $j_0 = \Phi_0/A$ from the equilibrium flux $\Phi_0 = M/\mathfrak{T}$ with the total number of permeant water molecules M and the effective pore cross section $A = \pi R_{\text{eff}}^2$ for pores or $A = L_x L_y$ for the bulk case, i.e. a simulation box of water with periodic boundary conditions. Permeant water molecules were defined as those whose pore entrance and exit z-coordinate differed (see Appendix D.2.2 for details). In addition, distributions of permeation times were computed.

3.3 Results and Discussion

Before investigating water in cylindrical pores in Section 3.3.2 it is worthwhile to look at the simplest inhomogeneous system: a planar surface in contact with liquid water. This will later allow us to compare the effect of different degrees of confinement. In cylindrical pores water is more strongly confined than near a surface. The former restricts movement of water molecules in the xy-plane (with the pore axis pointing along z) whereas the latter only restricts movement in one half-space along $-z$ (if the surface normal is taken to be parallel to $+z$).

3.3.1 Water near a hydrophobic surface

As described in Section 3.2.1, the hydrophobic surface comprises a crystal-like slab of methane molecules. These molecules (represented as united atoms with implicit hydrogens) oscillate about their equilibrium position with an RMSD of about 0.1 nm. The surface normal is parallel to the z direction, and all properties of interest are predominantly a function of z.

Density variations: Layering Simple, non-associating fluids like liquid argon or the theoretician's favourites, the hard sphere or the Lennard-Jones fluid,[101] exhibit a strongly layered structure near a repulsive or weakly attractive surface[194,207] with spatial oscillations in density which can span almost one order of magnitude. Associating liquids like water tend to show weaker layering (about 20% to 50% variation in density compared to the bulk value) as seen in simulations;[84,208–216] experimental evidence by surface force apparatus measurements,[217,218] atomic force microscopy (J. Pethica, personal communication),

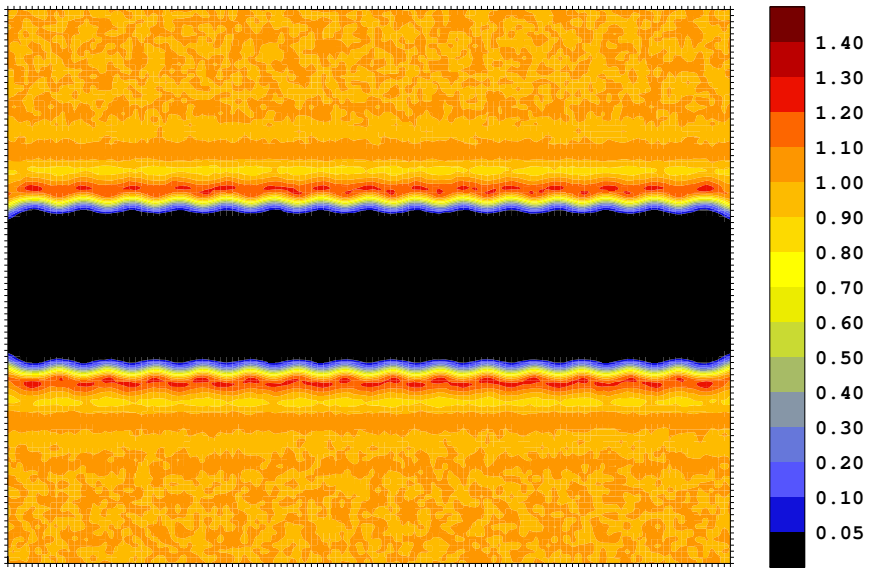


Figure 3.3: Water density near a hydrophobic surface. The density is averaged over the y direction. Weak layering is visible near the membrane mimetic slab. The interlayer distance is about 0.3 nm. The graph shows an average over a simulation time $\mathfrak{T} = 20$ ns; note that even after this comparatively long time the average density in the bulk has not completely converged (compare to the uniform density in Fig. 5.5 on page 116 over $\mathfrak{T} = 100$ ns). The density is in units of SPC bulk water $n_0 = 0.997 \times 10^3$ kg $\cdot$ m^{-3} at $T = 300$ K and $P = 1$ bar. See also the one-dimensional density profile in Fig. 3.4 on the following page.

and neutron diffraction[219] is indicative of surface layers near different kinds of surface. As shown in Fig. 3.3 and the bottom panel of Fig. 3.4 on the next page water near a hydrophobic slab shows marked layering. The 2D density shows that the influence of the surface roughness on the density is not large and only affects water directly adjacent to the surface. Two layers are visible, with the peak height not exceeding $1.2n_0$ (where $n_0 = 0.997 \times 10^3$ kg $\cdot$ m^{-3} denotes the density of SPC water at standard conditions). Between the first water layer and the slab surface there is a region of very low water density which is known as the depletion layer.[90,94,215,216,220] Its thickness of ca. 0.2 nm is in accord with other simulation studies of explicit water near hydrophobic surfaces.[215,216] The thickness of one water layer (as measured by the peak to peak distance between the two layers) is about 0.3 nm, which is only slightly larger than the position of the first peak in the oxygen-oxygen radial distribution function in bulk water at 0.28 nm. The influence of the wall, as far as the density is concerned, does not extend farther than two to three water layers, i.e. about 1 nm; this was already noted in early MD simulations[208,209] but at much shorter simulation times (72 ps). The range of layering effects suggests that the minimum thick-

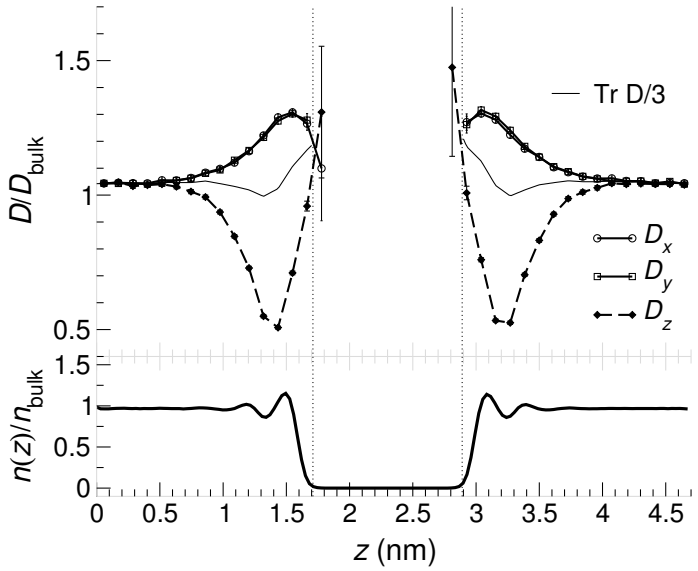


Figure 3.4: Diffusion coefficients of water near a hydrophobic surface. *Top panel*: Diffusion of SPC water molecules is measured in directions x and y parallel to the surface and in z perpendicular to it. D_i are normalised to the value of the isotropic diffusion coefficient of SPC water, $D_0 = 4.3\ \mathrm{nm}^2 \cdot \mathrm{ns}^{-1}$ at $T = 300$ K and $P = 1$ bar. The isotropic average $\frac{1}{3}\,\mathrm{Tr}\,\mathrm{D} = \frac{1}{3}(D_x + D_y + D_z)$ is plotted as a check (see text). Error bars indicate one standard deviation. Vertical lines show the van der Waals surface of the slab. *Bottom panel*: The one dimensional density profile $n(z)$ shows weak layering and a 0.2 nm depletion zone ($n_0 = 0.997 \times 10^3\ \mathrm{kg} \cdot \mathrm{m}^{-3}$).

ness of the water region around a solute should be larger than 1 nm so that the water at the edge of the simulation box has bulk characteristics. It should be noted that water close to a surface which is fluid rather than crystalline does not exhibit pronounced layering as was observed by analysis of water near lipid membranes and octane slabs (J. D. Faraldo-Gómez, personal communication). Presumably, the movements of the surface disrupt or smear out the layering, which is not too surprising as the difference in density of 20% relates to a free energy of layer stabilization of about $\beta^{-1} \ln 1.2 \approx 0.2\ kT$ (cf. Eq. (2.94)). This energy can be easily overcome by fluctuations in the liquid-liquid interface. (The fluctuations are capillary waves, which have a typical energy of 1 kT.[221,222]) However, the density is only one property, although a rather important one as far as equilibrium properties are concerned. For dynamical properties, the diffusion coefficient should be a good indicator of the influence of the surface on the liquid.

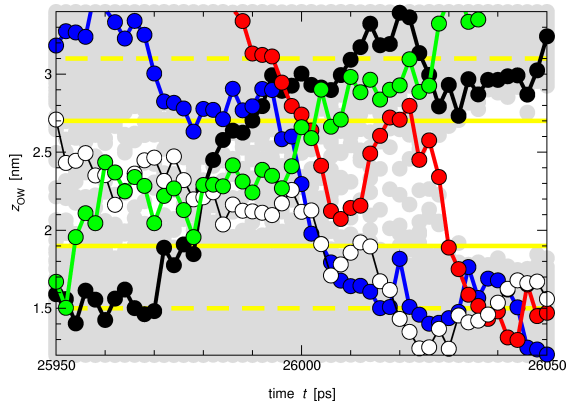


Figure 3.5: Snapshots of permeant water molecules in a $R = 0.55$ nm pore as it switches from the open to the closed state. z-coordinates of the water oxygen atoms are drawn every 2 ps. The mouth and pore region are indicated by horizontal broken and solid lines. Five trajectories are shown explicitly. The white water molecule permeates the pore within 54 ps whereas the black one only requires about 10 ps.

Dynamics The components of the self-diffusion coefficient tensor are shown in Fig. 3.4 as functions of z. The translational symmetry of the system is broken normal to the surface. Accordingly D_z, and D_x and D_y only vary with z. The components parallel to the surface retain their symmetry, i.e. $D_x(z) = D_y(z)$. As expected, D_z drops to half its bulk value, reflecting the fact that one half of the space is not accessible. It increases within the depletion layer, which indicates highly mobile and unbound water molecules that venture close to the methane molecules of the slab but are quickly expelled from the hydrophobic region. The parallel components increase to about $1.25D_0$ (where $D_0 = 4.34 \pm 0.01$ nm$^2 \cdot$ ns^{-1} is the bulk diffusion constant, obtained from an independent simulation of neat SPC water), i.e. within the first layer the water molecules are rather mobile parallel to the plane (which was also observed in other simulations[214]). Only very close to the surface the mobility decreases because the roughness of the slab impedes translational motion in the plane of the slab. For (isotropic) bulk water, all three components of D equal the bulk diffusion constant D, and their deviation from D_0 indicates the degree of non-bulk like behaviour. Similarly, the isotropically averaged diffusion constant $\frac{1}{3}(D_x + D_y + D_z)$ (see Eq. (3.18) on page 67), which measures effective diffusion and whose deviation from D also measure non-bulk like behaviour without regard to direction, only appreciably deviates from unity in the first layer and the depletion zone. The components D_i themselves, however, deviate from the bulk value over more than 1 nm although density variations are damped out before reaching this distance. Based on the diffusion coefficient, *bulk water* only begins at a distance > 1 nm from the membrane mimetic.

3.3.2 Water in cylindrical pores

The water density in the pore cavity oscillates between an almost empty (closed) and filled (open) state (Fig. 3.1 on page 59). We refer to the water-filled pore state as open because such a pore environment would favourably solvate an ion and

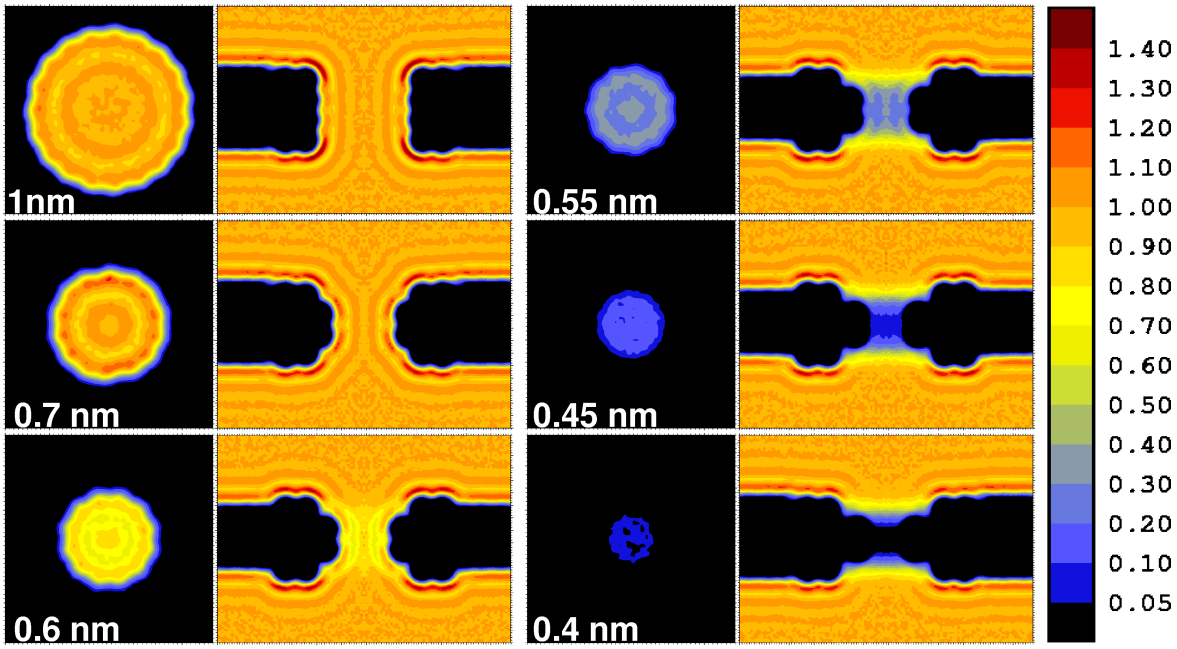


Figure 3.6: Water density in hydrophobic pores with radii ranging from 1.0 nm to 0.4 nm. Left column in each panel: density z-averaged over the length of the pore. Right column: radially averaged density. The density is in units of SPC bulk water at 300 K and 1 bar. Note that for the $R = 1.0$ nm, 0.7 nm, and 0.4 nm pore the membrane mimetic slab had a thickness of four methane layers but for technical reasons only three layers were used in the other cases. (Plots were prepared with xfarbe 2.5 (Ref. 223))

conceivably allow its permeation. Conversely, we assume that a pore that cannot sustain water at liquid densities will present a significant energetic barrier to an ion. As shown in Fig. 3.5 on the preceding page, water molecules can pass each other and often permeate the pore in opposite directions simultaneously.

Even though the oscillating behaviour was already suggested by earlier 1 ns simulations[62] only at these longer times do clear patterns emerge. The characteristics of the pore-water system change substantially with the pore radius. The oscillations (Fig. 3.1) depend strongly on the radius. The water density (Fig. 3.6) shows large pores to be water-filled and strongly layered at bulk density. With decreasing radius the average density is reduced due to longer closed states even though layer structures remain. The narrowest pores appear almost void of water.

The sudden change in behaviour is borne out quantitatively by the openness (Fig. 3.7), which indicates a sharp increase with increasing radius around $R = 0.55$ nm. Although the range of radii over which this transition takes place appears to be small (0.45 nm to 0.7 nm) the cross-sectional area doubles. The

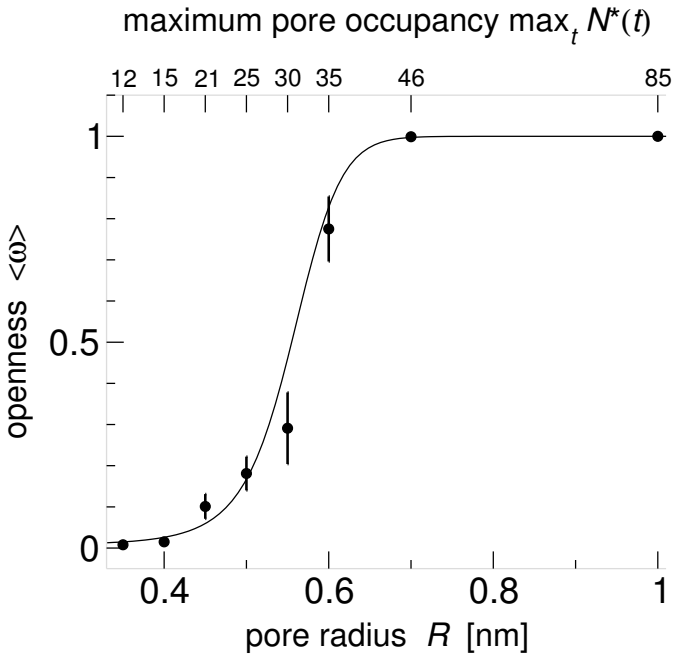


Figure 3.7: Openness $\langle\omega(R)\rangle$ of hydrophobic pores.. Wide pores are permanently water-filled ($\langle\omega\rangle = 1$) whereas narrow ones are predominantly empty ($\langle\omega\rangle \approx 0$). The continuous line is the function $\left(1 + \exp[-\beta\,\Delta\Omega_{\mathrm{eq}}(R)]\right)^{-1}$, with $\Delta\Omega_{\mathrm{eq}}(R)$ determined independently of $\langle\omega\rangle$. ($\Delta\Omega(R)$ is a quadratic function of R.) The maximum pore occupancy is the maximum number of water molecules that were recorded during simulations.

maximum number of water molecules found in the cavity in our simulations more than doubles from 21 to 46 in this range of R, so that the average environment which each water molecule experiences changes considerably.

Density The radial densities in Fig. 3.6 on the preceding page show destabilization of the liquid phase with decreasing pore radius. Above $R = 0.45$ nm distinctive layering is visible in the pore, and for the larger pores appears as an extension of the planar layering near the slab. For $R < 0.45$ nm no such features remain and the density is on average close to 0. The open state can be identified with liquid water and the closed state with water vapour. In the continuously open 1 nm-pore, the average density $\overline{n(t)}/n_0$ is 0.82, but 0.032 in the closed 0.35 nm-pore. Brovchenko et al.[224] carried out GEMC simulations of the coexistence of liquid TIP4P water with its vapour in an infinite cylindrical hydrophobic pore of radius $R = 1.075$ nm. At $T = 300$ K they obtained a liquid density of 0.81 and a vapour density close to 0, in agreement with the numbers from our MD simulations.

Analysis of the structure in the radial PMF (Fig. 3.8) lends further support to the above interpretation. Water molecules fill the narrow pores ($R \lessapprox 0.45$ nm) homogeneously as it is expected for vapour. For the wider pores, distinct layering is visible as the liquid state dominates. The number of layers increases from two to three, with the central water column being the preferred position initially. As the radius increases, the central minimum shifts toward the wall. For $R = 0.7$ nm the centre of the pore is clearly disfavoured by $0.2\,kT$. In the

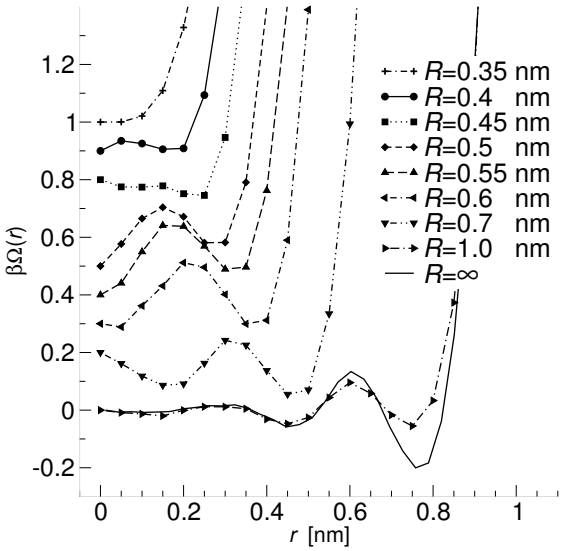


Figure 3.8: Radial potential of mean force of water $\Omega(r)$. Very narrow pores show a relatively featureless PMF, consistent with a predominantly vapour-like state. For larger pore radii, the liquid state, characterised by distinct layering, dominates. The PMF of the 1 nm pore is very similar to the one of water near a planar hydrophobic slab ($R = \infty$). PMFs are drawn with arbitrary offsets.

largest pore ($R = 1.0$ nm), the influence of curvature on the density seems to be already negligible as it is almost identical to the situation near a planar hydrophobic slab.

Kinetics Condensation (filling of the pore) and evaporation (emptying) occur in an avalanche-like fashion as shown in Fig. 3.5 on page 71. In our simulations both events take place within ca. 30 ps, roughly independent of R. However, the actual evaporation and condensation processes seem to follow different paths, as we can infer from the analysis of the kinetics of the oscillations. The time series of Fig. 3.1 reveals that the lifetimes of the open and closed state behave differently with increasing pore radius (Fig. 3.9 on the next page): In the range $0.45\text{ nm} \leq R \leq 0.6\text{ nm}$, the average time a pore is in the closed state is almost constant, $\tau_c = 1.40 \pm 0.37$ ns; outside this range no simple functional relationship is apparent. The average open time can be described as an exponential $\tau_o(R) = a \exp(R/\zeta)$ with $a = 1.3 \times 10^{-5}$ ns and $\zeta = 4.9 \times 10^{-2}$ nm for $0.35\text{ nm} \leq R \leq 0.7\text{ nm}$.

$1/\tau_o$ is related to the "survival probability" of the liquid state and $1/\tau_c$ to that of the vapour state. These times characterise the underlying physical evaporation and condensation processes. Their very different dependence on R implies that these processes must be quite different. The initial condensation process could resemble the evaporation of water molecules from a liquid surface. Evaporating molecules would not interact appreciably, so that this process would be rather insensitive to the area of the liquid-vapour interface $A = \pi R^2$ and hence R. The disruption of the liquid pore state, on the other hand, displays very strong dependence on the radius. Conceivably, the pore empties once a density fluctuation has created a vapour bubble that can fill the diameter of the

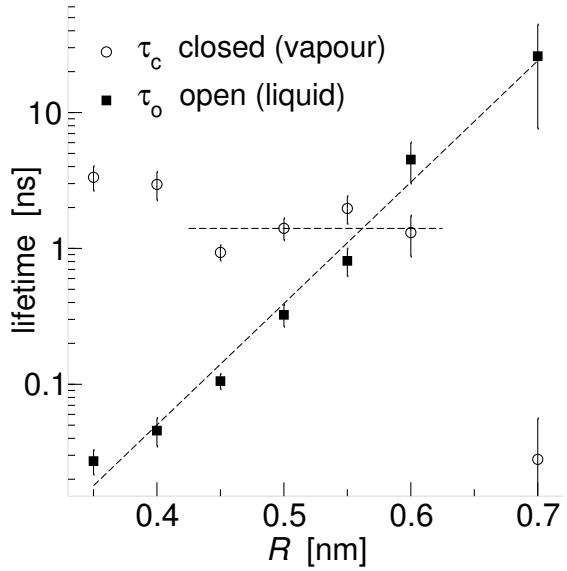


Figure 3.9: Kinetics *open⇌closed*. The average lifetime of the open state τ_o depends on the radius exponentially whereas τ_c is approximately constant in the two-state region (cf. Fig. 3.11) of radii.

pore and expose the wall to vapour, its preferred contact phase. The probability for the formation of a spherical cavity of radius λ with exactly N water molecules inside was determined by Hummer et al.[225]. From their study we find that the probability $p(\lambda; n)$ for the formation of a bubble of radius λ and density below a maximum density n is apparently an exponential. Once a bubble with $\lambda \approx R$ develops, the channel rapidly empties (much as in classical nucleation theory[99,100,113]) but this occurs with a probability that decreases exponentially with increasing R, which corresponds to the observed exponential increase in τ_o. In particular, for low density bubbles ($n < 0.2\, n_0$) we estimate the decay constant in $p(\lambda; n)$ as 2×10^{-2} nm, which is of the same order of magnitude as ζ. Waghe et al.[226] arrive at similar conclusions for the case of single-file carbon nanotubes.

From the equilibrium constant $K(R) = \mathfrak{T}_c(R)/\mathfrak{T}_o(R) = \exp[-\beta\,\Delta\Omega(R)]$ the free energy difference between the two states $\Delta\Omega = \Omega_c - \Omega_o$ can be calculated (Fig. 3.10). As will be shown in detail in Section 4.2.4 on page 89 in the next chapter [Eq. (4.5)], $\Delta\Omega$ is a quadratic function of R, $\beta\,\Delta\Omega(R) = a_2\,R^2 + a_1\,R$,* with the two parameters $a_2^{\text{kin}} = 0.584 \times 10^2$ nm^{-2} and $a_1^{\text{kin}} = -32.4$ nm^{-1}.† Together with $K(R)$, the gating behaviour of the pore is characterised.[204] In this

*In Ref. 96 we suggested that $\Delta\Omega(R)$ would be a linear function. At the time there were not enough data available to distinguish a quadratic polynomial from a linear one. Subsequently, thermodynamic models were found which predicted the quadratic relationship.[97,202]

†Note that this estimate for $\Delta\Omega(R)$ is based on fewer and shorter simulations than the data in Tab. 4.1 on page 93. Nevertheless, when a_2 and a_1 are expressed in terms of the parameters of the model Eq. (4.5), $\gamma_{lv} = a_2/2\pi = +9.3\ kT$ nm^{-2} and $\Delta\gamma_w = a_1/2\pi L = -6.4\ kT$ nm^{-2}, good agreement is found between them and the more extensive simulations ("hydrophobic" in Tab. 4.1).

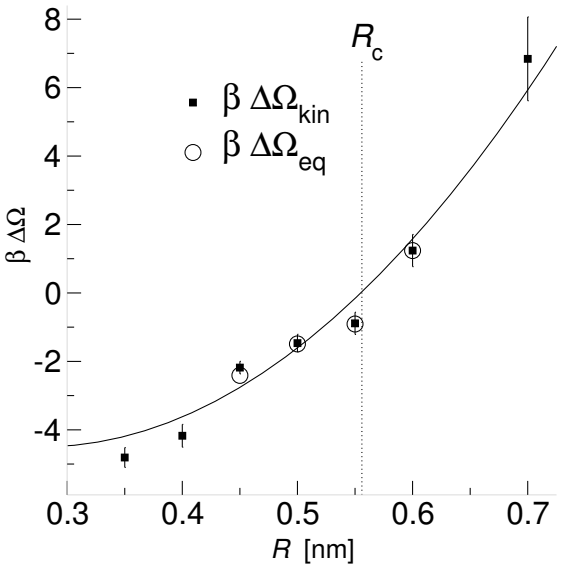


Figure 3.10: Free energy difference $\Delta\Omega = \Omega_c - \Omega_o$ between the open (liquid) and closed (vapour) state. Liquid is stable in the pore if $\Omega_o < \Omega_c$, i.e. $\Delta\Omega > 0$. The critical radius R_c denotes the coexistence pore size. $\Delta\Omega$ can be calculated from the kinetics Eq. (3.9) or the equilibrium distribution of states Eq. (3.10). The continuous line is a quadratic fit to $\beta\Delta\Omega_{kin}$, based on a surface-energy model [Eq. (4.5)]. Note that $\beta = 1/kT$ means that these energies are measured in kT.

sense, the MD calculations have related the input structure to a "physiological" property of the system.*

Free energy density The free energy density $\omega(n^*; R)$ displays one or two minima (Fig. 3.11): one for the vapour-filled pore ($n^* \approx 0$) and one in the vicinity of the bulk density. The 0.45 nm pore is close to a transition point in the free energy landscape: the minimum for the filled pore is very shallow and disappears at smaller radii ($R = 0.4$ nm and 0.35 nm). For very large and very small radii, only one thermodynamic stable state exists, either liquid or vapour. For intermediate radii, a metastable state appears. Near $R = 0.55$ nm both states are almost equally probable although they can not coexist spatially because the pore is finite and too short. In infinite pores *spatially* alternating domains of equal length would be expected[227] and were actually observed in MD simulations.[228] The oscillating states in short pores, on the other hand, alternate *temporally*, thus displaying a kind of "time-averaged" coexistence.

From the probability distribution $p(N^*)$ the free energy difference between closed and open state $\Delta\Omega(R)$ is calculated according to Eq. (3.10). Although only four data point are available for a two parameter fit, the result $a_1^{eq} = -27.3\ \mathrm{nm}^{-1}$ and $a_2^{eq} = 0.482 \times 10^2\ \mathrm{nm}^{-2}$ is consistent with the estimate from

*Note, however, that the time scales of ion channel gating (ca. 1 ms $= 1 \times 10^{-3}$ s) that are resolved in single-channel recordings[12,204] and of the oscillations observed here (about 1 ns $= 1 \times 10^{-9}$ s) differ by six orders of magnitude. The traces, though, look superficially similar. This appears to be a consequence of the fact that both are essentially two-state systems whcih switch between these two states (ignoring the complicated matter sub-states and desensitised states of ion channels[204]).

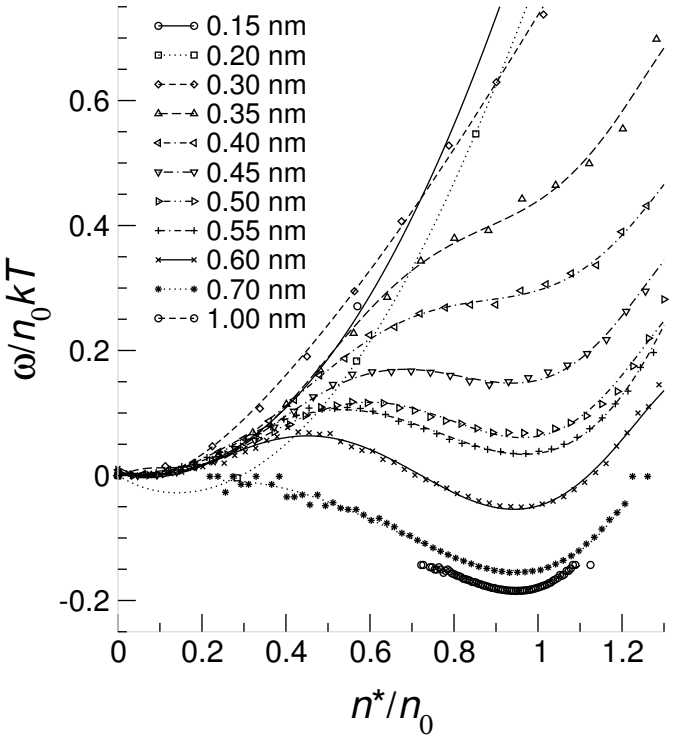


Figure 3.11: Constrained free energy density $\omega(n^*) = \Omega(n^*)/V$ at constant $T = 300$ K. n^* is the water density in the pore, normalised to bulk density n_0. ω is given in units of $n_0 kT = p_0^{\mathrm{id}}$, i.e. the pressure of the ideal gas with density n_0. Two minima correspond to the observed two-state behaviour. The vapour state becomes metastable with increasing radius and for $R > 0.55$ nm the liquid state is globally stable. Curves for $R \geq 0.7$ nm are drawn with an arbitrary offset. For completeness, some simulations from Chapter 4 were also included.

the kinetics. The graph of $\Delta\Omega(R)$ in Fig. 3.10 shows the transfer of stability from the vapour state for small R to the liquid state for large R. The progression from a stable vapour state through coexistence to a stable liquid state is also evident from Fig. 3.12. For small radii $R \leq 0.4$ nm the constrained grand potential $\Omega(n^*)$ only exhibits a minimum near the vapour density, though for 0.35 nm $\leq R \leq$ 0.4 nm already a dent develops, which hints at the liquid state. At intermediate radii ($0.45 \leq R \leq 0.6$ nm) two minima have developed, which belong to a stable and a metastable state. Below the coexistence radius $R_c = 0.56$ nm (cf. Fig. 3.10) liquid is metastable but at R_c both liquid and vapour could coexist in equilibrium ($\Delta\Omega(R_c) = 0$). At $R = 0.6$ nm, vapour has become metastable, but disappears completely for even larger radii.

Dynamics MD simulations not only allow us to investigate the thermodynamic properties of the system but also the dynamical behaviour of individual molecules. A few selected water molecules are depicted in Fig. 3.5 on page 71 shortly before the pore empties. They show a diverse range of behaviours and no single-file like motion of molecules is visible in the liquid state. On evaporation (and condensation) the state changes within ca. 30 ps.

The mean permeation time $\langle \tau_p \rangle$ in Tab. 3.1 on page 79 increases with the pore radius, i.e. water molecules permeate narrow hydrophobic pores faster than they diffuse the corresponding distance in bulk water (the reference value).

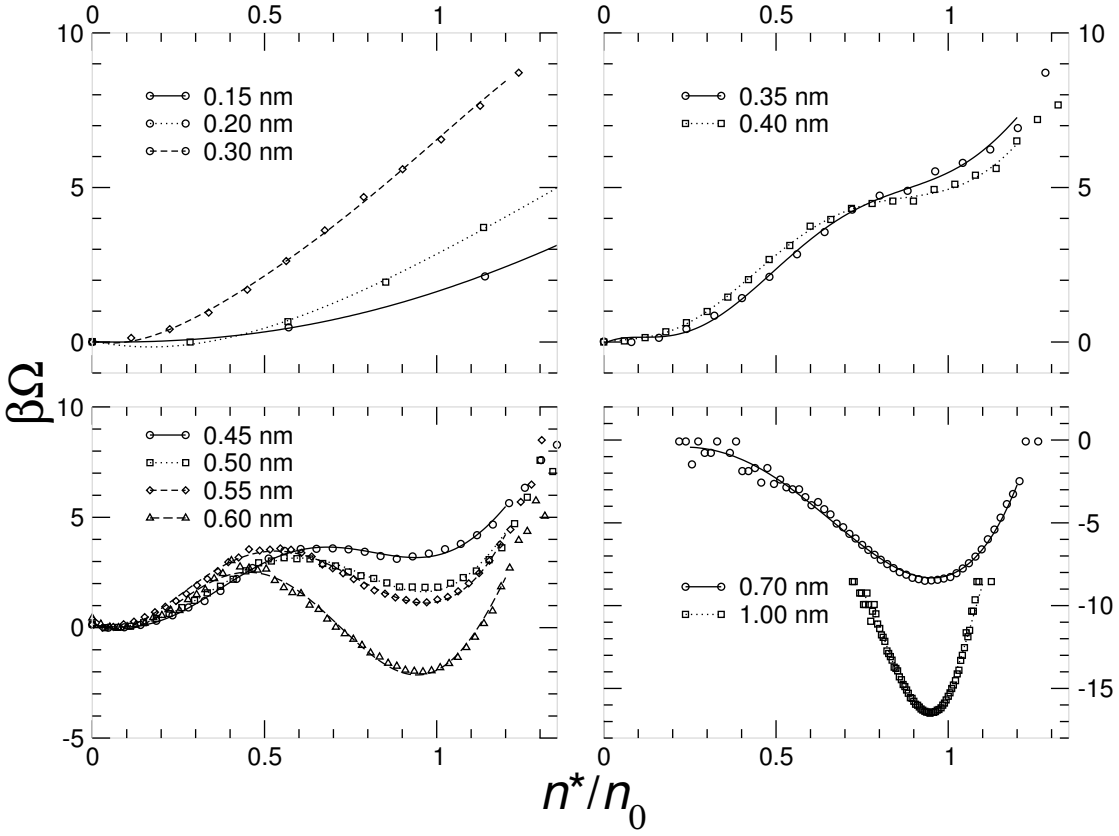


Figure 3.12: Constrained grand potential $\Omega(n^*)$ of water in hydrophobic pores. The free energy $\beta\Omega$ is displayed (in units of kT) over the water density n^* in the pore. The graphs are grouped according to the stability of the liquid or vapour state. n_0 is the bulk density of SPC water at 300 K and 1 bar. The same simulations as in Fig. 3.11 are shown. All curves are arbitrarily shifted to coincide in the vapour state.
It is no coincidence that the data always span about 10 kT. It is a consequence of the length of the equilibrium simulations. The largest free energy difference that can be estimated from a trajectory of length $\mathfrak{T}$, when frames of length Δt were saved, is $\Delta\mathcal{W}_{\max} = -kT \ln \Delta t/\mathfrak{T}$. For $\mathfrak{T} = 100$ ns and $\Delta t = 2 \times 10^{-3}$ ns, $\Delta\mathcal{W}_{\max} = 10.8\ kT$, which is the maximum $\Delta\Omega$ in the data presented here. The minimum Δt that can be useful is the correlation time of the system,[110] or rather, of the property of interest. Typical time scales in the water/pore systems: water reorientation time $\tau \approx 1$ ps, change of state (filling/emptying) ca. 30 ps; for the density fluctuations in the pore the correlation time has not been measured. Data points based on a single frame are not reliable, and these points tend to be rather noisy, so they were excluded from the polynomial fits of $\Omega(n^*)$ and $\omega(n^*)$.

Table 3.1: Dynamical properties of water in hydrophobic pores. R is the van der Waals pore radius, with $R = \infty$ denoting the bulk. $\langle\omega\rangle$ is the openness. The mean permeation time $\langle\tau_p\rangle$ is measured relative to the bulk value, $\tau_{p,bulk} = 29.9 \pm 0.1$ ps. The equilibrium current density j_0 is the total number of permeant water molecules per unit time and unit area ($j_{0,bulk} = 320 \pm 3$ ns^{-1} nm^{-2}). The diffusion coefficient along the pore axis D_z is normalised to the bulk value of SPC water at 300 K and 1 bar ($D_0 = 4.34 \pm 0.01$ nm^2 ns^{-1}). One standard deviation errors in the last decimals are given in parentheses.

R/nm	$\langle\omega\rangle$	$\langle\tau_p\rangle / \langle\tau_p\rangle_{bulk}$	$j_0 / j_{0,bulk}$	D_z / D_{bulk}
0.35	0.008(2)	0.482(61)	0.025(2)	
0.4	0.015(5)	0.421(25)	0.027(2)	2.87(9)
0.45	0.101(30)	0.629(12)	0.109(3)	2.27(4)
0.5	0.181(41)	0.729(10)	0.194(4)	1.91(3)
0.55	0.291(89)	0.786(8)	0.279(4)	1.87(3)
0.6	0.775(79)	0.833(5)	0.721(6)	1.32(1)
0.7	0.999(1)	0.799(3)	1.004(7)	1.25(0)
1.0	1.000(0)	0.819(2)	1.011(5)	1.18(0)
∞		1.000(3)	1.000(8)	1.00(0)

This is consistent with higher diffusion coefficients D_z in the narrow pores (up to almost three times the bulk value). The diffusion coefficient perpendicular to the pore axis, D_{xy}, drops to approximately half the bulk value. Martí and Gordillo[229] also observe increased diffusion in simulations on water in carbon nanotubes ($D_z \leq 1.65\, D_0$) and a corresponding decrease in D_{xy}. Experimental studies on water transport through desformyl gA[190] can be interpreted in terms of a D_z of five times the bulk value. Histograms for τ_p show that there is a considerable population of 'fast' water molecules (e.g. the black and the blue one in Fig. 3.5 on page 71 with τ_p between 2 and 10 ps, which is not present in bulk water. The acceleration of water molecules in the pore can be understood as an effect of 1D confinement. The random 3D motion is directed along the pore axis and the particle advances in this direction preferentially because there is only very weak attraction between particle and wall. The effect increases with decreasing radius, i.e. increasing confinement. The average equilibrium current density j_0 follows the trend of the openness closely but more detailed time-resolved analysis shows water translocation to occur in bursts in all pores. In narrow pores, bursts occurring during the "closed" state contribute up to 77% of the total flux. For single-file pores, simulations[191,197] and theory[230] also point towards collective motions of a number of water molecules as the predominant form of transport.

Capillary condensation The behaviour as described so far bears the hallmarks of capillary condensation and evaporation[194,207,231] although it is most often associated with physical systems which are macroscopically extended in at least one dimension such as slits or long pores. Capillary condensation can be discussed in terms of the Kelvin equation,[193]

$$\ln \frac{p}{p_0} = -\frac{\beta \gamma_{lv} v_l}{r}, \tag{3.20}$$

which describes how vapour at pressure p relative to its bulk-saturated pressure p_0 coexists in equilibrium with its liquid. Liquid and vapour are divided by an interfacial meniscus of curvature $1/r$ ($r > 0$ if the surface is convex); γ_{lv} is the surface tension between liquid and vapour and v_l the molecular volume of the liquid. Although the Kelvin equation is not expected to be quantitative in systems of dimensions of only a few molecular diameters it is still useful for obtaining a qualitative picture. Curvature $1/r$ and contact angle θ in a cylindrical pore of radius R are related by $R = r \cos\theta$. With the Young equation,

$$\gamma_{wv} = \gamma_{wl} + \gamma_{lv} \cos\theta, \tag{3.21}$$

Eq. (3.20) becomes

$$\ln \frac{p(R)}{p_0} = -\frac{\beta(\gamma_{wv} - \gamma_{wl}) v_l}{R}, \tag{3.22}$$

	Ref.	$p_f \times 10^{14}$ [$cm^3 s^{-1}$]	Φ_0 [ns^{-1}]
Aqp1	187	4.9	3.2
Aqp4	187	15	9.7
AqpZ	188	2.0	1.3
gA[a]	189	1.6	1.0
desformyl gA[b]	190	110	71
$R = 0.35$ nm		4.0	2.6
$R = 0.40$ nm		5.7	3.7
$R = 0.45$ nm		30.0	19.4
$R = 0.50$ nm		66.5	43.0
$R = 0.55$ nm		117	75.8
$R = 0.60$ nm		363	235
$R = 0.70$ nm		700	453
$R = 1.0$ nm		1480	956
carbon nanotube[c]	197	26.2	16.9
desformyl gA (DH)[d]	191	10	5.8

[a] bacterial peptide channel gramicidin A
[b] desformylated gramicidin A
[c] $(6,6)$ carbon nanotube, $R \approx 0.24$ nm
[d] desformyl gA in the double-helical conformation

Table 3.2: Osmotic permeability coefficient p_f and equilibrium flux Φ_0 of water in selected simulations and experiments. We used the relationship $p_f = \frac{1}{2}\Phi_0 v_l$ from Ref. 191 in order to compare non-equilibrium experiments (upper half of the table) with equilibrium molecular dynamics simulations (lower half). $v_l = 3.09 \times 10^{-23}$ cm^3 is the volume of a water molecule in the liquid state.

independent of the interface. For our system, the surface tension between liquid water and the wall, $\gamma_{wl} > 0$, and between vapour and the wall, $\gamma_{wv} > 0$, are fixed quantities. The hydrophobicity of the wall implies $\gamma_{wv} < \gamma_{wl}$, i.e. the wall is preferentially in contact with vapour; v_l can be considered constant. Hence, for a given pore of radius R there exists one vapour pressure $p(R) > p_0$ at which vapour and liquid can exist in equilibrium. Water only condenses in the pore if the actual vapour pressure exceeds $p(R)$. Otherwise, only vapour will exist in the pore. The effect is strongest for very narrow pores. Hence a higher pressure is required to overcome the surface contributions, which stabilize the vapour phase in narrow pores. The pressure fluctuates locally in the liquid bulk "reservoir." These fluctuations can provide an increase in pressure above the saturation pressure in the pore and thus drive oscillations between vapour and liquid.

Comparison with experiments, simulations, and a theoretical model Experiments on aquaporins[187,188] and gA[189,190] yield osmotic permeability coefficients of water, p_f, of the order of 10^{-12} to 10^{-14} $cm^3 s^{-1}$. We calculate

Table 3.3: Comparison of different studies of water in hydrophobic pores. The wall-atom density ρ_w is in units of nm^{-3}, the fluid-wall interaction ϵ_{fw} in $\mathrm{kJ \cdot mol^{-1}}$ and the effective interaction strength ϵ_{eff} in $\mathrm{kJ \cdot mol^{-1} \cdot nm^{-3}}$. The critical pore radius R_c is given in nm. The pore length was 0.8 nm in this work, 1.7 nm in the carbon nanotube simulations[197] and infinite in the mean field model.[232]

Ref.	ρ_w	ϵ_{fw}	ϵ_{eff}	R_c
this work	8	0.906493	7	≈ 0.56
197	50	0.478689	24	< 0.24
		0.272937	14	> 0.24
232	110	0	0	1500
		1.4	154	190
		1.45	160	0.35
		2.0	220	0

$p_f = \frac{1}{2}\Phi_0 v_l$ from the equilibrium flux of our MD simulations[191] and find that narrow ($R = 0.35$ nm and 0.4 nm), predominantly "closed" pores have $p_f \approx 5 \times 10^{-14}$ $\mathrm{cm^3 s^{-1}}$, that is, the same magnitude as Aqp1, AqpZ, and gA (see Tab. 3.2 on the previous page). As these pores are longer (ca. 2 nm) and narrower ($R < 0.2$ nm) than our model pores, strategically placed hydrophilic groups[72] seem to be needed to stabilize the liquid state and facilitate water transport in these cases.

Recently Giaya and Thompson[232] presented an analytical mean-field model for water in infinite cylindrical hydrophobic micropores. They predict the existence of a critical radius R_c for the transition from a thermodynamically stable water vapour phase to a liquid phase. The crucial parameter that R_c depends on is the wall-water interaction. We choose the effective fluid-wall interaction $\epsilon_{\mathrm{eff}} = \rho_w \epsilon_{fw}$, the product of the density of wall atoms with the well-depth of the fluid-wall interaction potential, as a parameter to compare different simulations because this seems to be the major component in the analytical fluid-wall interaction. As shown in Tab. 3.3, compared to carbon nanotube simulations our pore has a very small ϵ_{eff} and thus can be considered extremely hydrophobic. This explains why Hummer et al.[197] observe permanently water filled nanotubes with a radius of only 0.24 nm although their bare fluid-wall interaction potential is weaker than in our model.* The much higher density of wall atoms in the nanotube, however, more than mitigates this. Once they lower

*Hummer et al.[197] seem to express some surprise at water readily solvating the interior of carbon nanotubes; their title reads "Water conduction through the *hydrophobic* channel of a carbon nanotube" [emphasis added]. Apart from the considerations presented in this paragraph, it was also shown earlier that carbon nanotubes are rather hydrophilic[233] (similar to graphite) and can be wet by any material with a surface tension smaller than $200\,\mathrm{mN \cdot m^{-1}}$; water has a surface tension of about $72\,\mathrm{mN \cdot m^{-1}}$. Later experiments explicitly showed nanotubes filled by water.[234,235] Carbon in nanotubes is in an $\mathrm{sp^2}$ hybridisation state (as in graphite), which leads to π orbitals extending out of the surface. These are probably the source of the comparatively high hydrophilicity of graphite and carbon nanotubes.

their ϵ_{eff} to double of our value, they also observe strong evaporation. This suggests that the close packing of wall atoms within a nanotube may result in behaviour not seen in biological pores. The mean field model agrees qualitatively with the simulations as it also shows a sharp transition and high sensitivity to ϵ_{eff}.

The force field employed here and in Ref. 197 did not include polarisation effects. In the light of the preceding discussion one might anticipate that additional attractive interactions from the polarizability of the pore wall should have an influence on the liquid-vapour equilibrium. This is the case, as Hansen and coworkers showed in simulations of water in model pores which included a polarizable pore wall.[198,236] Although the effect on structural properties of the confined water is small, there is a marked shift in the critical radius with increasing dielectric constant of the wall. The net-effect of polarizable walls appears to be an increase in the effective wall-water interaction.

3.4 Conclusions

We have described oscillations between vapour and liquid states in short ($L_P = 0.8$ nm), hydrophobic pores of varying radius (0.35 nm$\leq R \leq$ 1.0 nm). Qualitatively, this behaviour is explained as capillary evaporation, driven by pressure/density fluctuations in the water "reservoir" outside the pore. Similar behaviour is found in simulations by different authors with different water models [SPC (this work), SPC/E,[198,202,236] TIP3P (data not shown), TIP3P,[197] TIP4P[199]] in different nanopores [atomistic flexible models (this work), carbon nanotubes,[197] spherical cavities,[1] and smooth cylinders[198,199,202]].

We presented a radically simplified model for a nanopore that is perhaps more hydrophobic than real proteins (although we note the existence of a hydrophobic pore in the MscS channel[55] which has been shown to exhibit dewetting in recent MD simulations[56]). From comparison with experimental data on permeability we conclude that strategically placed hydrophilic groups are essential for the functioning of protein pores. The comparatively high permeability of our "closed" pores suggests pulsed water transport as one possible mechanism in biological water pores. Local hydrophobic environments in pores may promote pulsatory collective transport and hence rapid water and solute translocation.

Our results indicate new, intrinsically collective dynamic behaviour not seen on simulation time scales currently considered sufficient in biophysical simulations. These phase oscillations in simple pores—a manifestation of the hydrophobic effect—require more than 50 ns of trajectory data to yield a coherent picture over a free energy range of 10 kT. We thus cannot safely assume that the behaviour of water within complex biological pores may be determined by

extrapolation from our knowledge of the bulk state or short simulations alone.

4 Hydrophobic gating

A hydrophobic constriction site is thought to act as an efficient barrier to ion and water permeation if its diameter is less than the diameter of an ion's first hydration shell (hydrophobic gating). Simplified pore models allow the investigation of the primary characteristics of a conduction pathway, namely its geometry, the chemical character of the pore wall surface, and its local flexibility and surface roughness. Extended molecular dynamic simulations show that a short hydrophobic pore is closed to water for radii smaller than 0.45 nm. By increasing the polarity of the pore wall (and thus reducing its hydrophobicity) closed pores can be opened until for hydrophilic pores liquid water is stable even in single-file pores. Ions behave similarly but the transition from conducting to non-conducting pores is even steeper and occurs at a radius of 0.65 nm for hydrophobic pores. The presence of water vapour in a constriction zone indicates a barrier for ion permeation. A thermodynamic model can explain the behaviour of water in nanopores in terms of the surface tensions, which leads to a simple measure of hydrophobicity—the classical contact angle. Increased local flexibility decreases the permeability of polar species and so does an increase in temperature. Both effects can be explained by a decrease in the effective solvent-surface attraction which in turn leads to an increase in the solvent-wall surface tension.

4.1 Geometry, surface character, and local flexibility

Based on a number of atomic resolution structures of ion channels[16,27–29,48,51,55,237] and aquaporins[69–71,238] an atomistic understanding of the transport and gating properties of these proteins is emerging. If we view transport proteins as molecular machines which are designed by evolution to perform selective and efficient transport that can be controlled, i.e. gated, then we can ask what the "building blocks" of these machines are. This involves not only identification and structural characterisation of the protein domains involved in gating, but also, at a more abstract level, understanding

the underlying physical principles.

Here we address the questions of what are physical properties that are important for the permeation of ions and water, and what is their effect? In particular we will focus on how the flow of ions and water can be controlled, i.e. we will investigate possible gating mechanisms. This interest is based on the observation that the putative gate in many known ion channel structures (in particular, the nicotinic acetylcholine receptor nAChR,[29] the bacterial potassium channels KcsA[16] and KirBac1.1,[48] and the mechanosensitive channels MscL[51] and MscS[55]) is formed by a constriction made from hydrophobic residues. For example, the closed state nAChR structure displays an ion pathway which is still wide enough (radius $R \approx 0.31$ nm) to admit three water molecules (radius of a water molecule $r_w = 0.14$ nm) or one potassium ion ($r_{K^+} = 0.133$ nm) with half of its first hydration shell intact. It is somewhat surprising that a pore does not have to be completely physically occluded to prevent the flow of ions. On the other hand ions can readily move through the KcsA[16,27] selectivity filter although its radius is less than 0.15 nm. When an ion enters the filter it has to shed its hydration shell at a high cost in free energy—the solvation free energy for a potassium ion is about $-320\,\mathrm{kJ \cdot mol^{-1}}$ (Ref. 239). The filter is lined by backbone oxygen atoms which coordinate the potassium ion and substitute for its hydration shell, thus reducing the desolvation barrier of $+320\,\mathrm{kJ \cdot mol^{-1}}$ to about $+12\,\mathrm{kJ \cdot mol^{-1}}$ (Ref. 240, 241). The putative gates differ from the selectivity filter in that they are lined by *hydrophobic* residues. For nAChR[29,65] and MscL[51,66] it was already hypothesised that these residues cannot substitute for water molecules so that the energetic cost of desolvation prevents the passage of the ion. Though not specifically in the context of gating, Brownian dynamics simulations (BD), carried out by Chung et al.[64], showed that ions would not traverse a non-polar $R = 0.4$ nm pore unless the polarity of the pore wall was enhanced by embedding dipoles. The authors acknowledged the problem of representing water as a continuum in the constriction site in BD, which supports the use of fully atomistic molecular dynamics simulations in the following.

It appears that the *geometry* of a pore (its radius, length and shape) and the *chemical character of the pore wall* have a great influence on the permeation of ions and water. In addition, *local flexibility* (i.e. fluctuations in the protein structure as opposed to concerted larger scale motions) of the pore lining might play a role, as seen in simulations of K^+ permeation through KcsA.[240]

It seems difficult to comprehend the dynamical nature of transport phenomena from static crystal structures alone. Computational methods can be used to complement the experimentally observed picture. In particular, classical molecular dynamics (MD) can be used to investigate the behaviour of water or ionic solutions in the environment presented by a protein. This realistic environment, however, makes it difficult to disentangle the contributions of

various pore properties. In order to reduce the number of parameters simplified pore models can be designed to capture the characteristics in question.[62,96,195,196,198,202,242] We focus on the influence of geometry, pore wall character and flexibility. In the last chapter we investigated the behaviour of water in these "nanopores". In particular, we found that below a critical radius liquid water becomes unstable in the pore and the pore is predominantly filled with water vapour.[96] We hypothesised that a pore environment which cannot sustain liquid water would also present a high energetic barrier to an ion.[62] Although it seems plausible that absence of water would imply the absence of ions it has not been demonstrated previously.

Here we explicitly consider ions in model pores and explore dimensions, pore wall character, and flexibility in more detail than in the last chapter. We also present a simple thermodynamic model based on surface energies that explains the observed behaviour of water in hydrophobic pores.

4.2 Methods and Theory

First, the setup of the model system and the parameters of the MD simulations are discussed. Then the state based analysis, that was already used implicitly in the last chapter, is formally introduced. Finally, a simple thermodynamic model for liquid-vapour equilibrium in cylindrical pores, based on surface tension arguments, is presented. It is applied in the Results section to explain the simulation data.

4.2.1 Model

The pore models were constructed as described Section 3.2.1. Briefly, they consist of concentric rings of methane-like pseudo-atoms of van der Waals radius 0.195 nm (Fig. 4.1 on the next page). These are held in their equilibrium position by harmonic restraints with spring constant $k_0 = 1000\,\mathrm{kJ \cdot mol^{-1} \cdot nm^{-2}}$. A pore consists of two mouth regions (length 0.4 nm, i.e. one layer of pseudo-atoms, and radius 1.0 nm) at either end of the constriction site of length L (varied between 0.4 nm and 2.0 nm) and radius R (varied between 0.15 nm and 1.0 nm). The pore was embedded in a membrane mimetic, a slab of pseudo-atoms held on a cubic lattice with unit cell length 0.39 nm with harmonic springs of strength k_0. Pores with a polar surface were created by placing partial charges of $\pm 0.38e$ on atomic sites 0.2 nm apart. The resulting dipole moments pointed parallel to the pore axis with a magnitude of 3.6 D each, which is comparable to the dipole moment of the peptide bond (ca. 3.7 D, Ref. 243).

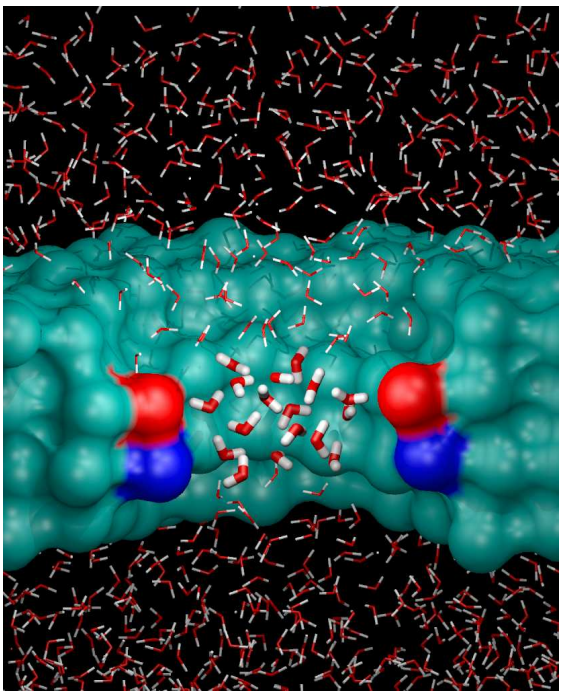


Figure 4.1: Simulation system: The solvent-accessible surface of a model pore with pore radius $R = 0.55$ nm is shown in light blue (parts of the pore close to the observer are removed for clarity). The pore is not completely hydrophobic due to two dipoles (red/blue) in the pore wall which have the same magnitude as the peptide bond dipole. Water molecules in the pore are depicted in licorice representation and as lines in the mouth region or the bulk (image created with vmd[43] and Raster3D[44]).

4.2.2 Molecular Dynamics

MD simulations were performed with gromacs version 3.1.4 (Ref. 127) and the SPC water model.[132] Parameters for the unified methane atoms and ions are described in Section 2.2.3 on page 38. The integration time step was 2 fs and coordinates were saved every 2 ps. With periodic boundary conditions, long range electrostatic interactions were computed with a particle mesh Ewald method (real space cutoff 1 nm, grid spacing 0.15 nm, 4th order interpolation[164,165]) while the short range van der Waals forces were calculated within a radius of 1 nm. The neighbour list (radius 1 nm) was updated every 10 steps. Weak coupling algorithms[166] were used to simulate at constant temperature ($T = 300$ K, time constant 0.1 ps) and pressure ($P = 1$ bar, compressibility 4.5×10^{-5} bar^{-1}, time constant 1 ps) with the x and y dimensions of the simulation cell held fixed at 4 nm. The total thickness of the water reservoir in the z-direction was 3.0 nm, ensuring bulk-like water behaviour far from the membrane mimetic. The initial system configuration was created by solvating the slab-embedded pore model with water. For simulations with ions, some water molecules were replaced with ions to reach the target concentration. Initially there were always ions present in pores with $R \geq 0.4$ nm. A typical simulation box measured $3.9 \times 3.9 \times 4.6$ nm^3 and contained about 1500 water molecules and 280 pseudo-atoms (and between 25 to 28 Na^+ and Cl^- ions each).

4.2.3 State based analysis

We define discrete states by mapping equilibrium states of the whole system (snapshots from the MD equilibrium trajectory) onto numbers ω_i. The same approach is used in statistical mechanics to map different microscopic states to one macroscopic state, characterised by the value of a state variable, here called ω. This approach is used to label the two phase states that the pore water exhibits in our simulations. We either find liquid-filled pores or vapour-filled ones. Due to the small pore volumes "vapour" typically refers to zero or one water molecule in the cavity. We use the density in the pore as an indicator of the phase state. We assign the state using a Schmitt-trigger procedure[204] in order to avoid spurious state changes due to fluctuations from the interfacial region. When the water density $n(t)$ rises above 0.65 of the density of bulk water $n_0 = 1.0\ \mathrm{g \cdot cm^{-1}}$ (at $T = 300$ K and $P = 1$ bar) the liquid state, i.e. $\omega = 1$, is assigned to the phase state at time t. When $n(t)$ drops below $0.25\, n_0$ the vapour state ($\omega = 0$) is assigned. The pore in the liquid-filled state is termed "open" because our simulations show that significant amounts of water pass through it; furthermore, we also demonstrate in this work that a pore that sustains a liquid environment potentially allows ions to permeate. A vapour-plugged pore, however, will prevent ion permeation and is said to be "closed".

Liquid-filled pores and vapour-filled pores are assumed to be in equilibrium. The "openness" or probability for the occurrence of the liquid, i.e. open, state is

$$\langle \omega \rangle = \frac{1}{\mathfrak{T}} \int_0^{\mathfrak{T}} dt\, \omega(t) = \frac{\mathfrak{T}_\mathrm{o}}{\mathfrak{T}}, \tag{4.1}$$

where $\mathfrak{T}$ denotes the total simulation time, $\mathfrak{T}_\mathrm{o}$ the total time that the pore is *open*, whereas $\mathfrak{T}_\mathrm{c}$ is the total time in the vapour or closed state. If $\langle \omega \rangle > \frac{1}{2}$ then equilibrium is on the side of liquid (and vapour is a meta stable state), otherwise the stable phase state is vapour. The equilibrium is governed by the equilibrium constant

$$K(R) = \frac{\mathfrak{T}_\mathrm{c}(R)}{\mathfrak{T}_\mathrm{o}(R)} = \frac{\mathfrak{T} - \mathfrak{T}_\mathrm{o}(R)}{\mathfrak{T}_\mathrm{o}(R)} = \langle \omega(R) \rangle^{-1} - 1, \tag{4.2}$$

which is trivially related to the openness. (Note that for $\mathfrak{T}_\mathrm{o} = 0$ or $\mathfrak{T}_\mathrm{c} = 0$ it is meaningless to compute K because it indicates that the simulation time was too short to sample any state changes.)

4.2.4 Model for liquid-vapour equilibrium in pores

We investigate a subsystem of the whole simulation system comprising of the pore of volume $V = L\pi R^2$. The subsystem can exchange water molecules with the bulk water outside the pore, which acts as a particle reservoir at average

chemical potential μ. μ is implicitly determined by the constant average density of water in the bulk system, which in turn is prescribed by simulating at constant pressure. Hence we can use the grand potential $\Omega(T, V, \mu) = -pV$ as the free energy to describe the subsystem in the closed and open state.[202] The equilibrium constant K is determined by the free energy difference between the closed (vapour) and open (liquid) state,

$$\beta\,\Delta\Omega(R) = \beta\,[\Omega_v(R) - \Omega_l(R)] = -\ln K(R) = -\ln\big(\langle\omega(R)\rangle^{-1} - 1\big). \qquad (4.3)$$

Following Evans[207], we write the free energy difference between the vapour and the liquid state with the corresponding surface contributions as

$$\begin{aligned}\Delta\Omega(R) &= \Omega_v(R) - \Omega_l(R) = \\ &- p_v(T,\mu)\,L\,\pi R^2 + 2\pi R\,L\,\gamma_{vw} + 2\pi R^2\gamma_{lv} - \big(-p_l(T,\mu)L\,\pi R^2 + 2\pi R\,L\,\gamma_{lw}\big).\end{aligned} \qquad (4.4)$$

Here the index w indicates the solid pore wall, l the liquid and v the vapour phase; for instance γ_{lv} is the surface tension or surface free energy per unit area of the liquid-vapour interface. The system is fairly close to bulk phase coexistence so we can expand the pressure p around the saturation chemical potential $\mu_{\text{sat}}(T)$ in a Taylor series

$$p(T,\mu) = p(T,\mu_{\text{sat}}) + (\mu - \mu_{\text{sat}})\,\frac{\partial p(T,\mu)}{\partial\mu}\bigg|_{\mu=\mu_{\text{sat}}} + \cdots,$$

where at saturation $p_l(T,\mu_{\text{sat}}) = p_v(T,\mu_{\text{sat}})$. This leads to a simple parabolic form for the free energy difference between the two states:

$$\begin{aligned}\Delta\Omega(R) &= \\ &\Big[2\gamma_{lv} - \big(\mu - \mu_{\text{sat}}\big)\big(n_v(T,\mu_{\text{sat}}) - n_l(T,\mu_{\text{sat}})\big)L\Big]\pi R^2 + 2\pi L(\gamma_{vw} - \gamma_{lw})R \\ &= 2\Big[\gamma_{lv} + \frac{1}{2}\Delta\mu\,\Delta n_{vl}\,L\Big]\pi\,R^2 + 2\pi L\,\Delta\gamma_w\,R,\end{aligned} \qquad (4.5)$$

where we define the distance of the state from saturation $\Delta\mu := \mu - \mu_{\text{sat}}$, the difference in densities $\Delta n_{vl} := n_l - n_v$ at saturation, and the difference in surface free energies of the two phases with the wall, $\Delta\gamma_w := \gamma_{vw} - \gamma_{lw}$. The term $\Delta\mu\Delta n_{vl}\,L$ is small for $L < 10$ nm as the system is close to phase coexistence (for $L \approx 1$ nm it is about 10^{-3} times smaller than $\gamma_{lv} = 17\,kT\,\text{nm}^{-2}$ when estimated from $\Delta n_{vl}\,\Delta\mu\,L \approx \Delta P\,L \approx 1\text{ bar} \times L = 2.4 \times 10^{-2}\ kT\ \text{nm}^{-3} \times L$, $T = 300$ K) and will be neglected. Only the difference between the surface free energies enters the model so we express it as the contact angle θ_e, using the macroscopic

definition from the Young equation $\gamma_{vw} - \gamma_{lw} = \gamma_{lv} \cos\theta_e$ (see, for instance, Ref. 92). Then Eq. (4.5) becomes

$$\Delta\Omega(R, L, \theta_e) = 2\pi R \gamma_{lv}(R + L \cos\theta_e). \tag{4.6}$$

(Eq. (4.6) is similar to the simple model derived by Allen et al.[202] but it includes $\gamma_{vw} - \gamma_{lw}$ and hence θ_e instead of just γ_{lw}.) For fixed pore length L and a given pore material, characterised by θ_e, the graph of $\Delta\Omega(R)$ over the pore radius R describes a parabola containing the origin. Although it is not *a priori* obvious that such a macroscopic treatment (Eq. (4.5) or Eq. (4.6)) of a nanoscale system is meaningful, the MD results for the free energy difference presented in Section 4.3.1 can be adequately explained within this model.

The connection between the model free energy $\Delta\Omega(R)$ and the behaviour of the system as observed in MD simulations, i.e. the openness $\langle\omega(R)\rangle$, is established by inverting Eq. (4.3), yielding

$$\langle\omega(R)\rangle = \frac{1}{1 + \exp[-\beta\,\Delta\Omega(R)]}. \tag{4.7}$$

For linear $\Delta\Omega(R)$ this represents a sigmoidal curve but non-linear terms (as in Eq. (4.5)) change its shape considerably.

4.3 Results and Discussion

First we investigate the behaviour of pure water in nanopores and quantify the influence of geometry, pore surface, and local flexibility on the equilibrium between liquid and vapour filled pores. Then simulations for a NaCl electrolyte are analysed and compared to the pure water case. The intention behind these studies is to explore more fully the ways in which a hydrophobic gate may be opened, and how other factors might modulate such hydrophobic gating. This is particularly important in the context of current models for the gating of nAChR[29,59] that suggest that the transition between the closed and open states of the pore involves both an increase in pore radius and in polarity. We also note that recent discussion of the state (open vs. closed) of MscS have suggested a hydrophobic gating mechanism,[56] similar to that proposed for nAChR.[62]

4.3.1 Pure water

A large number of simulations were run for water in model pores (total simulation time $> 5\ \mu$s) in order to investigate the influence of pore radius (0.15 nm $\leq R \leq$ 1 nm) and pore surface character (hydrophobic vs. amphipathic vs. hydrophilic). Local flexibility (wall atom positional root mean square deviations

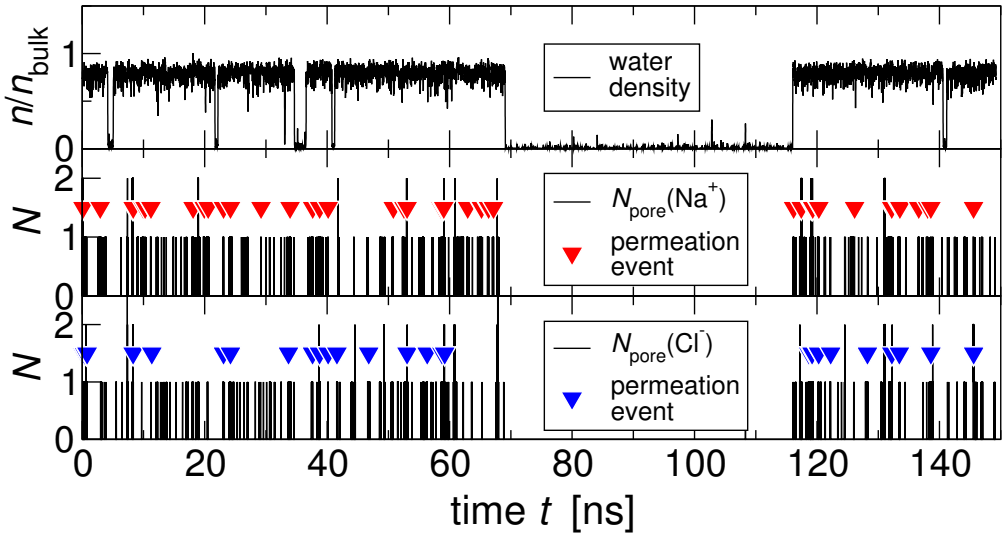


Figure 4.2: Liquid-vapour oscillations of water in a $R = 0.65$ nm apolar pore (top panel) when bathed in a 1.3 mol $\cdot$ l^{-1} NaCl solution. As indicated by the number N of ions within the pore, sodium (middle) and chloride ions (bottom) are only observed in the pore when there is also liquid water ($n/n_{\text{bulk}} \approx 0.8$) present. Permeation events, i.e. an ion entering on one and exiting on the other side, are indicated by triangles; ions do not permeate the pore during the vapour phases.

(RMSDs) 0.039 nm $\leq \rho \leq$ 0.192 nm) and temperature dependence in the range 273 K $\leq T \leq$ 450 K were examined for a hydrophobic $R = 0.55$ nm pore.

Pore dimensions and surface character

The water density in short nanopores oscillates between liquid and vapour on a nanosecond time scale as seen in Fig. 4.2, a manifestation of capillary evaporation and condensation at the nanoscale (see Chapter 3). In Fig. 4.3 the openness and the free energy difference $\Delta\Omega$ between vapour and liquid states is shown for different pore surfaces. For hydrophobic and amphipathic pores a strong dependence of the pore state on the radius is apparent. The stable thermodynamic state switches from vapour ($\Delta\Omega < 0$ or $\langle\omega\rangle < \frac{1}{2}$) to liquid ($\Delta\Omega > 0$ or $\langle\omega\rangle > \frac{1}{2}$) at a critical radius $R_c = -L\cos\theta_e$ (using $\Delta\Omega(R_c) = 0$ in Eq. (4.6)). R_c is 0.56 ± 0.03 nm for the hydrophobic pore and 0.29 ± 0.02 nm for the amphipathic one. The functional form Eq. (4.5) fits the data from the MD simulations (Tab. 4.1) well (the continuous lines in Fig. 4.3 on page 94). The coefficient of the quadratic term, $\gamma_{lv} + \frac{1}{2}\Delta\mu\,\Delta n_{vl}\,L$, is positive and similar for both the hydrophobic and the amphipathic (two dipoles, abbreviated *D2*) pores (Tab. 4.1 on the facing page), consistent with the model Eq. (4.5), which predicts this

Table 4.1: Parameters of the thermodynamic model Eq. (4.5) fitted to the MD results and resulting contact angle θ_e. Experimental values for $\gamma_{lv} = 72 \times 10^{-3}\,\mathrm{J m^{-2}} = 17\,kT\,\mathrm{nm^{-2}}$ and $\theta_e \approx 118^\circ$ for water on a flat methyl (–CH_3) terminated self assembled monolayer[244,245] indicate that the simple thermodynamic model gives the right order of magnitude results. (All values at room temperature.)

pore surface character			$\gamma_{lv} + \frac{1}{2}\Delta\mu\,\Delta n_{vl}\,L$ $[kT\,\mathrm{nm^{-2}}]$	$\Delta\gamma_w = \gamma_{vw} - \gamma_{lw}$ $[kT\,\mathrm{nm^{-2}}]$	θ_e $[^\circ]$
bare CH_4	*0*	"hydrophobic"	$+10\pm1$	-7.2 ± 0.2	134 ± 3
two dipoles	*D2*	"amphipathic"	$+8.2\pm0.5$	-3.0 ± 0.2	111 ± 2
four dipoles	*D4*	"hydrophilic"	$+2.8\pm0.5$	$+1.8\pm0.2$	51 ± 10

coefficient to be independent of the pore wall. For the polar pore (four dipoles, *D4*) the data are more ambiguous. Only if the first two data points at small radii are excluded from the fit as outliers is this coefficient positive (but still three times smaller than the one for the less hydrophilic pores). The outliers show that a high density of local charges leads to a higher probability of the pore being liquid-filled than predicted by the macroscopic model, possibly indicating a shortcoming of the model to subsume inhomogeneous potentials into the surface tension terms. For our system parameters, the coefficient is in fact dominated by the water liquid-vapour surface tension γ_{lv} (see Section 4.2.4 on page 89) and hence we will use its value as an approximation to γ_{lv}. $\Delta\gamma_w$, the difference in surface tensions between the wall and vapour or liquid, becomes more positive with increasing polarity of the pore wall. It effectively measures the hydrophobicity of the wall. This becomes even more apparent when the (macroscopic) contact angle $\cos\theta_e = \Delta\gamma_w/\gamma_{lv}$ is formally computed (Tab. 4.1). Macroscopically, a hydrophobic surface can be defined as one with $\theta_e > 90^\circ$. This allows us to call the apolar pore ($\theta_e = 134^\circ$) "hydrophobic" compared to the "amphipathic" pore *D2* ($\theta_e = 111^\circ$; still hydrophobic but with some "hydrophilic" patches). A "hydrophilic" pore like the *D4* system ($\theta_e \approx 51^\circ$) is characterised by $\Delta\gamma_w > 0$ or $\theta_e < 90^\circ$ and liquid is always the preferred phase in the pore, regardless of R.

Experimental macroscopic contact angles θ_e for water on flat methyl (–CH_3) terminated self assembled monolayers are reported up to 118°.[244,245] A "microscopic contact angle" of $135^\circ \pm 15^\circ$ was calculated for a droplet of 90 water molecules (radius $R \approx 0.75$ nm) on a methyl terminated film from MD simulations,[246] consistent with our result of $\theta_e = 134^\circ$. MD simulations seem to overestimate the experimental equilibrium contact angle. In this context we note that although the macroscopic equilibrium contact angle is only established about 10 nm from the contact line, the local contact angle θ_l in the core struc-

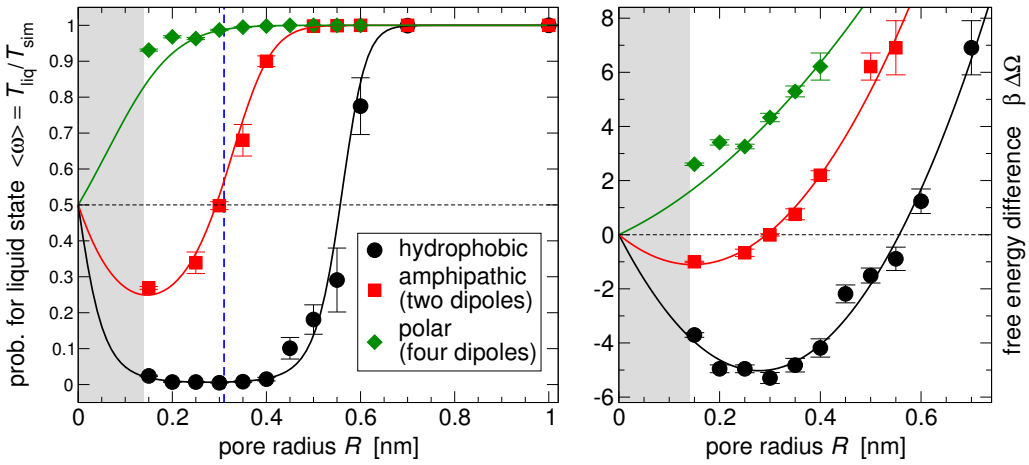


Figure 4.3: Water in model pores. Left: openness $\langle\omega\rangle$ over radius. Right: free energy difference $\Delta\Omega$ (in kT) between vapour and liquid state. The grey region indicates radii smaller than the radius of a water molecule (0.14 nm). Data points are obtained from MD simulations with the errors estimated from block averages[205,206]. The continuous lines are fits of the model Eq. (4.5) to the data points (right) or the openness computed from the model, Eq. (4.7). The vertical broken line indicates the radius of the closed nAChR gate ($R = 0.31$ nm).

ture of the contact line (in the region 3–10 nm from the contact line) in water (and electrolytes) is predicted to be *larger* than θ_e due to electrostatic screening effects.[92]

From the model free energy the openness can be calculated (Eq. (4.7)) and displayed together with the MD results (Fig. 4.3). Because the model effectively treats the liquid as a structureless continuum it is not meaningful to extrapolate to radii smaller than the radius of a water molecule (the shaded region in Fig. 4.3). Nevertheless, the MD results lie on the model curve even in the direct vicinity of this region. Further tests of Eq. (4.6) by varying the length of the pore together with the radius confirm the model only qualitatively. For a fixed radius of 0.55 nm short $L = 0.4$ nm pores are continuously filled whereas longer pores 1.2 nm $\leq L \leq$ 2.0 nm are empty. The filling kinetics become much longer (which has also been pointed out by Waghe et al.[226]), so that 100 ns equilibrium simulations are not sufficient to sample the system.

Our model implies that for both nano- and mesoscale pores the cost of creating the liquid-vapour interface is the only force driving the filling of a hydrophobic ($\Delta\gamma_w \leq 0$) pore. Little free energy $\Delta\mu\Delta n_{vl}\, L\pi R^2$ is gained by creating a bulk-like liquid in the pore instead of vapour.

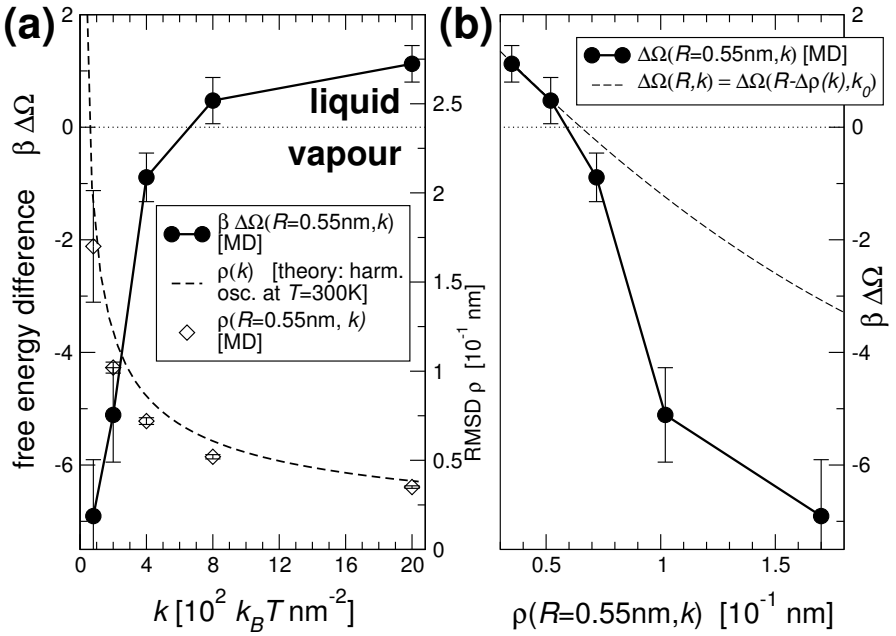


Figure 4.4: Influence of flexibility on the liquid-vapour equilibrium. (a) For a hydrophobic $R = 0.55$ nm pore the harmonic restraint force constant k of the pore wall atoms was varied, resulting in a range of RMSDs ρ, which show the expected behaviour of thermic harmonic oscillators. The free energy difference $\Delta\Omega$ between liquid and vapour states of water in the pore shows strong dependence on the local flexibility ρ; rigid pores contain liquid water whereas flexible ones favour vapour. The effect can not be explained by a reduction of the effective pore radius by ρ alone [broken line in (b)]. All data at $T = 300$ K.

Local flexibility

The influence of local fluctuations in protein structure can be modelled by changing the harmonic spring constant k that holds the pore wall atoms at their equilibrium positions. In a simplified picture of a harmonic oscillator with an average energy $k_\mathrm{B}T/2$ per degree of freedom the RMSD ρ is directly related to k by $\rho(k) = \sqrt{\langle \Delta r_{\max} \rangle^2} = \sqrt{3k_\mathrm{B}T/k}$. This simple model overestimates the measured RMSDs by only 10–20% (Tab. 4.2 on the next page), showing that the wall atoms behave like almost independent, thermally driven, harmonic oscillators (Fig. 4.4(a)). At a fixed radius, increasing the flexibility (smaller k, hence greater ρ) shifts the equilibrium towards vapour (Fig. 4.4). Conversely, a more rigid wall favours the condensation of water in the pore. Thus, water does not "push away" the pore walls to fill the pore but rather fluctuating pore walls appear to disfavour the formation of adjacent water layers. This effect cannot

Table 4.2: RMSD of wall atoms in flexible pore models of radius $R = 0.55$ nm and influence on the liquid-vapour equilibrium of water in the pore at $T = 300$ K.

restraint force constant k $[10^2 \times kT^{-2}]$	RMSD $\rho(k)$ $[10^{-1}$ nm] MD	theory	$\langle\omega\rangle$ MD	$\beta\,\Delta\Omega$ MD
0.8	1.7±0.3	1.92	0.001±0.001	−6.9±1.0
2	1.0±0.0	1.22	0.006±0.005	−5.1±0.8
4	0.72±0.02	0.86	0.291±0.089	−0.89±0.43
8	0.52±0.01	0.61	0.616±0.097	0.47±0.41
20	0.35±0.01	0.39	0.755±0.060	1.1±0.3

be explained with the assumption that the water molecules encounter a more narrow pore "on average". Even if one assumes that the pore is narrowed by the RMSD of an atom down to an effective radius $R(k) = R_0 - \rho(k)$, the effect is still much stronger. When the thermodynamic model and the parameters for the hydrophobic pore (from Tab. 4.1 on page 93 where $k_0 = 4 \times 10^2\ kT\ \mathrm{nm}^{-2}$) are used to predict $\Delta\Omega\big(R(k)\big)$ then the prediction compared to the simulation results underestimates the effect for the more flexible pores but is in agreement for more rigid pores (Fig. 4.4b).

The influence of the wall-water interaction We hypothesise that the random positional fluctuations of the wall atoms change the effective potential between the wall atom and water molecules by "smearing out" the attractive well of the interaction potential. This idea is illustrated by considering the Lennard-Jones potential Eq. (2.61) which is used for the water oxygen-methane interaction. (A generic LJ potential is shown in Fig. 4.5(a) on page 98.) Because in the MD simulation a water molecule only interacts with the methane molecules in the wall it follows that the surface tensions γ_{vw} and γ_{lw} will depend on the well depth ϵ (and to a lesser degree on the radius σ).* The dependence of the surface tension on the well depth ϵ of the interaction potential is explicit in Eq. (2.34) on page 24 combined with Eq. (2.35) (ϵ is one of the parameters λ_i of the potential). The external potential $u(\mathbf{x})$ in (2.34) is the sum of all the Lennard-Jones potentials Eq. (2.61), centred at the equilibrium position $\mathbf{R}_j$ of each wall atom.

$$u(\mathbf{x}) = \sum_{j=1}^{N_{\text{wall}}} V_{LJ}(\mathbf{x} - \mathbf{R}_j)$$

*Of course, as a "collective effect" the surface tensions also depends on the water-water interactions but for constant water-water interactions (i.e. a given water model) the surface tensions only depend on the wall-water interaction.

In order to appreciate the influence of flexibility or temperature we discuss a coarse grained version of the LJ potential. This is necessary because in Section 2.1.2 (in particular Eq. (2.15) and (2.19)) the wall was treated as an external field and hence the movement of the wall atoms was not accounted for. We want to derive a version of (2.61) which incorporates "on average" the effect of fluctuations in the wall, i.e. fluctuations in the positions of the wall atoms. We notice that for a rigid (or "frozen") wall the probability to find an atom near its equilibrium position $\mathbf{x}'$ is

$$p(\mathbf{x};\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}'),$$

i.e. it will only be found at $\mathbf{x}'$. For a harmonicly bound atom at finite temperature $T > 0$ (or $\beta < \infty$) the spatial probability is Gaussian distributed as

$$p(\mathbf{x};\mathbf{x}',\beta) = g(\mathbf{x}-\mathbf{x}',\beta) = \frac{1}{(2\pi)^{3/2}\rho(\beta)} \exp\left[-\frac{1}{2}\left(\frac{|\mathbf{x}-\mathbf{x}'|}{\rho(\beta)}\right)^2\right]$$

where the Gaussian g has a half width equal to the RMSD $\rho(\beta)$. This leads to the coarse grained potential

$$\hat{V}_{LJ}(\mathbf{x},\beta) = g \star V_{LJ} = \int d^3x'\, g(\mathbf{x}-\mathbf{x}',\beta)\, V_{LJ}(\mathbf{x}'). \tag{4.8}$$

Unfortunately, (4.8) diverges due to the divergence at the LJ centre. To overcome this problem we argue that the influence of flexibility is more strongly felt in the slowly varying, attractive component of the LJ potential. Following Weeks, Chandler, and Anderson[123] we decompose V_{LJ} in attractive and repulsive parts (Fig. 4.5(b)),

$$V_{LJ}(\mathbf{x}) = V^{\text{rep}}(\mathbf{x}) + V^{\text{attr}}(\mathbf{x}), \quad \text{with} \tag{4.9}$$

$$V^{\text{rep}}(\mathbf{x}) := \begin{cases} V_{LJ}(\mathbf{x}) + \epsilon & \text{if } |\mathbf{x}| < 2^{1/6}\sigma \\ 0 & \text{if } |\mathbf{x}| \geq 2^{1/6}\sigma \end{cases} \tag{4.10}$$

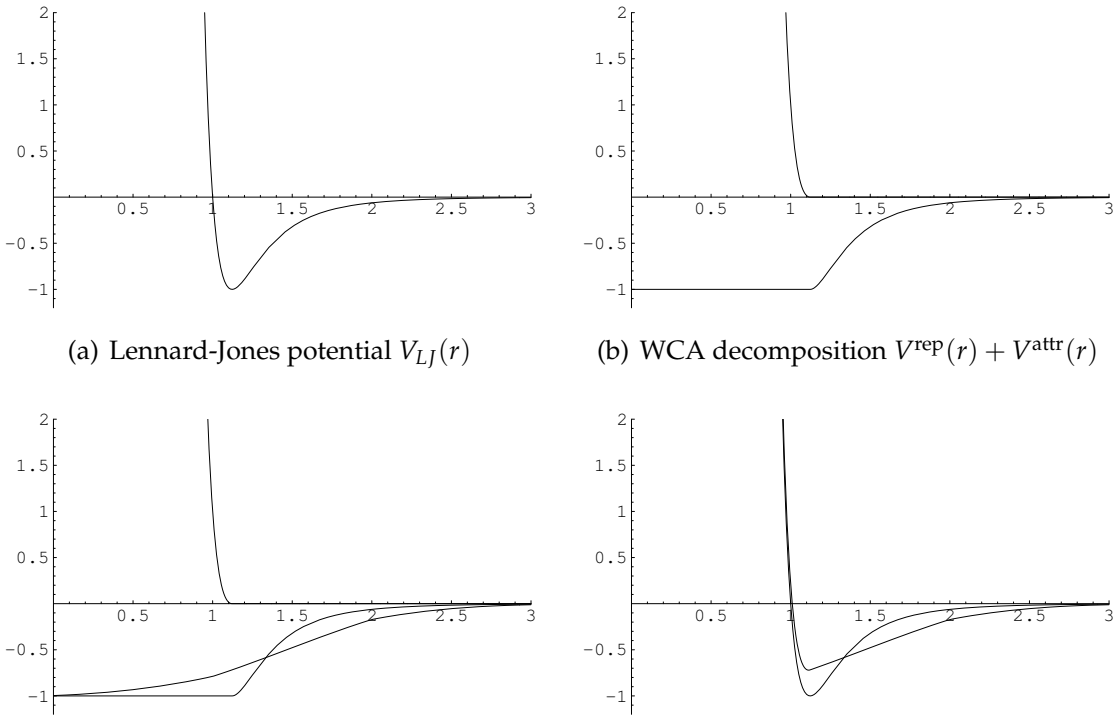
$$V^{\text{attr}}(\mathbf{x}) := \begin{cases} -\epsilon & \text{if } |\mathbf{x}| < 2^{1/6}\sigma \\ V_{LJ}(\mathbf{x}) & \text{if } |\mathbf{x}| \geq 2^{1/6}\sigma \end{cases} \tag{4.11}$$

where the centre of the potential is taken to be at $\mathbf{x} = \mathbf{0}$. Applying the thermal broadening to the attractive component V^{attr} yields

$$\hat{V}^{\text{attr}}(\mathbf{x},\beta) = g(\mathbf{x},\beta) \star V^{\text{attr}}(\mathbf{x}) = \int d^3\mathbf{x}'\, g(\mathbf{x}-\mathbf{x}',\beta)\, V_{LJ}(\mathbf{x}'), \tag{4.12}$$

which is shown for the 1D case in Fig. 4.5(c). Finally, the coarse grained LJ potential is taken as the sum

$$\hat{V}_{LJ}(\mathbf{x},\beta) = V^{\text{rep}}(\mathbf{x}) + \hat{V}^{\text{attr}}(\mathbf{x},\beta). \tag{4.13}$$



(a) Lennard-Jones potential $V_{LJ}(r)$

(b) WCA decomposition $V^{\text{rep}}(r) + V^{\text{attr}}(r)$

(c) Thermally broadened $\hat{V}^{\text{attr}}(r, \beta)$

(d) $V_{LJ}(r)$ and $\hat{V}_{LJ}(r, \beta)$

Figure 4.5: Thermally broadened Lennard-Jones potential. The figures illustrate the process of thermally coarse graining the Lennard-Jones water-wall interaction potential. Only the attractive part of the Weeks-Chandler-Anderson decomposition[123] is broadened by convolution with a thermal Gaussian, $\hat{V}^{\text{attr}}(r, \beta) = g(r, \beta) \star V^{\text{attr}}(r)$. The coarse grained LJ potential $\hat{V}_{LJ}(r, \beta) = V^{\text{rep}}(r) + \hat{V}^{\text{attr}}(r, \beta)$ is shallower than the original (though it is not differentiable at $2^{1/6}\sigma$ anymore). The graphs are only meant to illustrate the idea of thermal broadening of the attractive well; parameters are chosen for convenience ($\epsilon = 1, \sigma = 1, \beta = 1$).

Qualitatively it is obvious that the thermal broadening *decreases* the depth ϵ of the attractive well (Fig. 4.5(d)). As discussed above this leads to an *increase* in the surface tension γ_{lw} between the liquid and the wall.* Eq. (4.13) is not differentiable at the original minimum $|\mathbf{x}| = 2^{1/6}\sigma$ anymore but it suffices to illustrate the idea that motion of the wall atoms will lead to a decrease in well depth.

Evidence in favour of this explanation also comes from MD simulations[247] of a water droplet with 1000 molecules ($R \approx 2.5$ nm) on a graphite-like surface. By varying the well depth ϵ_{C-OW} the contact angle changed from 30° to 180°. Thus it is the well depth of the interaction potential which ultimately determines the thermodynamic equilibrium in the pore. This is made explicit by considering Eq. (4.5) and noting that $\Delta\Omega(R)$ depends on the water-wall surface tensions through $\gamma_{vw} - \gamma_{lw}$, which in turn depend on ϵ. For the case of water in confined geometries it was shown that capillary condensation and evaporation phenomena depend sensitively on ϵ[96,197,226,232] (or more precisely, on the effective well depth,[96] the product of the wall atom density with ϵ as discussed in Section 3.3).

4.3.2 NaCl electrolyte

It has been hypothesised that a local hydrophobic environment would present a significant desolvation barrier to ion permeation. Our simulations of model pores bathed in a 1.3 mol $\cdot$ l^{-1} NaCl solution exhibit a striking change in the behaviour of the ions for pore radii $R \geq 0.65$ nm. Fig. 4.6 on the following page shows that in pores of dimensions comparable to the closed nAChR pore ($R = 0.35$ nm) ions have a vanishing probability of entering the pore, but at the open-state radius $R = 0.65$ nm the pore density rises to half the bulk value. The average density of ions in the centre of the pore (Fig. 4.7 on page 101) also exhibits a sharp increase near $R = 0.65$ nm, similar to the behaviour of the openness (Fig. 4.3 on page 94).

The radially averaged densities show two pronounced water layers with an interlayer distance $d = 0.3$ nm near the membrane mimetic. The ionic density drops off sharply at the maximum of the outermost water layer and vanishes completely at the maximum of the water layer closest to the surface. Sodium and chloride ions are both present at the same distance from the surface. Because the membrane mimetic carries no charge and represents a low dielectric region (as do the less mobile bound water layers) it is energetically more fa-

*The vapour-wall surface tension will also increase but less so because the correlations in Eq. (2.35) vanish over much shorter distance in the vapour as in liquid; furthermore, at typical conditions the vapour density is much smaller than the liquid density, also contributing to a smaller integrand in (2.35).

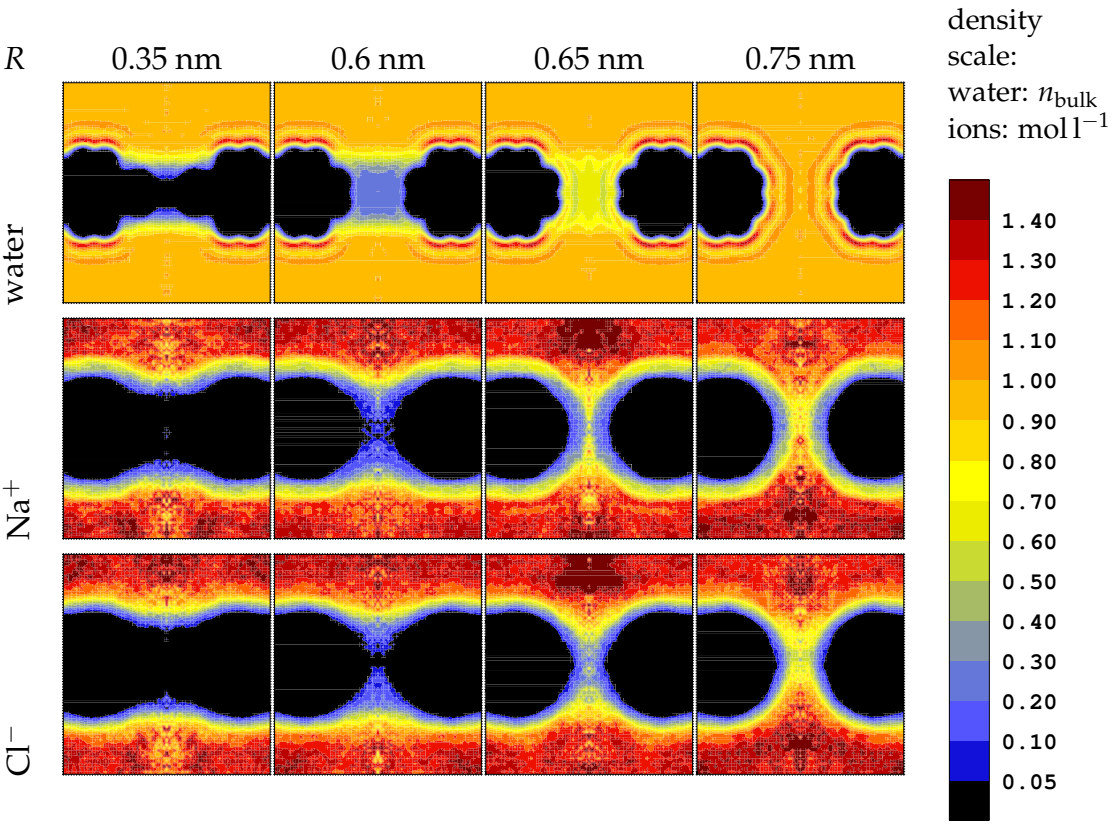


Figure 4.6: Equilibrium densities of water (top row), sodium ions (middle) and chloride ions (bottom) in hydrophobic model pores of radii $R = 0.35$ nm to 0.75 nm; the NaCl bulk concentration is 1.3 mol $\cdot$ l^{-1}. The narrow 0.35 nm pore mimics the *closed* gate of nAChR whereas the wide 0.65 nm pore approximates the *open* channel. Although wide enough to admit the ions physically, the narrow pore is effectively closed to ions ("hydrophobic gating"). The membrane mimetic is located in the horizontal region which appears black, i.e. void of water, in all images.

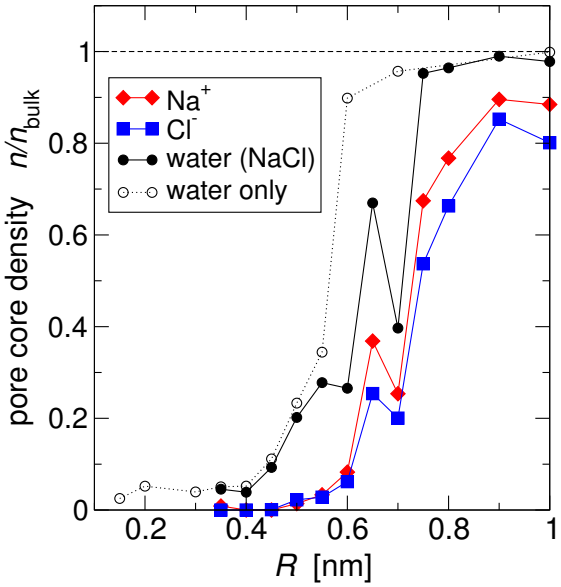


Figure 4.7: Ionic density in the core region of hydrophobic pores. The core is taken to have a radius of about one ionic radius (0.1 nm) and exhibits the highest density of ions. The core density increases markedly at the critical radius $R = 0.65$ nm. The core density follows the openness (as shown for water in Fig. 4.3). The dip in density at $R = 0.7$ nm is due to very few state changes in the 150 ns simulation and a subsequent bias towards a 70 ns vapour period; in general the state switching frequency in the electrolyte systems is reduced compared to the pure water case (see also Fig. 4.2). Bulk concentration of ions is 1.3 mol $\cdot$ l^{-1}. Lines are drawn as guides to the eye.

vourable for the ions to be solvated in the high dielectric region of the bulk-like water. The bound water layers are not a consequence of the presence of the ions as they also form in pure water near a hydrophobic surface with the same interlayer spacing of 0.3 nm (see Fig. 3.3 on page 69 and Fig. 3.4), and in this case even a third water layer is discernible with $d = 0.35$ nm. Water ordering in cylindrical and slit pores in the absence of ions was also observed in simulations by other groups.[195,196,202,210,213,248] Effectively, ions are excluded from a zone of thickness 0.6 nm (Na^+) or 0.7 nm (Cl^-) from the surface. The ion exclusion layer stretches into the pore without interruption so that at $R = 0.65$ nm there is only a narrow channel of easily displaceable water which can admit ions. For wider pores, the ion density in the pore approaches bulk values (Fig. 4.7).

Although the absence of ionic density can be taken as a strong indication for a closed pore, the presence of equilibrium density does not necessarily imply a functionally open state, as the latter also requires rapid permeation of ions through a pore. As MD simulations output the trajectories of water and ions, the equilibrium flux (Φ_0), i.e. the total number of particles per nanosecond which completely permeate the pore, can be measured. Fig. 4.8 on the next page shows Φ_0 for ions and water molecules (both for water in a 1.3 mol $\cdot$ l^{-1} NaCl electrolyte and for pure water) in hydrophobic pores of varying radius on a logarithmic scale. Water flux is about three orders of magnitude larger than ion flux and increases rapidly with increasing radii as long as R is smaller than the critical radius $R_c = 0.62$ nm. Beyond R_c (when liquid becomes the stable phase) the slope in the logarithmic plot decreases from 15 nm^{-1} to 2.4 nm^{-1}. The water current density $\Phi_0/(\pi R^2)$ reaches a constant value of 300 ns^{-1} nm^{-2} so the increase for $R > R_c$ is only due to the increase in pore diameter. The

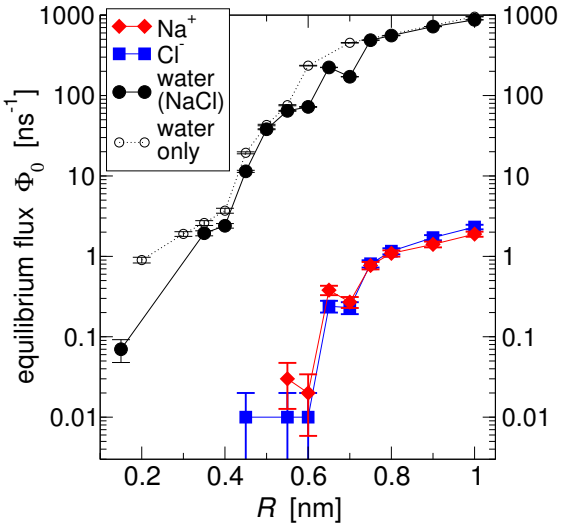


Figure 4.8: Equilibrium flux Φ_0 of water and ions through hydrophobic model pores. The total number of ions that successfully permeate the pore per nanosecond shows an increase over almost two orders of magnitude when the pore radius R is increased from 0.6 nm to 0.65 nm. Water flux increases with the radius. The most notable influence of the 1.3 mol $\cdot$ l^{-1} NaCl electrolyte solution is the shift in the transition radius from 0.6 nm (only water) to 0.75 nm (water with NaCl). Lines are drawn to guide the eye and errors are estimated from block averages.[205,206]

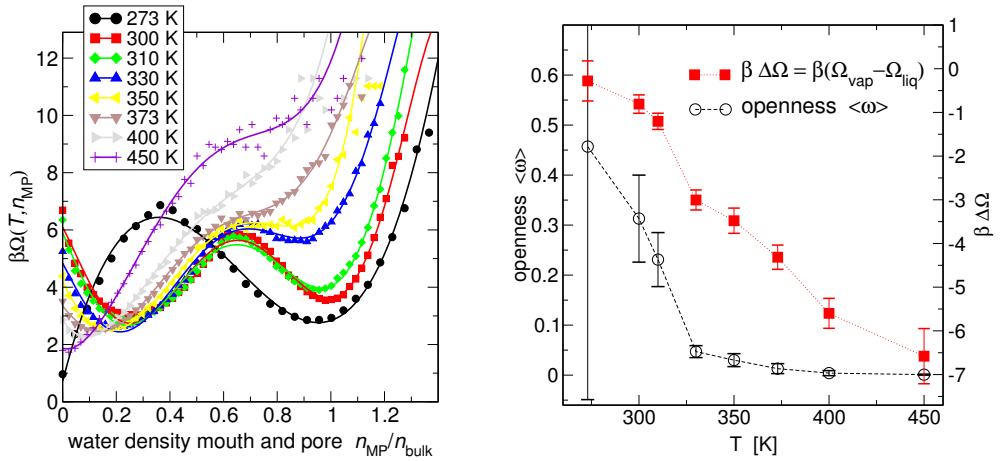
increase for $R < R_c$ correlates with the openness (Fig. 4.3 on page 94), indicating that the main contribution to the flux stems from the open-state periods (except for the very narrow pores, as described on page 80). The presence of ions slightly shifts R_c from 0.56 ± 0.03 nm for pure water to 0.62 ± 0.07 nm for 1.3 mol $\cdot$ l^{-1} NaCl electrolyte. Qualitatively, this is explained by the increase of surface tension with ionic concentration.[82] This leads to a higher contact angle θ_e,[249] which is also observed in our simulations where it increases from 134° to 140°. Hence, $R_c = -L \cos\theta_e$ is expected to increase for electrolytes compared to pure water, as hinted at by the simulations.

The total number of ions that successfully permeate the pore per nanosecond shows an increase over almost two orders of magnitude when the pore radius R is increased from 0.6 nm to 0.65 nm. Thus the pore "opens" at a radius much larger than the bare ionic radius. Ions permeate only when the pore is filled with liquid water (typical data in Fig. 4.2 on page 92) and Na^+ permeation events are not correlated with Cl^- ions passing through the pore.

Because the simulations are in thermodynamic equilibrium the net flux through the pore is zero for all R. In order to compare the MD equilibrium flux Φ_0, which is determined by the intrinsic free energy barrier to ion permeation, to experimentally measured non-equilibrium fluxes Φ of ions of charge q at a driving transmembrane voltage V we employ rate theory and estimate

$$\Phi(V;R) = \Phi_0(R) \sinh \frac{qV}{2kT}. \tag{4.14}$$

For $R = 0.65$ nm, $V = 100$ mV, $q = 1e$ and $T = 300$ K we obtain $\Phi \approx 1.2$ ns^{-1}, which is 40 times larger than the experimental value of 0.03 ns^{-1} for nAChR at 0.2 mol $\cdot$ l^{-1} ionic concentration (calculated from a conductance of ca. 45 pS (Ref. 12)). The discrepancy between our estimate and the experimental data is



(a) Change of the free energy landscape $\Omega(n, T)$ with T.

(b) Openness and free energy difference between states $\Delta\Omega$.

Figure 4.9: Temperature dependence of the liquid-vapour equilibrium in a hydrophobic $R = 0.55$ nm pore. In (a) the density is calculated over the pore and the mouth regions which act as a buffer of constant chemical potential except for $T = 273\ K$, where the mouth region also fills with vapour. Continuous lines are polynomial least square fits of order 4 (or order 6 for 300 K $\leq T \leq$ 330 K), reminiscent of a Landau free energy with the water density in the pore and mouth region n_{MP} as order parameter. The critical point of SPC water is near $T = 373$ K.

not unexpected as Eq. (4.14) has only qualitative character. Furthermore, nAChR presents a more complex pore lining surface than the model pore, and if the difference in bath concentrations and access resistance was taken into account the discrepancy would likely be reduced by an order of magnitude. Nevertheless, the estimate demonstrates that the intrinsic barrier to ion permeation is small beyond the critical radius.

4.3.3 Sensing external parameters

In the simplest model of sensing a system exists in an equilibrium with two states. External stimuli shift the equilibrium which elicits a cellular response. For example, a temperature-sensing channel at normal temperature might be in the closed state. Elevated temperature shifts the equilibrium to the open state and the influx of ions initiates a signalling cascade that terminates in the sensation of heat. However, the exact mechanisms of gating by temperature of e.g. TRP channels remain obscure,[21] though it is now known that TRP channels are voltage gated[250] and temperature shifts their voltage-dependent activation

curves.*

We tested the influence of temperature in the range from 273 K to 450 K on water in the $R = 0.55$ nm hydrophobic pore (data points for $T > 373$ K represent "supercritical" water because the critical point for the SPC water model is close to $T = 373$ K). The free energy landscape of the water-pore system depends on the temperature (Fig. 4.9(a) on the previous page). With increasing temperature the liquid state (i.e. near water bulk density) decreases in stability. Accordingly, the openness decreases with temperature so that the vapour state, i.e. the closed stated, dominates at higher temperatures (Fig. 4.9(b)). The decrease in openness with increasing T indicates that the difference of the wall surface tensions $\Delta\gamma_w$ must become even more negative, i.e. vapour in contact with the wall is increasingly favoured. The effect would be similar to the one discussed for increased flexibility of wall atoms as the major effect of temperature would be to smear out the interaction potential between water molecules and wall atoms. This example demonstrates in principle how an external signal, an increase in temperature from $T = 300$ K to 330 K, decreases the openness, and correspondingly the flux of water molecules, by a factor of six. Voltage-dependent TRP channels are probably not directly gated in this manner. Interestingly, their behaviour is described by a two-state model where the activation energies for the open-to-closed and closed-to-open transitions differ very much[250]—similar to a pore predominantly liquid or vapour filled. Perhaps the precarious stability of liquid water in hydrophobic cavities is used as a temperature sensor, e.g. by solvation and desolvation of voltage sensor domains and the change in dielectric properties.

4.4 Conclusions

We have explicitly demonstrated a hydrophobic gating mechanism for ions, using simplified hydrophobic pores as models for the closed gates of ion channels. There is a critical radius R_c above which a pore becomes effectively permeable to water or ions. $R_c \approx 0.56$ nm for water and $R_c \approx 0.65$ nm for ions, is much larger than the radius for a water molecule or a bare ion alone. This correlates nicely with the current view that the closed state structures of many ion channels contain gates formed by hydrophobic constriction sites. In nAChR the radius of the putative gate is 0.31 nm, 0.17 nm in MscL, and 0.13 nm in KcsA. Models of KcsA and of Kv channels in their open states, based on the MthK and KvAP structures respectively, have gate radii of about 0.6 nm (Ref. 251), whereas

*The voltage dependent activation curves of TRP channels are also shifted by ligands such as capsaicin (the "hot" ingredient of chilli peppers) or menthol.[250] Thus, the sensation of heat or cold elicited by some food is the same sensation that is felt when the temperature is increased or decreased.

the open-state nAChR structure opens up to about 0.65 nm. Our simulations show that the polarity of the pore wall can shift the critical radius considerably. This is reflected in the proposed gating mechanism of nAChR.[59] Not only does the radius of the constriction site increase but hydrophobic sidechains are also rotated out of the pore to expose the more polar, i.e. hydrophilic, peptide backbone. By combining hydrophobic gating with a change in surface polarity only a moderate change in radius is required to obtain a large physiological effect. Changes in local flexibility may also modulate the gating behaviour. It is too early to ascertain their importance as such changes in flexibility have not been widely investigated experimentally for channels (or any other proteins). There are suggestions of possible regulatory roles of changes in flexibility in e.g. KTN domains,[252] and in the modulation of binding of TRAP protein to the Trp operon mRNA.[253] Temperature can also affect gating, in a fashion similar to local flexibility, but temperature sensitive voltage-dependent TRP channels are most likely not directly gated by temperature in the manner described here, though water-solvation of the voltage sensors—and hence their dielectric environment—could be temperature dependent.

A thermodynamic model based on surface energies fits remarkably well the data of the atomic-scale MD simulation, even for very small radii. Though such a macroscopic treatment is not *a priori* expected to give a satisfactory description of a microscopic system there are other examples as for instance classical nucleation theory, which shares some similarities with our and others' models[94,202]. It generally agrees quite well with experiments on the condensation of droplets from vapour, which implies that the use of macroscopic surface tension is valid even for droplets of radii of about 1 nm (Ref. 113). It appears that in this case water structure (i.e. a hydrogen bond network) is only important insofar that it is responsible for different wall-fluid surface tensions. Also, Nicholls et al.[254] find in their investigations of the interfacial and thermodynamic properties of hydrocarbons that the "macroscopic concept of interfacial free energy is applicable at the molecular level." Our results seem to corroborate the conclusions of Maibaum and Chandler[255] that two-state behaviour of water in pores should only require a cold liquid close to phase coexistence and sufficiently different vapour-wall and liquid-wall surface tensions (which is, of course, where the "special" properties of water reside in such a model).

We have sketched out principles of gating mechanisms in ion channels, based on model channels. Ion channels in nature display complex conformational dynamics in relationship to gating. There is a need for better single molecule methods for ion channels to probe these phenomena experimentally. Combined electrical and optical methods look promising in this respect[256] although considerable improvements in time resolution are still required. In the meantime there is a continued role for simulations and theory to enable us to bridge between static structure and dynamic function.

In addition to hydrophobic gating *per se*, our studies suggest the possible importance of gate flexibility in regulation of biological activity. The relationship between static structures, conformational change, and intrinsic flexibility in relationship to proteins and signalling merits further active investigation. In a recent review of protein-protein interactions and conformational changes, Goh et al.[257] conclude that there is increasing support for a pre-existing equilibrium model. In such a model, proteins exist in a population of conformations, with ligand binding leading to a change in the probability distribution of the ensemble. Some evidence in support of this have been obtained from recent simulation studies of ligand binding proteins (e.g. Ref. 258) but further (single-molecule) experimental and computational studies of a wider range of proteins are needed to more firmly establish the general importance of changes in flexibility.

5 The dielectric barrier and the hydrophobic effect

A hydrophobic pore of sub-nanometre dimensions can appear impermeable to an ion even though its radius is still much wider than the one of the ion. The barrier to ion permeation is quantified by calculating the potential of mean force (PMF) from umbrella-sampled molecular dynamics simulations. It is compared to continuum-electrostatic Poisson-Boltzmann calculations. The latter fail to describe the ion barrier because they do not account for the properties of the confined water. The barrier originates from the energetic cost to desolvate the ion. Even in wide pores, which could accommodate an ion and its hydration shell, a barrier of several kT remains because the liquid water phase is not stable in the hydrophobic pore. Thus, the properties of the solvent play a crucial role in determining permeation properties of ions in confinement at the molecular scale. For water, the PMF depends strongly on the wall flexibility. With increasing pore length the barrier appears to reach a constant value.

5.1 A closer look at ions (and water) in hydrophobic pores

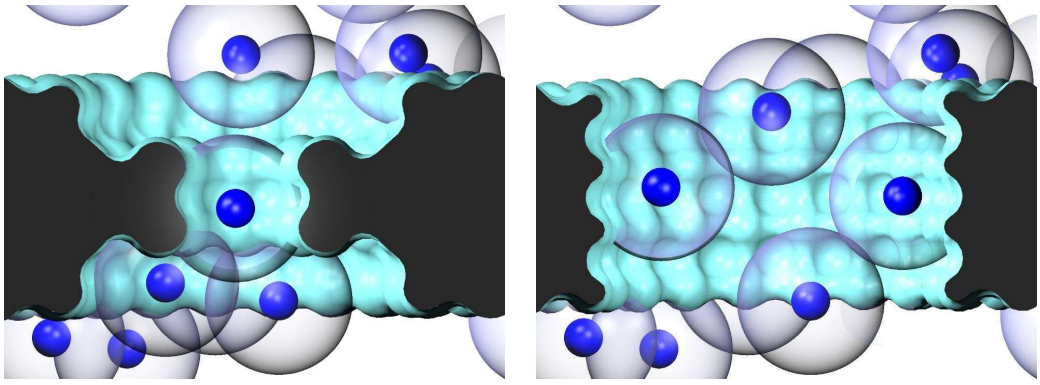
Confinement strongly affects the behaviour of water and ions and can lead to effects which are not anticipated from macroscopic descriptions.[96,193,194,197,218,234,248,259–261] One example is the exclusion of ions from pores that have radii much larger than the ionic radius as shown in the last chapter. Pores of molecular dimensions (radius $R \leq 1$ nm) can be found, for instance, in carbon nanotubes, zeolites, and ion channel proteins. In this chapter we show that an electrostatic continuum approach alone does not capture the essentials of ion permeation through pores in a low dielectric. A more detailed atomistic approach is required that incorporates both the interaction of the water molecules with the ion (the hydration shell) *and* the water-pore interaction (hydrophobic effects).

The atomistic molecular dynamics (MD) computer simulations of Chapter 4 indicated that short (length $L = 0.8$ nm) model nanopores with a methyl-

terminated surface would not allow the passage of ions below a critical radius $R \leq 0.6$ nm. Different effects must play a role in creating a barrier to ion permeation: For any charged species there is a purely electrostatic "dielectric barrier"[63] for transferring an ion from a high dielectric phase like water ($\epsilon = 78$) through a low dielectric membrane ($\epsilon \approx 2$), generally referred to as the Born energy.[82] An atomistic picture includes the interaction between individual water molecules and the solute, and the entropic cost of accommodating the solute cavity. The energetic cost for stripping the hydration shell off an ion is the free energy of solvation. When ions permeate narrow pores they will have to shed parts of their hydration shell. In the context of ion channel gating, it was hypothesised that this would create a barrier to ion permeation.[29,59,62,64,66]

Does the electrostatic continuum approach alone capture the essentials of ion permeation through low dielectric pores, as perhaps intuition and analytic models[63] suggest? Or is a more detailed atomistic approach required? If so, what is the magnitude of the effect?

In order to address these questions we calculate the potential of mean force (PMF; see Section 2.3.1 on page 46) of ions and water in the hydrophobic pore models that were also used in Chapter 3 and 4. The pore radius R was varied between 0.15 nm and 1.0 nm. These pores are very hydrophobic with the characteristics of a methyl-terminated surface. Hence there are no groups that could substitute for water molecules in the ion's hydration shell. The exclusion of ions (and water) from narrow pores (described in Section 4.3.2 on page 99) is quantified in terms of their free energy permeation barrier $\Delta G^{\ddagger}$, which is the maximum of the equilibrium free energy landscape as described by the one-



(a) $R = 0.35$ nm (b) $R = 1.00$ nm

Figure 5.1: Ions in pore models. Na^+ ions are depicted as blue spheres of radius 0.095 nm, with their first hydration shell shown as translucent spheres of radius 0.32 nm. The model pores are represented with their solvent excluded surface (images created with vmd[43] and Raster3D[44]).

dimensional PMF $\Delta G(z)$, i.e. the free energy required to position a particle at subsequent positions z along the pore axis.

The fully atomistic PMF is compared to continuum electrostatic PMF from Poisson-Boltzmann calculations. These are similar to the ones presented by Weetman et al.[262], who calculate the Born profile in cylindrical (radius $R = 0.25$ nm) and hour-glass shaped pores (minimum $R = 0.3$ nm).

5.2 Methods

Fully atomistic PMFs $\Delta G(z)$ are calculated from classical molecular dynamic (MD) trajectories (Section 2.3). Continuum electrostatic PMFs $\Delta G_{\mathrm{B}}(z)$ ("Born profile") are obtained as the Born energy by solving the Poisson-Boltzmann (PB) equation (Section 2.4).

5.2.1 Molecular dynamics

MD simulations were performed with gromacs version 3.2.1 (Ref. 127) and the SPC water model.[132] Parameters for the unified methane atoms and ions are described in Section 2.2.3 on page 38. Simulations were carried out for at least 100 ns. All other parameters were identical to the simulations described in Section 4.2.2 on page 88.

5.2.2 Potential of mean force

The PMF is related to the equilibrium density $n(\mathbf{x})$ through the density of the reference state in the bulk far from the pore n_0 (cf. Eq. (2.95) on page 48),

$$\beta\Delta G(\mathbf{x}) = -\ln\frac{n(\mathbf{x})}{n_0} + \beta(G_0 - \mu_0).$$

The free energy of the reference state G_0 is an unknown constant, but the normalisation with the excess chemical potential for the insertion of the permeant particle into the bulk liquid, μ_0, ensures that $G(z)$ equals 0 in the bulk.

Alternatively, one can bias the equilibrium distribution of a test ion to the energetically unfavorable regions using the technique of umbrella sampling[172] as described in Section 2.3.2 on page 48. We restrained a Na^+ ion or water molecule with a harmonic potential ($U(z) = k/2\,(z - z_i)^2$, $k = 7693$ kJ·mol^{-1}·nm^{-2}) at subsequent equally spaced positions z_i along the z axis of the pore (spacing $z_i - z_{i-1} = 0.025$ nm). This restricted its movement to the xy-plane while still allowing diffusion into adjacent windows. For each window a starting frame from an equilibrium trajectory was selected

which had the particle in question (ion or water molecule) close to z_i (maximum initial deviation was 0.3 nm, 80% of initial distances were smaller than 0.1 nm). One MD simulation of length 0.6 ns was carried out for each of the 101 windows, using the same parameters as above. The positional deviation $\zeta(t) = z(t) - z_i$ of the particle was recorded for every step (2 fs). A correlation time $\int dt \, \langle \zeta(t)\zeta(0)\rangle / \langle \zeta^2 \rangle$ of typically 0.2 ps indicates that ζ was actually oversampled by a factor of 0.2 ps/2 fs = 100.

The 101 biased distributions of z positions of the test particle (ion or water) were recombined and unbiased with the Weighted Histogram Analysis Method (WHAM)[174,175]. The first 0.1 ns of each window run were discarded as equilibration time, leaving a total of 0.5 ns per window. A tolerance of 10^{-5} kT in the self-consistent WHAM procedure lead to converged PMFs.

The PMF in the bulk and mouth region of the pore is obtained from the equilibrium MD trajectories. Umbrella sampled PMFs are only meaningful in regions where the diffusing of the particle is limited in the xy-plain, i.e. the pore and mouth region. Final PMF profiles were assembled by manually matching equilibrium profiles and umbrella sampled profiles in the mouth region where both methods are accurate.

For ions in pores with $R \geq 0.6$ nm only equilibrium MD is necessary to collect the complete PMF; sufficient sampling is achieved in 100 ns. Water PMFs could be obtained from equilibrium MD reliably for all radii except for the $R = 0.15$ nm pore but for comparison the PMF for a water molecule in the $R = 0.55$ nm pore was also calculated. The umbrella sampled water PMFs were symmetrized as they exhibited a small skew of about 1 kT (cf. Section 2.3.3).

5.2.3 Poisson-Boltzmann calculations

Poisson-Boltzmann calculations were carried out with apbs.[179] The test Na^+ ion was assigned a Born radius of 0.1680 nm (Ref. 181).* The radius of an implicit solvent molecule was set to 0.14 nm, the approximate radius of a water molecule. The electrostatic PMF or Born profile was calculated according to Eq. (2.115) on page 56,

$$\Delta G_B(z) = G^{el}_{pore+ion}(z) - G^{el}_{ion}(z) - G^{el}_{pore}. \tag{2.115}$$

In order to minimise grid artifacts, G^{el}_{ion} was calculated for every position z of the ion in the channel so that the ion was at the exact same grid position in the

*The Pauling radius of Na^+ is only 0.095 nm (Ref. 12) but as pointed out by Rashin and Honig[181], for electrostatic calculations one needs to consider the radius of the *cavity* within the solvent. For cations, the radius of the cavity is better approximated by the covalent radius, which is much larger than the ionic radius. Test calculations with the ionic radius did not reproduce the enthalpy of solvation for Na^+; the larger Born radius (including an empirical 7% correction[181]) was required.

calculations for the first two terms. $G^{\mathrm{el}}_{\mathrm{pore}} \equiv 0$ because the pore/slab system is uncharged.

Further parameters were chosen as to be consistent with the MD simulations. The temperature was 300 K, the ionic strength 1.3 M NaCl. The water dielectric constant ϵ_{w} was 63 (the value of the SPC water model[146]). The dielectric constant ϵ_{p} of the pore and slab material was set to 1 because in the MD simulations the material is modelled as simple van der Waals spheres without partial charges. These can not screen an electric field and hence are best described by $\epsilon_{\mathrm{p}} = 1$; a more realistic value for hydrocarbons would be $\epsilon_{\mathrm{p}} = 2$.

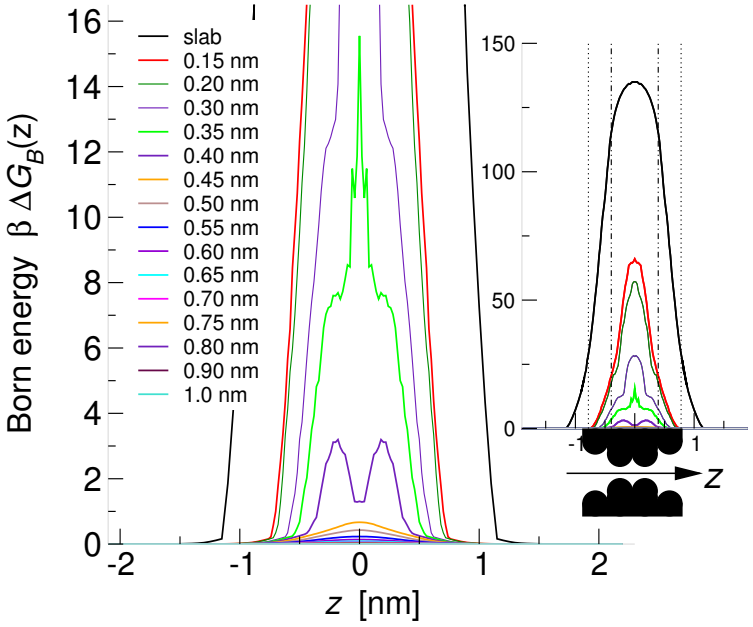
In addition to the calculations at a concentration of 1.3 $\mathrm{mol} \cdot \mathrm{l}^{-1}$ NaCl we also performed the same calculations at 0 $\mathrm{mol} \cdot \mathrm{l}^{-1}$ NaCl and 7 $\mathrm{mol} \cdot \mathrm{l}^{-1}$ NaCl. The binding energy ΔG_{B} shows negligible sensitivity on to the ionic strength. This is sensible because the test ion is the only charge in the system. Two sets of calculations with water dielectric coefficient at different values (63 for the SPC water model and 78.5 as the experimental value at 300 K) also indicated negligible sensitivity thereto. The calculations were sensitive to the grid discretization necessary to solve the Poisson-Boltzmann equation by a finite element scheme. Our tests on grid sizes ranging from 65 to 129 grid points indicated that our results are converged for grids of size $97 \times 97 \times 97$, using a focusing scheme (final grid resolution 0.031 nm) and a smoothed dielectric boundary.

5.3 Results and Discussion

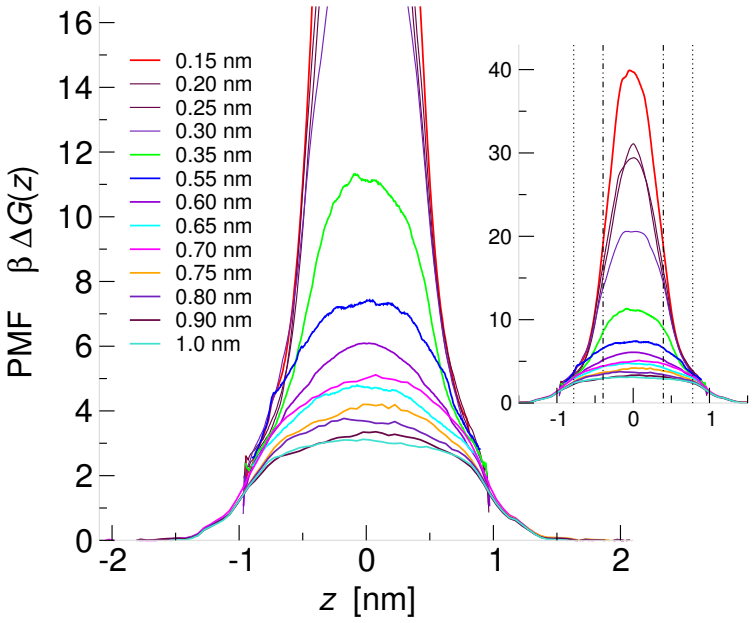
In the following we use the PMF as a quantitative measure for the influence of a range of parameters on the permeation properties of ions and water through hydrophobic pores. We begin by comparing the continuum picture (obtained from Poisson-Boltzmann calculations) to the fully atomistic one (based on umbrella-sampled MD). This will elucidate the rôle of the solvent in the process of ion permeation. Then we have a closer look at ions in wide pores ($R = 1$ nm) which, at first sight, should already provide a bulk-like environment, but fail to do so. Finally, for the case of water we look at varying the local flexibility of the wall and the length of the pore, at a fixed radius of $R = 0.55$ nm.

5.3.1 Continuum *vs* atomistic picture

First we turn to the continuum picture and consider a Born-type analysis of the barrier that a sodium ion has to overcome if it is to permeate a low dielectric



(a) Poisson-Boltzmann continuum electrostatics



(b) Molecular Dynamics (umbrella sampling and equilibrium simulations)

Figure 5.2: One dimensional PMF profile $\Delta G(z)$ for Na^+ ion permeation through hydrophobic pores of various radii. The dotted vertical line in the inset indicates the surface of the membrane mimetic slab. Mouth and pore region are indicated by dotted and dash-dotted lines; see Fig. 3.2(b) for details.

(ϵ_s) slab of thickness l. According to Parsegian[63] the Born barrier ΔG_B is

$$\beta\Delta G_B = \frac{\beta}{4\pi\epsilon_0}\left(\frac{e^2}{2\epsilon_s a} - \frac{e^2}{\epsilon_s l}\ln\frac{2\epsilon_w}{\epsilon_s+\epsilon_w}\right) \approx \frac{l_B}{2\epsilon_s a}\left(1-\frac{2a}{l}\ln 2\right), \quad \text{if } \epsilon_s \ll \epsilon_w. \tag{5.1}$$

a denotes the Born radius of the ion ($a = 0.1680$ nm for Na^+), ϵ_w the dielectric constant of the solvent (water), and $l_B = \beta e^2/4\pi\epsilon_0$ the Bjerrum length ($l_B = 56.0$ nm at $T = 300$ K). For $\epsilon = 1$ and thickness $l = 1.6$ nm Eq. (5.1) yields $\Delta G_B = 142\ kT$ (351 kJ $\cdot$ mol^{-1} at $T = 300$ K, almost the hydration energy 413 kJ $\cdot$ mol^{-1} of the sodium ion[181]). The explicit PB calculation gives a similar barrier of 135 kT [Fig. 5.2(a) on the preceding page]. This demonstrates that even a thin membrane creates a practically insurmountable barrier to ion permeation.

For a pore with $R = 0.55$ nm the continuum electrostatic calculations yield $\Delta G_B^\ddagger = 0.23\ kT$, i.e the barrier effectively vanishes. Fig. 5.2(a) shows electrostatic PMFs for pores of various radii R. Only for $R \leq 0.4$ nm is the Born barrier appreciable. For $R > 0.4$ nm it drops below 1 kT (Fig. 5.3 on the following page), ceasing to be a barrier for all practical purposes. This is in stark contrast to the results from the atomistic MD simulations [Fig. 5.2(b) on the preceding page]. For the wide $R = 0.55$ nm pore an ion still has to overcome a barrier of 7.4 kT, and even a pore with a diameter of 2 nm (more than the total thickness of the slab) will have an appreciable effect on ion permeation ($\Delta G^\ddagger = 3.2\ kT$).

The failure of the continuum model to describe the PMF in very narrow pores is not unexpected and has been noted before.[263,264] It is somewhat more surprising in the wider pores (about $R \geq 0.6$ nm) where a continuum description of the solvent electrolyte should become feasible. Investigation of the ionic density in the pore nevertheless shows a 0.6 nm depletion layer near the surface and a density in the wide pore barely reaching 70% of the bulk value (Fig. 5.5(b) on page 116). The strong disagreement between the electrostatic PMF and the atomistic one leads us to look for an explanation for the barrier which is not based on electrostatics.

The atomistic simulations include by their very nature entropic and short range interaction effects in addition to long range electrostatics, and they also account for the granularity of the solvent, thus admitting density effects. The situation in the narrow pores $R \leq 0.3$ nm can be interpreted in both pictures in a qualitatively similar way. The pore is too narrow to accommodate solvent together with the ion. In the Born picture this corresponds to an ion completely immersed in a low dielectric environment; in the atomistic picture it is related to the loss of the hydration shell (perhaps apart from two water molecules in axial positions—which can still contribute up to half of the total solvation energy[170]). In both cases it translates into a high energetic barrier. For $R \geq 0.4$ nm the ion is enveloped by a shell of high dielectric solvent of thicker than 0.1 nm. This is

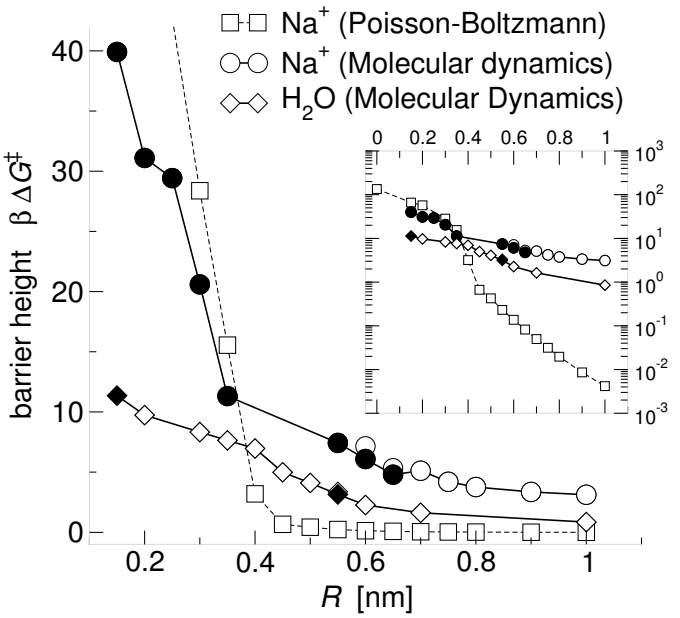


Figure 5.3: Comparison of the barrier height $\Delta G^{\ddagger}$ from the electrostatic $\Delta G_B(z)$ and atomistic $\Delta G(z)$ PMF profiles. Data points indicated by filled circles or diamonds were calculated from umbrella sampled MD trajectories whereas the empty variants were directly obtained from equilibrium MD. Empty squares were obtained from continuum-electrostatic Poisson-Boltzmann calculations.

already enough to lower the Born energy considerably. The continuum calculations assume that the solvent simply exists in the pore. There is no energetic cost associated with filling the pore with water. The PMF for water (Fig. 5.4 on the next page) shows that the free energy to place a water molecule at the centre of the pore is considerable (between 3 kT and 7 kT for medium sized pores, also see Fig. 5.3). This means it is difficult to maintain a solvent environment in the pore which in turn could successfully hydrate and stabilize the ion (although collective effects will reduce the barrier to filling the pore with water somewhat as hinted at in Chapter 3). Hence the reason for the high barrier for ions is not so much the electrostatic contribution of placing a charged particle at the centre of a aqueous pore through a low dielectric membrane but the inability of water to wet (or solvate) the pore itself.

5.3.2 Ions in wide pores

It seems surprising that ions are considerably affected even in a pore of 2 nm diameter, greater than its length of 1.6 nm, and about ten times the ionic radius. The radially averaged density from the MD simulations shows that water already behaves rather bulk-like in the pore (Fig. 5.5(a) on page 116). It fills almost the whole pore volume at bulk density. Near the pore surface the same characteristic water layers are visible that are also seen at the slab-reservoir interface. The ionic density, Fig. 5.5(b) on page 116, is concentrated in the central region and only reaches about 0.9 $\text{mol} \cdot \text{l}^{-1}$, i.e. 70% of the bulk value. This effect is not due to insufficient sampling time as investigation of the density plots

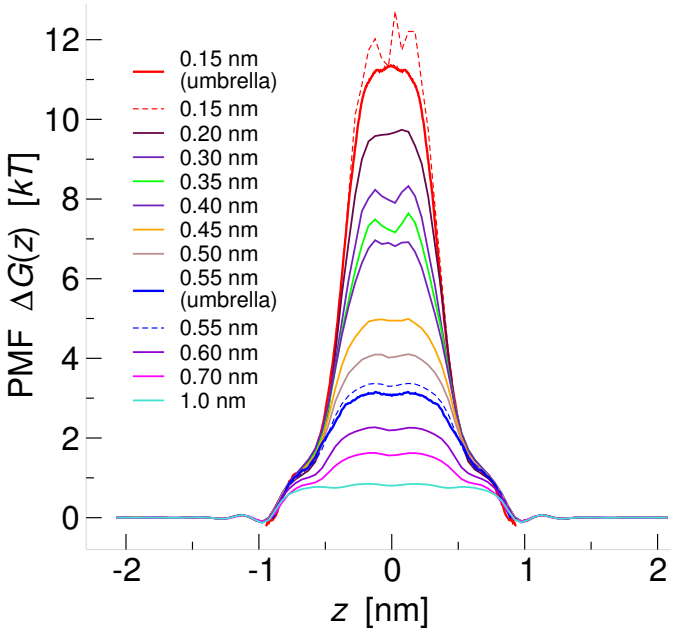
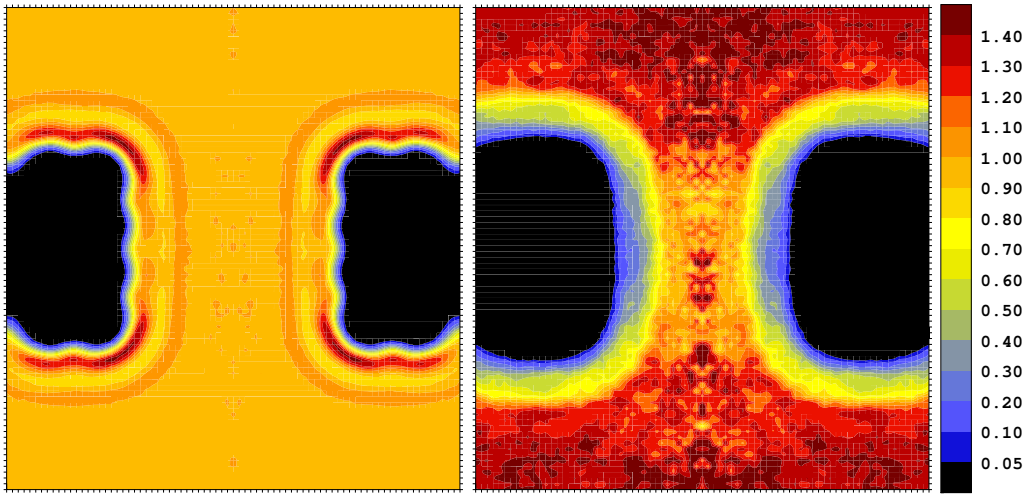


Figure 5.4: PMF profile for water. The PMF was derived from equilibrium simulations except for the $R = 0.15$ nm and 0.55 nm pore, where umbrella sampling was employed (heavy lines). For these two radii, the equilibrium trajectory derived PMF is also shown for comparison (thin lines).

from shorter stretches of the 100 ns trajectory indicates. (About 20 ns appear to be the minimum time necessary to obtain meaningful densities.)

The effect of the surface on the ions extends noticeably for at least 1 nm. This is also borne out by the radial distribution function $g(r; \mathrm{Na}^+, \mathrm{O_w})$ (Fig. 5.6 on page 117) from a homogeneous electrolyte (99 ns MD at 1 mol $\cdot$ l^{-1} NaCl concentration in a $4 \times 4 \times 4$ nm^3 simulation box, $T = 300$ K, $P = 1$ bar). The first water peak of height 7.59 is at a distance of 0.226 nm from the Na^+ centre but the second peak is still noticeable (1.44) at 0.447 nm. The first hydration shell of Na^+ contains ca. 4.6 water molecules in a radius of 0.32 nm, while the second shell adds another 15.6 within 0.53 nm (the number of water molecules $N(r)$ within a sphere of radius r at the minimum following a peak in $g(r)$; see the inset of Fig. 5.6 on page 117). Thus about 20 water molecules in a sphere of radius 0.53 nm are noticeably influenced by a Na^+ ion. On the other hand, water layering effects near the hydrophobic surface extend over two to three layers of water (Chapter 3 and 4), influencing water 0.6 nm afar from the surface. Combining these two ranges we estimate that the solvent-mediated wall-ion interaction can extend to about 1 nm. Thus ions will still be affected by the pore surface even if the pore is ten times as wide as the bare ion—an effect which is clearly not captured by the PB calculations, and which would be also absent in Brownian dynamics simulations of the system.



(a) Water (b) Na^+

Figure 5.5: Radially averaged density $n(r, z)$ in a hydrophobic pore with radius $R = 1.0$ nm. The colour scale indicates the local concentration in $mol \cdot l^{-1}$ for Na^+ or units of the bulk density of SPC water $n_0 = 1.0\ g \cdot cm^{-3}$ at $T = 300$ K and $P = 1$ bar. (The figures were mirrored on the pore axis to create the impression of a full pore; they were created with xfarbe, Ref. 223.)

5.3.3 Long or flexible pores

The previous results established that much can be learned about the permeation properties of ions by looking at the behaviour of the solvent, i.e. water. The water behaviour indicates how ions will behave because the ions rely on the water molecules for energetical stabilization. Qualitatively, water acts as a proxy for ions, hence simulations of water also predict the behaviour of ions in similar systems.

We have also investigated the PMF of water in longer pores, or pores with differing wall flexibility. Briefly, we touched on the latter subject in Section 4.3.1 on page 95. The question of longer pores has only been addressed theoretically before; in particular Eq. (4.6) on page 91 predicts that an increase in pore length has the same effect on the free energy difference between liquid and vapour states as has an increase in the contact angle, i.e. a decrease in the wall-liquid interaction. It turns out that the pore length L_p that was chosen on the basis of biological considerations (the length of the putative gate in nAChR) defines a system whose liquid-vapour oscillations are on the time scale of nano seconds. A shorter pore $L_p = 0.4$ nm is constantly filled, irrespective of radius, i.e. the free energy difference between vapour and liquid is very large and favours liquid. In longer pores $L_p \geq 1.2$ nm, on the other hand, the transition between

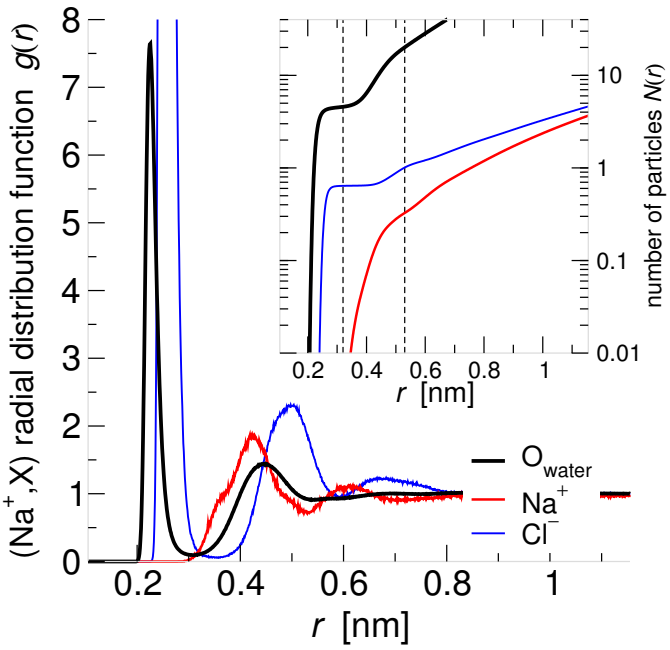


Figure 5.6: Radial distribution function $g(r)$, centred on Na^+ in 1 M NaCl bulk solution. Broken vertical lines indicate the boundary of the first and second hydration shell of a sodium ion, i.e. the minima in $g(r; Na^+, O_w)$. The inset shows the number of particles of a given type inside a sphere of radius r around a Na^+ ion, $N(r) = n_0 \int_0^r dr' \, 4\pi r'^2 \, g(r')$.

a system favouring liquid and one favouring vapour is much sharper. Essentially no oscillations are observed on the time scales of our simulations (100 ns) and pores are either liquid-filled or vapour filled for the whole simulation time. Hence the free energy difference according to Eq. (4.3) cannot be determined. This also means that the length dependence of Eq. (4.6) could not be verified by simulation data. Umbrella sampling of the PMF allows at least to determine the barrier height for a single water molecule though not the free energy difference, which would require the same kind of umbrella sampling simulations that were carried out by Allen et al.[202].

Dependence of the barrier on the wall flexibility It was already shown in Section 4.3.1 that *increased* wall flexibility, i.e. a softer wall, *decreased* the permeability for water molecules. In Fig. 4.4 a dramatic shift towards the vapour state was evident. The permeation barrier follows the very same trend as the water permeability. Fig. 5.7 shows results for a softer and a more rigid pore than the standard ones used throughout this work . The local flexibility of the wall atoms is measured by their positional root mean square deviation (RMSD). For the standard pore with $k = k_0 = 1.0 \times 10^3$ kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2} $= 4 \times 10^2$ kT nm, the RMSD is 0.072 nm. By increasing the local flexibility to an RMSD of 0.17 nm ($k = 0.2\,k_0$) the barrier to permeation increases substantially from $\Delta G^\ddagger(k = k_0) = 3.2$ kT by 7.2 kT to $\Delta G^\ddagger(k = 0.2\,k_0) = 10.5$ kT. However, making the pore wall twice as rigid (i.e. more crystalline, with RMSD 0.052 nm) has a negligible effect; the barrier only drops by 0.4 kT to $\Delta G^\ddagger(k = 2\,k_0) = 2.8$ kT.

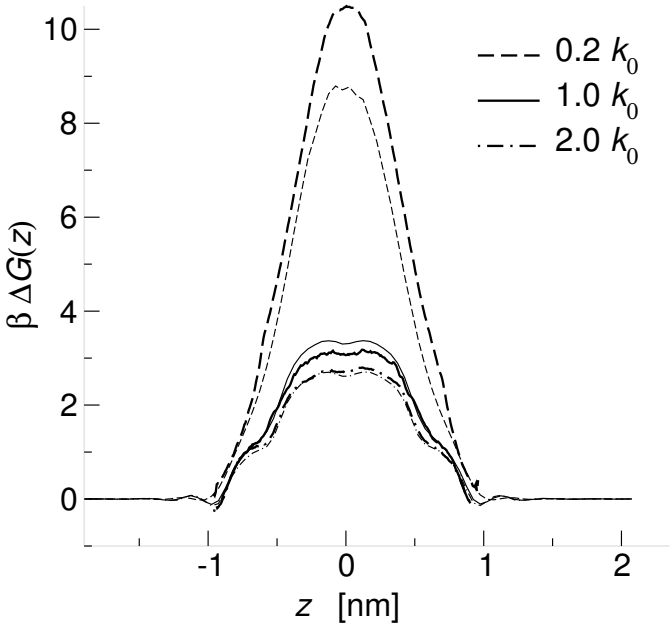


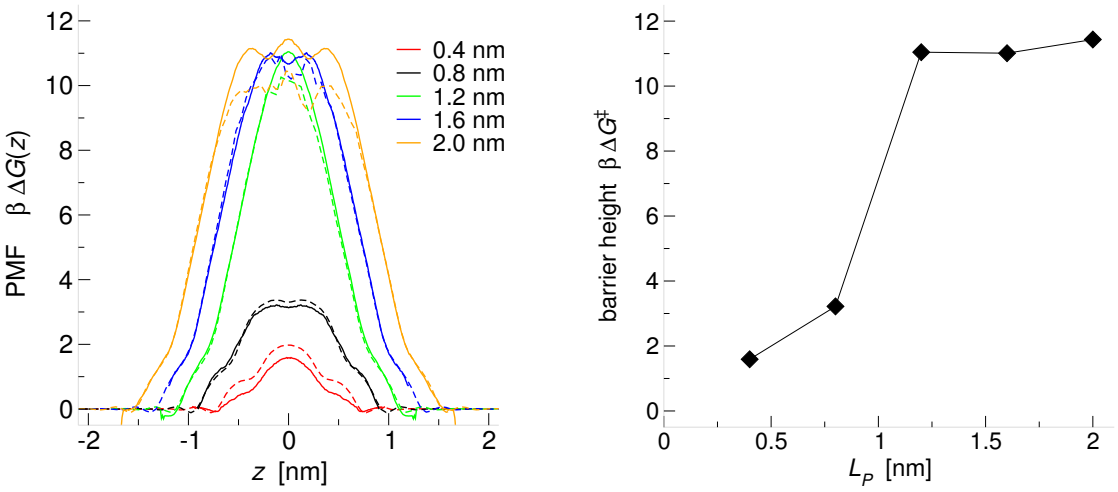
Figure 5.7: PMF of a water molecule in pores of varying local flexibility. The flexibility of the pore wall atoms is set through the harmonic force constant of the restraint force. $k_0 = 1.0 \times 10^3$ kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2}. The Umbrella-sampled PMF is given by the thick line; the PMF derived from the equilibrium density is shown as a thin line of the same line style.

The simulations show that the PMF reflects the profound influence that local flexibility can have on permeation properties of water molecules. Because ion permeation in hydrophobic pores requires the ion's water environment to be stable in the pore, an even stronger effect on the ion PMF is anticipated for softer pores.

Water PMF for differing pore length Umbrella sampled PMFs for a water molecule in hydrophobic pores of radius $R = 0.55$ nm and increasing length (0.4 nm $\leq L_p \leq$ 2.0 nm) are displayed in Fig. 5.8(a). When going from $L_p =$ 0.8 nm to 1.2 nm, the barrier height increases considerably from 3.2 kT to 11.0 kT (Fig. 5.8(b)) but then it seems to level off between 11 and 12 kT for longer pores. Because we have no theory that predicts the height of the barrier itself it is not clear what induces the sudden increase in $\Delta G^{\ddagger}$ or if the barrier approaches a constant value. However, these results demonstrate the importance of the pore length as another factor shaping the behaviour of water and—by extension—ions in narrow pores.

5.4 Conclusions

The calculations presented in this chapter suggest that hydrophobic pores can pose high barriers to the permeation of ions, even when the pore radius is considerably larger than the ionic radius. The barrier originates in the high energetic cost for an ion to shed its first or, in wider pores, its second hydration shell.



(a) PMF $\Delta G(z)$ (b) Barrier height $\Delta G^{\ddagger}$

Figure 5.8: Umbrella-sampled PMF of a water molecule in pores of differing length. The umbrella PMF is shown as a solid line, with the equilibrium density derived PMF in broken style. Umbrella sampled PMFs were symmetrized.

The barrier relates to the solvation energy for an ion/hydration shell complex in water. In the narrow selectivity filter of the potassium channels, nature has demonstrated how to overcome dehydration barriers by solvating ions by carbonyl oxygens, which seamlessly replace water molecules in the first hydration shell,[16,48] leading to a vanishing barrier for ion permeation.[240]

Umbrella sampling has been shown to be an appropriate method to quantify the effect of a change in radius, pore length, or wall flexibility on the permeation properties of water and ions. It allows to probe regions of parameter space that are inaccessible to equilibrium simulations.

The results obtained in this chapter are not (yet) embedded in a theory that is able to explain the shape and height of the barrier. It would be desirable to have such a theory though no attempt to formulate one will be made in this work. We simply note some final observations regarding such a theory. The comparison between continuum and atomistic calculations showed that the behaviour of the ions depends strongly on the behaviour of the solvent, and any theory would probably have to incorporate at least two effects: the dielectric barrier and the cost for water to solvate the pore.

6 A gate in the nicotinic receptor

The potential of mean force (PMF) of sodium and chloride ions permeating the nAChR pore (formed by the pore lining M2 helices from the 4 Å closed-state structure) was calculated by fully atomistic molecular dynamics equilibrium simulations and free energy calculations (umbrella sampling). It is found that ions would have to overcome a sizable free energy barrier of about 10 kT at the hydrophobic girdle between residues L9′ and V13′, which has been implicated in forming a "hydrophobic gate" at the centre of the pore. The other major constriction, close to the intracellular side, presents a lower barrier so that on the basis of these calculations it is suggested that the gate in the nicotinic receptor is located at the hydrophobic ring. Charged rings at both ends of the pore act as concentrators of ions up to about six times the bulk concentration; an effect which would increase the ion current in the open state. The influence of different MD force fields and the reduction of the protein to the M2 bundle is discussed; the general findings are robust with respect to different force fields though the PMFs differ in some regions.

The central question pursued so far has been how the pore environment influences the behaviour of water (and ions). The dimensions of the toy pore models of Chapter 4 were modelled after what was known from the lower resolution structure of the nicotinic acetylcholine receptor (nAChR) at the time.[58,59,65,265] The advent of the 4 Å structure of the transmembrane domain[29] in 2003 allowed us to compare the highly simplified toy model environment to the realistic one of the ion channel.

6.1 The nicotinic acetylcholine receptor

Nicotinic acetylcholine receptors are membrane channels that mediate signal transduction at the post-synaptic membrane of cholinergic synapses such as the neuromuscular junction.[266] They belong to the ligand-gated ion channels (LGIC).

6.1.1 The ligand gated ion channels

The ligand gated ion channels comprise of five different known families, four of which form a sequence-related superfamily (INTERPRO[267] entry IPR006029). These are the gamma-aminobutyric acid type A ($GABA_A$), nicotinic acetylcholine (ACh), glycine, and the serotonin (5-hydroxytryptamine, $5HT_3$) receptors. The ionotropic glutamate receptors (IPR001320) have a distinct primary structure.

The members of this superfamily share a common architecture. They are pentameric complexes of varying subunit composition (for instance, the muscle nAChR of the electric ray *Torpedo* contains four different subunits $\alpha_2\beta\gamma\delta$ whereas the $\alpha 7$ receptor is a homo pentamer). A pore runs through the centre of the complex. Each subunit contains a large N-terminal domain, which contains the ligand binding site, three hydrophobic transmembrane (TM) regions, a large intracellular region, and a fourth hydrophobic TM region at the C-terminus.[266,268] Based on the structure of the muscle nAChR from *Torpedo marmorata* one can confidently identify the TM regions with four membrane spanning α-helices M1–M4[29] and the ligand binding domain with a structure made up mainly from β-sheets.[60] The nicotinic receptor ligand binding domain is a specialised pocket of aromatic and hydrophobic residues formed at interfaces between protein subunits that changes conformation to convert agonist binding into gating of the ion channel.[29,269] The intracellular domain forms a sphere-like vestibule with five windows[59,265] but could not be resolved at atomic resolution.

An alignment of the transmembrane domain of the ligand gated ion channels was prepared, based on the Pfam ("protein family") PF02932 (Neurotransmitter-gated ion-channel transmembrane region), which contains 616 members (version 14.0 of PFAM[270]). After removal of fragmentary sequences 511 remained which were aligned with ClustalW.[271] A calculation of an average distance neighbour joining tree (using Jalview[272]) shows four distinct families (nAChR, $5HT_3$, $GABA_A$, GlyR) for the whole alignment. Focusing on the M2 region three distinct families remain: nAChR (234 sequences), $5HT_3$ (20 sequences), and GABA/GlyR (244 sequences). From the M2 sequences sequence logos[273] were created, which are shown in Fig. 6.1 on the next page. In the M2 region, the ligand gated ion channels have a number of features in common: two highly conserved hydrophobic residues at positions $8'$ and $9'$ (nAChR and $5HT_3$: LL, GABA/GlyR: VL; $L9'$ corresponds to αL251 in nAChR), basic residue at $-1'$ (nAChR: K, $5HT_3$ and GABA/GlyR: R) with a subsequent hydrophobic one ($V1'$) at the beginning of the M2 helix. The nAChR and $5HT_3$ receptors are cation selective ion channels. Their sequences exhibit features that are absent in the GABA/GlyR family, such as the long hydrophobic stretch $13'$–$18'$ (VFLLLI or VFLIIV), the intermediate acidic ring (E at $-2'$), and the extra-

(a) nAChR

(b) $5HT_3$

(c) GABA and GlyR

Figure 6.1: Sequence logos[273] for the ligand gated ion channels in the M2 region. Acidic residues are marked red, basic ones blue, polar ones green. Hydrophobic residues (and proline) are coloured black. Residues are counted within M2, following the standard "primed numbering scheme" for nAChR[274,275] where αM243 is designated $1'$. G$-3'$ in (a) only exists in the γ or ϵ subunits of nAChR; the other subunit sequences do not contain this particular residue. The information content $R(l)$ per residue l is measured in bits, $R(l) = \log_2 M - \left(-\sum_{i=1}^{M} p_i(l) \log_2 p_i(l)\right)$, where $p_i(l)$ is the probability to observe amino acid i at position l. As there are $M = 20$ amino acids the maximum information content per position is $\log_2 20 - 0 = 4.322$ bits.[273,276] The size of a letter indicates this amino acid's contribution to the information content at the position. Logos were created with WEBLOGO,[277] or with alpro and makelogo.[273]

cellular acidic rings (E or D at 20′ and 24′).

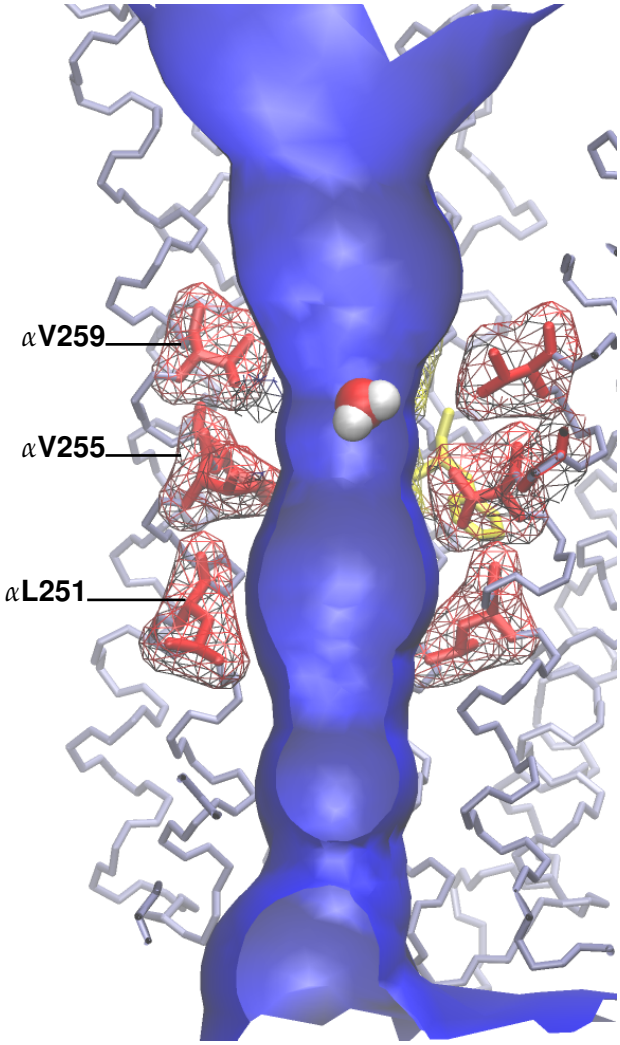
The nicotinic receptor is one of the best studied ion channels and serves as a pardigm for ligand gated ion channels. The muscle-type receptor facilitates signal transduction from motor neurons to muscle cells and is located in the post synaptic membrane of a muscle cell; neuronal nAChR such as the $\alpha 7$ receptor (discussed briefly in Appendix A) are found in brain synapses. nAChR is a cation selective ion channel with conductances of the order of 30 pS.[12] Its natural agonist is acetylcholine (ACh) but it binds nicotine at even higher affinities.

6.1.2 Structure of the pore

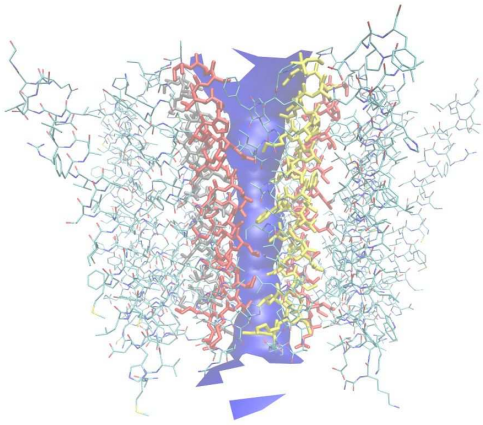
The nicotinic receptor is the only membrane channel of which a structure is known in both the open[58] and closed state.[29,65] Although the resolution of the open state structure is comparatively low at 9 Å, a wide pore of radius $R \approx 0.65$ nm is visible in the density. This is wide enough to let fully hydrated cations pass. It is also consistent with nAChR being an only weakly selective cation channel which is permeable to many inorganic cations and at least 41 organic cations and even non-polar molecules like urea, ethylene glycol, or glycerol.[12]

Because the open and closed structures are known one should be able to learn something about the gating mechanism of the receptor, and by similarity, about the gating mechanism in all related LGICs. The current view, based on the structure of the receptor itself[29,60,278] and the acetylcholine binding protein AChBP,[279] which is homologous to the ligand binding domain of nAChR, holds that a structural change in the ligand binding domain, induced by the binding of two molecules of ACh, is transmitted through the linker region[280] to the M2 helices of the transmembrane region. The M2 helices swing in hinges (in the M2–M3 and M1–M2 linker) to open or close the pore.[18,29,61,280] Many details of the gating process are still elusive. Even the seemingly trivial question where the flow of ions is blocked in the closed state—the gate—is debated as biochemical evidence[281,282] (summarised by Karlin[275]) clashes with structural[29] and other biochemical data.[283–287]

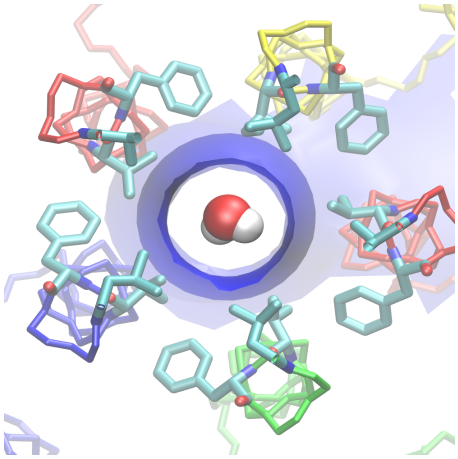
Where is the gate? The structure of Miyazawa et al.[29] (at 4 Å resolution) is thought to represent the closed state because no agonist was present during recording of the electron microscopy images. At this resolution the backbone could be built into the electron density although most of the sidechains had to be modelled. The structure unambiguously shows that the M2 helices are lining the ion pathway. The pore is also peculiarly feature-poor without any prominent sidechains that could act as an occlusion gate (Fig. 6.2(a) on the facing page). Furthermore, the lower resolution images[59,265,288] (at about 4.6 Å)



(a) Cut through the nAChR pore. Red and yellow: solvent accessible surface of pore lining hydrophobic residues.



(b) Pore lining M2 helices in the transmembrane region of nAChR, highlighted by subunit.



(c) Cut through the "hydrophobic gate" (approximately at the height of αV259)

Figure 6.2: Transmembrane spanning pore in nAChR. α subunits are coloured red, γ yellow, β green, and δ blue. The solvent accessible surface of the pore (created by hole[42]) is shown in blue. A water molecule is drawn to scale. Images were drawn from the cryo-electron microscopy structure 1OED,[29] using vmd.[43]

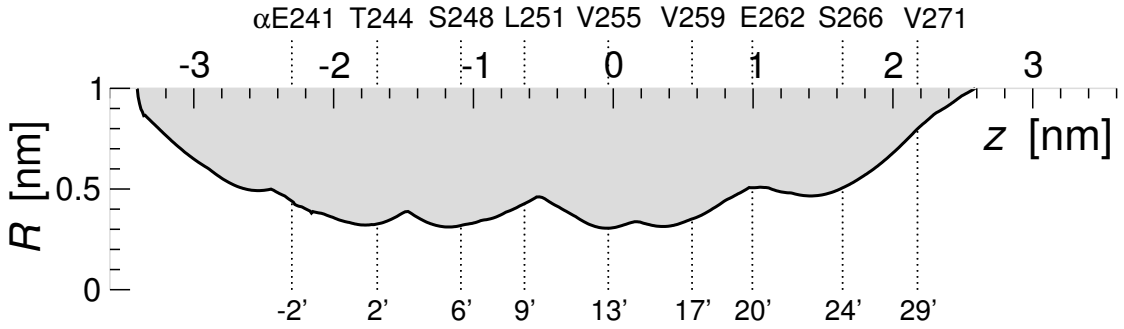


Figure 6.3: Pore profile $R(z)$ of the ion pathway through the closed nAChR transmembrane region. The radius represents the solvent excluded surface of the pore. Important pore lining residue in the α subunits are indicated: The two constriction sites at S248 and V255 are visible. The "hydrophobic girdle" is formed by L251/V255/V259. Charged rings are located at the height of E241, E262, and S266. The alternative "primed M2 notation" is also given. (Profile calculated with hole[42] with a water probe van der Waals radius of 0.14 nm.)

do not show any significant constriction sites in the extracellular ligand binding domain (pore radius > 1 nm). The intracellular vestibule, which is formed by a "hanging gondola" or "basket" with five windows of diameter 1 nm each, through which ions have to pass.[59,265] In the 5-HT_3 receptors these windows are lined by positively charged residues, which were shown to decrease the single-channel conductance[289] but other homologous ligand gated ion channels have acidic or uncharged residues in the corresponding positions. Hence it seems likely that the main function of the intracellular windows is to act as a static ion filter and attenuator of conductance.

Based on the structural evidence the gate is to be found somewhere in the transmembrane region of nAChR. In Fig. 6.2(a) the pore with the M2 helices is shown. Unwin and colleagues[29,61] hypothesise that the gate is located at the most constricted region of the channel in the vicinity of the highly conserved αL251 and the neighbouring αV255. The minimum pore radius in the closed state is $R = 0.31$ nm (cf. the pore profile $R(z)$ in Fig. 6.3) close to αV255; however there is a similar narrow region closer to the intracellular end at αS248. Although this represents a significant decrease in radius compared to the open state it is still a large hole in molecular terms. It allows the simultaneous passage of about three water molecules or a half-hydrated sodium or potassium ion. By comparison, the most constricted region in the K-channels[16,26–28] in the *open* state is the selectivity filter with a radius of less than 0.15 nm. Both nicotinic receptors and K-channels have typical single channel conductances of 30 to 50 pS (Ref. 12) but the K-channels conduct even though their open state pore is much narrower than nAChR closed state one. Thus, the dimension of

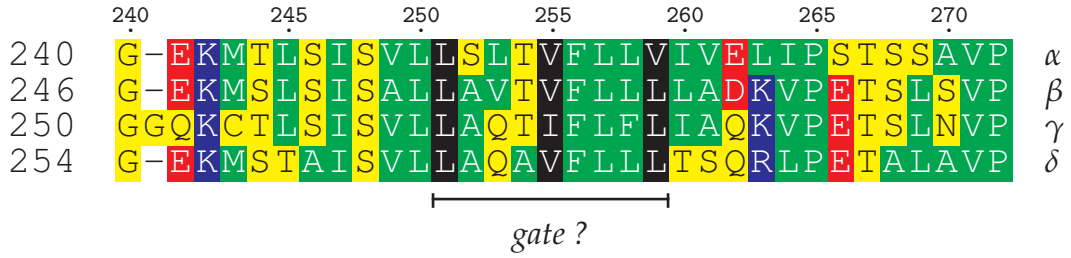


Figure 6.4: Alignment of the M2 stretches from the *Torpedo californica* nAChR subunits. The ruler counts residues in the α subunit. In the "primed M2 notation" of Lester[274], αM243 is 1′, the start of the M2 helix. αL251 corresponds to 9′ and αV255 to 13′. The gate as proposed by Miyazawa et al.[29] is indicated between αL251 and αV259. **Hydrophobic** residues are coloured in green, or black (as **pore lining hydrophobic** residues). **Acidic** (positively charged) residues are coloured red, **basic** (negatively charged) blue and **polar uncharged ones** yellow. (Alignment typeset with TEXshade[290])

the pore alone can not determine the conductance properties. Hille[12] sensibly suggests that the interaction of the ion with the pore wall must play a decisive role in shaping the conductance and selectivity properties of ion channels.

The K-channel selectivity filter is lined by eight backbone carbonyl oxygen atoms of the TVGYG signature motif.[16,26–28,237] These oxygens substitute for water molecules in the hydration shell of a permeating ion. The situation appears to be different for the nicotinic receptor. The sequence of the pore lining M2 helices (Fig. 6.4) shows a number of hydrophobic residues, and in fact large parts of the pore are exclusively lined by leucines and valines (Fig. 6.2(a)). Thus it appears unlikely that a similar water-substitution mechanism would help ions to move through the pore. Furthermore, as was demonstrated in the previous chapters for the case of hydrophobic model pores, the presence of a hydrophobic pore lining surface appears to *be* the mechanism which prevents ions from permeating the hydrophobic girdle. The MD simulations of this chapter are aimed at testing this hypothesis.

Not only is the central region of the pore lined by hydrophobic residues but also does the complete stretch L251–V259 consist of hydrophobic residues. These form a compact dielectric plug at the centre of the membrane, which might contribute to an increase in the Born barrier.

The TM region is flanked by rings of acidic residues. Three of these rings are clearly visible in the sequence logo for the nicotinic receptors, Fig. 6.1(a). The intracellular ring at position $-6'$ lies in the M2-M3 linker.[275] The intermediate

ring at $-2'$ and the extracellular ring at $20'$, however, are directly lining the pore (see also Fig. 6.4 for the identities of the corresponding residues in the different subunits). Some nicotinic receptors (notably the muscle nAChR and the $\alpha 7$ one) have a fourth charged ring at position $24'$. There are no charged residues within the membrane, where the position of the membrane is inferred from belts of aromatic residues (Fig. 6.5 on the next page shows the position of the charged residues relative to the aromatic belts and hence relative to the membrane). All charged sidechains are located in the lipid head group region or above. A highly negatively charged channel lumen is expected for a cation selective channel like the nicotinic receptor or the closely related 5-HT_3 receptors (cf. Fig. 6.1(b) and Fig. 6.1(a)); anion-selective channels like the GABA and Glycine receptors introduce positively charged rings at $19'$ and $24'$ instead (Fig. 6.1(c)). It is suggested that charged vestibules can act as ionic concentrators, which in turn increases conductance.[12]

6.1.3 Identifying the gate

Based on the knowledge of the structure of *Torpedo marmorata* nAChR in the closed state we wish to identify the gate, i.e. the region of the pore which prevents ions from permeating the pore. As argued in the previous section, we hypothesise that the gate is to be found in the transmembrane spanning pore. This inference is based on the knowledge of static structures. The experience with the dynamics of ion permeation in model pores, described in the previous chapters, strongly implicates the hydrophobic constriction site formed in the vicinity of αV255, βL261, γI264, and δL269. In this chapter we test our hypothesis by carrying out atomically detailed molecular dynamics simulations of ions in the nAChR pore. In particular, we compute the PMF of ions permeating the receptor pore; our calculations of the barrier in model pores in Chapter 5 demonstrated the applicability of this approach. Also, our previous calculations indicated that simulations of more than 50 ns (deemed "long" by current standards) are required for proper sampling of equilibrium ionic densities in confined geometries. In order to reach long simulation times we perform our calculations on a reduced model, consisting only of the M2 helices, which is described in detail in 6.2.2.

6.2 Methods

The complete transmembrane region of *Torpedo marmorata* nAChR (pdb code 1OED), which was solved by cryo electron microscopy (cryo-EM) to about 4 Å resolution,[29] was taken as the initial structure.

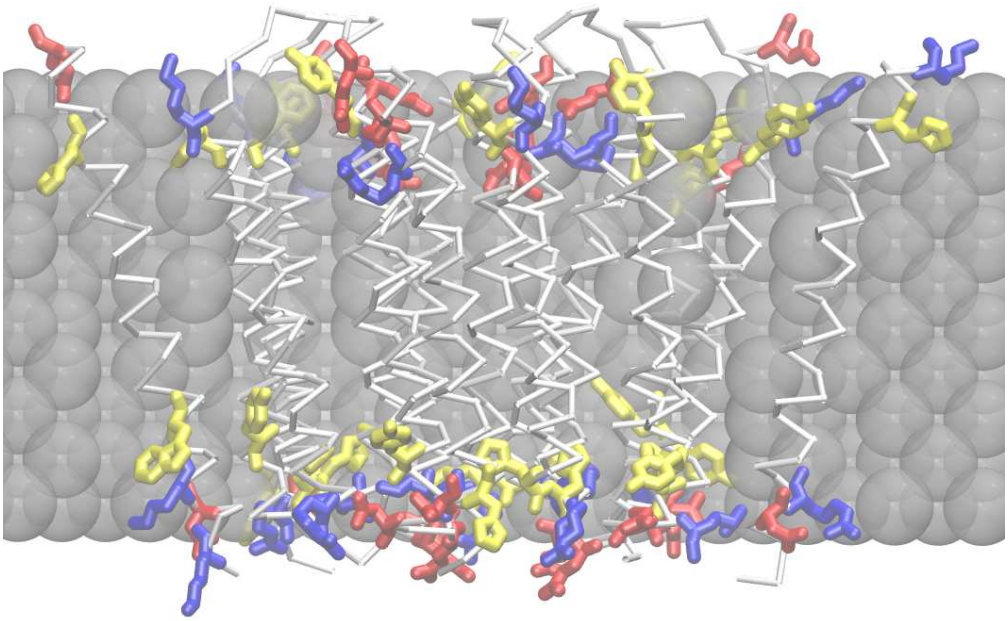


Figure 6.5: Transmembrane region of nAChR (1OED) embedded in a 0.3 nm thick membrane mimetic slab. "Aromatic belt" residues (Tyr, Trp) are coloured yellow, acidic groups (Glu, Asp) are red, and basic sidechains (Lys, Arg) blue.

6.2.1 pK_a calculations

First, the protonation states of titratable residues at pH 7.0 and ionic strength 0.144 mol $\cdot$ l^{-1} of monovalent ions were determined. The protein was embedded in a membrane mimetic slab of thickness 3.0 nm, which serves to exclude high dielectric solvent from the lipid-facing outer surface of the protein (Fig. 6.5). The slab was approximately aligned on the belts of the aromatic Tyr and Trp residues;[291] all acidic or basic sidechains on the outside of the protein are at the height of the interface or protrude into the solvent (Fig. 6.5).

The pK_a were computed with the help of Jens Nielsen's pK_a scripts.[184] These in turn drive WhatIf,[292] which employs DelPhi 4.0[186] as its Poisson-Boltzmann solver. WhatIf calculates a hydrogen bond network optimisation[185] although this is probably much more important when calculating accurate ionisation states from high resolution structures.

Because the pK_a module in WhatIf seems to have problems with N-terminal Pro and titratable C-terminal residues, all the N-terminal Pro were deleted from all five chains, and on chain C the C-terminal δPro485 was added after δLys484.*

*WhatIf was used for all molecular editing operations. The `CORALL` command was executed to add terminal oxygens and correct the structure (using a library of allowed conformations) although this was only necessary for the added Pro. This is to be expected as 1OED was run

Interactions between titratable groups were only computed if the electrostatic interaction in the fully charged state exceeded a "pair energy cutoff" of 4 kT.[184] Test calculations on other systems (OpcA and OmpX) showed that this is an acceptable simplification.

6.2.2 Models

As we want to investigate how the environment affects the behaviour of water and ions we are not concerned with large scale protein motions but only local properties of the water/wall interface. Thus major simplifications are carried out to reduce the system size, which is important to achive long simulation times for accurate sampling of water and especially ion trajectories.

1. The outer transmembrane helices M1, M3 and M4 (the "scaffold") were removed using WhatIf. Only the full length M2 helices with some residues from the M1-M2 and M2-M3 linker were retained (αE241 to αV271). The model contained 31 residues in each chain and stretched over a length of 5 nm. The resulting structure was called the *P1* model.*

2. Because the reduced pore model lacks stabilising contacts with the outer scaffold and with the ligand binding domain the backbone atoms were restrained to their position determined by cryo-EM,† similar to the approach adopted by Smith and Sansom[171] for PMF calculations in leucine-serine peptides.

3. Once the equilibrium position of the backbone is fixed we can also simplify the treatment of the lipid membrane. The M2 helices would not

through procheck[293] before publication and many sidechains were software-built, as can be inferred, for instance, from the REMARK section of 1OED.pdb:

> [...] These coordinates, including estimates of side chain positions, are being made available in the hope that they will be useful. Users should bear in mind that because of the limited resolution the conformations of the side chains and their atomic coordinates are not individually reliable. [...]

The edited structure was only used for the pK_a calculations.

*A so-called *P4* model was also constructed for initial test calculations. It consisted of the hydrophobic pore itself (from αI247 to αV259) (13 residues, length 2.1 nm). *P2* and *P3* were designed as intermediates but were never used.

†In fact, there are NMR studies of an M2 only channel in a bilayer[294] and corresponding MD simulations.[295,296] The MD simulations showed some significant changes in the pore geometry. For the purpose of the present study it was deemed necessary to keep the pore geometry well defined. In essence, it is assumed that the experimental structure is a good representation of the closed state and on the time scales accessible to MD no larger rearrangements or destabilisations would likely occur.

be in contact with lipids in the full receptor and any lipid-protein interactions would be countered by the backbone restraints. Although a real lipid membrane contains polar head groups, which may have a long range electrostatic effect on ion permeation this can likely be ignored in this particular case: In the complete receptor, the scaffold shields the permeation pathway from the head groups. The minimum distance for a lipid head group from the pore surface is 2 nm, as estimated from the solvent-accessible surface[121] of the structure 1OED in vmd.[43] It also contains cavities which are most likely be filled with water (see Section 6.3.3), which can screen electrostatic fields effectively. Thus it appears acceptable to leave out the polar head groups and model the membrane by a low dielectric slab made from methane molecules, as was done for the model pores in the previous chapters.

The final model is suitable to investigate the question how the environment of the closed nAChR pore affects ions and water under the assumption that the closed pore is correctly described by the cryo-EM structure 1OED.

6.2.3 Poisson-Boltzmann calculations

The Born profile for a Na^+ ion in the nAChR pore was calculated analogously to the model pores in Section 5.2.3 on page 110 with a few variations to account for the presence of the protein. Hydrogen atoms were added to the input structure (either 1OED or *P1*) with WhatIf[292]. The resulting structure was embedded in a membrane mimetic slab.* The ion was positioned on the pore axis (calculated with hole[42]) in steps of 0.02 nm and extended for 1.5 nm into the bulk. The PDB2PQR[297] server was used to build an input file for the Poisson-Boltzmann solver apbs.[179] Protein radii and partial charges were taken from the AMBER force field† (as provided by PDB2PQR) and a Born radius of 0.1680 nm was used for the Na^+ ion.[181] Poisson-Boltzmann calculations—as described in Section 2.4 on page 54—were carried out on a grid of dimensions $65 \times 65 \times 97$ with multiple focusing steps into an area of size 1 nm $\times$ 1 nm $\times$ 1 nm, centred on the position of the test ion. The temperature was set to $T = 300$ K, the dielectric constant of the solvent was taken to be $\epsilon = 78.5$ and $\epsilon_p = 2.0$ for the protein/membrane mimetic; the ionic strength of the solution was set to 1 mol $\cdot$ l^{-1} with Stern radii $r_{Na^+} = 0.095$ nm and $r_{Cl^-} = 0.181$ nm, and the radius of an implicit solvent

*Test calculations on a model of the $\alpha 7$ receptor actually showed negligible dependence on the presence of a membrane mimetic (Kaihsu Tai, personal communication).

†The OPLS-AA force field was in part derived from AMBER so choosing AMBER for the Poisson-Boltzmann calculations should facilitate comparison between the MD with OPLS-AA and the PB free energy profiles.

Table 6.1: Equilibrium simulations of nAChR M2 (*P1* model). GROMOS96 (G43A1) is an united atom force field, only including polar hydrogens, whereas OPLS-AA models all hydrogen atoms explicitly. k_{bb} is the strength of the restraints on the backbone of the M2 helices, measured in $k_0 = 1.0 \times 10^3$ kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2}. Simulations were typically carried out for a total run length of $\mathfrak{T}$ in a NaCl solution of nominal concentration c(NaCl). Each system contained N atoms in total (including hydrogens).

force field	water	k_{bb}/k_0	$c(\mathrm{NaCl})/\mathrm{mol}\cdot\mathrm{l}^{-1}$	$\mathfrak{T}$/ns	N
GROMOS96	SPC	0.2	1.0	80	10729
		1.0	—	80	10773
		1.0	0.2	80	10557
		1.0	1.0	80	10557
OPLS-AA	SPC	1.0	1.0	60	15298
	TIP4P	1.0	1.0	55	18002

molecule 0.14 nm. The Born energy was calculated according to Eq. (2.115) on page 56.

6.2.4 Molecular Dynamics

Equilibrium properties of ions and water were studied both by equilibrium MD and biased MD simulations (umbrella sampling along the pore axis) with the aim of determining the density $n(z)$ and the PMF $G(z)$.

MD simulations were performed with gromacs version 3.1.4 or version 3.2.1[127] and the SPC water model[132] or TIP4P.[144,145] Calculations were carried out with two different force fields. Initially, the united atom force field GROMOS96 (G43A1)[132,133] was employed for equilibrium and umbrella sampling simulations. However, the question arose if explicit representations of hydrogens might be important. Hence the previous simulations were repeated with the all atom OPLS-AA/L[134,140] force field. The integration time step was 2 fs.* In both cases, explicit hydrogens were "dummified", i.e. their position is calculated from the geometry of the surrounding heavy atoms and not by forces on the H themselves, thus removing fast degrees of freedom. Typical simulations employed the following parameters; for variations see the list of equilibrium simulations of the *P1* system in Tab. 6.1. Backbone atoms were

*Larger time steps were also explored. $\Delta t = 5$ fs steps with GROMOS96 run stably but the behaviour of the water in the pore was qualitatively different from the $\Delta t = 2$ fs runs. Water evaporated much more readily from the gate region in the 5 fs simulations. It was decided that accurate representation of the solvent is important in our simulations and hence a (canonical) $\Delta t = 2$ fs was used.

Table 6.2: Parameters of the Umbrella sampling simulations. k_{bb} is the spring constant of the backbone restraints on the *P1* model of the nAChR TM pore. Umbrella sampling of an ion or a water molecule proceeded for a time $\mathfrak{T}$ after an equilibration phase of length $\mathfrak{T}_{eq}$. Altogether N windows of length Δz were sampled with a harmonic potential of force constant k. $k_0 = 1.0 \times 10^3$ kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2}. All simulations employed the SPC water model.

force field	k_{bb}/k_0	species	$\mathfrak{T}$/ns	$\mathfrak{T}_{eq}$/ns	N	Δz/nm	k/k_0
GROMOS96	0.2	Na^+	0.5	0.1	101	0.0495	0.997
	1.0	water	0.5	0.1	101	0.0544	0.824
		Na^+	1.0	2.0	200	0.025	3.910
		Cl^-	1.0	2.0	101	0.0495	0.997
OPLS-AA	1.0	water	1.0	0.2	101	0.0396	1.558
		Na^+	1.0	0.2	101	0.0396	1.558
		Cl^-	1.0	0.2	101	0.0396	1.558

restrained[171] with $k_{bb} = k_0 = 4.0\ kT$ Å^{-2} = 1.0×10^3 kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2} (except for one simulation at $k_{bb} = 0.2\,k_0$), giving them roughly the same positional root mean square fluctuation as seen in simulations of potassium channels.[201] Coordinates were saved every 2 ps. With periodic boundary conditions, long range electrostatic interactions were computed with a particle mesh Ewald method (real space cutoff 1 nm, grid spacing 0.15 nm, fourth order interpolation[164,165]) while the short ranged van der Waals forces were calculated within a radius of 1.4 nm. The neighbour list (radius 1 nm) was updated every 10 steps. Weak coupling algorithms[166] were used to maintain constant temperature ($T = 300$ K, time constant 0.1 ps) and pressure ($P = 1$ bar, compressibility 4.5×10^{-5} bar^{-1}, time constant 1 ps) with the x and y dimensions of the simulation cell held fixed at 4.6 nm. The length in z was 8.5 nm (ensuring bulk-like water behaviour far from the membrane mimetic and the protein surface). The *P1* system contained about 500 methane pseudo-atoms (or methane molecules for the all atom force field), which form the membrane mimetic, 3000 SPC water molecules and about 1400 heavy atoms. It was simulated for 80 ns with GROMOS96 and between 55 ns and 60 ns for OPLS-AA (see Tab. 6.1 for details).

Umbrella sampling Umbrella sampling[172] was carried out as described in Section 2.3.2. Parameters of the different simulation runs are listed in Tab. 6.2. The biased histograms were unbiased with WHAM.[175] Umbrella-sampled PMFs and PMFs from long equilibrium simulations were manually combined so that the barrier height, i.e. the free energy difference between a particle in bulk and at the centre of the barrier, could be determined. The umbrella-

sampled PMFs were individually corrected for drift by a linear function as described in Section 2.3.3. The axial density required for the equilibrium PMF Eq. (2.94) had to be computed differently for water and ions to obtain agreement with the umbrella PMFs. The water density $n(z)$ was calculated as $N(z)/V$ in 1.0 nm $\times$ 1.0 nm $\times$ 0.05 nm square "disks". The ionic density was obtained from an average of the local density over the disk volume $\pi R_G \times 0.05$ nm (Eq. (2.49) on page 31), where R_G is the fluid radius of gyration (2.48). For comparison, the water density was also calculated in the same way as the ionic density and plotted in Fig. 6.9 and Fig. 6.13.

6.3 Results

First we list the protonation states obtained from the electrostatic calculations. They set the stage for the equilibrium MD simulations, where the influence of the protein environment on the electrolyte will become apparent (Section 6.3.3). The equilibrium calculations constitute the baseline for the free energy calculations of Section 6.3.4. The final section compares results obtained with two different force fields, the all atom force field OPLS-AA and the united atom one GROMOS96.

6.3.1 Protonation states

Protonation states were assigned after a pK_a calculation in 0.144 mol $\cdot$ l^{-1} salt. All titratable groups facing the pore lumen are in their default state at pH 7, resulting in strongly negatively charged rings at the top and at the bottom. The fully charged intermediate ring ($-2'$, using the notation in Ref. 275) agrees with substituted cysteine accessibility method (SCAM) data.[275]

It is somewhat surprising that the residues in the rings, which are close to each other, can actually maintain a charged configuration. This results in a total charge of -10 in the entrance/exit regions of the TM region (excluding the intracellular ring in the M2-M3 linker, which is not included in the *P1* model). It is consistent with a channel selective for cations. The intermediate ring is also stabilised by inter-subunit salt bridges to adjacent lysines (or arginines), which are highly conserved in all LGICs (Fig. 6.1). Only seven acidic residues in the outer scaffold adopt a non-default (neutral) charge state but for the M2-only *P1* model they play no role.

6.3.2 Influence of the outer scaffold

Reducing the transmembrane domain to just the M2 bundle is a potentially drastic simplification. There might exist long ranging interactions between the

Table 6.3: Protonation states of titratable residues in M2 at pH 7. There are four acidic rings at the height or above the bilayer surface. Negatively charged sidechains are set in **bold** type. Residues are counted within M2 by the primed numbering scheme which designates αM243 the number $1'$.[274,275] Numbering of residues is taken from the pdb file 1OED.[29] The second extracellular ring at $24'$ is specific to some nicotinic receptors like the muscle nAChR or the $\alpha 7$ receptor.

Description	Position	charge	α_δ	γ	α_γ	β	δ
(2nd EC ring)	$24'$	-3	S266	**E275**	S266	**E272**	**E280**
extracellular ring	$20'$	-3	**E262**	Q271	**E262**	**D268**	Q276
intermediate ring	$-2'$	-4	**E241**	Q250	**E241**	**E247**	**E255**
intracellular ring	$-5'$	-4	**D238**	Q246	**D238**	**D244**	**E252**

scaffold (the M1, M3, and M4) helices and the ion pathway, which affect permeation. (We are not concerned with interactions that position the M2 helices because this effect is approximated by restraining the M2 backbone to the EM structure positions.)

Long range effects would most likely be dominated by electrostatic effects. In order to assess the significance of the outer scaffold for ion permeation, the electrostatic PMF (or Born profile) of a Na^+ ion was calculated for the full TM domain and the M2 bundle (the *P1* model) only. As Fig. 6.6 shows, the difference in electrostatic free energy between the *P1* model and the full TM domain is negligible. The only region where differences are detectable is at the intracellular end where ions in the full TM domain are stabilised by about 0.5 kT compared to the M2 bundle. This is due to the intracellular ring (cf. Tab. 6.3), which carries a charge of $-4e$ in the full structure but which is absent in the M2 bundle. Thus, the influence of the scaffold is shown to be very small, which lends some justification to proceeding with the M2 helix bundle model.

Another potentially serious simplification is the omission of the synaptic (extracellular) ligand binding (LB) domain. For the *Torpedo* nAChR there is no model available that contains both TM and LB domains.* Recently a homology model of the $\alpha 7$ receptor (another member of the family of nicotinic receptors) was presented,[41] which only misses the intracellular domain. Preliminary calculations for the electrostatic PMF in the $\alpha 7$ receptor indicate that the LB domain does have an influence on the Born profile (see Appendix A). The barriers are consistently higher in the presence of the LB domain than in their absence (Fig. A.2). It is expected that this is also the case for *Torpedo* nAChR, even though

*The structure of the intracellular basket or hanging gondola is not known at atomic resolution though it is possible to capture general characteristics like dimensions, openings, or charge decoration in coarse grained models, based on pseudo-atoms (O. Beckstein, unpublished).

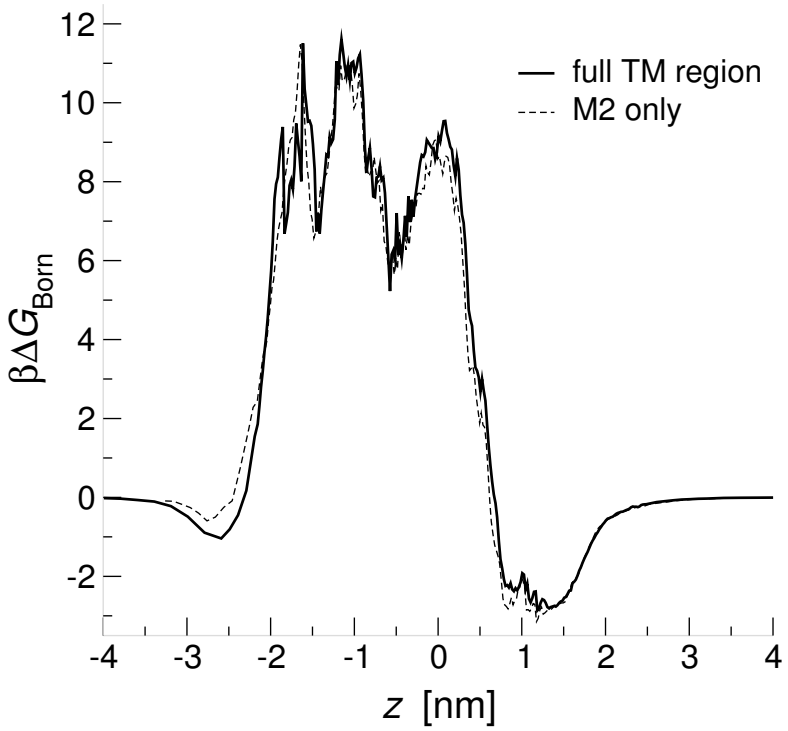


Figure 6.6: Born profile of Na^+ in nAChR TM and M2

the details of the PMF in the M2 regions differ. At this point it is not clear how big the effect in a full MD PMF would be although the main effect would likely be an overall *increase* in barrier height.

In passing we note that the Born calculation show an attractive well of about $-3\ kT$ at the extracellular side, which indicates strong stabilising effect exerted by the extracellular negatively charged rings on the ion. The overall barrier structure reflects the pore profile (compare Fig. 6.6 to Fig. 6.3 on page 126). The most narrow regions contain least solvent and hence the Born energy is highest—the same effect seen in the model pores on page 113. The results of Chapter 5, however, caution us not to take the electrostatic barriers at face value, and indeed, results from MD presented in the following sections do exhibit some important differences.

6.3.3 Equilibrium density

From the long equilibrium simulations (Tab. 6.1) we compute the density. The change in density compared to the bulk reveals to which degree the local protein environment influences solvent and ions.

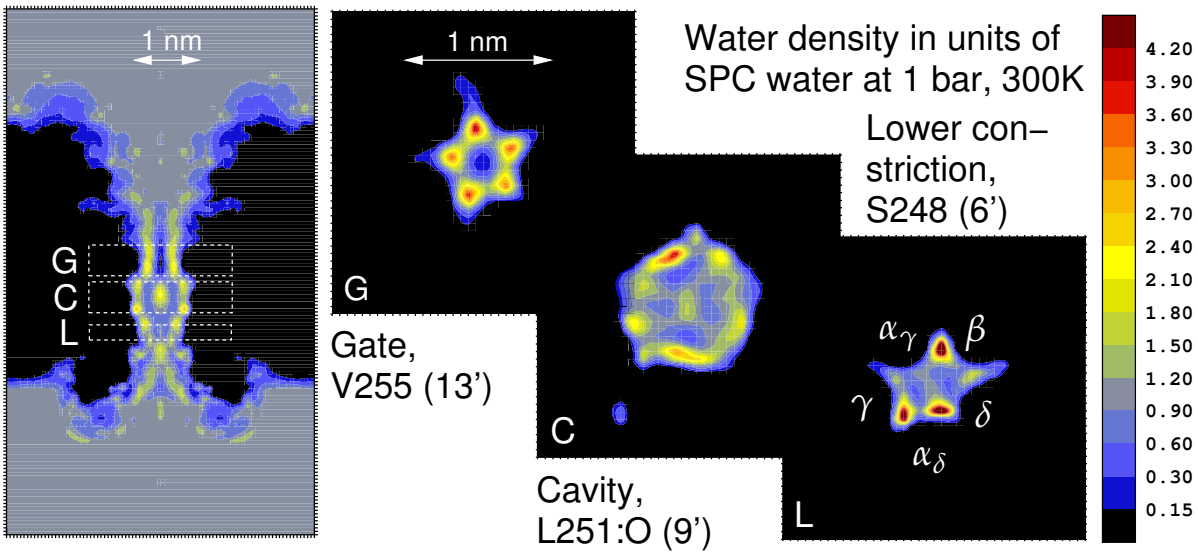


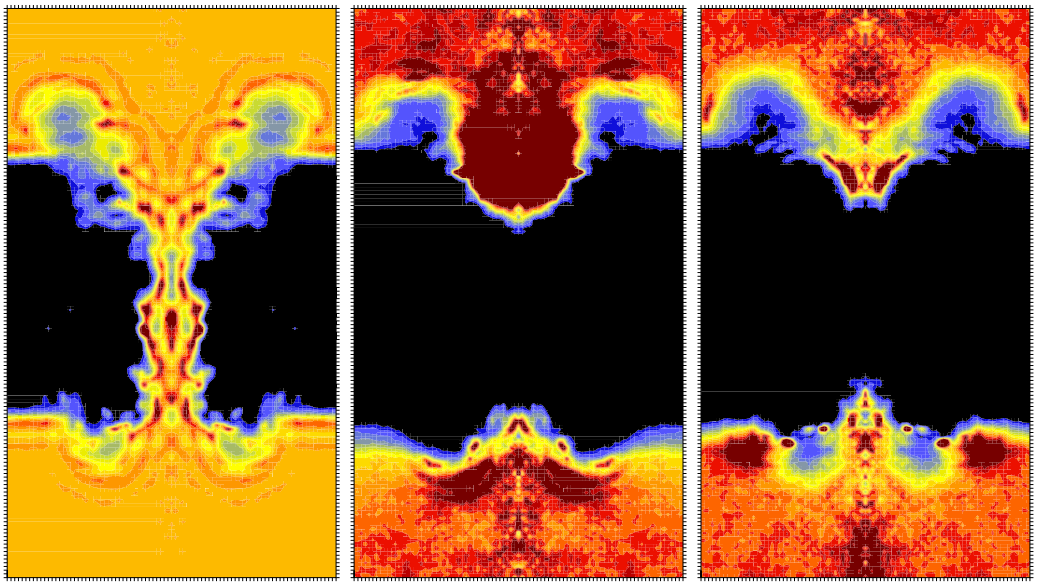
Figure 6.7: Water density in nAChR. The putative hydrophobic gate[29] *G* is formed by hydrophobic sidechains (αV255–αV259). *C* designates the "cavity", lined by backbone carbonyl group of L9′ (αL251), *L* the lower constriction site formed by Ser and Thr residues. (Force field GROMOS96 with similar results for OPLS-AA.)

Water Water in the closed receptor is shown in Fig. 6.7. The nAChR pore is much more polar than the hydrophobic model pores due to the presence of the backbone. So it is not too surprising that water can be seen in all regions of the pore.The putative *gate* region around αV255 is of particular interest. As seen in the cross section, water density is localised in a pentameric pattern, leaving a hollow tube.This hints at strong constraints for the water position: Water can only maintain its presence if it interacts in a fairly restrictive configuration.* However, these "rings" are not pentameric rings of water molecules hydrogen bonded to each other. The picture only emerges on average. Typically, a string of water molecules rotates in the pore and the string switches between five different favourable positions. Similar behaviour of water was observed in the putative gate region of MscS.[56] Inspection of the trajectories also shows evaporation and condensation events in the hydrophobic girdle although the vapour phases are shorter than in the case of the model pores.†

A slightly wider region (*cavity*), which is lined by the backbone oxygens of αL251, contains water at bulk density. The lower constriction site is seen to be the narrowest part (for water at least), where water molecules will only stay in very specific locations, namely close to βS250 (between α_γ and β, top of

*Cursory investigation suggests loss of hydrogen bonds above and below the gate.

†Anishkin and Sukharev[56] observe even more pronounced vapour phases in the putative hydrophobic gate of MscS and term it a "vapour-plug" or "vapour-lock" mechanism.[46,96].



(a) Water (b) Na^+ (c) Cl^-

Figure 6.8: Radially averaged density of 1.3 mol $\cdot$ l^{-1} sodium chloride solution in nAChR (M2 only, i.e. the *P1* model). OPLS-AA force field. which treats all hydrogens explicitly. The colour scale is the same as in Fig. 4.6 on page 100.

Fig. 6.7L) and in the vicinity of αT244 and γT253.

Ions Equilibrium MD simulations of electrolytes typically suffer from very long equilibration times. If there are also high barriers for the ions to overcome then they will not sample all available configurations sufficiently. By using a high salt concentration one can achieve more sampling in a given simulation time. Thus simulations were performed for a nominal 1 mol $\cdot$ l^{-1} NaCl solution; results for one at 0.2 mol $\cdot$ l^{-1} NaCl look similar albeit at a lower density, i.e. less frequent presence of ions in the pore. It turns out that the bulk density is *not* 1 mol $\cdot$ l^{-1} even though the number of ions per water molecule was adjusted to give this value. The densities from the simulations show a bulk density of about 1.3 mol $\cdot$ l^{-1} (Fig. 6.8). The reason lies in the formation of an ion exclusion layer near the membrane and protein surface and in the pore itself (cf. the situation encountered for the model pores in Section 4.3.2). Ions are forced into bulk regions, and, because of the limited size of the simulation box, increase the bulk density noticeably. This is not a problem as long as one uses the real bulk density $n_0 = 1.3$ mol $\cdot$ l^{-1} when calculating the equilibrium PMF Eq. (2.94).

The density of ions in the M2 pore (Fig. 6.8) vanishes in large parts of the pore region. Effectively, the pore is blocked for both ionic species, which is similar to

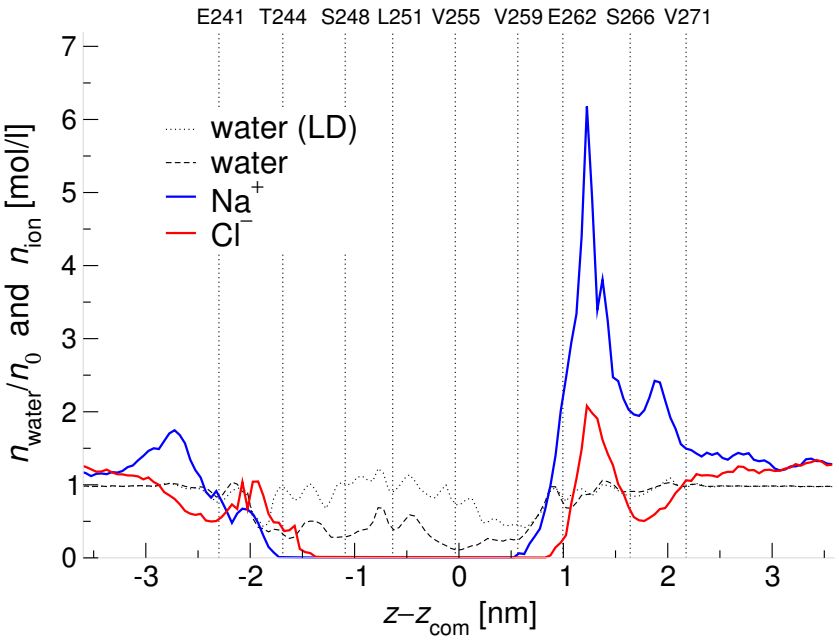
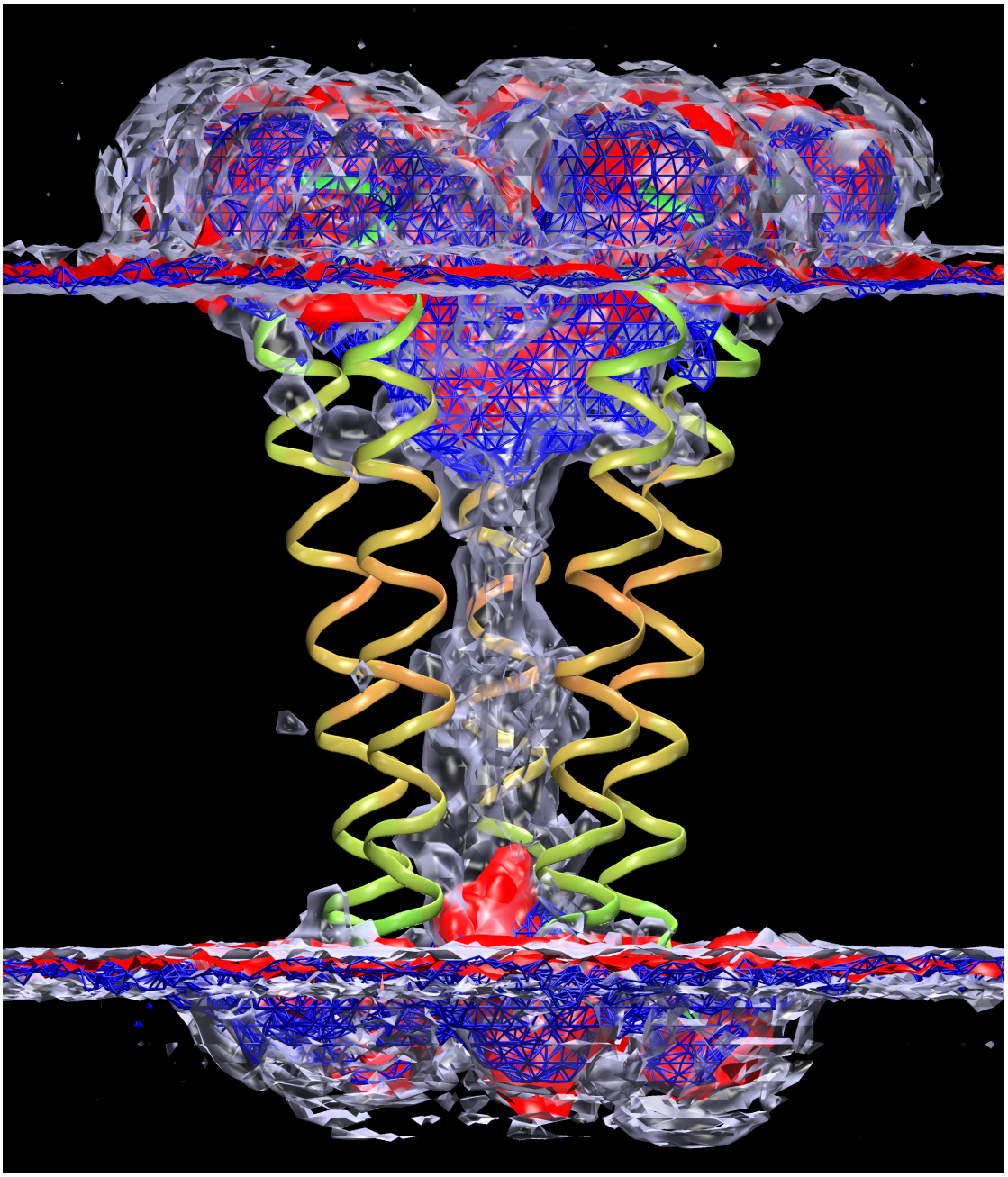


Figure 6.9: Axial density $n(z)$ in nAChR with the OPLS-AA force field and SPC water. For *ions* and *water (LD)*, $n(z)$ was computed from the local density $n(\mathbf{x})$ according to Eq. (2.49); for *water* it was calculated from averages in quadratic disks of dimensions 1 nm $\times$ 1 nm $\times$ 0.05 nm. Water density is given in units of the bulk density of SPC water n_0. Important pore lining residues of the *P1* model of nAChR are indicated. The hydrophobic girdle lies between V255 and V259 and the lower constriction site is located at S248. The charged extracellular rings lie at E262 and S266 and carry a total charge of $-6e$.

the situation encountered in the narrow model pores (Fig. 4.6 on page 100). The charged rings exert a large influence on the Na^+ ions with high concentrations in the extracellular vestibule and the intermediate ring. Cl^- ions are attracted to the intracellular ring of lysines and also to the extracellular vestibule. The situation is more clearly borne out in the axial density profile $n(z)$ (Fig. 6.9). The densities of both sodium and chloride ions drop to almost vanishing values in the TM region (stretching from T244 to V259). This is in stark contrast to up to six times the bulk concentration in the directly adjacent extracellular region where the extracellular, negatively charged rings concentrate Na^+ ions (similar to the results from the Poisson-Boltzmann calculations). The Cl^- concentration rises to twice its bulk value due to the presence of the Na^+ ions. The anionic density is hidden within the cationic one, thus forming a cylindrical double layer (Fig. 6.8 and Fig. 6.10).

Water is present in the pore at bulk density; the gate region (αV255–αV259), however, contains water in five high density columns whereas the centre ap-



pears void of water (Fig. 6.7G). There are also water bridges stretching between the M2 helices at the extracellular side (Fig. 6.10 on the preceding page). At a bulk concentration of 1.3 mol $\cdot$ l^{-1} NaCl, the ionic density in the gate drops to below 0.1 mol $\cdot$ l^{-1} during 60 ns of equilibrium MD.

The density profiles do not allow us to pinpoint the gate but they do show that there exists a barrier, and they can clearly show regions which attract ions. But in order to investigate the barrier, ions have to be forced to occupy these regions. In the next section the umbrella sampling technique is employed for this purpose. It allows us to probe the barrier structure of the closed nAChR pore.

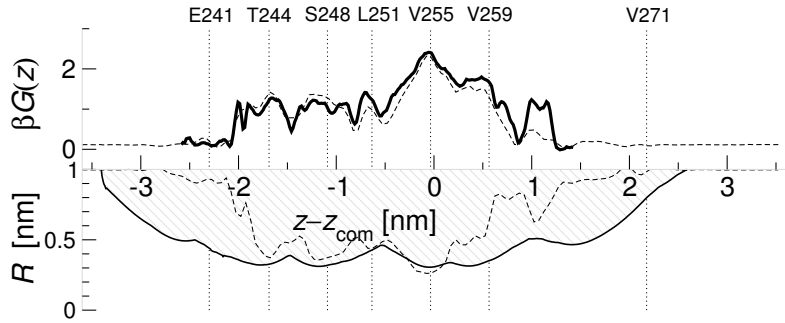
6.3.4 Potential of mean force

Water Unlike ions, water permeates the closed nAChR pore freely. Thus the average free energy of a water molecule along the pore (its potential of mean force) can be easily calculated from the equilibrium density as

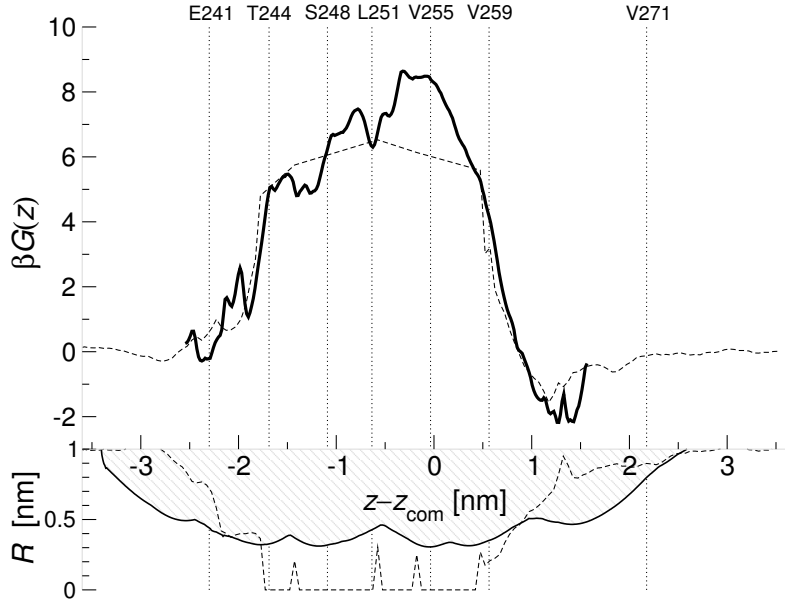
$$\beta G(z) = -\ln \frac{n(z)}{n_0} + \beta(G_0 - \mu_0) \tag{2.94}$$

throughout the whole pore region. (The normalisation with the excess chemical potential for the insertion of the water molecule into bulk, μ_0, defines $G(z)$ to be 0 in the bulk.) The umbrella-sampled PMF of a water molecule and the one derived from the equilibrium density should coincide. This is indeed the case, as Fig. 6.11(a) shows. The maximum barrier height is only about 2 kT. This is well within reach of thermal fluctuations and explains why the pore is solvated most of the time. The peak of the barrier is directly located on αV255 and extends through the hydrophobic girdle. The other constriction sites at αS248 and αT244 only command barriers of about 1 kT—they are effectively not noticed by water molecules which can hydrogen bond to the hydroxo groups of the side chains.

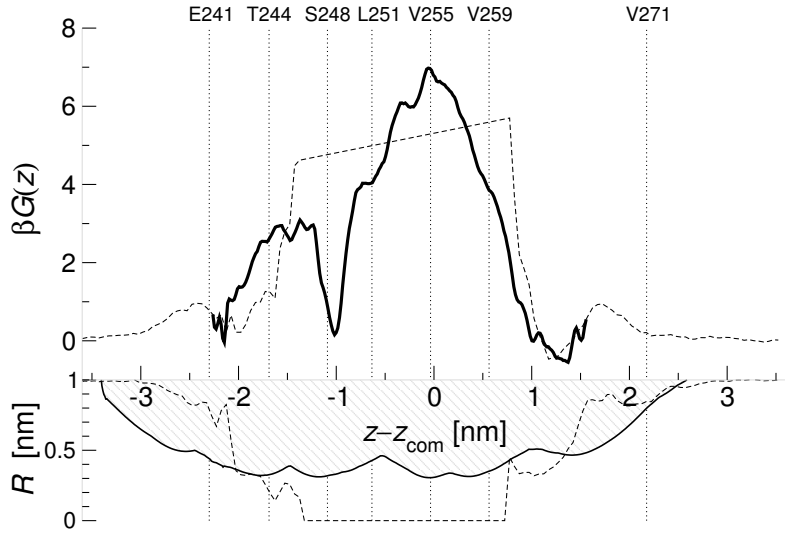
Figure 6.10: 3D density in the nAChR gate region from equilibrium MD simulations in 1.3 mol $\cdot$ l^{-1} NaCl. The M2 helices are rendered as ribbons. Water is contoured at 0.93 of the bulk density. The Na^+ density is rendered as a blue mesh at a contour of 0.1 mol $\cdot$ l^{-1}. The Cl^- density is contoured as a solid red surface at the same density. The pore is filled with water which displays a columnar structure. Ions do not permeate the pore so that it is not possible to pinpoint a location for the gate (the hydrophobic girdle is in the region of the red shading of the helices). In the vestibular region (top) the sodium density "envelops" the chloride density, creating a rolled up double layer. (Image created with vmd[43])



(a) Water



(b) Na^+



(c) Cl^-

From the water density one can also compute a ensemble pore profile, based on the (local) density (Eq. (2.48) on page 30). It describes up to which average radius (as a function of z) water fills the pore. It is shown in Fig. 6.11(a) along with the static hole-based profile. Both profile attribute the same radius to the constriction sites and the cavity (lined by αL251). For the intra- and extracellular vestibular regions the ensemble pore profile indicates radii larger than the static solvent excluded surface, typically by about 0.5 nm. This reflects the fact that water molecules are able to penetrate the inter helical spaces; the static profile does not allow for flexibility and thus misses this effect. Both calculations do agree on the central region of the pore, which indicates its compactness and well-defined shape.

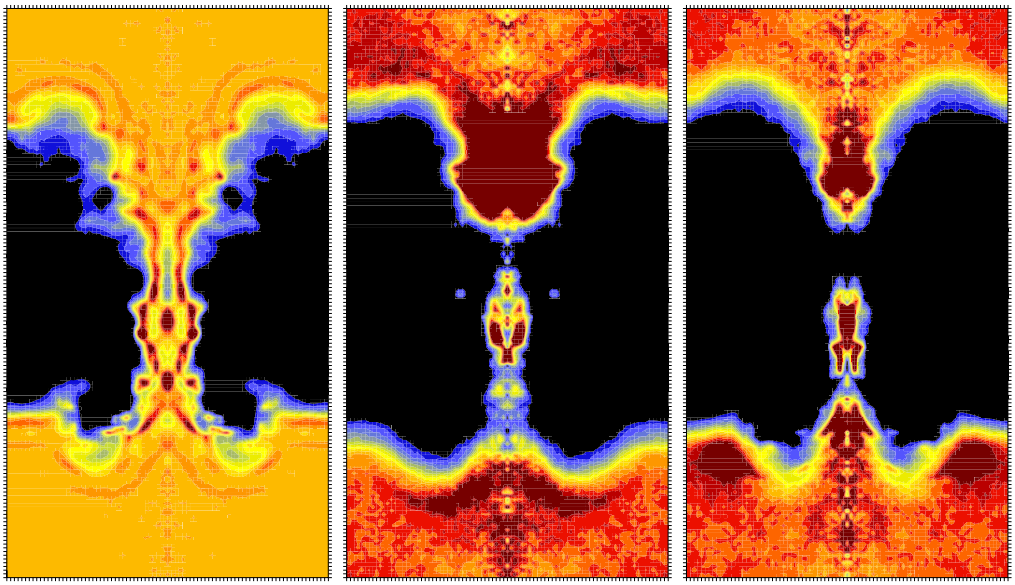
Ions The equilibrium density for ions is only well defined in the bulk and the vestibular regions of the pore. From the biased MD simulations PMFs are obtained in the low density/high free energy region, shown in Fig. 6.11. The equilibrium PMF and the umbrella-sampled PMF superimpose reasonably well in the vestibular regions but because the superposition procedure is carried out manually there is still room for some variation and hence one should expect overall errors of about 1 kT in the absolute height of the ionic PMFs relative to bulk.

For both Na^+ and Cl^- there are large barriers apparent. Na^+ ions would need to overcome a maximum barrier of 8.7 kT between αL251 ($9'$) and αV255 ($13'$, see Fig. 6.11(b)), whereas the Cl^- barrier is slightly smaller with 7.0 kT at αV255 (Fig. 6.11(c)).

The PMF for Na^+ at the αS248 constriction is 6.6 kT and only 5.1 kT at T244, even though the radius is also only 0.31 nm (Fig. 6.11(b)). The hydrophobic girdle, despite having the same pore radius as the intracellular constriction sites, has a more profound influence on ion permeation than the polar constrictions. The PMF suggests that the gate in the closed nAChR pore is located at the hydrophobic girdle.

On the extracellular side of the gate an attractive well for Na^+ can be seen. It has a depth of about -2 kT and is due to the extracellular acidic rings. This "moat" in front of the hydrophobic barrier increases the barrier for ions moving

Figure 6.11: PMF of ions and water in the nAChR pore. Bulk NaCl concentration is 1.3 $\mathrm{mol} \cdot \mathrm{l}^{-1}$. MD simulations and umbrella sampling were performed with the all-atom OPLS-AA force field. The equilibrium derived PMF is shown as a dashed line and the Umbrella sampled PMF as a heavy line. The pore radius (from hole) is shown as the filled outline, with the ensemble (or local density derived) pore profile, Eq. (2.48), as a dashed line. Some pore lining residues (on the α subunit) are indicated. The hydrophobic girdle stretches from αV255 to αV259; αL251 lines the cavity.



(a) Water (b) Na^+ (c) Cl^-

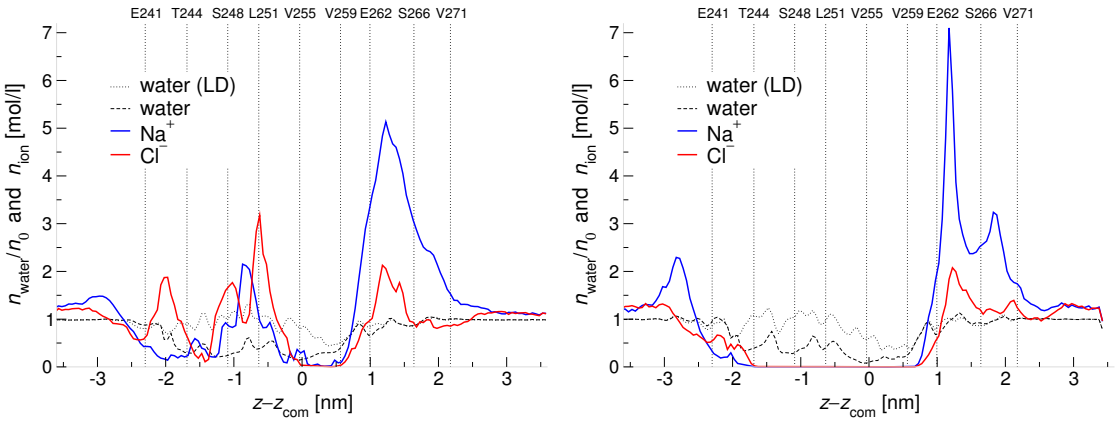
Figure 6.12: Radially averaged density of 1 mol $\cdot l^{-1}$ sodium chloride solution in nAChR (M2 only, i.e. the *P1* model). GROMOS96 force field, which only explicitly represents polar hydrogens and treats other hydrogens together with the heavy atoms as united atoms.

from the synaptic to the cytosolic end to an effective height of ca. 10 kT. The ensemble pore profile for ions is less useful than for water because the density is not well defined in the pore due to insufficient sampling. For the vestibular regions, though, it indicates that ions penetrate the protein wall much deeper than anticipated from the static radius.

The barrier to Cl^- permeation in the closed nAChR is smaller than the one for Na^+ ions, and there is even a weakly attractive well near αS248 visible. Cl^- are also attracted to the negatively charged extracellular rings, where there is an attractive well of -0.5 kT. As explained above, this is a double layer effect, depending on the presence of Na^+ ions. The closed nAChR state probably does not have to block anions as effectively as cations because anions cannot overcome the barrier at the selectivity filter. The position of the filter is not clear though it could be located in the windows of the intracellular "hanging-gondola" structure.[289]

6.3.5 Sensitivity to the force field and backbone restraints

The results of MD simulations depend on a number of parameters. The most influential is the set of parameters that describes the interactions between the



(a) GROMOS96, SPC water (b) OPLS-AA, TIP4P water

Figure 6.13: Axial density $n(z)$ in nAChR when using alternative force field parameters. See the caption of Fig. 6.9 on page 139.

atoms, i.e. the force field. It contains the parameters for the protein, the water model, and the ions. Ideally, all well-tested force fields should give the same answer for a given problem. Unfortunately, this is not always the case. PMFs in particular are known to be sensitive to the force field that is used.[298] In order to discuss the robustness of the results presented in the previous sections, which were obtained with the all-atom OPLS-AA force field, we also show simulations carried out with the united atom GROMOS96 (G43A1) force field. Briefly, some results for OPLS-AA with the TIP4P water model instead of SPC are also mentioned. Lastly, the effect of restraining the protein backbone of the M2 helices is discussed.

All-atom versus united atom Clear differences between the GROMOS96 and OPLS-AA simulations are already visible in the axial density $n(z)$, Fig. 6.13(a) compared to Fig. 6.9 on page 139. The all-atom simulations do not show any ionic density in the pore but with GROMOS96 high concentrations of both Cl^- (about 3 mol $\cdot$ l^{-1}) and Na^+ (2 mol $\cdot$ l^{-1}) are apparent in the cavity near αL251. The simulations started with ions in the pore but whereas they are expelled from the pore in the first 0.1 ns of the OPLS-AA simulations they seem to persist or even exchange in the united-atom simulations. Though the behaviour of water does not seem to differ with resepct to the density it was nevertheless observed by inspection of the trajectories that water (both SPC and TIP4P) with OPLS-AA is much more volatile in the hydrophobic gate region than with GROMOS96. It does exhibit liquid-vapour transitions though not as pronounced as in the model pores or as in the MscS gate.[56] SPC water with GROMOS96, on the other

hand, solvates the pore during the whole simulation time.

In both cases, though, the ionic density vanishes at the hydrophobic gate. If anything, the GROMOS96 density Fig. 6.13(a) shows even clearer the position of the gate.

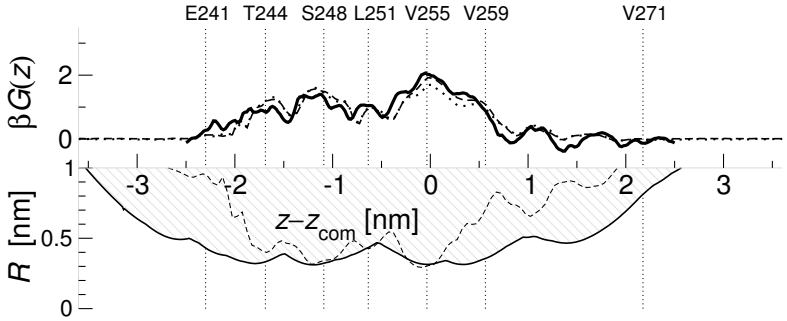
The higher density in the cavity region in GROMOS96 simulations does not appear to be an artifact of initial conditions. The Na^+ PMF exhibit a corresponding double peak structure with an attractive valley in the cavity (Fig. 6.14(b)). The barrier at the hydrophobic gate (7.2 kT) is about the same height as the second barrier between αT244 and αS248 with 6.8 kT (within estimated errors of 1 kT); both barriers are lower than the Na^+ barrier in OPLS-AA with 8.7 kT. The Cl^- barrier, Fig. 6.14(c), has one pronounced peak at the same location at αV255 as in the OPLS-AA simulations, but it is higher with 8.5 kT compared to 7 kT, and it lacks the attractive well near αS248.

Water model and ionic strength The use of the TIP4P model with OPLS-AA instead of SPC did not lead to dramatic changes in the density, as can be seen by comparing $n(z)$ (Fig. 6.13(b) with Fig. 6.9 on page 139) or the radial density $n(r, z)$ (Fig. 6.12(a) with Fig. 6.8(a) on page 138). This was taken as an indication that the PMF would probably not differ very much from the SPC/OPLS-AA case.

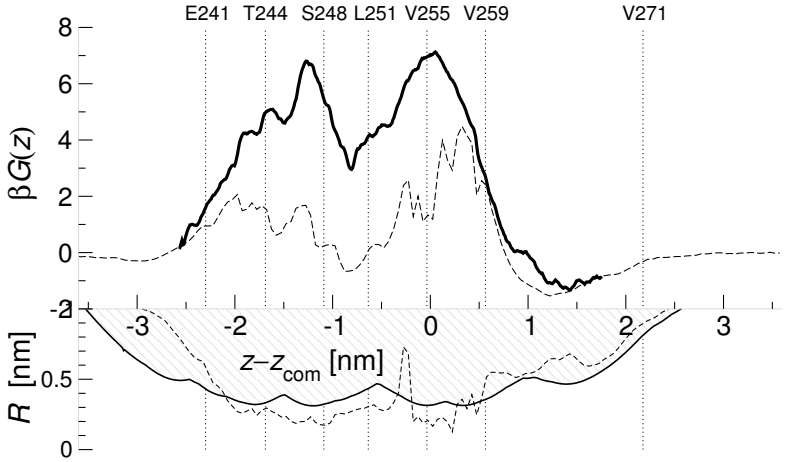
For the GROMOS96 case we also assessed the effect of the presence of ions on the water PMF. In Fig. 6.14(a) the equilibrium density derived PMF of water in 1.3 mol $\cdot$ l^{-1} NaCl is compared to the PMF of a water molecule in pure water. The difference is barely noticeable, which suggests that the influence of ions on water structure in the pore is actually very small.

Backbone restraints In Section 6.2.2 it was described how the full TM domain of nAChR was reduced to the *P1* model. In order to retain the experimental structure, the backbone atoms of the M2 helices were harmonicly restrained with force constant $k_{bb} = k_0$, which could be a source for artifacts in the PMF. To assess the influence of the restraints we also computed one PMF for a Na^+ ion at much weaker restraints $k_{bb} = 0.2k_0$. Both PMFs are very similar (Fig. 6.15), and within an estimated error of 1 kT indistinguishable. We believe

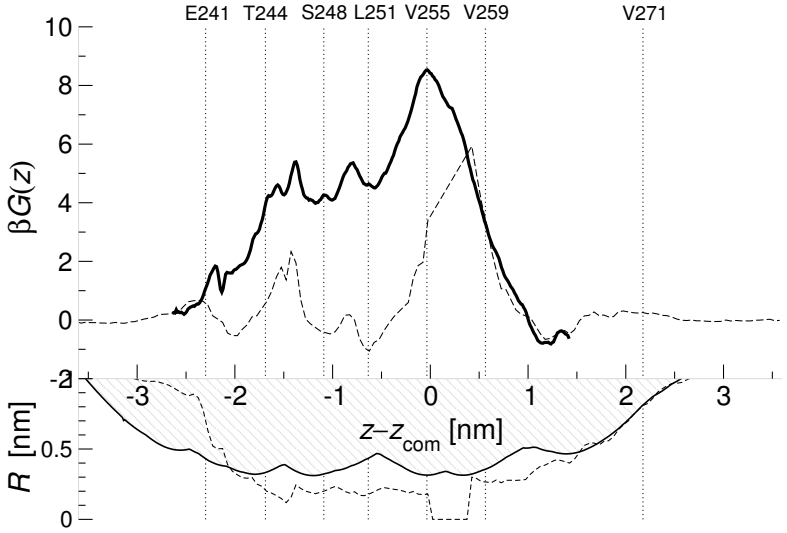
Figure 6.14: PMF of water and ions in the nAChR pore at a bulk NaCl concentration of about 1.3 mol $\cdot$ l^{-1}. MD simulations and umbrella sampling were performed with the GROMOS96 force field. The equilibrium derived PMF is shown as a dashed line and the Umbrella sampled PMF as a heavy line. For water (a), there is also shown the equilibrium density for nAChR solvated in pure water (dotted line). The pore radius (from hole) is shown below, together with the ensemble pore radius. Important pore lining residues (on the α subunit) are indicated.



(a) Water



(b) Na^+



(c) Cl^-

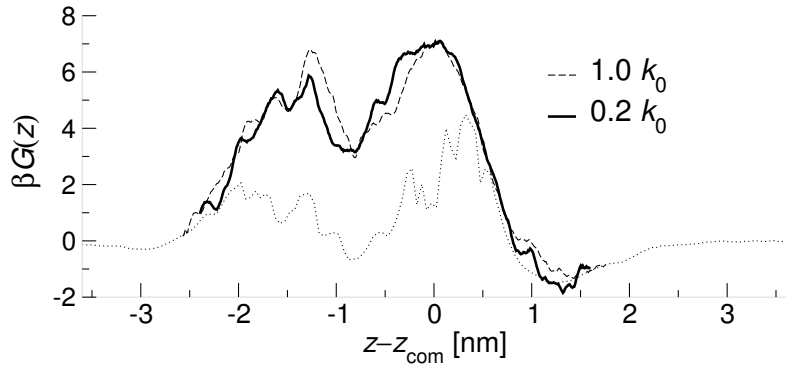


Figure 6.15: PMF of Na^+ in nAChR *P1* for two M2 backbone restraints. The PMF for weak backbone restraints with harmonic force constant $k_{bb} = 0.2k_0$ is very similar to the one with the standard restraints, $k_{bb} = k_0 = 1.0 \times 10^3$ kJ · mol^{-1} · nm^{-2}, as shown in Fig. 6.14(b). The dotted line indicates the equilibrium density derived PMF. (GROMOS96 force field with the SPC water model)

that this result is not primarily due to the relatively short simulation time (ca. 1 ns) per umbrella window (compared to the time needed for larger protein re-arrangements) because the starting frames for these windows were different snapshots from a 80 ns equilibrium simulation. Thus, for each window the ion does sample an environment characteristic of the more flexible model and not simply an environment which is initially similar to the less flexible *P1* model but does not change much during the sampling time. This justifies *a posteriori* the use of restraints. Furthermore, it indirectly indicates that the flexibility of the pore wall itself is accounted for by the movement of the side chains lining the ion pathway.

6.4 Discussion

All the preceding results are based on the assumption that the structure of the transmembrane domain of nAChR[29] faithfully represents the closed state. In fact, nAChR and the other LGICs do not only have one closed (resting) state but also a closed desensitized state[299] and Karlin and coworkers argue that the gate is in different positions in these two closed states.[281,282] The cryo-EM images were taken in the absence of ACh[29] so the EM structure should represent the resting state as desensitization requires prolonged exposure to agonist. The resolution of the structure is not very high at 4 Å—just good enough to trace the backbone with some confidence. But assuming that the backbone is correctly described, the MD simulations allow the sidechains to re-arrange themselves. Because the PMF is not sensitive to the strength of the backbone restraints it fol-

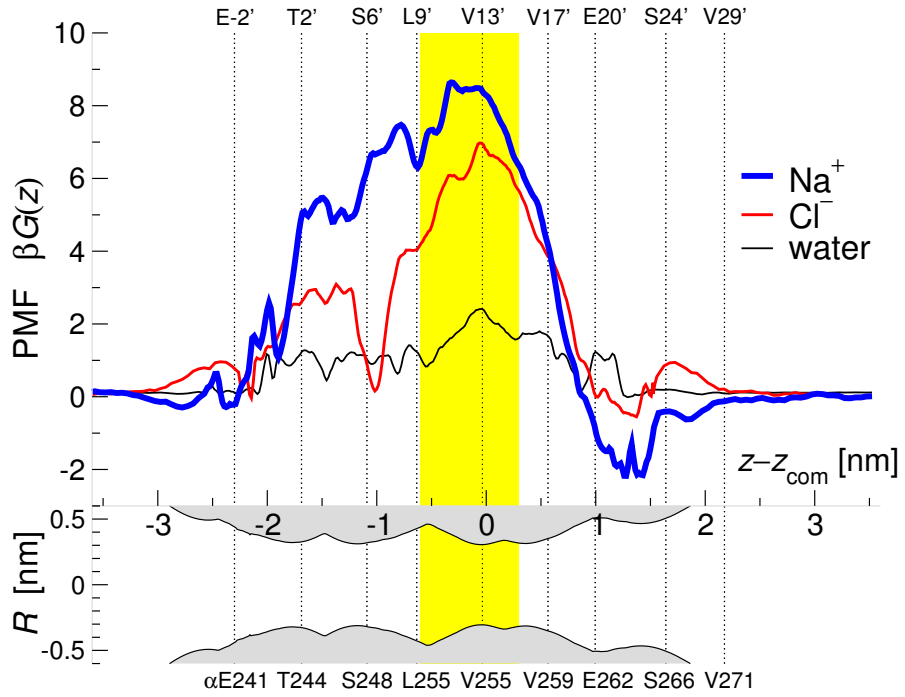


Figure 6.16: Potential of mean force of ions and water in nAChR. Na^+ ions encounter a wide barrier of up to 8.7 kT, which is peaked at the hydrophobic girdle (L9′–V13′)—the *hydrophobic gate* (marked in yellow). The barrier for Cl^- is somewhat lower which implies that the anion filter is located in other parts of the permeation pathway. Water solvates the pore because the barrier is not greater than 2 kT. The pore profile is shown in the lower panel, mirrored at the axis to create the impression of a section through the nAChR pore.

lows that the local pore environment, which is formed by the mobile sidechains, assumes a conformation independent of the details of the backbone motions. It appears that molecular dynamics fills in the functional details which could not be resolved through experiment.

The Na^+ PMF (Fig. 6.16) shows a pronounced barrier of about 8.7 kT at the hydrophobic girdle (between αL251 = 9′ and αV255 = 13′). Therefore, the hydrophobic girdle is identified with the hydrophobic gate of nAChR, as hypothesised by Unwin and colleagues,[29,61] who based their inference on the protein structure. This finding disagrees with the results of Karlin and coworkers[275,281,282] who predict the position of the (resting) gate at about αT244 = 2′, based on biochemical data (accessibility of substituted cysteines to small, positively charged, sulfhydryl-specific reagents). However, using the same method in the closely related $5HT_{3A}$ receptor, Slesinger et al.[286] find evidence for a gate between residues 9′ and 13′, and further experiments indicate that the narrow

constriction site near 2′ remains unchanged between the open and the closed state[287]—the very same region that Karlin[275,281,282] identifies as the gate. A gate at the position of the hydrophobic girdle is also consistent with experimental results by Auerbach and colleagues who describe gating in nAChR as a wave of conformational change[284] propagating from the ligand binding domain to the middle of the M2 segment, hence implying a gate near 9′.[285] Based on functional and behavioural studies of mutations in the muscle[283] ($\alpha_2\beta\gamma\delta$) and brain ($\alpha4\beta2$) nAChR[300,301] of mice, Lester and coworkers attribute a central role in gating to Leu9′. A Leu to Ser mutation in the $\alpha4$ receptor confers hypersensitivity to ACh and constant activation, leading to behavioural abnormalities and postnatal lethality in knock-in mice.[300] A less severe Leu to Ala mutation leads to a 50-fold increase in sensitivity to nicotine.[301] Mice bearing the mutation are otherwise healthy (in contrast to the Leu9′Ser mice). Not only do these gain-of-function experiments hint at an explanation of nicotine addiction (the $\alpha4\beta2$ nAChR, which is part of the dopaminergic reward system in the brain, is necessary and sufficient in mice for the behavioural effects of nicotine to be elicited) but also suggests that an increase in hydrophilicity in the gate region will lead to a greater likelihood of channel openings. This can be interpreted as a decrease in the barrier to ion permeation, which decreases with an increase in polarity (Ser, and to a lesser degree Ala) or with an increase in pore radius (Ala), which in turn provides more space for hydrating water molecules.

The question remains if the calculated barrier ($\Delta G^\ddagger = 22\ \mathrm{kJ\cdot mol^{-1}} = 8.7\ kT$ at $T = 300$ K) is actually large enough to prevent passage of ions. Allen et al.[170] computed K^+ PMFs in the gramicidin A channel and determined a barrier of 34 $\mathrm{kJ\cdot mol^{-1}}$ (12.4 kT at the temperature of the simulation, $T = 330$ K). They calculate a maximum conductance $g_{\max} = 0.8$ pS for the open channel.* We use their Eq. 5 (see also Eq. 18 in Ref. 169),

$$\begin{aligned} g_{\max} &= \frac{e^2}{kT}\frac{D_{\mathrm{eff}}}{L^2} \\ &= \frac{\beta e^2}{L^2}\left(L^{-1}\int_{\mathcal{P}} dz\, D(z)^{-1} e^{+\beta G(z)}\right)^{-1}\left(L^{-1}\int_{\mathcal{P}} dz\, e^{-\beta G(z)}\right)^{-1}, \end{aligned} \tag{6.1}$$

to estimate the maximum conductance of the *closed* nAChR pore. The spatial averages are carried out over the pore region $\mathcal{P}$ (from αE241 to αV271, $L = 4.5$ nm) with the Na^+ PMF $G(z)$ (Fig. 6.11(b)). The diffusion constant in the regions where the barrier is small ($G(z) < 1\ kT$) is taken as the experimental value of Na^+ at room temperature, $D(z) = D_0 = 1.33\ \mathrm{nm^2\cdot ns^{-1}}$ (Ref. 12). For $G(z) > 1\ kT$ (in the barrier region) we choose it to be a fraction α of D_0, $D(z) =$

*The experimental conductance for gA is 21 pS but Allen et al.[170] go to considerable length to demonstrate that their first-principles $g_{\max}$ of 0.8 pS should be considered a good estimate, an order of magnitude better than any previous ones.

αD_0. With a typical value of $\alpha = 0.5$,[95] we obtain $g_{\max} = 0.09$ pS, which is below the detection limit of current patch clamp single channel recordings.* At first glance it might seem surprising that although the maximum barrier height in nAChR (8.7 kT) is *lower* than in gA (12.4 kT), $g_{\max}$ is larger for open gA than for closed nAChR (0.8 pS and 0.09 pS). The difference lies in the shape and spatial extension of the PMFs. In gA, it is peaked at the centre and the region where $G(z) > 1\ kT$ extends over a length of only about $L = 1$ nm (Ref. 170, Fig. 2). The nAChR PMF (Fig. 6.11(b)) is much broader and extends over 2.7 nm which translates into a low conductance in Eq. (6.1). $g_{\max}$ is completely dominated by the average $\langle D(z)^{-1} \exp(+\beta G(z)) \rangle$, which in turn is determined by the region where the barrier is high. Because the barrier is large in the whole pore region this implies that our estimate $g_{\max}$ directly scales with the fraction of the diffusion constant, α. Even if we took $\alpha = 2$, i.e. diffusion were enhanced in the pore, the corresponding maximum conductance 0.18 pS would still be below the experimental noise limit. The rough estimate of the single channel conductance seems to indicate that the nAChR pore, as described by the Na^+ PMF, is indeed closed, especially when compared to the open state conductance of about 30 pS.[12]

For Cl^- the estimated maximum single channel conductance is $g_{\max} = 2.9$ pS (using $D_0 = 2.03\ \text{nm}^2 \cdot \text{ns}^{-1}$ (Ref. 12))—which shows that the barrier for chloride permeation is significantly lower than the one for sodium ions. However, the closed state should not allow the passage of either cations or anions. Because the preceding calculation shows that the closed state nAChR does not efficiently block Cl^- flow one must conclude that there exists *another* region in the receptor—the selectivity filter—which prevents permeation of Cl^-. This selectivity filter should be functional both in the closed and the open state of the channel. The most likely region for the filter are the charge-decorated windows in the intracellular vestibule.[289]

6.5 Conclusions

We tried to answer the question where the gate is located in the nicotinic receptor, i.e. we sought to identify the functional region which prevents ions from permeating the closed channel. We approached the question on the basis of the closed state cryo electron microscopy structure of Miyazawa et al.[29]

A direct computational method to find permeation barriers is to calculate the potential of mean force from fully atomistic molecular dynamics simulations, which, however, require significant computing resources. In order to make

*At a typical transmembrane voltage of $V = 100$ mV, a conductance of $g \approx 0.1$ pS would allow a current $I = gV = 0.01$ pA to flow (assuming ohmic behaviour). This is below 0.06 pA, the noise limit for current patch clamp experiments.[12]

the system tractable and achieve sufficient sampling we reduced the system size to the bare minimum. We argued that the bare minimum consists of only the conduction pathway lining M2 helices (with the backbone atoms harmonicly restrained to their experimentally determined position) and demonstrated the negligible influence of the outer helices on the (electrostatic) PMF. The ligand binding domain, which was not modelled in the PMF calculations, would probably increase the height of the barrier because it extends the low dielectric region in the vicinity of the pore, but more work is required to address this question. The cytosolic vestibule, the structure of which is not known at atomic resolution, has the shape of a sphere with inner radius of about 1 nm.[265] In the absence of any structure we can speculate that its main influence on ion permeation *in* the transmembrane domain would be similar to the one of the ligand binding domain, i.e. it would raise the existing barriers. Its particular shape and charge decoration might very well contribute significantly to overall conductance and selectivity.[289]

The Na^+ PMF shows a pronounced barrier of about 8.7 kT or 22 $kJ \cdot mol^{-1}$ at the hydrophobic girdle (αV255). In the absence of any other barriers, which, we argued, can only be located in the TM pore itself, this identifies the hydrophobic girdle with the hydrophobic gate of nAChR, consistent with structural[29,58] and most[283–287] but not all[281,282] biochemical evidence. The simulation results do not support the notion of a gate close to the intracellular mouth. Additional support for the plausibility of the PMF comes from a rough estimate for the single channel conductance of about 0.1 pS, which is very small compared to the experimental open state conductance of about 30 pS (Ref. 12) and below the detection limit of patch clamp recordings.

As a secondary result we established the function of the extracellular double ring and the intermediate/intracellular rings as concentrators of cations. By increasing the local ion concentration near the channel entrance, the protein lowers the effective access resistance and thus can increase the single-channel current once it opens.[12] Furthermore, the nAChR pore was shown to be predominantly filled with water in the closed state.

The shape of the PMF was found to be somewhat sensitive to the force field. We presented results for the all atom OPLS-AA force field because it is thought to better represent the pore environment. The results are also more in line with what one would expect from the studies on the simple hydrophobic pores, namely that for a given radius the hydrophobic region should be responsible for a higher permeation barrier than the polar one. Nevertheless, both force fields do exhibit a significant barrier at the hydrophobic girdle, which reinforces the main finding of this chapter: That there is a *hydrophobic gate* in the nicotinic receptor, and that we can understand its existence on the grounds of simple concepts.

7 Conclusions

Structural evidence suggests that the gate in many ion channels is formed by a hydrophobic constriction of the pore; for the known ion channels this appears to be the predominant form of the gate (the other being a block by a single sidechain). In the closed-state nicotinic receptor this constriction is wide enough to let three water molecules or a half-solvated sodium ion pass. The hydrophobic character of the sidechains lining the constriction appears to be responsible for the barrier to ion permeation and hence this was termed a *hydrophobic gating* mechanism. So far, evidence in the literature only consisted of static structure images[29,48,59,66] and the qualitative reasoning that the loss of the hydration shell of an ion would be energetically so costly as to prevent the ion from permeating.[12,67]

In this work only hydrophobic gating was considered; other authors have dealt with gates formed by local sidechains in the context of the chloride channels[31,302–307] and an outer membrane protein.[39] "Local gating" differs from hydrophobic gating in that only small structural rearrangements are required to open the gate, and the details of the local environment (the gating sidechain itself and local networks of hydrogen bonds and salt bridges) are crucial to explain the mechanism. Furthermore, permeating ions or protons themselves might be involved in the gating process. On the other hand, hydrophobic gating involves large scale movements of the protein to open the gate. The common feature of hydrophobic gates is a ring of hydrophobic residues but the exact composition is variable. This suggests that there is a common mechanism at work, a general principle, which has evolved in a wide range of systems. In this work an attempt was made to identify this mechanism, explain it, and quantify it, with the help of computational methods.

7.1 Water: Capillary effects at the atomic scale

In understanding ion permeation it is also necessary to understand the behaviour of the solvent, i.e. water. Water is not a simple liquid due to its strong hydrogen bonds (which, for instance, leads to the hydrophobic effect), and under standard condition it is also a liquid close to its liquid-vapour phase transition. In Chapter 3 and Chapter 4 water was studied confined to narrow hydrophobic pores of the dimensions of the gate of the nicotinic acetylcholine receptor. The radius of these pores ranged from the radius of a water molecule to seven times

as much. These model pores were kept conceptually and structurally simple in order to minimise the number of free parameters and so to be able to disentangle different contributions to a gating effect.

The water density in the pore did not gradually decrease with radius but two distinct states were found: The water was either liquid, comparable to bulk water, or it formed a vapour, but no intermediate state existed. Macroscopically, the phenomenon of capillary evaporation and condensation is well known for fluids confined to a pore. In this situation the interface between fluid and wall plays a more important role compared to volume properties. For a fluid close to a phase transition the interaction between the fluid molecules and the wall can drive the fluid across the phase transition. The observed oscillations between liquid water and water vapour in the microscopic pores could be explained as capillary evaporation/condensation, driven by thermal density fluctuations in the water outside the pore itself. Through a thermodynamic model the strength of the effect was related to the hydrophobicity of the surface, expressed as the contact angle or equivalently, the wall-water surface tension.

The simple model pores allowed to investigate different parameters that affected water transport, namely the geometry of the pore (radius and length), the chemical character of the surface (hydrophobic vs. hydrophilic), and the flexibility of the pore wall. Liquid water was forced out of pores with smaller radii, more hydrophobic and more flexible pore walls. Conversely, liquid water was stabilised in wide pores with hydrophilic and rigid pore walls. For the short (length 0.8 nm) hydrophobic pores studied here, the critical or transition radius between the vapour and the liquid regime was at about 0.56 nm while the wall atom position fluctuated by about 0.072 nm.

At least theoretically, this demonstrates that capillarity effects are present in atomic scale systems, such as fluid transporting conduits in nanotechnology, geology (water transport in rocks such as zeolites), and in biological protein channels. It seems that in the case of hydrophobic cavities *natura non abhorret vacuum*—or at least is content with filling it with the occasional vapour or gas molecule instead of with the surrounding water.

7.2 Ions: Dehydration barriers

Equilibrium molecular dynamics simulations of Na^+ and Cl^- ions in the model pores (Chapter 4) showed that ions would not permeate hydrophobic pores of radius less than about 0.6 nm over time-scales of about 100 ns, indicating that the hydrophobic constriction represents a high free energy barrier to ion permeation. This barrier was explicitly probed by calculating the potential of mean force for ion permeation from biased simulations (Chapter 5). For a pore of radius 0.3 nm, the barrier height measures almost $20\,kT$ (or about 50 kJ $\cdot$ mol^{-1}),

which renders the pore impermeable to ions having an average thermal energy of about $1\,kT$. Insight into the origin of this barrier comes from Poisson-Boltzmann calculations of the same system, in which the water is only treated as a high-dielectric continuum but which incorporate all electrostatic effects. For very narrow pores (up to twice the water radius) the barrier originates in the high electrostatic energy to place a bare charged particle at the centre of a low dielectric membrane (the Born barrier); because the pore is narrow, there is no space for water molecules to solvate the ion and hence the barrier becomes comparable in magnitude to the cost to completely dehydrate the ion. For wider pores the barrier is much higher than predicted on electrostatic grounds alone even though there would be room for almost one complete hydration shell around the ion. The reason is that water itself is not stable in the pore and has a tendency to evaporate from the pore due to the hydrophobic nature of the wall. This explains how hydrophobic pores as wide as 1 nm in diameter (and only 0.8 nm in length) can effectively be impermeable to ions. Hence, a hydrophobic pore need not be sterically occluded to act as an barrier to ion permeation.

The computational approach presented here provides strong evidence for a general hydrophobic gating mechanism. Major factors influencing the barrier to ion permeation are the chemical character of the pore wall, its radius, and also its flexibility or rigidity.

7.3 A hydrophobic gate in the nicotinic receptor

The model pores were designed as radically simplified versions of the nicotinic receptor gate. They were also more hydrophobic than any real protein. Even though hydrophobic sidechains present a surface similar to the methyl terminated surface of the model pores there is always at least the peptide bond dipole present so that the specific protein environment might mask or perhaps even nullify the hydrophobic gating effect. Hence, in Chapter 6 simulations were carried out to find the position of the gate in the closed-state nAChR pore, which contains a pronounced hydrophobic girdle with a pore radius of $R = 0.31$ nm but is generally much more complex than the model pores. Free energy calculations of the potential of mean force revealed a barrier to Na^+ permeation, peaked at the hydrophobic girdle formed at αL251 = L9′ and αV255 = V13′ with a height of 8.7 kT and extending over a length of more than 2 nm—almost the width of the membrane. Estimates of the single channel conductance are below the detection limit of experimental single channel recordings, which suggests that the hydrophobically sealed pore could function as a closed gate.

The calculations directly relate the structure of the nAChR channel to a functional characteristic. The observed gate can be rationalised on the basis of the calculations of the preceding chapters as a desolvation barrier to ion per-

meation, created by a hydrophobic constriction, i.e. a cylindrical pore with a high surface tension wall. Because the underlying principle should be generally applicable in a wider context and not just in the particular case of nAChR it is expected to play a role in other ion channels as well.

7.4 Confinement effects in other systems

The principles laid out in the present work have already been applied by other authors in a variety of biological systems. A brief outline of these authors' findings is presented here with reference to the results of the preceding chapters.

A gate in MscS The nicotinic receptor is not the only ion channel whose structure features a wide hydrophobic pore. The bacterial mechanosensitive channel of small conductance MscS displays a wide $R = 0.32$ nm hydrophobic constriction; on this ground it was originally believed to be captured in the open state.[55] Molecular dynamics simulations by Anishkin and Sukharev[56], however, showed that water evaporated from the MscS gate in the same way as seen in the hydrophobic model pores in Chapter 3. The authors termed this a "vapour-lock" mechanism, explicitly referring to results presented here.[62,96] Their work, together with the results for the muscle-type nAChR (Chapter 6) and the related $\alpha 7$ receptor (Appendix A), allows the tentative conclusion that hydrophobic gating is not dependent on a particular local protein environment but only relies on broad features, namely a non-polar constriction not much wider than 0.6 nm in diameter and about two α-helical turns in length.

Water transport through KcsA Liquid-vapour oscillations[96] were also invoked by Saparov and Pohl[308] to explain how the potassium channel KcsA can conduct water 100 times faster than the specialised water channel Aqp1. Water flux measurements under conditions of low K^+ concentrations ($< 0.2\ \mathrm{mol \cdot l^{-1}}$) yielded a very large water flux of $1.6 \times 10^2\ \mathrm{ns^{-1}}$. This flux could be gated by pH in the same way as the ionic flux is gated (KcsA opens below pH 4). The high turnover rate for water is—in the absence of K^+ in solution—independent of the osmotic driving force, i.e. an "all-or-none mechanism." Furthermore, addition of K^+ at concentrations greater than $0.2\ \mathrm{mol \cdot l^{-1}}$ limits water transport to the turnover rate of the ions ($< 0.1\ \mathrm{ns^{-1}}$); for each ion, two water molecules are dragged along. The turnover rate changes from fast water transport to slow (ion limited) transport over a very small K^+-concentration range (at ca. $0.2 \pm 0.01\ \mathrm{mol \cdot l^{-1}}$) in a step-like fashion. Even though the interior of KcsA is not especially hydrophobic, Saparov and Pohl[308] suggest that the observed

behaviour can be explained by capillary evaporation in the absence of K^+ in the channel and subsequent fast water transport in the vapour phase.

The amphipathic carbamate tunnel Kim and Raushel[309] discuss the significance of the surface character of the tunnel in the enzyme carbamoyl phosphate synthetase (CPS). A largely hydrophobic tunnel of length 10 nm connects three active sites, which catalyse the hydrolysis of glutamine to ammonia, and phosphorylation of bicarbonate and carbamate respectively. The reactive carbamate ion NH_2COO^- diffuses through the second half of the tunnel over ca. 5 nm. Five conserved glutamates with four bound water molecules provide hydrophilic patches in the tunnel. At least some of the glutamates are suspected to be in their protonated (uncharged) state as the hydrophobic environment would shift the local Glu pK_a to higher values, favouring protonation. Physiologically this is required because the intermediate carbamate needs to be protected from protonation (it is unstable at low pH) so the glutamate residues probably act as proton scavengers. Two arginines at either end control access to the tunnel; mutation to alanine also disrupts substantially the ability of CPS to catalyse the synthesis of carbamoyl phosphate. The authors interpret their findings in the light of the MD simulations of water in hydrophobic pores,[62] which were made increasingly hydrophilic through the addition of dipoles. They conclude that the carbamate tunnel contains a finely tuned set of polar residues in an overall hydrophobic environment which "opens" an otherwise "closed" hydrophobic tunnel.[309] The conserved Arg residues fulfil the role of gate-keepers, which steer the carbamate ion into or out of the tunnel.

"Are sobriety and consciousness determined by water in protein cavities?"[22] Ethanol (and butanol) are believed to produce their sedative-hypnotic effects in mammals by directly affecting ligand gated ion channels.[310] Alcohols are thought to bind in cavities which are normally water filled. Trudell and Harris[22] suggest that alcohols can only compete with cavity-water for binding because the enclosed water is—due to confinement—poised for evaporation.[96] That is, neurological states ultimately depend on the stability of liquid phase water in ion channels. Furthermore, it was shown that moderate increases in pressure (to 12 atmospheres) act as an antagonist to alcohols, reversing the effect of ethanol and butanol on $GABA_A$ and glycine $\alpha 1$ receptors, two ion channels of the central nervous system.[311,312] This result seems to be consistent with a view that increased external pressure stabilises liquid water in a cavity (or pore), as expressed in the Kelvin equation or Eq. (4.5). Quantitatively, an increase in pressure by 12 bar does not seem to be sufficient to change the equilibrium between liquid and vapour (the magnitude of the effect is only

about one tenth of the surface tensions involved).* However, the simulations of the nicotinic receptor showed that the particular protein environment can lead to some deviations from the ideal hydrophobic model pore behaviour, so it cannot be ruled out that pressure effects play a greater role in the case of water cavities in ligand gated ion channels.

7.5 Function follows from form

The American architect Louis Sullivan (1856–1924)[313] expressed the principle that design should follow from the purpose of the object:

> Whether it be the sweeping eagle in his flight, or the open apple-blossom, the toiling work-horse, the blithe swan, the branching oak, the winding stream at its base, the drifting clouds, over all the coursing sun, *form ever follows function*, and this is the law. Where function does not change form does not change.[314]

He intuitively recognised the principle that evolution shapes living organisms, and that their form expresses adaptation to a particular niche—from whole organisms down to the molecules that they are made from. He saw beauty in this concept and advocated an approach to design which is embodied in the phrase *form follows function*.

One of the fundamental aims of structural biology is to explain protein function based on protein structure. The hydrophobic gating effect, which was discussed in this work, is one example for such an explanation. It helps us to understand the position and the function of gates in a range of ion channel structures, i.e. we infer function from the structure, or—applying the inverse of Sullivan's principle—the function follows from the form.

*MD simulations of water in hydrophobic model pores at $p = 12$ bar did show a small but probably not significant shift towards the liquid state, with the main effect being a more rapid filling transition (data not shown).

Appendices

A The $\alpha 7$ receptor

The $\alpha 7$ receptor is a homopentameric neuronal nicotinic acetylcholine receptor.[25] Amiri et al.[41] constructed a model of the full receptor from homology models of its separate domains, based on the *Gallus gallus* (chick) sequence.* The transmembrane (TM) domain was modelled on the TM domain of the *Torpedo marmorata* muscle-type nAChR[29] (4 Å resolution) whereas the 2.7 Å structure of *Lymnea stagnalis* (snail) acetylcholine binding protein[279] (AChBP) was used for the ligand binding (LB) domain. The two domains were fused by an alignment procedure that accounts for stereochemical clashes, continuity of the amino acid chain, and closeness of functionally important residues.[41] The full model (Fig. A.1) exhibits a wide, aqueous pore in the LB domain. The TM pore resembles the one of muscle nAChR though some details are different. Most notably, the constriction at the $6'$ position ($R = 0.2$ nm) is narrower in $\alpha 7$ than in *Torpedo* nAChR ($R = 0.3$ nm; compare the pore profiles in Fig. A.2 to Fig. 6.11 on page 142).

Further details can be found in Amiri et al.[41]; in the following only the PMF calculations for a Na^+ ion in the $\alpha 7$ receptor are described.

A.1 Methods

Born energy calculations were performed as described in Section 6.2.3 at a concentration of 1 mol $\cdot$ l^{-1} NaCl. The bounding box for the whole receptor measured 9.6 nm $\times$ 9.6 nm $\times$ 21.0 nm and is considerably larger than the one used for *Torpedo* nAChR TM. In order to obtain a reasonable resolution, a grid of dimensions 97 $\times$ 97 $\times$ 193 cells was employed and focused into a volume of 1 nm $\times$ 1 nm $\times$ 1 nm, centred on the test ion. Electrostatic PMFs were calculated both for the full model (LB and TM domain) and the M2 bundle on its own (always embedded in a membrane mimetic slab); the same box size and grid dimensions were used in both cases.

Umbrella sampling of a Na^+ ion and 60 ns equilibrium molecular dynamics (MD) on the M2 helices of the chick $\alpha 7$ receptor model were carried out with the GROMOS96 (G43A1) force field and the SPC water model at a nominal concentration of 1 mol $\cdot$ l^{-1} NaCl. All other parameters are identical to the PMF

*The intracellular basket is missing from the "full" model because no atomic resolution structural data are available.

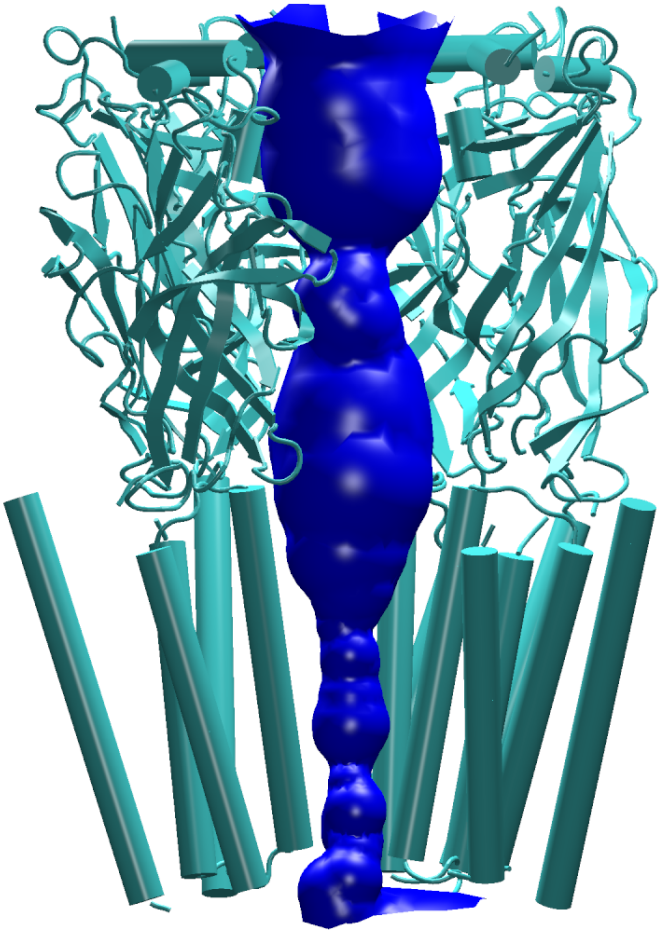


Figure A.1: $\alpha 7$ receptor with pore surface. Homology models of the ligand binding domain (top, mainly β-sheets) and the transmembrane domain (bottom, α-helical) were combined using stereochemical and functional data.[41] One subunit is omitted for clarity. The pore surface shows a wide, aqueous pore. (Pore surface created with hole[42] and rendered with vmd[43] and Raster3D[44])

calculations described in Section 6.2.4.

A.2 Results and Discussion

As shown in Fig. A.2, the MD PMF exhibits a 12.4 kT barrier at the narrowest constriction at the 6′ position (this corresponds to αS248 in *Torpedo*). In *Torpedo*, the GROMOS96 PMF, Fig. 6.14(b), has a height of 6.8 kT at 6′, and is slightly higher at the 13′ position in the hydrophobic girdle with 7.2 kT. At the same position in $\alpha 7$ there is a barrier of about 7.8 kT. It appears that in this $\alpha 7$ model the barrier-determining effect is strong low dielectric confinement at the rather narrow ($R = 0.2$ nm) 6′ constriction (similar to the $R < 0.3$ nm pores in Chapter 5).

The same pattern is also born out by the Born profile for the M2 helices of the TM domain. Although it predicts a much deeper valley between hydrophobic girdle 13′ and lower constriction site 6′ it agrees fairly well with the fully atomistic MD PMF. The electrostatic PMF for the *full* receptor exhibits the peaks in

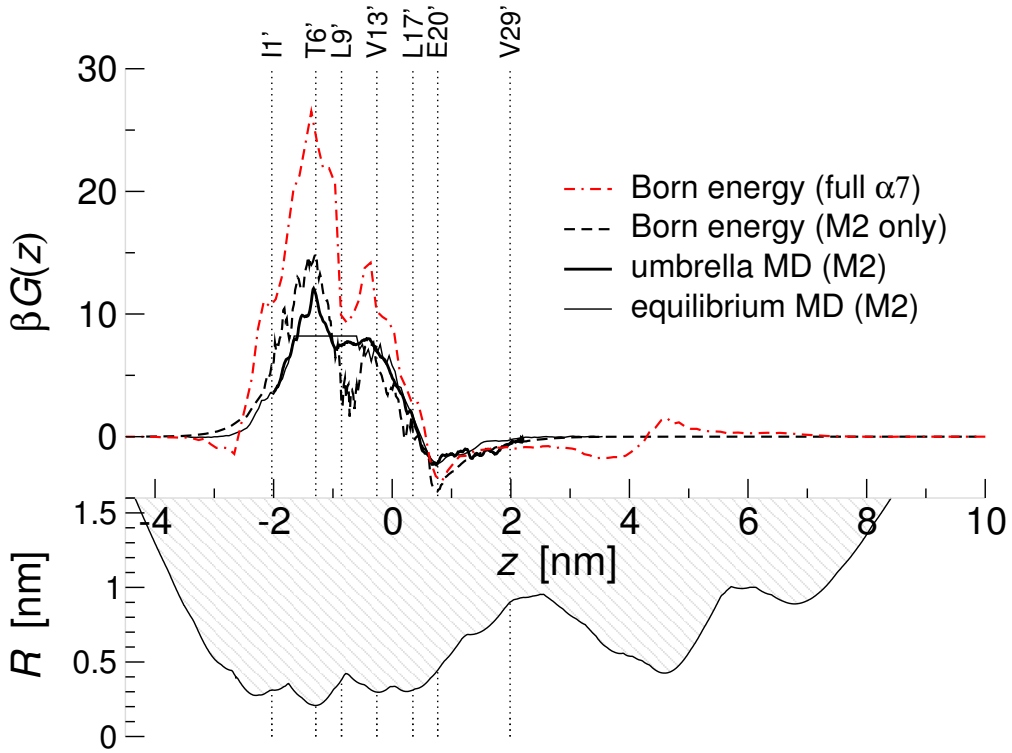


Figure A.2: PMF $G(z)$ in the $\alpha 7$ receptor. Equilibrium MD and umbrella sampling of a Na^+ ion was carried out in the backbone-restrained M2 bundle. The electrostatic PMF (Born profile) was calculated for the full $\alpha 7$ receptor and the M2 bundle, using the same parameters in the Poisson-Boltzmann calculations.

the PMF at the same positions as the M2-only profile. The peaks are, however, almost twice as high as in the profile with the LB domain absent.

The difference between the Born profile of the membrane embedded M2 bundle and of the full receptor is unexpected as it implies that the presence of a large protein mass away from the TM pore itself can actually influence the permeation of ions. The main effect of the LB domain is probably simply to displace a large volume of high dielectric solvent which would otherwise contribute to stabilising the ion in the pore. This effect is qualitatively different from the observation made in Section 6.3.2 where the influence of the full TM domain compared to only the M2 was shown to be negligible. In the latter case, parts of the low dielectric protein (M1, M3, M4, i.e. the scaffold) are replaced by the low dielectric membrane mimetic. Apart from minor changes in the shape of the dielectric boundary the only difference is the absence of partial charges from the region formerly occupied by the scaffold, which do not appear to be a major stabilising factor.

The maximum conductance $g_{\max} = 0.02$ pS (Eq. (6.1)) is smaller than for

Torpedo nAChR (0.1 pS, see Section 6.4), which is expected due to the larger barrier. It also indicates that the $\alpha 7$ receptor models a closed state.

It should be noted that the PMF and hence $g_{\max}$ depends rather sensitively on the geometry. Narrow constrictions will show up as large barriers in the PMF. These narrow constrictions might be a result of the homology modelling procedure, which are then frozen-in by the backbone restraints in the MD. It is possible that after extended free MD the pore opens up which in turn could lead to a geometry more similar to the *Torpedo* nAChR.

A.3 Conclusions

Preliminary calculations of the molecular dynamics and electrostatic PMF in a model of the $\alpha 7$ nicotinic acetylcholine receptor have been performed. When the electrostatic calculations are performed on the same reduced receptor model (M2 bundle) then both PMFs agree reasonably well, in contrast to the case of the model pores in Section 5.3.1. Either the agreement is fortuitous or the Poisson-Boltzmann calculation perform better in a system where there is more electrostatic influence than just the dielectric barrier. The Born profile for the transmembrane region *and* the extracellular ligand binding domain, however, is dominated by barriers almost twice as high as in the TM-only profile. More extensive (and much more costly in required computing resources) MD umbrella sampling calculations would need to be carried out to verify this effect. If shown to be true then it would have interesting implications on ion channel permeation. In particular, voluminous "protein-appendices" like the intracellular basket of the ligand gated ion channels could also function as attenuators of conductance, simply by their far reaching electrostatic influence.

Acknowledgements The $\alpha 7$ model was kindly provided by Shiva Amiri. The Born profile on the full receptor was calculated by Kaihsu Tai.

B A consistent definition of volume at the molecular scale

B.1 Introduction

The correct description of atomic scale phenomena is based on statistical mechanics and quantum mechanics. Nevertheless, classical descriptions of the interatomic forces, which parameterize the interatomic potentials as two, three, or four-body interactions, have been fairly successful.

It even appears that macroscopical concepts like surface tension can be applied down to the molecular scale.[113,254] This concept is employed in many empirical docking programs where the "hydrophobicity" of a sidechain is effectively parameterized as a surface tension times accessible surface area. Once one applies macroscopical concepts at the molecular scale one realises that some well-defined macroscopic variables like volume or surface area are ill-defined because the boundary of an atom or molecule is not sharp.[114] Atoms are not hard spheres but have a soft core (the Pauli exclusion principle, mimicked by the r^{12} term of the Lennard-Jones potential, Fig. B.1) and thus are—depending on their kinetic energy—able to approach arbitrarily close. But in the light of the success of macroscopic models it would be convenient to have some analogon of area and volume for microscopic system. Of course, "arbitrarily close" requires arbitrarily high energies, and thus some authors[202] define the distance of closest approach as this r where the repulsive core rises to 1 kT. Alternatively, atoms are treated as hard spheres with their van der Waals radii in order to calculate solvent accessible or solvent excluded surfaces.[117–119,121]

From a thermodynamic point of view the search for a definition of a surface is a non-question: a Gibbs dividing surface can be put anywhere between two phases and all thermodynamic properties have to be independent of the positioning of this surface.[100,113] In practical applications, however, surfaces are used, and defined often in an *ad hoc* manner, for instance, when determining the dielectric boundary in continuum electrostatic calculations.[179,183,186]

We propose a consistent definition of atomic volumes (and surfaces) based on a statistical mechanics approach which was initially inspired by Rowlinson[114], who noted that the integral

$$V_1 = \int d^3x \, e^{-\beta\psi(\mathbf{x})} \tag{B.1}$$

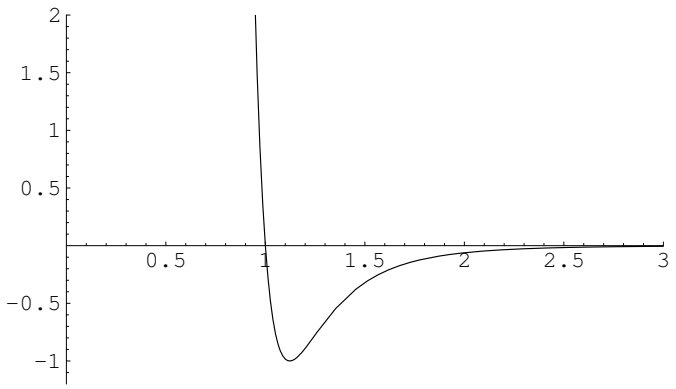


Figure B.1: The Lennard-Jones potential, the protoypical two-body interaction between two uncharged particles.

plays the role of an effective volume in the thermodynamics of small systems. In this approach we are concerned with particles (like water molecules or ions) that move around and through a host, which could be an ion channel protein or a zeolite. In the simplest picture the host provides a one-body potential $\psi(\mathbf{x})$ at each point $\mathbf{x}$ in space. ψ is essentially positive infinite in regions that are inaccessible to the particles, and in others it will be negative, denoting regions of attraction. The boundary surfaces on which $\psi(\mathbf{x})$ vanishes are an approximate measure of the shape of the host but there is still no well defined volume. The main problem with Eq. (B.1) is that it is not normalised and and does not yield absolute values. In the following we develop the idea of using expressions similar to V_1 as the volume of a pore or in general, the region accessible to solvent.*

B.2 Theory

The classically allowed region for a test particle with total energy E in an external potential $\psi(\mathbf{x})$ contains all points in space $\mathbf{x}$ where

$$E \geq \psi(\mathbf{x}), \text{ or } \quad E - \psi(\mathbf{x}) \geq 0 \tag{B.2}$$

(see Fig. B.2). Adding up all volume elements d^3x in the allowed region defines the classically allowed volume $\mathcal{V}$ for a particle with total energy E,

$$\mathcal{V}(E) = \int d^3x \, \Theta\big(E - \psi(\mathbf{x})\big), \tag{B.3}$$

*Similar concepts are also needed when calculating the partition function directly from molecular simulations.[315]

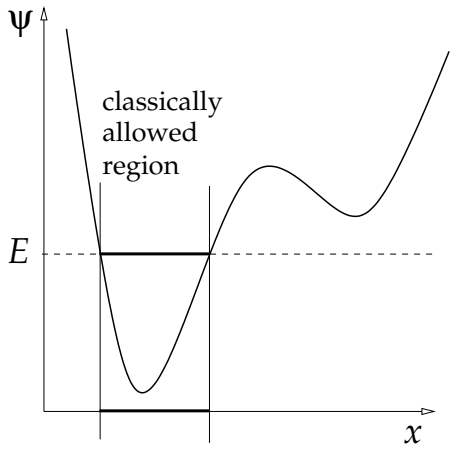


Figure B.2: Classically allowed region for a particle in a potential.

where Θ denotes the Heaviside step function. Note that $\mathcal{V}$ also depends on the test particle type because the interaction potential ψ is specific for the particle-host interaction. If the solvent particles were represented in a micro canonical ensemble with constant total E then $\mathcal{V}(E)$ would be a measure of the volume. Most of the time the particles are better represented in a canonical ensemble and only the average $\langle E \rangle$ is fixed through the temperature T. In this case one should compute the canonical expectation value

$$\mathcal{V}(\beta) = \int_{E_0}^{+\infty} dE\, \omega(E)\, e^{-\beta(E-E_0)} \int d^3x\, \Theta\big(E - \psi(\mathbf{x})\big), \tag{B.4}$$

where $\omega(E)$ is the density of states (DOS) and $E_0 = \min_{\mathbf{x}} \psi(\mathbf{x})$ the energy of the (classical) ground state; the Boltzmann factor gives the probability of finding a state with total energy E. Formally, the DOS is calculated as

$$\omega(E) = \int \frac{d^3p\, d^3x}{h^3}\, \delta(E - \mathcal{H}) \tag{B.5}$$

from the Hamiltonian $\mathcal{H} = \mathbf{p}^2/2m + \psi(\mathbf{x})$ of one particle in the external field $\psi(\mathbf{x})$. Alternatively, Eq. (B.4) can be written as

$$\mathcal{V}(\beta) = \int d^3x \int_{\psi(\mathbf{x})}^{+\infty} dE\, \omega(E)\, e^{-\beta(E-E_0)} \tag{B.6}$$

by changing the order of the integrations and evaluating the step function (note that $E_0 = \min_{\mathbf{x}} \psi(\mathbf{x})$). This can be interpreted as summing up the volume contributions at each point in space.

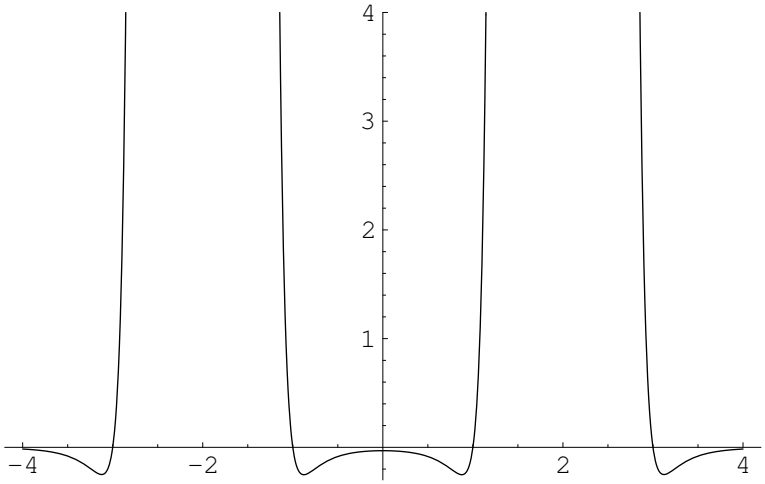


Figure B.3: Potential of a solvent particle between two Lennard-Jones spheres in one dimension.

B.3 Results and Discussion

As an example we calculate the 1-dimensional "volume" between two LJ-spheres (Fig. B.3) which is accessible to a water molecule. The interaction potential is given by

$$\psi(x) = V(x-d) + V(x+d) \quad \text{with } V(x) := \frac{c_{12}}{x^{12}} - \frac{c_6}{x^6} \tag{B.7}$$

(and $c_{12} = c_6 = 1$, $d = 2$ for the plots). The DOS for this particle will be approximated by the DOS for a free particle ($\mathcal{H} = \mathbf{p}^2/2m$, i.e. $\psi \equiv 0$). Otherwise one would have to solve the integral equations containing $\omega(E)$ in a self-consistent manner. The free particle DOS in a cylinder with a hard wall of radius R and height H is

$$\omega(E) = h^{-3} R^2 \pi H \int d^3p \; \delta\left(E - \frac{p^2}{2m}\right).$$

Because $\mathbf{p}$ is isotropic we can write $d^3p = 4\pi p^2 dp = \omega_p(E)dE$, and using $p^2 = 2mE$ and thus $\partial_E p \, dE = \sqrt{\frac{m}{2E}}\, dE$ we obtain

$$\omega_p(E) = 4\pi\sqrt{2m^3E}.$$

Hence the DOS

$$\omega(E) = R^2\pi H h^{-3} 4\pi\sqrt{2m^3E} = \frac{1}{4\pi\hbar^3} R^2 H (2m)^{3/2}\sqrt{E} =: \alpha\sqrt{E} \tag{B.8}$$

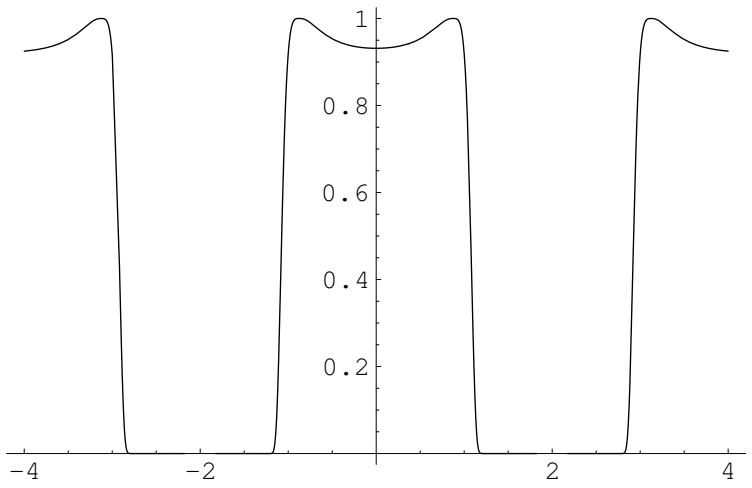


Figure B.4: $\chi(\mathbf{x}, \beta)$, the integrand of effective volume integral Eq. (B.6), plotted for a particle between two Lennard-Jones spheres (1D).

increases with the square root of the energy. When used as an approximation to the real DOS then one has to use

$$\omega(E - E_0) = \alpha\sqrt{E - \min_{\mathbf{x}} \psi(\mathbf{x})}. \tag{B.9}$$

First we will deal with the function $\chi(\mathbf{x}, \beta)$, which forms the integrand of the space integral in Eq. (B.6),

$$\begin{aligned}\chi(\mathbf{x}, \beta) &:= \int_{\psi(\mathbf{x})}^{+\infty} dE\, \omega(E - E_0)\, e^{-\beta(E-E_0)} = \alpha \int_{\psi(\mathbf{x})-E_0}^{+\infty} dE\, \sqrt{E}\, e^{-\beta E} \\ &= \frac{\alpha}{2\beta^{3/2}} \left[2\sqrt{\beta(\psi(\mathbf{x}) - E_0)}\, e^{-\beta(\psi(\mathbf{x})-E_0)} + \sqrt{\pi}\, \mathrm{erfc}\, \sqrt{\beta(\psi(\mathbf{x}) - E_0)} \right]\end{aligned} \tag{B.10}$$

As can be seen from Fig. B.4, the function $\chi(\mathbf{x}, \beta)$ acts as a softened shape function. Furthermore, it includes the effects of the attractive wells which show up as maxima. χ is not a true a volume but is (in principle) related to the partition function of one particle (or a non interacting gas of identical particles) in the potential $\psi(\mathbf{x})$, and hence is proportional to V_1 in Eq. (B.1). When the space integral is performed (here in one dimension),

$$\mathcal{V}(a, \beta) = \int_{-a}^{+a} dx\, \chi(x, \beta), \tag{B.11}$$

we obtain the volume as a function of the integration boundary a. This boundary should not necessarily extend to infinity because then the integrand could

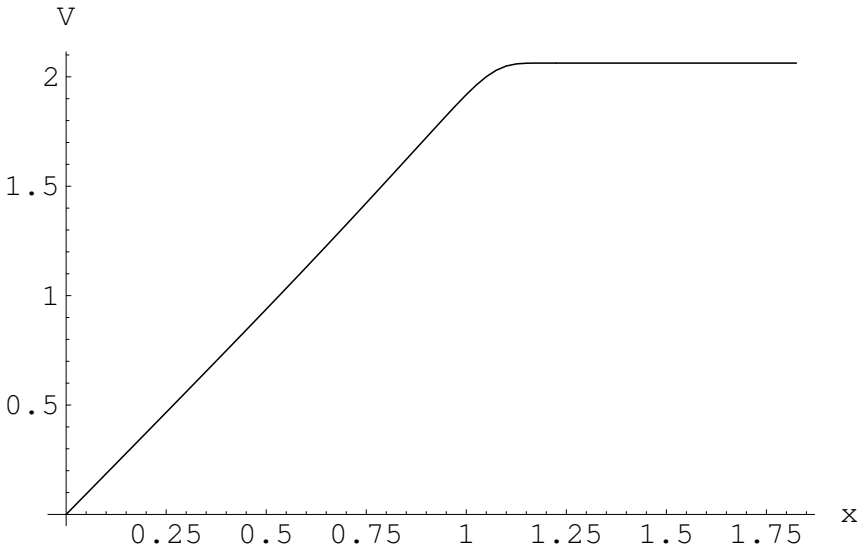


Figure B.5: Accessible volume in the Lennard-Jones 1D example. The integral is evaluated for increasing cylinder radii a. When a enters the hard core region of the LJ centre ($2^{1/6}\sigma \approx 1.12246$) the integral rapidly converges towards the true value at $d = 2$, the distance of a LJ centre from the origin.

pick up contributions from cavities which are not connected to the volume of interest. In general it would be a formidable task to find a topologically consistent description of the boundary.* In the case of simple, essentially one dimensional pores through a membrane it should be sufficient to use a cylinder as the integrating volume which encapsulates the pore. In our 1D example the cylinder reduces to the range $-a < x < a$. It is expected that Eq. (B.11) does not vary strongly for a sensible choice of the boundary.† As shown in Fig. B.5, $V(a, \beta)$ converges rapidly for $a > 2^{1/6}\sigma$ towards 2.06274 (for $\beta = 1; \sigma = (c_{12}/c_6)^{1/6} = 1$ and $\epsilon = c_6^2/4c_{12} = 1/4$). Because we are interested in the "volume" between the two spheres the maximum for the integration is $a = d$. Technically, the desired volume is thus

$$\lim_{a \to d} \mathcal{V}(a, \beta) = \mathcal{V}(\beta). \tag{B.12}$$

but due to the rapid decrease of χ in the LJ hard core region, the integration does not have to proceed very far into the hard core (i.e. far beyond $2^{1/6}\sigma$). For a cylindrical geometry, d would be set to the radius R of the enveloping

*By "topologically consistent" we mean to imply that the integral only extends over regions that are connected to the central volume by paths of sufficiently low barriers. More cautious qualifiers could be easily added—which mainly proves the difficulty of the problem.

†Or sidestepping the definition of "sensible": Perhaps a useful definition of the boundary would be "the surface a for which the variation $\delta\mathcal{V}([a], \beta)$ with δa vanishes."

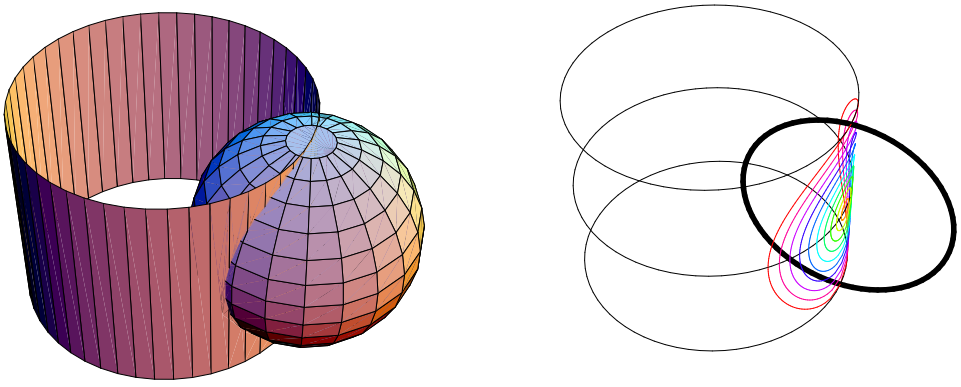
cylinder.

B.4 Conclusions

For simple systems the accessible volume can be calculated analytically but for complicated geometries and potentials this has to be carried out numerically. At this stage it is not clear if the effective volume differs significantly from conventional volume approximations. This method essentially calculates the partition function for an ideal gas confined to the host system. From then on it would be straightforward to derive other thermodynamic quantities as initial approximations to the more complicated real fluid. The method would also have to be extended to calculate surface areas. Finally, electrostatic effects would have to be considered. Clearly, the method presented here requires more work in order to make it potentially useful. Its current marginal merit lies in demonstrating the principle to define a volume based on statistical mechanics ideas.

C Analytical pore volume

The pore models used in Chapters 3 to 5 are geometrically simple objects so it should be possible to calculate their volume exactly, provided that they are assumed to be built from hard spheres with a defined radius. The "rugged volume" of the pore is understood as the space inside the cylinder which is not occupied by a wall atom. Apart from re-entrant surfaces this is the solvent excluded volume of the pore. A volume calculated in this way will be too large because it will also contain cusp regions that are not accessible to e.g. water molecules, which have a finite diameter. More elaborate schemes to calculate the solvent excluded surface (and thus the solvent excluded volume) take care of these re-entrant surfaces[118,121] but the following calculations will neglect them. Incorporating re-entrant surfaces tends to smooth the surface so that it will more resemble a smooth cylinder. Hence the shape of the pore is again approximately



(a) Solid rendering (b) Interfacial area

Figure C.1: Cylindrical pore intersecting a wall atom. (a) The pore volume can be calculated analytically as the volume of the enclosing cylinder minus the volume of the atomic sphere protruding into the cylinder. (b) The bounding surface of the intersected sphere (coloured contour lines) is related to *Viviani's* curve.[316]

cylindrical but with a reduced radius $R' \approx R - a$, where R is the radius of the cylinder on which the atomic spheres are located. and a is the van der Waals radius of such an atom.* This is the justification for using the scheme set out in Section 2.1.4 on page 29, which was employed in preference to the calculations set out in the following, not least because of its simplicity. Because the surface is calculated from hard spheres it suffers from the problems that motivated the approach in Appendix B.

C.1 Outline of the problem

The pore is thought to be made up from atomic spheres (radius a) with their centres lying on a cylinder surface (radius $R > a$). Fig. C.1(a) shows one sphere of a layer, together with the cylinder surface. The problem is to compute the volume of the cylinder minus the part of the spheres protruding into the cylinder. If the spheres do not overlap the problem reduces to calculating the volume $V(R, a)$ of the part of a *single* sphere within the cylinder. Then the volume in the pore for $\mathcal{N}$ spheres will simply be

$$V_{\mathrm{P}} = \pi R^2 L - \mathcal{N} V(R, a) \tag{C.1}$$

where the cylinder has length L.

The bounding surface of the sphere, which is created by the intersection with the cylinder, is not a simple surface (Fig. C.1(b)). In parametric form it is obtained from simultaneously solving the equations of the cylinder and of the sphere,

$$x^2 + y^2 - R^2 = 0 \qquad \text{(cylinder)} \tag{C.2}$$

$$(x - R)^2 + y^2 + z^2 - a^2 = 0 \qquad \text{(sphere at } (R, 0, 0)\text{)}. \tag{C.3}$$

This yields the surface as a function of z,

$$x(z) = \frac{1}{2R}\left(z^2 + (2R^2 - a^2)\right) \tag{C.4}$$

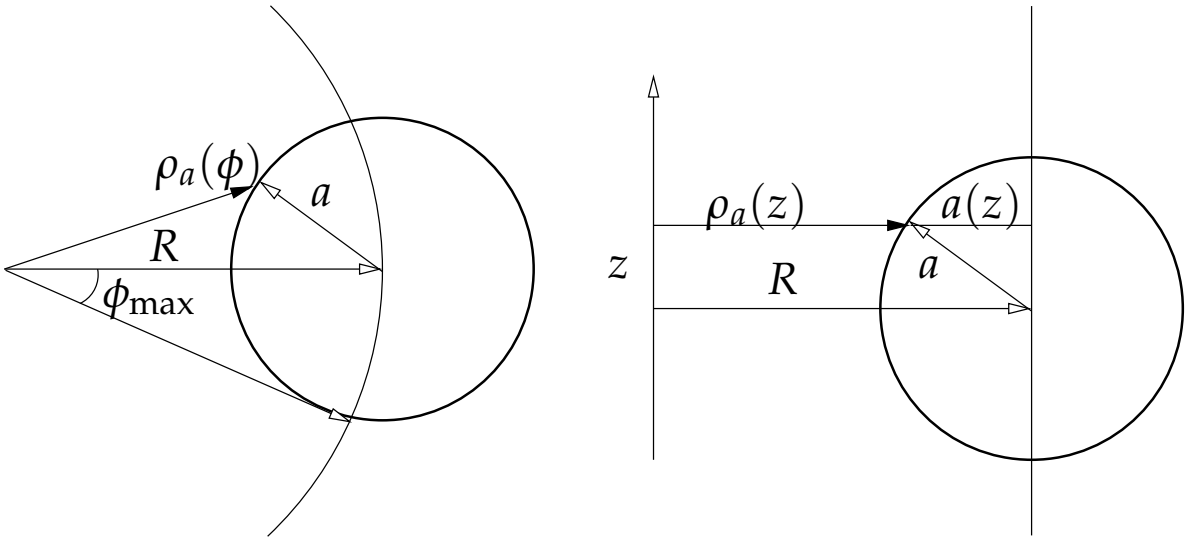
$$y(z) = \pm\frac{1}{2R}\sqrt{4R^4 - \left(z^2 + (2R^2 - a^2)\right)^2}, \tag{C.5}$$

where z runs from $-a$ to a (plotted in Fig. C.1(b) as coloured contour lines for $R = 1.3$, $a = 1$).

C.2 Calculation and Results

In order to calculate $V(R, a)$ we setup the problem in cylindrical coordinates (ρ, ϕ, z) with the z-axis parallel to the pore axis (Fig. C.2). The volume is calcu-

*In all other chapters the "pore radius R" is taken to be what is called R' in this appendix.



(a) View perpendicular to pore z-axis (b) View parallel to pore z-axis

Figure C.2: Geometry of an atomic sphere intersected by a cylindrical pore. The centres of the atomic spheres (one drawn as a circle) lie on the cylinder surface at radius R. The atomic hard sphere radius is a. In cylinder coordinates, the surface of a sphere is described by the radial vector $\rho(\phi, z)$.

lated as the integral

$$V(R,a) = \int_{-a}^{a} dz \int_{-\phi_{\max}(z)}^{\phi_{\max}(z)} d\phi \int_{\rho_a(\phi,z)}^{R} d\rho\, \rho \tag{C.6}$$

over the region bounded by the inner sphere surface $\rho_a(\phi, z)$ and the cylinder surface $\rho = R$. Elementary geometry yields

$$a(z) = \sqrt{a^2 - z^2} \tag{C.7}$$

$$\rho_a(\phi,z) = R\left(\cos\phi - \sqrt{\left(\frac{a(z)}{R}\right)^2 - \sin^2\phi}\right) \tag{C.8}$$

and for the maximum angle

$$\phi_{\max}(z) = \arcsin\frac{a(z)}{R}. \tag{C.9}$$

The evaluation of the Eq. (C.6) was performed in Mathematica[317] by successive integration. The radial integral becomes (writing $(a(z)/R)^2 = \sin^2\phi_{\max}$)

$$\begin{aligned} I_\rho(\phi, z) &= \int_{\rho_a(\phi,z)}^{R} d\rho\, \rho \\ &= \frac{R}{2}\left[1 - \cos 2\phi - \sin^2\phi_{\max}(z) + 2\cos\phi\,\sqrt{\sin^2\phi_{\max}(z) - \sin^2\phi}\right]. \end{aligned} \tag{C.10}$$

The angular integral (sweeping out the area at height z) is

$$\begin{aligned} I_\phi(z) &= \int_{-\phi_{\max}(z)}^{\phi_{\max}(z)} d\phi\, I_\rho(\phi, z) = 2\int_{0}^{\phi_{\max}(z)} d\phi\, I_\rho(\phi, z) \\ &= R^2\Bigg[\Bigg(\arctan\left(\frac{\sin\phi}{\sqrt{\sin^2\phi_{\max}(z) - \sin^2\phi}}\right)\,\sin^2\phi_{\max}(z) \\ &\quad + \phi\,(1 - \sin^2\phi_{\max}(z)) \\ &\quad + \left(\sqrt{\sin^2\phi_{\max}(z) - \sin^2\phi} - \sqrt{1 - \sin^2\phi}\right)\,\sin\phi\Bigg)\Bigg]_0^{\phi_{\max}(z)} \\ &= R^2\Bigg(\phi_{\max}(z) + \sin^2\phi_{\max}(z)\,\left(\frac{\pi}{2} - \phi_{\max}(z)\right) \\ &\quad - \sin\phi_{\max}(z)\,\sqrt{1 - \sin^2\phi_{\max}(z)}\Bigg). \end{aligned} \tag{C.11}$$

When the substitutions

$$\xi := \frac{R}{a} \quad \text{and } \zeta := \frac{z}{a} \tag{C.12}$$

$$\sin\phi_{\max}(z) = \frac{\sqrt{1 - \zeta^2}}{\xi} \tag{C.13}$$

are made then Eq. (C.11) becomes

$$\begin{aligned} I_\phi(z) = \frac{a^2}{2}\Bigg(&\pi(1 - \zeta^2) + 2(\zeta^2 + \xi^2 - 1)\,\operatorname{arccsc}\left(\frac{\xi}{\sqrt{1 - \zeta^2}}\right) \\ &- 2\sqrt{(1 - \zeta^2)(\zeta^2 + \xi^2 - 1)}\Bigg). \end{aligned} \tag{C.14}$$

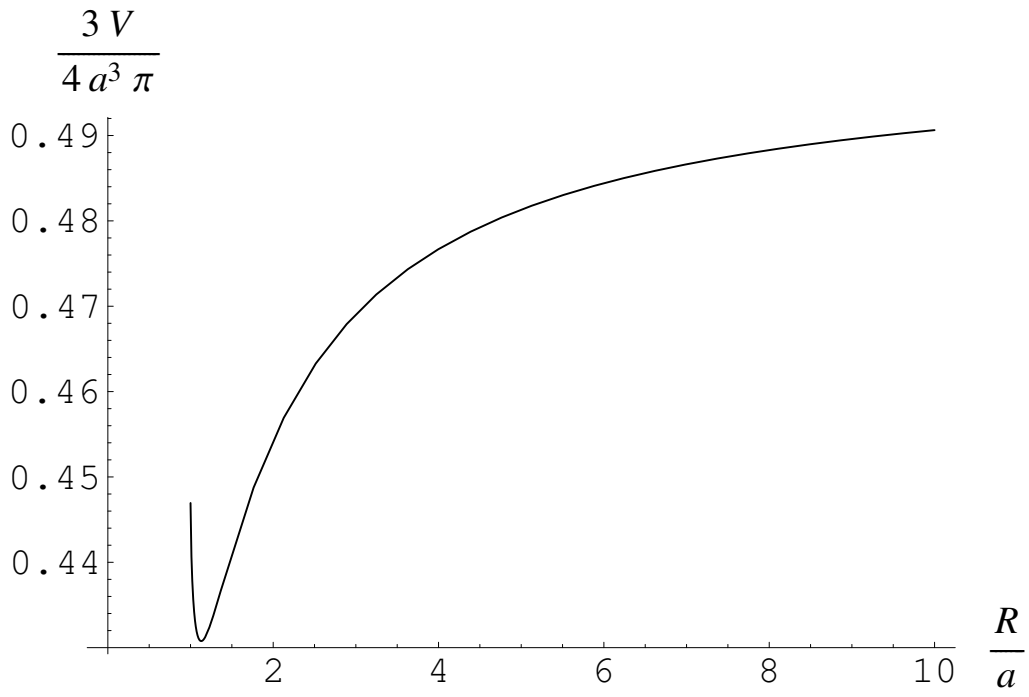


Figure C.3: Volume $V(R,a)$ of a sphere of radius a protruding into a cylinder of radius R, with the sphere centre lying on the cylinder surface. The volume of the part of the sphere *inside* the cylinder is shown for increasing cylinder radii. The volume is measured in the volume of the sphere $\frac{4}{3}\pi a^3$; it only depends on the cylinder radius, measured in the sphere radius, R/a. For large cylinders it approaches the expected result of half the sphere volume.

Integrating yields the volume

$$V(R,a) = \frac{4\pi a^3}{3} \times \frac{1}{6\pi}\left(3\pi + \sqrt{\xi^2-1}\,\left[(10\xi^2-11)\,E\left(\frac{1}{1-\xi^2}\right) - (10\xi^2-6)\,K\left(\frac{1}{1-\xi^2}\right)\right]\right) \tag{C.15}$$

K and E are complete elliptic integrals of the first and second kind* and $\xi = R/a$ [see (C.12)] is the cylinder radius, measured in units of the sphere radius. Eq. (C.15) is universal in that it only depends on the ratio R/a. As can be seen

*Elliptic integrals:[318] Complete elliptic integral of the first kind:

$$K(m) = \int_0^{\pi/2} \frac{1}{\sqrt{1-m\sin^2\theta}}\,d\theta, \quad -\infty < m < 1.$$

from the graph of $V(R,a) = V(R/a)$ in Fig. C.3, for large cylinder radii the cut-out volume approaches half the sphere radius because the cylinder surface increasingly resembles a plane.* Unexpectedly, there is a minimum at about $R/a \approx 1.1$. It is not relevant in our case, though, because $R/a \approx 1.1$ describes a pore which has barely enough space to accommodate two non-intersecting wall atoms. The R/a ratios of the model pores range from 1.769 to 6.128 for $a = 0.195$ nm and 0.345 nm $\leq R \leq$ 1.195 nm.†

C.3 Conclusions

The calculation shows that the exact volume of a cylinder with protruding, non-intersecting spheres can be easily calculated; for practical purposes this means reading off a correction factor to the atomic sphere volume from Fig. C.3. Conceivably, a polynomial could be fitted to Eq. (C.15) to facilitate the volume calculation. In any case, the error made by approximating the protruding volume of the atomic sphere as half the sphere volume is not greater than 10%. Also, inspection of 3D densities (the equivalent for the model pores to Fig. 6.10) indicated that water molecules trace out the solvent excluded MSMS surface,[121] which incorporates re-entrant surfaces. The MSMS surface in turn is not very rugged and thus well approximated by a cylinder. The MSMS/3D density surface and the calculations of this appendix justify the treatment of the model pores as simple cylinders, which can be characterised by a single radius for the purpose of volume calculations Section 2.1.4.

Complete elliptic integral of the second kind:

$$E(m) = \int_0^{\pi/2} \sqrt{1 - m\sin^2\theta}\, d\theta, \quad -\infty < m < 1.$$

*For $\xi \to \infty$ both $E((1-\xi^2)^{-1}))$ and $K((1-\xi^2)^{-1})$ become $\frac{\pi}{2}$; the limit $\lim_{\xi\to\infty} V(\xi) = \frac{2}{3}\pi a^3$, however, is non-trivial as the approach of K and E towards $\frac{\pi}{2}$ differs.

†Only in this chapter R denotes the radius of the cylinder surface. Previously R was the effective radius of the van der Waals surface bounded pore, $R - a$, with 0.15 nm $\leq R - a \leq$ 1.0 nm.

D Programs and scripts

This appendix describes some programs that were developed or modified in the course of this work. Where licensing permits, they are made available under the GNU General Public License[319] on the author's website `http://sbcb.bioch.ox.ac.uk/oliver/software/`.

D.1 Simulation setup

D.1.1 pgeom

Usage

`pgeom [OPTIONS] -R` R_{outer} `-P` R_{P} L_{P} `-M` R_{M} L_{M}

Synopsis

pgeom calculates positions of pseudo-atoms in a model for a transmembrane channel and writes the structure to the file `pore.pdb` in PDB format, and a gromacs topology to `pore.itp`. The pore is constructed from concentric, stacked rings of spherical pseudo atoms. Typically, no bonds are generated because the pore atoms are held at their equilibrium position by harmonic restraints (which are generated with the genpr program, which is part of gromacs).

The user determines the dimensions of the mouth region and the constriction region. The pore profile is linearly interpolated between the radius of the constriction R_{P} and the radius of the mouth R_{M}.

Options

`-h`		show help
`-v`		be verbose (= `-debuglevel 30`)
`-debug`	N	set debuglevel (0...100)
`-spec`	m_{spec}	atomic species for pore wall (see Tab. D.3)
`-showspec`		show hard coded atomic species (see Tab. D.3)
`-o`	`pore.pdb`	`pdb` coordinate file

-s	pore.itp	itp topology file
-R	R_{outer}	Outer radius of the model R_{outer}
-P	R L	Pore region: inner radius R and length L
-M	R_{M} L_{M}	Mouth region: largest inner radius R_{M} and length L_{M}
-b	$d_{\min}$ $d_{\max}$ $\gamma_{\min}$ $\gamma_{\max}$	Option is repeatable. Form bonds with angle γ when interatomic distance d satisfies $d_{\min} \leq d \leq d_{\max}$ (in Å) and angle $\gamma_{\min} \leq \gamma \leq \gamma_{\max}$, $0^\circ \leq \gamma \leq 90^\circ$
-c		only write connectivity to output, no bond length or k_{bond}, k_{angle}
-x		neither connectivity nor bond length to output (isolated atoms)
-kB	k_{bond}	force constant of bonds, in $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$
-kA	k_{angle}	force constant of angles, in $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{rad}^{-2}$
-cc		center coordinates on cavitybox, not on unitcell

As an experimental feature, the pore volume can be calculated as described in Appendix B; the Lennard-Jones parameters are taken from the GROMACS force field. Setting any of the switches for profile calculation also enables the volume calculations. All the following *lengths* are in *nano metre* (nm) not Ångstrom.

-volume		calculate pore volume
-profile	profile.dat	calculate the profile $R(z)$ in addition to the volume and write to file
-z1	z_1	profile between z_1 and z_2
-z2	z_2	
-Rmax	$R_{\max}$	integrate out to $R_{\max}$ (also use for the total volume integration if -profile is set)
-npoints	N	number of points per dimension in the integrals
-T	T	temperature in Kelvin [300]
-wca		if set, only use the repulsive part of the Lennard-Jones potential (split after Weeks, Chandler, and Andersen[123])
-plot		xfarbe output of the potential in z slices
-nzplot	N_{slice}	number of plot slices for -plot

m_{spec}	name	symbol	r/nm	q/e
0	methane	CH_4	0.195	0
1	carbon	C	0.140	+0.38
2	oxygen	O	0.150	−0.38

Table D.3: Hard coded atomic species in pgeom.

Example

The hydrophobic pores of radius R, that were investigated in Chapter 3 were generated with the command line

```
pgeom -R 18 -P R 8 -M 10 8 -x -o pore.pdb -s
pore.itp
```

where $R_{\text{outer}} = 1.8$ nm (which is not functionally important because the pore complex is embedded in a membrane mimetic slab), $R_{\text{M}} = 1.0$ nm, $L_{\text{M}} = L_{\text{P}} = 0.8$ nm (apart from the cases in Section 5.3.3 where L_{P} was varied), and the pore radius R. Pores with hydrophilic pore walls were created by adding charges to selected atoms in the topology file `pore.itp`.

D.1.2 prepumbrella.pl

Usage

```
prepumbrella.pl [OPTIONS] --topology=topol.top
--mdp=umbrella.mdp --index=refgroup.ndx
--confstart=startingconfigs.dat --type="SPECIES"
```

Synopsis

Create run input files (`tpr`) for umbrella sampling runs, using gromacs (version 3.2.1-LMB). Give the number of umbrella windows and either the length over which the windows are (equally) distributed or the minimum and maximum coordinate. At the moment this only caters for 1D PMFs, e.g. just along z in a pore or channel.

If only `--length` is given, the windows are created symmetrically to 0 (which is then the centre of mass of the reference index group (type '`com`' or '`com_t0`' in the `ppa`)). If `--min` and `--max` or `--length` and either one of `--min` or `--max` are given, the window centres are offset accordingly (still relative to the reference group's centre of mass). The window width Δz is calculated as length $L = \max z - \min z$ divided by the number of windows N, $\Delta z = L/N$.

Unless overridden from the commandline, the force constant is chosen so that an ion in window i can freely diffuse into an adjacent window $i+1$ or $i-1$. The

energetic barrier to this diffusion is given by `--barrier` (and is the energy of the umbrella potential at $z_i + \frac{3}{2}\Delta z$, i.e. the centre of the neighbouring window). Set the temperature to the one set for the simulations.

This script requires all the files needed to run grompp so that it can create the `tpr` files for all the windows and a special list of starting configurations, customarily named `conf.dat`. It creates a `ppa` file using a hard coded template. The grompp `mdp` file should have the proper run length (e.g. 1 ns) and typically neither `trr` nor `xtc` output are required.

Implicit assumptions

- Currently we assume that the direction of the sampled coordinate is $(0, 0, 1)$, i.e. the z-axis.
- the configuration list contains absolut positions
- the `center_of_mass` (com) must be supplied in the configuration list
- max, min, and length are relative to the com of the ref group

Options

Defaults in brackets, for most options abbreviations can also be used.

`--help`	this help
`--help-targets`	list of known target systems (for `--targetsystem`)
`--help-config`	format of the starting configuration file (for `--confstart`); see below
`--version`	version number of prepumbrella.pl (quote when submitting a bug report)
`--verbose`	chatty output
`--debug=INTEGER`	set debug level (> 9 switches off error checks, > 11 grompp is not run at all, only commandline shown)

Input files (all must exist):

`--topology=FILE`	gromacs topology [`topol.top`]
`--mdp=FILE`	grompp input with run parameters [`grompp.mdp`]
`--index=FILE`	index file containing the reference group [`pull.ndx`]

`--confstart=FILE`	list that describes available starting conformations (see below for format) [`conf.dat`]
`--np=INTEGER`	number of processors to write tpr files for [`1`]

Output:

`--Project=DIRECTORY`	place all window directories under `DIRECTORY` [`.`]
`--prefix=STRING`	output `tpr`s are named `STRINGn.nnnn.tpr` (where n.nnnn is the centre of the umbrella window in nm) if the standard `--format` string is used []
`--format=PRINTFSTRING`	pattern to create output filenames. May contain one '`%s`' (for prefix) and one '`%f`' for the umbrella centre (see printf(3)) [`%s%+.4f`]
`--qsuball=FILE`	script that contains qsub commands to submit all windows through a queuing system
`--targetsystem=SYSTEM`	create run script(s) for the given target systems; can be given repeatedly to produce more than one type of runscript. Use `--help-targets` for a list of valid `SYSTEM`s and the values that they define (with `--verbose`) [`local`]

Umbrella windows:

`--type=STRING`	name of the particle, eg `Na`, `Cl`, `NA`, ... (as used in the input files)
`--nwindows=INTEGER`	number of umbrella windows [`20`]
`--length=FLOAT`	length to cover with windows, symmetrically to the reference group's centre of mass. Alternatively, set `--min` and `--max`.
`--min=FLOAT`	relative to reference group centre of mass, in nm
`--max=FLOAT`	relative to reference group centre of mass, in nm
`--kUmbrella=FLOAT`	force constant of the harmonic umbrella potential in $\mathrm{kJ \cdot mol^{-1} \cdot nm^{-2}}$; overrides program choice
`--Temperature=FLOAT`	temperature of the simulation in Kelvin [`298.0`]
`--barrier=FLOAT`	barrier to move to the centre of an adjacent window in kT [`1.5`]

Format of the starting configuration file

The configuration file `conf.dat` contains a header section where parameters are set, and a list of possible starting configurations. prepumbrella.pl uses the list of starting configurations to find those which have a particle of interest close to the centre of an umbrella window. It then looks up the corresponding starting structure file (in `gro` format) from the list and uses it to create binary run input file (`tpr`) by running grompp.

```
; comment (same as in Gromacs mdp files)
; all entries must be on one line
; '@' lines set variables
@ var = value
filename   particle_type  atom_nr   pos_x pos_y pos_z
filename   particle_type  atom_nr   pos_x pos_y pos_z
...
```

`filename`	full path to the `gro` or `pdb` file
`particle_type`	name used in gromacs, eg `Na`, `Cl`, `SOL`
`atom_nr`	atom_id, number in `gro` or index file
`pos_x|y|z`	coordinates of the particle in nm

Required variables Variable assignments are introduced by an @-sign at the beginning of the line; spaces around '=' are mandatory. Required variables must be given in the configuration file.

`center_of_mass =` $x\ y\ z$	center of mass of the reference group

Optional variables If no values are given the program defaults are chosen (given in brackets) and the program issues a warning. A cylindrical confinement potential is only activated if both `RConfine` and `KConfine` are non-zero.

`reference_group = group`	reference group from index file [`protein`]
`reftype = com_t0|com`	center of mass from initial frame or always from the current frame [`com_t0`]
`RConfine =` R	use a confinement potential outside a radius R nm [`0.0`]
`KConfine =` k	and a harmonic force constant k kJ $\cdot$ mol^{-1} $\cdot$ nm^{-2} [`4500.0`]

prepconflist.pl (Appendix D.1.3) can be used for setting up the configuration list from `trr` frames.

Bugs and Limitations

- Runscripts are generated for different target system environments and the specifics are hard coded in the script itself. See the source itself and `--help-targets --verbose`. The target system complex should become configurable through an external XML file.
- Only single ions as pulled groups are tested.
- `conf.dat` should be XML.
- The umbrella `ppa` file should be a template and not hard coded.
- Dirty hack: if there is an index file named exactly '`index.ndx`' in the current directory then it is automatically picked up at the grompp stage; this allows for using non-standard groups in the `mdp` file.

Example

`conf.dat` contains all required variables and a list of starting configurations, with the starting configurations themselves accessible. The index file `refgroup.ndx` contains the index group `protein`, relative to whose centre of mass the umbrella windows are placed. The topology and related files are found two directories above the current one, with the main topology being `chl.top`. The molecular dynamics run input file `umbrella.mdp` is similar to one used for equilibrium simulations with the exception of a shorter run length (typically about 1 ns, no trajectory output, and when starting from structure files with velocities, `gen_vel = no` and `unconstrained-start = yes`). In the example, a Na^+ ion (with the general name `Na`) is to be umbrella sampled. The command line

```
source /opt/gromacs/3.2.1-LMB/i686-pc-linux-gnu/bin/GMXRC
prepumbrella.pl --topology=../../chl.top
--mdp=umbrella.mdp  --index=refgroup.ndx
--confstart=conf.dat  --Project=0R55/Na
--prefix=Na0R55z --type=Na --nwindows=101
--min=-0.9 --max=0.9  --Temperature=300
--targetsystem=workstations --qsuball=submitall.sh
```

will have prepumbrella.pl construct 101 separate input files for mdrun; each of these windows can be run independently from the others thus allowing

massively parallel execution on clusters or compute farms. Depending on the target system, run input files for a queuing system are also created, which allow submission of all 101 jobs through execution of a single script, `submitall.sh`. Note that the script requires a working gromacs environment to create the run input files.

After these jobs have finished, g_wham is run on all result files `*.pdo.gz` as described in Appendix D.2.5.

D.1.3 prepconflist.pl

Usage

```
prepconflist.pl [OPTIONS] file1.gro [file2.gro
...]
```

Synopsis

prepconflist.pl creates the file list which forms the major content of the input file (`conf.dat`) to prepumbrella.pl. It reads all files given on the command line and writes matches to the output file (with `--append` it appends to the output file).

Options

The options can be abbreviated; defaults are given in [].

`--help`	help
`--verbose`	(alias for `DEBUGLEVEL`= 3)
`--version`	print version id
`--debug=DEBUGLEVEL`	0 (almost) quiet, < 0 silent, $1 \ldots 3$ verbose, > 5 really debugging stuff (20 max) [`1`]
`--output=FILE`	output gro file [`conf.list`]
`--append`	append to output file (do not overwrite)
`--atomname=STRING`	name of the pulled atom in the gro file [`NA`]
`--z1=FLOAT`	only list starting positions with z coordinate
`--z2=FLOAT`	between z_1 and z_2 (box coordinates in nm).
`--radius=FLOAT`	Also, only list them when they are within a cylinder of radius R which is centered on the center of mass.
`--xcom=FLOAT`	ref. group centre of mass x-coordinate (nm)
`--ycom=FLOAT`	ref. group centre of mass y-coordinate (nm)

`--zcom=FLOAT`	ref. group centre of mass z-coordinate (nm), but this is not used at the moment as the cylinder is always taken to be parallel to the z-axis

Example

First we generate a **selection of equilibrium configurations** of the whole system by dumping snapshots from an equilibrium trajectory of the system.

```
mkdir start
echo -e '0\n'
| trjconv -tu ns -b 2 -sep -s ../../chl_md.tpr
-f ../../chl_lay_water_md.trr -o start/na.gro
```

Here a `trr` file is used, which allows to start with velocities. About 300 frames are typically required. The whole system is required (which is selected by choosing '0'). The index for the **reference group** is created through

```
make_ndx -f ../../chl_md.tpr -o refgroup.ndx
```

Typically, the whole protein is chosen. The **centre of mass** of the reference group can be calculated with

```
g_traj -com -f start/na_1.gro -s ../../chl_md.tpr
-n refgroup.ndx -ox com.xvg
```

and is found as the single entry in `com.xvg`.

Then prepconflist.pl goes through the configurations and generates a list relating the position of a particle type of interest (e.g. ion or water oxygen) to the file where it was found.

```
prepconflist.pl --xcom=2.2 --ycom=2.3
--radius=0.8  --z1=1.7 --z2=6.4 --atomname=Na
--output=conf.list --verbose  start/*.gro
```

To reduce the number of entries one can restrict the volume in which particles are searched for. The volume is a cylinder centered on the centre of mass of the reference group, extending from $z_1 = 1.7$ nm to $z_2 = 6.4$ nm with a radius $R = 0.8$ nm. The resulting `conf.list` can be appended to the header section of a `conf.dat` file (see Appendix D.1.2).

D.2 Analysis

D.2.1 g_count

Usage

```
g_count -f trajectory.xtc -s topol.tpr [OPTIONS]
```

Synopsis

g_count counts the numbers of molecules within a cylindrical region as a function of time. It takes an index file with atomnumbers and chooses the molecules that these atoms belong to (or optionally (with option `-nom`) it simply takes all atoms in the index) and generates an output file with the number of molecules/atoms which are present in the cavity a each time frame.

The output data file (option `-dat`) contains for each time step t the number of particles in the cavity (the pore occupancy $N(t)$), the concentration $c(t)$, and the mass density $\rho(t)$.

column:	1	2	3	4
quantity:	t	$N(t)$	$c(t)$	$\rho(t)$
units:	ps		$\mathrm{mol}\cdot\mathrm{l}^{-1}$	$\mathrm{g}\cdot\mathrm{cm}^{-3}$

Density and concentration are calculated from the volume of the cavity,

$$V = R_{\mathrm{eff}}^2 \pi (z_2 - z_1) \tag{D.1}$$

where R_{eff} is the apparent accessible radius (corrected with wall atom radius $r_A = \sigma_{AA}/2$ (approximately the van der Waals radius), water radius $r_{\mathrm{water}} \approx 0.14$ nm, and the distance between the centre of a wall atom A (typically, a united atom methane CH_4) and the water oxygen (O_W), $d_{\text{wall-water}} = \sigma_{AO_W}$. σ come from the Lennard-Jones potential of the wall atom and the water oxygen.

$$R_{\mathrm{eff}} = \hat{R} - d_{\text{wall-water}} + r_{\mathrm{water}}; \tag{D.2}$$

$\hat{R}$ is the radius of the pore, measured from the atomic center of a wall atom to the pore axis. The pore radius R (which is quoted throughout this work) is

$$R = \hat{R} - r_A, \tag{D.3}$$

and it is the radius of the cylinder formed by the solvent accessible van der Waals surface.* Thus, the radius correction could be calculated from the force

*Because of the atomic roughness of the pore wall it is not strictly correct to speak of a single pore radius—depending on the angle the distance from the pore axis to the surface varies. R is the shortest radial distance in the pore and hence represents a lower bound on the true average pore radius.. See also the discussion in Appendix C.3.

field (i.e. σ) as

$$\delta R = \frac{1}{2}\sigma_{AA} - \sigma_{AO_W} + r_{\text{water}} \tag{D.4}$$

This approach turns out to be somewhat imprecise and δR was then considered as an adjustable parameter which was determined as described in Section 2.1.4 on page 29 by considering a large pore and matching the volume resulting from the above calculation and an integration of the local density (produced by g_ri3Dc, described in Appendix D.2.3).

The `-track` index file contains the atom_ids of all atoms that were at least for `-tau` τ ps in the cavity (entry `[ *_tracked ]`). The entry `[ *_cavity ]` contains all atoms that were for at least one time frame in the cavity. Entries containing molecule numbers are titled `[ *_molecules ]`.

`-tdat` writes a data file containing all atoms or molecules in the cavity and their total residency time. Format:

column:	1	2	3	4	5
quantity:	atom_nr	molecule_nr	name	total residency time $\mathfrak{T}_{\text{res}}$	$\mathfrak{T}_{\text{res}}/\mathfrak{T}$
units:				ps	

Defaults If no values are given for the point on the axis (option `-cpoint`) then the center of the initial box is taken. If no radius `-R` R is given, the maximum radius fitting in the box is chosen. z_1 and z_2 default to the full box length in z-direction.

Files

`-f`	`traj.xtc`	Input	Generic trajectory: xtc trr trj gro g96 pdb
`-s`	`topol.tpr`	Input	Structure+mass(db): tpr tpb tpa gro g96 pdb xml
`-o`	`c_numbers.xvg`	Output	xvgr/xmgr file
`-dens`	`c_density.xvg`	Output, Opt.	xvgr/xmgr file
`-conc`	`c_conc.xvg`	Output, Opt.	xvgr/xmgr file
`-dat`	`cavity.dat`	Output	Generic data file
`-tdat`	`c_track.dat`	Output	Generic data file
`-n`	`index.ndx`	Input, Opt.	Index file
`-track`	`c_track.ndx`	Output, Opt.	Index file

Other options

`-h`	bool	`no`	Print help info and quit
`-nice`	int	`0`	Set the nicelevel

-b	time	-1	First frame (ps) to read from trajectory
-e	time	-1	Last frame (ps) to read from trajectory
-dt	time	-1	Only use frame when t MOD dt = first time (ps)
-w	bool	no	View output xvg, xpm, eps and pdb files
-m	bool	yes	index contains atoms, but g_count counts the molecules
-axis	vector	0 0 1	Vector pointing parallel to the pore axis
-cpoint	vector	0 0 0	Point on the pore axis
-R	real	-1	Radius of the cylindrical pore cavity
-z1	real	-1	Confine waters to have a z coordinate larger than z1, and
-z2	real	-1	smaller than z2
-tau	real	10	Total time (ps) that a particle has to be in the cavity at least so that it is tracked
-dR	real	0.03	Correct water-accessible cavity volume

Known problems and caveats

- `-m` behaves differently from the standard usage within the g_* programs—it figures out for itself what the molecules are and does not need *molecule* numbers but *atom_ids*.
- `-m` is the *default* behaviour!
- When counting *ions* you *must* use `-nom`!
- The density is calculated as $\langle M \rangle / [R_{\mathrm{eff}}^2 \pi (z_2 - z_1)]$ where $\langle M \rangle$ is the average mass in the pore cavity—so it makes only sense if this approximates the true cavity volume.
- The default volume/radius correction is specific for methane pseudo atoms (ffgmx, ffG43a1) and SPC water.
- The program is only tested with pore axis parallel to z-axis (0,0,1).

Example

Suggested use for water. Create an index file for the water molecules:

```
echo -e "keep 0\ndel 0\nr SOL\nq\n" _
| make_ndx -f in.pdb -o sol.ndx
```

Then analyse the water molecules in a trajectory traj.xtc with

```
g_count -m -f traj.xtc -s topol.tpr -n sol.ndx
```

D.2.2 g_flux

Usage

```
g_flux -f trajectory.xtc -s topol.tpr [OPTIONS]
```

Synopsis

g_flux calculates the flux of molecules (or atoms) through a cylindrical region as a function of time. It takes an index file with atomnumbers and chooses the molecules that these atoms belong to (or optionally (with `-nom`) it simply takes all atoms in the index) and generates an output file with the number of molecules/atoms at each time frame.

Definition of observables

Flux The flux Φ is the number of particles permeating the pore per unit time. It is calculated from the total number of permeating particles M during a simulation of length $\mathfrak{T}$ as

$$\Phi = \frac{M}{\mathfrak{T}}. \tag{D.5}$$

M counts particles in both directions ("up" or "+" and "down" or "−"), i.e. $M = M^+ - M^-$ where $M^+ \geq 0$ and $M^- \leq 0$. The fluxes are

$$\Phi^+ = \frac{M^+}{\mathfrak{T}} \geq 0 \tag{D.6}$$

$$\Phi^- = \frac{M^-}{\mathfrak{T}} \leq 0 \tag{D.7}$$

$$\Phi = |\Phi^+| + |\Phi^-|. \tag{D.8}$$

In equilibrium, the net flux $\Phi^+ + \Phi^-$ is zero because on average there is no net force favouring either direction. In practice, there tends to be a very small bias (of the order of 0.1% of Φ) towards negative fluxes which might be related to the rescaling of the particle positions due to the constant-pressure weak-coupling scheme.

The cumulative flux

$$\mathcal{F}(t) := \int_0^t dt'\, \Phi(t') = \sum_{t'=0}^{t} M(t')$$

is calculated from the instantaneous flux $\Phi(t) = M(t)/\Delta t$ over one time step Δt and counts the number of permeation events up to time t.

Permeation time The residency time $\tau_{\mathrm{p},i}(t)$ for a pore-permeating particle i that entered the pore time at t_i' and exited at t is

$$\tau_{\mathrm{p},i}(t) = t - t_i'. \tag{D.9}$$

The average permeation time $\tau_{\mathrm{p}}(t)$ for all $M(t)$ particles that permeated the pore between t_i' and $t > t_i'$ is calculated as

$$\tau_{\mathrm{p}}(t) = \frac{1}{M(t)} \sum_{i=1}^{M(t)} (t - t_i'),$$

i.e. at each time the length of the permeation events that completed at this step are averaged. The cumulative average permeation time is

$$\left\langle \tau_{\mathrm{p}}(t) \right\rangle = \frac{\sum_{t'=0}^{t} M(t') \tau_{\mathrm{p}}(t')}{\sum_{t'=0}^{t} M(t')}$$

and can be used to assess the convergence of the mean permeation time

$$\langle \tau_{\mathrm{p}} \rangle \equiv \left\langle \tau_{\mathrm{p}}(\mathfrak{T}) \right\rangle. \tag{D.10}$$

Algorithm

The algorithm to detect the permeating particles can be described as "check on exit." If 'O' denotes the space outside of the pore and 'I' the inside, then only trajectories O→I→O are counted as complete transitions through the pore; I→O→I are periodic boundary trajectories. Any enter event for a particle is recorded; on a later exit it is determined if it left the pore volume on the side opposite to the influx event (a permeation event) or on the same side (not a successful permeation).

In the following, the pore is assumed to be oriented parallel to the z-axis. The faces of the cylindrical pore are at $z_1 < z_2$.* For any given molecule i at position $\mathbf{x} = (x, y, z)$ compare the position at time step t and at the previous step $t - 1$.

1. A molecule **crosses a boundary** at z_m ($m = 1, 2$) if

   $$\min\big(z(t-1), z(t)\big) < z_m < \max\big(z(t-1), z(t)\big)$$

   is true.

*A generalisation to pore axes pointing along other vectors than $(0, 0, 1)$ is straightforward but was not necessary in the cases discussed here.

2. Given a boundary crossing at step t, determine the type of the crossing (enter or exit event). The type is recorded as the "divergence" D

$$D = \begin{cases} +1 & \text{if } z_1 < z(t) < z_2 \quad (\text{O}{\to}\text{I}, \textit{entering}) \\ -1 & \text{if } z(t) < z_1 \ \vee \ z(t) > z_2 \quad (\text{I}{\to}\text{O}, \textit{exiting}) \end{cases}$$

3. Determine the direction u of the crossing, relative to the pore axis $\mathbf{p}$

$$u = \operatorname{sgn}\big([\mathbf{x}(t) - \mathbf{x}(t-1)] \cdot \mathbf{p}\big),$$

   where $u = +1$ signifies a flow "up" and $u = -1$ a flow "down."

4. Characterise the crossing event, and determine permeation events:

   Influx event $D = +1$ Only trajectories O→I→O can represent permeation events, i.e. they always start with an influx event, so for the particle i that crossed a boundary at t store the influx event $(i, t, u(t))$.

   Efflux event $D = -1$ An efflux event signals a permeation event if the particle exits through the face opposite to the one it entered a previous time t'. Compare the direction of the efflux $u(t)$ with the one of the influx $u(t')$, which was stored as the influx event $(i, t', u(t'))$.

$$u(t) = u(t') \Rightarrow \text{ permeation event.}$$

   Then the recorded event $(i, t', u(t'))$ is erased.

Output

The `-dat` file contains instantaneous values at each time step t whereas the `-cdat` data file records the corresponding cumulative values up to t,

column:	1	2	3	4	5	6
`-dat`:	t	$\Phi^+(t) + \Phi^-(t)$	$\Phi^+(t)$	$\Phi^-(t)$	$\Phi(t)$	$\tau_\mathrm{p}(t)$
`-cdat`:	t	$\mathcal{F}^+(t) + \mathcal{F}^-(t)$	$\mathcal{F}^+(t)$	$\mathcal{F}^-(t)$	$\mathcal{F}(t)$	$\langle \tau_\mathrm{p}(t) \rangle$
units:	ps	ps^{-1}	ps^{-1}	ps^{-1}	ps^{-1}	ps

The index file `-ncr` lists the indices of all molecules that permeated the pore successfully. The data file `-res` contains a histogram of permeation times τ_p.

Options

Files

`-f`	`traj.xtc`	Input	Generic trajectory: xtc trr trj gro g96 pdb

-s	topol.tpr	Input	Structure+mass(db): tpr tpb tpa gro g96 pdb xml
-n	index.ndx	Input, Opt.	Index file
-o	flux.xvg	Output	xvgr/xmgr file
-dat	flux.dat	Output	Generic data file
-cdat	cflux.dat	Output	Generic data file
-ncr	crossed.ndx	Output	Index file
-res	residency.dat	Output	Generic data file

Other options

-h	bool	no	Print help info and quit
-nice	int	0	Set the nicelevel
-b	time	-1	First frame (ps) to read from trajectory
-e	time	-1	Last frame (ps) to read from trajectory
-dt	time	-1	Only use frame when t MOD dt = first time (ps)
-w	bool	no	View output xvg, xpm, eps and pdb files
-m	bool	yes	index contains atoms, but g_flux counts the molecules
-axis	vector	0 0 1	Vector pointing parallel to the pore axis
-cpoint	vector	0 0 0	[hidden] Point on the pore axis
-z1	real	-1	Confine waters to have a z coordinate larger than z1, and
-z2	real	-1	smaller than z2

Known problems and limitations

- -m behaves different from the standard usage within the g_* programs—it figures out for itself what the molecules are and does not need molecule numbers but atom_ids.
- -m is the *default* behaviour!
- Despite appearance it only makes sense to specify a pore axis approximately parallel to the *z*-axis because we only really base the definition of the boundaries on *z* coordinates.

Example

Suggested use for water: Create an index file for the water molecules:

```
echo -e "keep 0\ndel 0\nr SOL\nq\n"
| make_ndx -f in.pdb -o sol.ndx
```

Then run

```
g_flux -m -f traj.xtc -s topol.tpr -n sol.ndx
-cpoint 2.3.  2.3 2 -z1 1.9 -z2 2.7
```

and analyse the output in `flux.dat`, `cflux.dat`, and `residency.dat`.

D.2.3 g_ri3Dc

Usage

```
g_ri3Dc -f trajectory.xtc -s topol.tpr [OPTIONS]
```

Synopsis

g_ri3Dc places a 3D grid into a simulation box and counts the number of molecules (typically water) in each cell over a trajectory. The rectangular grid is large enough to encompass a cylinder of given radius and height (specify radius R, lower and upper z, and a point on the pore axis (which is always parallel to the z-axis). If no parameters are given for the cylinder, a cylinder is fit into the simulation box. The cylinder (and thus the grid) is fixed. The spatial resolution can be given either as one number for all three dimensions (cartesian coordinates; unit is nm) or separately as three (*Note*: (1) It is recommended to use at least the same grid spacing in x and y (which is compatible with a cylindrical symmetry along z)). (2) The exact dimensions of the grid (radius etc) are re-adjusted to integral numbers of the grid spacing.) At the end, a density map (occupancy time over total simulation time $\mathfrak{T}_{\text{cell}}/\mathfrak{T}$) is written to the grid file.

The output is the number density, averaged over the trajectory. Read this file into a_ri3Dc, the grid analysis program, and produce radial distribution functions, density plots, pore profiles etc. The grid data is written as a binary file (in the machine independent xdr format). a_ri3Dc has an option to write it out as ascii text.

Options

Files

`-f`	`traj.xtc`	Input	Generic trajectory: xtc trr trj gro g96 pdb
`-s`	`topol.tpr`	Input	Structure+mass(db): tpr tpb tpa gro g96 pdb xml
`-n`	`index.ndx`	Input, Opt.	Index file
`-grid`	`gridxdr.dat`	Output	Generic data file

Other options

`-h`	bool	`no`	Print help info and quit
`-nice`	int	`0`	Set the nicelevel
`-b`	time	`-1`	First frame (ps) to read from trajectory
`-e`	time	`-1`	Last frame (ps) to read from trajectory
`-dt`	time	`-1`	Only use frame when t MOD dt = first time (ps)
`-w`	bool	`no`	View output xvg, xpm, eps and pdb files
`-m`	bool	`yes`	index contains atoms, but g_ri3Dc counts the molecules
`-cpoint`	vector	`0 0 0`	Point on the central symmetry axis of the grid
`-R`	real	`0`	Maximum radius that should be contained in the grid
`-z1`	real	`0`	Center grid between z1, and ...
`-z2`	real	`0`	z2 (these boundaries are kept fixed!)
`-delta`	vector	`0.02 0.02 0.02`	Spatial resolution in X, Y, and Z (in nm)
`-dtweight`	bool	`yes`	[hidden] Weigh counts with the time step (yes) or count all equally (no)
`-subtitle`	string		Some text to add to the output graphs

Caveats and known limitations

- The program guarantees to use the user supplied grid spacing. If the other dimensions are incommensurable with Delta they are changed to comply.
- If you want radial distribution functions (yes, you do!) always use Delta[XX] == Delta[YY] to keep the cylindrical symmetry
- *z*-axis is the only allowed axis (and this will probably not change in the future)
- `-m` behaves different from the standard usage within the `g_*` programs – it figures out *for itself* what the molecules are and does not need *molecule* numbers but *atom_ids*.
- `-m` is the *default* behaviour. It works nicely with a SOL index file but more complicated solvents are untested.
- For **ions** you haved to use the `-nom` option!
- The XDR file is not compressed.
- Even the developer mistypes the name frequently

Example

Suggested use for water. Create an index file for the water molecules with

```
echo -e "keep 0\ndel 0\nr SOL\nq\n" \
| make_ndx -f in.pdb -o sol.ndx
```

Then run

```
g_ri3Dc -m -f traj.xtc -s topol.tpr -n sol.ndx
-delta 0.05 -grid grid.dat
```

(`-m` is the default but is shown here because the usage is different from the standard gromacs analysis programs). The 3D density is in file `grid.dat` and can be analysed with a_ri3Dc (see Appendix D.2.4).

D.2.4 a_ri3Dc

Usage

```
a_ri3Dc -grid gridxdr.dat [OPTIONS]
```

Synopsis

a_ri3Dc analyses a 3D grid produced by g_ri3Dc. The 2D projections are written in a format suitable for the fast 2D density plotter xfarbe.[223] It also writes a parameter file 'XFarbe' which can be used automatically by setting the environment variable '`XAPPLRESDIR=.`', and then running xfarbe, e.g. '`xfarbe rzp.dat` ' All geometry options (radius, z1, z2) that are not set and appear as '0' are set from the grid dimensions once it is read in. The `-dump` option converts a binary grid file into ascii txt. The `-plt` option produces a binary density file for gOpenMol, which you might have to rename to `xxx.plt`. vmd since version 1.8.2 can also read plt files and render them. The average density in the test cylinder is printed on std out, using the set units.

Output

density z-averaged	$n(x,y) = {L_z}^{-1} \int dz\, n(x,y,z)$
density y-averaged	$n(x,z) = {L_y}^{-1} \int dy\, n(x,y,z)$
density x-averaged	$n(y,z) = {L_x}^{-1} \int dx\, n(x,y,z)$
density radially averaged	$n(r,z) = (2\pi)^{-1} \int d\phi\, n(r,\phi,z)$
pore profile	$(z, R_G(z), R_G(z)/R)$
radial distribution function (rdf)	$n(r) = (2\pi\ L_z)^{-1} \int d\phi\, dz\, n(r,\phi,z)$
axial distribution function (zdf)	$n(z) = (L_x L_y)^{-1} \int dx\, dy\, n(x,y,z)$

$n(r)$ and $n(z)$ are averaged densities (normalisation is straightforward in order to turn them into 'real' probability distributions). The local density axial distribution function `-lzdf` averages over all occupied grid cells per z-slice and divides by an effective area which is determined from the 'radius of gyration' of the density

$$R_G^2(z) := \frac{2 \int_0^{2\pi} d\phi \int_0^{R'} dr\, r^3\, n(r,\phi,z)}{\int_0^{2\pi} d\phi \int_0^{R'} dr\, r\, n(r,\phi,z)} \tag{2.48}$$

as explained in Section 2.1.4. This radius is a good approximation to the pore profile. The density itself is in the `lzdf.xvg` file.

For diagnostic purposes one can also plot the radial distributions of the unoccupied cells (holes in the grid) in order to find suitable grid spacings.

Options

Files

`-grid`	`gridxdr.dat`	Input	Generic data file
`-profile`	`profile.xvg`	Output	xvgr/xmgr file
`-xyp`	`xyp.dat`	Output	Generic data file
`-xzp`	`xzp.dat`	Output, Opt.	Generic data file
`-yzp`	`yzp.dat`	Output, Opt.	Generic data file
`-rzp`	`rzp.dat`	Output	Generic data file
`-rdf`	`rdf.xvg`	Output	xvgr/xmgr file
`-zdf`	`zdf.xvg`	Output	xvgr/xmgr file
`-lzdf`	`lzdf.xvg`	Output	xvgr/xmgr file
`-hxyp`	`hxyp.dat`	Output, Opt.	Generic data file
`-hrzp`	`hrzp.dat`	Output, Opt.	Generic data file
`-hrdf`	`hrdf.xvg`	Output, Opt.	xvgr/xmgr file
`-dump`	`gridasc.dat`	Output, Opt.	Generic data file
`-plt`	`plt.dat`	Output, Opt.	Generic data file

Other options

-h	bool	no	Print help info and quit
-nice	int	0	Set the nicelevel
-R	real	0	Maximum radius that should be contained in the grid
-z1	real	0	Center grid between z1, and ...
-z2	real	0	z2 (these boundaries are kept fixed!)
-delta	vector	0.02 0.02 0.02	[hidden] Spatial resolution in X, Y, and Z for resampling (in nm)
-minocc	real	0	[hidden] The occupancy of a cell must be larger than this number so that it is counted as occupied when calculating the volume, effective radius and local density axial distribution -lzdf. This is given in the chosen units (see -unit).
-subtitle	string		Some text to add to the output graphs
-mirror	bool	yes	mirror the radial projection P(r,z) to create the impression of a full view of the pore
-unit	enum	unity	divide number density (in nm^3) by 1, the density of SPC water at 300K and 1 bar, in mol/l or in Angstrom^3. Allowed values: unity, SPC, molar or Angstrom
-xfarbe-maxlevel	real	48.4841	xfarbe will plot 15 equally space level up to this density (the unit must be the same as for the -unit option!) Default is 1.5 SPC bulk.
-holes	bool	no	[hidden] calculate all hole distribution functions, using default filenames
-radnr	int	0	[hidden] block-average the first rad_ba_nr radial bins...
-radstep	int	1	[hidden] ...in blocks of rad_ba_step

Known limitations

- The radial bin width DeltaR is fixed to (Delta[XX]+Delta[YY])/2). In any case one should never have different bin widths in X and Y.
- There are still a few hidden options of questionable usefulness. Resampling (=changing Delta) is not implemented yet.
- gOpenMol plt binary file comes out with wrong suffix
- note: -minocc also influences -lzdf

D.2.5 g_wham

Usage

```
g_wham [OPTIONS] file1.pdo.gz file2.pdo ...
```

Synopsis

g_wham is an analysis program that implements the Weighted Histogram Analysis Method (WHAM).[174,175] It is intended to analyze `.pdo` files generated by mdrun with umbrella sampling to create a potential of mean force (PMF).

Data files (`pdo`) may be gzipped and are simply listed on the command-line. Only frames between the begin time `-b` and the end time `-e` are included, giving the opportunity to discard an initial equilibration phase (-1 means 'no limit'). The program will discard any data that is outside of the interval $[\min z; \max z]$. The program will output the true lowest and highest coordinate values after completion, so these values can be used the next time or alternatively with `-noprof` only min and max are calculated in a first pass.

Output:

The first column in both data files is the position z along the sampled coordinate in nm. The histogram contains the data for each individual window. The profile `-o` contains the unbiased distribution $Z(z)$ and the PMF $G(z)/kT = -\ln Z(z)$:

column:	1	2	3
quantity:	z	$Z(z)$	$G(z)$
units:	nm		kT

Files

`-o`	`profile.xvg`	Output	xvgr/xmgr file
`-hist`	`histo.xvg`	Output	xvgr/xmgr file

Other options

`-h`	bool	`yes`	Print help info and quit
`-nice`	int	`19`	Set the nicelevel
`-b`	time	`-1`	First frame (ps) to read from trajectory
`-e`	time	`-1`	Last frame (ps) to read from trajectory
`-tu`	enum	`ps`	Time unit: `ps`, `fs`, `ns`, `us`, `ms`, `s`, `m` or `h`
`-min`	real	`0`	Minimum coordinate in profile
`-max`	real	`0`	Maximum coordinate in profile
`-bins`	int	`100`	Number of bins in profile
`-prof`	bool	`yes`	Calculate profile, when 'no' only calculate min and max

-temp	real	298	Temperature
-flip	bool	no	Combine halves of profile
-tol	real	1e-05	Tolerance

History

g_wham is part of the gromacs package. It was broken in version 3.2.1. It was modified to allow processing of output from umbrella sampling data from mdrun and the format version was incremented to 3.2, breaking any backward compatibility. The new version can read data files in uncompressed and GNUzipped format. It allows sub-selections of the data through the -b and -e flags. The code has been cleaned up and modularised, making future enhancements for higher dimensional WHAM easier.

Bugs and Limitations

- This version of g_wham reads pdo files with version number 3.2. It will *not* read version 3.0 (produced by the mdrun binary from the official gromacs 3.2.1 distribution).
- Only 1D PMF along a box axis implemented.
- Symmetric profiles (-flip) do not work.

Example

The first 100 ps of data are discarded as equilibration (this is necessary but one should check for each system by calculating PMFs for different time windows, using the -b and -e options; the equilibration phase can be significantly longer than the data collection, e.g. 2.5 ns equilibration and 0.5 ns data). First the minimum and the maximum z values need to be found because they are need for setting up the histogram. The -noprof option reads in all data and prints the limits on standard output:

```
find .  -name '*.pdo*'
| xargs g_wham -noprof -temp 300 -b 100
```

With these limits (in this example, -0.9 nm $\leq z \leq 0.9$ nm) the WHAM procedure is carried out with

```
find .  -name '*.pdo*'
| xargs g_wham -temp 300 -b 100 -min -0.9
-max 0.9 -tol 1e-5 -bins 500 -o pmf.xvg -hist
histogram.xvg
```

The tolerance for the self-consistent solution of the WHAM equations is important. 10^{-3} kT is not enough. 10^{-5} kT is the minimum that resolves a test-input step function (generated with fakepmf, see Appendix D.3.2). It is recommened to plot the PMF for different values of the tolerance and check its convergence. The PMF is found in the output file `pmf.xvg`, and the histograms from all input files are found in `histogram.xvg`. They can be displayed with xmgrace,

```
xmgrace -block pmf.xvg -bxy 1:3
xmgrace -nxy histogram.xvg
```

D.3 Trajectory generation

D.3.1 Confinement in mdrun

Synopsis

mdrun from version 3.2.1 of gromacs was modified in the parts concerned with umbrella sampling in order to make it work together with the WHAM analysis program g_wham and to add an experimental cylindrical confinement potential. The output format was changed (current version of the format is now 3.2 which is incompatible with the previous version 3.0 but can be read by the modified version of g_wham). mdrun and g_wham are now versioned 3.2.1-LMB.

Umbrella sampling is activated with the `-pi` option. An input file `pull.ppa` must be provided, which controls all aspects of umbrella sampling, an index file which defines the group that is sampled, and a standard run input `tpr` file. Umbrella sampling can be performed in parallel.

Confinement potential

A cylindrical confinement potential was added. Outside the pore region a sampled particle is not bounded in the plane perpendicular to the umbrella potential. Thus the 1D PMF is not well defined and any results from WHAM are meaningless.[169,170] With a confinement potential, starting from outside a confinement radius R_C in addition to the umbrella potential, it should be possible to obtain a defined PMF. The flat-bottomed confinement potential has the form

$$U_C(\mathbf{x}) = \tfrac{1}{2}\, k_C\, \Theta[\rho(\mathbf{x}) - R_C]\, (\rho(\mathbf{x}) - R_C)^2, \tag{D.11}$$

where $\rho(\mathbf{x})$ denotes the distance from the line along which the umbrella windows are distributed (typically, just the z-axis), and k_C is the harmonic force constant.

Input file format

The input `ppa` file for umbrella sampling needs to comply to the following example format (lines starting with ';' are comments and are ignored by the parser).

```
; Input for umbrella sampling
; version: Gromacs 3.2.1-LMB

verbose   = yes
runtype   = umbrella       ; umbrella sampling
pulldim   = N N Y          ; restrict in Z dimension
reftype   = com_t0         ; compute center of umbrella potential
                           ; relative to initial COM of
reference_group = Protein  ; the reference group

; Umbrella sampling
; group which is subjected to the umbrella potential
; (see the index file):
group_1   = NA_-0.337
K1        = 3909.50;            ; kJ / (mol nm^2)
Pos1      = 0.000 0.000 -0.337  ; centre of the umbrella potential
; confinement potential: flat-bottomed cylinder
;                        centered on ref_com
KConfine = 4500.000000 ; harmonic force constant, kJ/(mol*nm^2)
RConfine = 1.500       ; radius past which the potential is
                       ; switched on (nm)
```

Up to four group sections (`group_1` to `group_4`) can be added. A confinement potential is only switched on if both `RConfine` and `KConfine` are non-zero; `RConfine` defaults to `0`, thus leaving out these variables will automatically run without confinement. The index group mentioned in `group_1` must be defined in the accompanying index file.

Output file format

The output `pdo` file has a header which is versioned with the output format number, currently 3.2. Data follows after the last header line `#####`, with the first column being the time in ps and subsequent columns the deviation from the umbrella centre of each of the `N_pullgroups` groups (in nm).

```
# UMBRELLA      3.2
# pulldim  0 0 1
```

```
# nSkip 1
# reference_group 'Protein'
# N_pullgroups 1
# Group 1 'NA_-0.337'  Pos -0.337000  K 3909.500000
#####
0.000000        0.003102
0.002000        0.002794
0.004000        0.002398
...
2999.996094     -0.029982
2999.998047     -0.029695
3000.000244     -0.029211
```

D.3.2 fakepmf

Purpose

Instead of deducing a potential of mean force from umbrella sampling the aim is to generate data from a known PMF. Then these data can be processed by g_wham (or any other unbiasing code) to recreate the original PMF. Comparing the initial PMF and the recreated one allows assessment of the unbiasing procedure.

The output is a "time series" of $\xi = z - z_i$ values as if sampled by a particle restrained to an umbrella window at position z_i, given the underlying PMF $W(z)$. The simplest approach is to randomly draw z values from the underlying statistical distribution $p(z)$ [Eq. (D.12)]. Although subsequent values drawn will be uncorrelated—in contrast to values from a dynamical trajectory (where there is strong correlation between close-by time steps)—the data are suitable as test input for unbiasing methods like the weighted histogram analysis method (WHAM).

Usage

```
fakepmf [OPTIONS]
fakepmf --zeta W_0 [OPTIONS]
```

Synopsis

fakepmf creates data files that mimic an umbrella sampling run of a known PMF, which corresponds to a probability distribution of (modified) Gaussians whose width depends on the force constant k of the umbrella potential and the underlying PMF. The default is a flat PMF but if `--zeta` W_0 is selected, a step

function of height W_0 is simulated. The number of samples is determined by a fake run length (in ps) and a step size of 2 fs.

The output files have the same format as produced by mdrun (version 3.2.1-LMB) with umbrella sampling and can be directly fed into g_wham. Flat PMFs are stored in files with name `FLAT_zz`$_i$`.pdo.gz` whereas files with a step function PMF are named `STEP_zz`$_i$`.pdo.gz`.

Probability distributions

If the functional form of the PMF $W(z)$ is known then the probability distribution for the position of a particle z is

$$p_i(z) = \mathcal{Z}^{-1} \exp\left[-\beta\left(\tfrac{1}{2}\,k\,(z-z_i)^2 + W(z)\right)\right] \tag{D.12}$$

where k is the strength of the harmonic umbrella restraint, $\xi = z - z_i$ the deviation of the particle from the centre z_i of the i-th umbrella window, and $\mathcal{Z}^{-1}$ the normalisation

$$\mathcal{Z} = \int_{-\infty}^{+\infty} d\xi\, \exp\left[-\beta\left(\tfrac{1}{2}\,k\,\xi^2 + W(\xi + z_i)\right)\right]. \tag{D.13}$$

For a **flat PMF**

$$W(z) = W_0 = \text{const} \tag{D.14}$$

all p_i are Gaussians, centered at z_i with width $\sqrt{1/k\beta}$. The normalisation is trivial and the probability distribution for z becomes

$$p_i^{\text{flat}}(z; W_0) = \sqrt{\frac{\beta k}{2\pi}} \exp\left[-\tfrac{1}{2}\,\beta k\,(z-z_i)^2\right]. \tag{D.15}$$

For a **step function** of height W_0 at position z_0

$$W(z) = W_0\,\Theta(z - z_0) \tag{D.16}$$

the normalisation is

$$\mathcal{Z}^{\text{step}} = \sqrt{\frac{\pi}{2\beta k}}\left(1 + \operatorname{erf}[\sqrt{\tfrac{1}{2}\,\beta k}(z_0 - z_i)] + e^{-\beta W_0}\,\operatorname{erfc}[\sqrt{\tfrac{1}{2}\,\beta k}(z_0 - z_i)]\right)$$

and the probability distribution

$$p_i^{\text{step}}(z; z_0, W_0) = \frac{1}{\mathcal{Z}^{\text{step}}} \exp\left[-\beta\left(\tfrac{1}{2}\,k\,(z-z_i)^2 + W_0\,\Theta(z - z_0)\right)\right]. \tag{D.17}$$

Algorithm

The task is to draw values z according to a probability distribution $p_i(z)$ (D.12). The standard method is to use a uniform deviate x on the interval $0 < x < 1$ and transform it to yield a non-uniform deviate z obeying (D.12).[162] The random generator `ran1()` from Press et al.[162], p280 produces uniform deviates. For the flat PMF (D.14) the probability distribution (D.15) is a Gaussian, and the Gaussian deviate is easily obtained through a *Box-Muller* transformation (Ref. 162, p289) because the inverse of the cumulative distribution is available. The step function distribution Eq. (D.17) is more complictated but can be sampled using the rejection method (Ref. 162, p290). A comparison function $f(z)$ is required for which f-distributed deviates can be easily calculated and which envelopes the target distribution $p(z) \equiv p^{\text{step}}(z; z_0, W_0)$. The procedure is

1. to produce a f-distributed deviate z from a uniform deviate, then
2. to compare a second uniform deviate $0 < y < 1$ to the ratio $p(z)/f(z)$ and accept z as a p-distributed deviate if $y < p(z)/f(z)$.

The comparison function f was chosen as a non-normalised Gaussian

$$f(z) = 2\pi \max[p(z_0), p(z_i)] \exp[-\tfrac{1}{2}\beta k (z - z')^2], \text{ with}$$
$$z' = \begin{cases} z_0, & \text{if } p(z_0) \geq p(z_i) \\ z_i, & \text{if } p(z_0) < p(z_i) \end{cases}.$$

$f(z)$ is guaranteed to envelope $p(z)$ for all positions of the umbrella centre z_i, $1 \leq i \leq N$ and the position of the step barrier z_0 (see also Fig. 2.2 on page 49 for an illustration of $p(z)$).

Options

`-h,`	`--help`		this help
`-L,`	`--length`	L	Length of sampled stretch in nm [`1.9`]
`-N,`	`--nwin`	N	Number of windows [`101`]
`-k,`	`--forceconst`	k	Force constant in $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ [`7700.0`]
`-t,`	`--tsim`	$\mathfrak{T}$	'Simulation' time in ps, determines number of samples as the 'stepsize' is fixed at 2 fs [`1000.0`]
`-T,`	`--temperature`	T	temperature in K [`298.0`]
`-z,`	`--zeta`	W_0	height of step function in kT

E Publications

A number of publications resulted from the work on model pores and the nicotinic receptor. Chapters 3 to 5 are largely based on published articles, as detailed below. An article incorporating the material of Appendix A is submitted and Chapter 6 can be considered as "manuscript in preparation".

E.1 Research articles

- Amiri S, Tai K, Beckstein O, Biggin P C and Sansom M S P, 2005 The $\alpha 7$ nicotinic acetylcholine receptor: Molecular modelling, electrostatics and energetics of permeation. *Molec. Membr. Biol.* **22**, 1–11. (Ref. 41)

  PMF calculations of Na^+ in the $\alpha 7$ nAChR pore and influence of the extracellular ligand binding domain on permeation. (Appendix A)

- Beckstein O, Tai K and Sansom M S P, 2004 Not ions alone: Barriers to ion permeation in nanopores and channels. *J. Am. Chem. Soc.* **126**, 14694–14695. (Ref. 98)

  PMF calculations of Na^+ through model pores reveal the importance of treating water correctly in confined hydrophobic geometries; the free energy barriers support the hydrophobic gating hypothesis. (Chapter 5)

- Beckstein O and Sansom M S P, 2004 The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Physical Biology* **1** 42–52. (Ref. 97)

  A discussion of principles of (hydrophobic) gating, made explicit in a simple model system. (Chapter 4)

- Beckstein O and Sansom M S P, 2003 Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc. Natl. Acad. Sci. U.S.A.* **100** 7063–7068. (Ref. 96)

  Identification of capillary evaporation/condensation as a possible mechanism to explain solvent behaviour at the molecular scale, leading to liquid-vapour oscillations in hydrophobic pores. The article was accompanied by a commentary by H. Grubmüller[320]. (Chapter 3)

- Beckstein O, Biggin P C and Sansom M S P, 2001 A hydrophobic gating mechanism for nanopores. *J. Phys. Chem. B* **105** 12902–12905. (Ref. 62)

  Initial MD simulations of water in pores which mimic the gate region of nAChR (at a crude level because the atomic resolution structure[29] was only published in 2003) prompt us to state the "hydrophobic gating hypothesis" for ion channels.

E.2 Review articles

- Beckstein O, Biggin P C, Bond P, Bright J N, Domene C, Grottesi A, Holyoake J and Sansom* M S P, 2003 Ion channel gating: insights via molecular simulations. *FEBS Lett.* **555** 85–90. (Ref. 251)

  *MSPS is main author of this review

- Sansom M S P, Bond P, Beckstein O, Biggin P C, Faraldo-Goméz J, Law R J, Patargias G and Tieleman D P, 2002 Water in ion channels and pores—simulation studies, in *Ion channels: from atomic resolution physiology to functional genomics* (edited by G Bock and J A Goode), vol. 245 of *Novartis Foundation Symposia*, 66–78, Novartis Foundation, London (John Wiley & Sons, Chicester). (Ref. 321)

Bibliography

1. Brovchenko I, Paschek D and Geiger A, 2000 Gibbs ensemble simulation of water in spherical cavities. *J. Chem. Phys.* **113** 5026–5036.

2. Davies P, 2001 The origin of life I: When and where did it begin? *Sci. Progr.* **84** 1–16.

3. Davies P, 2001 The origin of life II: How did it begin? *Sci. Progr.* **84** 17–29.

4. Deamer D, Dworkin J P, Sandford S A, Bernstein M P and Allamandola L J, 2002 The first cell membranes. *Astrobiol.* **2** 371–381.

5. Monnard P A and Deamer D W, 2002 Membrane self-assembly processes: Steps toward the first cellular life. *Anat. Rec.* **268** 196–207.

6. Pohorille A, 2002 From organic molecules in space to the origins of life and back. *Adv. Space Res.* **30** 1509–1520.

7. Weiss T F, 1996 *Cellular Biophysics*, vol. 1 (MIT Press, Cambridge, Massachusetts, USA).

8. Berry S, 2002 The chemical basis of membrane bioenergetics. *J. Mol. Evol.* **54** 595–613.

9. Finkelstein A, 1987 *Water Movement Through Lipid Bilayers, Pores, and Plasma Membranes. Theory and Reality*, vol. 4 of *Distinguished Lecture Series of the Society of General Physiologists* (John Wiley & Sons, New York).

10. Preston G M, Carroll T P, Guggino W B and Agre P, 1992 Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* **256** 385–387.

11. King L S, Kozono D and Agre P, 2004 From structure to disease: the evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell Biol.* **5** 687–698.

12. Hille B, 2001 *Ion Channels of Excitable Membranes*, 3rd edn. (Sinauer Associates, Sunderland MA, U.S.A.).

13. Fujiyoshi Y, Mitsuoka K, de Groot B L, Philippsen A, Grubmüller H, Agre P and Engel A, 2002 Structure and function of water channels. *Curr. Opin. Struct. Biol.* **12** 509–515.

14. Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E, 2000 The Protein Data Bank. *Nucleic Acids Res.* **28** 235–242, `http://www.rcsb.org/pdb/`.

15. White S H, 2004 The progress of membrane protein structure determination. *Protein Sci.* **13** 1948–1949, `http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html`.

16. Doyle D A, Morais-Cabral J, Pfützner R A, Kuo A, Gulbis J M, Cohen S L, Chait B T and MacKinnon R, 1998 The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* **280** 69–77.

17. Aidley D J and Stanfield P R, 1996 *Ion Channels: Molecules in Action* (Cambridge University Press, Cambridge).

18. Lester H A, Dibas M I, Dahan D S, Leite J F and Dougherty D A, 2004 Cys-loop receptors: new twists and turns. *Trends Neurosci.* **27** 329–336.

19. Zhorov B S and Tikhonov D B, 2004 Potassium, sodium, calcium and glutamate-gated channels: pore architecture and ligand action. *J. Neurochem.* **88** 782–799.

20. Sukharev S and Anishkin A, 2004 Mechanosensitive channels: what can we learn from 'simple' model systems? *Trends Neurosci.* **27** 345–351.

21. Clapham D E, 2003 TRP channels as cellular sensors. *Nature* **426** 517–524.

22. Trudell J R and Harris R A, 2004 Are sobriety and consciousness determined by water in protein cavities? *Alcoholism Clin. Exp. Res.* **28** 1–3.

23. Hogg R C, Raggenbass M and Bertrand D, 2003 Nicotinic acetylcholine receptors: from structure to brain function. *Rev. Physiol. Biochem. Pharmacol.* **147** 1–46.

24. Corey D P, 2004 New TRP channels in hearing and mechanosensation. *Neuron* **39** 585–588.

25. Ashcroft F M, 2000 *Ion Channels and Disease: Channelopathies* (Academic Press, San Diego).

26. Jiang Y, Lee A, Chen J, Cadene M, Chait B T and MacKinnon R, 2002 The open pore conformation of potassium channels. *Nature* **417** 523–526.

27. Zhou Y, Morais-Cabral J H, Kaufman A and MacKinnon R, 2001 Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0 Å resolution. *Nature* **414** 43–48.

28. Jiang Y, Lee A, Chen J, Cadene M, Chait B T and MacKinnon R, 2002 Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417** 515–522.

29. Miyazawa A, Fujiyoshi Y and Unwin N, 2003 Structure and gating mechanism of the acetylcholine receptor pore. *Nature* **423** 949–955, http://dx.doi.org/10.1038/nature01748.

30. Dutzler R, Campbell E B, Cadene M, Chait B T and MacKinnon R, 2002 X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* **415** 287–294.

31. Dutzler R, Campbell E B and MacKinnon R, 2003 Gating the selectivity filter in ClC chloride channels. *Science* **300** 108–112.

32. Accardi A and Miller C, 2004 Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature* **427** 803–807.

33. Maduke M, Miller C and Mindell J A, 2000 A decade of ClC chloride channels: structure, mechanism, and many unsettled questions. *Annu. Rev. Biophys. Biomol. Struct.* **29** 411–438.

34. Dutzler R, 2004 The structural basis of ClC chloride channel function. *Trends Neurosci.* **27** 315–320.

35. Nikaido H, 2003 Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Molec. Biol. Rev.* **67** 593–656.

36. Koebnik R, Locher K P and Van Gelder P, 2000 Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37** 239–253.

37. Bond P J and Sansom M S P, 2004 The simulation approach to bacterial outer membrane proteins. *Mol. Membr. Biol.* **21** 151–161.

38. Pautsch A and Schulz G E, 1998 Structure of the outer membrane protein A transmembrane domain. *Nature Struct. Biol.* **11** 1013–1017.

39. Bond P, Faraldo-Gómez J D and Sansom M S P, 2002 OmpA: A pore or not a pore? Simulation and modeling studies. *Biophys. J.* **83** 763–775.

40. Benz J and Hofmann A, 1997 Annexins: from structure to function. *Biol. Chem.* **378** 177–183.

41. Amiri S, Tai K, Beckstein O, Biggin P C and Sansom M S P, 2005 The $\alpha 7$ nicotinic acetylcholine receptor: Molecular modelling, electrostatics and energetics of permeation. *Mol. Membr. Biol.* **22** 151–162, http://dx.doi.org/10.1080/09687860500063340.

42. Smart O S, Neduvelil J G, Wang X, Wallace B A and Sansom M S P, 1996 HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14** 354–360, http://hole.biop.ox.ac.uk/hole.

43. Humphrey W, Dalke A and Schulten K, 1996 VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **14** 33–38, `http://www.ks.uiuc.edu/Research/vmd/`.

44. Merritt E A and Bacon D J, 1997 Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277** 505–524, `http://www.bmsc.washington.edu/raster3d/raster3d.html`.

45. Spencer R H and Rees D C, 2002 The α-helix and the organization and gating of channels. *Annu. Rev. Biophys. Biomol. Struct.* **31** 207–233.

46. Betanzos M, Chiang C S, Guy H R and Sukharev S, 2002 A large iris-like expansion of a mechanosensitive channel protein induced by membrane tension. *Nature Struct. Biol.* **9** 704–710.

47. Eisenberg D and Kauzmann W, 1969 *The Structure and Properties of Water* (Oxford University Press, London).

48. Kuo A, Gulbis J M, Antcliff J F, Rahman T, Lowe E D, Zimmer J, Cuthbertson J, Ashcroft F M, Ezaki T and Doyle D A, 2003 Crystal structure of the potassium channel KirBac1.1 in the closed state. *Science* **300** 1922–1926.

49. Morais-Cabral J H, Zhou Y and MacKinnon R, 2001 Energetic optimization of ion conduction rate by the K^+ selectivity filter. *Nature* **414** 37–42.

50. Biggin P C and Sansom M S P, 2002 Open-state models of a potassium channel. *Biophys. J.* **83** 1867–1876.

51. Chang G, Spencer R H, Lee A T, Barclay M T and Rees D C, 1998 Structure of the MscL homolog from *Mycobacterium tubercolosis*: a gated mechanosensitive ion channel. *Science* **282** 2220–2226.

52. Yoshimura K, Batiza A, Schroeder M, Blount P and Kung C, 1999 Hydrophilicity of a single residue within MscL correlates with increased channel mechanosensitivity. *Biophys. J.* **77** 1960–1972.

53. Yoshimura K, Batiza A and Kung C, 2001 Chemically charging the pore constriction opens the mechanosensitive channel MscL. *Biophys. J.* **80** 2198–2206.

54. Sukharev S, Betanzos M, Chiang C S and Guy H R, 2001 The gating mechanism of the large mechanosensitive channel MscL. *Nature* **409** 720–724.

55. Bass R B, Strop P, Barclay M and Rees D C, 2002 Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science* **298** 1582–1587.

56. Anishkin A and Sukharev S, 2004 Water dynamics and dewetting transitions in the small mechanosensitive channel MscS. *Biophys. J.* **86** 2883–2895.

57. Perozo E, Kloda A, Cortes D M and Martinac B, 2002 Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nature Struct. Biol.* **9** 696–703.

58. Unwin N, 1995 Acetylcholine-receptor channel imaged in the open state. *Nature* **373** 37–43.

59. Unwin N, 2000 The Croonian Lecture 2000. Nicotinic acetylcholine receptor and the structural basis of fast synaptic transmission. *Phil. Trans. Roy. Soc. London B* **355** 1813–1829.

60. Unwin N, Miyazawa A, Li J and Fujiyoshi Y, 2002 Activation of the nicotinic acetylcholine receptor involves a switch in conformation of the alpha subunits. *J. Mol. Biol.* **319** 1165–1176.

61. Unwin N, 2003 Structure and action of the nicotinic acetylcholine receptor explored by electron microscopy. *FEBS Lett.* **555** 91–95.

62. Beckstein O, Biggin P C and Sansom M S P, 2001 A hydrophobic gating mechanism for nanopores. *J. Phys. Chem. B* **105** 12902–12905, `http://dx.doi.org/10.1021/jp012233y`.

63. Parsegian A, 1969 Energy of an ion crossing a low dielectric membrane: solution to four relevant electrostatic problems. *Nature* **221** 844–846.

64. Chung S H, Hoyles M, Allen T and Kuyucak S, 1998 Study of ionic currents across a model membrane channel using brownian dynamics. *Biophys. J.* **75** 793–809.

65. Unwin N, 1993 Nicotinic acetylcholine-receptor at 9 Å resolution. *J. Mol. Biol.* **229** 1101–1124.

66. Moe P C, Levin G and Blount P, 2000 Correlating a protein structure with function of a bacterial mechanosensitive channel. *J. Biol. Chem.* **275** 31121–31127.

67. Kipling R, 1902 *Just So Stories* (Project Gutenberg), `http://www.gutenberg.org/dirs/etext01/jusss10.txt`, print edition published in 1902, accessed online October 2004.

68. Stroud R M, Miercke L J W, Connell J O, Khademi S, Lee J K, Remis J, Harries W, Robles Y and Akhavan D, 2003 Glycerol facilitator GlpF and the associated aquaporin family of channels. *Curr. Opin. Struct. Biol.* **13** 424–431.

69. Sui H, Han B G, Lee J K, Walian P and Jap B K, 2001 Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414** 872–878.

70. Savage D F, Egea P F, Robles-Colmenares Y, O'Connell III J D and Stroud R M, 2003 Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z. *PLoS Biol.* **1** 334–340.

71. Fu D, Libson A, Miercke L J, Weitzman C, Nollert P, Krucinski J and Stroud R M, 2000 Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290** 481–486.

72. Tajkhorshid E, Nollert P, Jensen M Ø, Miercke L J, O'Connell J, Stroud R M and Schulten K, 2002 Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* **296** 525–530.

73. Tournaire-Roux C, Sutka M, Javot H, Gout E, Gerbeau P, Luu D T, Bligny R and Maurel C, 2003 Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. *Nature* **425** 393–397.

74. Gonen T, Sliz P, Kistler J, Cheng Y and Walz T, 2004 Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature* **429** 193–197.

75. Khademi S, O'Connell Joseph I, Remis J, Robles-Colmenares Y, Miercke L J W and Stroud R M, 2004 Mechanism of ammonia transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science* **305** 1587–1594.

76. Raushel F M, Thoden J B and Holden H M, 2003 Enzymes with molecular tunnels. *Acc. Chem. Res.* **36** 539–548.

77. Hyde C C, Ahmed S A, Padlan E A, Miles E W and Davies D R, 1988 Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* **263** 17857–17871.

78. Binda C, Bossi R T, Wakatsuki S, Arzt S, Coda A, Curti B, Vanoni M A and Mattevi A, 2000 Cross-talk and ammonia channeling between active centers in the unexpected domain arrangement of glutamate synthase. *Structure* **8** 1299–1308.

79. Doukov T I, Iverson T M, Seravalli J, Ragsdale S W and Drennan C L, 2002 A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* **298** 567–572.

80. Darnault C, Volbeda A, Kim E J, Legrand P, Vernède X, Lindahl P A and Fontecilla-Camps J C, 2003 Ni-Zn-[Fe_4-S_4] and Ni-Ni-[Fe_4-S_4] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. *Nature Struct. Biol.* **10** 271–279.

81. Montet Y, Amara P, Volbeda A, Vernede X, Hatchikian E C, J. F M, M. F and Fontecilla-Camps J C, 1997 Gas access to the active site of Ni-Fe hydrogenases probed by X-ray crystallography and molecular dynamics. *Nature Struct. Biol.* **4** 523–526.

82. Dill K A and Bromberg S, 2003 *Molecular Driving Forces* (Garland Science, New York).

83. Lazaridis T, 2001 Solvent size vs cohesive energy as the origin of hydrophobicity. *Acc. Chem. Res.* **34** 931–937.

84. Pratt L R and Pohorille A, 2002 Hydrophobic effects and modeling of biophysical aqueous solution interfaces. *Chem. Rev.* **102** 2671–2692.

85. Chandler D, 2002 Two faces of water. *Nature* **417** 491.

86. Hummer G, Garde S, García A E, Paulaitis M E and Pratt L R, 1998 Hydrophobic effects on a molecular scale. *J. Phys. Chem. B* **102** 10469–10482.

87. Jorgensen W L, Gao J and Ravimohan C, 1985 Monte carlo simulations of alkanes in water: hydration numbers and the hydrophobic effect. *J. Phys. Chem.* **89** 3470–3473.

88. Humphrey W, Dalke A and Schulten K, 1996 VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **14** 33–38, `http://www.ks.uiuc.edu/Research/vmd/`.

89. Wallqvist A and Berne B J, 1995 Computer simulation of hydrophobic hydration forces on stacked plates at short range. *J. Phys. Chem.* **99** 2893.

90. Lum K, Chandler D and Weeks J D, 1999 Hydrophobicity at small and large length scales. *J. Phys. Chem. B* **103** 4570–4577.

91. Huang D M and Chandler D, 2000 Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **97** 8324–8327.

92. de Gennes P G, 1985 Wetting: statics and dynamics. *Rev. Mod. Phys.* **57** 827–863.

93. Stillinger F H, 1973 Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled particle theory. *J. Solution Chem.* **2** 141–158.

94. Huang X, Margulis C J and Berne B J, 2003 Dewetting-induced collapse of hydrophobic particles. *Proc. Natl. Acad. Sci. U.S.A.* **100** 11953–11958.

95. Tieleman D P, Biggin P C, Smith G R and Sansom M S P, 2001 Simulation approaches to ion channel structure-function relationships. *Quart. Rev. Biophys.* **34** 473–561.

96. Beckstein O and Sansom M S P, 2003 Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc. Natl. Acad. Sci. U.S.A.* **100** 7063–7068, `http://www.pnas.org/cgi/content/abstract/100/12/7063`.

97. Beckstein O and Sansom M S P, 2004 The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Physical Biology* **1** 42–52, `http://stacks.iop.org/1478-3975/1/42`.

98. Beckstein O, Tai K and Sansom M S P, 2004 Not ions alone: Barriers to ion permeation in nanopores and channels. *J. Am. Chem. Soc.* **126** 14694–14695, `http://dx.doi.org/10.1021/ja045271e`.

99. Chaikin P M and Lubensky T C, 2000 *Principles of condensed matter physics* (Cambridge University Press, Cambridge).

100. Rowlinson J S and Widom B, 2002 *Molecular Theory of Capillarity* (Dover Publications, Mineola, New York), reprint of Ref. 231.

101. Hansen J P and McDonald I R, 1990 *Theory of Simple Liquids*, 2nd edn. (Elsevier Academic Press, Amsterdam).

102. van Swol F and Henderson J R, 1986 Wetting at a fluid-wall interface. *J. Chem. Soc., Faraday Trans. 2* **82** 1685–1699.

103. van Swol F and Henderson J R, 1989 Wetting and drying transitions at a fluid-wall interface: Density-functional theory versus computer simulation. *Phys. Rev. A* **40** 2567–2578.

104. Nijmeijer M J P, Bruin C, Bakker A F and van Leeuwen J M J, 1991 Determination of the location and order of the drying transition with a molecular dynamics simulation. *Phys. Rev. B* **44** 834–837.

105. Allen M P and Tildesley D J, 1987 *Computer Simulations of Liquids* (Oxford University Press, Oxford).

106. McQuarrie D A, 1976 *Statistical Mechanics* (HarperCollins, New York).

107. Deutsch D, 1997 *The Fabric of Reality* (Penguin Books, London).

108. Huang K, 1987 *Statistical Mechanics*, 2nd edn. (John Wiley & Sons, New York).

109. Tegmark M, Strauss M A, Blanton M R, Abazajian K, Dodelson S, Sandvik H, Wang X, Weinberg D H, Zehavi I, Bahcall N A, Hoyle F, Schlegel D, Scoccimarro R, Vogeley M S, Berlind A, Budavari T, Connolly A, Eisenstein D J, Finkbeiner D, Frieman J A, Gunn J E, Hui L, Jain B, Johnston D, Kent S, Lin H, Nakajima R, Nichol R C, Ostriker J P, Pope A, Scranton R, Seljak U, Sheth R K, Stebbins A, Szalay A S, Szapudi I, Xu Y, Annis J, Brinkmann J, Burles S, Castander F J, Csabai I, Loveday J, Doi M, Fukugita M, Gillespie B, Hennessy G, Hogg D W, Ivezic Z, Knapp G R, Lamb D Q, Lee B C, Lupton R H, McKay T A, Kunszt P, Munn J A, O'Connell L, Peoples J, Pier J R, Richmond M, Rockosi C, Schneider D P, Stoughton C, Tucker D L, Berk D E V, Yanny B and York D G, 2004 Cosmological parameters from SDSS and WMAP. *Phys. Rev. D* **69** 103501.

110. Zwanzig R and Ailawadi N K, 1969 Statistical error due to finite time averaging in computer experiments. *Phys. Rev.* **182** 280–283.

111. Chandrasekhar S, 1943 Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15** 1–89.

112. Einstein A, 1905 Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. d. Phys.* **17** 549–560.

113. Adamson A W, 1990 *Physical Chemistry of Surfaces*, 5th edn. (John Wiley & Sons, Inc, New York).

114. Rowlinson J S, 1986 The statistical mechanics of small systems. *J. Chem. Soc., Faraday Trans. 2* **82** 1801–1815, Faraday Symposium 20, The Lennard-Jones Lecture.

115. Barrat J L and Hansen J P, 2003 *Basic Concepts for Simple and Complex Liquids* (Cambridge University Press, Cambridge).

116. Chalikian T V, 2003 Volumetric properties of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32** 207–235.

117. Richards F M, 1977 Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6** 151–176.

118. Connolly M L, 1983 Solvent-accessible surfaces of proteins and nucleic-acids. *Science* **221** 709–713.

119. Connolly M L, 1985 Computation of molecular volume. *J. Am. Chem. Soc.* **107** 1118–1124.

120. Lee B and Richards F M, 1971 The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55** 379–400.

121. Sanner M F, Olson A J and Spehner J C, 1996 Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **38** 305–320.

122. Greer J and Bush B L, 1978 Macromolecular shape and surface maps by solvent exclusion. *Proc. Natl. Acad. Sci. U.S.A.* **75** 303–307.

123. Weeks J D, Chandler D and Andersen H C, 1971 Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.* **54** 5237–5247.

124. Born M and Oppenheimer R, 1927 Zur Quantentheorie der Molekeln. *Ann. d. Phys.* **84** 457–484.

125. Schwabl F, 1993 *Quantenmechanik*, 4th edn. (Springer, Heidelberg).

126. Goldstein H, 1980 *Classical Mechanics*, 2nd edn. (Addison-Wesley, Reading, Massachusetts).

127. Lindahl E, Hess B and van der Spoel D, 2001 Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7** 306–317, `http://www.gromacs.org`.

128. Berendsen H J C, van der Spoel D and van Drunen R, 1995 GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **91** 43–56.

129. van der Spoel D, Lindahl E, Hess B, van Buuren A R, Apol E, Meulenhoff P J, Tieleman D P, M.Sijbers A L T, Feenstra K A, van Drunen R and Berendsen H J C, 2004 *Gromacs User Manual version 3.2*, The Gromacs development team, `http://www.gromacs.org`.

130. Harding J, Heyes D M, Smith W, Howlin B, Anwar J, Wilding N, Refson K and Lindan P, 2002 The CCP5 molecular simulation summer school 2002, King's College, London.

131. Mackerell A D, 2004 Empirical force fields for biological macromolecules: Overview and issues. *J. Comp. Chem.* **25** 1584–1604.

132. Hermans J, Berendsen H J C, van Gunsteren W F and Postma J P M, 1984 A consistent empirical potential for water-protein interactions. *Biopolymers* **23** 1513–1518.

133. van Gunsteren W F, Billeter S R, Elsing A A, Hünenberger P H, Krüger P, Mark A E, Scott W R P and Tironi I G, 1996 *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Biomos and Hochschulverlag AG an der ETH Zürich, Zürich and Groningen.

134. Jorgensen W L, Maxwell D S and Tirado-Rives J, 1996 Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118** 11225–11236.

135. McDonald N A and Jorgensen W L, 1998 Development of an all-atom force field for heterocycles. properties of liquid pyrrole, furan, diazoles, and oxazoles. *J. Phys. Chem. B* **102** 8049–8059.

136. McDonald N A and Jorgensen W L, 1998 Development of an all-atom force field for heterocycles. properties of liquid pyridine and diazenes. *Theochem. J. Mol. Struc.* **424** 145–155.

137. Rizzo R C and Jorgensen W L, 1999 OPLS all-atom model for amines: Resolution of the amine hydration problem. *J. Am. Chem. Soc.* **121** 4827–4836.

138. Price M L P, Ostrovsky D and Jorgensen W L, 2001 Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field. *J. Comp. Chem.* **22** 1340–1352.

139. Watkins E K and Jorgensen W L, 2001 Perfluoroalkanes: Conformational analysis and liquid-state properties from ab initio and Monte Carlo calculations. *J. Phys. Chem. A* **105** 4118–4125.

140. Kaminski G A, Friesner R A, Tirado-Rives J and Jorgensen W L, 2001 Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105** 6474–6487.

141. Jorgensen W L and Jenson C, 1998 Temperature dependence of TIP3P, SPC, and TIP4P water from *NPT* Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comp. Chem.* **19** 1179–1186.

142. van der Spoel D and Lindahl E, 2003 Brute-force molecular dynamics simulations of Villin headpiece: Comparison with NMR parameters. *J. Phys. Chem. B* **107** 11178–11187.

143. Berendsen H J C, Postma J P M, van Gunsteren W F and Hermans J, 1981 Interaction models for water in relation to protein hydration, in *Intermolecular Forces* (edited by B Pullman), 331 (D. Reidel Publishing Company, Dordrecht, Holland).

144. Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W and Klein M L, 1983 Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79** 926–935.

145. Jorgensen W L and Madura J D, 1985 Temperature and size dependence for Monte-Carlo simulations of TIP4P water. *Mol. Phys.* **56** 1381–1392.

146. van der Spoel D, van Maaren P J and Berendsen H J C, 1998 A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **108** 10220–10230.

147. Mark P and Nilsson L, 2001 Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **105** 9954–9960.

148. Guillot B, 2002 A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liquids* **101** 219–260.

149. Åqvist J, 1990 Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* **94** 8021–8024.

150. Chandrasekhar J, Spellmeyer D C and Jorgensen W L, 1984 Energy component analysis for dilute aqueous solutions of Li^+, Na^+, F^-, and Cl^- ions. *J. Am. Chem. Soc.* **106** 903–910.

151. Marcus Y, 1988 Ionic radii in aqueous solutions. *Chem. Rev.* **88** 1475–1473.

152. Verlet L, 1967 Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **159** 98–103.

153. Frenkel D and Smit B, 2002 *Understanding Molecular Simulations*, 2nd edn. (Academic Press, San Diego).

154. Hockney R W, Goel S P and Eastwood J W, 1974 Quiet high-resolution computer models of a plasma. *J. Comp. Phys.* **14** 148–158.

155. Hess B, Bekker H, Berendsen H J C and Fraaije J G E M, 1997 LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18** 1463–1472.

156. Ryckaert J P, Ciccotti G and Berendsen H J C, 1977 Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys.* **23** 327–341.

157. Miyamoto S and Kollman P A, 1992 SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comp. Chem.* **13** 952–962.

158. Darden T A, 2001 Treatment of long-range forces and potential, in *Computational Biochemistry and Biophysics* (edited by O M Becker, J Alexander D. Mackerell, B Roux and M Watanabe), 91–114 (Marcel Dekker, Inc., New York).

159. Feller S E, Pastor R W, Rojnuckarin A, Bogusz S and Brooks B R, 1996 Effect of electrostatic force truncation on interfacial and transport properties of water. *J. Phys. Chem. A* **100** 17011–17020.

160. Ewald P P, 1921 Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. d. Phys.* **64** 253–287.

161. Deserno M and Holm C, 1998 How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.* **109** 7678–7693.

162. Press W H, Teukolsky S A, Vetterling W T and Flannery B P, 1992 *Numerical Recipes in C*, 2nd edn. (Cambridge University Press, Cambridge), `http://www.library.cornell.edu/nr/bookcpdf.html`.

163. Hockney R W and Eastwood J W, 1988 *Computer Simulation Using Particles* (Institute of Physics, Bristol).

164. Darden T, York D and Pedersen L, 1993 Particle mesh Ewald – an $N\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98** 10089–10092.

165. Essman U, Perela L, Berkowitz M L, Darden T, Lee H and Pedersen L G, 1995 A smooth particle mesh Ewald method. *J. Chem. Phys.* **103** 8577–8592.

166. Berendsen H J C, Postma J P M, DiNola A and Haak J R, 1984 Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81** 3684–3690.

167. Harvey S C, Tan R K Z and Cheatham T E, 1998 The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comp. Chem.* **19** 726–740.

168. Kirkwood J G, 1935 Statistical mechanism of fluid mixtures. *J. Chem. Phys.* **3** 300–313.

169. Roux B, Allen T, Bernèche S and Im W, 2004 Theoretical and computational models of biological ion channels. *Quart. Rev. Biophys.* **37** 15–103.

170. Allen T W, Andersen O S and Roux B, 2004 Energetics of ion conduction through the gramicidin channel. *Proc. Natl. Acad. Sci. U.S.A.* **101** 117–122.

171. Smith G R and Sansom M S, 2002 Free energy of a potassium ion in a model of the channel formed by an amphipathic leucine-serine peptide. *Eur. Biophys. J.* **31** 198–206.

172. Valleau J P and Torrie G M, 1977 A guide to Monte Carlo for statistical mechanics: 2. Byways, in *Statistical Mechanics. Part A: Equilibrium Techniques* (edited by B J Berne), vol. 5 of *Modern Theoretical Chemistry*, chap. 5, 169–194 (Plenum Press, New York).

173. Torrie G M and Valleau J P, 1977 Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comp. Phys.* **23** 187–199.

174. Roux B, 1995 The calculation of the potential of mean force using computer simulations. *Comp. Phys. Comm.* **91** 275–282.

175. Kumar S, Bouzida D, Swendsen R H, Kollman P A and Rosenberg J M, 1992 The Weighted Histogram Analysis Method for free-energy calculations on biomolecules. I. The method. *J. Comp. Chem.* **13** 1011–1021.

176. Roux B and Karplus M, 1991 Ion transport in a model gramicidin channel: Structure and thermodynamics. *Biophys. J.* **59** 961–981.

177. Debye P and Hückel E, 1923 Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen. *Phys. Z.* **24** 185–206.

178. Honig B and Nicholls A, 1995 Classical electrostatics in biology and chemistry. *Science* **268** 1144–1149.

179. Baker N A, Sept D, Joseph S, Holst M J and McCammon J A, 2001 Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98** 10037–10041, `http://agave.wustl.edu/apbs/`.

180. Sharp K A and Honig B, 1990 Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **94** 7684–7692.

181. Rashin A A and Honig B, 1985 Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **89** 5588–5593.

182. Holst M, Baker N and Wang F, 2000 Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comp. Chem.* **21** 1319–1342.

183. Baker N, Holst M and Wang F, 2000 Adaptive multilevel finite element solution of the Poisson-Boltzmann equation II. Refinement at solvent-accessible surfaces in biomolecular systems. *J. Comp. Chem.* **21** 1343–1352.

184. Nielsen J E and McCammon J A, 2003 Calculating pKa values in enzyme active sites. *Protein Sci.* **12** 1894–1901, `http://enzyme.ucd.ie/Science/pKa/`.

185. Nielsen J E and Vriend G, 2001 Optimizing the hydrogen-bond network in Poisson-Boltzmann equation-based pKa calculations. *Proteins: Struct. Funct. Genet.* **43** 403–412.

186. Rocchia W, Alexov E and Honig B, 2001 Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B* **105** 6507–6514, `http://trantor.bioc.columbia.edu/delphi/`.

187. Yang B, van Hoek A N and Verkman A S, 1997 Very high single channel water permeability of aquaporin-4 in baculovirus-infected insect cells and liposomes reconstituted with purified aquaporin-4. *Biochemistry* **36** 7625–7632.

188. Pohl P, Saparov S M, Borgnia M J and Agre P, 2001 Highly selective water channel activity measured by voltage clamp: analysis of planar lipid bilayers reconstituted with purified AqpZ. *Proc. Natl. Acad. Sci. U.S.A.* **98** 9624–9629.

189. Pohl P and Saparov S M, 2000 Solvent drag across gramicidin channels demonstrated by microelectrodes. *Biophys. J.* **78** 2426–2434.

190. Saparov S M, Antonenko Y N, and Pohl P, 2000 Desformylgramicidin: a model channel with an extremely high water permeability. *Biophys. J.* **79** 2526–2534.

191. de Groot B L, Tieleman D P, Pohl P and Grubmüller H, 2002 Water permeation through gramicidin A: desformylation and the double helix: a molecular dynamics study. *Biophys. J.* **82** 2934–2942.

192. de Groot B L and Grubmüller H, 2001 Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science* **294** 2353–2357.

193. Christenson H K, 2001 Confinement effects on freezing and melting. *J. Phys.: Condens. Matter* **13** R95–R133.

194. Gelb L D, Gubbins K E, Radhakrishnan R and Sliwinska-Bartkowiak M, 1999 Phase separation in confined systems. *Rep. Prog. Phys.* **62** 1573–1659.

195. Lynden-Bell R M and Rasaiah J C, 1996 Mobility and solvation of ions in channels. *J. Chem. Phys.* **105** 9266–9280.

196. Allen T W, Kuyucak S and Chung S H, 1999 The effect of hydrophobic and hydrophilic channel walls on the structure and diffusion of water and ions. *J. Chem. Phys.* **111** 7985–7999.

197. Hummer G, Rasaiah J C and Noworyta J P, 2001 Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* **414** 188–190.

198. Allen R, Melchionna S and Hansen J P, 2002 Intermittent permeation of cylindrical nanopores by water. *Phys. Rev. Lett.* **89** 175502.

199. Brovchenko I and Geiger A, 2002 Water in nanopores in equilibrium with a bulk reservoir—Gibbs ensemble Monte Carlo simulations. *J. Mol. Liquids* **96–97** 195–206.

200. Ashcroft N W and Mermin N D, 1976 *Solid State Physics* (Saunders College Publishing, Fort Worth).

201. Domene C, Grottesi A and Sansom M S P, 2004 Filter flexibility and distortion in a bacterial inward rectifier K^+ channel: Simulation studies of KirBac1.1. *Biophys. J.* **87** 256–267.

202. Allen R, Hansen J P and Melchionna S, 2003 Molecular dynamics investigation of water permeation through nanopores. *J. Chem. Phys.* **119** 3905–3919.

203. Hunter III J E and Reinhardt W P, 1995 Finite-size scaling behavior of the free energy barrier between coexisting phases: Determination of the critical temperature and interfacial tension of the Lennard-Jones fluid. *J. Chem. Phys.* **103** 8627–8637.

204. Sakmann B and Neher E, eds., 1983 *Single-Channel Recordings* (Plenum Press, New York).

205. Flyvbjerg H and Petersen H G, 1989 Error estimates on averages of correlated data. *J. Chem. Phys.* **91** 461–466.

206. Hess B, 1999 *Stochastic concepts in molecular simulation*, Ph.D. thesis, Rijksuniversiteit Groningen, Groningen, `http://www.ub.rug.nl/eldoc/dis/science/b.hess/`.

207. Evans R, 1990 Fluids adsorbed in narrow pores: phase equilibria and structure. *J. Phys.: Condens. Matter* **2** 8989–9007.

208. Lee C Y, McCammon J A and Rossky P J, 1984 The structure of liquid water at an extended hydrophobic surface. *J. Chem. Phys.* **80** 4448–4455, `http://link.aip.org/link/?JCP/80/4448/1`.

209. Lee S H and Rossky P J, 1994 A comparison of the structure and dynamics of liquid water at hydrophobic and hydrophilic surfaces—a molecular dynamics simulation study. *J. Chem. Phys.* **100** 3334–3345.

210. Spohr E, Trokhymchuk A and Henderson D, 1998 Adsorption of water molecules in slit pores. *J. Electroanalyt. Chem.* **450** 281–287.

211. Gallo P, Rovere M and Spohr E, 2000 Glass transition and layering effects in confined water: A computer simulation study. *J. Chem. Phys.* **113** 11324–11335.

212. Freund J B, 2002 Electro-osmosis in a nanometer-scale channel studied by atomistic simulation. *J. Chem. Phys.* **116** 2194–2200.

213. Striolo A, Chiavolo A A, Cummings P T and Gubbins K E, 2003 Water adsorption in carbon-slit nanopores. *Langmuir* **19** 8583–8591.

214. Kalra A, Garde S and Hummer G, 2003 Osmotic water transport through carbon nanotube membranes. *Proc. Natl. Acad. Sci. U.S.A.* **100** 10175–10180.

215. Jensen M Ø, Mouritsen O G and Peters G H, 2004 The hydrophobic effect: Molecular dynamics simulations of water confined between extended hydrophobic and hydrophilic surfaces. *J. Chem. Phys.* **120** 9729–9744.

216. Mamatkulov S I, Khabibullaev P K and Netz R R, 2004 Water at hydrophobic substrates: Curvature, pressure, and temperature effects. *Langmuir* **20** 4756–4763.

217. Klein J and Kurnacheva E, 1998 Simple liquids confined to molecularly thin layers. I Confinement-induced liquid-to-solid phase transitions. *J. Chem. Phys.* **108** 6996–7009.

218. Raviv U, Laurat P and Klein J, 2001 Fluidity of water confined to subnanometre films. *Nature* **413** 51–54.

219. Dore J, 2000 Structural studies of water in confined geometry by neutron diffraction. *Chem. Phys.* **258** 327–347.

220. Wallqvist A, Gallicchio E and Levy R M, 2001 A model for studying drying at hydrophobic interfaces: Structural and thermodynamic properties. *J. Phys. Chem. B* **105** 6745–6753.

221. Aarts D G A L, Schmidt M and Lekkerkerker H N W, 2004 Direct visual observation of thermal capillary waves. *Science* **304** 847–850.

222. Mecke K, 2004 Mikroskopische Kapillarwellen sichtbar gemacht. *Physik Journal* **3** 19–20, see also Aarts et al.[221].

223. Preusser A, 1989 Algorithm 671—FARB-E-2D: Fill Area with Bicubics on Rectangles—A contour plot program. *ACM Trans. Math. Softw.* **15** 79–89, `http://www.fhi-berlin.mpg.de/grz/pub/xfarbe/`.

224. Brovchenko I, Geiger A and Oleinikova A, 2001 Phase equilibria of water in cylindrical nanopores. *Phys. Chem. Chem. Phys.* **3** 1567–1569.

225. Hummer G, Garde S, García A E, Pohorille A and Pratt L R, 1996 An information theory model of hydrophobic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **93** 8951–8955.

226. Waghe A, Rasaiah J C and Hummer G, 2002 Filling and emptying kinetics of carbon nanotubes in water. *J. Chem. Phys.* **117** 10789–10795.

227. Privman V and Fisher M E, 1983 Finite-size effects at first-order transitions. *J. Stat. Phys.* **33** 385–417.

228. Peterson B K, Gubbins K E, Heffelfinger G S, Marconi U M B and van Smol F, 1988 Lennard-Jones fluids in cylindrical pores: Nonlocal theory and computer simulation. *J. Chem. Phys.* **88** 6487–6500.

229. Martí J and Gordillo M C, 2001 Temperature effects on the static and thermodynamic properties of liquid water inside nanotubes. *Phys. Rev. E* **64** 021504.

230. Berezhkovskii A and Hummer G, 2002 Single-file transport of water molecules through a carbon nanotube. *Phys. Rev. Lett.* **89** 065403.

231. Rowlinson J S and Widom B, 1982 *Molecular Theory of Capillarity*, no. 8 in The International series of monographs on chemistry (Clarendon Press, Oxford).

232. Giaya A and Thompson R W, 2002 Water confined in cylindrical micropores. *J. Chem. Phys.* **117** 3464–3475.

233. Dujardin E, Ebbesen T W, Hiura H and Tanigaki K, 1994 Capillarity and wetting of carbon nanotubes. *Science* **265** 1850–1852.

234. Bekyarova E, Hanzawa Y, Kaneko K, Silvestre-Albero J, Sepulveda-Escribano A, Rodriguez-Reinoso F, Kasuya D, Yudasaka M and Iijima S, 2002 Cluster-mediated filling of water vapor in intratube and interstitial nanospaces of single-wall carbon nanohorns. *Chem. Phys. Lett.* **366** 463–468.

235. Naguib N, Ye H, Gogotsi Y, Yazicioglu A G, Megaridis C M and Yoshimura M, 2004 Observation of water confined in nanometer channels of closed carbon nanotubes. *Nano Lett.* **4** 2237–2243.

236. Allen R, Melchionna S and Hansen J P, 2003 Permeation of nanopores by water: the effects of channel polarization. *J. Phys.: Condens. Matter* **15** S297–S302.

237. Jiang Y, Lee A, Chen J, Ruta V, Cadene M, Chait B T and MacKinnon R, 2003 X-ray structure of a voltage-dependent K^+ channel. *Nature* **423** 33–41.

238. Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, Heymann J B, Engel A and Fujiyoshi Y, 2000 Structural determinants of water permeation through aquaporin-1. *Nature* **407** 599–605.

239. Lyubartsev A P, Førrisdahl O K and Laaksonen A, 1998 Solvation free energies of methane and alkali halide ion pairs: An expanded ensemble molecular dynamics simulation study. *J. Chem. Phys.* **108** 227–233.

240. Bernèche S and Roux B, 2001 Energetics of ion conduction through the K^+ channel. *Nature* **414** 73–77.

241. Noskov S Y, Bernèche S and Roux B, 2004 Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* **431** 830–834.

242. Crozier P S, Henderson D, Rowley R L and Busath D D, 2001 Model channel ion currents in NaCl-Extended Simple Point Charge water solution with applied-field molecular dynamics. *Biophys. J.* **81** 3077–3089.

243. Eisenberg D and Crothers D, 1979 *Physical Chemistry with Applications to the Life Sciences* (The Benjamin/Cummings Publishing Company, Menlo Park, California).

244. Wu Y, Sugimura H, Inoue Y and Takai O, 2002 Thin films with nanotextures for transparent and ultra water-repellent coatings produced from trimethylmethoxysilane by microwave plasma CVD. *Chem. Vapor Depos.* **8** 47–50.

245. Whitesides G M and Laibinis P E, 1990 Wet chemical approaches to the characterization of organic surfaces: Self-assembled monolayers, wetting, and the physical-organic chemistry of the solid-liquid interface. *Langmuir* **6** 87–96.

246. Hautman J and Klein M L, 1991 Microscopic wetting phenomena. *Phys. Rev. Lett.* **67** 1763–1766.

247. Lundgren M, Allan N L, Cosgrove T and George N, 2002 Wetting of water and water/ethanol droplets on a non-polar surface: A molecular dynamics study. *Langmuir* **18** 10462–10466.

248. Mashl R J, Joseph S, Aluru N R and Jakobsson E, 2003 Anomalously immobilized water: A new water phase induced by confinement in nanotubes. *Nano Lett.* **3** 589–592.

249. Williams R, 1975 Interfacial free energies between polymers and aqueous electrolyte solutions. *J. Phys. Chem.* **79** 1274–1276.

250. Voets T, Droogmans G, Wissenbach U, Janssens A, Flockerzi V and Nilius B, 2004 The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels. *Nature* **430** 748–754.

251. Beckstein O, Biggin P C, Bond P, Bright J N, Domene C, Grottesi A, Holyoake J and Sansom M S P, 2003 Ion channel gating: insights via molecular simulations. *FEBS Lett.* **555** 85–90, http://dx.doi.org/10.1016/S0014-5793(03)01151-7.

252. Roosild T P, Miller S, Booth I R and Choe S, 2002 A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell* **109** 781–791.

253. McElroy C, Manfredo A, Wendt A, Gollnick P and Foster M, 2002 TROSY-NMR studies of the 91 kDa TRAP protein reveal allosteric control of a gene regulatory protein by ligand-altered flexibility. *J. Mol. Biol.* **323** 463–473.

254. Nicholls A, Sharp K A and Honig B, 1991 Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11** 281–296.

255. Maibaum L and Chandler D, 2003 A coarse-grained model of water confined in a hydrophobic tube. *J. Phys. Chem. B* **107** 1189–1193.

256. Harms G S, Orr G, Montal M, Thrall B D, Colson S D and Lu H P, 2003 Probing conformational changes of gramicidin ion channels by single-molecule patch-clamp fluorescence microscopy. *Biophys. J.* **85** 1826–1838.

257. Goh C S, Milburn D and Gerstein M, 2004 Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* **14** 104–109.

258. Pang A, Arinaminpathy Y, Sansom M S P, and Biggin P C, 2003 Interdomain dynamics and ligand binding: Molecular dynamics simulations of glutamine binding protein. *FEBS Lett.* **550** 168–174.

259. Hirama Y, Takahashi T, Hino M and Sato T, 1996 Studies of water adsorbed in porous vycor glass. *J. Colloid Interface Sci.* **184** 349–359.

260. Koga K, Gao G, Tanaka H and Zeng X C, 2001 Formation of ordered ice nanotubes inside carbon nanotubes. *Nature* **412** 802–805.

261. Levinger N E, 2002 Water in confinement. *Science* **298** 1722–1723.

262. Weetman P, Goldman S and Gray C G, 1997 Use of the Poisson Boltzmann equation to estimate the electrostatic free energy barrier for dielectric models of biological ion channels. *J. Phys. Chem. B* **101** 6073–6078.

263. Corry B, Kuyucak S and Chung S H, 2000 Invalidity of continuum theories of electrolytes in nanopores. *Chem. Phys. Lett.* **320** 35–41.

264. Edwards S, Corry B, Kuyucak S and Chung S H, 2002 Continuum electrostatics fails to describe ion permeation in the gramicidin channel. *Biophys. J.* **83** 1348–1360.

265. Miyazawa A, Fujiyoshi Y, Stowell M and Unwin N, 1999 Nicotinic acetylcholine receptor at 4.6 Å resolution: Transverse tunnels in the channel wall. *J. Mol. Biol.* **288** 765–786.

266. Wagner K, Edson K, Heginbotham L, Post M, Huganir R L and Czernik A J, 1991 Determination of the tyrosine phosphorylation sites of the nicotinic acetylcholine receptor. *J. Biol. Chem.* **266** 23784–23789.

267. Mulder N J, Apweiler R, Attwood T K, Bairoch A, Barrell D, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley R R, Courcelle E, Das U, Durbin R, Falquet L, Fleischmann W, Griffiths-Jones S, Haft D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Lonsdale D, Silventoinen V, Orchard S E, Pagni M, Peyruc D, Ponting C P, Selengut J D, Servant F, Sigrist C J, Vaughan R and Zdobnov E M, 2003 The InterPro database, 2003 brings increased coverage and new features. *Nucleic Acids Res.* **31** 315–318, `http://www.ebi.ac.uk/interpro/`.

268. Kofuji P, Wang J B, Moss S J, Huganir R L and Burt D R, 1991 Generation of two forms of the gamma-aminobutyric acid A receptor gamma 2-subunit in mice by alternative splicing. *J. Neurochem.* **56** 713–715.

269. Sine S M, 2002 The nicotinic receptor ligand binding domain. *J. Neurobiol.* **53** 431–446.

270. Bateman A, Coin L, Durbin R, Finn R D, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer E L, Studholme D J, Yeats C and Eddy S R, 2004 The Pfam protein families database. *Nucleic Acids Res.* **32** D138–D141, `http://www.sanger.ac.uk/Software/Pfam/`.

271. Thompson J D, Higgins D G and Gibson T J, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22** 4673–4780.

272. Clamp M, Cuff J, Searle S M and Barton G J, 2004 The Jalview Java alignment editor. *Bioinformatics* **20** 426–427, `http://www.jalview.org`.

273. Schneider T D and Stephens R M, 1990 Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18** 6097–6100, `http://www.ccrnp.ncifcrf.gov/~toms/logoprograms.html`.

274. Lester H A, 1992 The permeation pathway of neurotransmitter-gated ion channels. *Annu. Rev. Biophys. Biomol. Struct.* **21** 267–292.

275. Karlin A, 2002 Emerging structure of the nicotinic acetylcholine receptors. *Nature Rev. Neurosci.* **3** 102–114.

276. Shannon C E, 1948 A mathematical theory of communication. *Bell System Technical Journal* **27** 379–423, 623–656, `http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html`, reprinted with corrections.

277. Crooks G E, Hon G, Chandonia J M and Brenner S E, 2004 WebLogo: a sequence logo generator. *Genome Res.* **14** 1188–1190, `http://weblogo.berkeley.edu/`.

278. Unwin N, 1995 Acetylcholine-receptor channel imaged in the open state. *Nature* **373** 37–43, `http://dx.doi.org/10.1006/jmbi.1993.1107`.

279. Brejc K, van Dijk W J, Klaassen R V, Schuurmans M, van der Oost J, Smit A B and Sixma T K, 2001 Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411** 269–276.

280. Bouzat C, Gumilar F, Spitzmaul G, Wang H L, Rayes D, Hansen S B and Sine P T S M, 2004 Coupling of agonist binding to channel gating in an ACh-binding protein linked to an ion channel. *Nature* **430** 896–900.

281. Wilson G G and Karlin A, 2001 Acetylcholine receptor channel structure in the resting, open, and desensitized states probed with the substituted-cysteine-accessibility method. *Proc. Natl. Acad. Sci. U.S.A.* **98** 1241–1248.

282. Wilson G G and Karlin A, 1998 The location of the gate in the acetylcholine receptor channel. *Neuron* **20** 1269–1281.

283. Labarca C, Nowak M W, Zhang H Y, Tang L X, Deshpande P and Lester H A, 1995 Channel gating governed symmetrically by conserved leucine residues in the M2 domain of nicotinic receptors. *Nature* **376** 514–516.

284. Grosman C, Zhou M and Auerbach A, 2000 Mapping the conformational wave of acetylcholine receptor channel gating. *Nature* **403** 773–776.

285. Cymes G D, Grosman C and Auerbach A, 2002 Structure of the transition state of gating in the acetylcholine receptor channel pore: A Φ-value analysis. *Biochem.* **41** 5548–5555.

286. Panicker S, Cruz H, Arrabit C and Slesinger P A, 2002 Evidence for a centrally located gate in the pore of a serotonin-gated ion channel. *J. Neurosci.* **22** 1629–1639.

287. Panicker S, Cruz H, Arrabit C, Suen K F and Slesinger P A, 2004 Minimal structural rearrangement of the cytoplasmic pore during activation of the 5-HT_{3A} receptor. *J. Biol. Chem.* **279** 28149–28158.

288. Unwin N, 1996 Projection structure of the nicotinic acetylcholine receptor: Distinct conformations of the α subunits. *J. Mol. Biol.* **257** 586–596.

289. Kelley S P, Dunlop J I, Kirkness E F, Lambert J J and Peters J A, 2003 A cytoplasmic region determines single-channel conductance in 5-HT_3 receptors. *Nature* **424** 321–324.

290. Beitz E, 2000 TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. *Bioinformatics* **16** 135–139, `http://homepages.uni-tuebingen.de/beitz/txe.html`, TeXshade version 1.5a of 01/03/08.

291. White S H and Wimley W C, 1999 Membrane protein folding and stability: Physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28** 319–365.

292. Vriend G, 1990 WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8** 52–56, `http://www.cmbi.kun.nl/whatif/`.

293. Laskowski R A, MacArthur M W, Moss D S and Thornton J M, 1993 PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26** 283–291.

294. Opella S J, Marassi F M, Gesell J J, Valente A P, Kim Y, Oblatt-Montal M and Montal M, 1999 Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nature Struct. Biol.* **6** 374–379.

295. Law R J, Forrest L R, Ranatunga K M, La Rocca P, Tieleman D P and Sansom M S P, 2000 Structure and dynamics of the pore-ligning helix of the nicotinic receptor: MD simulations in water, lipid bilayers, and transbilayer bundles. *Proteins: Struct. Funct. Genet.* **39** 47–55.

296. Law R, Tieleman D P and Sansom M S P, 2003 Pores formed by the nicotinic receptor M2δ peptide: A molecular dynamics simulation study. *Biophys. J.* **84** 14–27.

297. Dolinsky T J, Nielsen J E, McCammon J A and A. B N, 2004 PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32** W665–W667, `http://agave.wustl.edu/pdb2pqr/`.

298. Allen T W, Başţuğ T, Kuyucak S and Chung S H, 2003 Gramicidin A channel as a test ground for molecular dynamics force fields. *Biophys. J.* **84** 2159–2168.

299. Katz B and Thesleff S, 1957 A study of 'desensitisation' produced by acetylcholine at the motor end-plate. *J. Physiol.* **138** 63–80.

300. Labarca C, Schwarz J, Deshpande P, Schwarz S, Nowak M W, Fonck C, Nashmi R, Kofuji P, Dang H, Shi W, Fidan M, Khakh B S, Chen Z, Bowers B J, Boulter J, Wehner J M and Lester H A, 2001 Point mutant mice with hypersensitive α4 nicotinic receptors show dopaminergic deficits and increased anxiety. *Proc. Natl. Acad. Sci. U.S.A.* **98** 2786–2791.

301. Tapper A R, McKinney S L, Nashmi R, Schwarz J, Deshpande P, Labarca C, Whiteaker P, Marks M J, Collins A C and Lester H A, 2004 Nicotine activation of $\alpha 4^*$ receptors: Sufficient for reward, tolerance, and sensitization. *Science* **306** 1029–1032.

302. Bostick D L and Berkowitz M L, 2004 Exterior site occupancy infers chloride-induced proton gating in a prokaryotic homolog of the ClC chloride channel. *Biophys. J.* **87** 1686–1696.

303. Cohen J and Schulten K, 2004 Mechanism of anionic conduction across ClC. *Biophys. J.* **86** 836–845.

304. Corry B, O'Mara M and Chung S H, 2004 Conduction mechanisms of chloride ions in ClC-type channels. *Biophys. J.* **86** 846–860.

305. Faraldo-Gómez J D and Roux B, 2004 Electrostatics of ion stabilization in a ClC chloride channel homologue from *Escherichia coli*. *J. Mol. Biol.* **339** 981–1000.

306. Miloshevsky G V and Jordan P C, 2004 Anion pathway and potential energy profiles along curvilinear bacterial ClC Cl^- pores: electrostatic effects of charged residues. *Biophys. J.* **86** 825–835, `http://www.biophysj.org/cgi/content/full/86/2/825`.

307. Yin J, Kuang Z, Mahankali U and Beck T L, 2004 Ion transit pathways and gating in ClC chloride channels. *Proteins: Struct. Funct. Genet.* **57** 414–421.

308. Saparov S M and Pohl P, 2004 Beyond the diffusion limit: Water flow through the empty bacterial potassium channel. *Proc. Natl. Acad. Sci. U.S.A.* **101** 4805–4809.

309. Kim J and Raushel F M, 2004 Access to the carbamate tunnel of carbamoyl phosphate synthetase. *Arch. Biochem. Biophys* **425** 33–41.

310. Harris R A, 1999 Ethanol actions on multiple ion channels: which are important? *Alcoholism Clin. Exp. Res.* **23** 1563–1570.

311. Davies D L, Trudell J R, Mihic S J, Crawford D K and Alkana R L, 2003 Ethanol potentiation of glycine receptors expressed in *Xenopus* oocytes antagonized by increased atmospheric pressure. *Alcoholism Clin. Exp. Res.* **27** 743–755.

312. Davies D L, Crawford D K, Trudell J R, Mihicà S J and Alkana R L, 2004 Multiple sites of ethanol action in α1 and α2 glycine receptors suggested by sensitivity to pressure antagonism. *J. Neurochem.* **89** 1175–1185.

313. "Sullivan, Louis", Encyclopædia Britannica, `http://search.eb.com/eb/article?tocId=9070265`, retrieved 2004-10-29 from Encyclopædia Britannica Online.

314. Sullivan L H, 1896 The tall office building artistically considered. *Lippincott's Magazine* `http://www.njit.edu/old/Library/archlib/pub-domain/sullivan-1896-tall-bldg.htm`, accessed online 2004-10-29.

315. Baranyai A, 2002 Direct estimation of the partition function from computer simulations. *Phys. Rev. E* **65** 026110.

316. Weisstein E W, Viviani's curve, From *MathWorld*—A Wolfram Web Resource., `http://mathworld.wolfram.com/VivianisCurve.html`, accessed 2004-10-19.

317. Wolfram Research, Inc., 2003 *Mathematica* (Wolfram Research, Inc., Champaign, Illinois), version 5.0.

318. Good R H, 2001 Elliptic integrals, the forgotten functions. *Eur. J. Phys.* **22** 119–126.

319. Free Software Foundation, 1991, GNU General Public License, `http://www.gnu.org/copyleft/gpl.html`.

320. Grubmüller H, 2003 What happens if the room at the bottom runs out? A close look at small water pores. *Proc. Natl. Acad. Sci. U.S.A.* **100** 7421–7422.

321. Sansom M S P, Bond P, Beckstein O, Biggin P C, Faraldo-Goméz J, Law R J, Patargias G and Tieleman D P, 2002 Water in ion channels and pores—simulation studies, in *Ion channels: from atomic resolution physiology to functional genomics* (edited by G Bock and J A Goode), vol. 245 of *Novartis Foundation Symposia*, 66–78, Novartis Foundation, London (John Wiley & Sons, Chicester), `http://doi.wiley.com/10.1002/0470868759.ch6`.

Index