Using Simulation to Probe Physiology and Pharmacology of Ion Channels

Bjarne Feddersen

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Declaration

I declare that this thesis is entirely my own work and describes my own research unless otherwise stated.

Bjarne Feddersen

Abstract

In ion channels, fenestrations are lateral openings that connect the membrane to the protein's ion-conducting pore. They are well known to act as access pathways for lipophilic drugs in voltage-gated sodium channels and have been implicated in inheritable diseases. Despite their clear functional importance in many ion channel families, no such fenestrations have previously been reported in the superfamily of pentameric ligand-gated ion channels (pLGICs). This channel family has been the subject of in-depth study for decades, but many questions on the precise details of channel gating and pharmacology still remain. The glycine receptor (GlyR) is an inhibitory pLGIC that is responsible for regulation of motor control in the spinal cord. Its activation and drug interactions are the focus of this thesis.

Here, I developed an improved umbrella sampling workflow aimed at improving convergence speeds of potentials of mean force (PMF). I benchmarked this workflow on test systems of drug molecules permeating through lipid bilayers of different lipid compositions. I found that my improved workflow outperforms common practices in umbrella sampling, and also quantified errors arising from boundary conditions commonly applied to the calculation of PMFs.

The main finding of this DPhil is the identification of a state-dependent, drug-permeable fenestration in GlyR. I analysed 55 published GlyR structures in different functional states for potential fenestrations and consistently detected a state-dependent fenestration near an anaesthetics binding site at each subunit interface. Observed to be almost exclusively occluded in the closed state, the fenestration widens considerably in open state structures, exhibiting a similar 'bottleneck' radius across structures. I then employed my improved umbrella sampling workflow to quantify the translocation energetics of fipronil. Comparison of the resulting potential of mean force (PMF) profiles shows that fenestration-based entry of fipronil is energetically favourable to membrane permeation. My results thus provide the first characterisation of a drug-permeable fenestration in GlyR.

Additionally, through investigations of precisely how binding of the positive allosteric modulator ivermectin destabilises the orthosteric inhibitor strychnine, allosteric networks in GlyR that tie conserved ligand binding loops to the transmembrane domain were described. Finally, intermediate activation states of the *Gloeobacter* ligand-gated ion channel (GLIC) were probed in computational electrophysiology. We demonstrate that intermediate, asymmetric activation states of the channel already display partial conductivity.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof Phil Biggin for his scientific mentorship and advice. My thesis would not have been possible without his support and expertise. I have to thank Phil for giving me the flexibility to change projects when my original work was not panning out, and I am hugely grateful for his support when I lost an entire year to illness.

In this context, I also want to thank Louise Samson and Prof Matt Higgins, as well as the Wellcome Trust, for finding ways to extend my funding when I needed it so that I could actually finish this degree. My period of illness was stressful enough without funding-related worries, and it was a relief that I was able to take my time to recover.

Throughout my work here, I learned so much from Dr Afroditi Maria Zaki, Dr Irfan Alibay, and Asst Prof Keith Cassidy. Thank you all for taking this wet-lab biochemist and turning him into a halfway passable computational biophysicist. In addition, I want to thank all members of SBCB for forming such a wonderful, welcoming, and caring community, particularly Johnny, Daniel, Charlie, Simon, Will, Lauren, Matt, Franco, Mina, Cyril, Bertie, and Astrid.

Early in my DPhil, I participated in the Google Summer of Code programme. Thank you to my mentors, Hugo MacDermott-Opeskin, Irfan Alibay, Oliver Beckstein, and Fiona Naughton. As a self-taught programmer, I benefited tremendously from learning from you. Whoever has to pick up my scripts in the future should thank you, as well.

Thank you to the Biochemistry IT team, in particular Borut Rozman, for being an actual IT wizard and keeping our clusters in excellent shape.

Huge thanks to my friends who kept me sane throughout these (sometimes very long-feeling) five years. Charlie, Will, Rob, and Danny, I'm so sorry you had to live with me for so long. I hope we can stay friends. Thank you, Molly, Lauren, Anna, Daniel, and Mark. Thanks to my friends from home, particularly Nils, Ines, Yannic, Bene, Franzi, Feli, Dennis, and Niklas, and thanks to my St Andrews friends who made me feel so welcome in the UK I decided to stay for an entire PhD. Iain, Haley, Marta, Montse, and Zac, we may all have moved away, but Farmers we shall remain.

Claire, I really could not have done this without you. Thank you so, so much for your unwavering support. I might not have eaten or rested as the deadline

approached had it not been for you. You are my rock and you made the final few months tolerable, and I'm sorry I repaid you by making you proofread my thesis. Thank you for always being in my corner, regardless.

Schließlich geht großer Dank an meine Familie. Oma, Oma und Opa, Svenja und Lars (und Fiete, der diese Arbeit ja vielleicht auch eines Tages liest), und vor allem Mama und Papa, ohne eure jahrzehntelange Unterstützung hätte ich es nicht geschafft, euch jetzt meine Doktorarbeit in die Hände drücken zu können. Tausend Dank für eure bedingungslose und uneingeschränkte Unterstützung, den moralischen Beistand, und die Zuflucht.

Contents

List of Figures

List of Tables

List of Abbreviations

CHAP . . . . . Channel Annotation Package

Chol . . . . . Cholesterol

CPU . . . . . Central processing unit

cryo-EM . . . . cryogenic electron microscopy

DOPC 1,2-dioleoyl-sn-glycero-3-phosphocholine

DOPE 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine

ECD . . . . . Extracellular domain

FFT . . . . . Fast Fourier Transform

GABA . . . . . Gamma-amino butyric acid

GLIC . . . . . *Gloeobacter* ligand-gated ion channel

GlyR . . . . . Glycine receptor

GPU . . . . . Graphics processing unit

ICD . . . . . Intracellular domain

LJ . . . . . Lennard-Jones

MBAR . . . . . Multistate Bennet Acceptance Ratio

MD . . . . . Molecular dynamics

Na_v . . . . . Voltage-gated sodium channel

PBC . . . . . Periodic boundary conditions

PDB . . . . . Protein Data Bank

pLGIC . . . . . Pentameric ligand-gated ion channel

PME . . . . . Particle mesh Ewald

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

POPE 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

POPS 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine

QM . . . . . Quantum mechanics

RMSD . . . . . Root mean square deviation

RMSF . . . . . Root mean square fluctuation

SSRI . . . . . Selective serotonin reuptake inhibitor

STeUS . . . . . Simulated tempering-enhanced umbrella sampling

TMD . . . . . Transmembrane domain

US . . . . . Umbrella sampling

WHAM . . . . . Weighted Histogram Analysis Method

WL . . . . . Wang-Landau algorithm

. . . . . .

1 Introduction

Contents

Life is sustained chemical nonequilibrium. The state of equilibrium is only reached in death[1,2]. This sustained nonequilibrium is what allows the work that makes life possible to be carried out. Ion channels are one of the strategies that arose in evolution to harness it[3,4]. Biological membranes are the focus points of electrochemical gradients that arise from imbalances of ion concentrations on either side. Cells expend metabolic energy (in the form of adenosine triphosphate, ATP) maintaining these gradients[5]. Ion channels can then open as a reaction

to, for example, ligand binding[6], membrane deformation[7], or changes in the membrane potential[8], to allow ions to flow down their electrochemical gradient. They fulfil a variety of roles in the cell, including signal transduction[9,10], cell volume regulation[11,12], and pH control[13,14]. One aspect that makes ion channels such crucial components is the extremely high speeds of signal transduction they allow—ion channel gating occurs on millisecond timescales, while phosphorylation cascades can take minutes[15] and nuclear receptors with genomic responses often operate on timescales of hours[16,17]. In this thesis, pentameric ligand-gated ion channels (pLGICs) were studied with computational methods. A particular focus lay on the glycine receptor (GlyR).

1.1 Pentameric ligand-gated ion channels

The family of pentameric ligand-gated ion channels (pLGICs) is a group of proteinaceous ion channels that mediate fast synaptic transmission[18]. In reference to a conserved disulphide bond in the extracellular domain, it is also referred to as the Cys-loop receptor family[18]. This family is a major component of the synaptic integration machinery[19]. It has excitatory and inhibitory members, and the combined sum of their effects—plus the effects of other ion channels—determines whether sufficient depolarisation of the membrane potential occurs to trigger an action potential. Its members differ by ligand sensitivity and ion selectivity, with acetylcholine-gated and serotonin-gated channels displaying cation selectivity and gamma-amino butyric acid (GABA) and glycine receptors being selective for anions[20–22]. The structures of these four neurotransmitters are shown in Fig. 1.1. In pLGICs, selectivity for anions or cations is primarily determined through electrostatic interactions between the permeating ions and ionisable residues lining the channel pore, particularly at the intracellular end of the pore[23,24].

A large degree of structural and functional similarity among members of the pLGIC family facilitated their study, with results obtained for one member informing our understanding of the family as a whole. In particular, work on

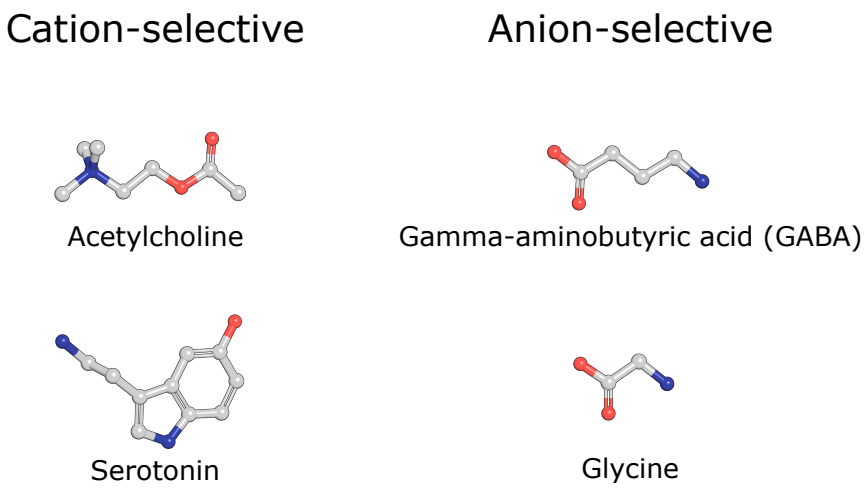

Figure 1.1: Structures of four neurotransmitters that act on pentameric ligand-gated ion channels. Acetylcholine and serotonin bind to cation-selective channels. GABA and glycine activate anion-selective channels.

nicotinic acetylcholine receptors (nAChR) and bacterial ion channels such as the *Gloeobacter* and *Erwinia* ligand-gated ion channels (GLIC and ELIC) provided a structural and functional framework to contrast and compare the different channels. The high degree of homology allowed the structures of nAChR[25] (2004), ELIC[26] (2008), and GLIC[27] (2009) to inform our understanding of the glycine receptor, as well[28,29]. GLIC in particular continues to be a source of insight on pentameric ligand-gated ion channels in general. Chapter 6 reports the results of a study in which intermediate activation states of GLIC were investigated. The overall structure of pLGICs is explained on the example of the glycine receptor in Section 1.2.1 and shown in Figs. 1.2 and 1.3.

Physiological roles of pLGICs include, but are not limited to, signal transmission at the neuromuscular junction[30], regulation of nausea[31], contributions to addiction[32], fast inhibitory transmission in the central nervous system[33], anxiety[34], and memory functions[34]. Mutations in these receptors are associated with a variety of diseases, including epilepsy[35], congenital myasthenic syndrome[36], and hyperekplexia[37].

1.2 The Glycine receptor

The glycine receptor is an inhibitory pLGIC that is primarily expressed in the spinal cord, where one of its main functions is the regulation of motor control[38,39]. Binding of the neurotransmitter triggers a structural rearrangement, which leads to the opening of an ion-conducting pore and the influx of chloride ions into the cell. This influx increases polarisation of the membrane potential, making action potentials less likely[40]. Study of this ion channel has been ongoing for decades. The earliest evidence that glycine can act as an inhibitory neurotransmitter dates back to the late 1960s[41]. Glycine receptors were then purified and later cloned in the 1980s[42,43]. In 2015, the first structures of the glycine receptor were published[44,45].

1.2.1 Overall structure and function

While there is a significant degree of variety between subunits of pentameric ligand-gated ion channels, many elements are highly conserved. This allows a similar nomenclature to be used in discussions of different members of this channel family.

The structure of a glycine receptor subunit is shown in Fig. 1.2. It is broadly similar to structures of other pLGICs. The classical pLGIC subunit has an extracellular domain (ECD) formed of an extended β-sheet of ten β-strands (β1–10). This ECD connects to a transmembrane domain (TMD) which consists of four transmembrane α-helices (M1–M4). The linkers between helices M1 and M2 and between M2 and M3 are relatively short. M3 and M4, on the other hand, are separated by a large loop that forms the subunit's intracellular domain (ICD). These loops are the least conserved regions between different subunits and appear to be structurally flexible[46,47] — to date, no structure with a fully resolved ICD has been published.

These subunits associate as pentamers to form functional channels (Fig. 1.3 A). GlyR has five orthosteric binding sites. These sites are located at each subunit interface in the ECD. Ligand recognition occurs through six conserved binding loops A–F[48]. Each subunit carries all six of these loops, but loops A–C and D–F

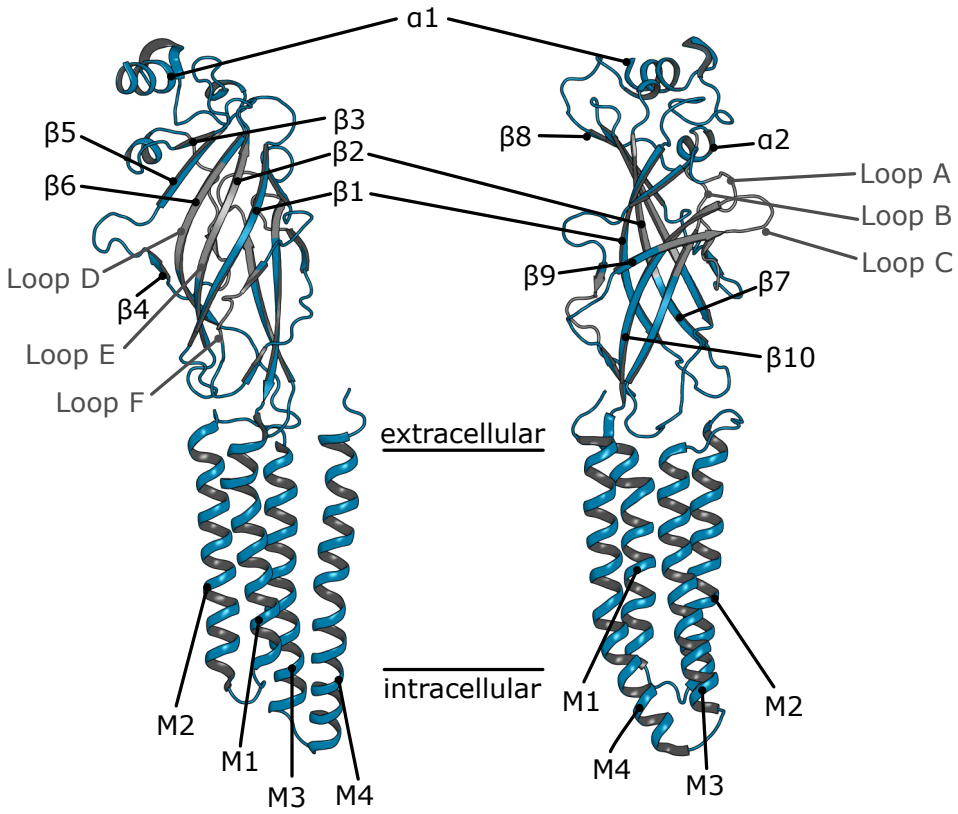

Figure 1.2: Labelling of the secondary structure elements and ligand binding loops (loop A–F, grey) of a *Danio rerio* glycine receptor α1 subunit (PDB 3JAF[44]).

are located on opposite sides (Fig. 1.2). The binding site is thus formed of loops A, B, and C of one and loops D, E, and F of the other subunit at the interface (Fig. 1.3 B). Loop C is the most critical binding loop as it physically covers the binding site[49]. To distinguish between the two subunits at each interface, that which contributes its loop C is referred to as the principal (or +) subunit. Its counterpart is called the complementary (or –) subunit.

The extracellular domain is linked to the transmembrane domain, both directly via the β10–M1-linker and through an allosteric network, a particular focus point of which is the M2-M3 linker[50]. Conformational changes in the ECD, caused for example by ligand binding, trigger rearrangements in the TMD (and vice versa) through a 'conformational wave'[51–53]. Crucially, this communication occurs rapidly, conferring sub-millisecond-timescale gating cycles on the channel[54]. This allosteric network is subject of study in chapter 5 of this thesis, where the opposing

effects of an inhibitor binding in the ECD and an activator binding in the TMD are investigated.

Following ligand binding in the ECD, the conformational wave arrives in the TMD. Here, the channel pore is formed by the M2 helix of each subunit (Fig. 1.3 C). It is between these M2 helices that the narrowest constriction in the channel lies and that gating occurs. M2 has thus seen particular interest. In nomenclature that originated in the study of nAChRs[55], its residues are numbered in a prime notation. The then-presumed intracellular end of M2 was designated as residue 1′ and the immediately following residues were numbered accordingly. Nowadays, negative numbering in this notation is used to indicate residues on M2 that are located below the originally presumed intracellular end of M2 (Fig. 1.3 D).

Two residues on M2 in particular are responsible for the gating behaviour of the channel. The five L9′ residues form the activation gate[57–59], and the five P-2′ form the so-called desensitisation gate[56,60]. The rearrangements of M2 during the gating cycle are shown in Fig. 1.4. In the closed state, the constriction at L9′ acts as a hydrophobic gate[25,61] — geometrically wide enough to accommodate ions, the hydrophobicity in this region would require the stripping of the hydration shell which imposes too large an energetic penalty for permeation[62–64]. As discussed above, ligand binding in the ECD leads to a widening of the pore in the TMD. The open state is short-lived, however, and desensitisation quickly sets in[65]. The desensitised state's M2 helices adopt a V-shaped configuration: the extracellular half of the pore is similarly wide open as that of the open state, but the intracellular end constricts at P-2′. After ligand unbinding and recovery from desensitisation, GlyR returns to the resting state[66].

1.2.2 Subunit combinations

Pentameric ligand-gated ion channels can exist as hetero- or homopentamers of different subunits. There are 17 different nAChR subunits (α1–10, β1–4, γ, δ, and ε) in vertebrates[68,69]. The $GABA_A$-receptor shows even more complexity, with a total of 19 different subunits (α1–6, β1–3, γ1–3, δ, ε, ϑ, π, ϱ1–3)[70].

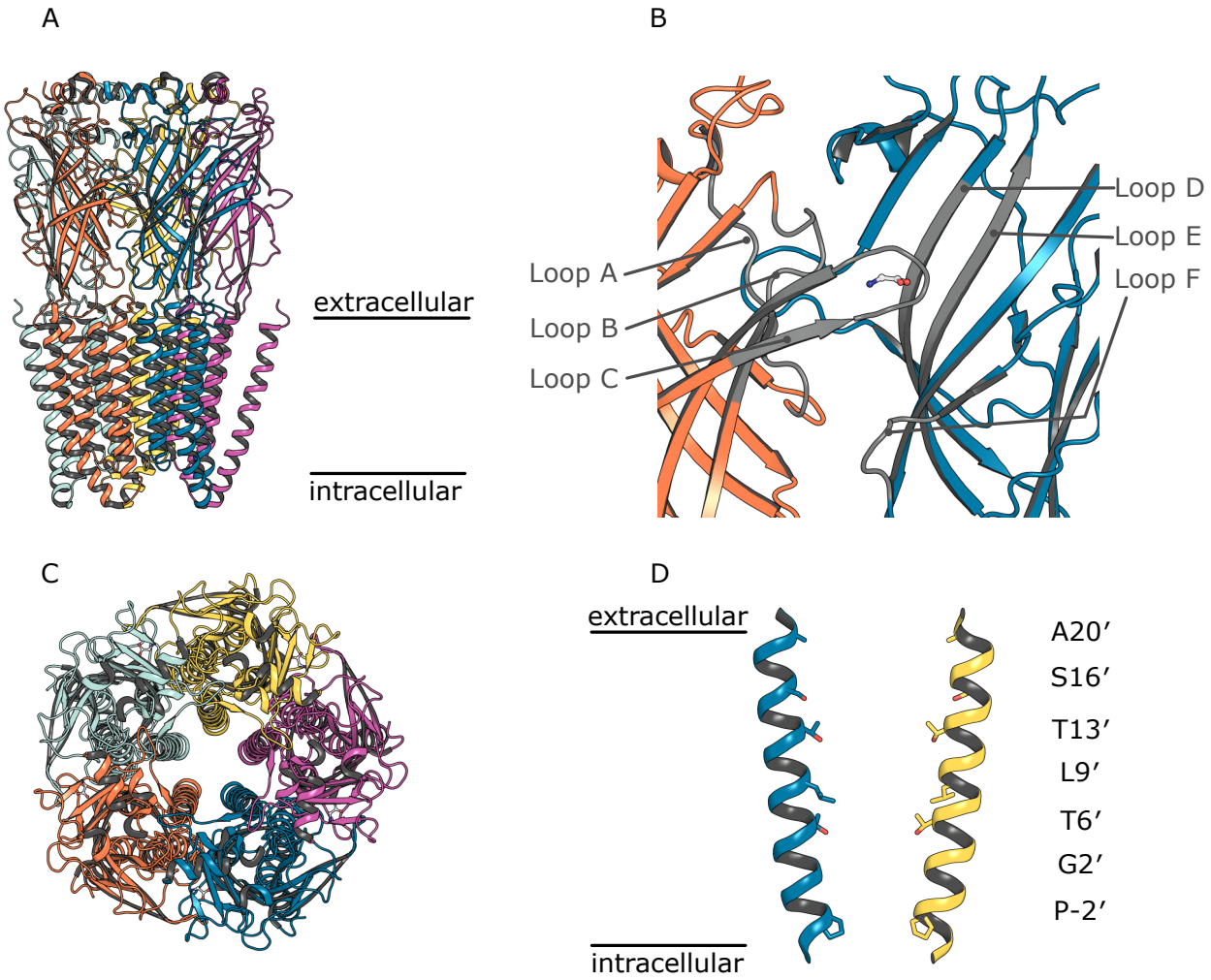

Figure 1.3: A: Side view of a glycine receptor pentamer (PDB 3JAF[44]). B: Close-up view of one of the five orthosteric ligand binding sites, located at the subunit interface in the extracellular domain. Glycine is bound between the principal (orange) and the complementary subunit (blue) (glycine coordinates from PDB 6UD3[56]). C: Extracellular view onto the pentamer shown in A. The view is centred on the ion-conducting pore. The five helices in the centre are each subunit's M2 helix. D: Illustration of the M2 helices of two opposing subunits. The pore-lining residues are shown as sticks and labelled with their prime notation[55].

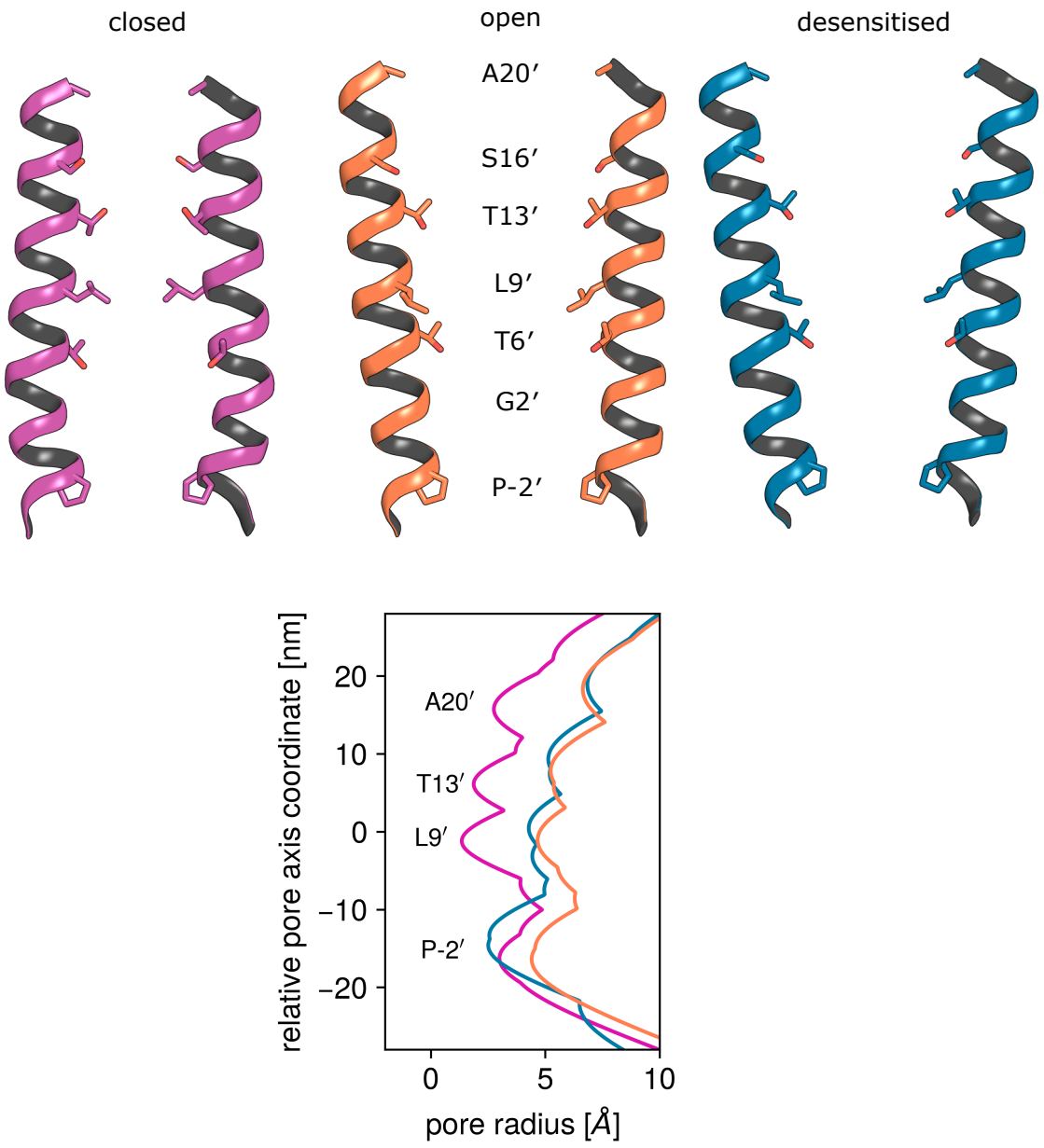

Figure 1.4: Changes of M2 helix orientation during the gating cycle. The constriction sites in the pore lie at P-2′, L9′, T13′, and A20′, with L9′ forming the most constricted site in the closed state (magenta, PDB 3JAD). In the transition to the open state (orange, PDB 3JAE), the M2 helices move apart from one another, and key residues like L9′ turn away from the pore. This results in an expansion of the pore throughout its entire length. During desensitisation (blue, PDB 3JAF), the M2 helices come together at the intracellular end of the pore, and the P-2′ constriction stops ion flow. NB: the open-state structure used here has been challenged as artificially wide open[59]. This figure is replicated in the appendix with an MD-stabilised open-state structure (Fig. A.1). The observations made here stand, regardless of which of the two open-state structures is used. Radius profiles were calculated with HOLE2[67].

The glycine receptor, on the other hand, is comparatively simple: only 5 subunits have been identified, α1–4 and β. The majority of glycine receptor pentamers shares the same subunit composition: prenatally, the predominant form of GlyR is the α2-homopentamer. Adult receptors, on the other hand, are mostly α1β-heteropentamers[71]. These heteropentamers consist of four α1 subunits and one β subunit. The latter is required for synaptic localisation as it is responsible for interactions with the anchoring protein gephyrin[72]. Incorporation of β-subunits into GlyR pentamers additionally reduces the channel's sensitivity to picrotoxin[73].

1.2.3 Pharmacology

The glycine receptor is an important drug target and has multiple binding sites for agonists, positive allosteric modulators (PAMs), and inhibitors. The Glycine Receptor Allosteric Ligand Library (GRALL) is a database of such compounds and their known or suspected modes of action[74]. A number of ligand binding sites throughout the extracellular and transmembrane domains of the channel have been described (Fig. 1.5). Of these binding sites, the orthosteric binding site and the overlapping ivermectin / alcohol binding site were investigated further in this thesis.

The orthosteric binding site

The orthosteric binding sites lie at the subunit interfaces in the extracellular domain and are formed of six conserved binding loops A–F (Fig. 1.3 B). Besides glycine, β-alanine and taurine are endogenous agonists of the glycine receptor[79].

This site is also where a competitive inhibitor of GlyR—strychnine—binds. Strychnine is a highly selective, high-affinity inhibitor that was invaluable in the early advancement of our understanding of glycine receptors[80] that, for example, saw use as a radiographic probe[81].

Loop C plays a large role in the differential effects of orthosteric agonists and inhibitors. When an agonist enters the binding site, this loop covers the binding

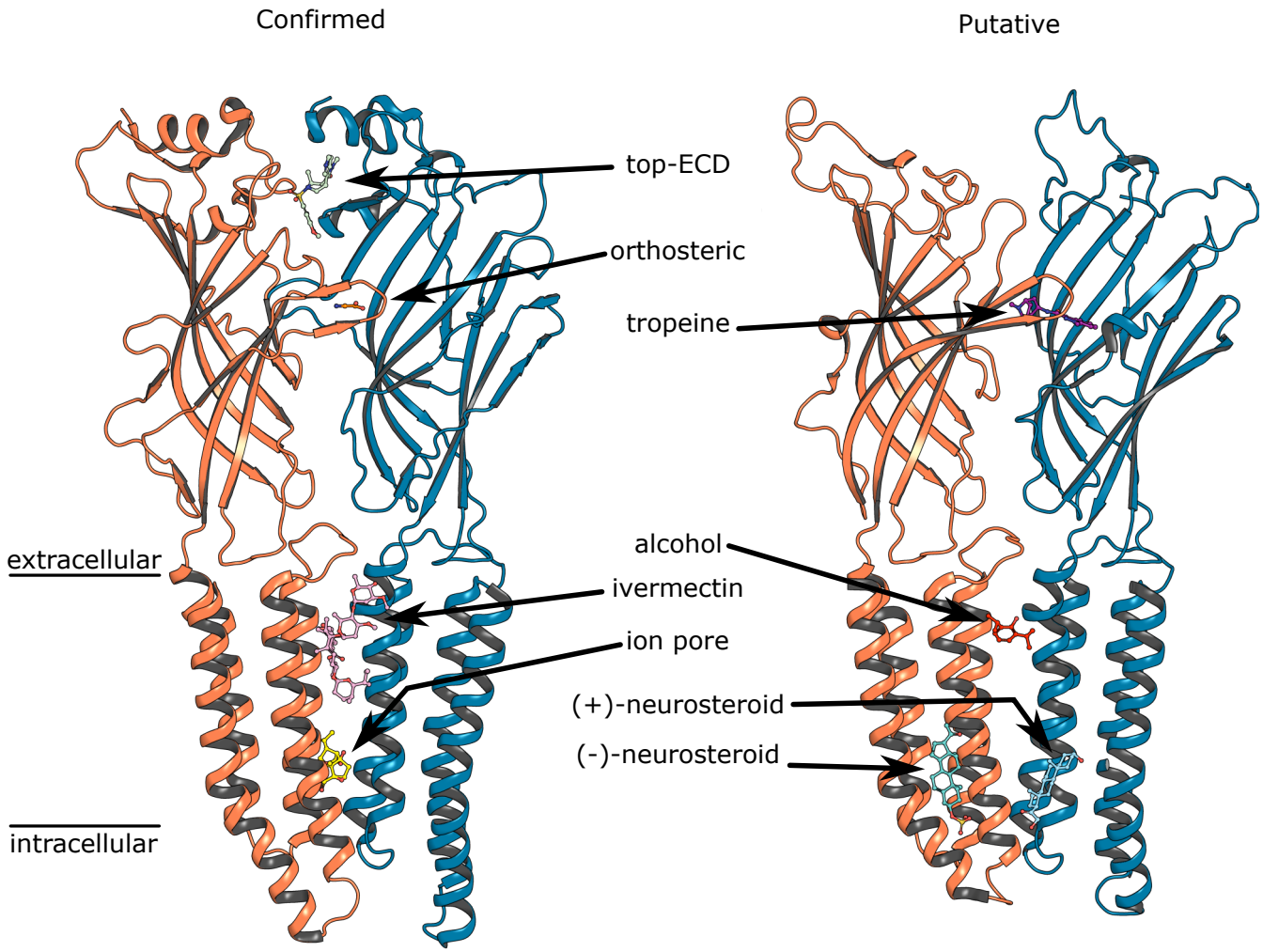

Figure 1.5: The GRALL database categorises ligands by strength of evidence. The strongest form of evidence in GRALL is direct, high-resolution structural evidence of binding at the glycine receptor. Such evidence is available for some compounds in the so-called top-ECD, orthosteric, ivermectin, and ion pore binding sites. Shown here, in order, are AM-3607, glycine, ivermectin (all PDB 5VDH[75]) and picrotoxin (PDB 6HUG[76]). Less strong evidence of binding at the glycine receptor is available for the tropeine, alcohol, and activating (+) or inhibiting (-) neurosteroid binding sites. Such weaker evidence could be a high-resolution structure of a different pLGIC combined with concordant evidence in GlyR, for example. Shown in these putative binding sites here are, in order, granisetron (PDB 6NP0, 5HT3A-receptor[77]), propofol (PDB 5MVM, GLIC), tetrahydrodeoxycorticosterone (PDB 5OSB, GLIC-$GABA_AR$-chimera[78]), and pregnenolone sulfate (PDB 5OSC, GLIC-$GABA_AR$-chimera[78]).

site more closely. Antagonist binding, on the other hand, generally causes Loop C to adopt a more open conformation[49,82].

The competing effects of strychnine and ivermectin (described below) are investigated in chapter 5.

The ivermectin / alcohol binding site

High-resolution structural information locates the ivermectin binding site on GlyR in the subunit interface within the transmembrane domain[44]. Strong evidence suggests that an alcohol / anaesthetics binding pocket overlaps with the confirmed ivermectin binding site[83–86]. These compounds are all activators of GlyR in one way or another. Ivermectin, for instance, is a potentiator of glycine-induced currents when applied at low concentrations, and acts as a full, current-inducing agonist when applied at higher concentrations[87].

In addition to its role as a drug target, this site has also been shown to interact with lipids in structural studies[88] and simulation[89].

In a 2012 study, Islam and Lynch reported that the insect-GABAR pore blocker fipronil is also an inhibitor of the human glycine receptor. They suggest that in addition to its pore blocking site, it might interact at the alcohol binding site[90]. This claim is further discussed in chapter 4, where it is developed into the first description of fenestration-based drug access in a pLGIC family member.

1.3 Importance of membrane permeability of small molecules in drug development

Chapters 3 and 4 of this thesis are concerned with the study of how ligands reach their site of action in the cell. In chapter 3, I developed an improved umbrella sampling workflow and tested it on the problem of membrane permeation of small molecules.

The ability to cross biological membranes is a critical requirement for most drugs[91,92] as it is a large determinant of their absorption, distribution, metabolism,

and excretion parameters (ADME)[93]. As a criterion to quickly assess likely bioavailability due to passive membrane permeability, Lipinski's rule of five[94] can be used. This rule of thumb states that a bioavailable compound is unlikely to have more than 5 hydrogen bond donors, more than 10 hydrogen bond acceptors, a molecular mass more than 500 daltons, or an octanol-water partition coefficient (logP) of more than 5. Often, only compounds passing this initial hurdle are considered further, e.g. for chemical synthesis and *in vitro* tests such as the parallel artificial membrane permeation assay (PAMPA)[95] or Caco-2 cell-based assays[96].

logP calculation is thus an important part of the early drug discovery process. It is particularly useful in high-throughput applications due to its ease of computation, but it comes with an accuracy trade-off. If quantitative permeability estimates are required, logP values alone do not suffice. They also fail to capture complex membrane compositions. Molecular dynamics simulations can prove useful here[97]. In MD, the desired system composition can be precisely defined and permeation data can then be obtained, for example through umbrella sampling simulations[98,99]. Using the inhomogeneous solubility-diffusion model (ISD)[100], an estimate of the permeability coefficient P can be obtained from the potential of mean force (PMF) and the diffusivity D of a compound at any given reaction coordinate ξ in the membrane system as

$$\frac{1}{P} = \int_{\xi_1}^{\xi_2} \frac{e^{\beta\ PMF(\xi)}}{D(\xi)} d\xi, \tag{1.1}$$

where β is the inverse temperature $1/k_BT$. Permeability coefficients obtained from simulation data in this way can then in theory be compared to experimental PAMPA or Caco-2 data[91], with the caveat that perfect agreement is not expected due to the more complex nature of the experimental assays.

MD simulation can thus be a useful tool in drug discovery and permeability prediction. However, obtaining converged PMF data has historically been challenging[101–103]. In the course of this thesis, I encountered these same convergence problems when I tried to calculate PMFs for a complicated system (chapter 4). To overcome these issues, I developed an improved umbrella

sampling workflow with enhanced window generation, simulation parameters, and statistical estimation of PMFs. This workflow is thoroughly tested and benchmarked in chapter 3, where I break down the individual and combined effects of its three constituent methods.

1.4 Fenestrations in membrane proteins

In 1977, Bertil Hille carried out functional studies on voltage-gated sodium channels (Na_vs)[104]. His results led him to hypothesise on the presence of a membrane-exposed access pathway for hydrophobic drugs, decades before the first structures of ion channels were published. It was not until 1998 that Doyle et al. reported the first structure of the KcsA potassium channel[105], and not until 2011 that William Catterall's group published the first structure of a voltage-gated sodium channel[106]. In this 2011 publication, the authors report the presence of hydrophobic tunnels that laterally connect the ion-conducting pore to the channel's membrane environment. They coined the term "fenestration" for such tunnels and refer back to Hille's work, believing to have identified the hypothesised drug access pathway. Since this seminal work, fenestrations have been identified and characterised in a variety of different ion channel families. Fenestration-based drug access has also been referred to as membrane to channel inhibition (MCI)[107]. The functional importance of this mechanism is demonstrated by the recent identification of disease-related mutations in fenestration regions, a discovery that gave rise to the term fenestropathy[108].

1.4.1 Voltage-gated sodium channels

Voltage-gated sodium channels (Na_vs) were the original focus of the study of fenestrations. It has long been understood that resting-state block of Na_vs is mediated through them[104]. Since the initial publication of the bacterial Na_vAb structure[106] (Fig. 1.6 A), ever more structural data on the sodium channel and

its fenestrations has become available[109–111]. One early study tested the resting-state block caused by lidocaine, benzocaine, and flecainide[112] (Fig. 1.7). The authors found that mutating a phenylalanine residue at the fenestration to the larger tryptophan reduced the drugs' affinities for Na_vAb. Mutation to alanine, on the other hand, increased the affinity — most dramatically of the largest of the compounds, flecainide, which the authors argue is most restricted by the fenestration radius.

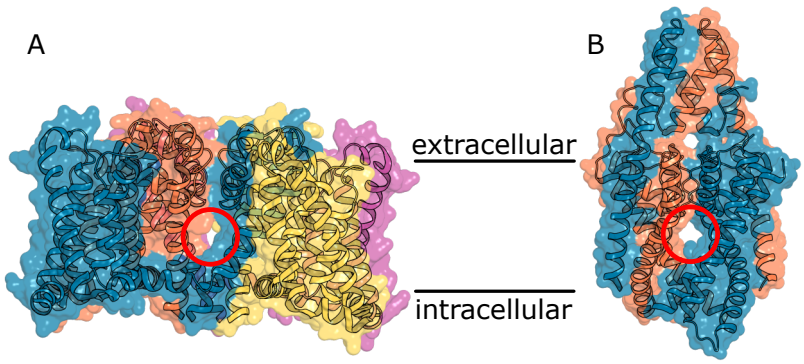

Figure 1.6: Example illustration of fenestrations. A: The bacterial voltage-gated sodium channel Na_vAb (PDB 3RVY[106]) has four fenestrations, one at each subunit interface in the pore module. B: The TREK-2 K2P channel (PDB 4XDK[113]) has two fenestrations in the transmembrane domain.

The importance of fenestration radii is intuitive. It was nicely illustrated through work on the sodium leak channel NALCN carried out by Schott et al[114]. This channel is structurally related to Na_vs, but it is known for its pharmacological resistance. Its fenestrations are obstructed by four key bulky residues. The authors tested the efficacy of common channel blockers on the NALCN wild-type and a mutant in which one bulky residue for each of its four fenestrations was mutated to alanine. They found that these mutations significantly increase the channel's susceptibility to phenytoin (Fig. 1.7). This study demonstrates that fenestration dimensions are indeed important determinants of drug access.

The fenestrations of voltage-gated sodium channels have been extensively studied with molecular dynamics simulations[115,116]. These studies report fenestration radii in Na_vs during simulations that mostly fall between 1 and 3 Å and vary slightly between fenestration and channel types. They also provide

evidence of and insight into fenestration permeation and describe differing binding affinity through the use of metadynamics. Other studies carried out µs-timescale simulations and demonstrated fenestration permeation without any bias[117,118].

While the relative importance of fenestration-based compared with pore-based drug access is still under discussion[109,119], the fact that drug access through fenestrations does occur is undisputed. Outside of voltage-gated sodium channels, drug-permeable fenestrations have been described in other channel families, as well. The following is by no means a complete list and is only intended to serve as further evidence of the importance of fenestrations across ion channel families. A selection of known fenestration-permeating drugs is shown in Fig. 1.7 to illustrate the wide range of sizes and chemical diversity found in these compounds.

1.4.2 Potassium channels

The family of potassium channels is large and functionally and structurally diverse[124]. Interactions between drugs and fenestrations have been described in the voltage-gated human ether-à-go-go-related gene (hERG)[125] and $K_v1.5$[126] channels. An interaction-dependent fenestration has been identified in KCNQ1, where it only opens when the channel associates with its auxiliary KCNE1 subunit and undergoes a structural rearrangement[127].

Among two pore domain potassium channels, the mechanosensitive TREK-2[113,121,122,128] and TRAAK[129] channels have been confirmed to have fenestrations. The former is a target of norfluoxetine (Fig. 1.7), while lipid occupancy of fenestrations in the latter has been reported. It is possible that lipid-fenestration interactions contribute to mechanosensation[130]. Fenestrations in the TWIK-1 channel have been probed with structural methods and MD simulations[131,132]. As an example, the TREK-2 structure with its fenestration is shown in Fig. 1.6 B.

1.4.3 Voltage-gated calcium channels

The presence of fenestrations in voltage-gated calcium channels is understood as a conserved structural feature[133]. They are critical for closed-state inacti-

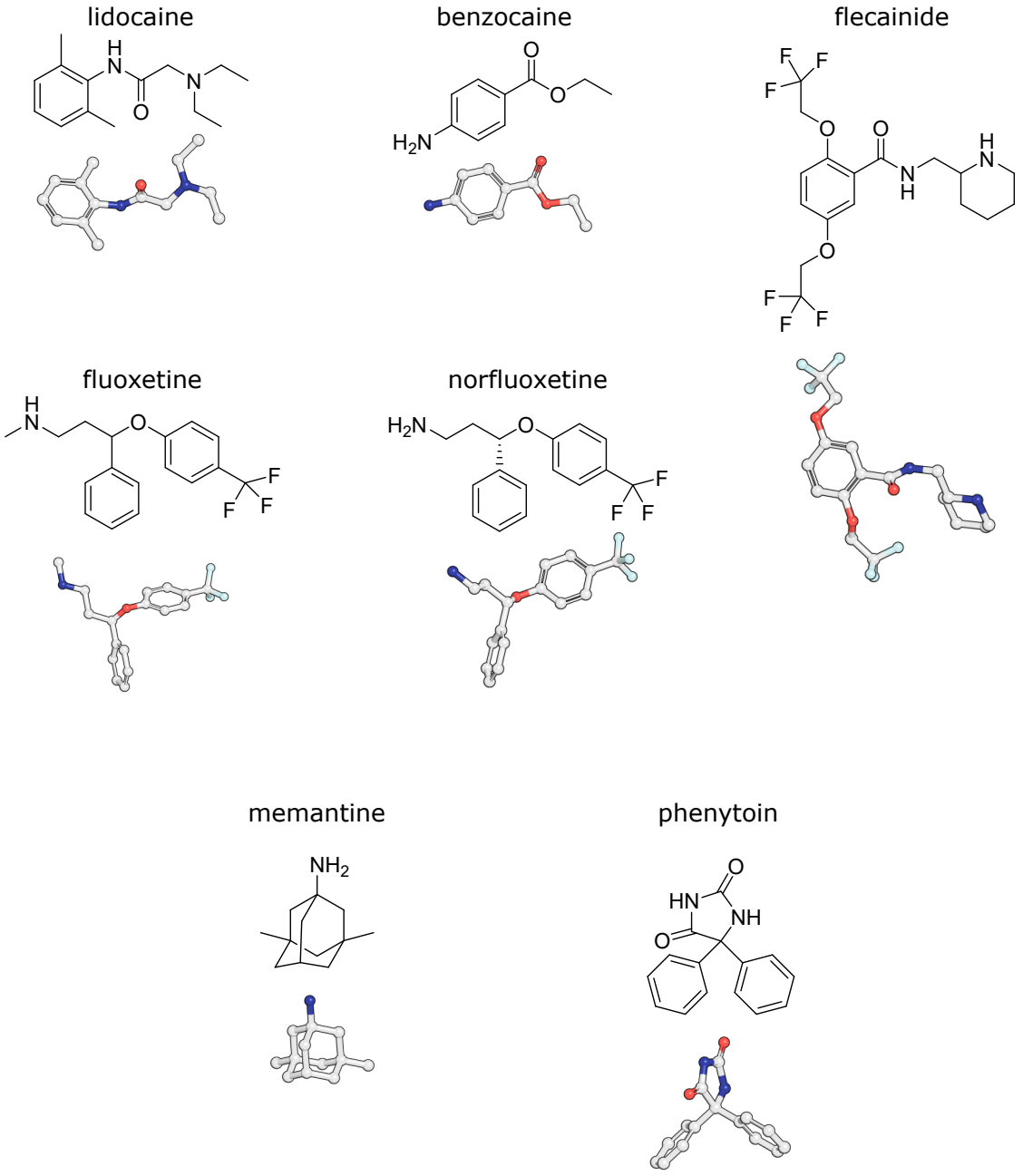

Figure 1.7: Compounds that are either known or suspected to interact with ion channel fenestrations, shown as chemical formula and in liquorice representation. Lidocaine, benzocaine, and flecainide are inhibitors of voltage-gated sodium channels[112]. (Nor)fluoxetine's fenestration interaction has been demonstrated in the AMPA glutamate receptor[120] and the TREK2 K2P potassium channel[121,122]. Memantine is a fenestration-going inhibitor of NMDA glutamate receptors[123]. Phenytoin has been shown to inhibit the sodium leak channel NALCN if bulky, fenestration-blocking residues are mutated to alanine[114]. Demonstrably, molecules of a diverse range of sizes and functional groups have the potential for fenestration-based pore access.

vation of $Ca_v2.2$, for example[134], and were also implicated in dihydropyridine binding in general[135].

1.4.4 Glutamate receptors

Fenestrations have also been found in ionotropic glutamate receptors. To wit, the NMDA receptor was found to be susceptible to memantine, with strong evidence that it interacts with fenestrations in the channel[123]. The importance of fenestration-based membrane-to-channel inhibition is well-understood in this family[107]. Another glutamate receptor, the AMPA receptor, was also shown to be a target of fenestration-going fluoxetine[120] (Fig. 1.7).

1.4.5 Pentameric ligand-gated ion channels

This section has reported just some of the available evidence of fenestrations in the better-studied ion channel families. One family for which evidence of fenestrations is conspicuously unavailable is that of the pentameric ligand-gated ion channels. It comprises nicotinic acetylcholine receptors, 5-HT_3-receptors, $GABA_A$-receptors, glycine receptors, and glutamate-gated chloride channels, covering a vast range of physiological roles. It is thus somewhat surprising that drug-permeable fenestrations have not been discovered or discussed in this entire channel family—fenestrations are only mentioned as extensions of the ion-conduction pathway in the extracellular domain[136].

The discussion of how ligands reach their site of action starts in chapter 3, where the membrane permeation of small molecules is studied. This discussion continues in chapter 4. Here, I aim to lay the foundation for future discussions of fenestrations in pLGICs. Through the use of MD simulations, I provide further context for electrophysiology data on the glycine receptor that appears congruent with membrane-to-channel inhibition. Structural data and simulation results provide the first evidence of a state-dependent, drug permeable fenestration in glycine receptors.

1.5 Aims of this thesis

This thesis is aimed at the study of pentameric ligand-gated ion channels, particularly the glycine receptor, with computational methods. The key research question of chapter 4 was whether or not fenestration-based drug access in the glycine receptor is possible or likely. Addressing this question required a large number of difficult-to-converge umbrella sampling simulations to be run. Limiting the amount of simulation time needed thus became a second aim of the thesis and led to the development of an improved umbrella sampling workflow that combines three previously described components. Chapter 3 describes the detailed benchmarking of this workflow that I have carried out on test cases of drug-membrane permeation. In addition to testing the performance of this workflow, a goal of this project was to quantify the error that is introduced in potential of mean force calculations when symmetry or periodicity are enforced as boundary conditions, as they commonly are.

In a collaborative project, results from simulation methods and wet-lab biochemistry were combined to further our understanding of allosteric signalling in the glycine receptor. Using molecular dynamics simulations, I helped answer the question of how ivermectin binding in the transmembrane domain of GlyR is able to disrupt strychnine binding in the orthosteric site. This is reported in chapter 5 of this thesis.

A second collaboration aimed to elucidate the activation cycle of the *Gloeobacter* ligand-gated ion channel (GLIC) further. While it is useful to discuss channel states as either open, closed, or desensitised, in reality, there is large structural variety within and between these groupings. Our collaborators employed a specialised cryo-EM data analysis pipeline to solve structures of GLIC in intermediate activation states. My simulations on these structures aimed to answer the question of how open-like a structure has to be to allow the flow of ions through the channel. These results are reported in chapter 6.

2 Methods

Contents

2.1 Molecular dynamics

As atoms are quantum objects, quantum mechanics provides the most correct description of molecular systems. However, due to their theoretical and mathematical complexity, QM calculations become prohibitively expensive for large systems. With current hardware, systems larger than a few hundred atoms cannot readily be studied with quantum mechanical methods. Molecular mechanics offers a way to study larger systems by employing a few approximations. A classical Newtonian description of the system is built in which the potential energy is calculated as a function of nuclear coordinates only, while electronic positions are assumed to average out around the nuclei in accordance with the Born-Oppenheimer approximation. This reduced level of detail allows simulations of systems that are millions of atoms in size on microsecond timescales, but it also means that electronic effects such as polarisation, ionisation, and the breaking and formation of covalent bonds are beyond the scope of a pure molecular mechanics-based approach.

The theory underlying molecular dynamics simulations is detailed here following the examples of Leach[137] and Allen and Tildesley[138].

The objective of molecular dynamics simulations is to apply Newton's equation of motion to a molecular system to calculate the evolution of atomic coordinates over time. To this end, MD simulations utilise the molecular mechanics-based description of the system of interest to calculate its potential energy from its atomic positions. The parameters at the foundation of this calculation are collectively referred to as a force field and usually include terms to describe bond stretching, angle bending and torsional motions as well as parameters to calculate potentials between atoms not linked with covalent bonds. The force on each atom is obtained as the derivative of the potential energy with respect to the atomic positions, which via Newton's second law $F = ma$ yields accelerations of each atom. Through integration over a finite, small time step the next set of coordinates is obtained.

This process is repeated until a trajectory covering the desired amount of time has been calculated.

All simulations in this work are carried out with GROMACS version 2021.7[139–145].

2.1.1 Force fields

Force fields are the collections of parameters and functional forms that are used to calculate the potential energy from the atomic positions. The limitations inherent in the molecular mechanical approach make a universal force field that could describe every chemical species in every conceivable condition impossible. Instead, force fields are carefully parametrised to reproduce a subsection of experimental findings (i.e. heats of vaporisation or bond lengths) for specific subsets of chemical space (i.e. proteins or DNA), and care has to be taken that a suitable force field is chosen for the research question at hand.

A large number of force fields has been developed over the years. The following functional form (Eq. 2.1) is frequently used,

$$\begin{aligned} U_{MolecularMechanics} = & \sum_{bonds} U_{bonds} \\ & + \sum_{angles} U_{angles} \\ & + \sum_{torsions} U_{torsions} \\ & + \sum_{pairs} U_{electrostatic} \\ & + \sum_{pairs} U_{van\ der\ Waals} \end{aligned} \tag{2.1}$$

where the potential energy U is described as a sum of various energy terms arising from bonded and non-bonded interactions that are considered independently. These energetic contributions will be discussed in detail in the following.

Bonded interactions

Bond stretching

In molecular mechanics, molecules are approximated as spheres connected by

springs. These springs can be described by Hooke's law, which relates the force on a spring to its extension and its spring constant. In the molecular mechanics analogy, the extension of the spring is the deviation from a reference bond length x_0 and the spring constant is a measure of the strength or stiffness of the bond. This formulation treats bond stretching as a form of harmonic oscillation (Fig. 2.1).

$$U_{bond}(x) = \frac{k_{bond}}{2}(x - x_0)^2 \tag{2.2}$$

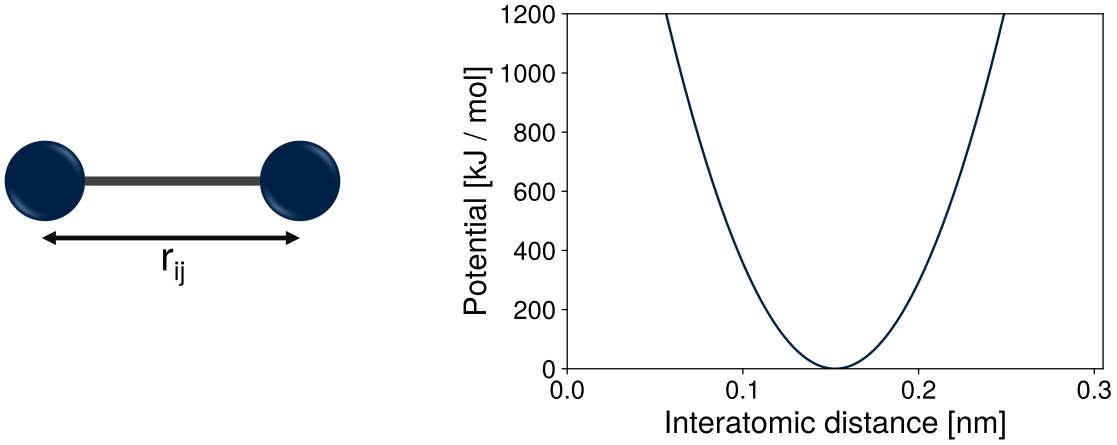

Figure 2.1: The bond-stretching potential of a C-C bond as defined in the AMBER99SB-ILDN force field with its reference bond length of 0.1525 nm and its spring constant of 259 408 kJ mol^{-1} nm^{-2}.

Covalent bonds are strong interactions and the spring constants are correspondingly large. As a consequence, deviations from the reference bond length x_0 are usually small, but frequencies of oscillation are very high. This has consequences for the time step that can be used in integration. This will be discussed further in Section 2.1.4.

Angle bending

The energy required to bend a bond angle is implemented analogously to the bond stretching terms. A harmonic potential term is used.

$$U_{angle}(\theta) = \frac{k_{angle}}{2}(\theta - \theta_0)^2 \tag{2.3}$$

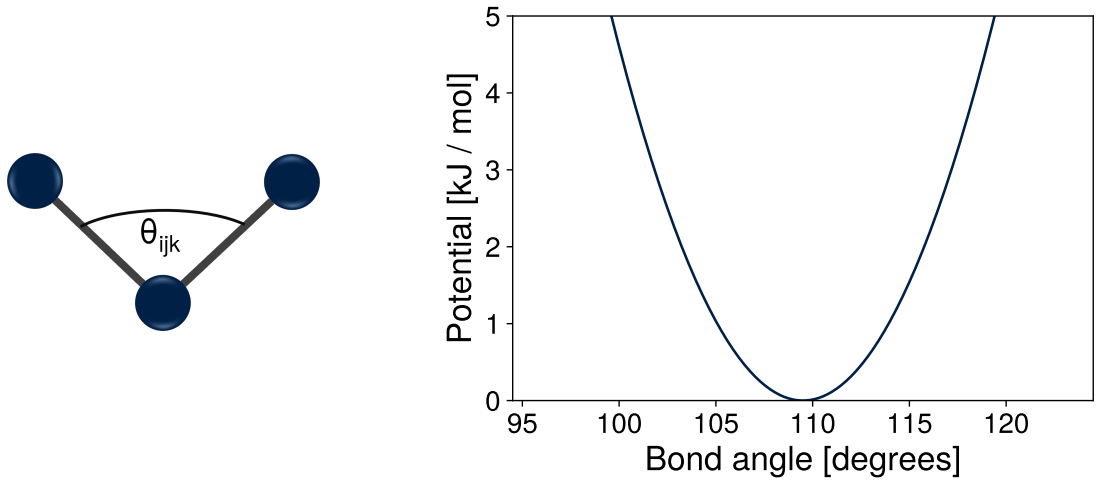

Figure 2.2: The angle-bending potential of an N-C-O angle as defined in the AMBER99SB-ILDN force field with its reference angle of 105.9° and its spring constant of 334.72 kJ mol^{-1} rad^{-2}.

It is of note that substantially less energy is required for small angle deformations than for bond stretching. This is due to the much smaller force constants used to accurately reproduce experimental observations of molecular behaviour through, for example, infrared spectroscopy or X-ray crystallography.

Torsional potential terms

The torsional angle ϕ_{ijkl} of four sequentially bonded atoms i, j, k, and l is defined as the angle between the two planes formed by i, j, k and j, k, l, respectively. Since, generally, a full rotation of 360° about the j, k bond is possible, the harmonic potentials used for bonds and angles are not suitable descriptors of these dihedral angles. The functional form of the dihedral potential instead makes use of the inherent periodicity of the trigonometric functions (Eq. 2.4), where k_ϕ relates to the barrier height, n is the multiplicity, and ϕ_s determines the minima of the torsional potential.

$$U_{dihedral}\left(\phi_{ijkl}\right) = k_\phi\left(1 + \cos\left(n\phi_{ijkl} - \phi_s\right)\right) \tag{2.4}$$

To allow a better description of torsions in asymmetrical systems, Eq. 2.4 can be expanded as the sum of torsional contributions with individual values for k_ϕ, n, and ϕ_s. As an example of this, the torsional potential of the OS-CT-CT-OS dihedral as defined in the AMBER99SB-ILDN force field[146] and its individual contributions are shown in Fig. 2.3. The two contributions' parameters are listed in Table 2.1.

Table 2.1: Parameters of the two contributions to the OS-CT-CT-OS dihedral of the AMBER99SB-ILDN force field.

Contribution	k_ϕ	n	ϕ_s
1	0.6025	3	0
2	4.91620	2	0

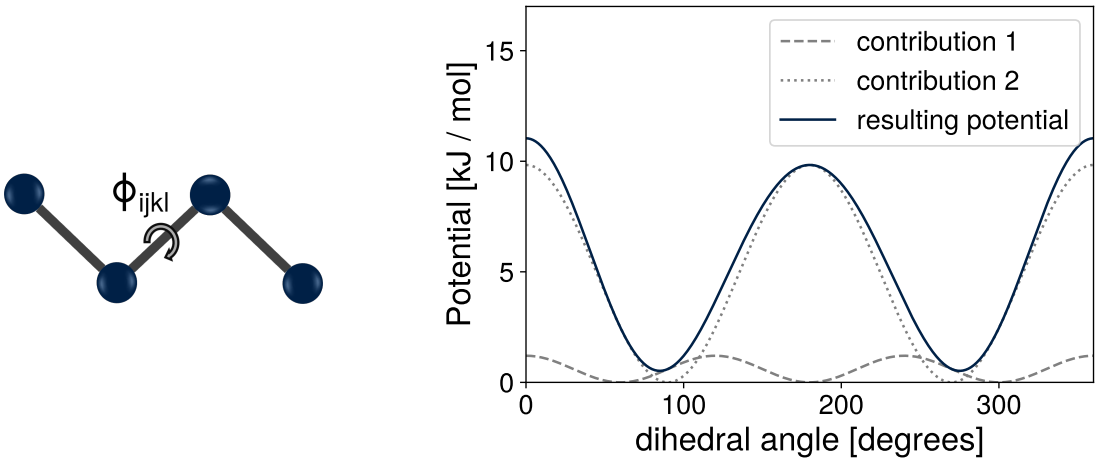

Figure 2.3: The dihedral potential of an OS-CT-CT-OS dihedral as defined in the AMBER99SB-ILDN force field. It is formed as a sum of two contributions of the functional form of Eq. 2.4 with the parameters in Table 2.1.

Accurate modelling of torsional degrees of freedom is especially important. Bond stretching and angle bending are relatively hard degrees of freedom due to their implementation as harmonic oscillators with large force constants. On the other hand, a wide range of dihedral angles is usually explored in simulations because the energy barriers involved in changing dihedral angles are often comparable to the thermal energy kT at the temperature of interest. For example, $kT \approx 2.58\,\mathrm{kJ\,mol^{-1}}$ at 310 K. This means that rotations about single bonds in particular are responsible for a major part of structural variability.

Nonbonded interactions

Electrostatic interactions

Electrostatic interactions between two charged atoms 1 and 2 with charges q_1 and q_2 at a distance r_{12} are modelled through Coulomb's law:

$$U_{Coul} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_{12}}, \tag{2.5}$$

with the vacuum permittivity ϵ_0 of 8.8542 × 10^{-12} F nm^{-1}. Even at distances beyond 1 nm, electrostatic interactions still contribute significantly to the overall potential (Fig. 2.4). This is problematic: applying a cut-off and only calculating the electrostatic potential to nearby atoms introduces large errors, but considering each pair of atoms in the simulated system is expensive as it scales with the square of the number of atoms in the system. A solution to this problem is discussed in Section 2.1.7.

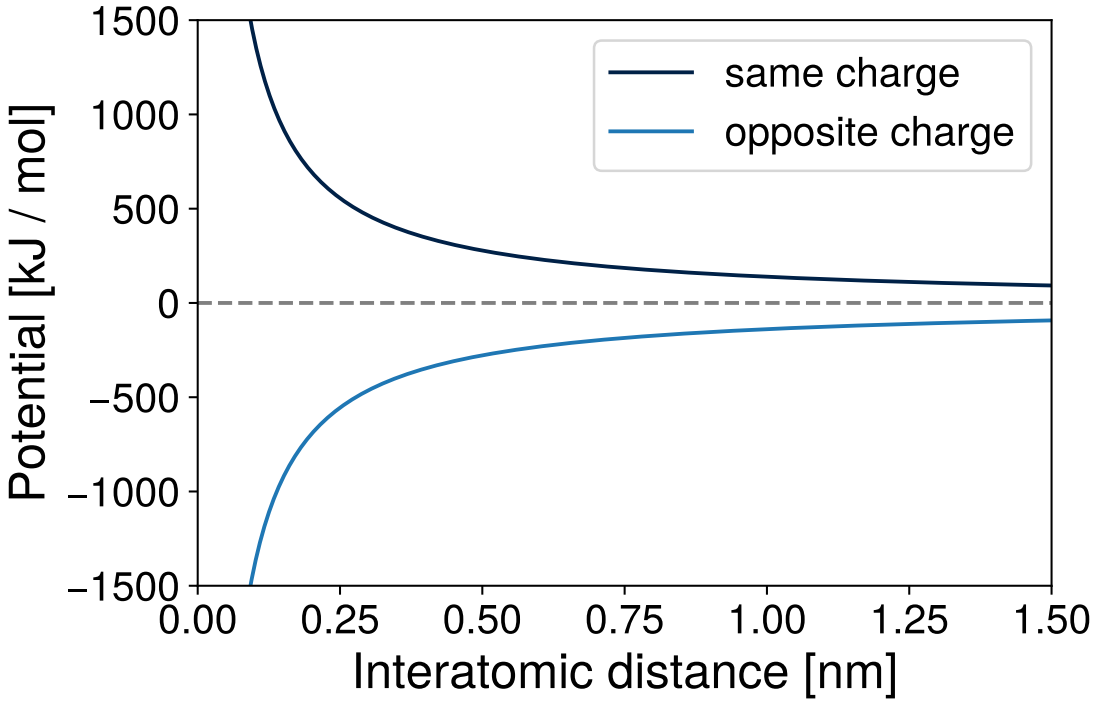

Figure 2.4: The Coulomb potential between two singly charged ions with identical (repulsive force) or opposing charges (attractive force) in water.

This implementation of the electrostatic potential is not only applied to nominally charged species like ions. Instead, to reflect the fact that some elements are

more electronegative than others and will more strongly attract the electrons that form covalent bonds, each atom is assigned a partial charge. The actual electronic distribution in a molecule is then represented as a distribution of partial point charges, and electrostatic potentials between point charges are calculated in the same way with Eq. 2.5.

Van der Waals interactions

Even between non-bonded, uncharged atoms there are interactions. This is made clear in the case of noble gases, for example, which otherwise could not exist as a liquid phase even near absolute zero. These interactions are due to van der Waals forces and have attractive and repulsive components. At long distances, an attractive force acts between particles because of dispersion, where instantaneous dipoles are formed by random fluctuations in the electron cloud around each atom. These dipoles can be self-reinforcing between close atoms, giving rise to the attractive component. The strongly repulsive force on two atoms that approach one another closely results from a violation of the Pauli exclusion principle, as when the atoms approach one another their electron orbitals start to overlap. No electrons in an orbital may share the same set of quantum numbers, resulting in a large energetic penalty for decreasing interatomic distance.

For computational efficiency – and because electronic effects cannot be described in molecular mechanics – these van der Waals interactions are routinely described by the Lennard-Jones potential,

$$U_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right], \tag{2.6}$$

which has a strongly repulsive component that decreases with r^{12} and a weakly attractive component that decreases with r^6. Thus, for two atoms approaching one another, the Lennard-Jones potential will decrease to a minimum at $r_{min} = 2^{1/6}\sigma$ before rapidly increasing (Fig. 2.5). The formulation in Eq. 2.6 combines the attractive and repulsive contributions into a single, simple equation that relies on two parameters that have to be empirically determined for each atom type

in a force field: σ, the distance between two atoms at which the Lennard-Jones potential is 0, and ϵ, the depth of the energy well at r_{min}.

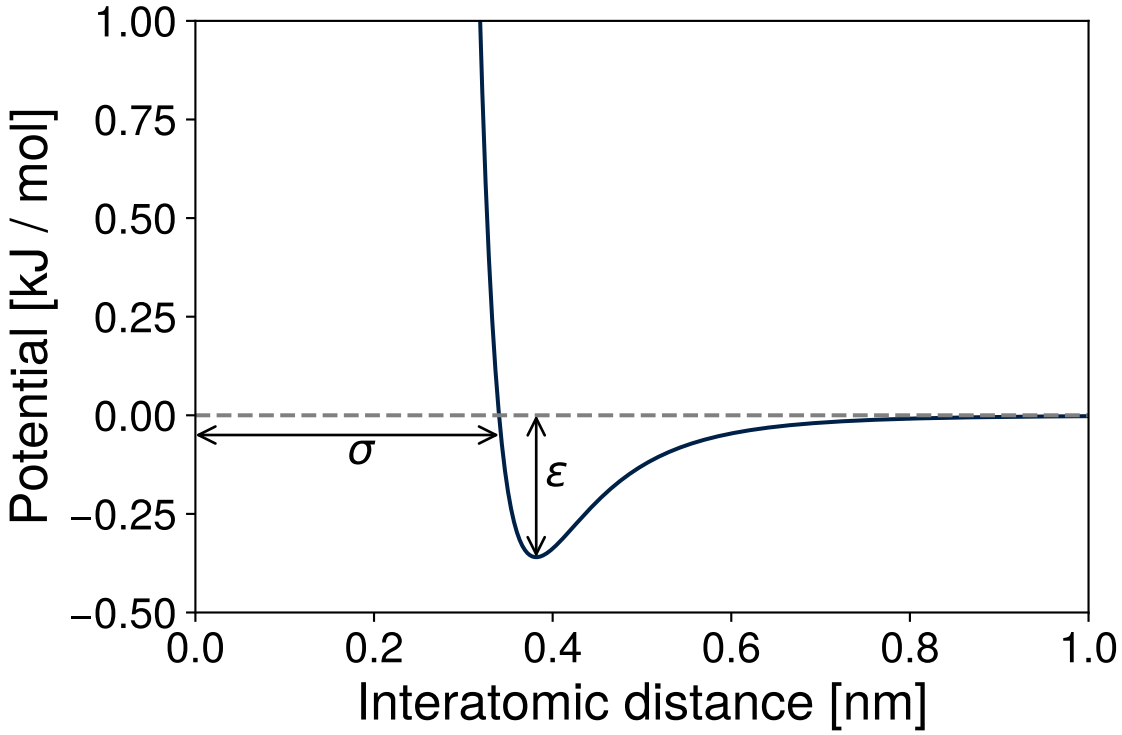

Figure 2.5: The functional form of the Lennard-Jones potential with its r^{-12} repulsive contribution and its r^{-6} attractive contribution. ϵ is the depth of the energy well at the minimum of the potential energy, and σ is the distance when the potential is zero.

If different atom types are present in a simulation, the Lennard-Jones parameters for their interaction can be obtained from the individual parameters via combination rules. For example, for two species A and B, the parameters for the A-B Lennard-Jones interaction can be obtained via the Lorentz-Berthelot combination rules:

$$\sigma_{AB} = \frac{1}{2}(\sigma_A + \sigma_B)$$

$$\epsilon_{AB} = \sqrt{\epsilon_A \epsilon_B}$$

It is important to note that the formulation of the Lennard-Jones potential as a pairwise potential is an approximation. In reality, multi-body effects contribute significantly to the potential energy in condensed phases. Their inclusion is

computationally expensive, however, so most force fields instead modify the parameters of their pair potentials slightly to account for these multi-body effects. This improves the accuracy in simulations of condensed phases, but as a consequence, these parameters are not suitable for the description of an interaction between an isolated pair of atoms.

2.1.2 Integration methods

The force field is used to calculate the potential energy of the system as a function of the nuclear coordinates. The forces F on each atom are then obtained as the derivative of the potential with respect to atomic coordinates. From here, MD simulations produce trajectories by calculating the evolution of atomic coordinates over time utilising Newton's equations of motion. Integration of these equations gives time-dependent expressions for the velocities and coordinates. An iterative procedure is used, where previous coordinates and velocities are used to determine a new set of coordinates after a short time step Δt. For the first simulation step starting from a static structure, when no initial velocities are present, they can be chosen randomly from a Maxwell-Boltzmann distribution at the desired temperature (Fig. 2.6).

Two different integration methods were used in this work.

The leap-frog integrator

The leap-frog integrator uses the positions **x** and accelerations **a** at time t and the velocities **v** at time $t - \frac{1}{2}\Delta t$ to update the positions and velocities as such:

$$\mathbf{a}(t) = \frac{\mathbf{F}(t)}{m}$$

$$\mathbf{v}\left(t + \frac{1}{2}\Delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\Delta t\right) + \Delta t\, \mathbf{a}(t)$$

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta t \mathbf{v}\left(t + \frac{1}{2}\Delta t\right),$$

where m is the atom mass. The velocities at time $t + \frac{1}{2}\Delta t$ are calculated first from those at $t - \frac{1}{2}\Delta t$ and the accelerations at time t, and are then used with

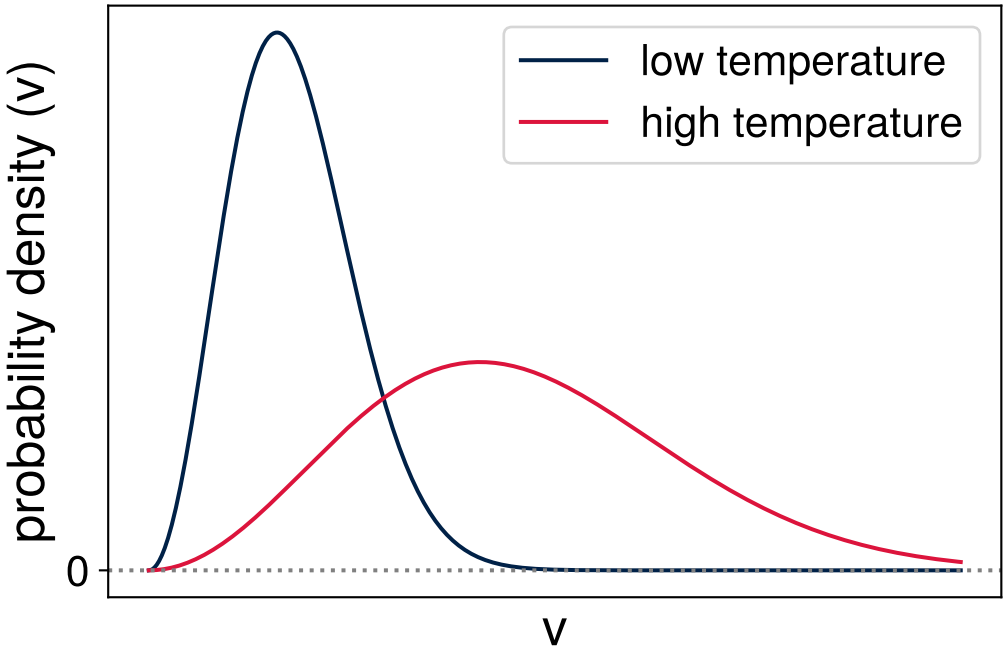

Figure 2.6: The Maxwell-Boltzmann distribution is a probability distribution of molecular speed v. At low temperatures, most particles have low speeds, and the speed distribution is narrow. At higher temperatures, molecular speeds are more broadly distributed and cover a wider range of speeds. The highest-probability speed is higher, but a smaller number of particles have this particular speed.

the positions at time t to calculate the positions at $t + \Delta t$. As such, positions are available at each time step, and velocities are available at each half time step — they leap-frop over one another as the simulation progresses. The leap-frog integrator is computationally efficient, but has the disadvantage that positions and velocities are not synchronised. The kinetic energy is not calculated for the same time point as the potential energy, which prohibits the use of simulation methods that need to process the total energy as the simulation is ongoing. One example of this is the expanded ensemble methods, for which the velocity-verlet integrator has to be used (see Section 2.2).

The velocity-verlet integrator

The velocity-verlet integrator is different from leap-frog in that positions, velocities, and accelerations are all given for the same time t. This is achieved by first

calculating the next set of positions as

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta t \mathbf{v}(t) + \frac{1}{2}\Delta t^2 \mathbf{a}(t).$$

To then obtain the velocities at $t + \Delta t$, the accelerations at both t and $t + \Delta t$ are needed. $\mathbf{a}(t + \Delta t)$ is calculated from the positions $\mathbf{x}(t + \Delta t)$ calculated before. With these accelerations, the new velocities are then

$$\begin{aligned}\mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{1}{2}\Delta t \left[\mathbf{a}(t) + \mathbf{a}(t + \Delta t)\right] \\ &= \mathbf{v}\left(t + \frac{1}{2}\Delta t\right) + \frac{1}{2}\Delta t \mathbf{a}(t + \Delta t).\end{aligned}$$

Since the velocities and positions are determined at each whole time step, the total energy of each time step is known as the simulation progresses. The downside of the velocity-verlet integrator is that, in its current GROMACS implementation, it cannot fully utilise GPU acceleration. It is therefore only employed here when necessitated by the use of, for example, an expanded ensemble method (see Section 2.2).

2.1.3 Energy minimisation

Generally, before a simulation can be started, the input coordinates have to be slightly modified in a process called energy minimisation. The goal is to find a minimum on the potential energy surface from which to start the simulation. This is necessary because large potential energy values yield large forces, which then result in large movements of atoms in a single time step. These large movements can then cause more unfavourable configurations of the system; for example, atoms might end up very close to one another, resulting in large repulsive Lennard-Jones potentials. This way, the simulation can become increasingly unstable, and with larger and larger atom displacement per time step is said to 'blow up'. When atom overlap is severe, potentials may become too large to be stored in a single-precision 32 bit floating-point number, causing abrupt termination of the simulation.

The most commonly used energy minimisation algorithm is *steepest descent*. It is a gradient-based method, meaning the derivatives of the potential energy—the forces—are used to inform small movements of atoms in the search of an

energy minimum. In the iterative procedure, the coordinates $\mathbf{x}_{n+1}$ at step $n+1$ are calculated from the forces $\mathbf{F}_n$ on each atom as

$$\mathbf{x}(n+1) = \mathbf{x}(n) + \frac{\mathbf{F}_n}{\max\left(|\mathbf{F}_n|\right)} h_n, \tag{2.7}$$

where $\max\left(|\mathbf{F}_n|\right)$ is the largest scalar force on any atom at step n, which is used as a normalisation factor. h_n is a variable step size that has to be defined for the first minimisation step and is then updated throughout the run of the algorithm. The potential energy based on the new set of coordinates is calculated, and the move is accepted if the new potential energy is lower than before and rejected if it is not. When a move is accepted, the step size h_n is increased by 20 %. If it is rejected, the step size was too large and the system crossed over a local energy minimum. h_n is then multiplied by 0.2 to reduce the step size. The algorithm stops when either a user-defined maximum number of steps has taken place, the potential energy has decreased below a user-defined threshold, or when an energy minimum has been found and the potential energy no longer changes.

2.1.4 Bond constraints

The choice of time step is important: smaller time steps mean the calculation of a trajectory covering a certain time takes longer, but too large a time step makes the system become unstable due to the effects described in Section 2.1.3. Therefore, the time step should be chosen to be as large as is possible while still maintaining a stable simulation.

Looking at this problem from another perspective: interesting conformational changes of a molecule such as bond rotations are low-frequency motions. On the other hand, bond stretching terms have high-frequency motions, but the bond lengths are of less interest because they oscillate very close to their reference values. However, the maximum length time step that can be chosen must still be able to resolve the highest frequency motions in the system. This conflict — uninteresting motions of the system forcing the use of a small time step and thus a slow simulation — can partially be solved by the use of constraint algorithms.

Bond constraints can be used to force bond lengths to their reference value. This eliminates their bond vibration and thus reduces high frequency motions in the system. Commonly, all bonds between heavy atoms and hydrogen atoms in a simulation system are constrained. This allows the use of a longer time step, frequently of 2 fs instead of 1 fs. In this work, the LINCS algorithm[147] was used to constrain all bonds between heavy atoms and hydrogen atoms except in the case of water, which was constrained with the SETTLE algorithm[148]. After initial, unconstrained coordinates for the next time step have been calculated through integration of Newton's equation of motion (see Section 2.1.2), these algorithms correct the coordinates to enforce the constraints.

2.1.5 Periodic boundary conditions

Correct treatment of boundary conditions is crucial for the accurate calculation of thermodynamic properties from simulation. Most of the time, we are interested in bulk properties, not interactions with system boundaries or surfaces. However, simulation systems would have to be enormously large for a significant number of molecules in the centre of the system not to interact with the boundaries. Periodic boundary conditions enable the study of bulk phenomena in small systems, saving computational time.

In a simulation under periodic boundary conditions, the simulation system is replicated in space, surrounding the system with identical copies of itself (Fig. 2.7). Mirror particles directly copy the actual system's particle's motion, and if a particle leaves the central system, it is replaced by a mirror particle entering from the opposite end. In this way, bulk behaviour can be simulated even in relatively small systems. A minimum size of the simulation system arises as a consequence of having to prevent a particle from interacting with its own mirror images. This is done by implementing a cut-off for short-ranged non-bonded interactions beyond which their potentials are set to zero, and following the minimum image convention under which only the closest copy of a particle is considered for these interactions. This closest particle need not be in the same copy of the simulation system. Thus,

the box length has to be at least twice the cut-off distance. Efficient implementation of the cut-offs is done using neighbour lists (see Section 2.1.6). Long-ranged non-bonded interactions have to be treated differently (see Section 2.1.7).

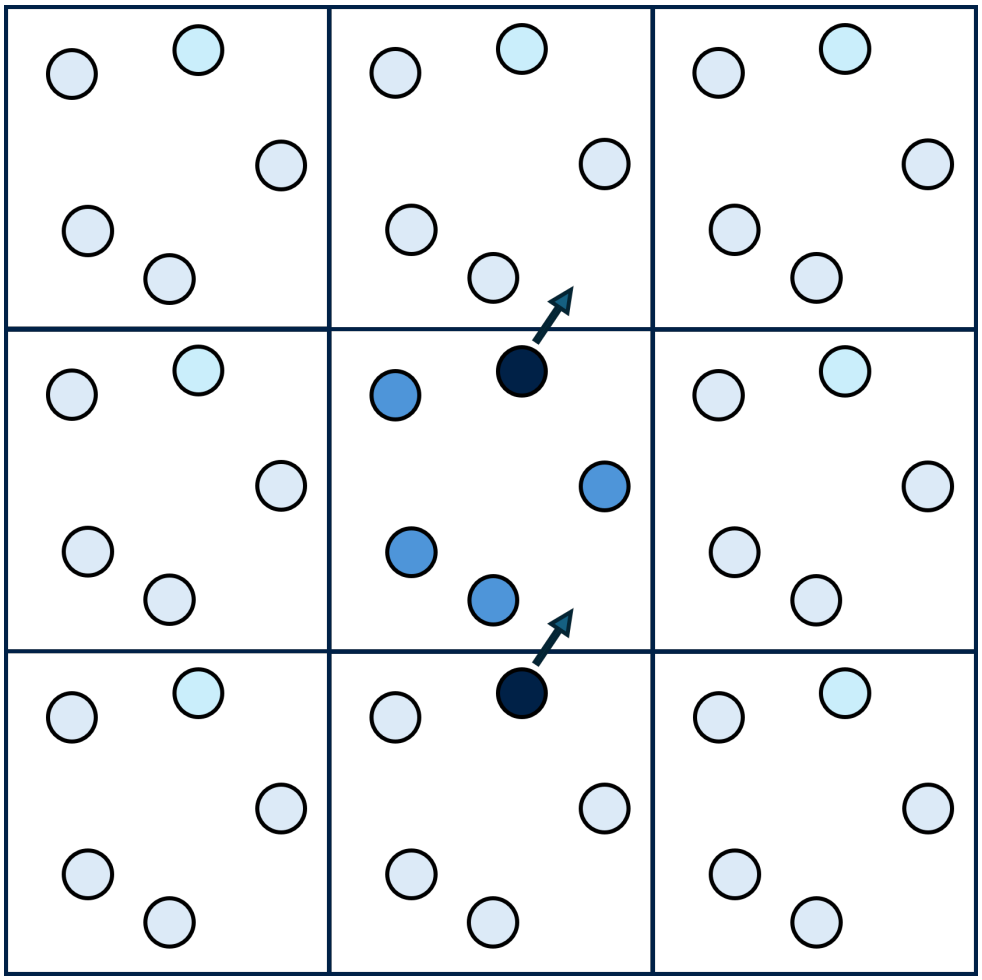

Figure 2.7: Illustration of periodic boundary conditions. The actual simulation system in the centre is surrounded by identical copies of it. If a particle leaves the box, its mirror image enters the box on the opposite side from such a copy.

2.1.6 Cut-offs and neighbour lists

Non-bonded potential terms trail off with increasing distance. This is especially true for the Lennard-Jones potential, the magnitude of which decreases with r^6. However, during the force evaluation part of an MD step, non-bonded interactions of all pairs of atoms are calculated. The potential energy of a pair of atoms at considerable distance will be very close to zero, and in a large enough system,

most atoms will be so far away from most other atoms as to barely interact with them. To reduce the number of force calculations necessary, a cut-off distance for non-bonded interactions is defined beyond which the potential energy is set to zero. This reduces the number of necessary force evaluations, but by itself does not meaningfully speed up the overall calculation. The reason is that the distance between each pair of atoms would still have to be calculated and compared to the cut-off distance to decide whether or not the force has to be determined. Neighbour lists were implemented to reduce the number of distance calculations that need to be carried out during a simulation.

Neighbour lists are based on the argument that in condensed phases, atoms cannot dramatically change their positions in a few femtoseconds. Therefore, it is assumed that each atom's neighbours will remain the same for a brief period of time. For each atom in the system, a neighbour list holds a reference to all atoms that are within and slightly beyond the distance cut-off. Only these atoms are then considered for non-bonded interactions. The neighbour list has to be updated at certain intervals to account for atoms entering and leaving the cut-off radius. This way, the expensive calculation of all $N(N-1)$ distances in the simulation box does not have to be carried out at every step.

The sudden switch from a finite potential just below the cut-off to a potential of zero just above the cut-off introduces artefacts. One common way to address this is to shift the potential energy by addition of a constant that ensures the potential reaches zero at the cut-off distance. This improves energy conservation and has no effect on the forces, as the constant term disappears in differentiation.

2.1.7 Long-ranged interactions

Electrostatic interactions modelled with Coulomb's law only decrease with r^1 as compared to the Lennard-Jones potential's r^6. These interactions are usually still significant at distances near the cut-off used in the calculation of Lennard-Jones interactions (Fig. 2.4), and implementing a plain cut-off introduces significant artefacts. Instead, electrostatic interactions are split: a short-ranged component

benefits from the performance improvements of the cut-off and neighbour list, and a long-ranged component ensures that electrostatic interactions beyond the cut-off are still considered.

The short-ranged electrostatic component is calculated as described in Section 2.1.1. This calculation occurs in direct space. The long-ranged component, on the other hand, is usually calculated with the smooth particle mesh Ewald (SPME) method[149], which handles its calculations in reciprocal space. This approach employs fast Fourier transforms (FFTs) to efficiently calculate the electrostatic potential, allowing PME to scale as $\mathcal{O}(n \log n)$.

PME enables more accurate treatment of electrostatic interactions and is computationally efficient. However, since it considers all periodic images in the force evaluation and does not use the minimum image convention, it can artificially impose long-ranged order on the system that may not be desired. Another point of note is that PME requires a charge-neutral system and will create a uniformly charged background to neutralise any charge imbalances, if present. Systems should therefore always be created with neutralising ions in the solvent.

2.1.8 Temperature and pressure coupling

Natively, molecular dynamics simulations sample from the microcanonical ensemble, in which the number of particles N, volume V, and total energy E are constant. However, if results from simulations are to be compared to experimental findings, it makes sense to match experimental conditions as closely as possible. Most experiments are not carried out under conditions of constant volume and total energy, but instead under constant temperature T and constant pressure P. Thus, simulating in this isothermal-isobaric NPT ensemble requires control of both temperature and pressure during the simulation.

Temperature coupling is achieved by manipulating the velocities of the system. This is based on the relation of temperature T and velocity v via the kinetic energy E_{kin} as

$$\frac{3}{2} N k_B T = E_{kin} = \frac{1}{2} m v^2,$$

so changes in the overall velocities directly correspond to changes in the instantaneous temperature and vice versa. Immediately changing the velocities to match the desired temperature at every integration step, however, would result in unphysical behaviour of the system and prevent proper sampling of the ensemble. Instead, the simulation system is coupled to a heat bath of the desired temperature T_{bath} in such a way that the change of the system's temperature over time is

$$\frac{dT(t)}{dt} = \frac{1}{\tau_T}\left(T_{bath} - T(t)\right),$$

where τ_T is a coupling constant that determines how closely the the system's temperature is coupled to the heat bath's. This is the formulation of the Berendsen thermostat[150]. During one time step of Δt, the temperature changes by

$$\Delta T = \frac{\Delta t}{\tau_T}(T_{bath} - T(t)),$$

which gives a scaling factor λ to apply to the velocities to reach the new temperature as

$$\lambda^2 = 1 + \frac{\Delta t}{\tau_T}\left(\frac{T_{bath}}{T(t)} - 1\right).$$

In this work, the velocity-rescale thermostat[151] was used. It is a modified version of the Berendsen thermostat that includes a stochastic term. This is necessary for proper sampling of thermal fluctuations.

Pressure coupling is generally achieved by varying the size of the simulation box. A scaling factor λ has to be determined, which can then be applied to the box vectors and atomic coordinates. In the Berendsen barostat[150], a 'pressure bath' is defined in analogy to the heat bath used for temperature coupling. The rate of change of the pressure is

$$\frac{dP(t)}{d(t)} = \frac{1}{\tau_P}(P_{bath} - P(t)),$$

with a coupling constant τ_P defined analogously to τ_T for temperature coupling. From the rate of change of the pressure, the scaling factor for the simulation box volume is obtained as

$$\lambda = 1 - \kappa\frac{\Delta t}{\tau_P}(P - P_{bath}),$$

where κ is the isothermal compressibility of the simulation box. In cases where the simulation system is mostly made up of water, a value for κ of 4.5 × 10^{-5} atm^{-1} is reasonable. From the scaling factor λ the new atomic positions x_i' are obtained as

$$x_i' = \lambda^{1/3} x_i.$$

Pressure coupling can be applied isotropically, semi-isotropically, or anisotropically. In the isotropic case, the same scaling factor is applied in all three spatial dimensions. In the anisotropic case, each dimension can have its own scaling factor and can change independently. The semi-isotropic case is most relevant to the study of membrane systems. Here, it is assumed that the membrane lies in the xy plane, and the same scaling factor is used for the x and y components, while the z component can be scaled independently to more accurately mirror conditions in a membrane system.

In this work, the C-rescale barostat[152] was used. It is an evolution of the Berendsen barostat that includes a stochastic term in addition to the deterministic term. This allows correct modelling of pressure fluctuations in the system.

2.1.9 Simulating electric fields

Biological membranes generally have ion gradients across them. From this charge gradient results an electric membrane potential. It can be useful to simulate the effects of this membrane potential, which is possible through simulation of an external electric field. This electric field can be imagined as the result of placing the plates of a parallel plate capacitor at opposite ends of the simulation box and applying a voltage to them. In practice, the strength of the simulated electric field E in volts per nanometre can be defined as an input parameter to the simulation. The force experienced by charged species i with charge q_i in the simulation box as a result of this field is calculated during the force evaluation step as

$$F_{electric\ field} = q_i E,$$

and is then used for updating the coordinates the same way forces obtained from application of the force field are.

The external field is specified via its field strength, but comparisons to the membrane potential are usually more useful. The strength of the simulated potential can be obtained from the field strength by multiplying it with the length of the simulation box along which the field is applied. This implies, however, that the potential changes with box size fluctuations that result from the pressure coupling in simulations in the NPT ensemble. In principle, then, the NVT ensemble should be used when precise knowledge of the electrical potential in the simulation box is needed, but the fluctuations are rather minor in practice. A more serious problem arises due to the use of periodic boundary conditions. Dipoles in the system will partially align with the electric field, creating an internal field that opposes the external field. This causes a screening effect, reducing the effect of the external field in the simulation box, and is one of the reasons why much higher potentials than measured in live cells are often used in simulations.

2.2 Expanded ensemble methods

Expanded ensemble methods are related to replica exchange methods, and both can be used to enhance sampling in a simulation by exploring a variety of conditions that might differ in temperature, force field parameters, umbrella potentials, alchemical states, or other factors. In replica exchange, these different states are sampled in parallel. One simulation per set of parameters is running at all times, and exchanges of coordinates between adjacent conditions are periodically attempted. Expanded ensemble methods are a serialised version of this parallel workflow with only one simulation system that periodically attempts to move to adjacent conditions. One advantage of expanded ensemble methods over replica exchange methods is that the hardware requirements are more flexible. In parallel applications, each condition has to be simulated at the same time, and communication between the simulation ranks has to be fast. Cross-node or

cross-GPU communication can be a serious bottleneck, limiting the number of conditions that can be tested in parallel. In the expanded ensemble approach, on the other hand, this limitation does not apply.

Expanded ensemble methods have successfully been used for alchemical transformations[153,154] and umbrella sampling[155], with modified Hamiltonians and umbrella potentials, respectively. Simulated tempering[156] is the application of the expanded ensemble framework to different temperatures (Fig. 2.8). As in the parallel-tempering replica exchange approach, the goal is to overcome potential energy barriers through sampling at higher temperatures. This allows sampling of more regions of phase space that would be inaccessible if only the actual temperature of interest were sampled. Spending simulation time at higher temperatures allows faster crossing of energy barriers.

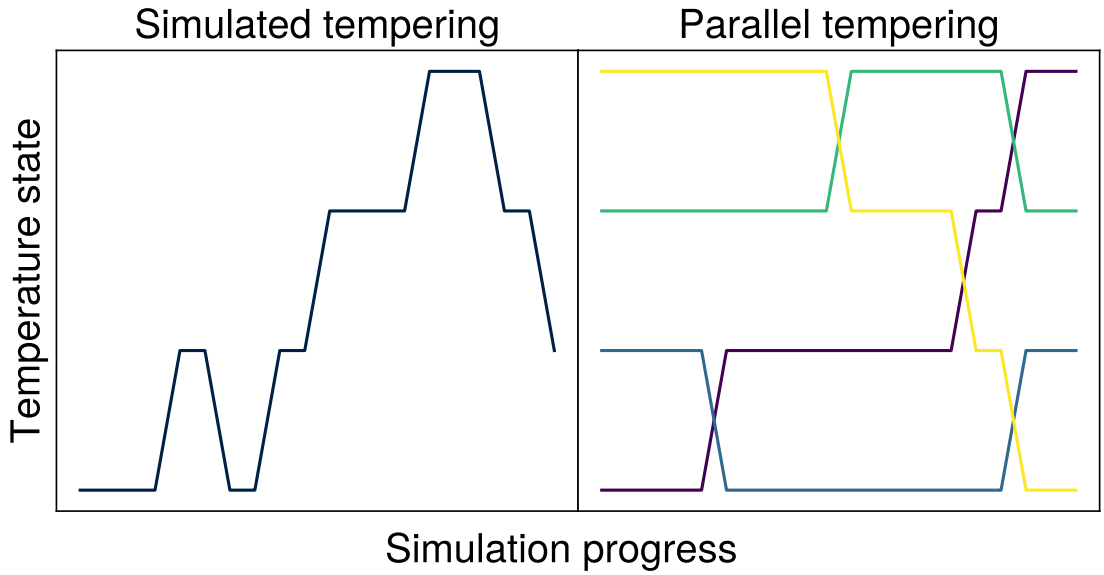

Figure 2.8: Overview of how simulated tempering and parallel tempering collect samples from a range of temperatures. In simulated tempering, a single system is simulated and moves to adjacent rungs on a temperature ladder are periodically attempted. Parallel tempering simulates one replica per temperature and attempts coordinate exchanges between adjacent temperatures at certain intervals.

2.2.1 Theory

In a regular MD simulation, the probability P of observing the system in a particular configuration **x** is proportional to the Hamiltonian of the system in its particular ensemble via

$$P(\mathbf{x}) \propto e^{-H(\mathbf{x})}.$$

The expanded ensemble quite literally expands this proportionality by providing more ensembles n to be sampled from, and $P(\mathbf{x})$ becomes

$$P(\mathbf{x}, n) \propto e^{-H(\mathbf{x},n)}. \tag{2.8}$$

The different states n can vary in a number of ways as described above. In the case of simulated tempering, the states differ in temperature. For example, a simulation in the NPT ensemble could be expanded by a series of NPT ensembles at different constant temperatures T, which are sampled by changing the reference temperature of the thermostat. The Hamiltonian in Eq. 2.8 is then given by

$$H(\mathbf{x}, n) = \beta_n H(\mathbf{x}) - g_n, \tag{2.9}$$

where $\beta_n = 1/k_B T_n$ is the inverse temperature of state n and g_n is the weight of the state. The temperatures of each state remain constant throughout the simulation, while the weights of each state are updated throughout the simulation to ensure equal sampling across all conditions.

A simulation runs in a certain state n for a set number of simulation steps, after which a move to the adjacent states $n+1$ and $n-1$ is attempted and accepted or rejected with a Metropolis criterion defined as

$$w\left(T_n \rightarrow T_{n\pm1}\right) = \min\left(1, e^{\left(\beta_{n\pm1} - \beta_n H(\mathbf{x}) - (g_{n\pm1} - g_n)\right)}\right). \tag{2.10}$$

In this way, a single simulation can visit all temperature states if suitable weights g_n are chosen. Importantly, correct ensembles are sampled and equilibrium and detailed balance are maintained by this approach. The Hamiltonian $H(\mathbf{x})$

contains kinetic and potential energy contributions and thus depends on the velocities and positions of each atom in the system, which have to be available at the same time as a move is attempted. For this reason, the velocity-verlet integrator (Section 2.1.2) is used for expanded ensemble simulations.

2.2.2 Obtaining initial weights

The weights g_n are updated throughout the simulation until all states are evenly sampled. The weights are then equal to the energy difference between the states. This relation is used to help choose suitable starting weights, which in turn helps accelerate the weight update procedure. The following approach was used in this thesis after first being described by Park and Pande[157].

Based on the facts that the weights are related to the energy difference between the states, and that optimal weights should result in equal probabilities of the moves $n \rightarrow n+1$ and $n+1 \rightarrow n$, Park and Pande obtained initial guesses for the weights g_n. Equal probabilities can only be achieved if the differences between the Hamiltonians are equal, i.e. $\Delta H_{n\rightarrow n+1} = \Delta H_{n+1\rightarrow n}$. Approximating the Hamiltonian through the potential energy U of the system gives

$$\Delta H_{n\rightarrow n+1} \approx \Delta\beta_{n\rightarrow n+1} U_n - \Delta g_{n\rightarrow n+1}$$

and

$$\Delta H_{n+1\rightarrow n} \approx \Delta\beta_{n+1\rightarrow n} U_{n+1} - \Delta g_{n+1\rightarrow n},$$

which, when equated, give an expression for the approximate relative weight between two states as

$$\Delta g_{n\rightarrow n+1} \approx (\Delta\beta_{n\rightarrow n+1}) \frac{U_n + U_{n+1}}{2}. \tag{2.11}$$

An initial guess of the weights g is obtained by running short simulations — 500 ps per temperature suffice — at each of the temperatures and calculating the average potential energy U_n of the system at condition n. The relative weight between two adjacent states is calculated from the temperature difference and the mean of the potential energies through Eq. 2.11. Since only relative weights

are needed for the Metropolis criterion in Eq. 2.10, the absolute values of the weights are of no consequence. However, since each state must have a weight value assigned to it in the GROMACS implementation, the first weight should be set to an arbitrary value, with the remaining values being calculated from the relative weights obtained through Eq. 2.11.

2.2.3 Updating the weights

The initial guess of the weights is then updated throughout the simulation. This is done using the Wang-Landau algorithm[158], which works by building a histogram h_n counting how often each state n is visited. At each attempted move t, the histograms and weights are updated by incrementing the histogram count of the new state and by adding the Wang-Landau increment δ to its weight g as

$$h_n^{t+1} = h_n^t + 1$$

and

$$g_n^{t+1} = g_n^t + \delta.$$

The increment δ is a positive energy value (in units of $k_B T$) with the purpose of encouraging the system to move away from the current state. By slightly increasing the energy of state n, moves to state $n \pm 1$ are made more likely. Over time, this causes the simulation to visit all states more or less equally, yielding a flatter histogram h_n. Flatness of the histogram is judged by comparing the ratio of the histogram counts h_n to the mean $\bar{h} = \sum_n h_n/N$ with the Wang-Landau ratio η and the total sum of histogram counts N. When the criterion $\eta < h_n/\bar{h} < \eta^{-1}$ is met for all states n, the histogram is considered flat, the Wang-Landau increment δ is scaled by a factor $\alpha < 1$, and the histogram counts are reset. The procedure is then repeated until δ shrinks below a user-defined threshold, at which point the weights are considered equilibrated and are no longer updated. At this point, all states are equally likely to be sampled.

2.3 Umbrella sampling

The free energy surface of a biomolecular simulation system is highly multidimensional and exceptionally complex, making free energy changes in processes of interest difficult to calculate. The dimensionality of the free energy surface can be reduced by choosing adequate collective variables (CVs) to describe the process of interest. Suitable choices of collective variables depend on these processes — the distance of a ligand from a membrane bilayer, a torsion about a bond, or angles between protein subdomains are all possible examples, as are combinations of such factors. The free energy surface along such collective variables, known as the potential of mean force (PMF), is more readily calculable.

However, 'more readily' still means complicated in all but the simplest cases. Calculating the PMF of a process from an unbiased simulation is problematic because the probability of sampling higher energy regions on the potential energy surface is naturally much lower than that of regions of lower energy, meaning that less data is available on energy barriers. As a result, unbiased simulations usually do not generate enough data to accurately calculate PMFs.

Umbrella sampling aims to overcome this issue of data generation in high-energy regions. This is done by taking the region of interest of the reaction coordinate and dividing it into a discrete set of windows. One simulation system per window is then generated. For example, to calculate the PMF of a full rotation around the central C-C bond in butane, 24 windows could be generated in which the torsional angle is offset by 15° each to cover a full rotation of 360°. A restraint potential is then applied to ensure that each window samples its assigned region of the reaction coordinate. The most commonly used functional form of this restraint potential is that of a harmonic restraint. For each simulation window, additional potential energy terms U_{bias} that result from this harmonic restraint enter into the force calculation as

$$U_{bias}(\xi) = \frac{k_{bias}}{2} \left(\xi - \xi_0\right)^2 , \tag{2.12}$$

where k_{bias} is the force constant of the biasing potential, ξ is the reaction coordinate and ξ_0 is the centre point of the umbrella window to which the reaction coordinate is restrained. While these umbrella potentials ensure that data points from all regions along the reaction coordinate are collected (Fig. 2.9), the PMF of interest is that of the unbiased system. The biases applied to improve sampling thus complicate calculation of the PMF. They need to be accounted for and their effects have to be removed from the raw data before the PMF can be obtained.

This section will briefly discuss current approaches to umbrella sampling, including how to obtain the starting configurations, how to calculate the PMFs from the simulation data, and how to assess convergence.

2.3.1 Window generation methods

The reaction coordinate is divided into smaller windows and the system is restrained to sample these subregions. Before data collection in each window can begin, the system must be made to adopt the desired configuration. In very simple cases no special care is required to generate appropriate starting conditions. Umbrella sampling simulations of the backbone torsions of the alanine dipeptide, for example, can all be started from a single input structure. The system will then quickly adopt the configuration to which it is restrained, and the simplicity of the system means undesired structural distortions are unlikely.

In larger systems, more control over the starting configurations is generally required. A standard approach is to use a non-equilibrium method like steered MD to force the system to move along the reaction coordinate. Starting configurations for umbrella sampling simulations can then be extracted from the resulting trajectory. This approach has successfully been used in a number of studies[159,160], but it can be problematic. The forces exerted on the system during non-equilibrium pulling can cause deformations that alter the chemical environment and can thus influence the PMF. The PMF therefore depends not only on the reaction coordinate but also on the parameters of the pull. As a consequence, a selection of pull

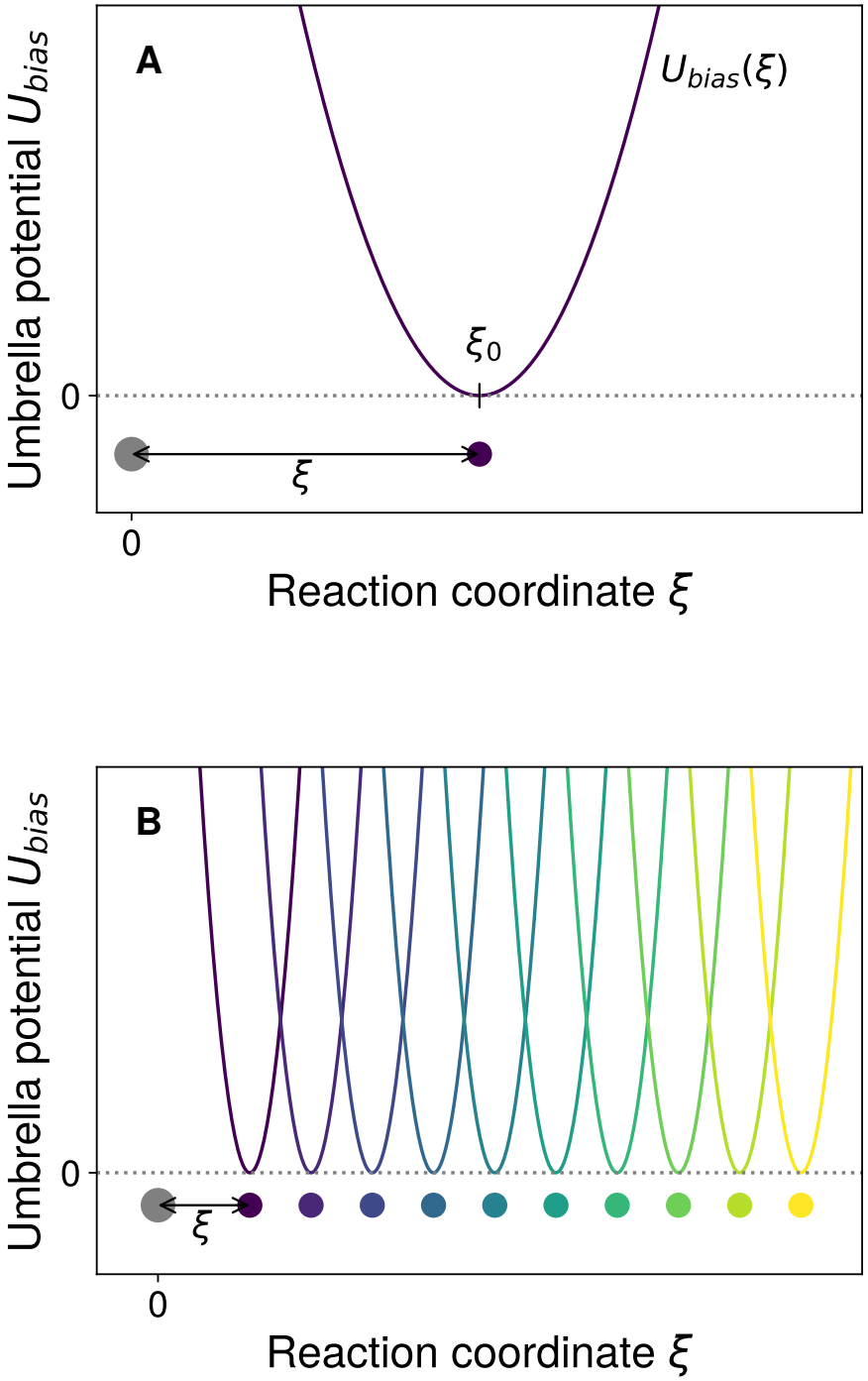

Figure 2.9: Overview of the role of U_{bias} in umbrella sampling, explained for the example of a simple one-dimensional distance. A: The system is restrained to a reaction coordinate of ξ_0 by $U_{bias}(\xi)$ (Eq. 2.12). As the distance ξ deviates from ξ_0, a harmonic potential is applied. B: To sample all regions along the reaction coordinate, independent windows with individual ξ_0 values are set up and run. The system is thus restrained to different values along ξ and samples from all regions are collected.

parameters such as pull speed or direction should be tested and the resulting PMFs compared to one another.

An alternative approach was used by Cetin et al.[161] in a recent study to calculate PMFs of ions passing through the pore of the nicotinic acetylcholine receptor. To obtain configurations with ions at the desired position along the reaction coordinate, the authors manually modified the coordinates of an ion to place it onto its starting point along the pore axis. Since space at this location would already be occupied by water molecules, this manual placement was followed by two rounds of energy minimisation during which strong position restraints were placed on the ion. Overlapping water molecules were thus pushed out of the way and starting configurations for umbrella sampling were obtained. One advantage of this method is that it is not sensitive to a pull direction as the steered MD approach is. It is only feasible for very small molecules of interest, however, as larger molecules would have more overlap with other atoms which could not be resolved by simple energy minimisation.

2.3.2 Choice of umbrella potentials

The umbrella potential restrains the system to a subsection of the reaction coordinate. Both its strength and the number of windows have to be chosen carefully because the calculation of the PMF depends on sufficient overlap between neighbouring umbrella windows. If the force constant of the harmonic restraint is set too high, or if the windows are spaced too far apart, the regions sampled by adjacent windows will not overlap sufficiently and the calculated PMF will be incorrect. If the harmonic restraint is too weak, however, it might not be able to restrain the system to regions of high energy along the reaction coordinate, causing both problems of overlap between windows and coverage of the reaction coordinate. The ideal umbrella potential is thus strong enough to restrain the system to every point along the reaction coordinate while not being too strong to prevent sampling a broad enough range around the reference value of the restraint (ξ_0 in Eq. 2.12).

2.3.3 Calculation of the potential of mean force

During the simulation of each umbrella window, the values of the reaction coordinate are saved to a file. These coordinate values ξ combined with the centre of the restraint ξ_0 of the window and the strength of the umbrella potential k_{bias} allow the calculation of the force necessary to restrain the system to a particular configuration along the reaction coordinate. Based on this information from biased simulations, numerical approaches can determine the unbiased free energy surface along the reaction coordinate *without* the restraints. Two such approaches will briefly be discussed here.

Weighted histogram analysis method

The application of the weighted histogram analysis method (WHAM) to umbrella sampling was first described by Kumar et al.[162]. It works by building a histogram $h_i(\xi)$ from each of the N_w biased umbrella simulations, splitting the reaction coordinate into discrete bins. These histograms are the results of biasing the underlying probability distribution by $U_{bias}(\xi)$, and the goal is to obtain the original unbiased probability distribution $P(\xi)$. This is done by iteratively solving the WHAM equations,

$$P(\xi) = \frac{\sum_{i=1}^{N_w} h_i(\xi)}{\sum_{j=1}^{N_w} n_j e^{-\beta\left(U_{bias,\ j}(\xi) - f_j\right)}}$$

and

$$e^{-\beta f_j} = \sum_{\xi} P(\xi) e^{-\beta U_{bias,\ j}(\xi)},$$

where n_j is the number of data points in histogram h_j and the N_w f_j are free energy constants. After solving this equation system, the PMF can be calculated from the unbiased probability distribution via

$$\mathrm{PMF}(\xi) = -k_B T \ln \frac{P(\xi)}{P(\xi_0)},$$

where ξ_0 is an arbitrarily chosen reference point at which the PMF is set to zero.

WHAM is a computationally efficient method, but its reliance on binning can be problematic as it introduces a bias dependent on user input.

Multi-state Bennet acceptance ratio

The multi-state Bennet acceptance ratio (MBAR) of Shirts et al.[163] is a newer estimator that does not rely on binning the data along a reaction coordinate and instead uses the raw data directly. It can be considered an application of WHAM with a bin width of zero.

MBAR works directly on the configurations **x** generated in the K umbrella windows by considering the reduced potential energy $u_i(\mathbf{x})$ of these configurations. The reduced potential energy is defined as

$$u_i(\mathbf{x}) = \beta_i \left(U_i(\mathbf{x}) + p_i V(\mathbf{x}) + \mu_i n(\mathbf{x}) \right),$$

with β_i being the inverse temperature of state i. U_i is the potential energy function of the system and here includes the umbrella potentials U_{bias} of window i. p_i is the pressure of state i and μ_i the vector of chemical potentials, which are needed for simulations run in the NPT and μVT ensemble, respectively. Each of the K umbrella windows contributes N_k pairs of independent configurations **x** and corresponding reduced potential energies u_i. From these, the free energies f_k of each of the K states are calculated by iteratively and self-consistently solving the MBAR equations

$$\hat{f}_i = -\ln \sum_{j=1}^{K} \sum_{n=1}^{N_j} \frac{e^{-u_i(\mathbf{x}_{jn})}}{\sum_{k=1}^{K} N_k e^{\hat{f}_k - u_k(\mathbf{x}_{jn})}}.$$

Once the free energies of each of the biased states are known, the PMF can be calculated by reweighting to the unbiased state. At this point, MBAR also allows straightforward temperature reweighting, enabling the use of data from simulations carried out at different temperatures to inform the PMF at a desired temperature. One downside of this is that it can quickly become computationally expensive, with available memory becoming the limiting factor to how many states can be included.

2.3.4 Convergence criteria

It is desirable to keep simulations as short as possible to save computational resources. So when is an umbrella sampling simulation long enough? In general, a sensible end point to aim for is convergence of the PMF profile. Convergence can be assessed by calculating the PMF over subsets of the available simulation data. Starting with few data points and gradually including more data in the PMF calculation should initially lead to stark changes in the PMF. The calculation approaches convergence once inclusion of more data points no longer significantly changes the PMF. It is important to note, however, that convergence does not necessarily imply correctness of the PMF, which makes other criteria useful. These criteria depend on the exact nature of the system at hand. In the case of the translocation of a small molecule through a symmetrical lipid bilayer, for example, one would expect the PMF in bulk solvent on either side of the membrane to have the same value, and for the PMF to reflect the symmetry of the system. If these criteria are not met, longer simulations or enhanced sampling schemes should be considered.

2.4 Homology modelling

Homology modelling is a technique used to generate a target protein structure from its sequence based on a sequence alignment to a related protein of known structure. It is a useful technique when structures of closely related proteins are available, for example of a homologue from a different species with high sequence identity. Even though artificial intelligence-based structure prediction tools like AlphaFold[164] have found widespread use in recent years, homology modelling remains an important technique.

In this work, the MODELLER software[165] was used to build connecting loops into unresolved regions in structures obtained through cryo-electron microscopy. For this purpose, target and template sequence only differ in the addition of the

connecting linker, and care was taken that atoms resolved in the structure remain invariant during the relaxation steps run by MODELLER.

2.5 Measuring of cavities in protein structures

The properties of solvent- and small molecule-accessible pockets, tunnels, pathways and pores in proteins are of special interest. Permeability of ion channels and accessibility of drug binding sites depend on their physical dimensions. One of the earlier tools for the calculation of ion channel radius profiles was HOLE[67]. It was originally intended for use on static structures, but has recently been adapted for use on trajectories as part of an effort to incorporate HOLE2 in MDAnalysis[166,167]. The Channel Annotation Package CHAP[168] is a more powerful tool that natively handles GROMACS trajectories and calculates pore hydrophobicity and water number density as well as pore radius to allow study of the phenomenon of hydrophobic gating.

HOLE and CHAP are well-suited to the study of pores, but they struggle when one end of the cavity of interest lies within the protein. CAVER[169] is useful in these cases. It employs a Voronoi diagram-based approach to describe tunnels connecting the surface of the protein to regions in its core. Originally intended for the study of buried ligand binding sites, it has been used for the description of fenestrations in this work. These are tunnels that connect the membrane environment of an ion channel directly to its ion conducting pore.

The accessible volume of ligand binding pockets is an important readout in ligand binding studies. In this work, PyVOL[170] was used to measure pocket volumes at different stages of a simulation pipeline.

3

A novel umbrella sampling workflow

Contents

3.1 Introduction

Most drugs have to pass cell membranes on the way to their site of action[171–173]. Hydrophobicity and membrane partitioning behaviour strongly influence a drug's absorption, distribution, metabolism and excretion properties (ADME) and need to be considered early on in drug development[174,175]. The importance of these factors was distilled into Lipinsky's "rule of 5"[94], which offers criteria to predict membrane permeability based on a compound's size, number of hydrogen donor and acceptor groups, and octanol-water partition coefficient (logP). logP values are useful predictors and can readily be calculated from chemical structures. However, they are often insufficient determinants of membrane permeation[176,177]. Membranes are highly complex assemblies whose precise composition can vary between species and cell types[178]. Lipids affect overall membrane behaviour through their different head groups, fatty acid chain lengths, degrees of saturation, and rigidity. It is thus no surprise that octanol-water partition coefficients often do not adequately represent membrane partitioning behaviour.

Molecular dynamics simulations have been used in the past to more accurately study the effects of membrane composition on solute partitioning[91,179], but slow convergence of free energy calculations has been an issue in the past. It is thus common practice to enforce symmetry and/or periodicity of the potential of mean force (PMF) as a boundary condition. However, there are undoubtedly errors that arise from these constraints[180,181]. As their magnitude is hard to determine, these errors are difficult to account for, meaning that the use of these constraints should ideally be avoided.

In this chapter, I develop a workflow for umbrella sampling simulations that shows accelerated convergence of PMFs and reduced hysteresis than commonly used methods. I benchmark its three components — non-standard window generation, Simulated Tempering-enhanced Umbrella Sampling[182], and temperature reweighting for maximal data utilisation — individually to determine the

effectiveness of each by calculating the PMFs of permeation of eight different small molecules through two different lipid bilayers. A focus lies on obtaining symmetrical and periodic PMFs without the need to enforce these characteristics as boundary conditions.

3.2 Systems and molecules studied

This section will introduce the two lipid bilayer systems and the eight small molecules that were simulated to create the dataset for this study.

3.2.1 Lipid bilayers

Permeation through two different lipid bilayers was tested in this study: a pure POPC bilayer, and a cholesterol-doped POPC bilayer with 30 mol% cholesterol (Fig. 3.1 and Table 3.1).

Table 3.1: Composition of the two membrane systems investigated in this study.

System	POPC	Chol-doped POPC
# POPC	48	42
# Chol	—	18
# Na^+	7	8
# Cl^-	7	8
# H_2O	2916	3270
# Atoms	15194	16786
Box size [nm]	4 x 4 x 10	4.25 x 4.25 x 10

POPC

1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC, Fig. 3.2) is a phosphatidylcholine phospholipid.

Phosphatidylcholine lipids are the most common class of phospholipids in mammalian cell membranes[183]. POPC is a useful model lipid because it combines a *cis*-unsaturated (oleic acid, 18:1) and a saturated (palmitic acid, 16:0) fatty acid, giving it intermediate behaviour in terms of membrane rigidity and fluidity compared

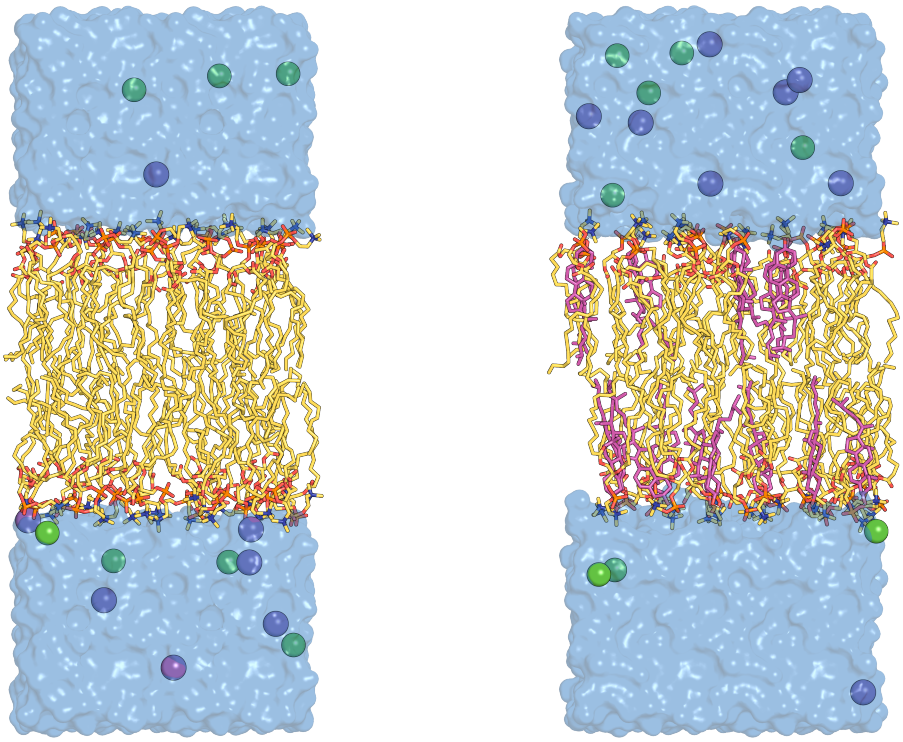

Figure 3.1: Visualisation of the two simulation systems used in this study. The bilayers are formed of POPC (yellow) and cholesterol (magenta). They are surrounded by water molecules (blue surface) and 0.15 M sodium chloride (purple and green spheres, respectively).

with fully saturated and poly-unsaturated lipids[184]. While fully saturated DPPC with two palmitoyl groups only undergoes phase transition at 41.3 °C, POPC has a gel-to-liquid crystalline (L_β-to-L_d) phase transition temperature of −2.5 °C[185]. Since many membranes are known to display liquid-crystalline behaviour overall[186] and phosphatidylcholines are the most common phospholipid in the membrane, pure POPC bilayers can be appropriate simplified membrane models. They are thus commonly used in MD studies[187–190].

Cholesterol-doped POPC

Cholesterol (Fig. 3.3) is a sterol lipid. It has a polar head group in the form of a hydroxyl group, which is attached to a rigid tetracyclic structure that also

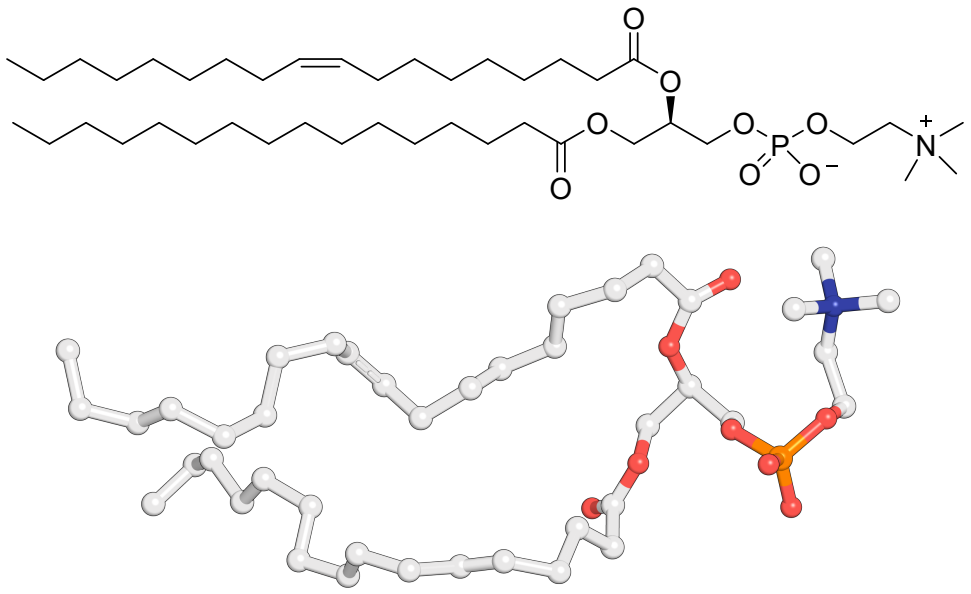

Figure 3.2: Structure of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC). The phosphocholine head group is shown on the right side of either representation, the oleyl component makes up the upper left section, and the palmitoyl component is on the bottom left.

carries an aliphatic chain. Cholesterol constitutes 20 to 50 mol% of mammalian membranes[191–193]. As such a significant component, it strongly affects membrane behaviour. It intercalates between the phospholipids, orienting its hydroxyl group toward their head groups (Fig. 3.1). In this arrangement, it enforces increased order of the phospholipid tails through the rigidity of its sterol tetracycle, while the lateral diffusion seen in liquid-crystalline (L_d-phase) bilayers is maintained. This causes the bilayer to form a distinct liquid-ordered (L_o) phase. The enforced order on the lipid tails increases membrane thickness and reduces the area per lipid, increasing the barrier of permeation of solutes[191,194].

In this study, a POPC bilayer with 30 mol% cholesterol was used as a consensus value from the range of cholesterol contents that is reported for various cell types.

3.2.2 Compounds

Eight compounds of varying chemical complexity were tested in this study.

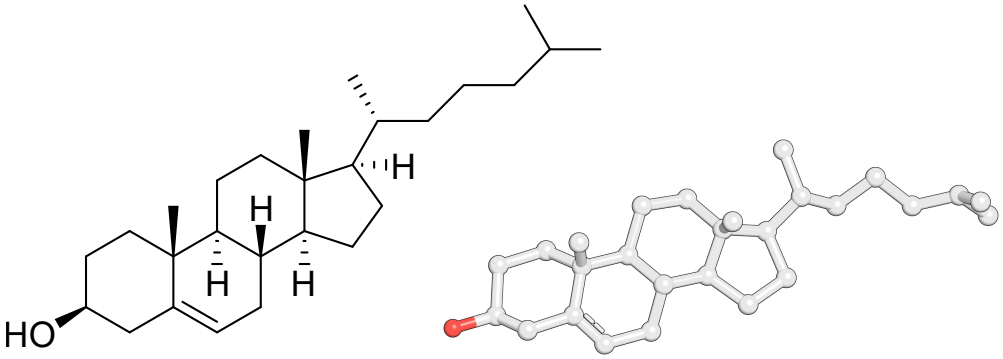

Figure 3.3: Chemical structure and liquorice model of cholesterol. All hydrogens are omitted from the model representation for visual clarity.

Benzene

The simplest molecule investigated in this study is benzene (Fig. 3.4). It is included as a baseline compound: small and rigid, it should sample water and membrane phases well. A dependence of its partition coefficient based on membrane composition has previously been reported[195], with reduced partition coefficients at higher cholesterol concentrations. Thus, a difference in PMFs collected in the POPC and POPC+Chol systems is expected.

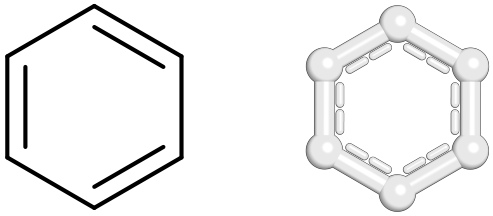

Figure 3.4: Chemical structure and liquorice model of benzene

Lindane

Lindane (γ-hexachlorocyclohexane, Fig. 3.5) is a neurotoxin that was used as an insecticide until its agricultural use was banned in 2009[196]. It is an inhibitor of insect GABA receptors[197,198]. The reduction of its partitioning coefficient with

increasing cholesterol content of membranes was described by Antunes-Madeira and Madeira[199].

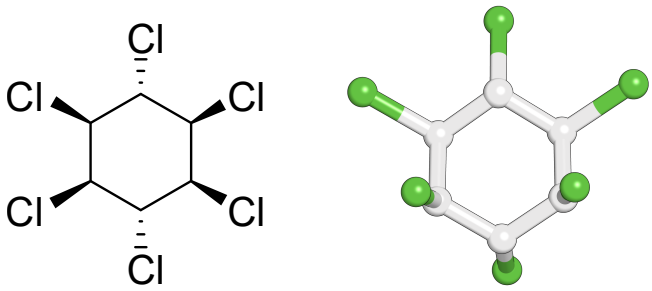

Figure 3.5: Chemical structure and liquorice model of lindane

Isoflurane

Isoflurane (Fig. 3.6) is an inhalational anaesthetic. Its precise mode of action is unclear, though it has been shown to activate glycine and GABA receptors[200], TRPA1[201], and RyR1[202], and inhibit ELIC[203] and voltage-gated sodium channels[204]. Dickinson et al. described reduced partition coefficients of isoflurane in bilayers that contain cholesterol[205].

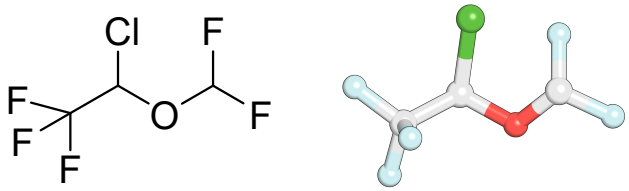

Figure 3.6: Chemical structure and liquorice model of isoflurane

Thiopental

Thiopental (Fig. 3.7) is a barbiturate modulator / agonist of $GABA_A$ receptors and has seen use as a general anaesthetic[206], for lethal injections[207], and as "truth serum"[208]. Its partition coefficient was determined for different membrane compositions by Korten et al[209]. They found partition coefficients to decrease

with increasing cholesterol content. The secondary amines act as weak acids, and the compound is commercially available as its sodium salt. However, with a pK_a value of 7.5[210], more than half the molecules are expected to be protonated and uncharged at physiological pH. It is this uncharged form of the drug that was tested in this study.

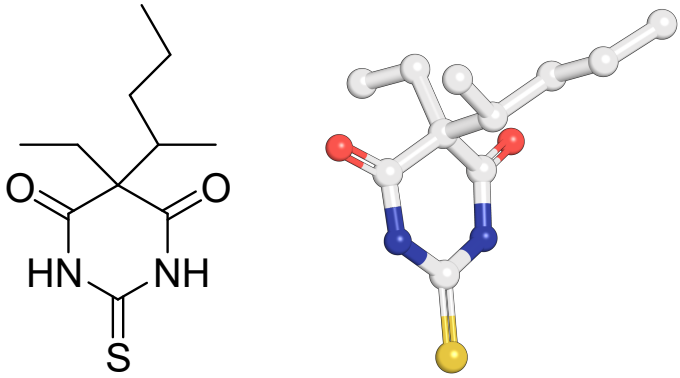

Figure 3.7: Thiopental

Fluocinolone Acetonide

Fluocinolone acetonide* (Fig. 3.8) is a corticosteroid used to treat skin conditions. Takegami et al demonstrated that its partition behaviour is influenced by cholesterol[177].

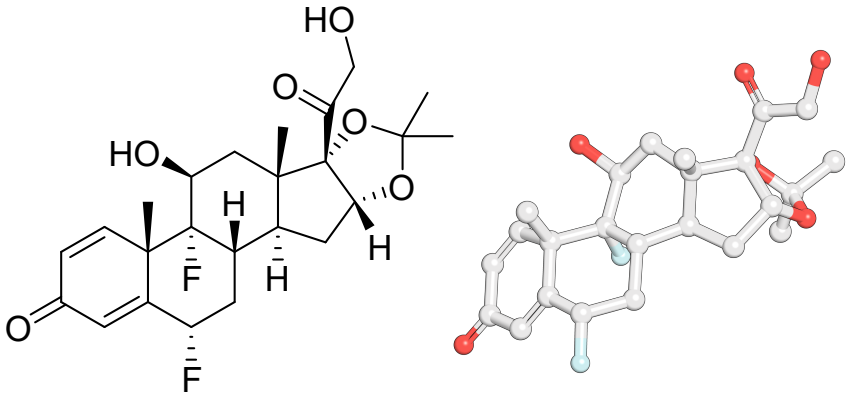

Figure 3.8: Chemical structure and liquorice model of fluocinolone acetonide

*Throughout this study, "fluocinolone" is used as a shorthand for fluocinolone acetonide.

Chlorpromazine

Chlorpromazine (Fig. 3.9) was one of the first antipsychotic drugs. Its adaptation was a strong contributor to the phasing out of lobotomy procedures[211]. Luxnat et al. determined partition coefficients of chlorpromazine drug with different membrane compositions and found a strong effect of cholesterol content[212].

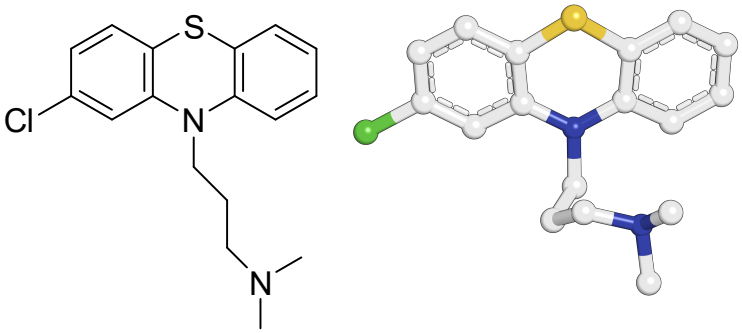

Figure 3.9: Chemical structure and liquorice model of chlorpromazine

Accurately capturing the protonation state of chlorpromazine is a challenge, as it is an amphipathic compound and can be protonated at the terminal amino group. A vast range of pK_a values of the molecule is reported, with values ranging from around 9.3 in aqueous solution[213] to in 7.8 Triton X-100[214]. Its logP value of 5.4[213] shows that the hydrophobic features of chlorpromazine dominate its overall behaviour. However, it is possible that chlorpromazine could integrate into a lipid bilayer like a phospholipid, with its hydrophobic component embedded in the core of the membrane and its charged head group at the membrane-water interface. This would make membrane permeation unlikely while still allowing a high logP value. Alternatively, counter ions could move with the compound, allowing it to pass through the bilayer[215]. In this study, I have chosen to simulate the deprotonated form of chlorpromazine. This decision is based on a study by Elferink carried out on chlorpromazine and a structural relative with a quaternary ammonium instead of the tertiary amine head group[216]. Elferink found that unlike the quaternary ammonium derivative, chlorpromazine is able to permeate through lipid bilayers. Since the quaternary ammonium derivative would be able to associate with a

counter ion, but cannot lose its charge outright, this suggests that loss of the charge is the determinant for membrane permeability, and that chlorpromazine is deprotonated while inside the bilayer. A likely model for chlorpromazine membrane permeation follows from pK_a value determinations by Chen et al.[214]: protonated chlorpromazine embeds into the outer layer of the membrane with its hydrophobic core while its head group is oriented towards the membrane-water interface. In this new environment, it has a lower pK_a value of 8.0, thus suggesting that about 25 % of all molecules should be deprotonated. This deprotonated form is then able to cross the bilayer. However, fine-tuning protonation states of these compounds is beyond the scope of this work. As the focus lies on the impact of cholesterol content on permeation PMFs, the most likely protonation state to encounter cholesterol — the deprotonated state — was chosen for all windows. It must be noted, however, that this is a clear limitation of this study, and further work would be required to arrive at values that match experimental predictions.

Paroxetine

Paroxetine (Fig. 3.10) is a selective serotonin reuptake inhibitor-type (SSRI) antidepressant. It is a base with a pK_a of 9.9[217], meaning it should be protonated at physiological pH. However, Ngo et al. have demonstrated that paroxetine partitions into large unilamellar vesicles formed of DOPC[218]. Though there is less evidence than for chlorpromazine, it appears likely that similar considerations on the protonation state apply here. As such, paroxetine was modelled in the deprotonated state to capture the most likely interaction between it and cholesterol. Again, it has to be acknowledged that this simplified treatment of protonation states is problematic and limits how meaningful comparisons to experimental data are. This trade-off was made here as the focus of this study is on enhanced sampling and improved convergence in umbrella sampling rather than the reporting of accurate partition coefficients.

Ngo et al. also found paroxetine's partitioning into membranes to increase with temperature and decrease with cholesterol content[219].

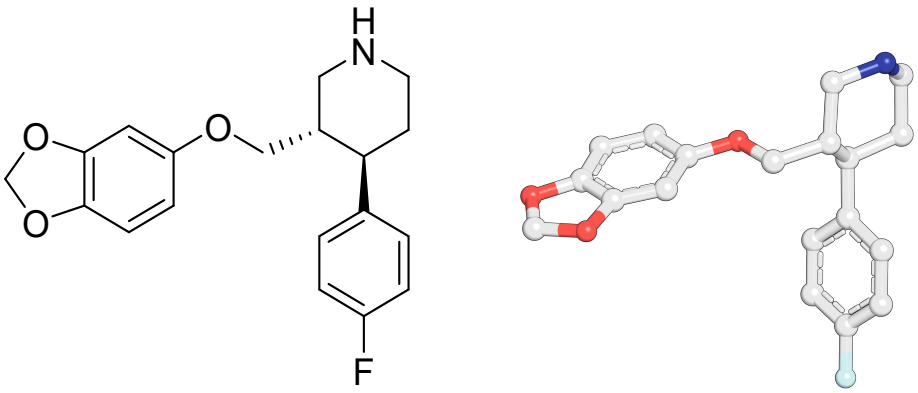

Figure 3.10: Chemical structure and liquorice model of paroxetine

Sertraline

Sertraline (Fig. 3.11), like paroxetine, is an SSRI-type antidepressant. It was the second compound studied by Ngo et al. and was also found to have temperature-dependent and cholesterol-dependent changes of its partition coefficient[219]. In spite of its pK_a of 9.16[220], it was modelled in the neutral, deprotonated state based on evidence of membrane partitioning[218] and following the same line of argument as for chlorpromazine and paroxetine.

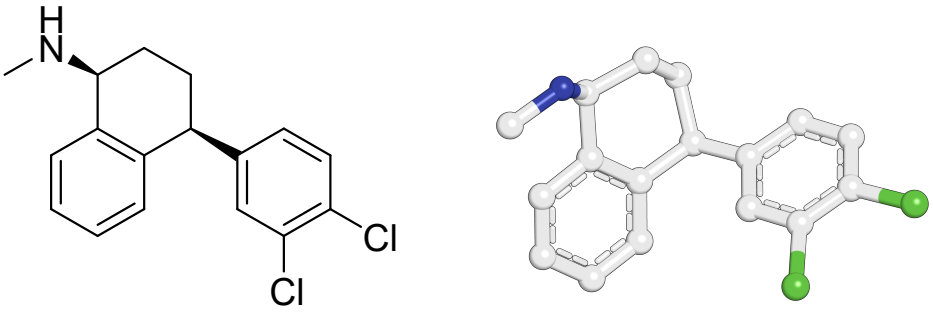

Figure 3.11: Chemical structure and liquorice model of sertraline

3.3 Methods

For all systems, the umbrella sampling reaction coordinate was defined as the centre-of-mass distance between membrane and small molecule (Fig. 3.12).

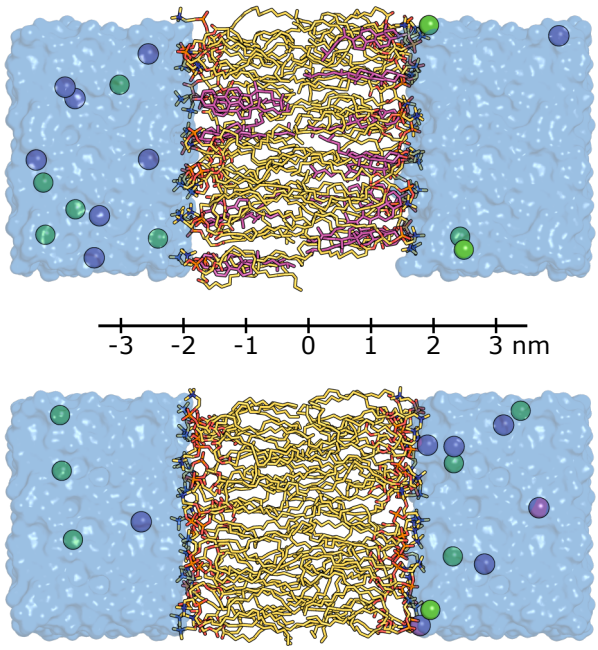

Figure 3.12: Illustration of the reaction coordinate used in this study. Top: Cholesterol-doped POPC bilayer. Bottom: Pure POPC bilayer.

3.3.1 Shared MD parameters

For this study, the chosen force fields are SLipids for POPC and cholesterol[221,222], TIP3P for water[223], and GAFF2 for the eight small molecules[224]. The LINCS algorithm[147] was used to constrain bonds to hydrogen in all molecules except for water, which was constrained with SETTLE[148]. A timestep of 2 fs was used for integration of the equations of motion. Long-ranged electrostatic interactions were calculated with the smooth particle mesh Ewald approach[149], while a Verlet cut-off scheme was applied to short-ranged interactions. A pressure of 1 bar was maintained in all simulations using a semi-isotropic C-rescale barostat[152] with a coupling constants of 5 ps. Temperature was maintained at variable values in a range from 310 to 358 K with the V-rescale thermostat[151], as indicated later, using a 1 ps coupling constant. During all umbrella sampling simulations, EDR energy files and reaction coordinates were written every 0.1 ps. The centre of mass-distance in the membrane-normal direction between the lipid bilayer and the small molecule was chosen as the reaction coordinate. An umbrella potential

of 1000 kJ mol^{-1} nm^{-2} was applied to harmonically restrain the system to certain values along this reaction coordinate.

3.3.2 Bilayer preparation

The two bilayers used in this study — pure POPC and POPC with 30 mol% cholesterol (Fig. 3.1) — were assembled with the CHARMM-GUI Membrane Builder[225–229]. Final system contents are listed in Table 3.1. These systems were then energy minimised with the steepest-descent algorithm and equilibrated in a procedure that saw stepwise release of position restraints (Table 3.2).

Table 3.2: Overview of the parameters used in each equilibration step of the bilayer systems. Head group restraint force constants are in units of kJ mol^{-1} nm^{-2}, while the dihedral restraints have force constants in units of kJ mol^{-1} rad^{-2}

Step	n_{steps}	timestep [fs]	Restraint force constants	
			head groups	POPC tail dihedrals
1	125000	1	1000	1000
2	125000	1	400	400
3	125000	1	400	200
4	250000	2	200	200
5	250000	2	40	100
6	250000000	2	0	0

Step 6 constitutes 500 ns of simulation without any restraints or bias on the system to ensure proper equilibration and phase behaviour of the lipid bilayer. The output of this equilibration step is then used to generate the individual umbrella windows.

Cholesterol restraints

During the analysis of the initial set of simulations on the cholesterol bilayer, bilayer asymmetry as a result of cholesterol flip-flopping was observed in some umbrella windows. An MDAnalysis-based analysis utilising `LeafletFinder`[166] was employed to quantify flip-flopping. This analysis confirmed that a significant degree of cholesterol flip-flopping occurs, particularly at the higher temperatures

sampled in STeUS simulations. Because of this, a new set of simulations with harmonic flat-bottom restraints was run. These restraints applied a force of 1000 kJ mol^{-1} nm^{-2} to the oxygen atom of the cholesterol head group if it moved more than 1.3 nm away from its starting position in the membrane-normal direction. These restraints still allowed cholesterol to move freely within and slightly beyond its leaflet while preventing a full flip-flop.

3.3.3 Window generation

The individual umbrella windows were obtained by generating conformations in which the ligands are placed along the reaction coordinate across the bilayers. The effectiveness of two different approaches is compared in this study.

Alchembed

In Section 2.3.1, I briefly described a window generation method employed by Cetin et al. in their study of nicotinic acetylcholine receptors[161]. The authors manually positioned ions along the channel pore and resolved van der Waals clashes with energy minimisation coupled with strong position restraints on the ion of interest. This approach works well, but is limited by the size of the group of interest, as larger molecules are more likely to have irresolvable overlap. We were inspired by this approach and work carried out by Jefferys et al. on Alchembed, a method for the embedding of proteins into lipid bilayers[230]. Here, the authors used a soft-core implementation of the Lennard-Jones (LJ) potential to gradually switch on van der Waals interactions between a protein and the membrane in which it was to be embedded. In combination, these two papers suggested the use of a soft-core LJ potential for the generation of umbrella sampling windows. We will refer to this method as Alchembed-based window generation.

Though it is possible to scale the standard implementation of the LJ potential by some alchemical variable λ as

$$U_{LJ}(r, \lambda) = 4\epsilon\lambda \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right], \tag{3.1}$$

where σ and ϵ are the LJ parameters of an interaction pair and r is the distance, this does not resolve the problem of atom overlap giving rise to forces that are too large for the integrator to handle. This is because the repulsive r^{-12} component assumes very large values even when multiplied by small scalar values of λ (Fig. 3.13 A). If a soft-core implementation of the form

$$U_{LJ,softcore}(r,\lambda) = 4\epsilon\lambda\left[\frac{1}{(\alpha(1-\lambda)+(r/\sigma)^6)^2} - \frac{1}{\alpha(1-\lambda)+(r/\sigma)^6}\right] \tag{3.2}$$

is used instead, scaling by λ yields a continuously and gradually increasing repulsive potential (Fig. 3.13 B). This gives rise to moderate forces which are handled well by the integrator and gently resolve atom overlaps. Here, α is a soft-core parameter that controls how quickly the scaled potential approaches the fully switched-on LJ potential with λ.

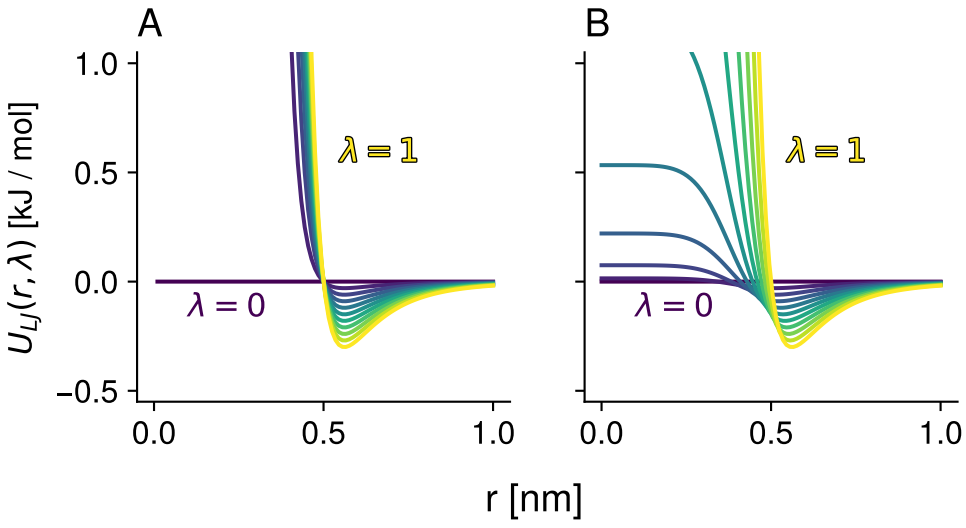

Figure 3.13: Visualisation of the effects of λ-scaling on the Lennard-Jones potential in its standard (A, Eq. 3.1) and soft-core (B, Eq. 3.2, with $\alpha = 1$) implementations. The repulsive contribution of the standard implementation is barely affected, while it scales continuously with λ in the soft-core implementation. At $\lambda = 0$ and $\lambda = 1$, both potentials are identical.

Experimental details To create umbrella windows with Alchembed, initial configurations are obtained by positioning the target compound roughly 3.75 nm above the membrane bilayer and, for each of 75 further windows, translating it

by 0.1 nm along the membrane-normal. The clashes in these 76 configurations are resolved by brief simulations of 1000 steps with a time step of 1 fs. In these simulations, the Lennard-Jones interactions of the target compound are coupled to λ. The simulations start at $\lambda = 0$, and λ was increased by $\Delta\lambda = 0.001$ each simulation step, while α was set to 0.1. Thus, λ reached 1 in 1000 steps. At this value, Lennard-Jones forces are fully switched on. When these simulations are coupled with harmonic position restraints on the target compound, they allow a gentle resolution of clashes while the reaction coordinate values are maintained (Fig. 3.14).

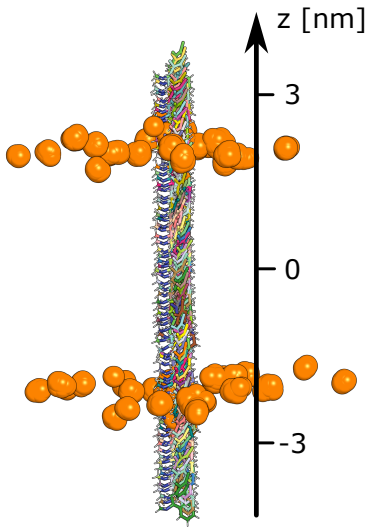

Figure 3.14: Alchembed-generated windows for sertraline. All atom overlaps are effectively resolved while the spacing along the reaction coordinate is well-maintained. Note that phosphorus atoms of all POPC molecules from all 76 windows are shown, but only very minor movements are necessary to resolve clashes.

We reasoned that this approach would reduce hysteresis artefacts that are often observed when umbrella sampling is started from steered MD simulations[231–233].

Steered MD

To evaluate the effectiveness of Alchembed-based window generation in the reduction of hysteresis artefacts, additional umbrella windows were generated

from a steered MD simulation for comparison. Starting from the first Alchembed-based window with the ligand roughly 3.75 nm above the membrane, 100 ns-long steers were started with a pull rate of 0.1 nm/ns. The pulls were automatically stopped once the reaction coordinate reached a value of more than 0.51 times the box vector to ensure the minimum image convention was maintained. During these simulations, flat-bottom restraints were applied to the ligands to restrain them to a cylinder parallel to the membrane-normal. The trajectory was analysed using MDAnalysis[166,167], dividing the reaction coordinate space visited in the simulation into equally spaced windows and writing out the frames of the trajectory closest to each of these windows. Artefacts of the pull are visible in slight disturbances of the membrane (Fig. 3.15).

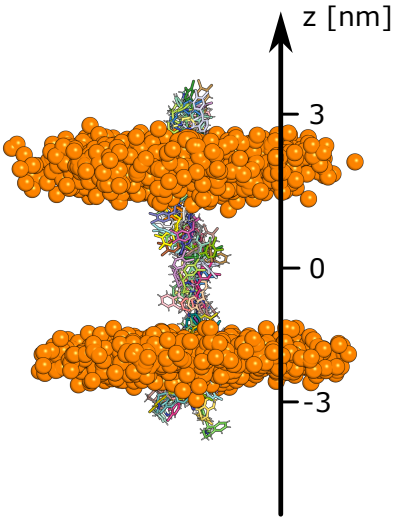

Figure 3.15: Steered-MD-generated windows for sertraline. During the steer, the ligand moves freely within the cylindrical restraint. As in Fig. 3.14, phosphorus atoms from all POPC molecules of all 76 windows are shown. The membrane appears thicker here for two reasons. The box volume fluctuates because the simulation is carried out in the NPT ensemble. This causes slight up-and-down movements of the membrane. Secondly, individual lipids can be influenced by the steer and are slightly pulled along with sertraline as it moves through the bilayer.

3.3.4 Data collection

For each compound, two different sets of windows were created as input for umbrella sampling simulations. Both sets of windows were then simulated under two different umbrella sampling protocols: standard umbrella sampling as implemented in GROMACS through the pull code, and Simulated Tempering-enhanced Umbrella Sampling (STeUS), a method that was developed by Sousa et al[182]. All umbrella windows were simulated for 100 ns.

Standard Umbrella Sampling

Standard umbrella sampling simulations were carried out at 310 K with the parameters described in Section 3.3.1. The leap-frog integrator was used to solve the equations of motion, which allowed offloading the constraint solution and coordinate update steps to the GPU for enhanced performance.

Simulated Tempering-enhanced Umbrella Sampling

The theory behind simulated tempering in general was explained in Section 2.2. In their 2023 paper *Simulated Tempering-Enhanced Umbrella Sampling Improves Convergence of Free Energy Calculations of Drug Membrane Permeation*, Sousa et al. report on an improved method to apply simulated tempering to umbrella sampling[182]. Simulated tempering is implemented natively in GROMACS as a special case of the expanded ensemble free-energy code and can be combined with the pull code to enable umbrella sampling windows to also sample a temperature ladder. The inclusion of high-temperature windows allows the system to overcome energy barriers more readily, which improves sampling both along the reaction coordinate and along orthogonal degrees of freedom which often hinder convergence[234]. However, while this approach does help overcome energy barriers, it at the same time reduces the number of data points available for the analysis. As the PMF is temperature-dependent[235,236], only data points collected at the same temperature can be combined to inform the PMF. Normally, simulated tempering simulations enforce even sampling of each temperature rung. This

means that for every additional rung, the number of data points sampled at the temperature of interest is reduced. This is where the contribution of Sousa et al. lies: the authors provide a patch for the GROMACS source code that modifies the implementation of the Wang-Landau algorithm.

This modification reads the content of the user-defined `SIMTEMP_OCCUP_STATE0` environment variable, in which the desired relative occupancy of the ground state (between 0 and 1) can be set. The updated algorithm ensures that the ground state is sampled at the desired, higher percentage, and all higher temperature states are sampled evenly. For example, for a temperature ladder with nine states and a specified ground state occupancy of 20 %, the ground state would be sampled at the specified 20 % and the eight higher-temperature states at 10 % each. It is this modification that makes simulated tempering useful for umbrella sampling, allowing both the overcoming of energy barriers and the collection of sufficient simulation data in the ground state.

Experimental details Before STeUS simulations can be run, initial guesses of the weights for each state have to be provided. As described in Section 2.2.2, these guesses can be obtained from the average potential energies extracted from short simulations of the system run at each of the temperatures.

As an example, to run a simulated tempering simulation that samples 310, 332.5, 355, 377.5, and 400 K, the following mdp options could be used.

```
free-energy = expanded
simulated-tempering = yes
sim-temp-low = 310   ; for example
sim-temp-high = 400  ; for example
simulated-tempering-scaling = linear
temperature-lambdas = 0 0.25 0.5 0.75 1  ; for example
init-lambda-weights = 0 200 400 600 800  ; for example
nstexpanded = 500
lmc-move = metropolis
```

```
lmc-stats = wang-landau
init-wl-delta = 1
lmc-weights-equil = wl-delta
wl-ratio = 0.7
wl-scale = 0.8
wl-oneovert = yes
integrator = md-vv
```

These commands instruct GROMACS to run an expanded ensemble simulation, namely the special case of a simulated tempering simulation. It will linearly interpolate between 310 and 400 K using `temperature-lambdas`. The initial guesses of the weights are provided as a vector to `init-lambda-weights` that has to match `temperature-lambdas` in length. Moves between adjacent temperatures will be attempted every 500 steps and are accepted or rejected based on the Metropolis criterion. The weights will be updated throughout the simulation using the (STeUS-modified) Wang-Landau (WL) algorithm with an initial WL-δ of 1 that is scaled by a factor $\alpha = 0.8$ when the histogram counts are sufficiently flat as compared to the WL-ratio $\eta = 0.7$ (see Section 2.2.3). To help prevent "burning in" to suboptimal weights, the WL-δ is set to scale with the reciprocal of the number of steps (or simulation time, hence "oneovert") in the large sampling limit. The velocity-verlet integrator has to be used for simulated tempering, which unfortunately limits some forms of GPU utilisation in current versions of GROMACS. This might be improved in future versions.

Unless otherwise stated, these mdp parameters were used throughout this study, with a temperature ladder of nine temperature steps that bridged 310 K and 358 K in steps of 6 K and a ground state occupancy of 20 or 40 %.

3.3.5 Calculation of PMFs

Statistical estimators are needed to obtain an estimate of the true underlying potential of mean force (PMF) from the finite, biased simulation data. In this

study, the Weighted Histogram Analysis Method (WHAM)[162,237] and the Multistate Bennet Acceptance Ratio (MBAR)[163] estimators were used.

WHAM

GROMACS's own WHAM implementation — `gmx wham`[180] — was used here to calculate the PMF based on the runs' TPR input files and the reaction coordinate output files. In the case of STeUS simulations, the output files were preprocessed and only the subset of data collected at the temperature of interest was used in WHAM.

To quantify the errors that result from the common strategy of enforcing symmetry and periodicity boundary conditions on the PMF, some PMF calculations were carried out with the `-sym` and `-cycl` flags passed to `gmx wham`.

MBAR

Data collected at higher temperatures can inform the PMF at the ground state temperature through statistical reweighting[238,239]. To evaluate the effectiveness of temperature reweighting to the STeUS simulations carried out here, the `pymbar`[163] implementation of MBAR was used to calculate PMFs at the ground state temperature using all available simulation data. The form of the reduced potential u_{kn} used for the MBAR calculations was

$$u_{kn} = \beta_k(U_{kn} + B_{kn} + pV_{kn}), \tag{3.3}$$

where β_k is the inverse temperature $1/k_B T_k$, U_{kn} is the potential energy, B_{kn} is the umbrella bias, p is the pressure (held at 1 bar by the barostat), and V_{kn} is the volume of the system. The indices n indicate the uncorrelated configurations of the system which are extracted from the trajectory after estimation of the autocorrelation time[240]. The indices k denote the simulation conditions, which are specified by the centre of the umbrella bias of each window and the temperature of the system. For a STeUS study that samples 9 temperatures and has 76 umbrella windows, there are $k = 9 \times 76 = 684$ states to consider in MBAR.

Effectiveness of temperature reweighting The potential energy of the system U is a key component of the reduced potential energy. If reweighting between different umbrella windows sampled at the same temperature is carried out, the βU terms cancel out. However, if different temperature states are sampled, the difference in potential energy between the two states has a large influence on the size of the weight. Systems with larger numbers of atoms have a steeper temperature dependence of the potential energy U (Fig. 3.16). The larger ΔU is, the smaller the weight of data collected at other temperatures becomes. Thus, if the system size is too large, the potential energy differences even between minor temperature steps are too large for temperature reweighting to meaningfully inform the PMF of the target state.

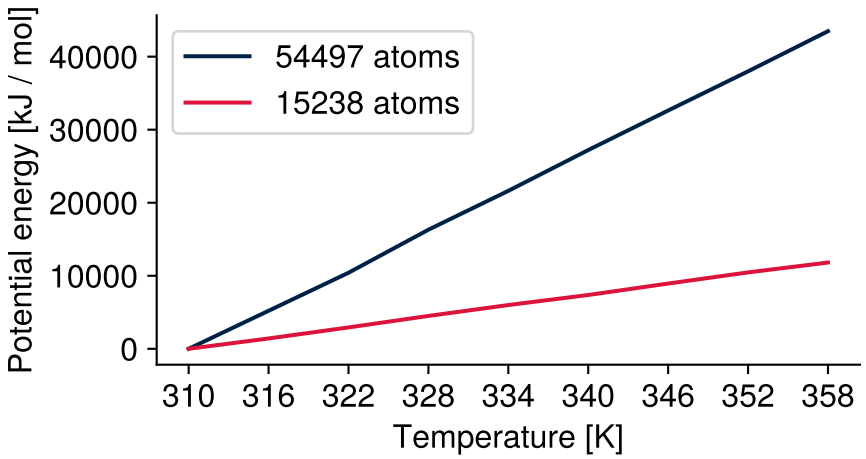

Figure 3.16: Temperature reweighting is only effective in small systems. As the system size increases, the gradient of the potential energy with respect to the temperature increases, as well. The larger the potential energy difference between two states is, the smaller the weight becomes in reweighting. Trying to apply reweighting across wide temperature gaps in large systems is not beneficial as the weights of data points from the other state are all effectively zero.

Error definitions

Two different PMF errors are considered in this study. As the simulation systems are symmetric, the potentials of mean force should be symmetric about the centre of the system at a reaction coordinate value $\xi = 0\ \mathrm{nm}$, as well. This symmetry

is quantified through the mean absolute error of symmetry, MAE_{sym}, as

$$MAE_{sym} = \frac{1}{N} \sum |PMF(\xi) - PMF(-\xi)|, \tag{3.4}$$

where N is the total number of reaction coordinate pairs $[\xi, -\xi]$. PMF profiles do not necessarily have perfectly symmetrical spacing in reaction coordinate space. To ensure meaningful MAE_{sym} values even in these cases, PMF values for symmetrical reaction coordinate pairs were obtained through linear interpolation with `numpy.interp`[241].

The second error considered here is the difference in PMF between the two bulk water components. Though already captured by MAE_{sym}, it is also independently considered because it is visually striking. The PMF at the minimum reaction coordinate value $PMF(\xi_{min})$ is defined as 0 kcal mol^{-1}. The sum of all sampling errors can be estimated in the offset of $PMF(\xi_{max})$ from this value. Since this bulk offset can give immediate clues about the degree of correctness at a glance, it is chosen as the second error under consideration in this study.

3.3.6 Calculation of partition coefficients

Partition coefficients are calculated from the PMF profiles obtained in this study. These coefficients are not expected to match experimental results: since the aim of this study is the improvement of convergence speeds rather than the correctness of PMFs, no optimisation of ligand parameters was carried out. Instead, they are calculated here to facilitate comparisons between results collected in the cholesterol-free and cholesterol-doped POPC bilayers. The free energy of partitioning ΔG_{part} into the region of the reaction coordinate ξ between ξ_{low} and ξ_{high} can be obtained from the PMF via integration[182] as

$$\Delta G_{part} = -\beta \ln \frac{1}{\xi_{high} - \xi_{low}} \int_{\xi_{low}}^{\xi_{high}} e^{-\beta PMF(\xi)} d\xi, \tag{3.5}$$

which can then be used to calculate the partition coefficient K_p as

$$K_p = e^{-\beta \Delta G_{part}}. \tag{3.6}$$

In this work, the integration is carried out with the trapezoidal rule as implemented in `numpy.trapezoid`. ξ_{low} and ξ_{high} are chosen as −1.5 and 1.5 nm, respectively, which roughly correspond to the hydrophobic core of the membrane.

3.4 Results

The goal of this study is to compare convergence behaviour of potentials of mean force (PMF) obtained with different methods and in different bilayers. Convergence and correctness of PMFs will be assessed as follows. In these umbrella sampling set-ups, the compounds' starting and ending positions lie on opposite sites of the membrane in bulk solvent. As the simulations are run under periodic boundary conditions, there is only one aqueous compartment, and the environment of the compound at either extreme of the reaction coordinate is chemically identical. The PMF should therefore be identical here, as well, and a bulk offset of the PMF is a clear sign of error. Furthermore, as the membrane bilayers under study are symmetrical and have identical leaflet composition, the PMF profile as a whole should be symmetrical.

How symmetrical is symmetrical enough? In computational chemistry, "chemical accuracy" is an often-used target. A commonly cited value for chemical accuracy is 1 kcal / mol[242]. Applied here, a PMF is considered correct when both its bulk offset and its mean symmetry error (Eq. 3.4) are less than 1 kcal mol^{-1}. We define the convergence speed of a PMF as the amount of simulation data needed to meet this criterion. The effect of each constituent method of the umbrella sampling workflow — window generation, data acquisition, and statistical estimation — on convergence speed will be discussed here.

As an example, Fig. 3.17 contains data for two different ligands tested on the POPC bilayer with steered MD-based windows and standard umbrella sampling. Lindane showed fast convergence behaviour, while chlorpromazine is slow to converge.

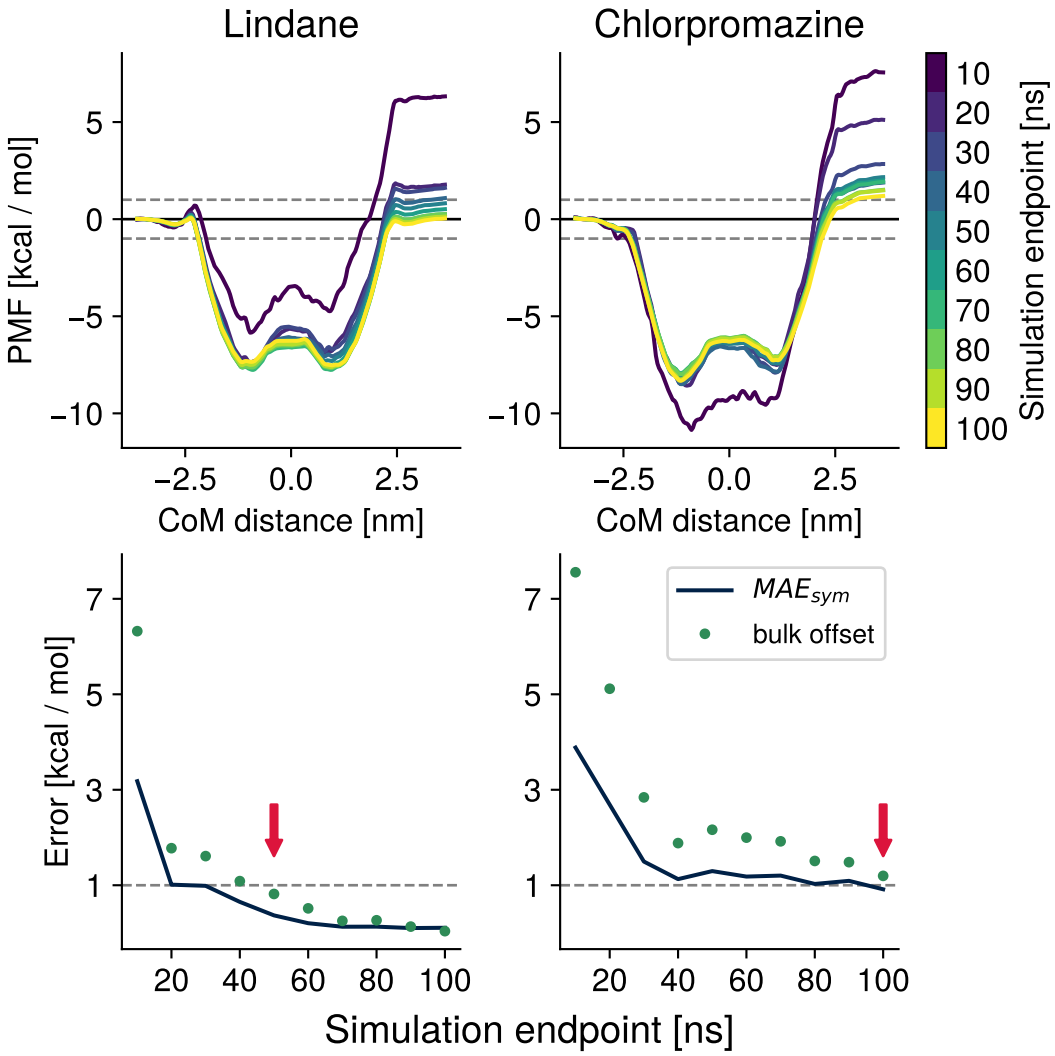

Figure 3.17: Convergence criteria used in this study, explained on a fast-converging (lindane) and a slow-converging (chlorpromazine) ligand. The effect of extending the umbrella sampling simulations is shown in the top row. The PMF profiles become progressively more symmetric. The plots on the bottom row display the evolution of the symmetry errors with increasing simulation time. For lindane, both errors drop below 1 kcal mol^{-1} after 50 ns of simulation time, while even 100 ns are insufficient to reduce the bulk offset of chlorpromazine below that threshold (red arrows).

All PMF convergence plots (like those shown in the top row of Fig. 3.17) are included in the appendix (Section B). To facilitate the comparison of convergence speeds between different conditions, each system's convergence information will be summarised by a single value: the simulation endpoint at which both MAE_{sym} and the bulk offset are less than 1 kcal mol^{-1}. This abstraction eases convergence comparisons significantly.

In this study, the convergence speed of each ligand in two bilayers was determined. The effects of the choice of umbrella window generation method, umbrella simulation method, and statistical estimator are compared, and a definitive recommendation is made based on the data.

3.4.1 Individual impacts on convergence speeds

The umbrella sampling workflow has three components: window generation, data acquisition, and PMF calculation. In this study, different options were tested for each of these components. Their individual effects are outlined below in detail on the example of the pure POPC bilayer. This detailed description is followed by an overview of convergence speed impacts of each workflow component on both bilayers.

Window generation: Alchembed vs. sMD

The first component of the umbrella sampling workflow is window generation. Fig. 3.18 shows PMFs obtained for each of the eight compounds on the POPC bilayer with standard umbrella sampling and WHAM at the 100 ns endpoint.

After 100 ns of simulation time per window, the differences between PMFs are mostly minor. Benzene, lindane, isoflurane, and sertraline appear very similar in behaviour between both window generation methods. Thiopental and paroxetine appear to have slightly less symmetrical PMFs calculated from Alchembed-based windows, whereas fluocinolone's and chlorpromazine's Alchembed-based PMFs are more symmetrical. Their steered-MD based PMFs are not yet converged at 100 ns as is visible from their bulk offset values beyond 1 kcal mol^{-1}. It is

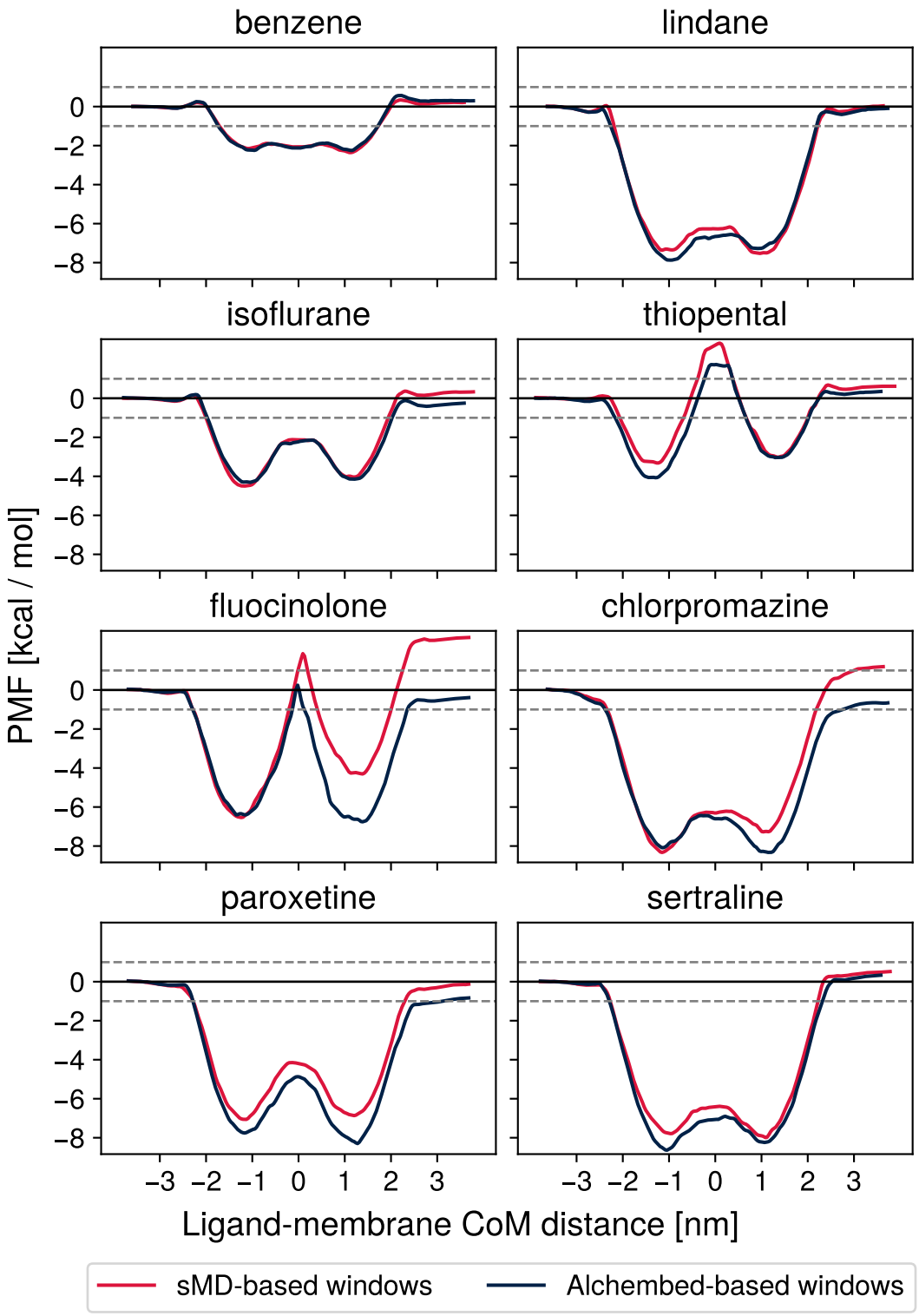

Figure 3.18: Comparison of PMFs for the 100 ns endpoint obtained with alchembed-based or sMD-based windows in the POPC bilayer. Data was acquired with standard US, and PMFs were calculated with WHAM.

encouraging to see that 100 ns of simulation time was sufficient for the Alchembed-based PMFs of all eight compounds to converge.

To determine if one of the two window generation methods leads to faster convergence, the error analysis described in Fig. 3.17 is carried out.

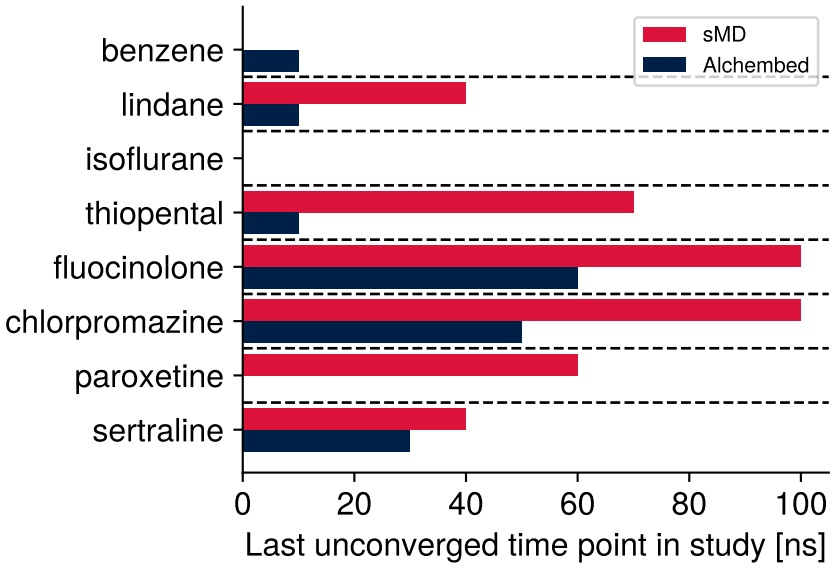

Figure 3.19: Amount of simulation time needed for PMF convergence according to the two error metrics described in Fig. 3.17.

This error analysis revealed that the convergence speed for all compounds except for benzene is as high or higher when Alchembed-based windows are used (Fig. 3.19). Lindane, thiopental, fluocinolone, chlorpromazine and paroxetine benefit dramatically, reducing the required simulation times by tens of nanoseconds per window.

The impact that the choice of window generation by itself has on convergence speeds in both bilayers is shown in Fig. 3.20. Across both bilayers, Alchembed-based windows perform as good or better than sMD-based windows in 13 out of 16 comparisons.

Data acquisition: Standard umbrella sampling vs. STeUS

STeUS can improve sampling in cases where convergence is slowed down by orthogonal degrees of freedom. If, however, these orthogonal energy barriers are

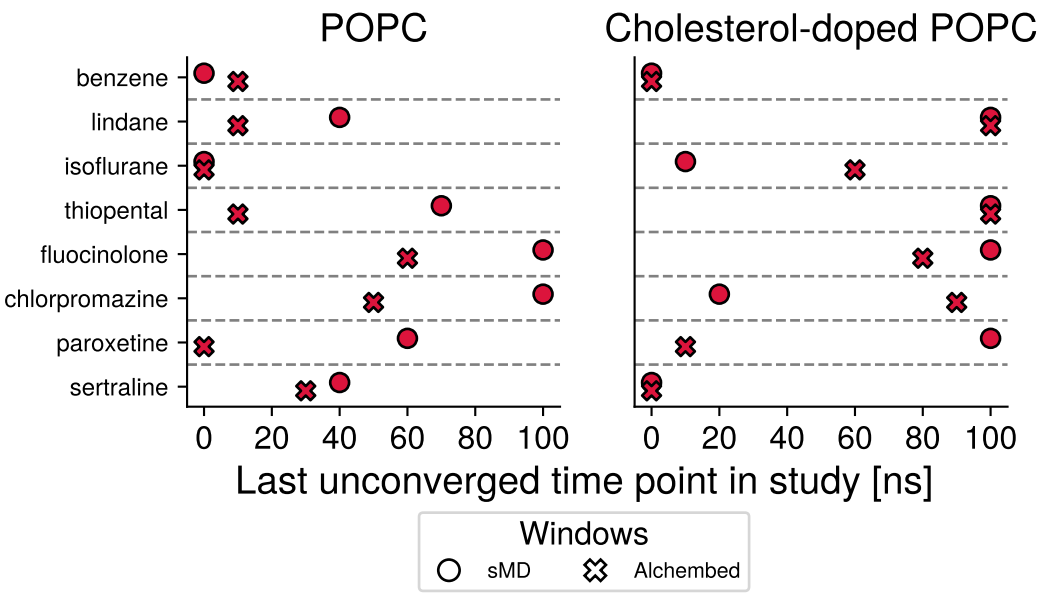

Figure 3.20: Impact of the window generation method on convergence speeds across all compounds and both bilayers. Alchembed- or sMD-derived windows were simulated with the standard umbrella sampling protocol, and PMF profiles were calculated with WHAM. Convergence was assessed with the error metric described in Fig. 3.17, and the last unconverged time point of each compound-membrane combination is shown.

small, simply more simulation time is needed, and visiting higher temperatures that do not inform the PMF can be actively harmful for convergence speed. The original authors of the method thus state that the choice of ground state occupancy needs to be made carefully[182]. As a rule of thumb, compounds which are expected to have easily-converging PMFs should be sampled at a higher ground state occupancy, while slowly-converging compounds should spend more time at higher temperatures to help overcome these energy barriers. In this study, ground state occupancies of 20 and 40 % were tested. The comparison of the resulting PMFs is shown in Fig. 3.21.

These PMFs are all calculated from sMD-based windows. For comparison, the PMF obtained with standard umbrella sampling from Fig. 3.18 is replicated here as the dashed red line. The importance of choosing an appropriate ground state occupancy is made particularly clear in the cases of thiopental, fluocinolone, and chlorpromazine, where inappropriate values prevent convergence even after 100 ns of simulation time. The comparison of convergence speeds is

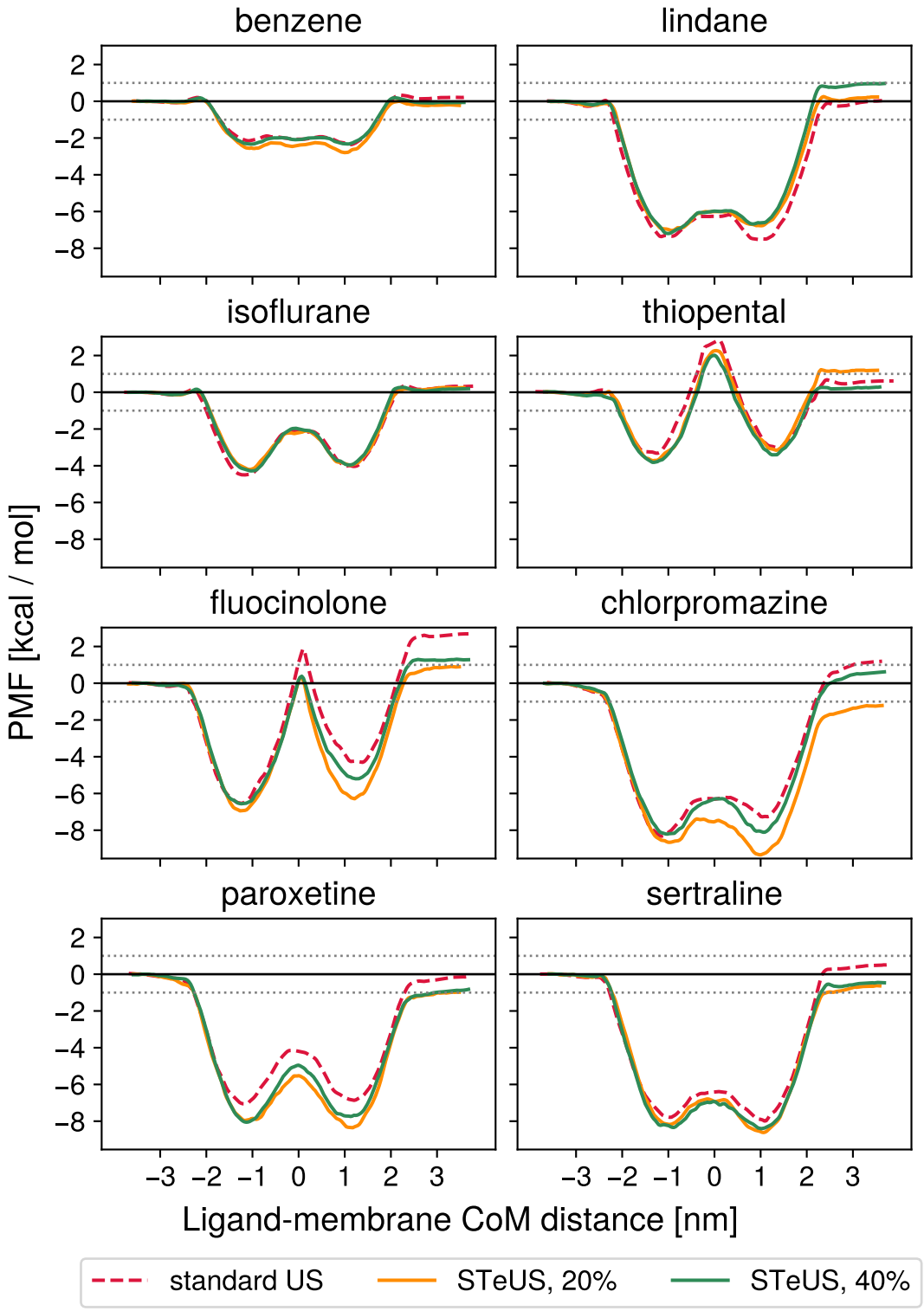

Figure 3.21: Comparison of PMFs for the 100 ns endpoint obtained with standard US or STeUS in the POPC bilayer. Different ground state occupancies were tested. Windows are sMD-based. PMFs were calculated with WHAM.

shown in Fig. 3.22.

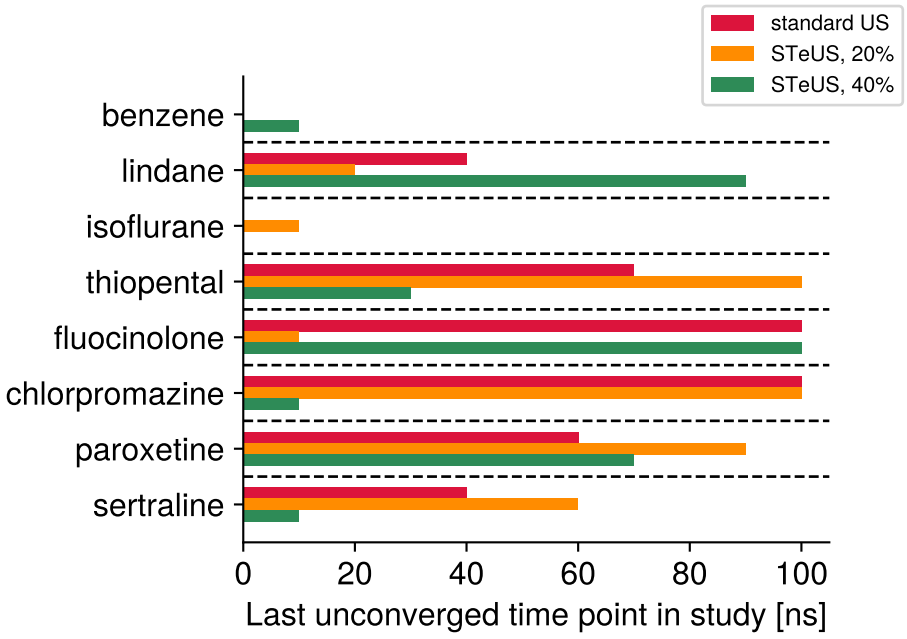

Figure 3.22: Amount of simulation time needed to converge PMFs based on the simulation method used.

The strong effect of the choice of ground state occupancy is striking. For most of the compounds tested here, STeUS will perform significantly better than standard US with one value of ground state occupancy and worse with the other value. Which value performs better depends on the compound and thus requires testing at the outset of new studies. STeUS is found to be a powerful tool to improve convergence, as reported by its original authors[182].

Starting from sMD-based windows, STeUS was found to be a strong accelerant of convergence in both bilayers (Fig. 3.23). With an appropriate choice of ground state occupancy, STeUS performed as good or better than standard umbrella sampling in 13 out of 16 comparisons.

MBAR vs WHAM

One downside of the STeUS methodology in its original description is that the majority of simulation data is collected at temperatures other than the temperature of interest and is not used to inform the PMF. However, this is theoretically

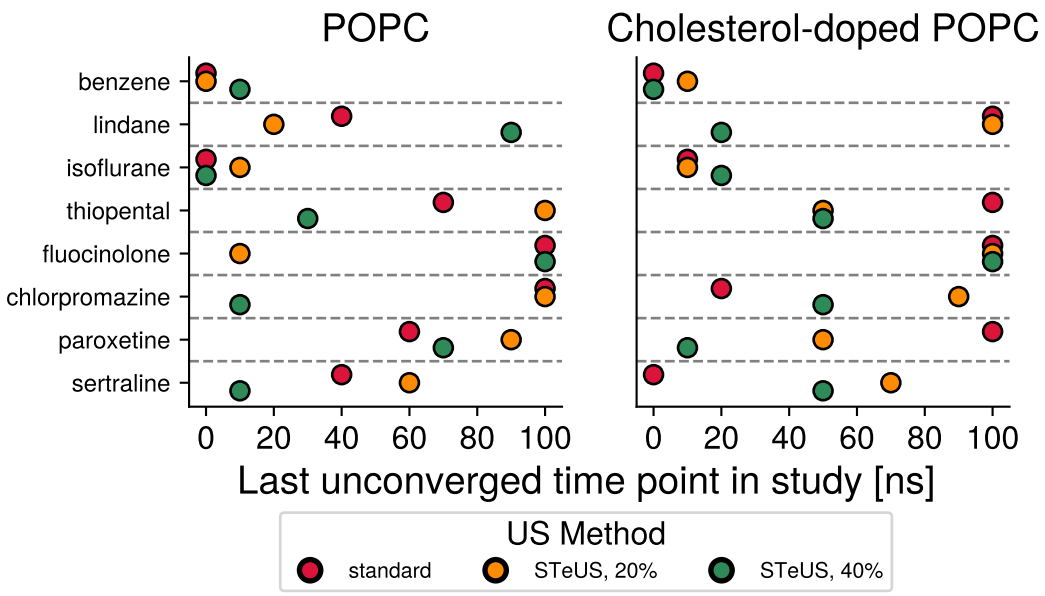

Figure 3.23: The individual impact of the choice of umbrella sampling protocol on convergence speeds of sMD-derived windows. STeUS can be a powerful accelerant of convergence—in 13 out of 16 comparisons, a STeUS method is among the fastest converging systems.

possible through temperature reweighting, though its effectiveness is often limited in practice (see Section 3.3.5). To test if including information recorded at higher temperatures can help inform the ground state PMF for the systems simulated in this study, PMFs of STeUS simulations were also calculated with the MBAR estimator[163] to carry out temperature reweighting. The MBAR-calculated PMFs are compared to their WHAM counterparts in Fig. 3.24.

At the 100 ns endpoint, MBAR with temperature reweighting performs as well or better than WHAM without it. The convergence speed analysis comparison for MBAR and WHAM is shown in Fig. 3.25.

This analysis reveals that the systems studied here are small enough for temperature reweighting to have a noticeable effect. PMFs calculated with temperature-reweighting MBAR appear to require up to tens of nanoseconds of simulation time less per window.

In both bilayers, temperature reweighting of STeUS simulations was found to have a positive impact on convergence speeds (Fig. 3.26). In 31 out of

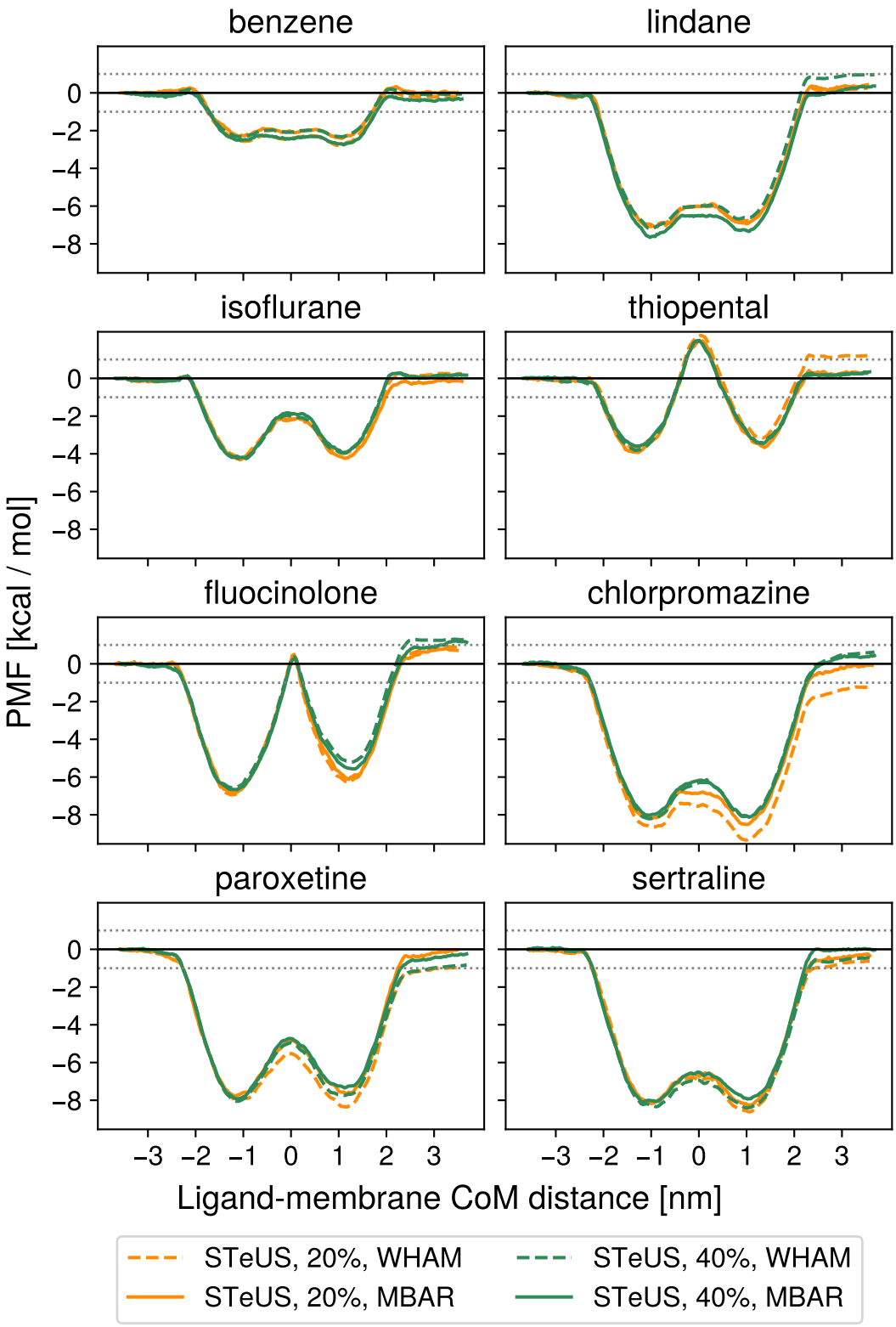

Figure 3.24: Comparison of PMFs obtained at the 100 ns endpoint for sMD-based umbrella windows sampled with STeUS, calculated with WHAM (dashed lines) or MBAR (solid lines).

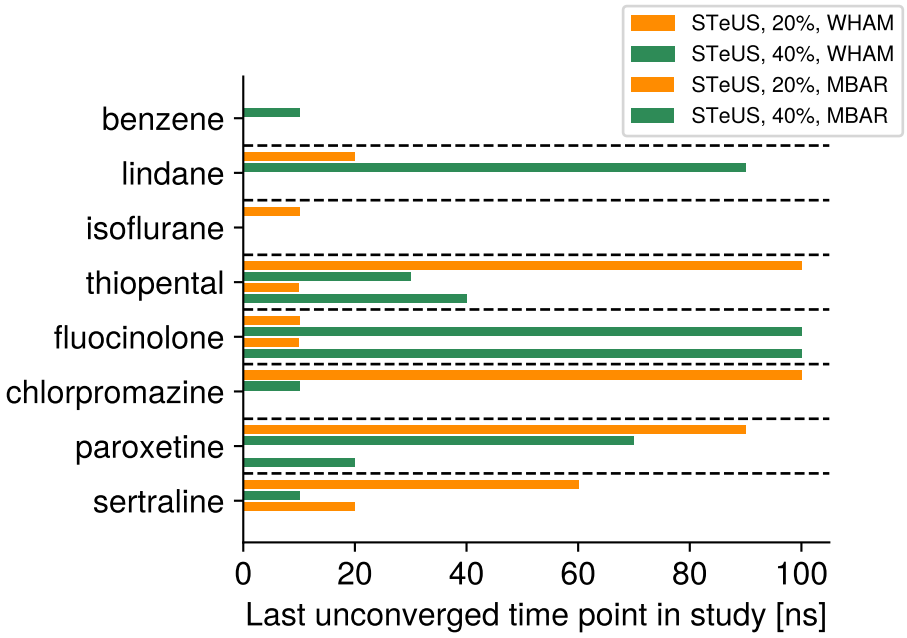

Figure 3.25: Amount of simulation time needed to converge PMFs. The comparison between WHAM and MBAR is shown for both ground state occupancies of the STeUS simulations. Temperature reweighting is highly effective in these systems.

the 32 WHAM-MBAR comparisons carried out on STeUS simulations of sMD-based windows, MBAR with temperature reweighting leads to as fast or faster convergence.

3.4.2 Effects of combinations of workflow components on convergence speeds in the POPC bilayer

Thus, Alchembed-based window generation, STeUS, and temperature reweighting with MBAR have all been demonstrated to be beneficial for PMF convergence individually. There are 10 possible combinations of these components. The convergence speed analysis for all combinations of components on the POPC bilayer is presented in Fig. 3.27.

This rather dense plot contains the results of 360 µs of simulation data. Each marker on the plot represents one of the ten combinations of window generation method, umbrella sampling method, and statistical estimator and shows how quickly the PMF for one of the eight compounds tested in this study converges with this method combination. For example, the plot shows that the choice of

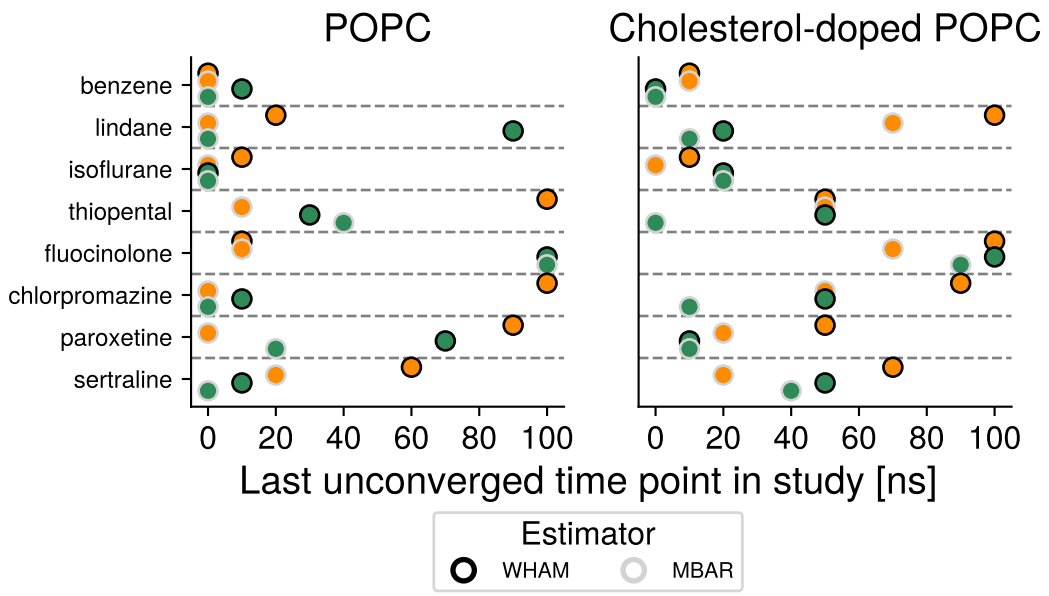

Figure 3.26: Comparison of convergence speeds of PMF calculations of STeUS simulations, carried out with MBAR or WHAM. MBAR allows straightforward temperature reweighting during the calculation of PMFs from STeUS simulation data. This was found to be highly effective, with in some cases drastic improvements of convergence speeds.

components barely matters for benzene, as, even in the worst case, the PMF converges after 10 ns per window. On the other hand, the choice of method has a huge impact in the case of chlorpromazine. Its convergence speeds range from less than 10 ns for STeUS simulations with 40 % ground state occupancy to more than 100 ns when standard umbrella sampling or STeUS with 20 % ground state occupancy are performed on sMD-based windows.

Alchembed-based windows perform as well or better than sMD-based ones 25 out of 40 times. The case of fluocinolone is of particular interest here: Alchembed-based windows improve standard umbrella sampling, but detract from STeUS simulations (with 20 % ground state occupancy). Fluocinolone is the most complicated ligand to converge. A possible explanation for this behaviour is that, while Alchembed-based windows generally improve PMFs, the additional variation in window starting positions introduced by sMD (see Fig. 3.15) is actually more beneficial as it might help rigid fluocinolone initially break alignment with the membrane leaflets.

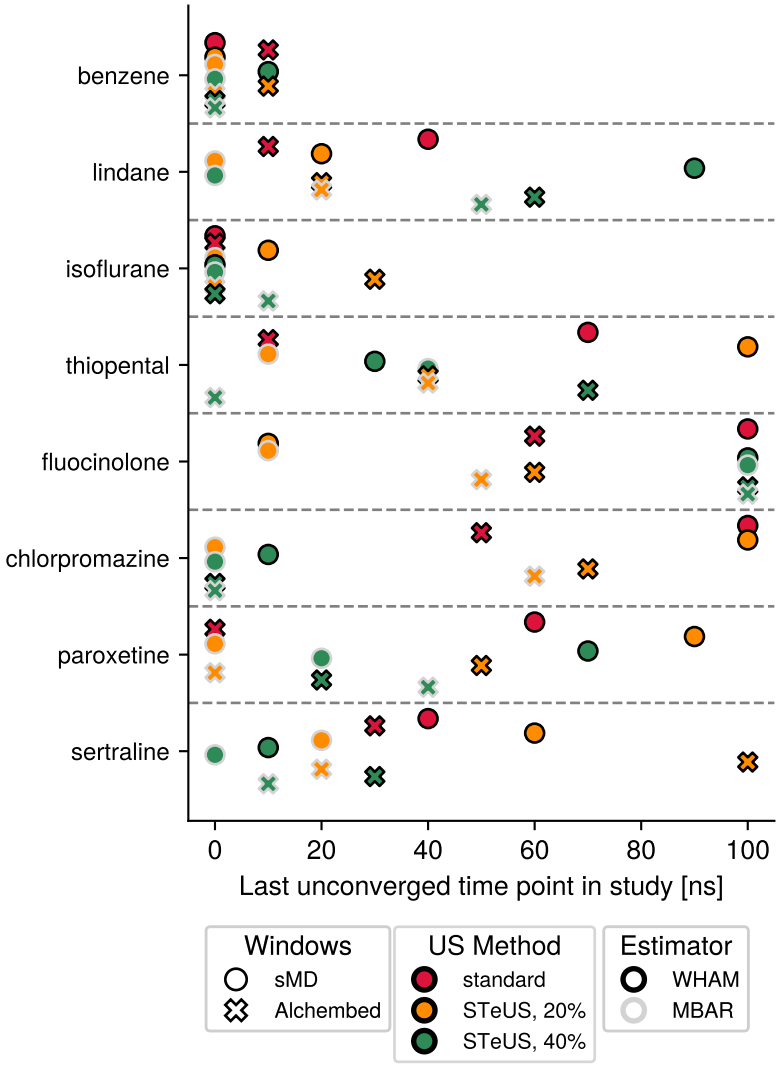

Figure 3.27: Convergence speed comparison across all methods and compounds for data collected in the POPC bilayer. For each compound and each combination of umbrella sampling workflow components, the last time point at which the resulting PMF had not converged according to the criteria described in Fig. 3.17 is marked. The shape, colour, and outline of each marker on the plot denote the workflow components used to calculate this PMF. Raw PMF profiles for each data point are included in the appendix (Section B). This plot expands on the information contained in Figs. 3.20, 3.23, and 3.26 by including all possible combinations of workflow components.

For all 8 compounds tested in this study, a STeUS simulation is in the group of best-performing methods. A clear result of this study is that STeUS always performs as well as or better than standard umbrella sampling, provided an appropriate ground state occupancy is chosen. In the case of sertraline, for example, STeUS with a ground state occupancy of 40 % outperforms standard umbrella sampling. If a ground state occupancy of 20 % is set instead, then standard umbrella sampling performs better than STeUS unless temperature reweighting with MBAR is carried out. The situation is reversed for fluocinolone, where the best results are obtained with a ground state occupancy of 20 %, while none of the simulations with 40 % have converged after 100 ns. Thus, STeUS is shown to be superior to standard umbrella sampling, provided that the initial time investment to find suitable parameters is made.

For a system of this size (see Table 3.1), temperature reweighting with MBAR is possible and overall beneficial for convergence. In 29 out of the 32 STeUS simulations tested in the POPC bilayer, temperature reweighting made the PMF converge as fast or faster. If the simulation system of a STeUS simulation is small enough, temperature reweighting with MBAR should be used for further performance gains.

3.4.3 Cholesterol flip-flop complicates PMF calculation in the cholesterol-doped POPC bilayer

During the analysis of PMFs in the cholesterol-doped membrane systems, it became apparent that membrane fluidity might prove problematic to obtaining converged, symmetrical PMFs. This is because cholesterol, with its rigid, bulky core and its charge-neutral head group, has a much lower energy barrier to movement between membrane leaflets than phospholipids such as POPC. The movement of cholesterol between leaflets ("flip-flopping") is a well-studied, physiologically relevant phenomenon[243–246], so on the one hand, capturing it in these simulations is encouraging. However, umbrella sampling relies on a sufficient degree of similarity between the individual windows. This similarity is not maintained in

these simulations, as the number of flip-flop events between windows is not consistent (Fig. 3.28 A). On average, ≈8 flip-flops occur in each STeUS window. The highest total count observed in a single window is 20. Considering that the system only contains 18 cholesterol lipids, these numbers are surprisingly large.

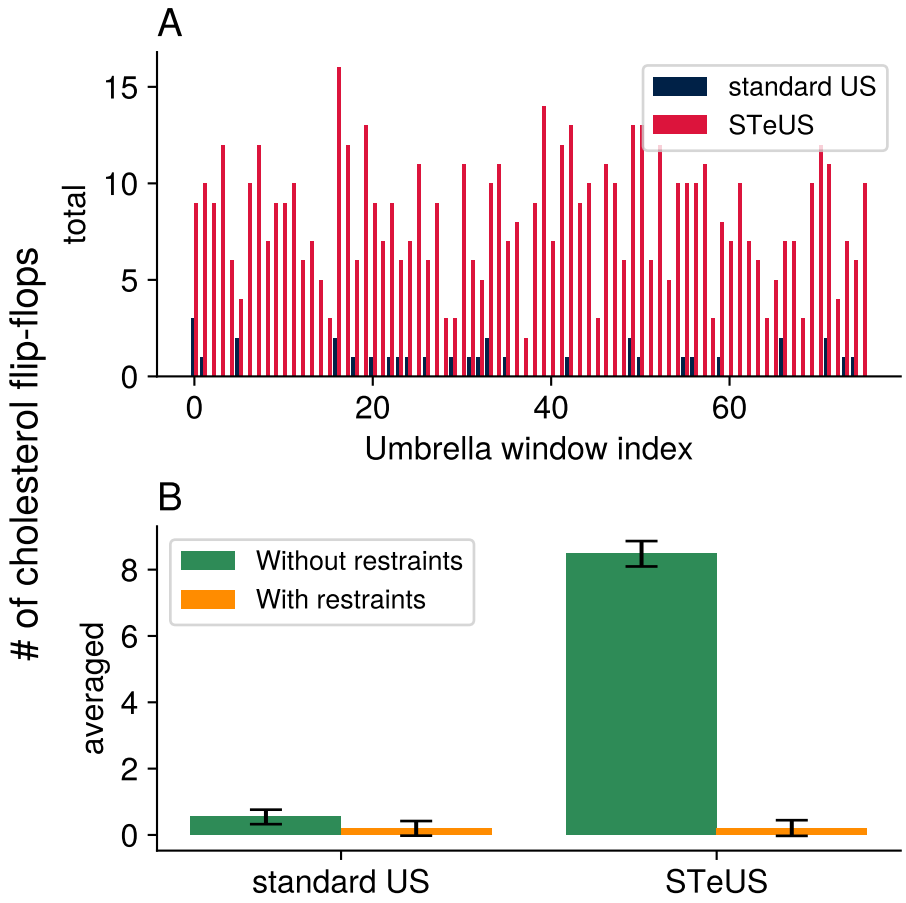

Figure 3.28: Flip-flopping is rare at 310 K, but its probability is increased at the higher temperatures sampled in STeUS simulations. A: Cholesterol flip-flop counts for each umbrella window of an example system run with standard umbrella sampling and STeUS. B: Average number of cholesterol flip-flop events across all umbrella windows and compounds. The application of loose flat-bottom restraints on the cholesterol head groups all but prevents flip-flopping. Thus, stable membrane systems that facilitate PMF calculations are obtained. Error bars indicate standard deviations.

For this reason, it was decided to limit the range of motion of cholesterol. Loose harmonic flat-bottom restraints were applied to the oxygen atom of the cholesterol head group, applying a force only when it moved more than 1.3 nm away from its starting coordinate along the membrane normal. This restraint

allowed unrestrained and unbiased diffusion of cholesterol within its starting leaflet and into the membrane centre while preventing outright flip-flops (Fig. 3.28 B). PMFs collected in the thus restrained system are discussed below. The overall analysis graph of the unrestrained system is included in the appendix (Fig. B.1).

3.4.4 Convergence speeds in cholesterol-doped POPC

The graphical analysis of the convergence speeds of the different methods is shown in Fig. 3.29. The comparison to convergence speeds in the cholesterol-free POPC bilayer (Fig. 3.27) shows that, overall, the presence of cholesterol reduces convergence speeds.

Alchembed-based window generation is again found to be beneficial, performing as well or better than steered MD-based windows in 27 out of 40 comparisons. In this bilayer, fluocinolone appears to benefit from Alchembed-based window much more than in the cholesterol-free POPC bilayer, while the benefits appear less clear in the cases of isoflurane and chlorpromazine. This variation in behaviour should be studied further, for example through the generation of random ligand orientations during Alchembed-based window generation.

STeUS proves itself to be a powerful tool for the acceleration of PMF convergence on this lipid bilayer, just as it did in the case of the cholesterol-free POPC membrane. As before, a STeUS method is in the best-performing group of methods for every tested compound. Lindane, thiopental, and fluocinolone appear to particularly benefit.

Although it is 10 % larger than the cholesterol-free POPC bilayer, this system is still small enough to benefit from temperature reweighting with MBAR. MBAR performs as well as or better than WHAM in 27 out of 32 comparisons.

3.4.5 Comparison of PMFs between both bilayers

We have thus conclusively shown that Alchembed-based window generation, combined with enhanced sampling through the use of STeUS, and the maximisation of data utilisation through temperature reweighting with MBAR is a powerful

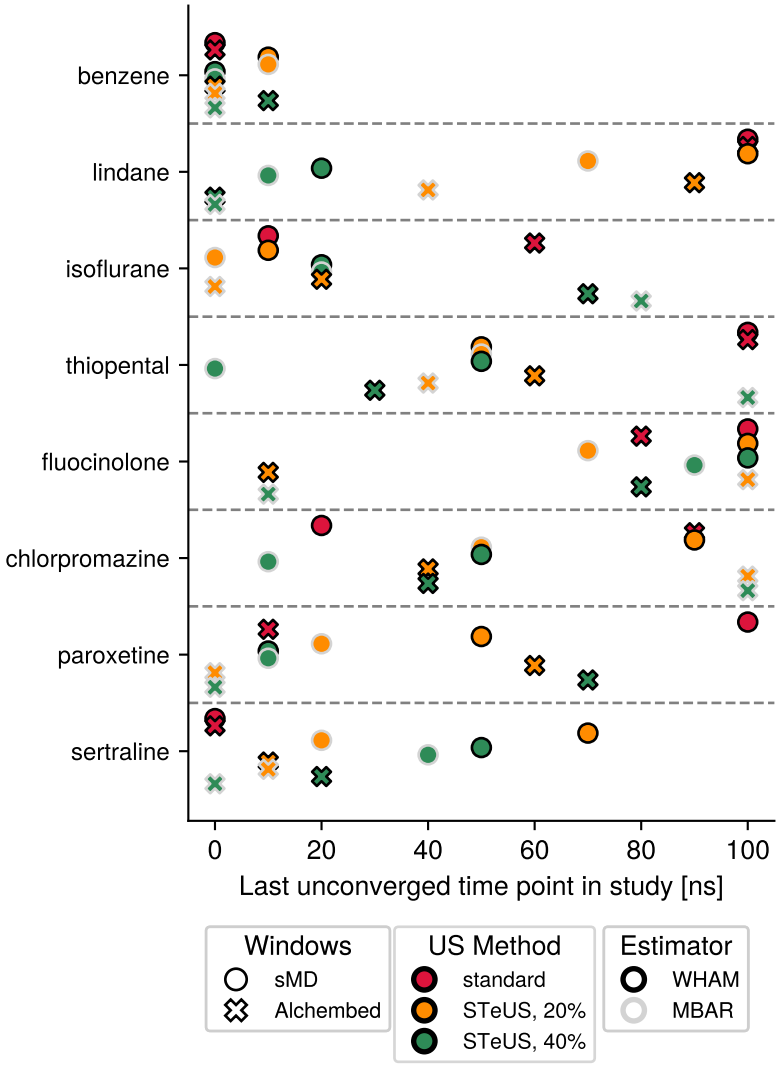

Figure 3.29: Convergence speed comparison across all methods and compounds for data collected in the cholesterol-doped POPC bilayer. For each compound and each combination of umbrella sampling workflow components, the last time point at which the resulting PMF had not converged according to the criteria described in Fig. 3.17 is marked. The shape, colour, and outline of each marker on the plot denote the workflow components used to calculate this PMF. Raw PMF profiles for each data point are included in the appendix (Section B). This plot expands on the information contained in Figs. 3.20, 3.23, and 3.26 by including all possible combinations of workflow components.

method to accelerate convergence of PMFs. Converged, symmetrical PMFs for all eight ligands and both lipid bilayers were obtained (Fig. 3.30).

The effects of cholesterol's presence in the membrane outlined in Section 3.2.1 are visible in these PMFs. Cholesterol intercalates between the phospholipids and enforces higher order on their tails, which leads to a thickening of the lipid bilayer. This thickening is observed in all eight PMF pairs. Similarly, all eight comparisons show an increased energy barrier to membrane permeation in presence of cholesterol, as predicted.

To quantify the effects of the presence of cholesterol in the bilayer, partition coefficients between the hydrophobic core of the membrane (orange shading in Fig. 3.30) and bulk solvent were calculated. They are listed in Table 3.3 and compared to predicted logP values reported in PubChem. As correctness of the partition coefficients was not a goal of this project, no optimisation of ligand forcefield parameters was carried out. Instead, we were interested in differences in partition coefficient based on the lipid composition, as the partitioning behaviour of all eight compounds tested in this study was shown to depend on cholesterol content (see Section 3.2.2).

Table 3.3: Comparison of logP values obtained from the PMFs of each compound going through the cholesterol-free and the cholesterol-doped POPC bilayers. Shown are the average logP values calculated from all converged conditions (see Figs. 3.27 and 3.29) and their standard deviations. Here, n denotes the number of converged calculations. As expected, partition coefficients for the latter bilayer are consistently lower. $logP_{o\text{-}w}$ denotes the octanol-water partition coefficients calculated with XLogP3[247] and reported by PubChem.

compound	$logP_{o\text{-}w}$	$logP_{POPC}$	$logP_{POPC+Chol}$
Benzene	2.1	1.55 ± 0.13 (n=10)	1.42 ± 0.13 (n=10)
Lindane	3.8	4.95 ± 0.25 (n=10)	4.65 ± 0.16 (n=7)
Isoflurane	2.1	2.62 ± 0.08 (n=10)	2.40 ± 0.16 (n=10)
Thiopental	2.9	2.13 ± 0.21 (n=9)	1.87 ± 0.21 (n=7)
Fluocinolone	2.5	4.29 ± 0.12 (n=5)	3.83 ± 0.34 (n=6)
Chlorpromazine	5.2	5.48 ± 0.15 (n=8)	4.84 ± 0.12 (n=8)
Paroxetine	3.5	5.01 ± 0.24 (n=10)	4.51 ± 0.19 (n=9)
Sertraline	4.8	5.46 ± 0.15 (n=9)	5.24 ± 0.21(n=10)

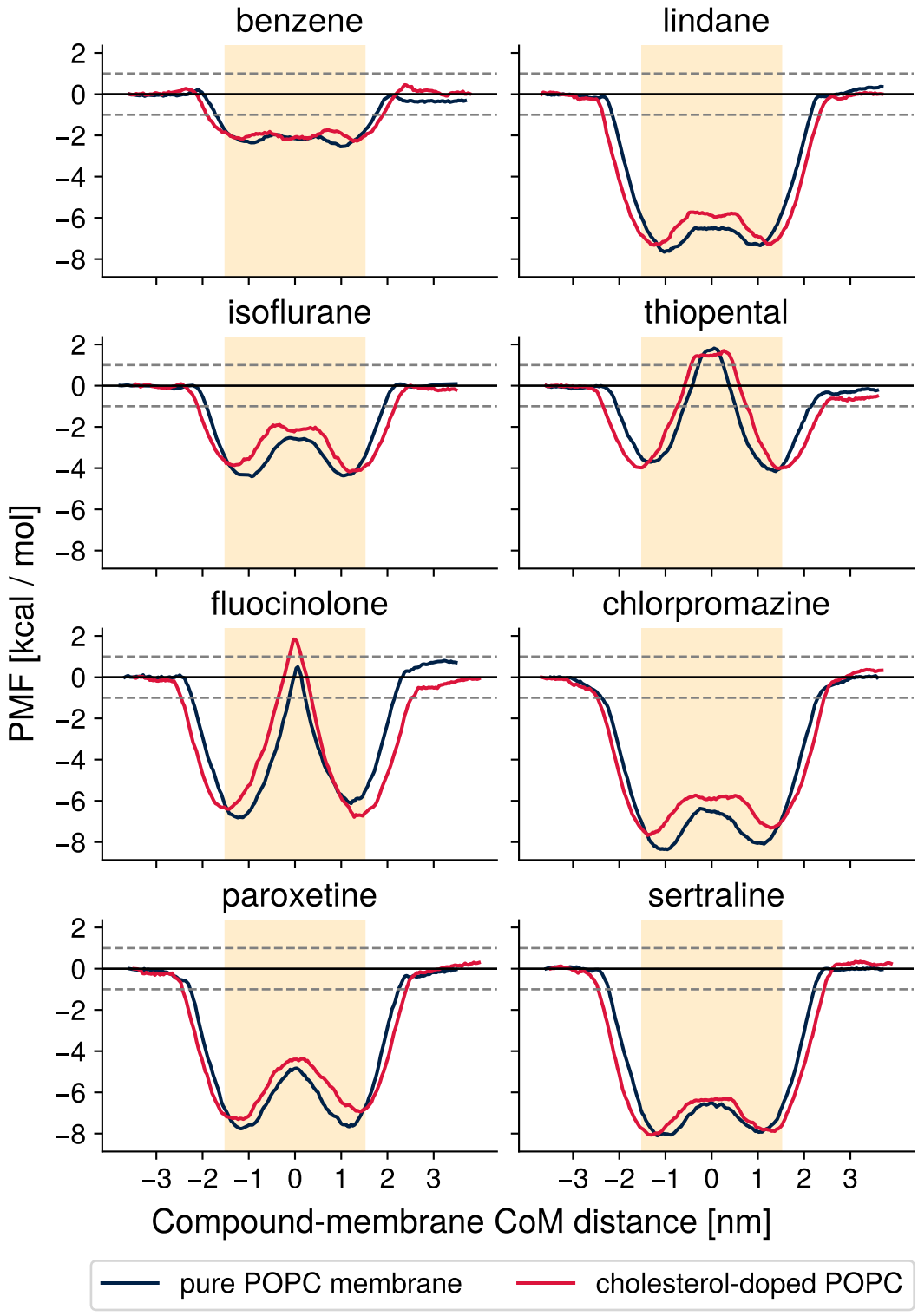

Figure 3.30: Final, fully converged and highly symmetrical PMFs for all eight compounds and both lipid bilayers tested in this study. The effects of cholesterol are visible through a broadening of membrane region and an increase of the energy barriers. Partition coefficients between water and the hydrophobic core, highlighted in orange, were calculated (Table 3.3). Multiple conditions lead to convergence. The combination of workflow components that underlies the profiles shown here is listed in Table 3.4.

Table 3.4: Chosen workflow components of the PMF profiles displayed in Fig. 3.30 and Fig. 3.31. The profile resulting from any workflow that leads to PMF convergence (see Figs. 3.27 and 3.29) could have been chosen for illustration. The combinations listed here were chosen for display in Fig. 3.30 because they correspond to the fastest converging combination of workflow components for this ligand and membrane.

Compound	Pure POPC bilayer	Cholesterol-doped POPC bilayer
Benzene	Alchembed, STeUS (40 %), MBAR	Alchembed, STeUS (40 %), MBAR
Lindane	sMD, STeUS (40 %), MBAR	Alchembed, STeUS (40 %), MBAR
Isoflurane	Alchembed, STeUS (40 %), WHAM	sMD, STeUS (20 %), MBAR
Thiopental	Alchembed, STeUS (40 %), MBAR	sMD, STeUS (40 %), MBAR
Fluocinolone	sMD, STeUS (20 %), MBAR	Alchembed, STeUS (40 %), MBAR
Chlorpromazine	Alchembed, STeUS (40 %), MBAR	sMD, STeUS (40 %), MBAR
Paroxetine	sMD, STeUS (20 %), MBAR	Alchembed, STeUS (20 %), MBAR
Sertraline	sMD, STeUS (40 %), MBAR	Alchembed, STeUS (40 %), MBAR

The partition coefficients obtained for the cholesterol-doped POPC bilayer are consistently lower than those for the cholesterol-free bilayer. This is in agreement with expectations, as cholesterol is known to reduce partitioning for many compounds, including all eight molecules used in this study. This demonstrates the usefulness of umbrella sampling for the determination of partition coefficients for precisely defined, heterogeneous chemical environments in contrast to predictions of octanol-water partition coefficients.

3.4.6 Errors associated with symmetry and periodicity constraints

This study was motivated by the fact that obtaining converged PMFs from umbrella sampling can be a computationally expensive process, requiring a large number of windows and long simulation times. This problem has previously been addressed through the use of boundary conditions in the PMF calculation[179,248–250]. PMFs can be forced to be symmetrical about the origin of the reaction coordinate and/or to be periodic so that the first and last window are considered to be identical[180]. These boundary conditions artificially cause the PMF to look converged by introducing errors along its entire length while enforcing symmetry and periodicity. The workflow described here is able to generate converged PMFs with reduced

computational cost and without the introduction of systematic errors brought about by these constraints.

To quantify the errors that are introduced by the application of these constraints, the PMF profiles presented in Fig. 3.30 are compared to PMFs obtained through WHAM calculated from 10 ns of standard umbrella sampling on sMD-based windows, computed with enforced symmetry and periodicity. As recommended in its original publication, autocorrelation times are estimated by and supplied to `gmx wham`[180]. The PMF comparison is shown in Fig. 3.31.

The data shown in Fig. 3.31 reveal that the differences between constrained and unconstrained PMFs can appear minor, particularly in cases of easily converging compounds like benzene. The large differences in PMFs of chlorpromazine and fluocinolone, on the other hand, demonstrate the risk of this approach. By employing these boundary conditions, PMFs that are far from converged can be made to look convincing.

The apparent partition coefficients between the hydrophobic core and bulk solvent are listed in Table 3.5. A graphical comparison of logP values obtained from either constrained or unconstrained, fully converged PMF profiles is shown in Fig. 3.32.

Table 3.5: Partition coefficients calculated for the constrained PMFs shown in Fig. 3.31, and their value relative to the average logP values obtained for all fully converged PMFs listed in Table 3.3.

compound	$\log P_{POPC,constr.}$	% of $\log P_{POPC}$	$\log P_{POPC+Chol,constr.}$	% of $\log P_{POPC+Chol}$
Benzene	1.53	99	1.23	87
Lindane	5.33	108	4.54	98
Isoflurane	3.05	116	2.06	86
Thiopental	1.75	82	1.84	98
Fluocinolone	4.71	110	1.87	49
Chlorpromazine	8.76	160	5.39	111
Paroxetine	4.95	99	5.16	114
Sertraline	6.30	115	4.23	81

This comparison further makes the case against the use of symmetry- and periodicity-enforcing boundary conditions. Such PMFs no longer reliably report

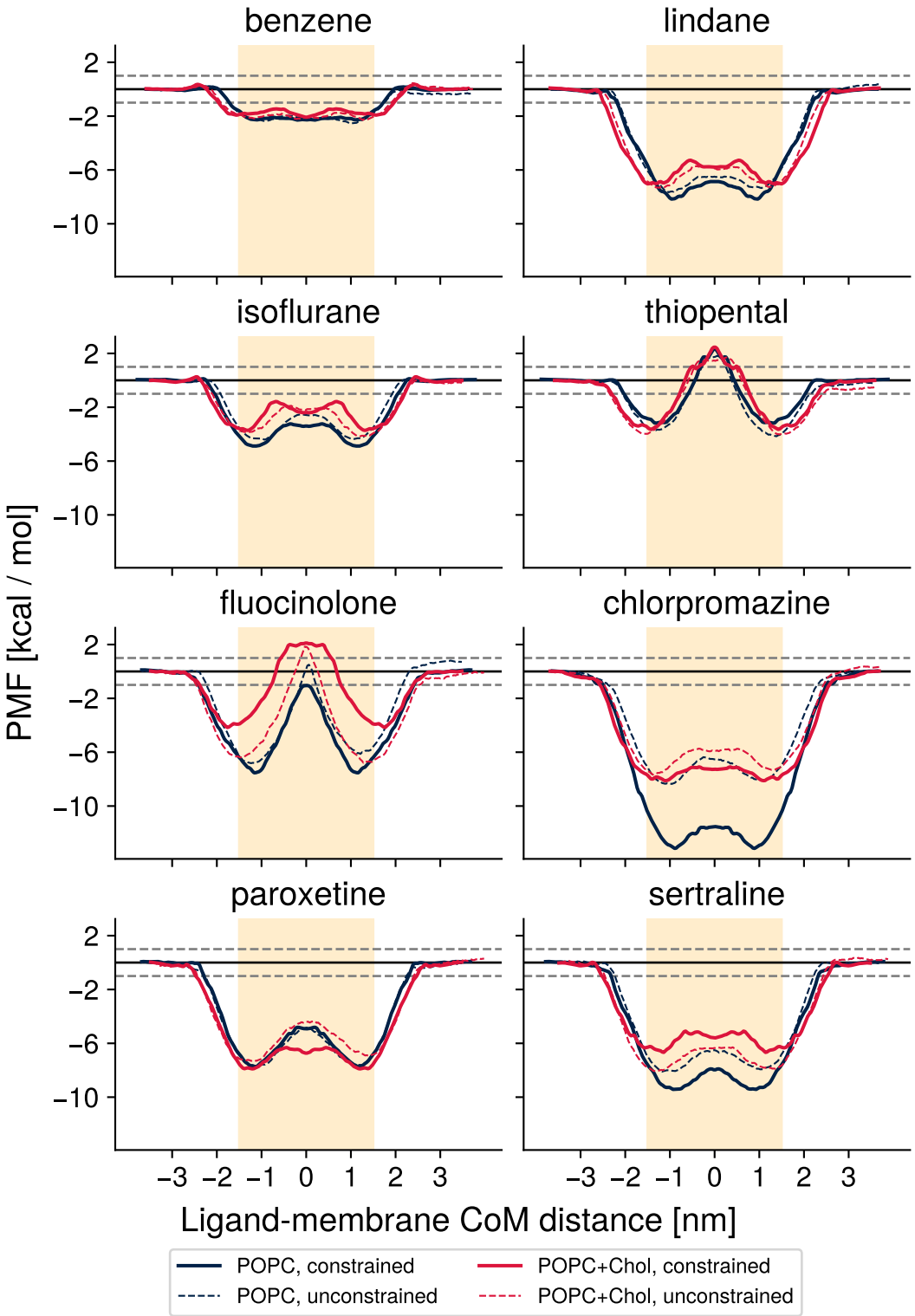

Figure 3.31: Comparison between the converged PMFs from Fig. 3.30 (dashed lines) and PMFs calculated by WHAM with enforced periodicity and symmetry, based on 10 ns of standard US data collected on sMD-based windows (solid lines). The orange region marks the definition of the hydrophobic core used in the calculation of partition coefficients. The combination of workflow components underlying each unconstrained profile is listed in Table 3.4. The workflow underlying constrained PMF calculations is sMD, standard, WHAM.

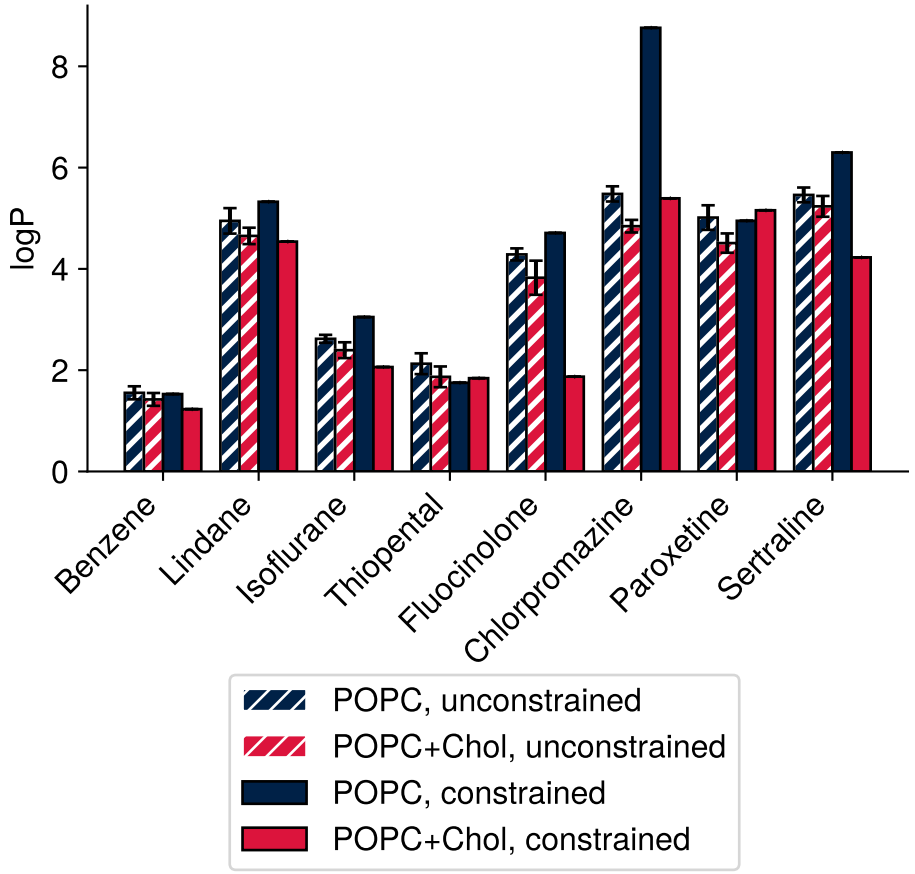

Figure 3.32: Comparison of logP values calculated from unconstrained and constrained PMF profiles. Values shown for unconstrained calculations are average values and standard deviations of all converged calculations (hatched bars, see Table 3.3). Solid bars represent logP values calculated from PMF profiles obtained with WHAM under symmetry- and periodicity constraints, based on 10 ns of standard MD on steered MD-based umbrella windows (Table 3.5).

a lower partition coefficient in cholesterol-doped POPC bilayers. Comparing these partition coefficients to those obtained from well-converged PMFs shows that the error is particularly large for fluocinolone, chlorpromazine, paroxetine, and sertraline. There is no consistent over- or underestimation of the partition coefficient, complicating the discussion of constrained PMFs even when acknowledging the underlying error. As a result, we argue that potentials of mean force should not be calculated with these boundary conditions. Instead, enhanced sampling methods such as the one described here should be used to ensure correctness of the reported data.

3.4.7 A failed attempt to circumvent the system size limit for reweighting

In order to circumvent the system size limit for effective reweighting with MBAR, I attempted the implementation of an alchemical version of STeUS. This was inspired by the *Replica Exchange with Solute Tempering* and *Replica Exchange with Solute Scaling* methods (REST and REST2)[251,252]. These methods were developed to solve the equivalent problem in parallel tempering replica exchange simulations: too many replicas are needed to both cover a sufficiently wide temperature range and ensure a sufficient exchange probability. REST2 addresses this problem by performing Hamiltonian replica exchange instead of parallel tempering. Instead of changing the reference temperature, the replicas differ by scaling factors applied to the dihedral and non-bonded interaction terms of the group of interest. These scaled potentials mimic the effect of higher temperatures without increasing the overall potential energy U of the system as much as higher temperatures would. Thus, fewer replicas are needed to span the same "effective temperature" range and enhance sampling of the group of interest.

Since REST2 is a replica exchange method, it places certain constraints on the hardware configuration it can be run on. One of the advantages of STeUS is that each window is truly independent: simulations can be run serially, across computational resources, and individual windows can be added or extended as needed. However, once simulation systems become large enough to prevent temperature reweighting, a large amount of simulation time is "lost".

Because of this, I attempted a combination of these ideas and developed an expanded ensemble method in which the system moves along an alchemical ladder instead of a temperature ladder. Each rung on said ladder sampled the system with a differently scaled Hamiltonian. Early tests looked promising: all states were visited as expected, the potential energy was similar across all Hamiltonians, and the method was able to sample the alanine dipeptide torsion fully. However, this approach was found to have an unacceptable performance impact.

In simulated tempering, the temperature of the system is changed by simply updating the thermostat's reference temperature. To move in an expanded ensemble between different Hamiltonians, on the other hand, the free energy code of GROMACS is employed. The free energy kernel unfortunately carries a significant performance impact per perturbed atom[253]. Since this method would scale the interaction parameters of the ligand and the entire membrane, it is at present prohibitively expensive. In future versions of GROMACS, or in fact in other MD codes, this method might be more feasible.

3.5 Discussion and future directions

The calculation of free energy profiles along reaction coordinates is an important application of computational methods, but it is often challenging to obtain converged results. Umbrella sampling is one method with which these profiles can be calculated. In this study, the individual effects of three components of the umbrella sampling workflow on overall convergence speed were systematically studied and compared.

Umbrella sampling windows can be generated in a variety of ways. Here, the comparison between steered-MD based windows — which are extracted from a pulling trajectory — and Alchembed-based windows — in which ligands are alchemically grown along the reaction coordinate — was carried out. We hypothesised that hysteresis might be a larger problem in steered-MD based windows as the directionality of the initial pulling simulation can introduce larger starting condition-dependent disturbances which can take a long time to be resolved through equilibration. Indeed, we find that umbrella sampling carried out on Alchembed-based windows converges faster in the majority of the cases in both bilayers. Its acceleration is twofold: the initial pulling simulation is not needed, and each window requires less simulation time to reach convergence. We thus recommend Alchembed-based window generation be attempted at the outset of future umbrella sampling studies.

One limitation of the Alchembed-based window generation method as tested here is that the ligand orientation is precisely matched in all windows. Steered-MD based windows, on the other hand, have inherent variability of ligand orientations between windows. During the pull, more favourable orientations might be sampled that are not explored by growing the ligand in place. In future work on this project, the effect of randomising ligand orientation during Alchembed-based window generation to mimic this behaviour of sMD-based windows should be explored.

Simulated Tempering-enhanced Umbrella Sampling[182] is a method to enhance sampling in umbrella sampling studies. In this work, we independently verified its authors' claims of enhanced sampling and found that STeUS is indeed superior to standard umbrella sampling. However, on GPU-based systems, it comes with a significant performance impact because it requires the velocity-verlet integrator which does not support some forms of GPU acceleration (see Section 2.1.2). Thus, we conclude that STeUS is at its most advantageous when PMFs are particularly slow to converge, or when run on hardware that is fully or predominantly CPU-based. Future versions of GROMACS may improve GPU utilisation of expanded ensemble methods, at which point we would unreservedly recommend STeUS over standard umbrella sampling for most applications.

Another downside of STeUS as described in its original paper is that simulation data collected at higher temperatures is not used to inform the PMF. With the ground state occupancies trialled in this study, this would mean discarding 60 to 80 % of the sampled data. Here, we attempted to maximise data utilisation of STeUS simulations. We evaluated the effectiveness of temperature reweighting of high-temperature information to inform the PMF at the temperature of interest. We report that for the system sizes studied here, temperature reweighting with MBAR is effective and reduces the amount of simulation time needed to reach convergence. System size should be minimised as much as possible to maximise simulation performance. These results suggest that a second reason for system size minimisation is the improvement of temperature reweighting. We fully recommend the use of temperature reweighting, for example with MBAR,

in the case of STeUS simulations performed on small enough systems, as this offers an essentially free way to improve convergence speed.

3.5.1 Future work

The study already demonstrates the superiority of this novel workflow. However, it would benefit from a larger sample size. In the future, repeats of each condition should be run. With careful simulation set-up, most conditions will not require the full 100 ns of simulation time, making the repeats significantly less computationally expensive than the initial work reported here.

The reported flip-flopping behaviour of cholesterol should be investigated further. It might be necessary to further refine the restraints that were applied to the lipid here. Additionally, living cells expend energy to maintain a degree of lipid leaflet asymmetry. Determining PMFs in more realistic biological membranes would involve accounting for this asymmetry, as well.

The chemical environment on either side of the membrane was identical in the simulations carried out here. The effect of the many concentration gradients that are maintained across biological membranes was not investigated in this study. This could be done, however, in future studies where the simulation system contains two lipid bilayers. Such systems are used in computational electrophysiology[254], for example, and effectively simulate an intracellular and an extracellular compartment. A system with true concentration gradients and asymmetrical membranes would give more realistic results of membrane permeation in biological systems.

A related problem to membrane permeation of compounds is the investigation of interactions between lipids and membrane proteins[255–257]. The method developed here could also be tested on such problems.

The three components tested in this study account for a lot of variation in umbrella sampling studies. Two important parameters that were not varied here are the umbrella window spacing and the restraint force constants. It is conceivable that STeUS would benefit from a more in-depth study of the effect of

these parameters, as a wider range of the reaction coordinate can be sampled at higher temperatures.

If future versions of GROMACS provide a more efficient implementation of the free energy kernel, revisiting the failed attempt at Hamiltonian expanded ensemble umbrella sampling described here is advisable. It might then be possible to make the advantages of reweighting with MBAR available to larger systems, as well.

3.6 Conclusion

In this work, the individual components of umbrella sampling workflows were benchmarked. The combination of Alchembed-based window generation[230], Simulated Tempering-enhanced Umbrella Sampling[182], and PMF calculation using temperature reweighting with MBAR[163] was identified as showing the fastest overall convergence of PMF profiles. At the outset of this chapter, the limitations of using octanol-water partition coefficients as a stand-in for membrane permeability were discussed. The workflow described here is capable of screening the energetics of membrane permeation relatively quickly and is sensitive enough to distinguish between different membrane compositions. Unfortunately, it is much, much slower than cheminformatics methods like XLogP3[247] and is thus not well-suited for high-throughput applications. However, in studies of a limited number of complicated systems, this workflow could prove very helpful. For example, it could help in the study of skin permeation when applied to the skin models of Lundborg et al[173,258]. I developed this workflow for the study of drug-permeation of fenestrations, which will be discussed in the following chapter.

4

Fenestrations in the Glycine Receptor

Contents

4.1 Introduction

Although the importance of drug-permeable fenestrations is well-accepted in the ion channel field in general[108], they have not previously been reported in the superfamily of pentameric ligand-gated ion channels (pLGICs). In this project, electrophysiology data of the glycine receptor is reinterpreted through the lens of fenestration-based drug access to the channel pore. Through MD simulation and analysis of available structural data, evidence of a state-dependent, drug permeable fenestration in the glycine receptor is provided for the first time.

4.1.1 Motivation

This project was inspired by electrophysiology data reported in "Mechanism of action of the insecticides, lindane and fipronil, on glycine receptor chloride channels" by Robiul Islam and Joseph W. Lynch, published in the British Journal of Pharmacology in 2012[90]. Although fipronil is typically thought of as an insect GABA receptor inhibitor (see Section 4.1.2), the authors were motivated to investigate the action of fipronil on the human glycine receptor following a toxicological study that suggested glycine receptor-linked pathology in *Danio rerio*[259].

Islam and Lynch thus tested the fipronil-mediated inhibition of human glycine receptors in whole-cell patch-clamp electrophysiology. Their results provide clues regarding the possibility of fenestration-based access of fipronil into the channel pore. Such cavities, laterally connecting the pore to the ion channel's surroundings, have been well-studied in a number of ion channel families (see Section 1.4), but in the superfamily of pentameric ligand-gated ion channels, no drug-permeable fenestrations have been identified thus far. Islam and Lynch's data suggested the possibility of such a fenestration, and this project was aimed at exploring this possibility further. This chapter reports the results of computational analyses carried out to this end.

4.1.2 The phenylpyrazole insecticide fipronil

Fipronil (Fig. 4.1) is a phenylpyrazole insecticide. Its insecticidal activity is a result of its ability to act as a channel blocker of pentameric ligand-gated ion channels[260], most prominently the "resistance to dieldrin" (RDL) GABA-gated chloride channel[198]. Its use as a pesticide is enabled by its lower affinity for vertebrate $GABA_A$-receptors compared with insect RDL receptors (IC_{50}s: 240 nM for RDL, 1.6 µM for rat $GABA_AR$[198]). However, its use in the field is complicated by its toxicity to useful pollinators like the honeybee[261,262].

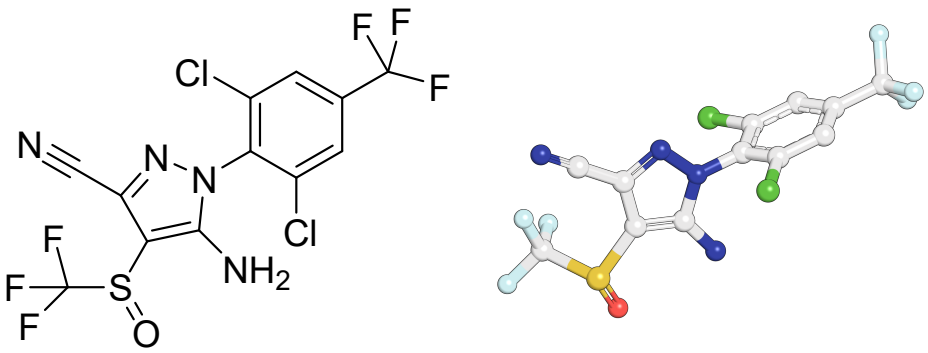

Figure 4.1: The phenylpyrazole insecticide fipronil.

Even though the primary amine substituent on the pyrazole ring might appear like a candidate for protonation, it does not act as a base under physiological conditions. One reason for this is that the nitrile substituent on the pyrazole ring exerts both a negative inductive (-I) and a negative mesomeric (-M) effect. This causes the nitrogen atom's lone pair of electrons to be localised away from it and towards the pyrazole ring, rendering it unavailable for protonation. As a result, fipronil does not show dissociation, and no pK_a value is reported[263]. In fact, its already poor water solubility is slightly lower still at pH 5 than at pH 9 (1.9 mg / mL vs 2.4 mg / mL)[264]. Its hydrophobicity is also seen in its logP of 4.0[264].

4.1.3 Fipronil is an inhibitor of human glycine receptors

Islam and Lynch demonstrated that a number of different subunit compositions of the human glycine receptor are inhibited by fipronil. A brief 5 s application

of fipronil in the open state caused rapid channel closure (Fig. 4.2 A). The majority of data points reported in this paper were obtained from closed-state drug applications. Here, fipronil was applied in the absence of glycine for 30 s. This resulted in inhibition of the channel; its sensitivity to following applications of the neurotransmitter was reduced (Fig. 4.2 B). As a result, the currents I elicited by various concentrations of glycine relative to the uninhibited current I_{max} were reduced after incubation with fipronil (Fig. 4.2 C). Islam and Lynch have thus provided evidence of inhibition of the human glycine receptor mediated by this inhibitor of the insect GABA receptor.

4.1.4 Fipronil binds in the channel pore

Having established that fipronil is an inhibitor of human glycine receptors, the authors went on to narrow down its binding site on the channel through site-directed mutagenesis. The 2′ and 6′ positions on the intracellular half of the M2 segment were considered for mutation. Various mutations of these residues made the glycine currents insensitive to fipronil (Fig. 4.2 D, E, F), suggesting in agreement with other studies[265,266] that the fipronil binding site is located within the intracellular half of the pLGIC pore.

4.1.5 Evidence for the importance of the alcohol binding site

The inter-subunit interface in the transmembrane domain (Fig. 4.3 A) is the location of the binding site of avermectins and various anaesthetics (see Section 1.2.3). These groups of compounds are activators of the glycine receptor, potentiating and/or outright inducing chloride currents and thus acting as inhibitors of action potentials. Islam and Lynch noticed that on addition of fipronil to open-state glycine receptors, a small increase in current was visible before inhibition set in (Fig. 4.2 A). This was particularly apparent when lower glycine concentrations are applied (Fig. 4.2 G).

In another experiment, fipronil was applied in the closed state for only 5 instead of 30 s (Fig. 4.2 H). It was then washed out and immediately replaced

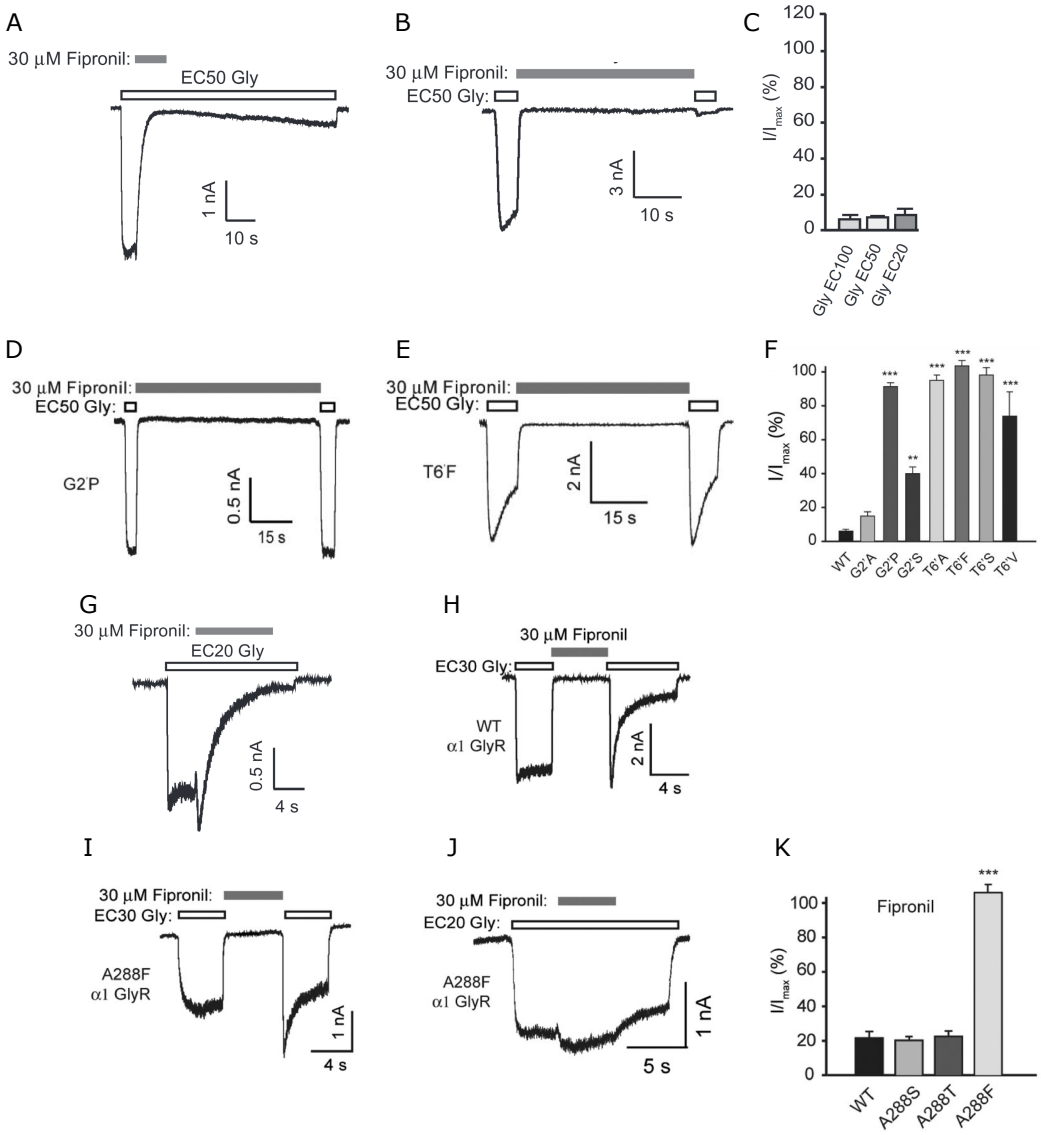

Figure 4.2: Whole-cell patch-clamp electrophysiology data reported by Islam and Lynch[90]. The individual plots are discussed in the main text. Reproduced here with permission from the rights holder, John Wiley and Sons, Inc.

with glycine. This 5 s treatment was insufficient for channel inhibition — glycine currents of similar magnitude as before were elicited. However, the channel was then rapidly inhibited, even though fipronil was no longer present in the external solution. Islam and Lynch concluded that fipronil likely has a secondary, non-blocking binding site, from where it can reach the inhibitory binding site in the pore once the channel opens.

Combined, the observation of transient strengthening of currents and the

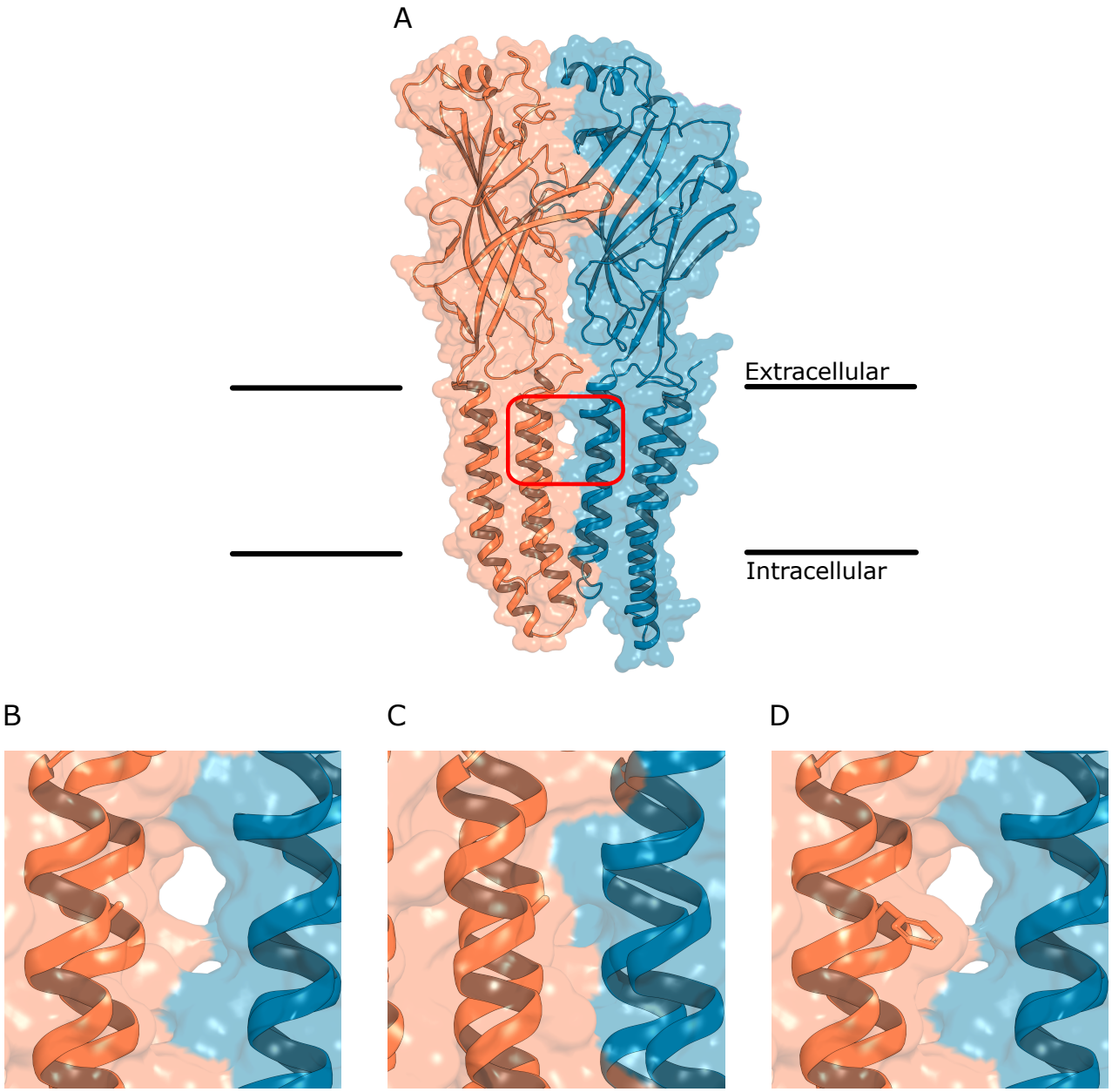

Figure 4.3: The location of the anaesthetics binding site on the glycine receptor. A: The binding site is located at the subunit interface in the transmembrane domain. The section marked with the red box is shown close-up for the desensitised state of the channel (PDB 3JAF, B), the closed state of the channel (PDB 3JAD, C), and the desensitised state with the A288F mutation (D, mutation introduced to 3JAF in PyMOL).

hypothesis of a secondary binding site led Islam and Lynch to further investigate the alcohol binding site near A288* (Fig. 4.3 B). The authors suggest that fipronil might bind at this site in the closed state and move from there into the pore as the channel opens, causing the transient current potentiation on the way. Further hints for the importance of this binding site were reported in a docking study by Law and Lightstone[267]. This paper was also referenced by Lynch's group in a prior publication where they first proposed, but did not test, this mechanism[268].

To test whether fipronil might bind in the alcohol binding site and move into the channel pore from there, Islam and Lynch tested point mutations of A288 in electrophysiology. While A288S and A288T showed no effect, the A288F mutation (Fig. 4.3 D) made the glycine receptor seemingly resistant to fipronil (Fig. 4.2 K). Brief 5 s applications in both the closed (Fig. 4.2 I) and open state (Fig. 4.2 J) failed to reduce currents. However, both traces still showed transient current potentiation, which led the authors to conclude that transition into the pore is blocked by the mutation, while interaction with the alcohol binding site is still possible. Data for longer applications like the 30 s closed-state application in Fig. 4.2 B is not reported for A288 mutations, all data in Figs. 4.2 I–K is based on 5 s applications.

4.1.6 Aims of this study

Islam and Lynch's hypothesis was not tested further in the years since its original publication. Their results are consistent with a re-interpretation of the important ivermectin and alcohol binding site as a fenestration, however. Structural evidence presents a large opening at the subunit interface in the transmembrane domain that connects the channel pore to its membrane environment (Fig. 4.3). In a structure of the homologue GluCl channel, this opening is fully occupied by a phospholipid protruding into the channel pore (PDB 4TNW[88]), showing its size is

*Throughout this chapter, references to residue A288 of the human glycine receptor α1 or its mutation A288F will be made. This designation by Islam and Lynch[90] is based on UniProt entry P23415 but does not include the 28 amino acid-long N-terminal signal peptide. To facilitate comparisons between different references, note that "A288" is equivalent to A316 in UniProt entry P23415, A304 in the structures of *Danio rerio* GlyR-α1 used in this study (PDB 3JAF, 3JAD[44]), and A312 in the corresponding UniProt entry O93430 for the zebrafish subunit.

sufficient to accommodate small molecules along its entire length and highlighting a potential role of lipid regulation of this channel family. Mutation of A288 (Fig. 4.3 B) to phenylalanine (Fig. 4.3 D) leaves the glycine receptor still functional and responsive to glycine but makes it seemingly resistant to fipronil (Figs 4.2 I–K). In light of this evidence, it is plausible that fipronil might predominantly or exclusively enter the channel pore through this fenestration, but is blocked from doing so by the bulky aromatic side chain of the A288F mutant. This is of high interest because, even though fenestrations are known to be important drug access pathways in a variety of ion channels (see Section 1.4), this mechanism has to date not been described in the pentameric ligand-gated ion channel superfamily.

Based on the available data, it is unlikely that fipronil would reach its binding site by moving down through the channel pore from the extracellular domain. If that were the primary access pathway, then no inhibition should have been observed after a short closed-state application of fipronil followed by wash-out and addition of glycine (Fig. 4.2 H), and the A288F mutation should not confer resistance.

In this study, the role of these possible fenestrations in the glycine receptor is investigated further with computational methods. The aims are to combine structural analyses with free energy estimates obtained from umbrella sampling to further support or refute the hypothesis of fenestration-based drug access first formulated by Islam and Lynch in 2012. Two different pathways are considered: Access to the pore through the fenestration, or membrane permeation followed by pore entry from within the cytoplasm (Fig. 4.4).

4.2 Methods

4.2.1 CAVER analysis of fenestration radii

55 glycine receptor structures were downloaded from the Protein Data Bank (PDB) and their fenestrations were measured with CAVER version 3.0[169]. The first step was the alignment of these structures, which was carried out through a PyMOL python script. Different labs use different naming conventions for

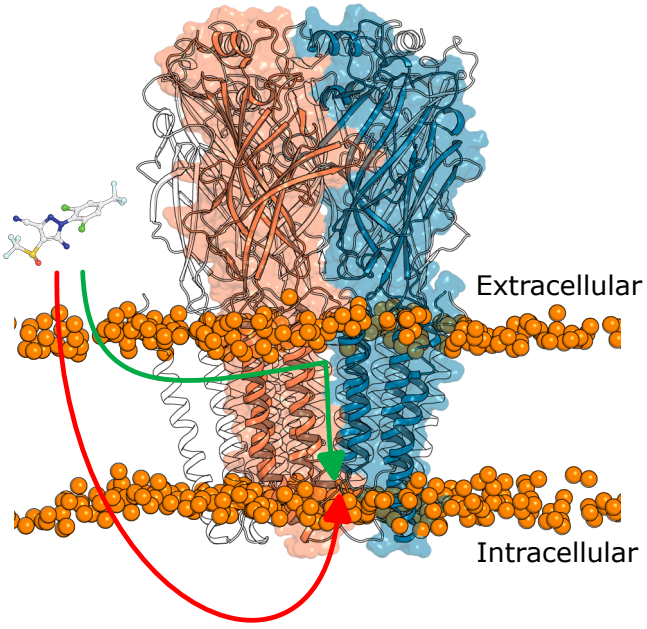

Figure 4.4: The two pathways considered in this study along which fipronil (not to scale) could enter the channel pore. It could either permeate through the membrane and reach its binding site from within the cell (red arrow) or, once in the membrane, go through the fenestration to reach the pore (green arrow). Phosphorus atoms of POPC are shown as orange spheres.

the five subunits of GlyR, labelling them A through E in either the clockwise or anticlockwise direction. As a result, neither the `super` nor the `align` commands in PyMOL produce good results across all structures, because both rely heavily on chain-based alignments. PyMOL's `cealign` command, which is based on the combinatorial extension algorithm[269] and does not consider chain indices, was used instead.

The structural alignment was a prerequisite for automated identification of tunnels with CAVER. The software allows specification of the origin of calculation through atom or residue indices or through Cartesian coordinates, but the former two would require extensive manipulation of the PDB files to ensure matching numberings. Instead, for each of the five fenestrations, Cartesian coordinates of a point in its centre were supplied as origins (Fig. 4.5 A, C). Due to the structural alignment, these initial coordinates were sufficient for application across all structures.

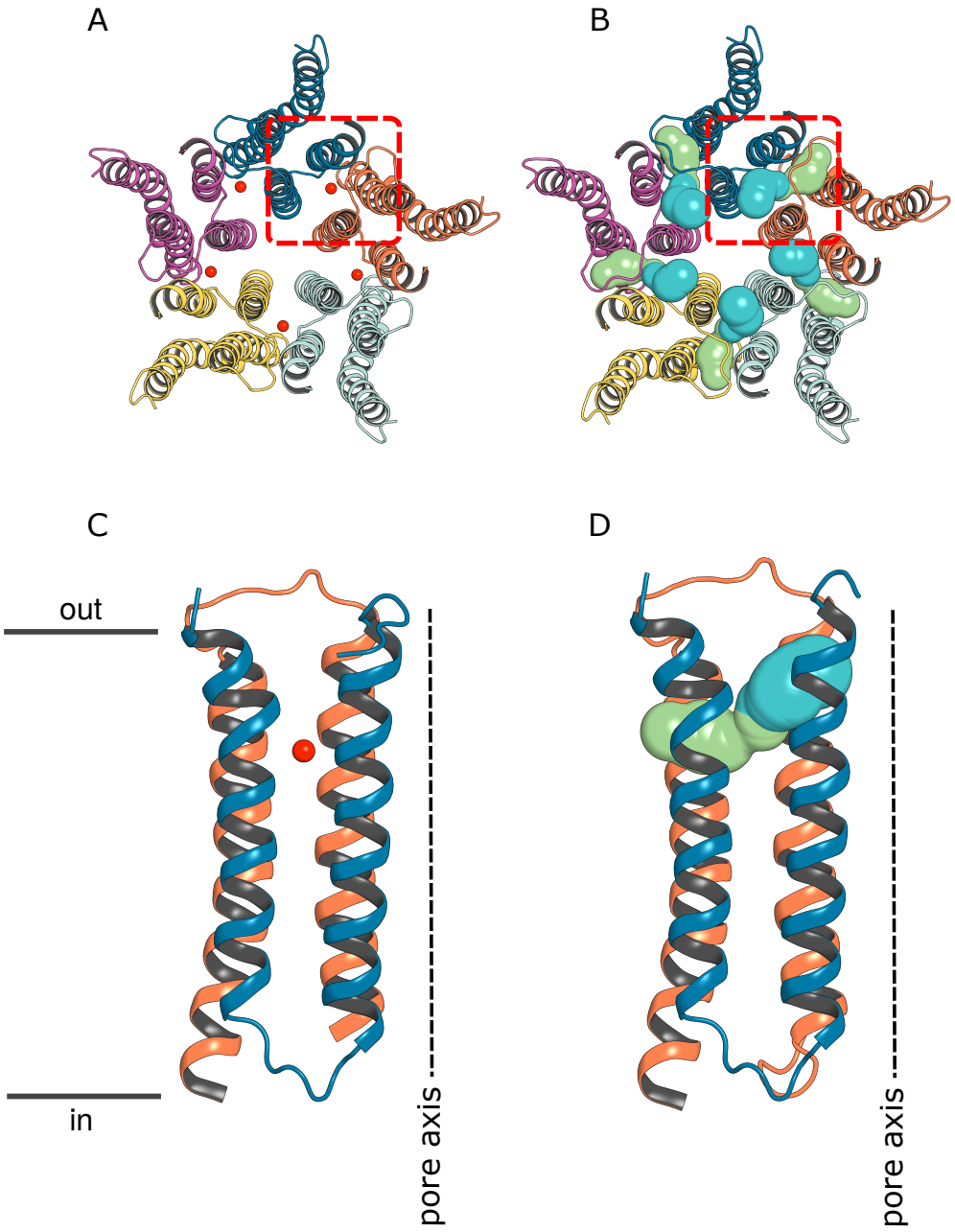

Figure 4.5: Illustration of the use of CAVER for the measurement of fenestrations. A: Top view onto the transmembrane domain of 3JAF. Coordinates in the centre of each fenestration were supplied as origins for the calculations (red spheres). These initial coordinates are then subject to optimisation. C: Side view of the highlighted section in A, showing the height of the fenestration in the subunit interface. B: Caver then identifies a number of tunnels from the origin to the pore (blue) and to the outside edge of the protein (green). The correct tunnels were visually identified and combined for analysis. D: Side view of the highlighted section in B, showing the extent of the fenestration CAVER identified.

CAVER uses a Voronoi diagram-based algorithm to identify tunnels in proteins. The Cartesian starting coordinates were subject to optimisation, in which a starting Voronoi vertex with a maximum distance from the origin of 3 Å and a desired cavity radius of 5 Å was chosen. From this starting vertex, the tool identified tunnels to the surface using a probe radius of 1.2 Å, a shell radius of 5 Å, and a shell depth of 4 Å. These parameters were chosen so that the channel pore would also be identified as protein surface, terminating the calculation here.

The algorithm is most robust when all spheres are the same size. To still properly account for different atomic radii, the space-filling characteristics of larger atoms were approximated by 20 overlapping spheres with the radius of the smallest particle in the system — hydrogen. Redundant tunnels were removed by distance-based clustering, and only unique tunnels at each fenestration were reported. Their radii were then measured by growing a sphere at each Voronoi vertex until van-der-Waals overlaps with the protein occurred. The tunnels that best match the fenestration were visually identified in PyMOL (Fig. 4.5 B, D). Generally, from the origin in the centre of the fenestration, one tunnel towards the pore and one towards the outer edge of the protein had to be combined to adequately describe the fenestration.

CAVER returns tunnel radius profiles, a list of tunnel-lining residues, and information about the narrowest point along the tunnel: the bottleneck. Comparisons across the 55 GlyR structures were then made based on the bottleneck radii of each of their five fenestrations. Here, the smaller bottleneck radius of each fenestration's two constituent tunnels was taken to be the fenestration's overall bottleneck radius.

4.2.2 System preparation

Two structures were used in simulations for this study, both reported in a 2015 study from the Gouaux lab[44]. To model the channel with open fenestrations, the desensitized-state structure 3JAF, which was solved in complex with glycine and ivermectin, was used. The strychnine-bound closed-state structure 3JAD was

used as a model for closed fenestrations. These structures' average bottleneck radii are 1.50 and 0 Å, respectively.

Both 3JAF and 3JAD are structures of the α1 glycine receptor subunit of *Danio rerio* (UniProt accession number O93430). This gene includes an N-terminal signal peptide of 24 amino acid residues (MFALGIYLWETIVFFSLAASQQAA) that is cleaved during cellular localisation and is not resolved in the structures. The authors used a slightly modified gene construct in their work: the numbering starts with the signal peptide, and the first resolved residue is A25. Residues R26–P33 were deleted from the construct, as were the C-terminal I435–Q444. Instead, residue K434 is followed by an LVPRS thrombin cleavage site, which is resolved until the arginine. In addition, the largely disordered intracellular loop between transmembrane segments M3 and M4 (UniProt Q334–R400) is replaced by an AGT tripeptide linker. As a result, the residue numbering differs between these structures from the Gouaux lab, the UniProt sequence of the protein, and other structures of the channel. For clarity, residue numbers will be given for both the numbering used in the structures and in UniProt entry O93430.

Two different system scales were investigated. In full-size simulations, all residues present in 3JAD and 3JAF were included in simulation systems (roughly UniProt O93430 A25–K434, but with the deletions and substitutions outlined above), while simulations of the transmembrane domain only included residues of 3JAD and 3JAF from Q235 onwards (UniProt O93430 Q243). The N- and C-terminal residues of both structures and both system sizes were capped with acetyl and N-methyl groups, respectively, to avoid introducing terminus charges in locations where there are none in the full-length protein. Some residues were modelled incompletely, and the AGT linker is missing in the 3JAD structure. These gaps were filled in with PDBFixer[270].

Both full-length and TMD-only proteins were embedded in a pure POPC bilayer consisting of 512 lipids with the inflateGRO method[271] and solvated with water containing 0.15 M NaCl to form rectangular boxes. After steepest-descent energy minimisation and equilibration in the NPT ensemble at 310 K and 1 bar for 100 ns,

simulations of the full-length system were started with and without glycine present in its binding site (Fig. 4.6 A, C). The TMD-only systems were modified further: without the extent of the protein away from the membrane, less membrane area is needed as well. To maximise simulation performance, the rectangular system was trimmed down to a hexagonal prism using in-house MDAnalysis[166,167] scripts (Fig. 4.6 B, D). This effective reduction of system size helped significantly improve performance of the umbrella sampling simulations without sacrifices. System compositions are listed in Table 4.1.

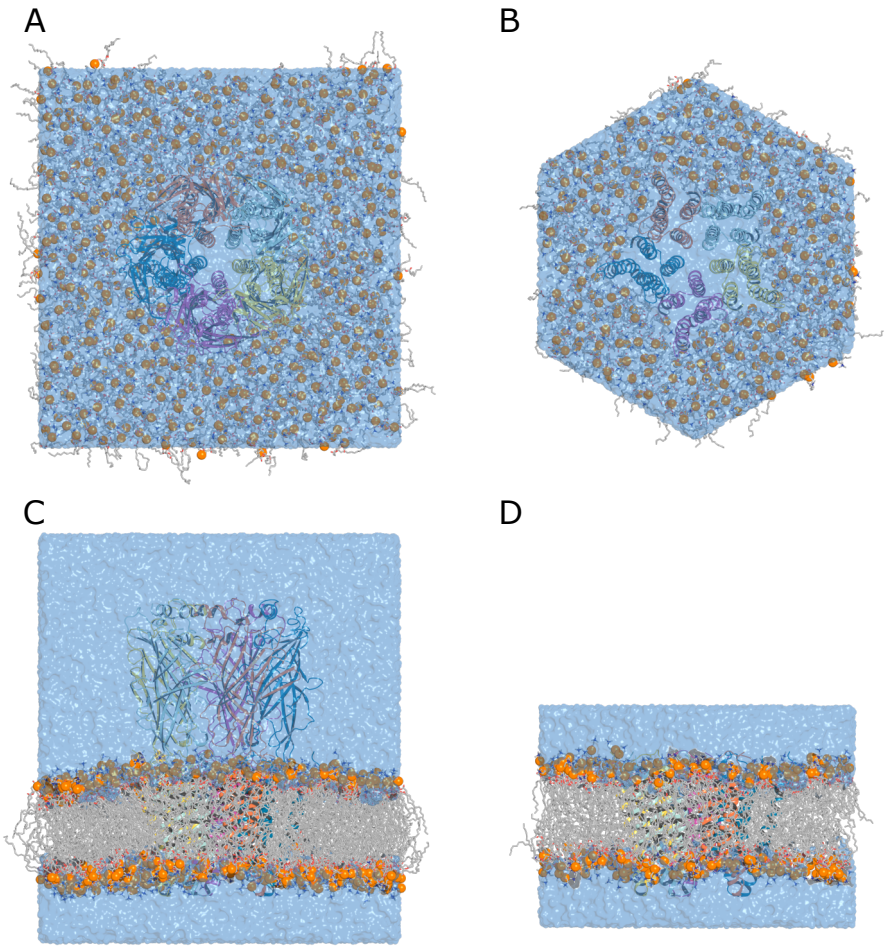

Figure 4.6: Examples of simulation system scales investigated in this study. Full-length (A, C) and TMD-only (B, D) systems were used. Shown are top-views (A, B) and side-views (C, D).

Table 4.1: Composition of the systems used in this study.

	3JAF full	3JAF TMD	3JAF-A288F TMD	3JAD TMD
# POPC	500	281	279	285
# H_2O	65,518	16,922	16,898	16,471
# Na^+	272	100	100	100
# Cl^-	282	140	140	135
# Glycine	5 or 0	0	0	0
# Fipronil	0	1	1	1
# Atoms	291,888	99,590	99,300	98,493
Box volume [nm^3]	2,875	969	967	954

Umbrella window generation

The windows of the initial umbrella sampling attempt were extracted from steered MD trajectories with in-house MDAnalysis[166,167] scripts.

After that approach failed, umbrella windows of fipronil going through two different fenestrations were set up for each TMD-only system with Alchembed. Vectors going through each fenestration were derived in PyMOL (Fig. 4.7 A). Fipronil was then aligned to these fenestration vectors and translated to the vector origin in the channel pore (Fig. 4.7 B). Care was taken that the rotation of fipronil about the fenestration vector matched between different fenestrations. The umbrella windows were then generated by translating fipronil along the fenestration vector in steps with a stride of 1 Å (Fig. 4.7 C). Finally, the Alchembed-based method (see Section 3.3.3) was used to resolve any clashes between fipronil and the system (Fig. 4.7 D).

For the determination of a membrane-permeation PMF, fipronil was placed along the membrane-normal of a pure POPC bilayer using the Alchembed-based method and a stride of 1 Å.

4.2.3 Simulation parameters

All simulations in this study were carried out in the NPT ensemble. A temperature of 310 K and a pressure of 1 bar were maintained with the V-rescale thermostat[151] and the C-rescale barostat[152] with coupling constants of 1 and 5 ps, respectively,

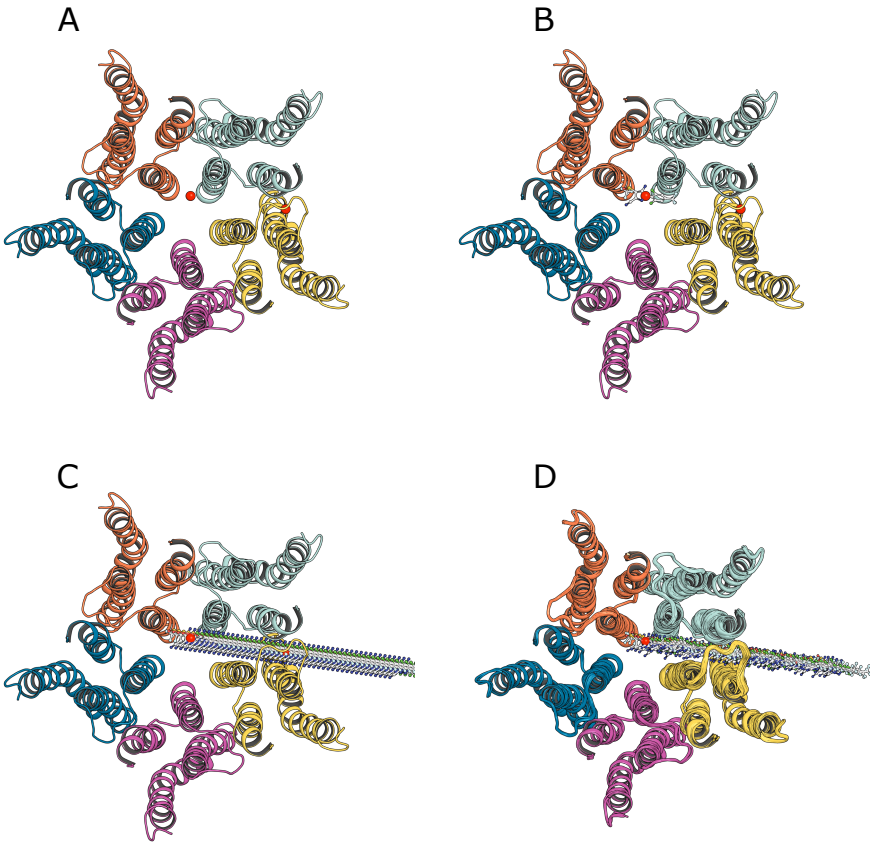

Figure 4.7: To generate the umbrella windows for fipronil passing through fenestrations, two points on the fenestration vector are identified in PyMOL (A, red spheres). Fipronil is aligned to this vector and translated onto the origin (B). Windows are generated by translating fipronil along the vector in increments of 1 Å (C). Finally, Alchembed simulations are run to resolve the clashes. Only minimal movements of fipronil and the rest of the system are necessary to arrive at starting configurations for umbrella sampling (D). Subfigure D shows an overlay of all 47 umbrella windows — the variations between windows are small.

unless otherwise stated. The chosen force fields were AMBER99SB-ILDN for protein and ions[146], TIP3P for water[223], SLipids for POPC[221,222], and GAFF2 for fipronil and the zwitterionic glycine ligand[224]. Bonds to hydrogen atoms were constrained with either LINCS[147] or, in the case of water, the SETTLE algorithm[148] to allow the use of an integration time step of 2 fs. The smooth particle-mesh Ewald approach[149] was used for treatment of long-ranged electrostatic interactions. Short-ranged interactions were calculated under a Verlet cut-off scheme.

Unbiased simulations of full-length channels

Unbiased simulations of full-length 3JAF were carried out with and without glycine in its binding sites (Table 4.1). In accordance with its pK_as of 2.35 for the carboxylic acid group and 9.78 for the ammonium group[272], it was modelled as a zwitterion. 1 µs simulations were carried out in triplicate.

Steering fipronil through a fenestration

Initial probing of the permeability of these fenestrations to fipronil was carried out with steered MD. For this purpose, the full-length channel was oriented such that one of its fenestrations was exactly parallel to the x-axis. Fipronil was positioned about 3 nm away from the channel pore on the x-axis in the membrane using the energy minimisation method previously used by Cetin et al.[161] (See Section 2.3.1). A steered MD trajectory was then produced, in which fipronil was pulled towards the centre of the channel pore with a pull rate of 0.01 nm ns^{-1} and a force constant of 1000 kJ mol^{-1} nm^{-2}. During the steer, a cylindrical flat-bottom restraint was applied to fipronil that added a harmonic restraining force of 1000 kJ mol^{-1} nm^{-2} if it deviated from the pull axis by more than 1 nm. The centres of mass of the five protein subunits were restrained to their starting value with GROMACS's pull code, applying harmonic restraints of 1000 kJ mol^{-1} nm^{-2}. This was necessary to prevent rotations of GlyR within the membrane that would move the fenestration off the pull axis. Simulations of reverse pulls that steered fipronil back out of the channel pore were started from the endpoints of successful forward pulls.

Umbrella sampling

Initial umbrella sampling attempts were started from configurations extracted from the steered MD simulations. For this purpose, the range of the reaction coordinate that was sampled in the trajectory was divided into 80 evenly spaced windows and those frames closest to each window were extracted. Each umbrella window was sampled with standard umbrella sampling for 20 ns during which

the system was restrained to the window's starting reaction coordinate with a harmonic umbrella potential of 1000 kJ mol^{-1} nm^{-2}.

The above approach was quickly abandoned. Instead, Simulated Tempering-enhanced Umbrella Sampling (STeUS, see Section 3.3.4) simulations of Alchembed-based umbrella windows was carried out. A temperature ladder spanning 310 to 342 K in 4 K increments was used. Care was taken that this choice of highest temperature did not cause the protein to unfold, but it was noticed that the protein undergoes significant rotations in the membrane at these temperatures. To prevent this, the centres of mass of each protein subunit were restrained to their starting position as described above.

The reaction coordinate was defined in reference to the fenestration vectors used for window generation by setting `pull-coord-geometry = direction` and supplying the fenestration vector as `pull-coord-vec`. Since the length of the projection of fipronil onto this vector is thus defined as the reaction coordinate, the compound must not move too far away from this axis. This is ensured through cylindrical flat-bottom restraints, but GROMACS can only apply such restraints along the three base coordinate axes of the system and not along arbitrary vectors. Instead, the PLUMED package[273,274] was used for restraint application. This was achieved by defining a vector between the protein's centre of mass and a `FIXEDATOM`-type dummy atom with coordinates in the fenestration centre. The `PROJECTION_ON_AXIS` function was used to calculate the extension of fipronil from this vector, and `UPPER_WALLS` was used to restrain this extension to below 1 nm with a force constant of 1000 kJ mol^{-1} nm^{-2}. To sufficiently sample each system, umbrella sampling of fipronil through two different fenestrations and in two ligand orientations was carried out. The simulation data of these four runs was combined and used for PMF calculations with WHAM. Unfortunately, the systems studied here are too large to meaningfully benefit from temperature reweighting with MBAR (see Section 3.3.5).

4.3 Results

4.3.1 The fenestration is state-dependent

As a first step, a better structural understanding of the fenestrations was required. To this end, the fenestrations of 55 glycine receptor structures were investigated with CAVER version 3.0. The workflow is illustrated in Fig. 4.5. These 55 structures were solved with a variety of activating and inhibiting ligands bound and thus cover a broad range of activation states. They were grouped into four distinct activation state groups — closed, open, desensitized, and wide-open — as reported by their original publications. The distribution of average bottleneck radii is shown in Fig. 4.8. The complete bottleneck radius data is listed in SI Table C.1.

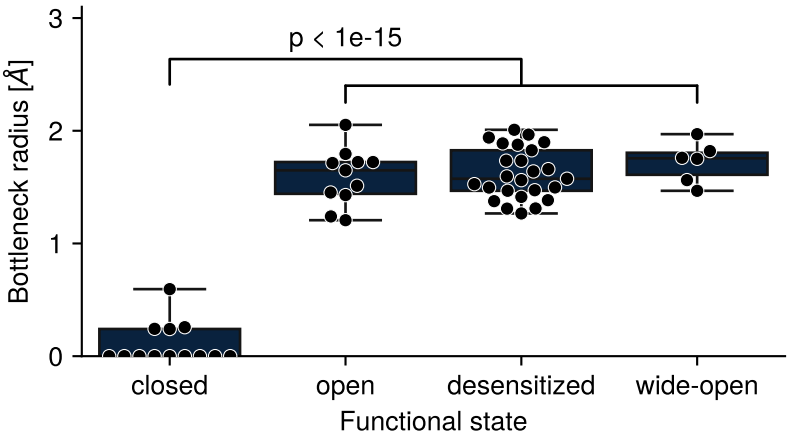

Figure 4.8: Box plot showing the distributions of the average bottleneck radius by functional state of the channel, with individual data points marked as dots. A clear, bi-modal state dependence is apparent: the fenestrations in closed-state structures are closed, while open and desensitized structures have fenestrations with bottleneck radii between 1.5 and 1.9 Å. This difference is highly statistically significant as determined by Tukey's HSD test[275].

Strikingly, a strong state-dependence of the fenestration radii is apparent. While displaying similar radii in open, wide-open, and desensitised states (Fig. 4.3 B), the fenestration all but disappears when the channel is in the closed state (Fig. 4.3 C). The structural rearrangements that occur during channel activation to connect ligand binding in the ECD to pore widening in the TMD are now shown to

also open this fenestration. Statistical significance was determined with Tukey's HSD test[275] and strongly confirmed.

The context of state-dependence of these fenestrations reinforces the hypothesis of fenestration-based access formulated from Islam and Lynch's results. Their electrophysiology data for brief applications of fipronil in the closed state showed that the compound cannot reach its binding site in this condition: on addition of glycine, the channel initially opens fully. However, it then rapidly closes, even though fipronil is no longer in the surrounding medium. As we have now shown, when the channel is in the closed state, the fenestrations are also closed. Our hypothesis for Islam and Lynch's results thus is that fipronil, being very hydrophobic, gets stuck in the membrane and is slow to be washed out in electrophysiology. It can only reach its binding site when glycine is added and the channel opens, as this causes the fenestrations to open as well. In the time it takes fipronil to permeate the fenestration, GlyR currents are observed, which then rapidly decrease as fipronil reaches its binding site.

4.3.2 The channel is stable in unbiased MD simulations

The simulations of the full-length glycine receptor showed that the protein is stable at these temperatures, pressure, and membrane composition. Overall, glycine remains stably bound. It was thus deemed suitable for simulations probing the interaction of fipronil with the fenestrations.

4.3.3 The fenestration allows fipronil permeation in steered MD

Having established their state-dependence, we set out to test the fenestrations for permeability. Initial investigations of fipronil permeation through the fenestrations were carried out using steered MD simulations. A moving harmonic bias was applied to fipronil to pull it through the fenestration, both from outside the channel towards the pore (Fig. 4.9) and out of the channel back into the membrane environment (Fig. 4.10).

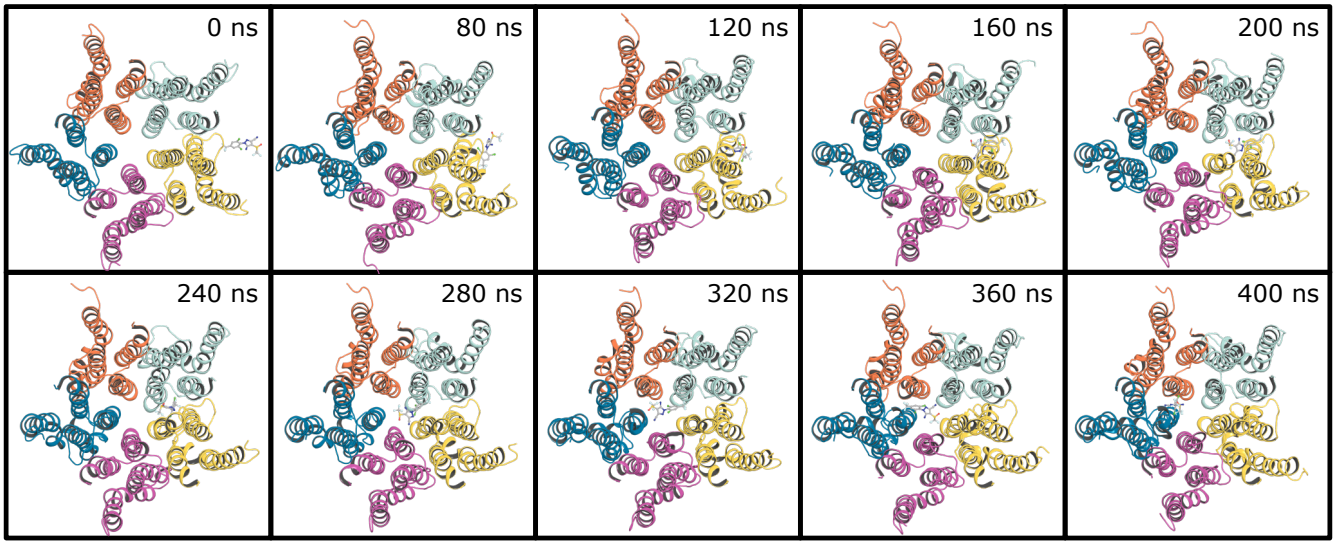

Figure 4.9: Snapshots from a steered MD trajectory in which fipronil was pulled through a fenestration towards the channel pore. Fipronil is able to pass through the fenestration without overly disturbing the overall structure of GlyR.

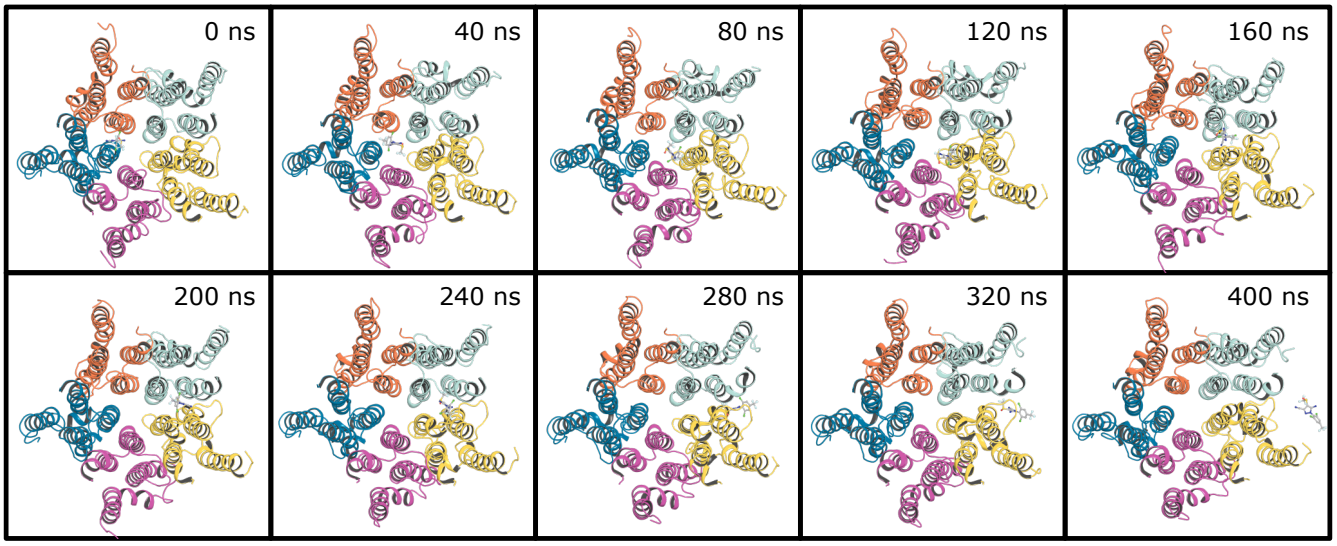

Figure 4.10: Snapshots from a steered MD trajectory that pulled fipronil out of the fenestration. This simulation was started from the final frame of the simulation shown in Fig. 4.9. The overall GlyR structure remains largely undisturbed by fipronil passing through the fenestration.

The centre of mass of each protein subunit is harmonically restrained to its starting value, but no other position restraints acted on the protein atoms during these simulations. Nonetheless, the overall protein structure remained remarkably stable, indicating that the fenestration is able to accommodate fipronil without requiring major structural reorientations. This constitutes the first hint from simulations in this work that fenestration-based drug access might play a role in the glycine receptor.

4.3.4 Steered MD-based umbrella sampling suffers from hysteresis and convergence issues

The steered MD simulations suggested the absence of major hurdles to fipronil-permeation of the fenestration. Umbrella sampling was used to quantify the energy barrier of permeation. Windows extracted from steered MD show a degree of structural heterogeneity while maintaining the overall conformation of the channel (Fig. 4.11).

PMFs obtained from the application of WHAM to 20 ns of simulation data from each window of the inwards steer-based simulations yield a profile that steadily increases the closer fipronil gets to the centre of the pore, with an overall energy difference of roughly 50 kcal mol^{-1} (Fig. 4.12, blue line). This is very high compared to kT at 310 K ($\approx$ 0.62 kcal mol^{-1}), thus suggesting a very low probability of fipronil permeation of the fenestrations.

However, significant systematic issues are apparent in this data. For one, umbrella sampling simulations started from the reverse steered-MD run similarly show an energy difference of roughly 50 kcal mol^{-1} but in the opposite direction (Fig. 4.12, red line). This very high degree of hysteresis reveals underlying sampling issues. In addition, a convergence analysis showed that 20 ns of simulation time per window are insufficient and the PMFs are not yet converged (Fig. 4.13).

Combined, the hysteresis and convergence issues made it clear that the data shown in Fig. 4.12 are not the true potentials of mean force. Much more simulation data was evidently needed, but these simulations are expensive. The problem

Figure 4.11: Overlay of the 80 equidistant umbrella windows extracted from the inwards pull (Fig. 4.9). The reaction coordinate is defined as the centre of mass distance between the protein and fipronil along the x axis, to which the fenestration is aligned. Note that there is a much larger degree of between-window variation than in the Alchembed-based windows used later in this study (Fig. 4.7).

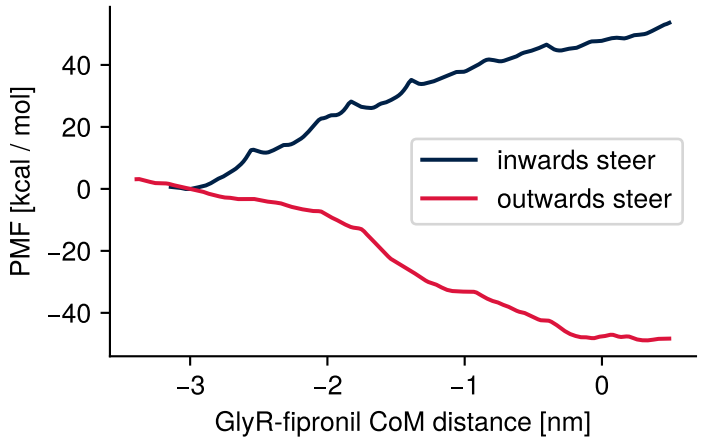

Figure 4.12: Comparison of the potentials of mean force obtained from steered MD-based umbrella windows from an inwards-pulling and an outwards-pulling trajectory. Strong hysteresis is apparent. A difference in PMF in the channel pore (at a distance of 0 nm) of roughly 100 kcal mol^{-1} is observed.

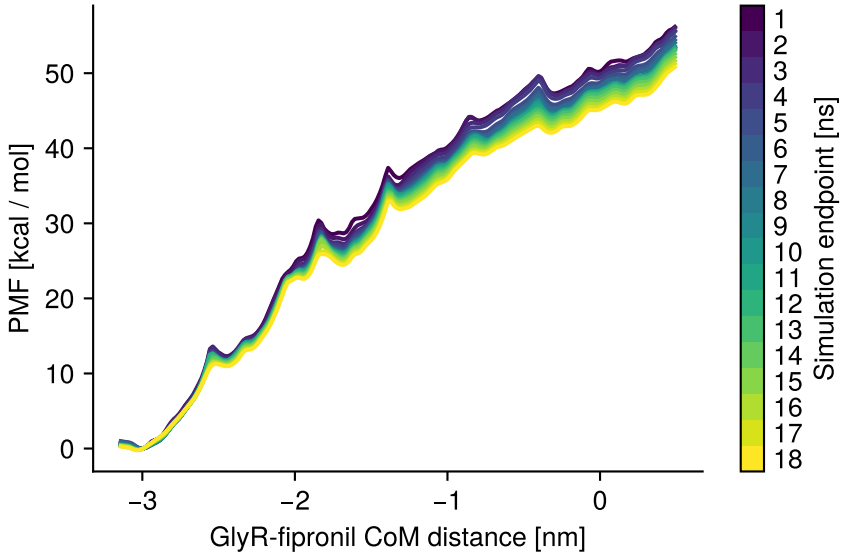

Figure 4.13: Convergence analysis of the inward pull-based umbrella sampling simulations. Each window was simulated for 20 ns. The first 2 ns were discarded in WHAM to allow for equilibration of the system. The colour bar shows how many nanoseconds of simulation data were included in the PMF calculation after these initial 2 ns. This analysis revealed that the PMF had not converged after 20 ns of simulations.

was approached from two directions. For one, simulation performance was improved by reducing the system size to its minimum (Fig. 4.6). More importantly, the slow convergence speed observed in this system formed the motivation for the development of the improved umbrella sampling workflow that is described and tested in chapter 3.

4.3.5 Membrane-permeation of fipronil

The improved umbrella sampling workflow quickly proved successful. One of its first applications was the calculation of the PMF of fipronil permeating through a POPC bilayer. This PMF is the baseline against which PMFs for fenestration permeation need to be compared, as the two most likely paths for fipronil to reach its binding site are to enter via the fenestration or to permeate the membrane and access the channel pore from the intracellular end (Fig. 4.4).

Fipronil is a hydrophobic molecule, so an energy cost when leaving the membrane for an aqueous environment is expected. The PMF obtained with

the improved umbrella sampling method shows that this energy cost amounts to 10 kcal mol^{-1} (Fig. 4.14). Comparisons between methods show the superiority of the improved workflow (Fig. C.1), and the PMF is well-converged (Fig C.2).

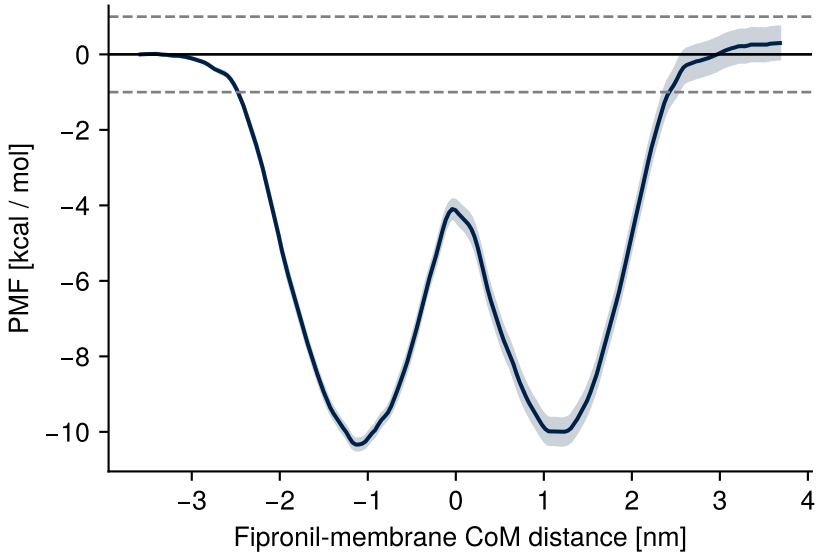

Figure 4.14: Potential of mean force of fipronil permeating through a pure POPC bilayer. The shaded area is an error estimate obtained from 200 bootstraps. The grey dashed line indicates chemical accuracy. The symmetry of the profile overall and the similar values obtained for bulk water on either end suggest a high-quality result. The PMF is well-converged (Fig. C.2).

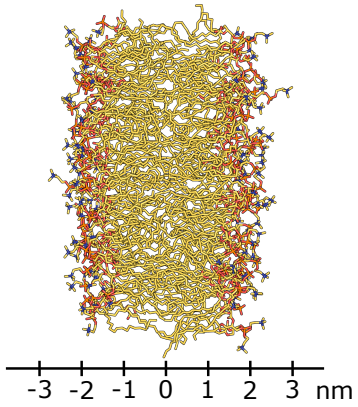

Figure 4.15: Reaction coordinate of the PMF shown in Fig. 4.14

Thus, if an energy barrier to fenestration-based entry of less than 10 kcal mol^{-1} is found, then drug access through the fenestration is predicted to be energetically favourable and thus more likely.

4.3.6 Fenestration-based PMFs of fipronil

The improved umbrella sampling workflow was then applied to the three TMD systems listed in Table. 4.1. The aim was to obtain potentials of mean force for fipronil going through open and closed fenestrations and through open fenestrations that carry the A288F mutation found to be less sensitive to fipronil by Islam and Lynch[90]. These PMFs can then be compared with one another and with the available experimental data.

In all three cases, fipronil's permeation through the fenestrations formed by the transmembrane domains of chains C and D and chains E and A was simulated. The reaction coordinate is visualised in Fig. 4.16 A. Each fenestration has two main constriction sites, one formed from each subunit's M2 helices, and one from the M3 and M1 segments. The fenestration is located roughly 1 nm above the centre of the membrane, placing fipronil close to its minimum-energy position within the membrane (Fig. 4.14).

Open fenestrations

The PMF of fipronil going through the open fenestrations of desensitised *Danio rerio* GlyR (PDB 3JAF) is shown in Fig. 4.17.

Due to the complicated nature of this system, the PMF has unfortunately not yet converged, as can be seen from the offset from zero at a distance of 4.5 nm. Nonetheless, conclusions can be drawn from this profile already. Approaching the channel pore from the membrane environment at ±4.5 nm, fipronil does not have to pass any notable energy barriers. The only minor barrier is seen at the M2-M2 constriction of the CD fenestration at −1 nm. However, at 1.2 kcal mol^{-1}, it is on the order of $2\ kT$ and thus is not expected to meaningfully slow down fipronil permeation, especially since it follows from a drop of $\approx\ 16\ kT$. This result suggests that permeation of fipronil through open fenestrations does not have an energy cost.

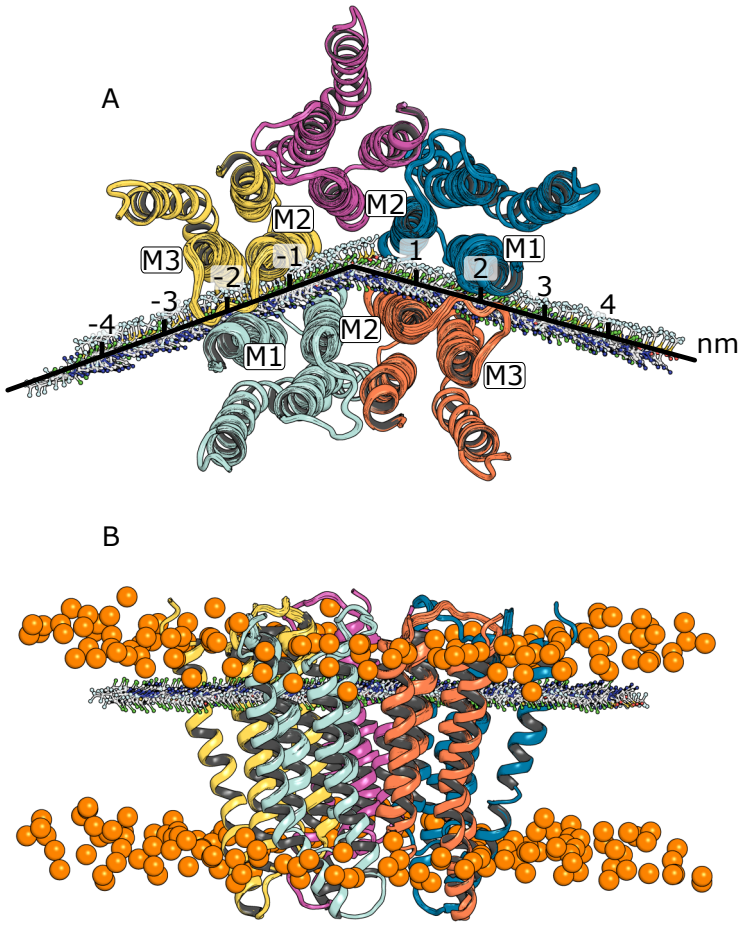

Figure 4.16: A: Reaction coordinate of the fenestration system. Negative reaction coordinate values cover the fenestration at the CD interface, positive values the EA fenestration. The M2-M2 constriction lies at about ±1 nm and the M3-M1 constriction at ±2 nm. B: Side view onto the reaction coordinate. The fenestration is located above the membrane centre at the equivalent of ±1 nm in the membrane PMF system (Figs. 4.14 and 4.15), close to fipronil's minimum energy value.

Closed-state fenestrations

The umbrella sampling results for fipronil passing through closed-state fenestrations of *Danio rerio* (PDB 3JAD) is shown in Fig. 4.18.

Like the PMF for open fenestrations (Fig. 4.17), this one has not converged yet, but careful interpretation of the data collected thus far is still possible. Strikingly, energy barriers are identified in both M3-M1 and M2-M2 constrictions that were not seen in the case of open fenestrations. The largest energy barriers for the CD and EA fenestration are read as 5.2 kcal mol^{-1} ($\approx 8\ kT$) and 5.6 kcal/mol ($\approx 9\ kT$),

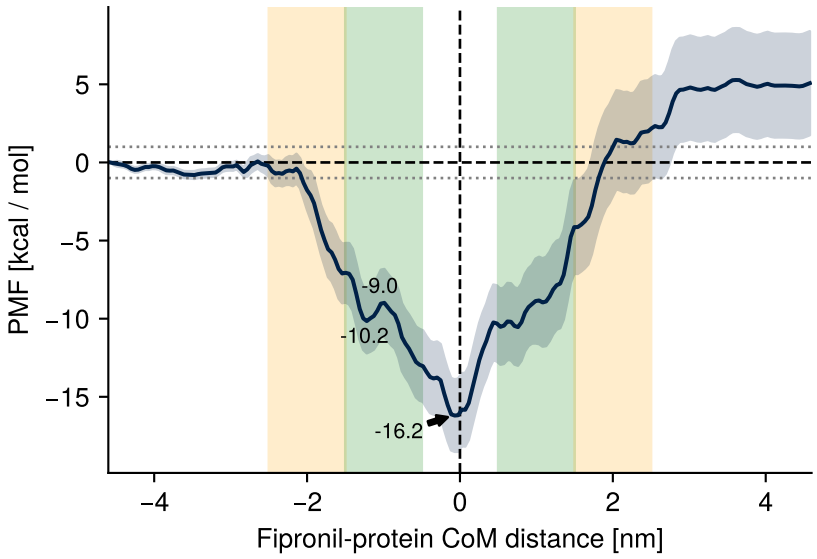

Figure 4.17: Potential of mean force of fipronil permeating through open fenestrations of the glycine receptor. The vicinity of the M2-M2 constriction and the M3-M1 constriction are indicated in green and yellow, respectively. The shaded area represents an error estimate obtained from 200 bootstraps. Numbers on the plot are PMF values in kcal / mol at that point.

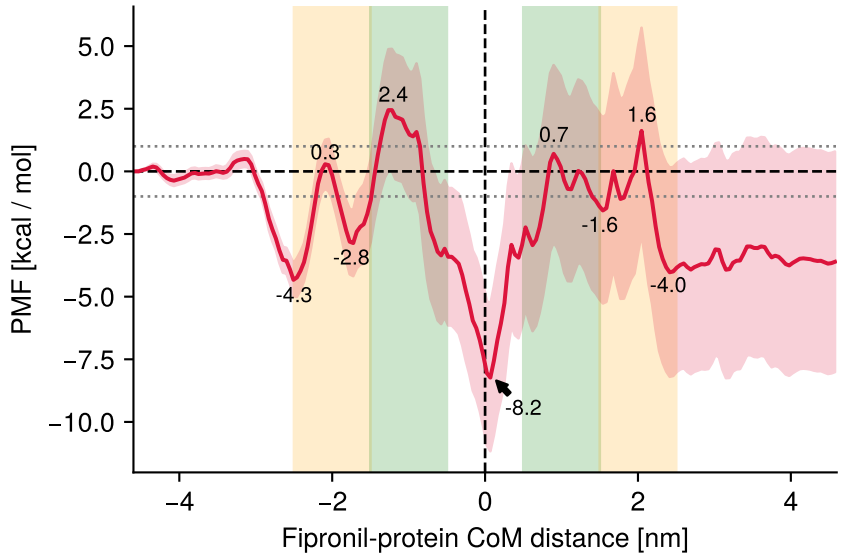

Figure 4.18: Potential of mean force of fipronil permeating through closed fenestrations of the glycine receptor. The vicinity of the M2-M2 constriction and the M3-M1 constriction are indicated in green and yellow, respectively. The shaded area represents an error estimate obtained from 200 bootstraps. Numbers on the plot are PMF values in kcal / mol at that point.

respectively. In addition, the energy well in the channel pore is shallower than that of the desensitised state by 8 kcal mol^{-1}, demonstrating the unfavourable effect of increased pore constriction. Though the absolute values of these energies will change as the PMF converges, a qualitative comparison with the PMF for open fenestrations clearly shows that the closed state's fenestrations are less accessible for drug permeation.

The effect of A288F in open fenestrations

The PMF obtained for open fenestrations with the A288F mutation is less readily interpretable (Fig. 4.19).

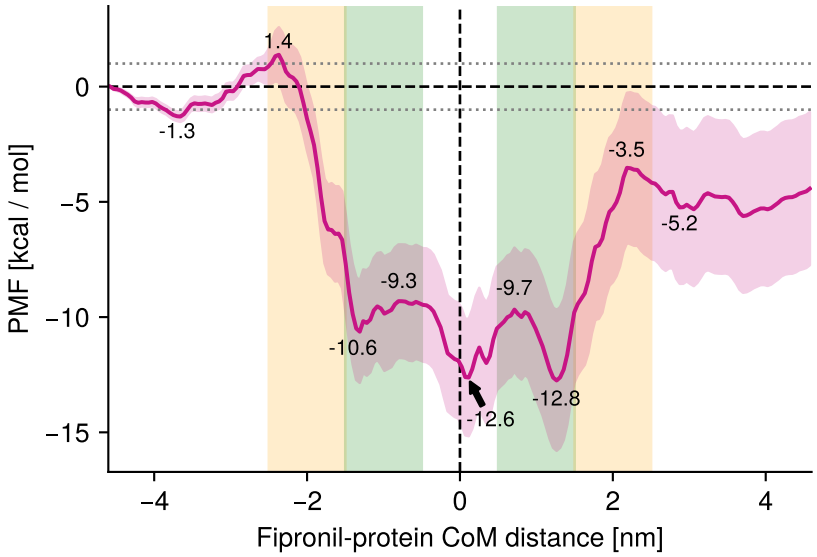

Figure 4.19: Potential of mean force of fipronil permeating through open-state fenestrations of the glycine receptor with the A288F mutation. The vicinity of the M2-M2 constriction and the M3-M1 constriction are indicated in green and yellow, respectively. The shaded area represents an error estimate obtained from 200 bootstraps. Numbers on the plot are PMF values in kcal / mol at that point.

The preliminary, unconverged PMF problematically predicts that the global energy minimum in this system is between the M2-M2 and M3-M1 constrictions of the EA fenestration rather than in the channel pore. However, encouragingly, energy barriers at A288F near the outer edge of the M3-M1 constriction are observed. The apparent energy barriers for either fenestration at A288F are

2.7 kcal mol^{-1} ($\approx 4\ kT$) and 1.7 kcal mol^{-1} ($\approx 3\ kT$) for CD and EA, respectively. However, the apparent dominant energy barrier in the EA fenestration is actually in the M2-M2 constriction at 3.1 kcal mol^{-1} ($\approx 5\ kT$). The PMF estimate in the pore differs from that obtained for the wild type by 3.6 kcal mol^{-1}, which is within the error estimate obtained through bootstrapping.

Diagnosis of convergence issues

To diagnose the convergence issues further, histogram counts and histogram overlap were determined with Dr Owen Vickery's `umbrella_sampling` analysis script[276]. The analysis is shown in Fig. 4.20.

The aim of this analysis is to reveal regions of the reaction coordinate that have been undersampled thus far. The reaction coordinate is divided into 94 windows with 1 Å spacing, and each window has been run for 150 ns. As it was expected that fipronil would be unable to freely rotate within the confinements of the fenestrations, the innermost 26 windows were run twice, with opposite relative fipronil orientation in each set of windows. The effect of this is seen clearly in Fig. 4.20 as reduced count and overlap values for reaction coordinates outside of the range of −2.5 to 2.5 nm. Strikingly, while the desensitised 3JAF-WT system shows good overlap for the entire inner section, 3JAF-A288F and, in particular, the closed-state 3JAD show stretches of low overlap even in this doubly sampled region.

To improve convergence of these three potentials of mean force, a three-pronged approach will be employed in the future. The simulation time of all windows will be increased. The outer regions of the reaction coordinate will be treated like the inner region and doubly sampled with flipped initial ligand orientations. And additional windows with higher restraining potentials will be started at regions of the reaction coordinate that show strongly reduced overlap.

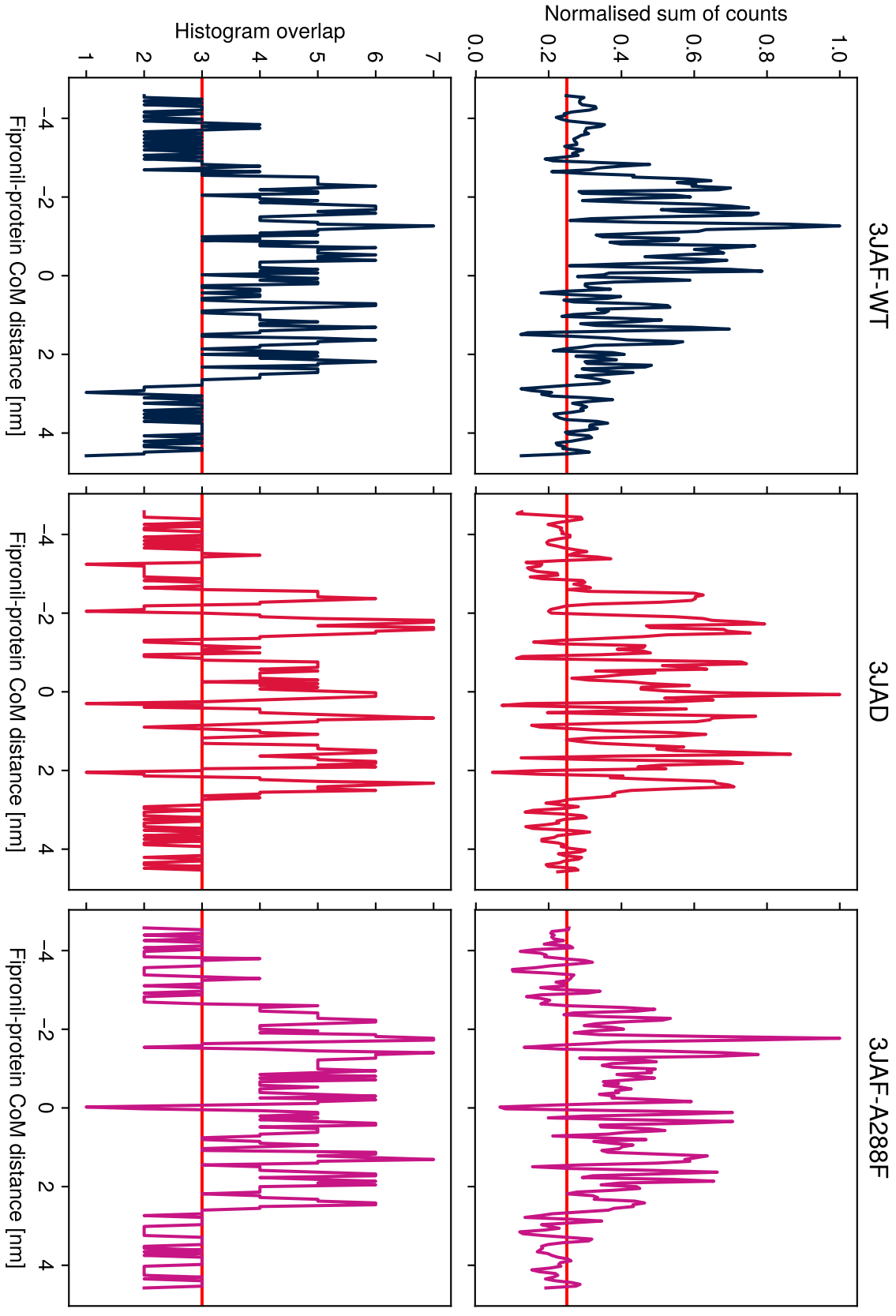

Figure 4.20: Analysis of histogram counts and overlaps of the PMFs in Figs. 4.17–4.19. Analysis script adapted from Dr Owen Vickery's `umbrella_sampling` script[276].

Overall comparison and analysis

Islam and Lynch's results show that in the closed state, fipronil can still reach its inhibitory binding site in the channel pore, but that it does so more slowly than when the channel is in the open state. This is demonstrated by a 30 s closed-state application of fipronil being sufficient for channel blockage (Fig. 4.2 B) while a 5 s closed-state application is insufficient (Fig. 4.2 H). The A288F mutation was shown to prevent fipronil from reaching the pore after a 5 s closed-state application (Fig. 4.2 I–K) — unlike in the wild type, no inhibition sets in after the channel opens. However, Islam and Lynch unfortunately did not provide data for a 30 s closed-state application for the A288F system. It is thus unknown if the mutation mimics the closed-state behaviour and merely slows down fenestration permeation or if it makes it impossible outright. Preliminary and unconverged as they are, the PMFs calculated to date allow comparisons of the three systems tested in this study, both with one another and with Islam and Lynch's experimental data.

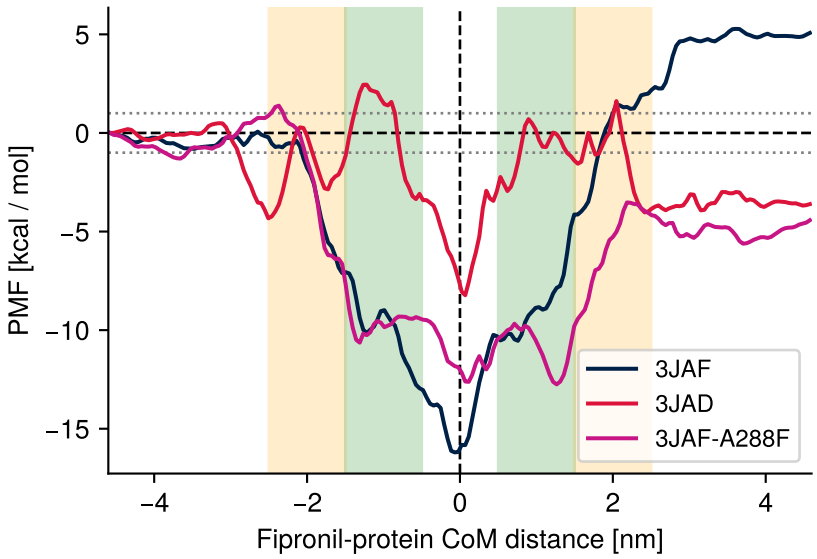

Figure 4.21: Comparison of the three fenestration-going PMFs collected in this study. The vicinity of the M2-M2 constriction and the M3-M1 constriction are indicated in green and yellow, respectively. Error indications are omitted for visual clarity.

The three fenestration PMFs are shown in Fig. 4.21. The error estimates are omitted for visual clarity. Under the assumption that the process of fenestration

permeation is adequately described by the Arrhenius equation,

$$k = Ae^{-E_a/k_BT}, \tag{4.1}$$

where k is the rate constant, A is the pre-exponential factor, and E_a is the activation energy of the reaction, the reduction of the rate constant due to an increased energy barrier can be estimated as

$$\frac{k_2}{k_1} = e^{(E_{a,1}-E_{a,2})/k_BT}. \tag{4.2}$$

The application of Eq. 4.2 to the energy barrier values read from the PMF plots is shown in Table 4.2.

Table 4.2: Estimation of the reduction of the rate constant of fenestration permeation through the application of Eq. 4.2. Here, k_B is used in the form of 1.987 × 10^{-3} kcal mol^{-1} K^{-1} and a temperature T of 310 K is supplied.

comparison 2 vs 1	$E_{a,2}$ [kcal mol^{-1}]	$E_{a,1}$ [kcal mol^{-1}]	k_2/k_1
closed vs open	$5.2 - 5.6$	$0 - 1.2$	$(1.13 - 15.1) \times 10^{-4}$
A288F vs open	$2.7 - 3.1$	$0 - 1.2$	$(6.52 - 87.6) \times 10^{-3}$
closed vs A288F	$5.2 - 5.6$	$2.7 - 3.1$	$(9.02 - 33.1) \times 10^{-3}$

Assuming Arrhenius-style kinetics apply and the pre-exponential factor A is the same for all systems, the PMFs estimate the rate of permeation through closed-state fenestrations is between 0.01 and 0.15 % the rate of permeation through open fenestrations. The rate of permeation through open A288F fenestrations is larger, at 0.6 to 8.7 % of the rate of wild type open fenestrations. As none of the three PMFs have converged yet, all values in Table 4.2 are estimates and are liable to change.

The fenestration's location in the membrane bilayer is very close to the minimum of the PMF for membrane permeation. Fipronil will naturally tend to accumulate at the height of the fenestration. The energy barrier of leaving the membrane for the intracellular aqueous compartment to then enter the pore from there is 10 kcal mol^{-1} (Fig. 4.14). Compared with the energy barrier estimation of 0 to 1.2 kcal mol^{-1} for permeation of the open fenestration, the latter is both

energetically favourable and an estimated 1.6×10^6 to 1.1×10^7 times as fast. These results demonstrate that fenestration-based pore access is preferable to membrane permeation.

4.4 Discussion and future directions

The work outlined in this chapter is the first demonstration of state-dependent, drug-permeable fenestrations in the glycine receptor. The state-dependence of these fenestrations was shown through an analysis of 55 different structures of the channel. Binary behaviour of the fenestrations was described, in which they disappear in the closed state and open as a result of structural rearrangements during activation (Fig. 4.8). State-dependent drug binding has important implications for drug safety and specificity[277–280], making the state-dependence of these fenestrations an interesting and useful discovery.

Steered MD simulations revealed that the open-state fenestration is able to accommodate fipronil without requiring structural deformation (Figs. 4.9 and 4.10. This is an important result as this knowledge was a prerequisite for further study with umbrella sampling — had the channel undergone drastic conformational changes or even fallen apart at this stage, further investigation would have been moot.

Umbrella sampling simulations were then attempted, but this system proved difficult to study adequately in free energy simulations. Initial umbrella sampling results with poor quality (Fig. 4.12) provided the motivation to develop the enhanced umbrella sampling workflow that was introduced and benchmarked in chapter 3. In the end, several techniques were combined for the umbrella sampling simulations run in this study. The size of the system was minimised by using a hexagonal prism box shape. Centre of mass restraints on the individual protein subunits were applied with GROMACS's pull code to prevent rotations in the membrane that would disturb the reaction coordinate. On the ligand level, cylindrical flat-bottom restraints were applied with PLUMED to prevent fipronil from distancing itself from

the vector that defines the reaction coordinate. And finally, Alchembed-based window generation and STeUS were combined for enhanced sampling. This combination makes it possible to study this system with tolerable computational effort and the results are much improved compared to the initial attempts. Even so, the PMFs obtained to date only allow qualitative comparisons because they have not converged yet. Further simulations as described in Section 4.3.6 are required and will be carried out.

Nonetheless, the qualitative comparison is illuminating in its own right. The preliminary energy barriers calculated thus far suggest that fenestration-based pore access of fipronil is energetically favourable and orders of magnitude faster than membrane permeation. The comparison also provides early evidence in agreement with Islam and Lynch's results. Increased energy barriers for closed-state fenestrations and the A288F mutation match reduced fipronil inhibition after brief closed-state applications (Fig. 4.2 H, I).

It must be noted that the structure of the A288F mutant in this study is dubious. It was obtained by applying the point mutation to the desensitised-state structure (PDB 3JAF) used in this study *in silico*. However, there is no evidence to indicate that the mutant receptor's structure would match this conformation. Results on the A288F system must be read with this caveat in mind.

4.4.1 Future work

This project will be pursued further to prepare it for publication. In a first instance, the simulations will be extended as discussed in Section 4.3.6 to obtain converged PMF profiles.

The PMF of a minor component is still missing for the membrane-permeation pathway (red arrow in Fig. 4.4): the process of fipronil entering the channel pore from the intracellular aqueous surroundings should be simulated. As the energy barrier for membrane permeation is much larger than that for fenestration permeation, this is not strictly necessary to determine which pathway is more favourable. However, it would be useful for completeness as it would allow a

more like-for-like comparison of PMFs. It would also give access to cycle closure as another convergence criterium — the PMF differences between pore and membrane should be the same in both pathways.

An alternative or supplemental approach to the umbrella sampling employed here could be metadynamics. MetaD has successfully been employed in the study of fenestrations in the voltage-gated sodium channel $Na_v1.5$[116].

Molecular dynamics simulations of the closed-state channel might provide additional insights on the behaviour of the fenestrations in the closed state. Combined with CAVER-based analysis of trajectories, fenestration fluctuations can be detected and quantified. It would be interesting to see if compounds that interact with the channel here such as fipronil, ivermectin, or propofol exploit natural fluctuations of the fenestration radius to bind or if they induce the formation of their binding site themselves.

In Islam and Lynch's paper, data for a second pesticide — lindane (Fig. 3.5) — is reported. It is found to be an inhibitor of the human glycine receptor as well. It has not been considered here thus far because the authors found it was unaffected by the A288F mutation. However, in future investigations, it would be interesting to probe why that is. It could be that lindane permeates the lipid bilayer more readily, or that it is small enough to be fenestration-permeable even in the closed state and with the A288F mutation.

In the future, it might be interesting to study this fenestration in systems with more complex lipid bilayers. There are hints of the possibility of lipid regulation of the glycine receptor[88] via the fenestration, but this aspect has not been considered in this study at all thus far. A more realistic membrane composition would be a requirement in the pursuit of this research question.

The importance of tunnels that connect to active sites buried deeply within enzymes is increasingly being studied, recently leading to the formulation of the keyhole-lock-key model of enzymatic regulation[281,282]. This model attributes part of what determines selectivity between ligand and enzyme to the tunnel through which the ligand must reach its active site. In this study, only homomeric *Danio*

rerio α1 glycine receptors were simulated, meaning that all fenestrations probed here were formed by the same protein residues. Beyond α1, there are three more α subunits α2-α4 (though α4 is a pseudogene in humans[283]) and one β subunit. Glycine receptors can form α-homomeric channels or 4α1β-heteromers[284]. The keyhole-lock-key model could also be a suitable description of non-enzymatic proteins such as ion channels. Mutations along the "keyhole" component that affect inhibitor activity have been shown both here by Islam and Lynch and on the sodium leak channel[114]. Further study on the fenestrations on GlyRs with different subunit composition might open the door to the development of subtype- or tissue-specific inhibitors. Perhaps more promising in this regard than the glycine receptor is the $GABA_A$-receptor. With a total of 19 subunits to form heteropentamers from[70], it has a larger potential of significant fenestration differences between subtypes. Its γ2 subunit could be of particular interest as it is required for synaptic clustering of $GABA_A$-receptors[285].

4.5 Conclusion

The importance of fenestrations in voltage-gated sodium channels is well-understood. They are known to modulate drug and lipid binding as well as channel gating. Until now, they had not been reported in pentameric ligand-gated ion channels. Here, we identified the presence of a state-dependent and drug-permeable fenestration in the glycine receptor for the first time. This discovery will open the door for further investigations in the crucial Cys-loop receptor superfamily and might be particularly useful in the development of subtype-specific compounds.

5

Action of Strychnine and Ivermectin on the Glycine Receptor

Contents

5.1 Motivation

Molecular dynamics simulations are a useful tool in addressing biochemical questions, generating models of dynamics and mechanisms from structural information. MD simulations thus form a natural complement to structural biology both retro- and prospectively. Retrospectively, simulations can help interpret and contextualise experimental results, as they are able to provide insight into conformational changes, allostery, and ligand binding, among others. Prospectively, simulations can be used to predict, for example, the effect of mutations or ligand binding affinities.

In the following two chapters, I will discuss two collaborations I was involved in during my DPhil and will show how MD simulations acted as a complement to experimental data.

Disclaimer: Parts of the work described in this section are included in our manuscript "Structural Basis for Domain Coupling in Heteromeric Glycine Receptors Revealed by an Atypical Allosteric Agonist"[286] (accepted for publication at Science Advances). All investigation, writing and figures in this paper and chapter were produced by me with feedback from the co-authors, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

The structure, physiology, and pharmacology of the glycine receptor (GlyR) in general is described in detail in Section 1.2.

5.2 Introduction

In this study, the competing activities of the orthosteric inhibitor strychnine and the allosteric activator ivermectin on heteromeric GlyR are investigated. Strychnine (Fig. 5.1) is an alkaloid inhibitor of the glycine receptor with low nanomolar affinity[287] and competes with glycine for the orthosteric binding site[288]. Ivermectin (Fig. 5.2) is an anthelmintic drug whose main target is the invertebrate glutamate-gated chloride channel. It potentiates glutamate-induced currents

when applied in low-nanomolar concentrations and is able to activate the channel outright at concentrations above ≈140 nM[289]. Glutamate-gated chloride channels are not found in mammals[290] and ivermectin activity on the structurally related mammalian glycine receptor is much weaker. Triggering GlyR activation requires low-micromolar concentrations of ivermectin[291], making it a safe drug in mammals. Notably, ivermectin-induced GlyR currents are distinct from glycine-induced ones in that they are insensitive to strychnine[87]. In fact, it has previously been shown that ivermectin is able to protect mice from strychnine poisoning, highlighting the independent modes of action of these two compounds[292] and the "tug of war" they engage in over the channel behaviour.

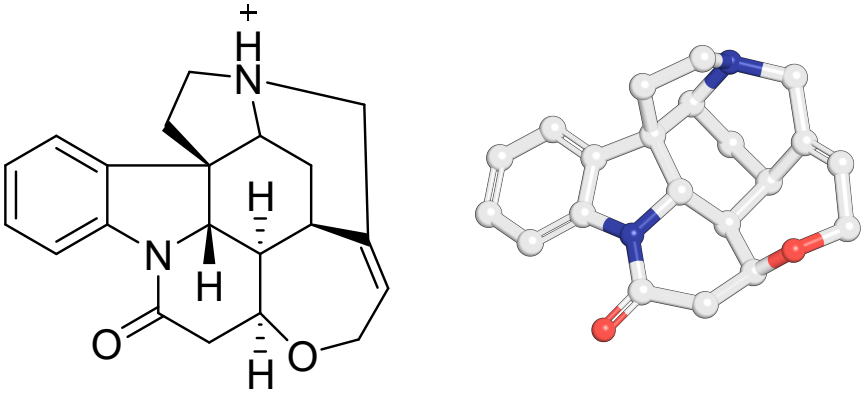

Figure 5.1: Strychnine

The orthosteric binding site is located at the subunit interface in the extracellular domain (ECD) (Fig. 5.3). It is formed by conserved Loops A–F, three of which are located on either subunit at the interface. Loop C is of particular importance as it covers the binding site like a lid. Ivermectin, on the other hand, binds at the transmembrane domain (TMD) between helices M3 and M2 of one subunit and M1 and M2 of the adjacent subunit (Fig. 5.3). The five M2 helices line the channel pore in the TMD. Ivermectin acts by pushing them apart, thus widening the pore sufficiently for ion permeation.

In this study, we set out to investigate the competing effects of strychnine and ivermectin on the glycine receptor. We hypothesized that structures solved in the presence of both compounds might reveal intermediate states that would

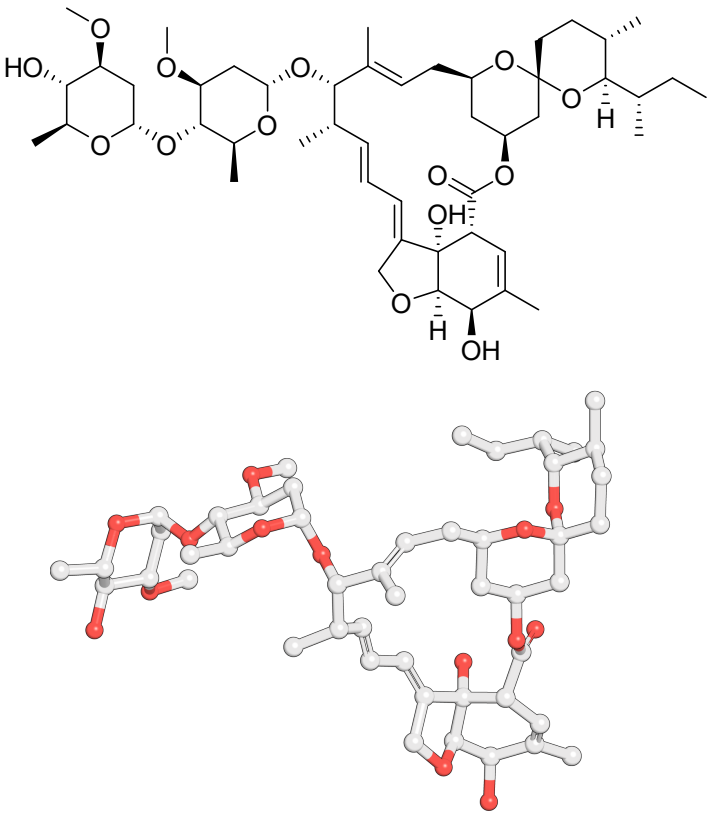

Figure 5.2: Ivermectin

help elucidate channel activation. Since ivermectin can activate the channel independently of glycine, and since ECD and TMD can undergo independent rearrangements, we hoped that this approach would provide insight into the question of whether ivermectin-dependent channel activation follows similar activation pathways as glycine-dependent activation. In this enquiry, we also sought to find whether our structural data, combined with MD simulation, could provide an explanation for the observed protective effect of ivermectin against strychnine.

5.3 Methods

5.3.1 Structure preparation and simulation parameters

Our collaborators on this paper provided cryo-EM structures of heteropentameric zebrafish (*Danio rerio*) glycine receptors consisting of four α1 subunits and one β subunit (GlyRα1β) with a constant, saturating concentration of strychnine (200 µM)

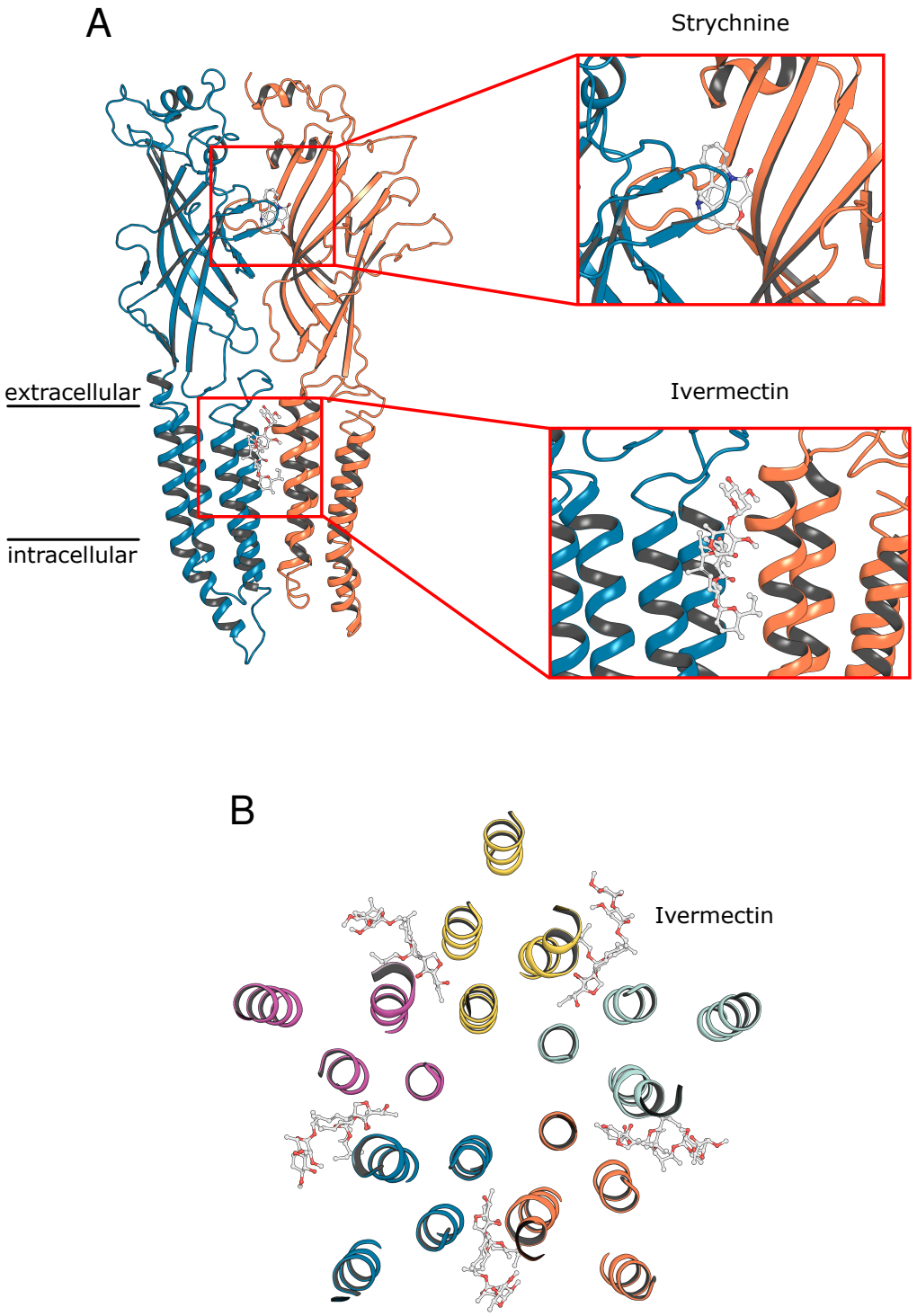

Figure 5.3: A: Binding sites of strychnine and ivermectin at the glycine receptor subunit interface. B: Top view onto the transmembrane domain with its five ivermectin binding sites.

and increasing ivermectin concentrations (0.5, 2, and 20 µM; PDB ID 9PKX, 9PKY, and 9PKZ, respectively). Aside from the missing intracellular domain (ICD) and few unresolved residues at either terminus, no gaps in the structure are present. As in previous studies[56], each segment's missing ICD was replaced with an AGT tripeptide linker using MODELLER[165] version 9.25, and acetyl- and N-methyl caps were added to the N- and C-termini respectively. The three structures are henceforth referred to as GlyRα1β-0.5Ivm200Stry, GlyRα1β-2Ivm200Stry, and GlyRα1β-20Ivm200Stry. Electron density for strychnine was observed in all five binding sites in all three structures, while ivermectin density was only seen in its five binding sites in GlyRα1β-2Ivm200Stry and GlyRα1β-20Ivm200Stry. The protein-ligand complexes were embedded in a POPE bilayer using the inflateGRO method[271], solvated with TIP3P water[223], and neutralised with 150 mM NaCl. Protein and lipids were described with the CHARMM36m force field[293,294], while ligands were parametrised with CGenFF[295]. Bonds to hydrogen were constrained with the LINCS algorithm[147] or, in the case of water, with SETTLE[148]. A timestep of 2 fs was used for integration of the equations of motion. Long-ranged electrostatic interactions were calculated with a smooth particle mesh Ewald approach[149], while short-ranged interactions were calculated under a Verlet cut-off scheme. The V-rescale thermostat[151] and semi-isotropic C-rescale barostat[152] with coupling constants of 1 ps and 5 ps, respectively, were used to maintain a temperature of 310 K and a pressure of 1 bar throughout the simulations.

5.3.2 Equilibration procedure and production simulations

After steepest-descent energy minimisation, a multi-step equilibration procedure was employed, during which position restraints were gradually released (Table 5.1).

Initial production simulations were started from equilibration step 5, but the system was found to be insufficiently equilibrated. Steps 6–8 were introduced to combat these issues. For these later simulations, the endpoint of equilibration step 6 is shared between the three independent production runs. After step 6, the

Table 5.1: Overview of the parameters used in each equilibration step of the three GlyRα1β systems.

Step	n_{steps}	timestep [fs]	Restraint force constant [$\mathrm{kJ\ mol^{-1}\ nm^{-2}}$]			
			backbone	sidechain	lipid	ligands
1	125000	1	4000	2000	1000	1000
2	125000	1	2000	1000	400	1000
3	125000	1	1000	500	400	1000
4	250000	2	500	200	200	1000
5	25000000	2	500	0	0	1000
6	25000000	2	0	0	0	1000
7	25000000	2	0	0	0	500
8	25000000	2	0	0	0	200

velocities for each replicate were randomly drawn from the Maxwell-Boltzmann distribution and equilibrations 7 and 8 as well as the 500 ns production simulations were run in sequence.

5.3.3 Calculation of RMSD and RMSF quantities

In-house MDAnalysis[166,167] scripts were used to calculate root mean square deviation (RMSD) and root mean square fluctuation (RMSF) quantities. For ligand RMSD quantities, all ligand atoms were used, while protein RMSD values were determined for backbone atoms of either the entire structure or of Loop C (Fig. 5.4). RMSF values of Cα atoms of key residues in the strychnine binding pocket were calculated with respect to the average structure of each production run.

5.3.4 Ligand pose cluster analysis

To supplement ligand RMSD data, clustering of strychnine binding poses was carried out using the hierarchical clustering implementation in SciPy[296]. Strychnine is a relatively rigid molecule, so differences in binding poses will in large part be determined by the relative orientations of strychnine in the binding pocket rather than by torsions within the molecule. For computational efficiency, ligand poses were thus described by the centres of mass of five ring groups (Fig. 5.5).

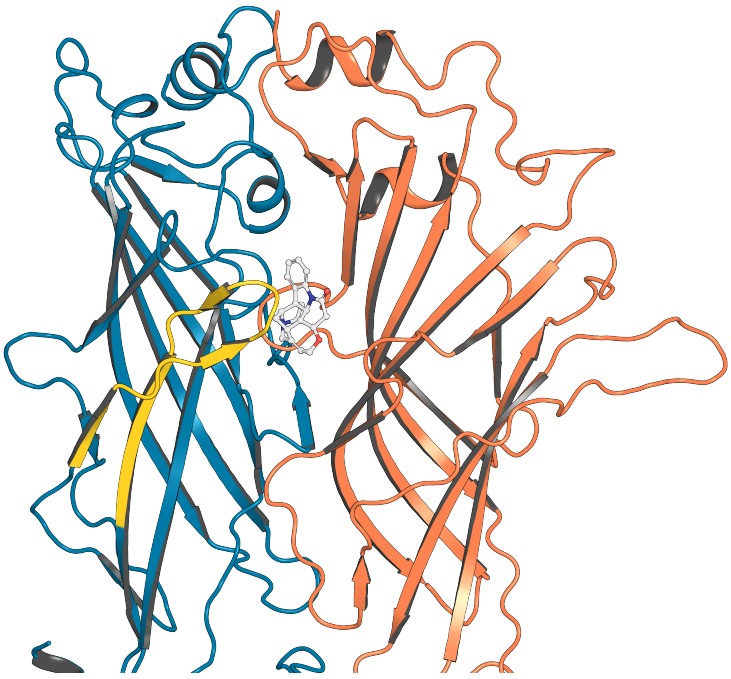

Figure 5.4: Loop C was defined as residues 218–238 (238–258 for GlyRβ) for the purpose of Loop C RMSD calculation. Loop C, shown here in gold, covers the orthosteric binding site. As such, its movements are strong determinants of ligand stability.

This reduction of detail allowed much faster calculation while still accurately capturing molecular orientation.

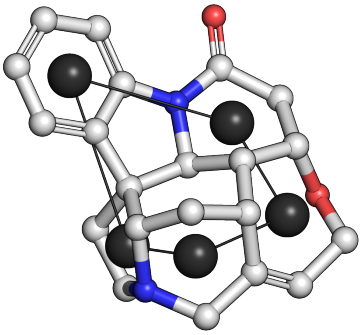

Figure 5.5: Definition of centres of mass (black) used to define strychnine orientations for the purpose of hierarchical clustering.

The distances of each centre of mass from its starting position were calculated for each strychnine molecule throughout the trajectory and this distance data was used for clustering with `scipy.cluster.hierarchy.linkage` based on a

Euclidean distance metric and the median method for assigning new centroids of merged clusters. Poses within 2 Å of each other were considered to be part of the same cluster.

5.3.5 Protein-ligand interaction analysis

Production simulations were run in triplicate on three different GlyRα1β structures, each of which had five strychnine molecules modelled into its binding sites, meaning that interaction data for a total of 45 strychnine molecules was available. To find key residues involved in stable binding of strychnine, protein-ligand interaction fingerprints were calculated with ProLIF[297] for all strychnine molecules. Strychnine molecules were then classified as either stably or unstably bound based on ligand RMSD. Contact persistence — the fraction of simulation frames in which ProLIF detects an interaction with a particular residue — was chosen as the target metric and calculated for all residues identified in the interaction fingerprints. The distribution of these contact persistence values from stably and unstably bound ligands was then compared to identify crucial residues for stable ligand binding.

5.3.6 Measuring binding pocket volumes

The volume of strychnine binding pockets at the subunit interface was determined with the PyVOL[170] plug-in for PyMOL[298]. Volumes in the original cryo-EM structures and in the final snapshot of the production simulation runs were measured. Parameters used are listed in Table 5.2. After all pockets at a given subunit interface were identified by PyVOL, the strychnine binding pocket was visually identified and its volume read from the console output.

Table 5.2: Parameters used for binding pocket volume calculations with PyVOL. *: appropriate chain IDs for each interface specified in turn.

PyVOL Parameter	Chosen value
Protein Pymol Selection	chain A* or chain B*
Minimum Radius	1.9 Å
Maximum Radius	3.4 Å
Pocket Selection	all
Minimum Volume	200 Å^3

5.4 Results

5.4.1 An extended equilibration procedure is necessary for ligand stability

The main aim of this study was to determine the effects of ivermectin on the binding of strychnine at the heteromeric glycine receptor (Fig. 5.6). Initial simulations followed on from an equilibration scheme (steps 1–5 in Table 5.1) that sees a gradual release of position restraints on different parts of the system. Such schemes ensure that the desired starting conformation of the protein is not distorted by the equilibration of the solvent and the lipid bilayer around it before production simulations begin. However, we discovered that this initial equilibration procedure was insufficient. Production simulations started from these equilibrations showed clearly apparent defects. The extracellular domains undergo unrealistic "breathing" and tumbling motions and multiple ligands unbind (Fig. 5.7).

Strychnine has low-nanomolar affinity for the glycine receptor[287]. As such, observing seven ligands unbind across all runs, mostly within the first 200 ns, suggested the presence of a systematic error. Furthermore, the large movement of subunits away from one another broadened the gap at the interface unrealistically, making any further analysis of strychnine binding impossible.

The solution to these problems was found in an extended equilibration scheme. We reasoned that the rapid onset of ECD deformations in production simulations suggested the systems had not yet properly equilibrated. The additional equilibration steps' goal was to allow the protein to relax around strychnine to prevent

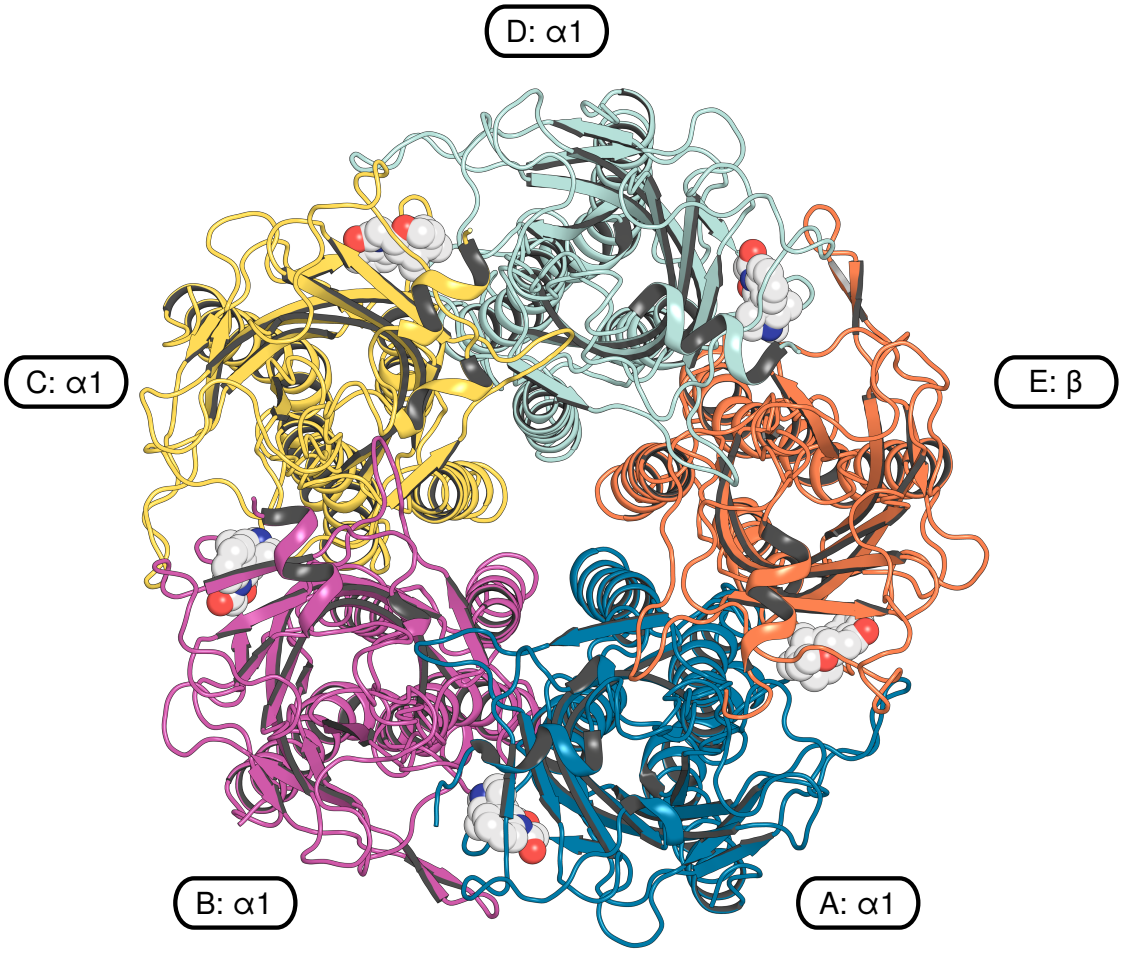

Figure 5.6: Heteromeric GlyR structures consisting of four α and one β subunits were investigated in this study. The mapping of chain index to colour shown here was used throughout. Strychnine is shown in sphere representation for clarity. In descriptions of interfaces, the principal subunit is listed first — for example, the AE interface has a principal α subunit and a complementary β subunit. This figure shows a top view onto the extracellular domain of the receptor. The five helices in the centre are the pore-lining M2 transmembrane helices.

it being driven apart at the subunit interfaces as soon as all restraints are lifted. We found that a stepwise, gradual release of position restraints on the ligand atoms while the protein was unrestrained (Steps 6–8 in Table 5.1) was able to achieve this goal. All results reported on from now were obtained from simulations started from this extended equilibration scheme.

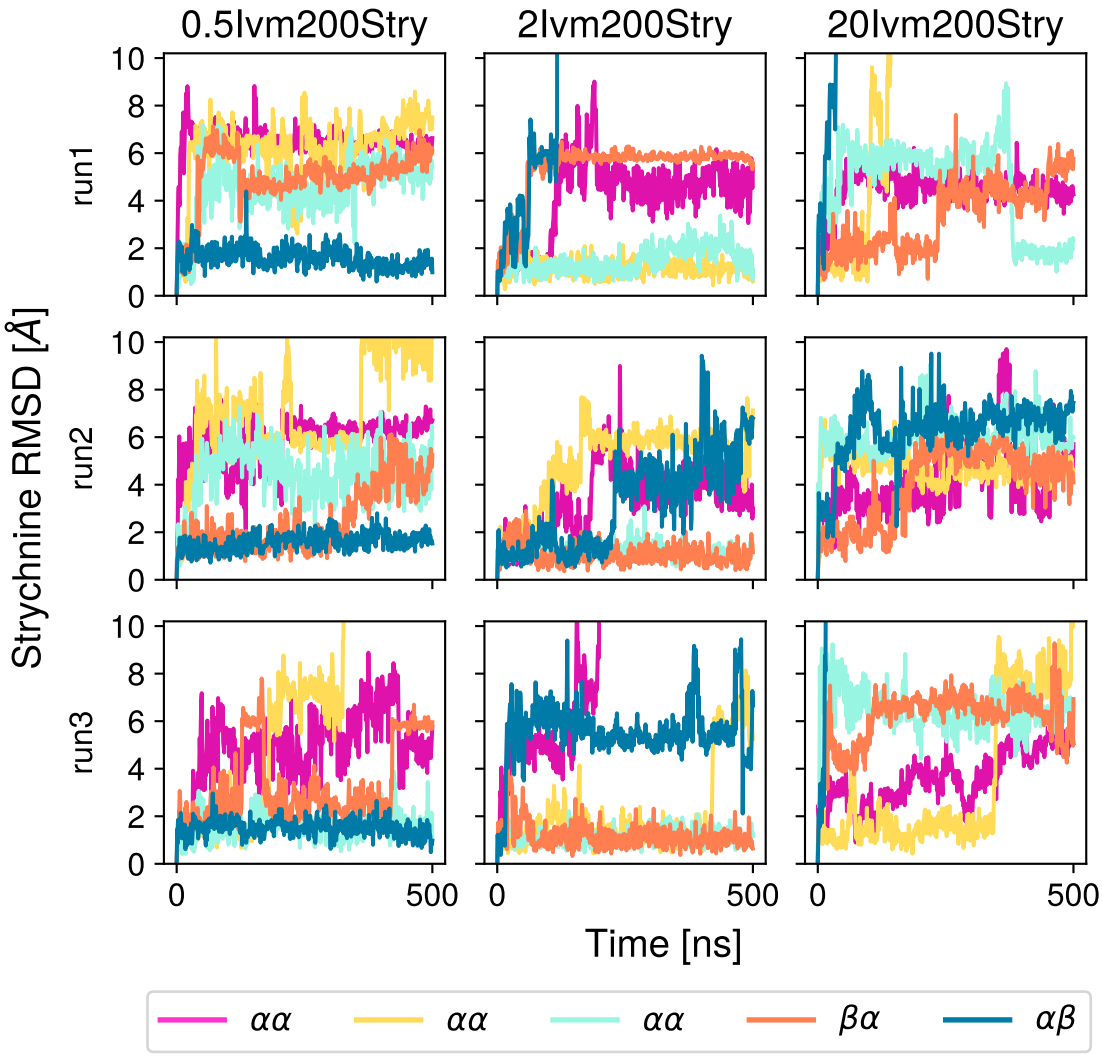

Figure 5.7: Strychnine RMSD values throughout the original production simulations. Strychnine showed a high degree of instability across all runs and interfaces. Several ligands left their respective binding sites and dissociated from the protein. The colour scheme matches that shown in Fig. 5.6, with the colour of the plot matching the colour of the interface's principal subunit.

5.4.2 GlyR-stability is reduced at high ivermectin concentrations

Having obtained stable protein systems, the effects of increasing ivermectin concentrations on the glycine receptor could be studied. In a first instance, the backbone RMSD of the whole protein was calculated (Fig. 5.8, SI Fig. D.1).

GlyRα1βb-0.5Ivm200Stry and GlyRα1βb-2Ivm200Stry show very similar backbone dynamics, even though the former system does not contain ivermectin in

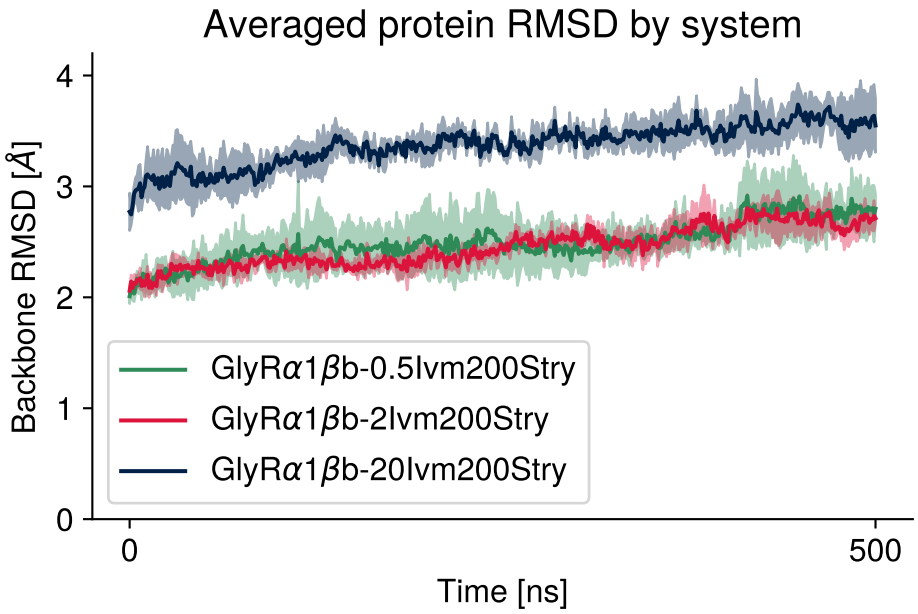

Figure 5.8: Protein backbone RMSDs of the three simulation systems. Shown are the means (solid lines) and standard deviations (shaded areas) of three 500 ns repeats. It is clearly apparent that the protein backbone of GlyRα1βb-20Ivm200Stry is more flexible than that of systems with lower ivermectin concentrations, even though the RMSDs have not plateaued yet. While a slight drift is evident, the difference between the systems is clear.

its binding site. GlyRα1βb-20Ivm200Stry displays distinctly higher flexibility in its backbone as shown by the higher RMSD across all replicates. To determine what influence this reduced protein stability has on strychnine binding, its binding site was investigated in greater detail.

Measurements of the strychnine binding pocket volume with PyVOL mirror this finding of increased instability of GlyRα1βb-20Ivm200Stry (Fig. 5.9). While binding site volumes in systems with low ivermectin concentrations remain stable, a statistically significant widening is evident for this system, with pocket size increasing by roughly a factor of two.

To identify which regions of the protein move to lead to this increase in binding site volume, we determined the Cα-RMSF of key binding residues (Fig. 5.10). This analysis revealed two regions in particular that are more flexible in GlyRα1βb-20Ivm200Stry than in systems with lower ivermectin concentrations: Loop C, as represented by Y226, T228, and F231 (GlyRβ: Y246, T249, Y252), and the β5-β6

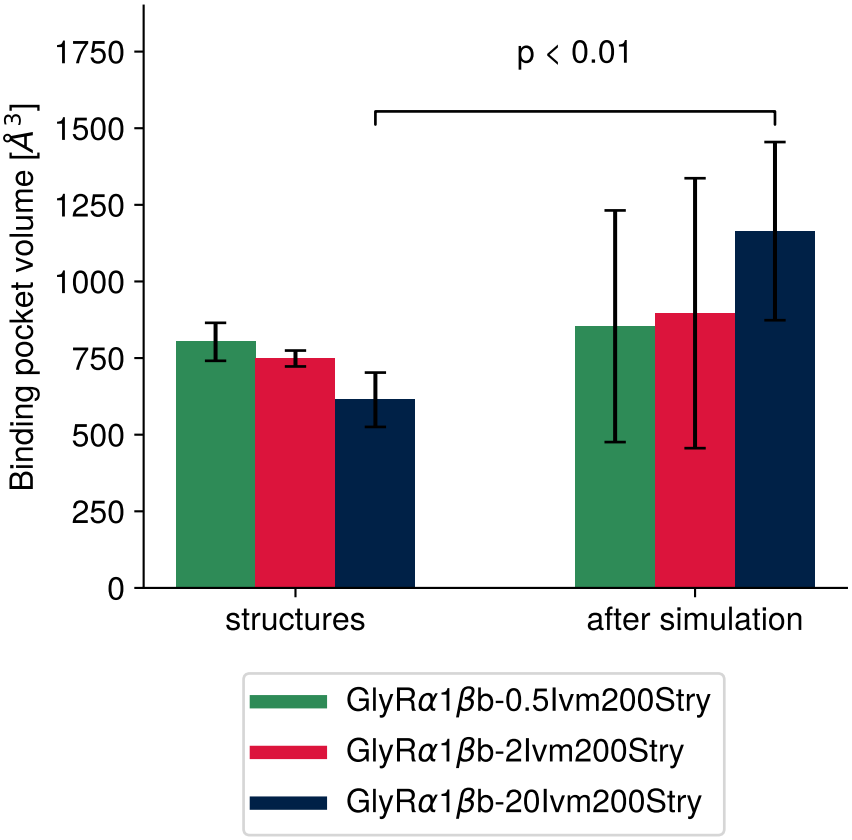

Figure 5.9: Average binding pocket volumes as determined by PyVOL. Volumes in the original cryo-EM structures are compared to those after 500 ns of simulation. While there is no significant difference in binding pocket volume for GlyRα1βb-0.5Ivm200Stry and GlyRα1βb-2Ivm200Stry, the volume increases in size significantly in GlyRα1βb-20Ivm200Stry as determined by Student's *t*-test.

hairpin containing L141, R143, and L151 (GlyRβ: L161, F163, L171).

Loop C was further investigated, and the per-interface RMSD of its constituent residues was calculated (Fig 5.11). This conserved ligand binding loop is crucial for ligand stability because it laterally covers the binding site (Fig. 5.4). This analysis revealed significantly increased RMSD values for the αα and βα and interfaces of GlyRα1βb-20Ivm200Stry.

In summary, an increased concentration of ivermectin is correlated with increased instability of the glycine receptor, particularly at the orthosteric binding site.

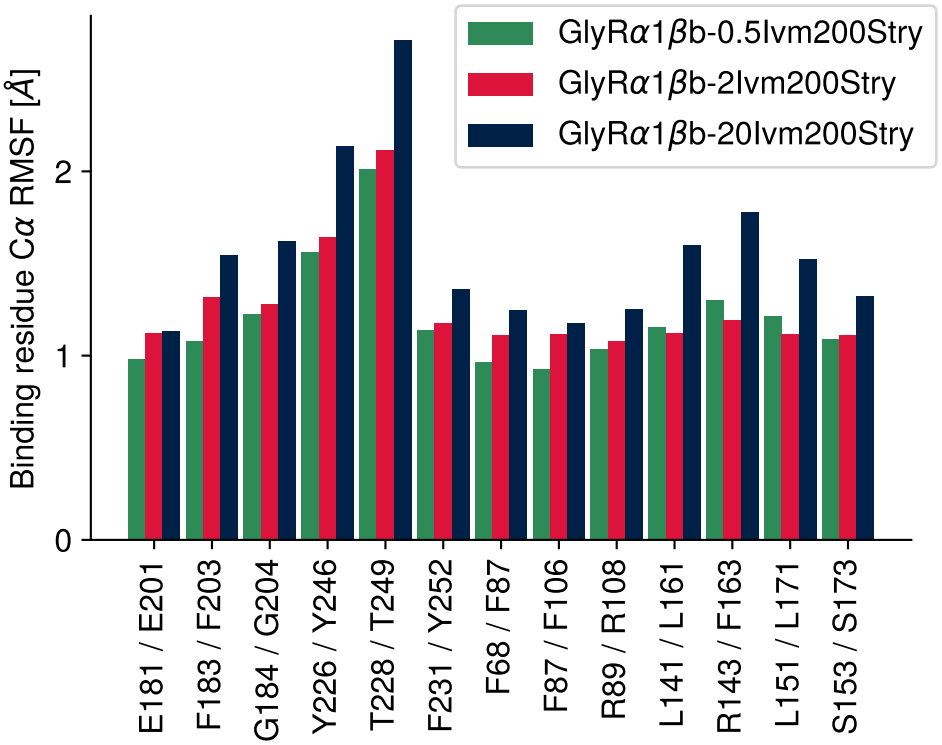

Figure 5.10: The Cα RMSF was determined for key strychnine-binding residues. In general, RMSF values for residues of GlyRα1βb-20Ivm200Stry are the highest. The difference is particularly noticeable in residues on Loop C (Y226, T228, and F231 or Y246, T249, and Y252 for β subunits), and on the β5-β6 hairpin (L141, R143, and L151 or L161, F163, and L171 for β subunits), suggesting that these are key regions determining ligand stability or lack thereof.

5.4.3 Strychnine is destabilised by ivermectin at high doses

Having established that the orthosteric binding site is destabilised by increased ivermectin concentrations, we turned our attention to the binding behaviour of strychnine. In the first instance, the calculation of ligand RMSD values first shown in Fig. 5.7 was repeated for these new, stable simulations. It is evident that the extended equilibration scheme led to improved ligand stability overall (Fig. 5.12). Strychnine shows bimodal behaviour, where it either remains stably bound (RMSD < 2 Å) or becomes unstable (RMSD ≈ 6 Å). Only one molecule unbinds outright, compared to seven ligands in the original runs. The number of stable and unstable strychnine molecules per condition is reported in Table 5.3.

The strong destabilisation of the protein in GlyRα1βb-20Ivm200Stry is mirrored

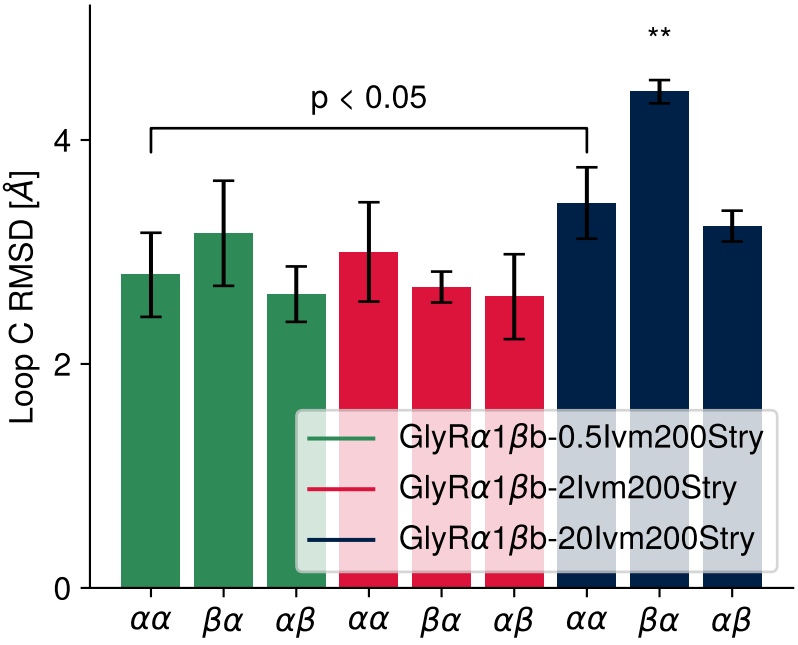

Figure 5.11: Per-interface analysis of Loop C RMSD reveals that, in GlyRα1βb-20Ivm200Stry, Loop C is more flexible than in systems with lower ivermectin concentrations. For each interface, the average Loop C RMSD across the trajectory is first calculated, and the average of these averages and their standard deviation is reported here. **: Statistically significantly different from every other interface. Significance was probed with Tukey's HSD test[275]. For $\alpha\alpha$, $n = 9$, otherwise $n = 3$.

Table 5.3: Numbers of stably and unstably bound strychnine molecules read from Fig. 5.12.

System	Stable strychnine molecules	Unstable strychnine molecules
0.5IVM200Stry	10/15	5/15
2IVM200Stry	9/15	6/15
20IVM200Stry	1/15	14/15

in less stably bound strychnine ligands compared to the other systems. However, in an instance of the Anna Karenina principle, while all stably-bound ligands are alike, each unstably-bound ligand is unstable in its own way. RMSD as a metric cannot by itself distinguish between various distinct binding pose as it increases. So while most of the unstable strychnine molecules appear to plateau at RMSD values of approximately 6 Å, an unknown number of distinct binding conformations could be obscured by that value. For a more detailed analysis of unstable binding poses, hierarchical clustering of the ligand poses based on an

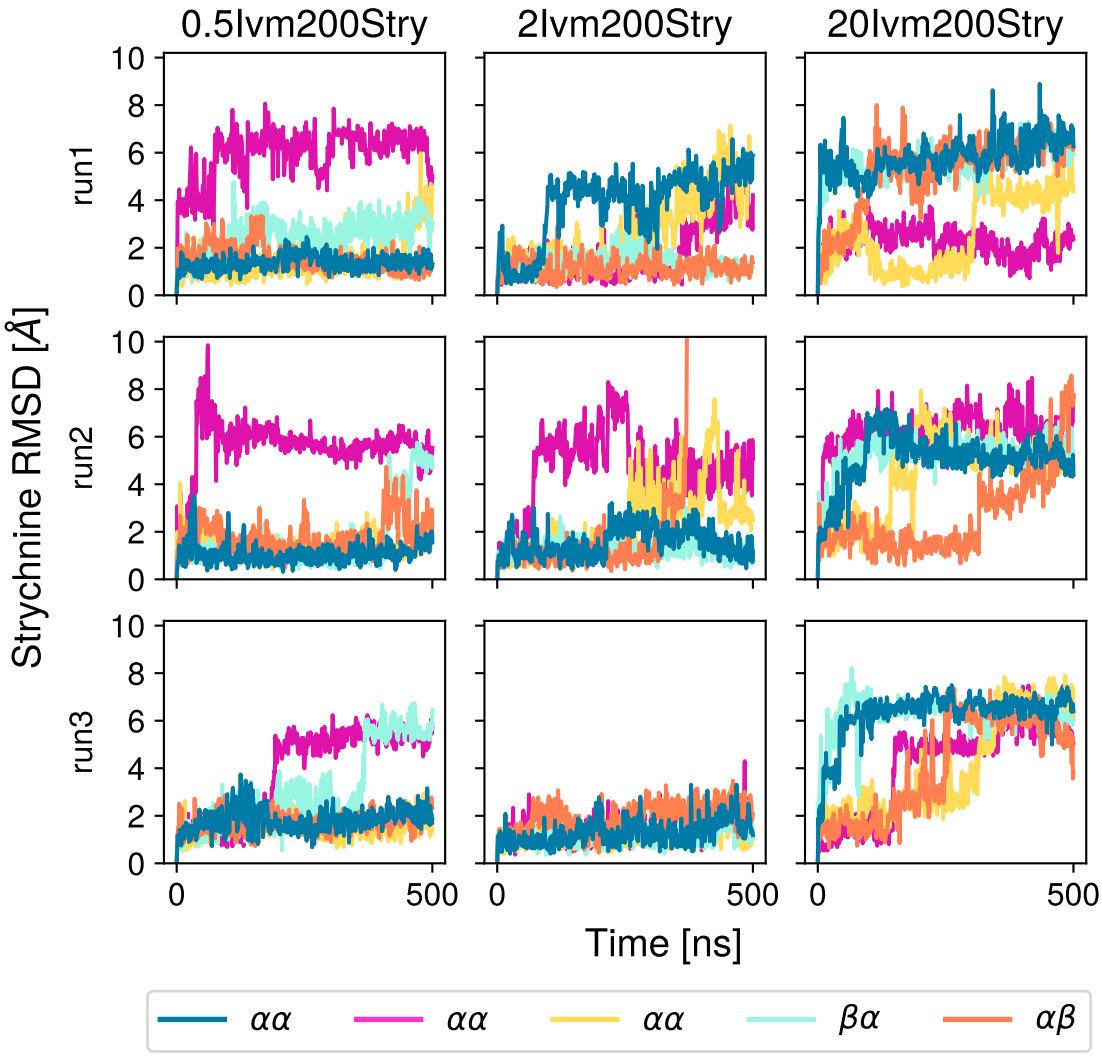

Figure 5.12: Strychnine RMSD values after an extended equilibration scheme. Comparison with Fig. 5.7 shows the additional equilibration had a strongly stabilising effect. Only one ligand leaves its binding site outright (that of the αβ-interface in run 2 of 2Ivm200Stry), while all other ligands appear to follow a bimodal behaviour — either remaining stably bound with RMSDs < 2 Å, or becoming unstable and reaching RMSDs ≈ 6 Å. The colour scheme matches that shown in Fig. 5.6, with the colour of the plot matching the colour of the interface's principal subunit.

abstraction of ligand coordinates (see Fig. 5.5) was run. This analysis enabled a better description of ligand instability. Structural snapshots were grouped into clusters by a distance criterion (Fig. 5.13 A, B). The number of clusters as well as their relative populations then revealed key unstable ligand poses (Fig. 5.13 C, D) that a plain RMSD value cannot resolve. Further analysis enabled us to extract representative poses of unstable strychnine from the largest clusters to demonstrate the orientational heterogeneity (Fig. 5.14). Finally, an analysis of the clustering data overall again pointed to 20Ivm200Stry being the least stable system (Fig. 5.13 E). This is shown by the ligands in this system having the largest total numbers of clusters, and by two metrics of cluster spread: the number of clusters needed to describe 80 % of the simulation data, and the number of clusters that contain at least 5 % of the poses. If the former measure is larger than the latter — as is the case for 20IVM200Stry — this is a sign of large heterogeneity in the data: it is so spread out that even relatively small clusters are needed to describe the system adequately.

Combined, the RMSD and pose clustering data collected on the ligand level match the results of protein-level analysis (Section 5.4.2). Having firmly established that GlyRα1βb-20Ivm200Stry has less stably-bound strychnine than the other tested systems, we now turn our attention to the reasons for this reduced stability.

5.4.4 Protein-strychnine interaction fingerprints reveal key binding interactions

To understand the differences between stable and unstable strychnine binding, interaction fingerprints were calculated with ProLIF[297]. These fingerprints hold time-resolved information on which ligand atoms interact with which protein residues through which interactions. This information can then easily be displayed as a barcode (Fig. 5.15). Comparison between barcodes showing stable and unstable binding reveals that some interactions appear to be crucial for stable ligand binding. Contact persistence — the fraction of simulation frames in which an

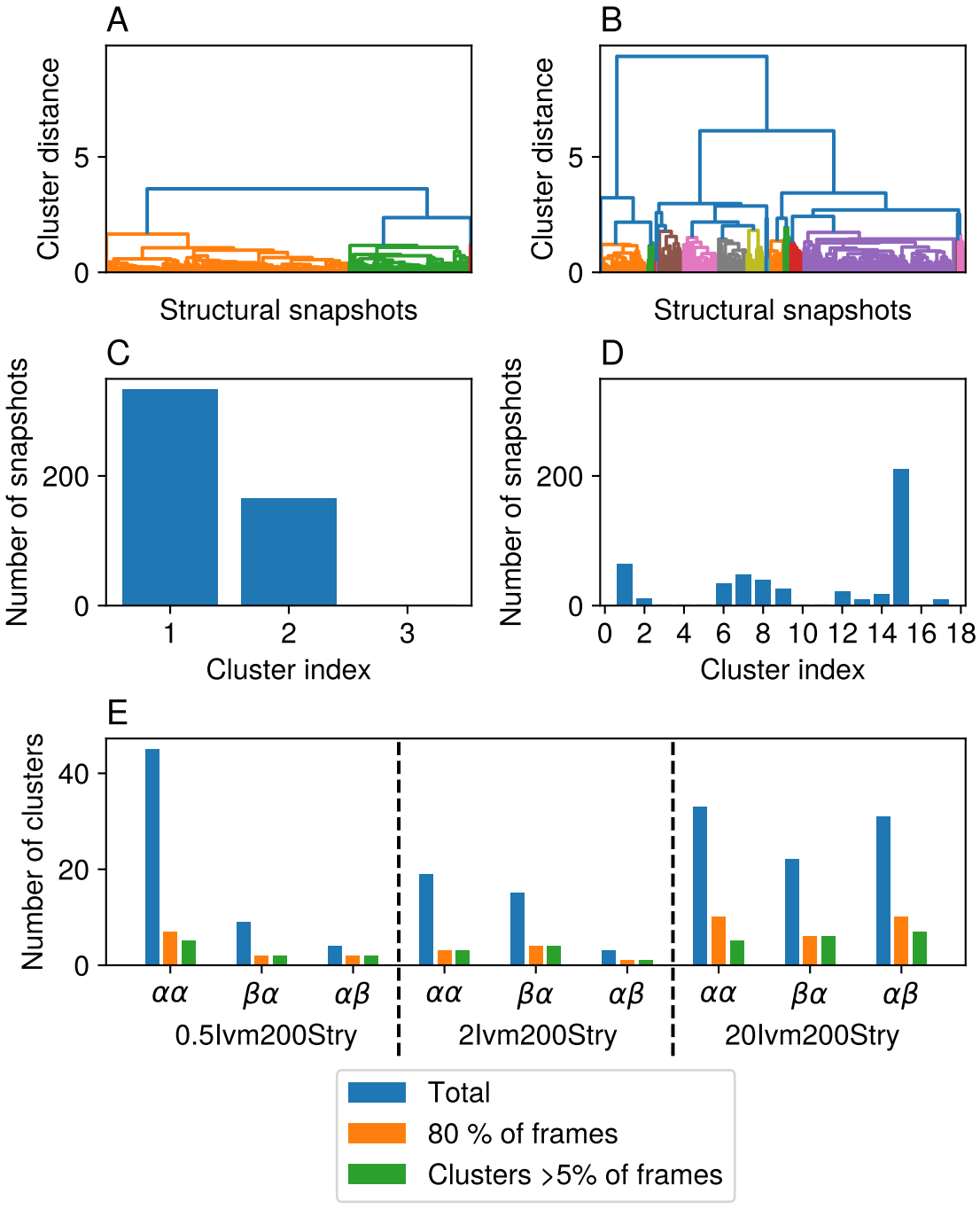

Figure 5.13: Overview of strychnine binding pose clustering. Dendrograms of a stably (A) and unstably (B) bound strychnine molecule. A much larger conformational space is sampled by unstably bound ligands. This is also seen in the cluster populations for the stable (C) and unstable (D) ligand. An overall analysis of the clustering results (E) reinforces the notion that GlyRα1βb-20Ivm200Stry has the least stably bound strychnine molecules. This can be seen in the higher total number of clusters and the larger number of significantly populated clusters. The large number of total clusters for $\alpha\alpha$ of 0.5Ivm200Stry is a result of the outlier behaviour of the BA interface in run 2 (cf. magenta line in the corresponding plot in Fig. 5.13).

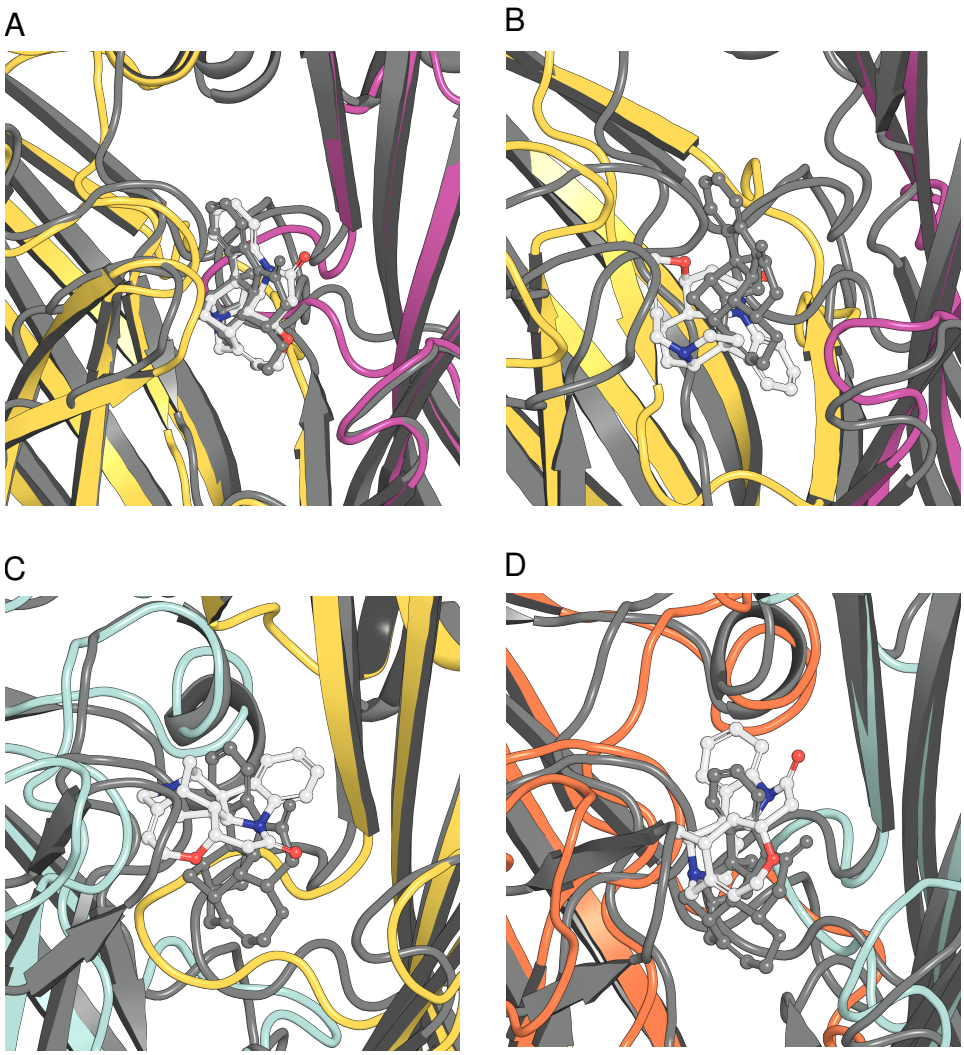

Figure 5.14: Comparing representative cluster structures (colour) to the original cryo-EM structures (grey). Examples of stable (A) and unstable (B, C, D) strychnine binding are shown. While the binding pose is well-preserved in stable binding, strychnine explores a larger configurational volume in unstable binding. The cases shown here can most easily be distinguished by the orientation of the aromatic ring.

interaction is detected — is a useful metric here. The barcodes in Fig. 5.15 show that contact persistence with F87, L141, Y226, and F231 is at or near 100 % in this stably-bound case, while these residues are clearly less-frequently contacted by the unstable strychnine. The barcodes also show when reorientations of the ligand occur through shifts in the interaction patterns. The stable example undergoes a minor rearrangement at roughly 370 ns, where the interaction with T228 is lost and partially replaced by F123. In the unstable example, multiple reorientations are recognisable, most clearly at 100 ns and 320 ns.

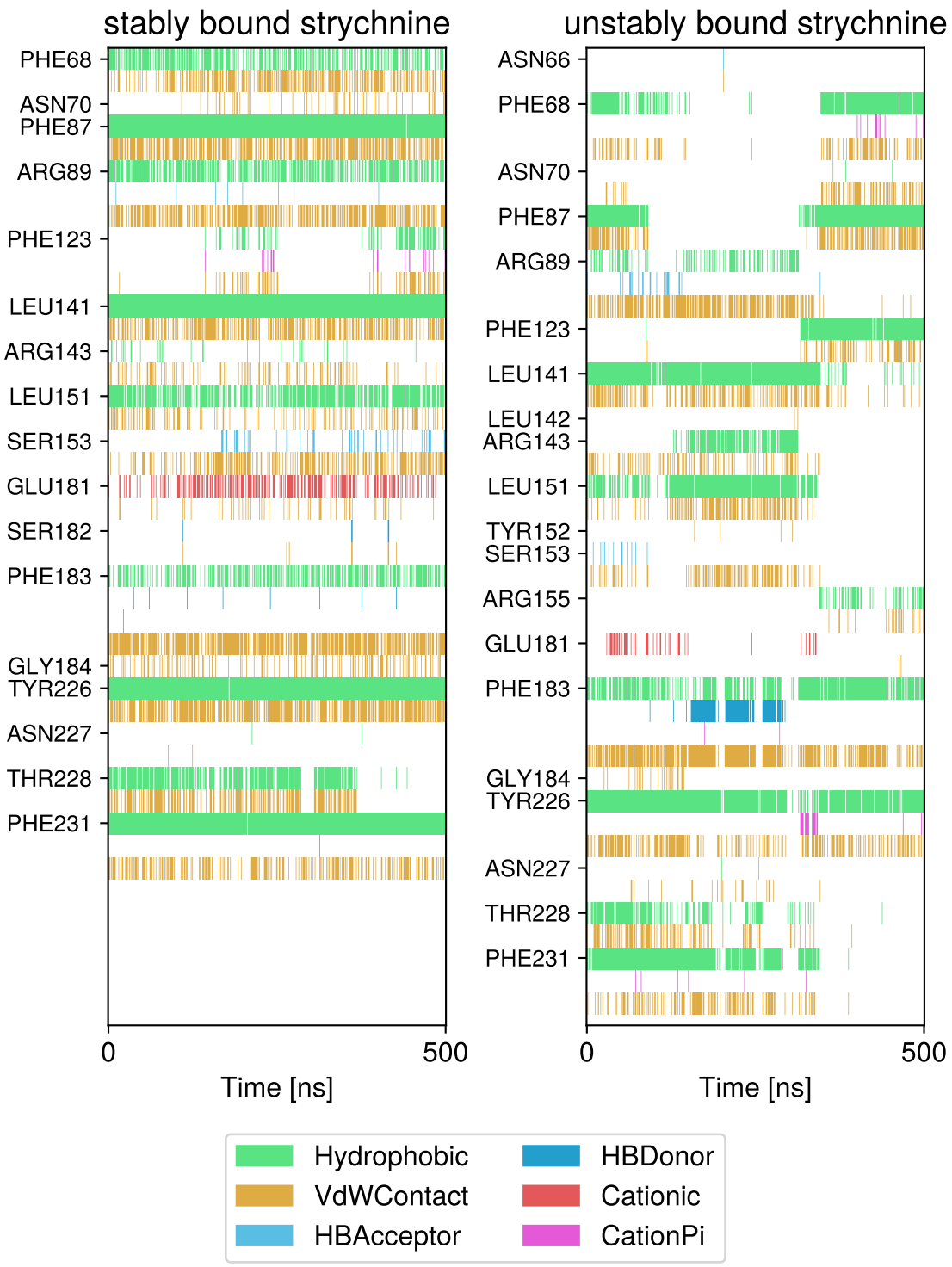

Figure 5.15: ProLIF-generated barcode plots depicting the protein-strychnine interaction fingerprints for a stably-bound and for an unstably-bound ligand. The unstably-bound molecule interacts with a larger number of residues, and many interactions are more intermittent.

The barcodes are useful visualisations for single interfaces but do not facilitate comparisons across the entire dataset. For this purpose, the contact persistence for each strychnine and all residues it interacts with was calculated. The distribution of this contact persistence data, broken down by ligand stability as determined in Table 5.3, was plotted. This plot clearly reveals the residues that are crucial for stable ligand binding (Fig. 5.16). These key residues are F87, F231, L141, Y226, and, to a lesser extent, T228 (F106, Y252, L161, Y246, and T249 in GlyRβ). When contact to these residues is lost, strychnine becomes unstable and starts contacting other residues in a larger volume within the binding site. These residues are, among others, F123, N70, and R155 (F143, N89, and R175 in GlyRβ). A third category consists of residues that show no clear difference in contact persistence, such as F183 or E181 (F203 or E201 in GlyRβ). Example snapshots of these residues interacting with stable and unstable strychnine are shown in Fig. 5.17, with residues coloured green, red, or yellow based on if they are preferentially contacted by stable or unstable strychnine, or do not show a clear difference, respectively.

Combining this contact persistence information with the data on Loop C RMSD (Fig. 5.11) and binding residue RMSF (Fig. 5.10) refines the conclusions. The Loop C residues Y226, T228 and F231 have all been identified as important binding residues, but the contact persistence analysis revealed that Y226 is the most important residue of the three, while T228 at the tip of Loop C displays larger flexibility. On the opposite face of the binding site, the β5-β6 hairpin was identified by RMSF analysis as another key region for ligand binding. This view was refined by the contact persistence analysis, which identified L141 as a crucial residue in stable binding while reporting less importance of R143 and L151.

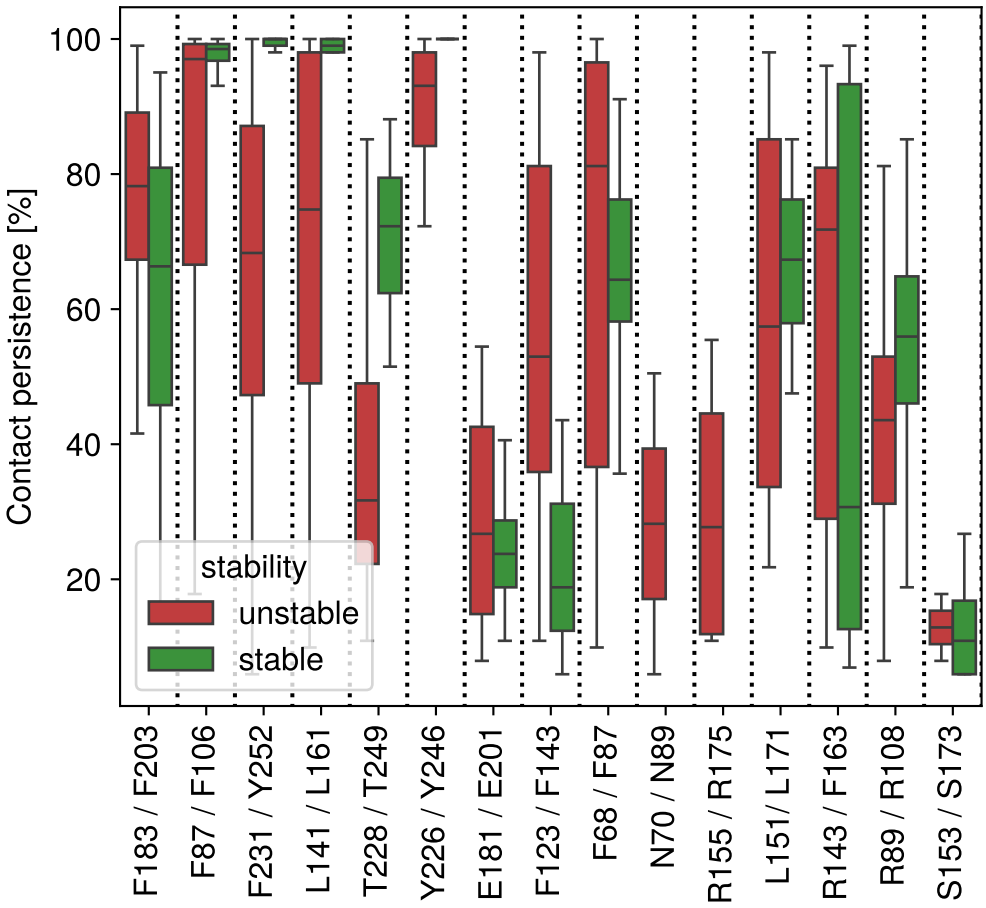

Figure 5.16: The distribution of contact persistence values for each residue calculated for every strychnine molecule, broken down by ligand stability. Key determinants of stable strychnine binding are F87, F231, L141, Y226, and, to a lesser extent, T228 (F106, Y252, L161, Y246, and T249 in GlyRβ). Shown here is a subset of all contact persistence data; some residues only briefly encountered by unstably bound strychnine are omitted for clarity.

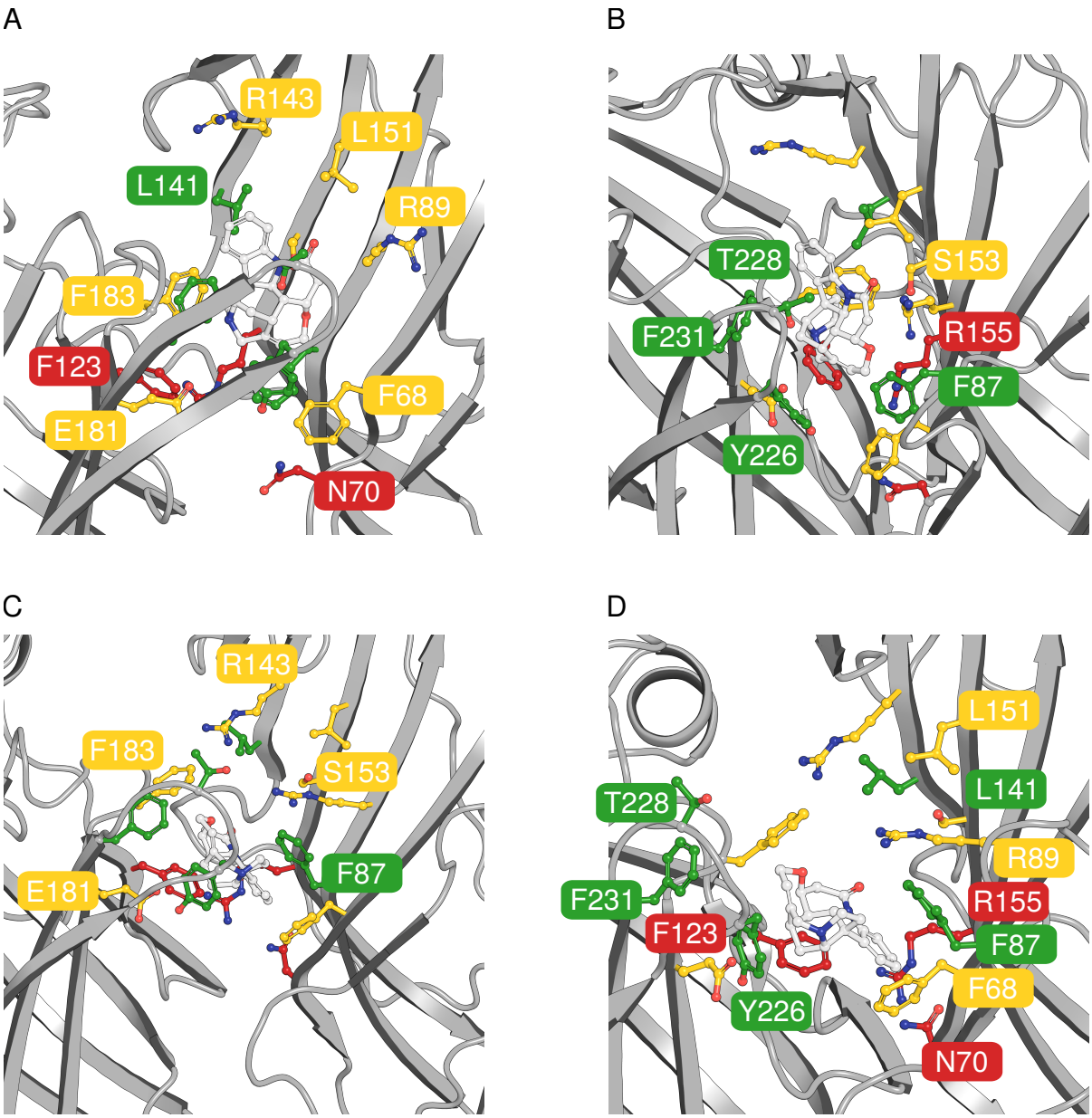

Figure 5.17: Structural overview of residues described in Fig. 5.16. Residues in green are more frequently contacted during stable strychnine binding, while red residues are contacted more by unstably-bound strychnine. Yellow residues showed no clear trend. Example structures for stably (A, B) and unstably-bound (C, D) strychnine are shown.

5.5 Discussion and future directions

It has previously been shown that ivermectin is able to cause glycine receptor activation even in the presence of strychnine[87]. In fact, ivermectin has been reported to act as an antidote to strychnine poisoning in mice[292]. This opposite action of the two ligands suggested that structures solved with both compounds present might reveal intermediate activation states. Within the larger scope of a project aimed at describing such states of the heteromeric glycine receptor and untangling the competing effects of ivermectin and strychnine on the channel, the question we wanted to address with molecular simulation is how ivermectin is able to disrupt strychnine inhibition. To this end, we collected a total of 4.5 µs of simulation data on GlyR-structures solved with varying ivermectin concentrations and a constant, saturating strychnine concentration. Our collaborators observed that the electron density for strychnine became more diffuse with increasing ivermectin concentration. Through analysis of our simulation data, we were able to identify Loop C and the β5-β6 hairpin as the regions that most strongly react to the presence of ivermectin. This provided the atomistic insight needed to understand this piece of the problem, as we identified the key residues in the orthosteric binding side that mediate ivermectin-dependent strychnine resistance through destabilisation of the latter compound.

This destabilisation of strychnine is an example of allostery. Allosteric effects are difficult to probe in simulation as they are exceedingly subtle, often requiring enhanced sampling techniques and/or long simulation timescales[299–301]. This difficulty was demonstrated here by the clear dependence on an appropriate equilibration scheme. Taken in isolation, our individual results reported here have little explanatory power. However, when combining the results from RMSD-, RMSF-, clustering-, binding pocket volume-, and contact persistence analyses, several subtle hints reinforce one another and form strong evidence. In further studies of the effects of counteracting ligands, dynamic network analysis might provide additional insight[302]. This method calculates a mutual information estimate

from a trajectory to build a network representation of the protein and find groups with correlated motions, and to find optimal pathways connecting different regions of the protein. Based on these pathways, it might be possible to describe in further detail exactly how ivermectin binding destabilises the regions in the orthosteric binding site we identified here. Further, nodes with high degrees of network centrality could be tested in mutagenesis studies, potentially allowing the decoupling of the strychnine and ivermectin binding sites.

Our simulation work on the glycine receptor is based on newly solved structures and combines evidence obtained from different analyses to pinpoint loci of allosteric action. It is thus a good example of how molecular dynamics and experiment can be each other's complement.

6

Probing intermediate configurations of a prototypical ion channel

Contents

Disclaimer: Parts of the work described in this section are part of "Asymmetric gating of a homopentameric ion channel GLIC revealed by cryo-EM"[303], published in PNAS. All investigation, writing and figures in this paper and chapter were produced by me with feedback from the co-authors, and will, with adaptations, be used in part here under the CC BY 4.0 license. Images and phrasing may therefore be similar or partially identical to my published work.

6.1 Introduction

The gating cycle of pentameric ligand-gated ion channels (pLGICs) is structurally and mechanistically complex. The three main functional states of these channels — resting, open, and desensitised— are formed of, as well as connected by, a plethora of structural and functional intermediates. Due to their transient nature, these intermediates are exceptionally difficult to resolve in structural experiments, but without further structural detail the mechanism of activation will remain elusive.

The *Gloeobacter violaceus* ligand-gated ion channel (GLIC) is a bacterial pentameric ion channel. It is a close structural homologue of vertebrate pentameric ligand-gated ion channels, with both the extracellular and transmembrane domain structures being highly conserved. Unlike its ligand-gated vertebrate counterparts, GLIC activation is pH-controlled. At low pH, key acidic residues in the ECD are hypothesised to be protonated, which triggers ECD and then TMD rearrangements that ultimately lead to channel opening. Because bacterial proteins are generally easier to express and study *in vitro*, because pH is relatively easy to control experimentally, and because of the high degree of structural similarity, GLIC has frequently been used as model system to study pLGICs[304–308].

In this channel family, activation in the extracellular domain is translated into protein motions that lead to a widening of the pore in the transmembrane domain, particularly in the upper, extracellular half of the M2 helix and the M2-M3 linker. In the inactivated state, this section leans inward towards the pore (*i*), and during channel activation transitions to an outwards-leaning conformation (*o*) (Fig. 6.1). Recent publications suggested that pLGIC activation can occur independently in each of the five subunits[309–312].

In this study, we report the structure of such intermediates of the *Gloeobacter violaceus* ligand-gated ion channel (GLIC). Our contribution to this paper was the functional annotation of GLIC structures of intermediate degrees of activation. Our collaborators developed a data analysis pipeline for cryo-EM data to extract asymmetric structures from single-particle data and arrived at an ensemble of

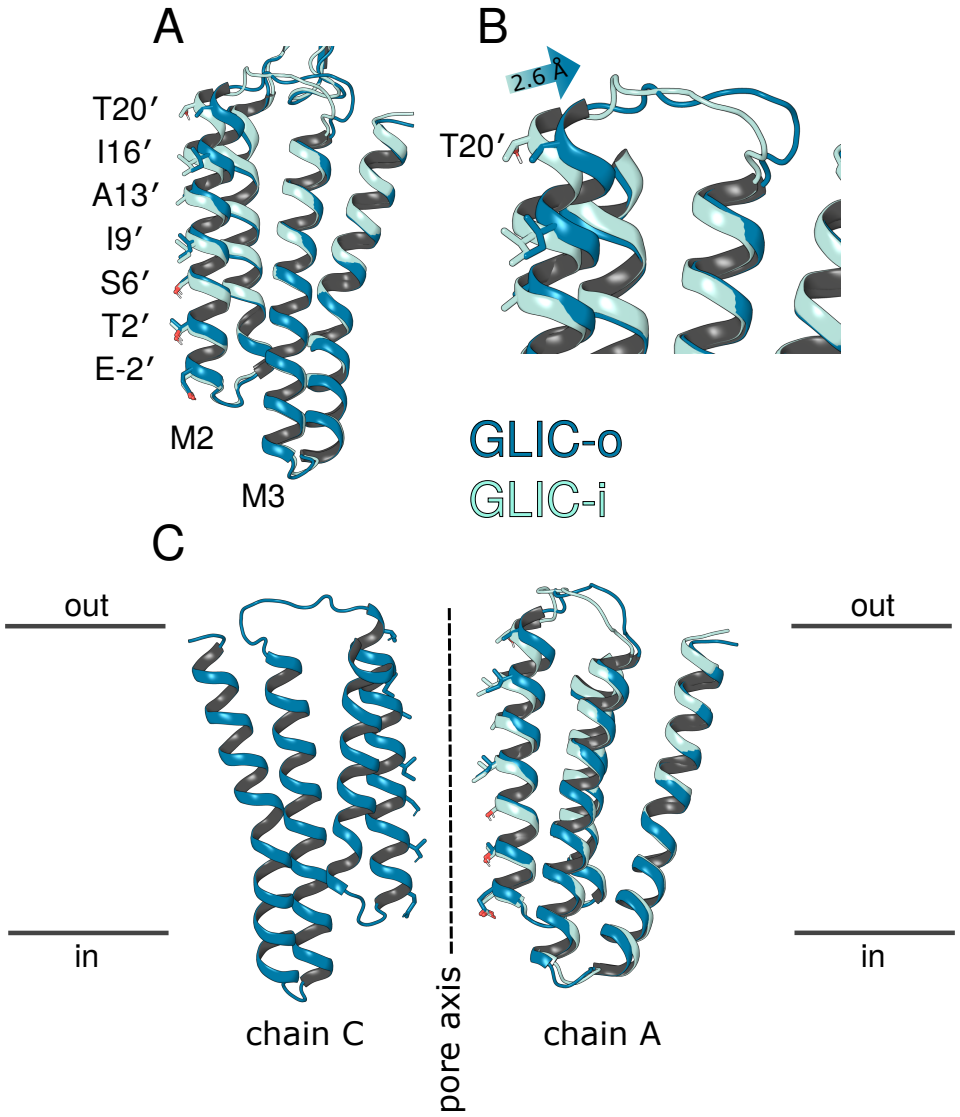

Figure 6.1: Comparison of inwards- (*i*) and outwards-leaning (*o*) M2-M3 linkers. A: Overlay of a subunit in *i* conformation onto an *o*-subunit. The pore lining residues are shown as sticks. B: Close-up view of the M2-M3 linker. The C-α of T20′ moves by 2.6 Å. C: An opposing subunit in *o*-conformation is shown to more clearly demonstrate the location of the channel pore.

structures of different activation states. These states are distinguished by the conformation of the M2-M3 linker and range from the most inactivated *iiiii* to the most activated *ooooo*, each letter referring to the conformation of one of the five subunits of GLIC (Fig. 6.1). This description distingishes a total of eight states — *iiiii*, *iiiio*, *iiioo*, *iioio*, *ioioo*, *iiooo*, *ioooo*, and *ooooo* — all of which were observed by our collaborators. *ioooo* is of particular interest in this study — it is only one subunit-activation away from a fully open channel, and its channel radius suggests it might already be permeable to ions. We set out to probe this structure as well as the next-closest structures *iiooo*, *ioioo*, and *ooooo* with MD simulations to assess their conductance.

6.2 Methods

6.2.1 Structure preparation and simulation parameters

To capture intermediate states with asymmetric activation levels, our collaborators solved cryo-EM structures of GLIC at pH 4.0. This pH value is above the pH_{50} of 2.9 and thus represents a sub-activating condition. Structures of four intermediate states with different numbers of inwards- (i) or outwards-leaning (o) M2-M3 linkers of the GLIC channel were chosen for further study in MD (Fig. 6.2).

These structures — *iiooo*, *ioioo*, *ioooo*, and *ooooo*, PDB IDs 9LAK, 9LBC, 9LB9, and 9LBA, respectively (Fig. 6.2) — were prepared for simulation with CHARMM-GUI[225–229]. Two different protonation states were investigated: in initial simulations, a previously used protonation scheme[313] was applied and E26, E35, E67, E75, E82, D86, D88, E177, and E243 were protonated while H277 was doubly protonated (Fig. 6.3). This protonation scheme proved problematic in initial simulations, however: an unexpected accumulation of chloride ions in the extracellular domain and a strong rectification effect were observed. After consulting pK_a values predicted with PROPKA3[314,315], an alternative protonation scheme in which only E26, E35, E75, E177, E243 and H277 were protonated (doubly for H277) was also considered.

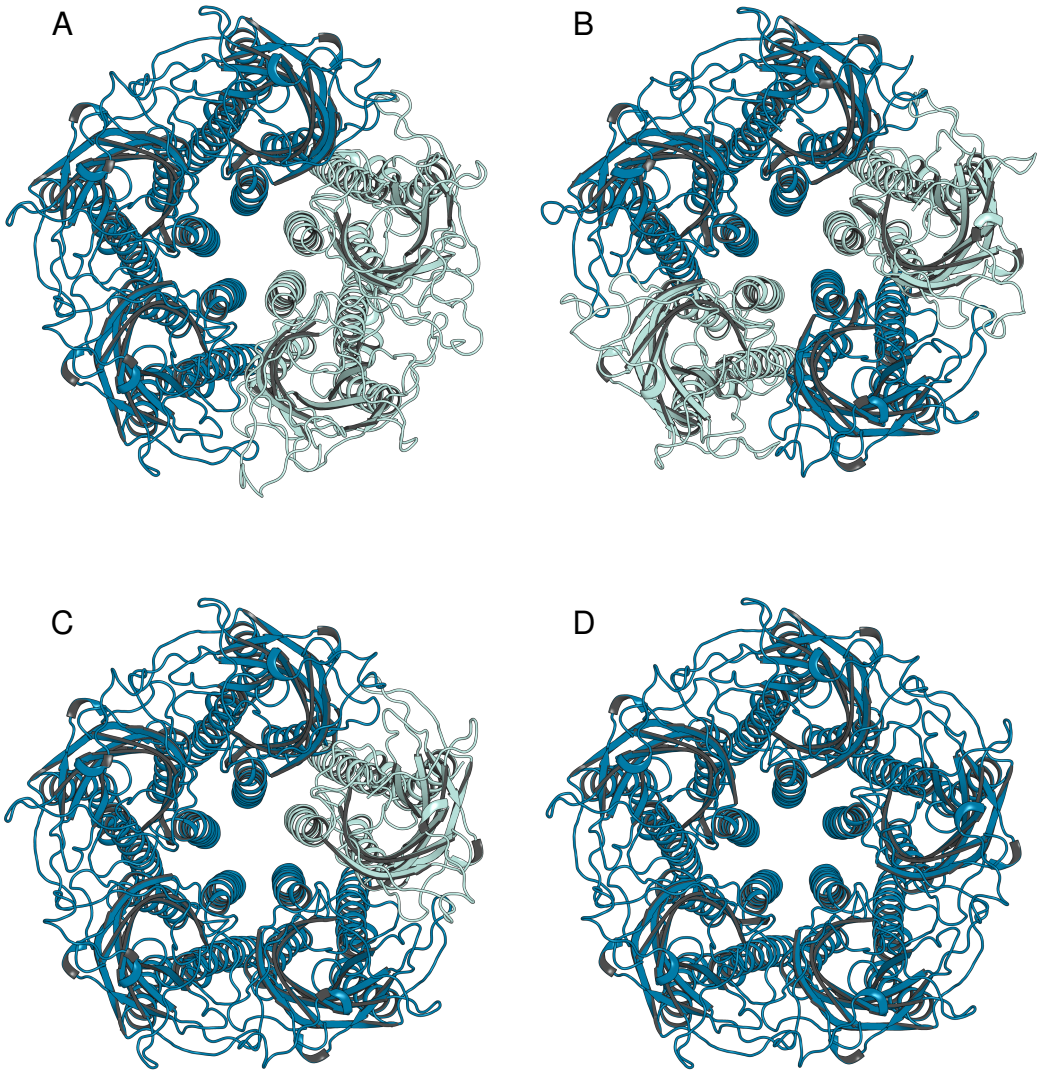

Figure 6.2: View onto the extracellular domain of the four GLIC structures used in this study. A: *iiooo*. B: *ioioo*. C: *ioooo*. D: *ooooo*. The colouring indicates the *i* (light blue) or *o* (dark blue) conformation.

The proteins were embedded in lipid bilayers consisting of DOPE, POPC, and POPS in a 2:1:1 ratio and solvated with water containing neutralising 0.15 M NaCl. The CHARMM36m force field[293,294] was chosen for proteins and lipids and the TIP3P water model[223] was used. Water molecules were constrained with the SETTLE algorithm[148], and all other bonds to hydrogen atoms were constrained with LINCS[147] to allow the use of a timestep of 2 fs. Long-range electrostatics were handled with the smooth Particle Mesh Ewald[149]. Simulations were run in the isothermal-isobaric ensemble at 310 K and 1 bar through the use of the

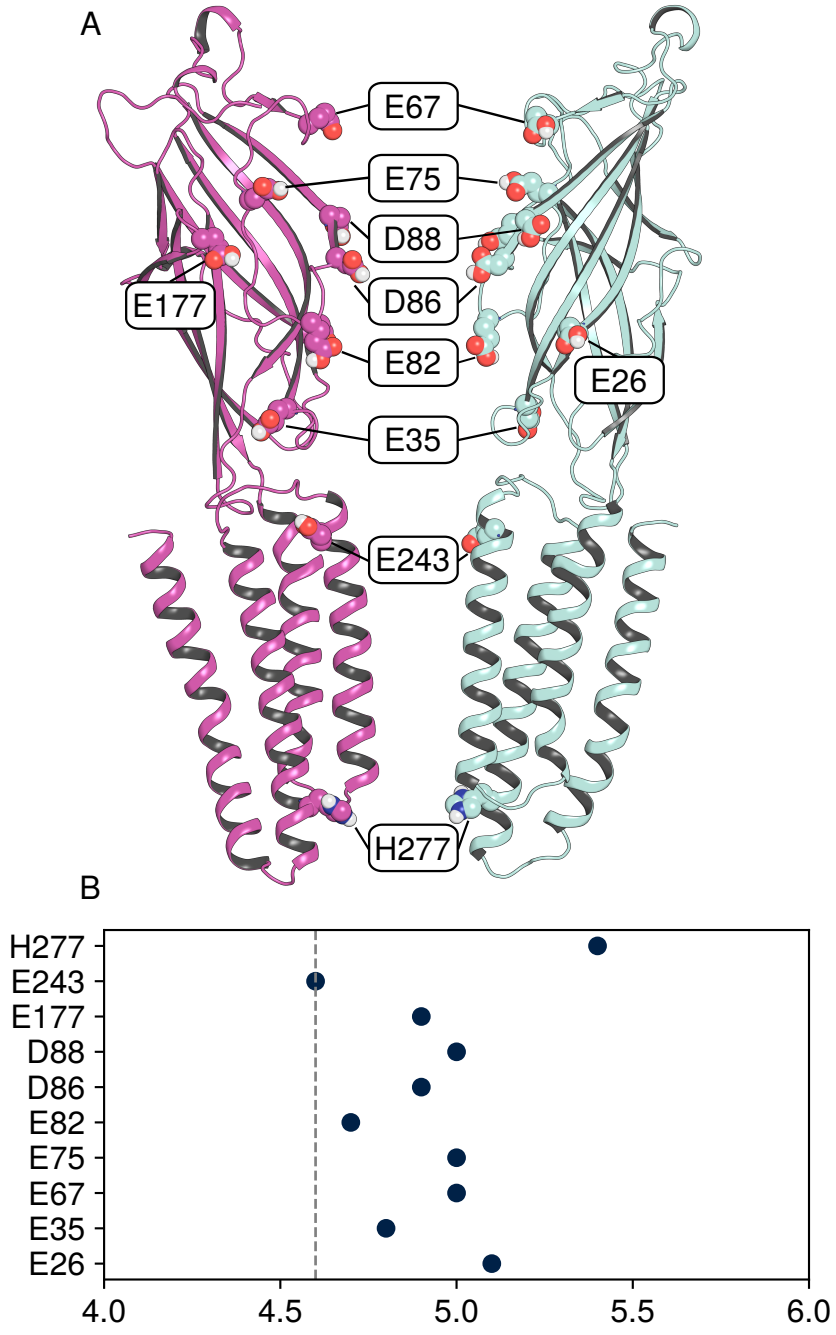

Figure 6.3: Original protonation state used in this study, going back to Nury et al.[311]. A: Visualisation of protonated residues. E67, D88, D86, and E82 are the most pore-facing. Only the more visible copy of E177 and E26 is indicated, but both residues were protonated in all subunits. B: Reported pK_a values of the ten residues chosen for protonation[311]. The experimental pH of 4.6 is highlighted with the dashed line.

V-rescale thermostat[151] and the C-rescale barostat[152], with coupling constants of 1 ps and 5 ps, respectively.

6.2.2 Equilibration procedure

After a round of steepest-descent energy minimisation, a stepwise equilibration strategy was used (Table 6.1).

Table 6.1: Overview of the parameters used in each equilibration step of the four GLIC systems.

Step	n_{steps}	timestep [fs]	Restraint force constant [$\mathrm{kJ\ mol^{-1}\ nm^{-2}}$]		
			backbone	sidechain	lipids
1	125000	1	4000	2000	1000
2	125000	1	2000	1000	400
3	125000	1	1000	500	400
4	250000	2	500	200	200
5	250000	2	500	50	40
6	25000000	2	500	0	0

This procedure is modified from the default CHARMM-GUI equilibration parameters in that the restraints on backbone atoms are not lowered below 500 kJ mol^{-1} nm^{-2}. This is because an additional equilibration step is carried out. Inspired by work done by Dämgen and Biggin[59] on the glycine receptor, harmonic, flat-bottom distance restraints were applied between Cα atoms of key pore-lining residues (Fig. 6.4). These restraints prevented the pore from closing, but left the protein free to move otherwise. This additional equilibration step has previously been shown to stabilise open pore conformations and was run here for 200 ns with restraints between Cα atoms of T20′, I16′, A13′, I9′, S6′, T2′, and E-2′. The flat-bottom restraints were set to act repulsively with a force constant of 10 000 kJ mol^{-1} nm^{-2} if two opposing Cα atoms moved closer than in the original cryo-EM structure.

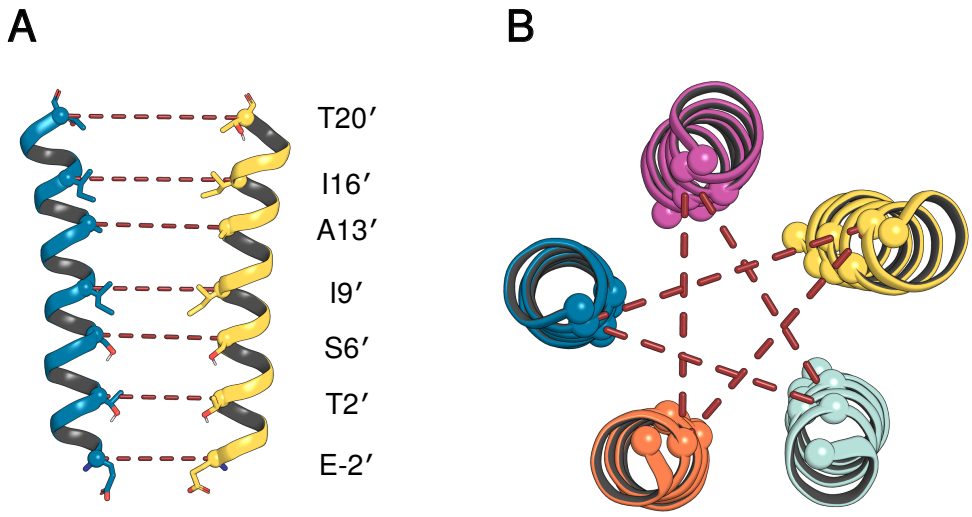

Figure 6.4: Application of cross-pore restraints as inspired by Dämgen and Biggin[59]. A: The flat-bottom restraints act the Cα atoms of pore-facing residues on M2. B: For each of these residues, the restraints are applied to its two opposing copies, resulting in a star-shaped restraint pattern.

6.2.3 Production simulations

From the final equilibrated structure, independent 200 ns production simulations with velocities randomly drawn from the Maxwell-Boltzmann distribution were started in triplicate. Data from unbiased MD simulations was used to determine pore radii, pore hydration, and water free-energy profiles with the Channel Annotation Package (CHAP)[168]. The trajectories were analysed at 100 ps intervals with a 0.14 nm kernel bandwidth for water density estimation.

Ion permeation through the channels was investigated in simulations with a uniform external electric field (see Section 2.1.9). Conductance was probed in both directions, and membrane potentials of ±500 mV were applied. To determine the strength of the simulated external electric field in $V\,nm^{-1}$ to be specified, the desired membrane potential was divided by the length of the simulation system in the membrane-normal direction. Ion conductance was then assessed by plotting the coordinates of ions in the membrane-normal direction over time.

6.3 Results

6.3.1 Intermediate conductance and strong rectification of initial simulations

We began our investigation of intermediate GLIC structures with simulations of *iiooo*, *ioioo*, and *ioooo*. Unbiased simulations following on from 200 ns equilibration simulations with cross-pore restraints (Fig. 6.4) were run to assess stability of pore conformations and pore hydration to test for the presence of hydrophobic gates[316]. Trajectories from these simulations were analysed with the Channel Annotation Package (CHAP)[168] (Fig. 6.5).

These plots show a trend of pore openness. *ioooo* has the most open pore of these three structures, as is visible in water number densities and water free energies, but it is less hydrated than a previously published open state structure[313]. Both *iiooo* and *ioioo* show clear dewetting at the I9′ gate. Of the two, *ioioo* is more strongly dewetted, with the standard deviation of the water number density being very nearly zero at the I9′ gate. While it has to be noted that in the absence of any bias or restraint, the pores do contract to a radius below that of a hydrated sodium ion, a trend of pore openness (with *ioooo* being most open-like and *ioioo* being most closed-like) is visible in the CHAP data. We assigned *ioooo* as "partially open".

Channel conductance was tested in simulations with an external electric field applied to systems after equilibrations with cross-pore restraints. By applying effective membrane potentials of ±500 mV to backbone-restrained structures, ions can be driven through the channel. These transition events can then be counted to arrive at channel conductance estimates. *ioooo* showed by far the most number of ion transitions, followed by *iiooo*. Consistent with the CHAP analysis, *ioioo* showed intermittent dehydration of the pore as well as little to no conductance (Fig. 6.6 and SI Figs. E.1–E.3).

For easier comparison, the ion permeation counts are also listed in Fig. 6.7. The *ioooo* structure has significantly higher conductance than *iiooo* and *ioioo*, and

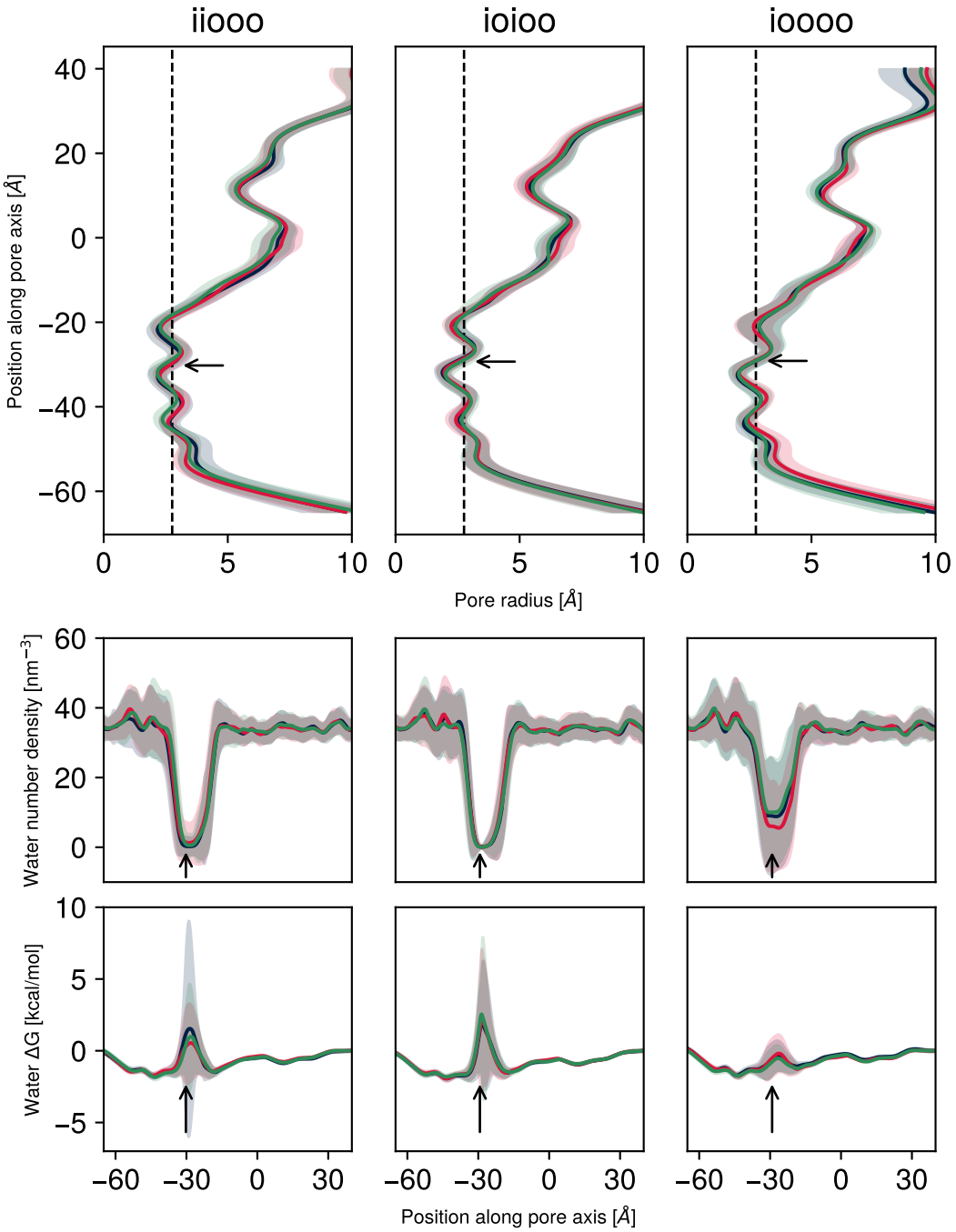

Figure 6.5: CHAP results showing pore radius profiles (top), number densities of water molecules in the pore (middle row), and water free energies (bottom) of three independent simulations of the *iiooo* (left), *ioioo* (middle column), and *ioooo* (right) systems. Shown are the average values of each quantity across entire trajectories, the shaded areas signify the standard deviations. The arrows mark the position of I9′ in the last simulation frame, and the dashed line shows the radius of a sodium ion with its hydration shell.

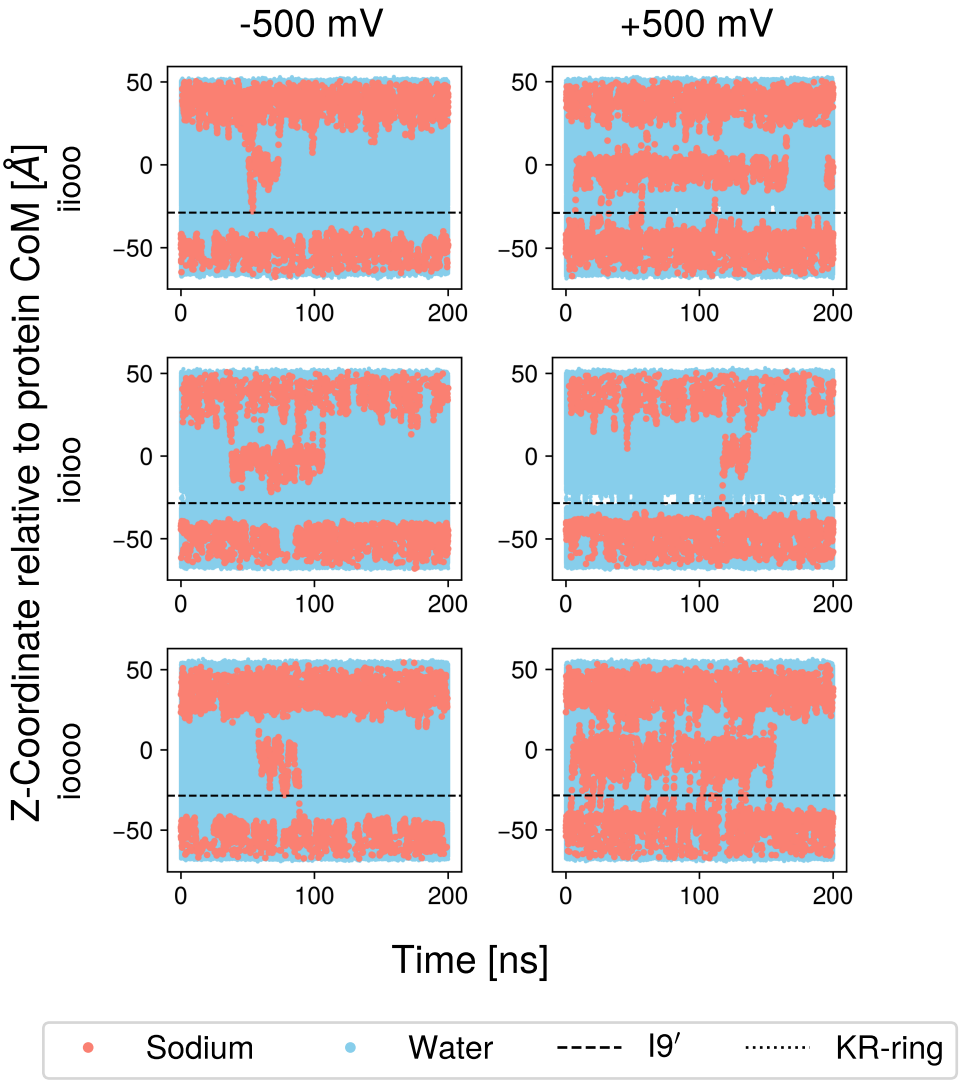

Figure 6.6: Analysis of ion permeation events in GLIC systems *iiooo*, *ioioo* and *ioooo*. The z-coordinate of sodium ions and water molecules in the pore is plotted over time to visualise conduction events. A trend in conductance is visible, with *ioooo* being the most and *ioioo* being the least conductive structure. Across all systems, more conduction events occur with a simulated positive membrane potential.

it is significantly sodium-selective. However, these results revealed a number of problems with our simulations. For one, a strong rectification effect is apparent: the simulations suggest that the channel is only conductive under positive membrane potentials, driving sodium ions out of the cell. However, even though the physiological role of GLIC in *Gloeobacter* is unclear, the electrochemical gradient across the membrane suggests that opposite currents should also be possible. As such, the fact that but a few permeation events were observed under negative

membrane potentials is concerning. Additionally, GLIC is known to be strongly cation-selective[317], but the simulation shows only a 2:1 preference for sodium over chloride ions. As the simulations showed unexpected behaviour in terms of rectification and ion selectivity, we decided to reassess our initial assumptions, starting with the protonation state of the channel.

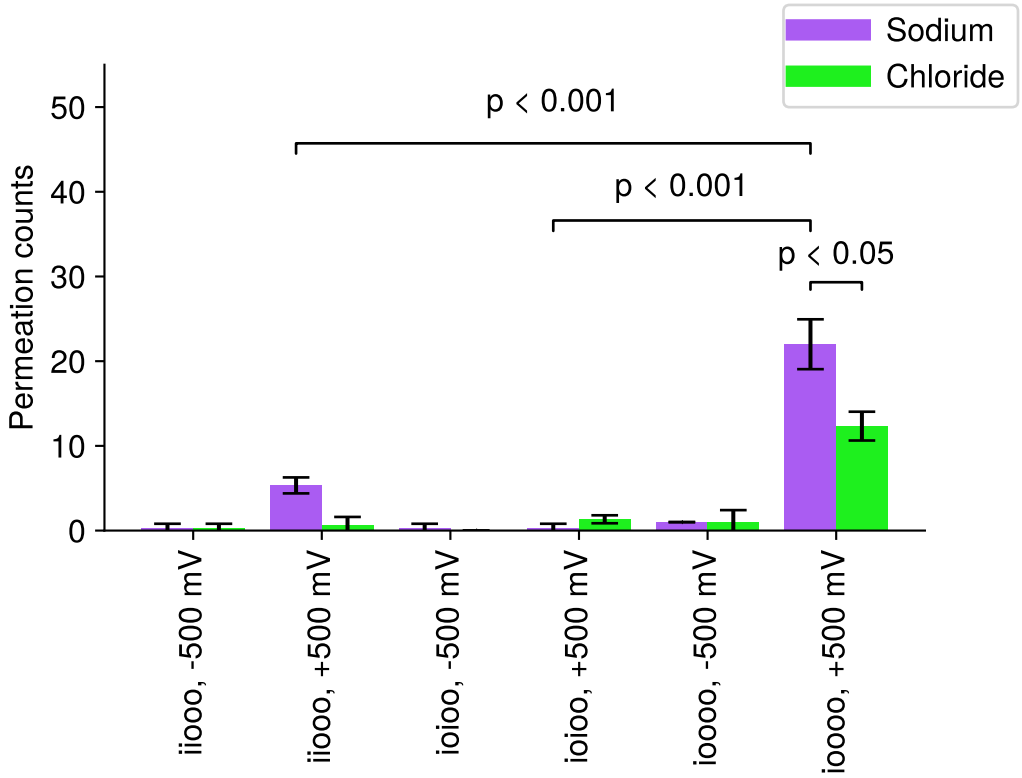

Figure 6.7: Counts of ion permeation events across all tested systems and field directions. *iooooo* is the most conductive system and displays sodium-selectivity. Shown are means ± standard deviations from three repeats. Statistical significance was tested with Welch's *t*-test[318].

6.3.2 Deliberations on protonation states

To address the issue of rectification, we closely examined our simulation data. The above permeation analysis plots suggest the presence of a second permeation barrier in addition to I9′, located at roughly 10 Å. It became apparent that a ring of positively charged residues formed of K38 and R109 — from now on dubbed the KR-ring — is located at this position and likely is the reason for this observed

barrier (Fig. 6.8). In fact, the KR-ring and the hydrophobic gate at I9′ appear to form the boundaries between distinct reservoirs of ions within GLIC. Such reservoirs had first been observed by Nury et al.[311]. It appeared that the KR-ring posed a rather formidable barrier to permeation in our simulations. Since GLIC is a proton-gated channel, we reasoned that protonation states of residues in the vicinity of K38 and R109 might contribute to this behaviour.

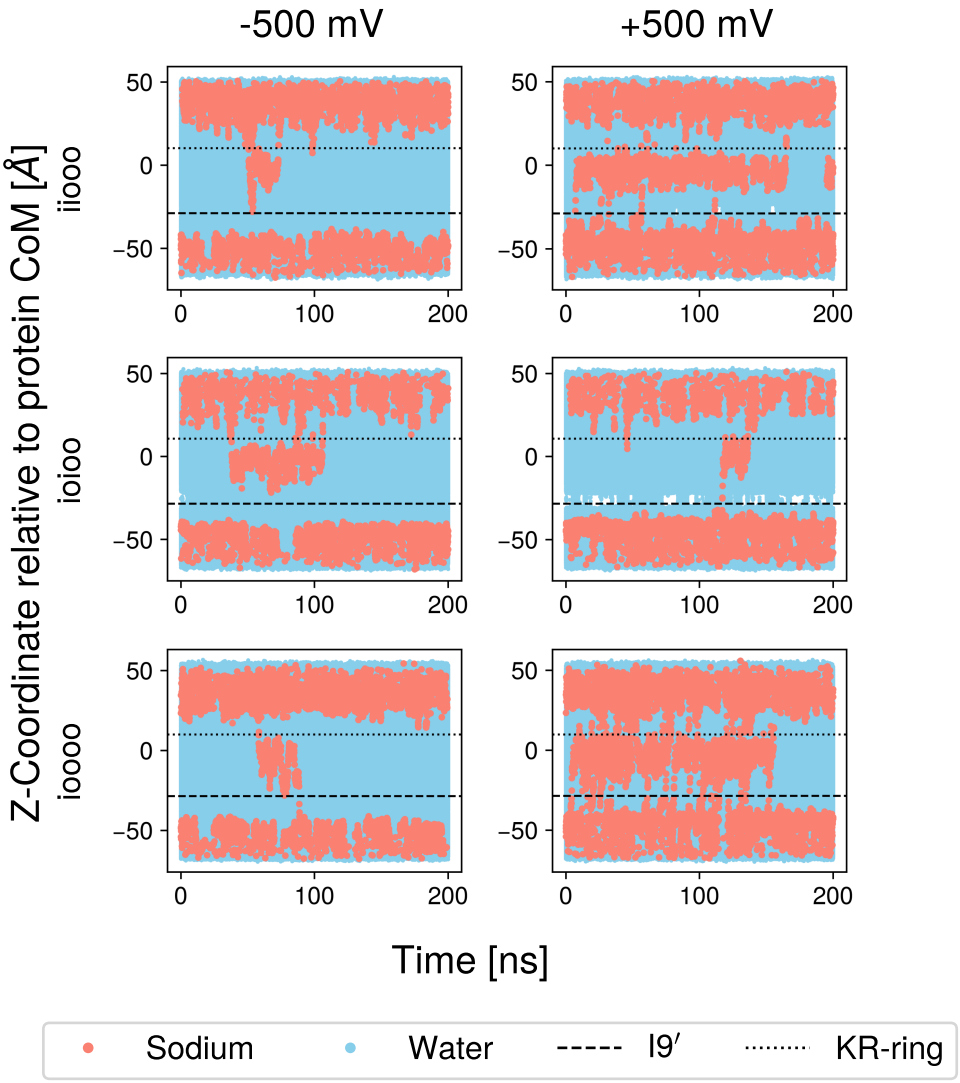

Figure 6.8: Representative plots from Fig. 6.6 are reproduced. In addition to I9′, the position of K38 and R109 ("KR-ring") is also shown. Both groups of residues clearly form barriers to permeation in both directions.

A prediction of pK_a values of acidic residues in GLIC revealed that the protonation scheme going back to Nury et al.[311] requires further refinement (Fig. 6.9).

Some of the pK_a values reported back then cannot be reproduced here, leaving room for adjustments of the protonation states as determined through a comparison of pK_a and pH. The Henderson-Hasselbalch equation

$$\mathrm{pH} = \mathrm{p}K_a + \log_{10}\left(\frac{[\mathrm{A}^-]}{[\mathrm{HA}]}\right), \tag{6.1}$$

where HA is a weak acid and A^- is its deprotonated form, allows estimation of the protonation state of a weak acid. If $\mathrm{pH} = \mathrm{p}K_a$, 50 % of the acid molecules will be protonated. If $\mathrm{pH} - \mathrm{p}K_a = 1$, protonation will be ≈91 %, and at $\mathrm{pH} - \mathrm{p}K_a = 2$, protonation increases to ≈99 %. Based on this estimation, while the protonation state of E26 or E243 is undisputed, a more nuanced behaviour is to be expected from other residues we previously protonated.

We then re-visualised the systems with the effect of the KR-ring in mind and found that E82, D86, and D88 are all in proximity to the KR ring and were all protonated in the first group of simulations (Fig. 6.10). In addition, E67 is located at the top of the ECD and, if deprotonated, might help attract sodium ions into the pore.

Based on the pK_a value prediction (Fig. 6.9) and the observation that D86, D88, E67, and E82 are pore-lining ECD residues, they were chosen to be left deprotonated in new simulations. In addition to stabilisation of the KR-ring in a conformation conducive to ion conduction, we hoped the tight clustering of the negative charges of D86 and D88 might repel chloride ions that were observed to cluster in the ECD vestibule in previous simulations. A strong effect is visible as early as the equilibration stage (Fig. 6.11), where chloride ions indeed no longer accumulate to the same extent under the new protonation scheme. The same figure also shows sodium ions entering the ECD in the deprotonated state. Thus, this new protonation scheme under which only E26, E35, E75, E177, E243 and H277 were (doubly) protonated was chosen for a new set of simulations.

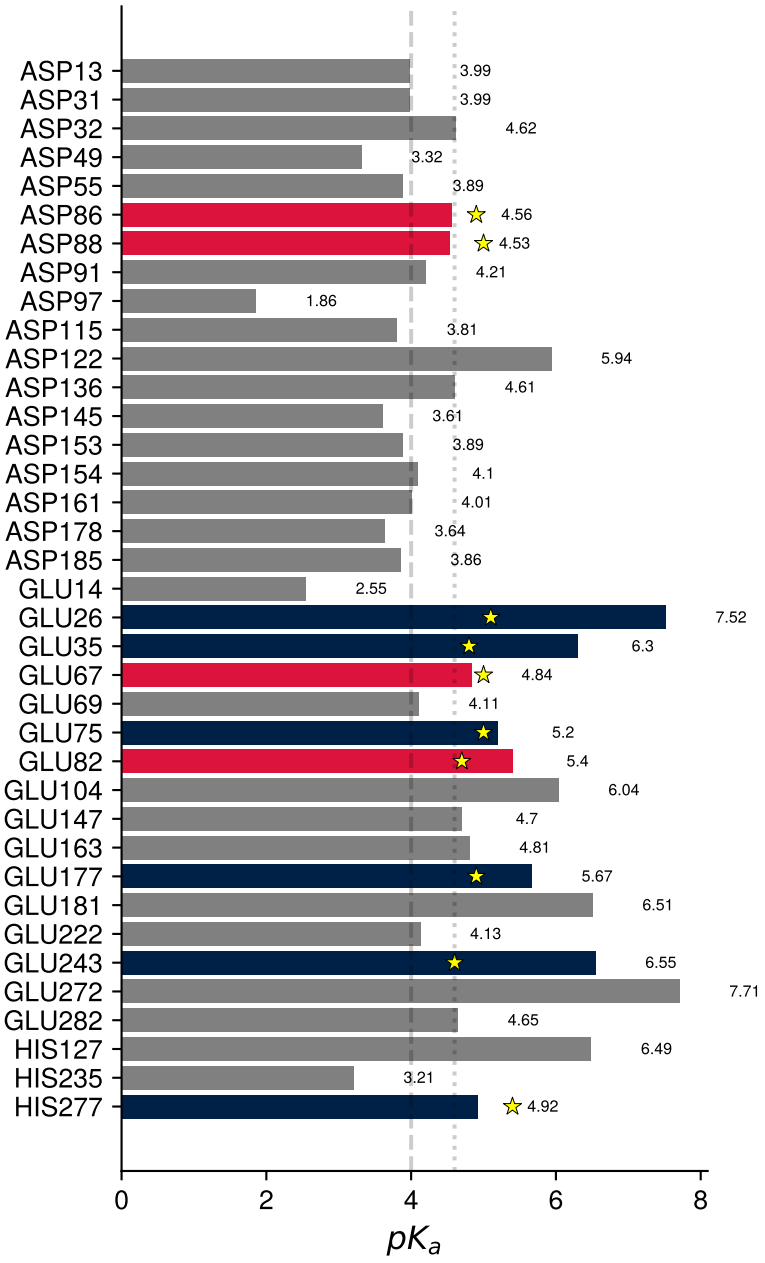

Figure 6.9: PROPKA 3 prediction[314] of pK_a values of all aspartates, glutamates, and histidines in GLIC, performed on the *ooooo* structure. We cannot reproduce some of the values reported by Nury et al.[311], indicated as yellow stars. Based on their suggestion, residues shown with blue or red bars were protonated in previous simulations, but there are inconsistencies. Comparing these pK_a predictions to pH 4 (dashed line, used by our collaborators) or pH 4.6 (dotted line, used for structures simulated by Nury et al.) suggests the protonation pattern is not optimal. Having established that the protonation pattern still required refinement, we decided to leave D86, D88, E67, and E82 (red bars) deprotonated in new simulations. These residues line the pore in the ECD (Fig. 6.3) and are likely to affect ion selectivity. Based on the pK_a predictions here, leaving them deprotonated is no less realistic than leaving, say, D122 or E181 deprotonated.

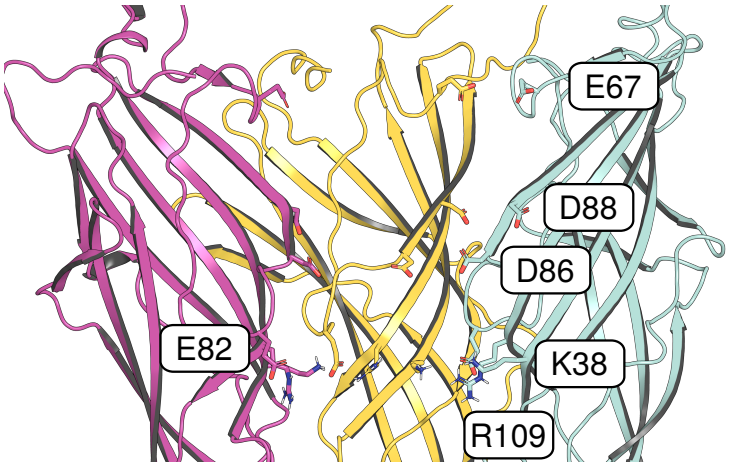

Figure 6.10: Visualisation of D86, D88, E67, and E82. They are left deprotonated in a new protonation scheme. D86 and D88 form a ring of negative charges that might help repel chloride ions. E67 might capture sodium ions at the opening of the ECD. E82 is now able to form salt bridges with residues K38 and R109 of the KR-ring that was seen to be a barrier to permeation (Fig. 6.8.)

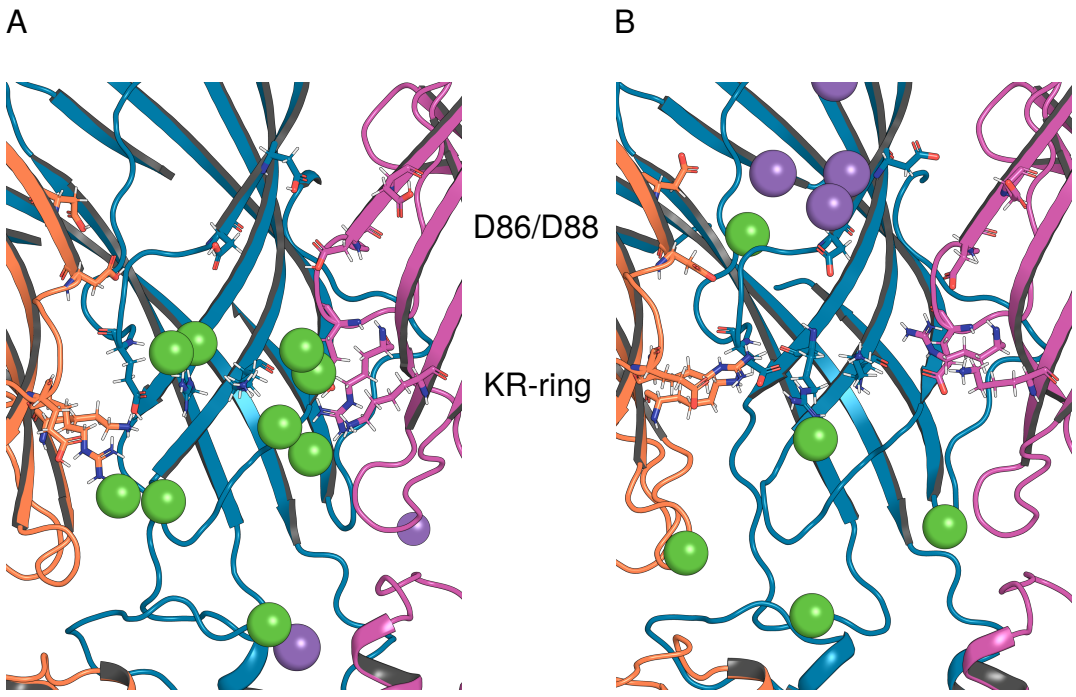

Figure 6.11: Comparison of the final snapshot of the last equilibration step of *iiooo* under the old (A) and new (B) protonation schemes. With E67, E82, D86 and D88 deprotonated, chloride ions accumulate at the KR-ring and the vestibule between it and the transmembrane domain. This accumulation is strongly reduced under the new protonation scheme, and sodium ions enter the ECD more readily.

6.3.3 Providing evidence of partial conductance

A new set of simulations of *iiooo*, *ioooo*, and *ooooo* under the new protonation scheme was carried out. The goals for these simulations were threefold. First, to determine if the rectification observed before can be influenced by the protonation state. Initial observations of ion behaviour suggested this might indeed be the case (Fig. 6.11). Second, to investigate the dependence, if any, of ion selectivity on the protonation state. Third, to show that *ioooo* is a partially conductive structure by comparing its behaviour to that of fully-open *ooooo*.

For the analysis of pore hydration with CHAP in these new runs, the trajectories from the cross-pore restraint equilibration simulations were used for CHAP calculations. This has the advantage of not requiring a second batch of simulations and being more directly comparable to the permeation plots below, but means that comparison between the data reported in Fig. 6.5 and Fig. 6.12 is not useful. The CHAP results show that all three structures in this simulation setup are fully hydrated throughout the pore and remained open throughout the simulations (Fig. 6.12).

Simulated electric fields were able to elicit permeation events in all three systems, but a strong rectification effect is still visible (Fig. 6.13). Strikingly, *iiooo* now shows much higher conductance than it did under the old protonation scheme. The KR-ring still acts as a barrier to permeation (Fig. 6.14), meaning that the new protonation states failed to make it more permeable by stabilising more conductive orientations. However, comparison of Figs. 6.8 and 6.14 shows that sodium ions are now able to approach the KR ring closer from the top of the ECD than they were before. This is likely due to deprotonated D86 and D88 attracting sodium ions and repelling chloride ions (Fig. 6.11), suggesting that, while the issue of rectification is not solved here, protonation states are indeed likely to play a part.

Statistical analysis of permeation counts from all replicates (Figs. E.4–E.6) shows that the all-open *ooooo* state has a statistically significantly higher conductance than the next-most open *ioooo* (Fig. 6.15). This is clear evidence of an intermediate conductance, allowing the annotion of *ioooo* as "partially conductive".

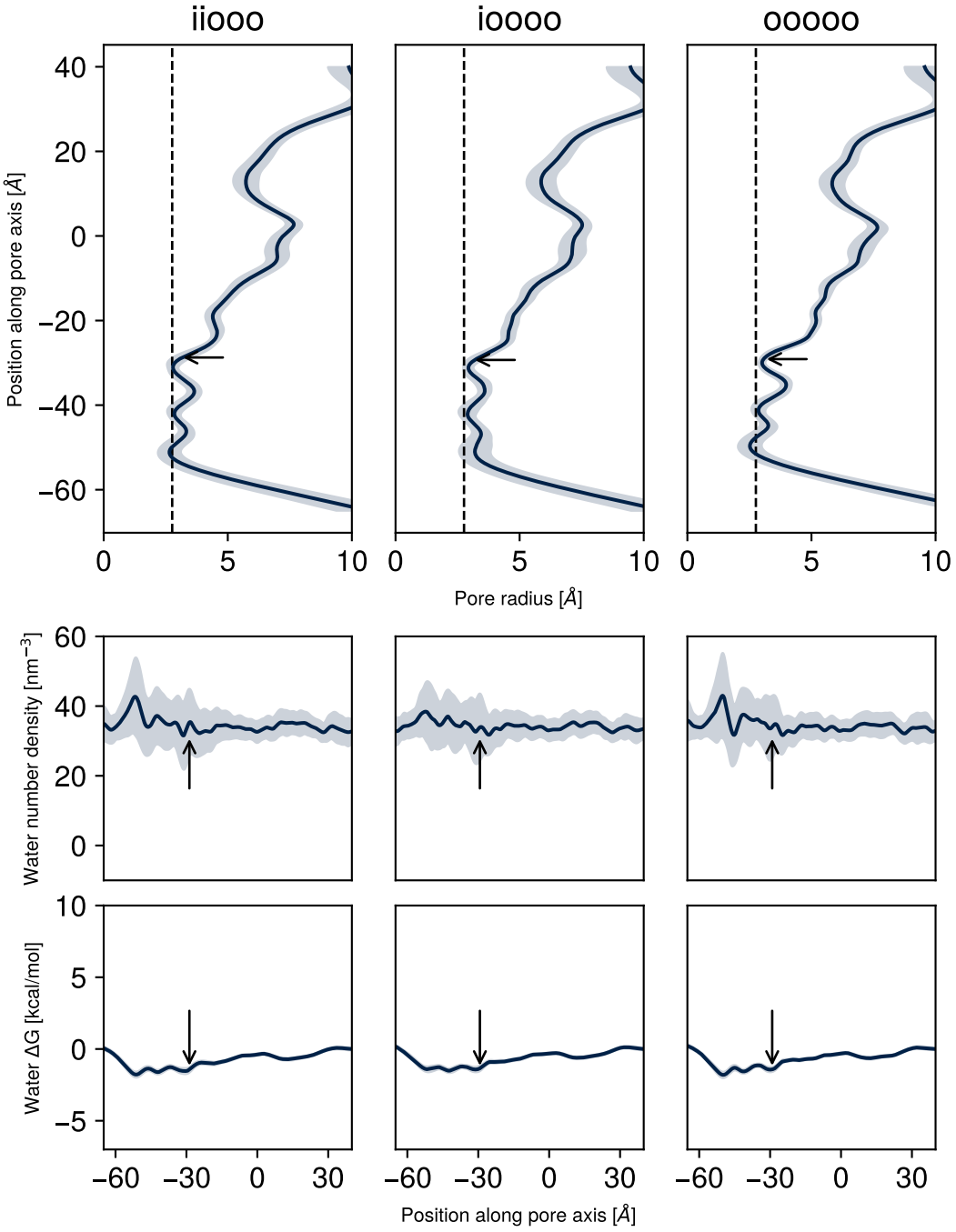

Figure 6.12: CHAP results for trajectories collected with the new protonation scheme during the cross-pore restraint equilibration run. Pore radius profiles (top), number densities of water molecules in the pore (middle row), and water free energies (bottom) of three independent simulations of the *iiooo* (left), *ioooo* (middle column), and *ooooo* (right) systems. The arrows mark the position of I9′ in the last simulation frame.

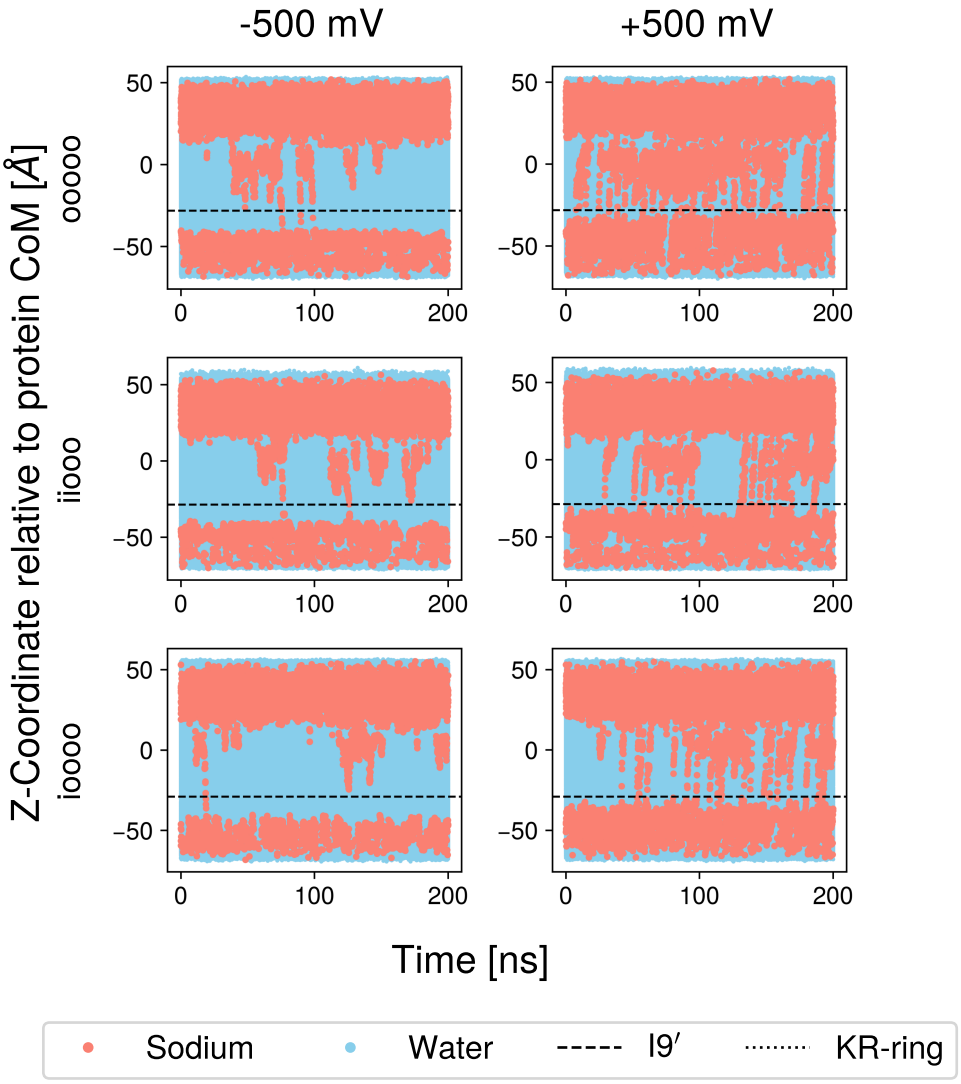

Figure 6.13: Permeation plots for *ooooo*, *iiooo*, and *ioooo* with the new protonation states. The rectification behaviour is still clearly visible. In these runs, *iiooo* is more conductive than in runs using the old protonation states (Fig. E.1).

This analysis also shows cation selectivity more clearly. However, the rectification behaviour remains an obvious problem. Furthermore, *iiooo* is now shown to be equally as conductive as *ioooo*, whereas there was a significant difference between their conductances under the old protonation scheme (Fig. 6.7). Of the three goals for these simulations, improving the selectivity and providing evidence for partial conductance of *ioooo* were achieved, while the rectification behaviour requires further study.

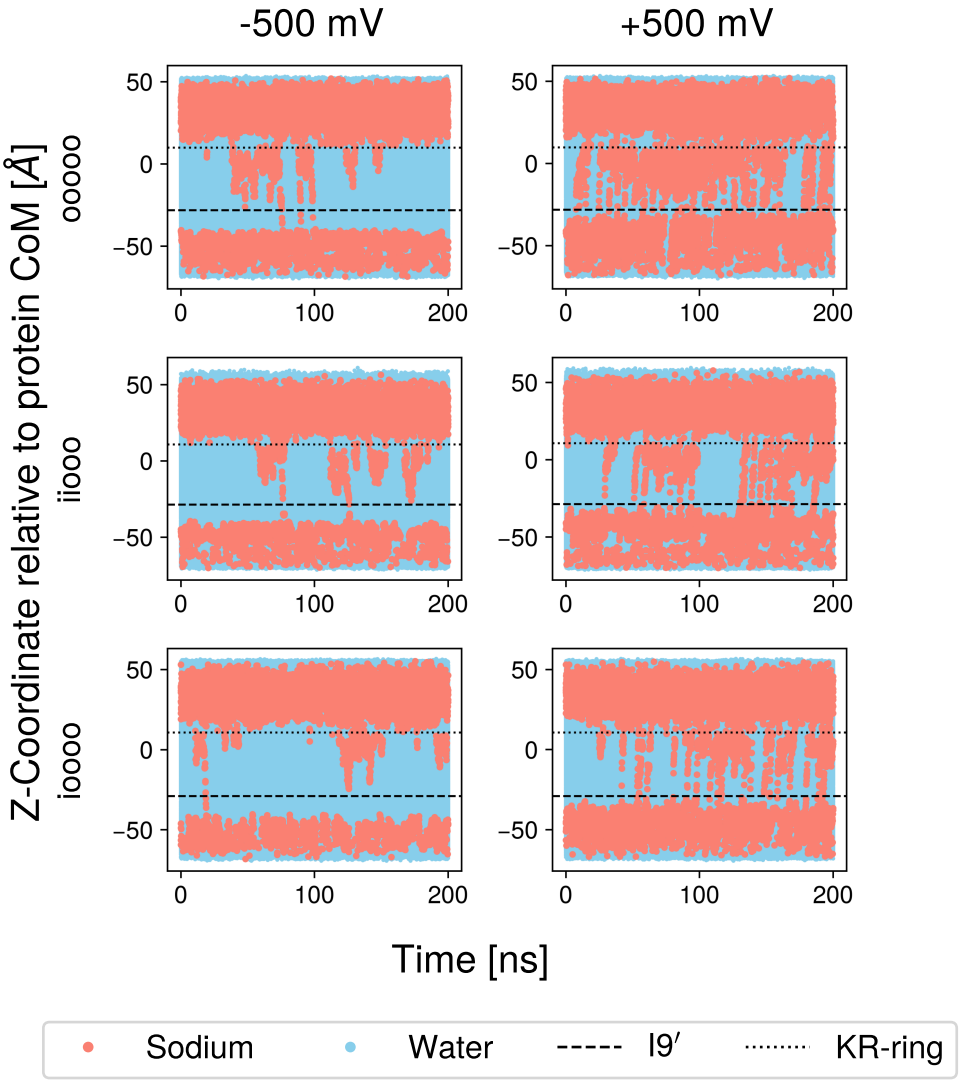

Figure 6.14: Permeation plots from Fig. 6.13 with the location of K38 and R109 indicated. Even when E82 is deprotonated and available for interactions with R109 (as seen in Fig. 6.10), the KR-ring still poses a barrier to permeation in both directions.

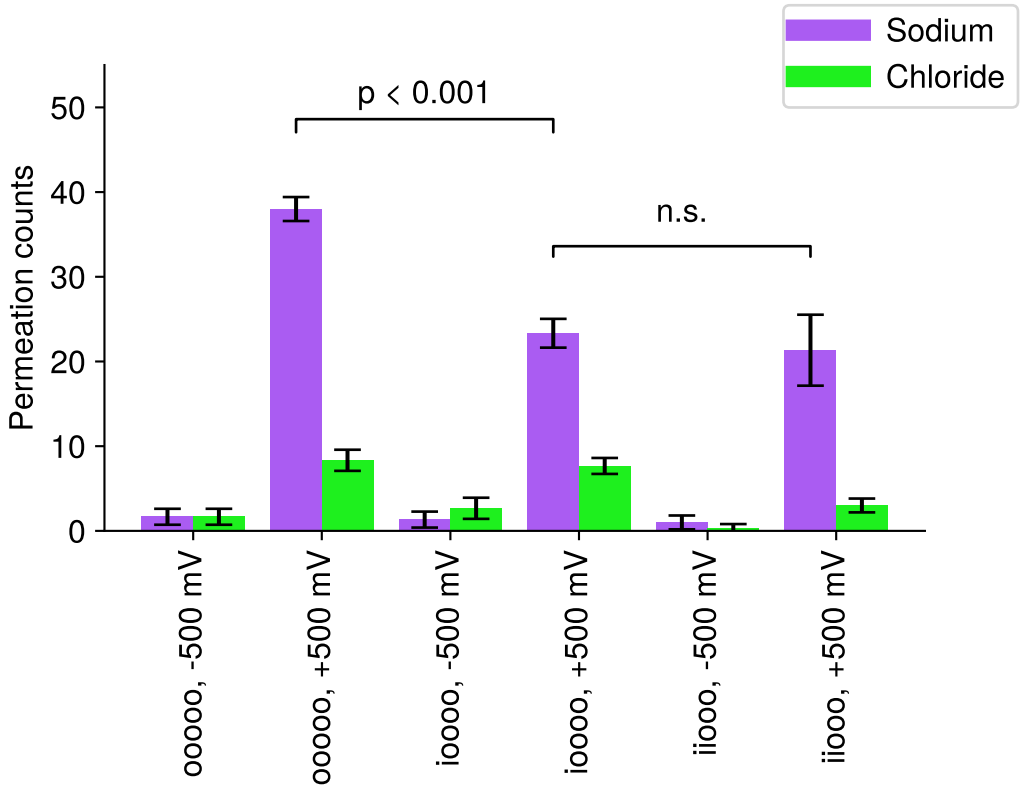

Figure 6.15: Permeation count statistics of simulations with the new protonation scheme. *ooooo* is significantly more conductive than *iooooo*, justifying the label of "partially conductive" for the latter. A clear sodium selectivity is observed. Shown are means $\pm$ standard deviations from three repeats. Statistical significance was tested with Welch's *t*-test[318].

6.4 Discussion and future directions

The *Gleobacter violaceus* ligand-gated ion channel (GLIC) is considered a simplified model system of eukaryotic pentameric ligand-gated ion channels. However, while that might be true to the experimentalist, it has continually proved to be a complicated system to study with molecular simulation. We carried out functional annotation of the intermediate *ioooo* structure and discovered it was partially conductive compared to the fully-open state. Even though we achieved this goal, we are left with a number of questions. The strong rectification effect that we have observed in all conditions here and that have previously been reported by others[308,313] is surprising. While the physiological role of GLIC is unknown, we can speculate as to its function in the bacteria. *Gleobacter violaceus* are gram-negative cyanobacteria that do not have thylakoids. Instead, their photosynthesis machinery is embedded in the inner cell membrane, building a proton gradient across it by pumping protons into the periplasm[319]. The GLIC channel is likewise embedded in the inner membrane, with its "extracellular" domain located within the periplasm. It is conceivable that GLIC evolved to react to the proton gradient created by the photosystem, potentially as a way to limit the electrochemical gradient across the inner membrane. Whatever its physiological role may be, the channel's localisation reinforces the intuition that cations should flow into the cell, as the proton gradient across the plasma membrane strengthens the electrochemical potential that would drive cations into the cell even under non-energised conditions. This makes our observation of rectification in the opposite direction puzzling. To study this phenomenon in detail, further collaboration between experimentalists and computational scientists offers the most promising route. Patch clamp electrophysiology in the inside-out and outside-out configurations could provide experimental data for multiple ionic currents at different pH values in both directions, which would then be combined with further MD simulation data. It appears clear that protonation states of key residues in the ECD need to be a main focus of such simulations. In simulations

performed here, protonation was carried out in an all-or-nothing manner, where each of the five copies of a residues had the same protonation state. However, to better mirror predictions based on Henderson-Hasselbalch estimations (Eq. 6.1), choosing a subset of residues to (de-)protonate might prove beneficial. For example, if pH is 0.6 below pK_a, then Eq. 6.1 predicts that ≈81 % of residues should be protonated. Randomly choosing one residue to leave deprotonated would better match this prediction, however, this requires a refined estimate of pK_a values. This approach would also drastically increase the already sizeable number of combinations of protonation states to test, as many residues in Fig. 6.9 would need to be varied to solve the underlying combinatorial problem. As an alternative approach, this problem might be suitable for study with constant pH simulations[320]. Here, titratable residues are allowed to change between protonation states through λ-dynamics. This could yield a better representation of the system under a given pH value than an all-or-nothing assignment of protonation states.

It is surprising to see the behaviour of *iiooo* change so drastically between protonation states. Future studies should reconsider the use of the cross-pore restraint equilibration step described in Fig. 6.4. The differences between *i* and *o* configuration are subtle (Fig. 6.1), and while the equilibration protocol described by Dämgen and Biggin[59] is suited for the stabilisation of open states of pentameric ligand-gated ion channels, it cannot prevent the loss of these subtle structural distinctions between i and o. As such, it is unclear whether the different conductance of *iiooo* is due to the protonation scheme or subtle differences in the equilibration trajectories. The intermediate structures described by our collaborators are clearly but subtly different members of a structural ensemble. Use of backbone restraints throughout the stimulation pipeline as opposed to equilibration under cross-pore restraints might be justified to guarantee maintaining these distinct states.

Despite the problems outlined above, the key finding and contribution of this work remain sound. *ioooo* shows the same conductance in both protonation states, and is shown to be significantly more conductive than the less open-like *iiooo*

structure in the old protonation state and significantly less conductive than the fully-open *ooooo* structure. As such, our simulation results support labelling *ioooo* as a partially conductive pre-open state, providing further evidence for asymmetric gating behaviour and mechanistic insight into pore opening in the GLIC channel.

And here I am, for all my lore, the wretched fool I was before

— Heinrich Faust, in Johann Wolfgang von Goethe's *Faust, der Tragödie erster Teil*

7 Conclusion

7.1 Overview

Ion channels are indispensable components of the nervous system due to their rapid signal transmission[321]. Pentameric ligand-gated ion channels (pLGICs) in particular are involved or implicated in, among others, synaptic inhibition[33,322], cognition[323,324], nociception[325], neuromuscular transmission[326], epilepsy[327], and hyperekplexia[328].

This broad range of functions is mediated through subunit diversity, but a significant degree of structural similarity between subunits means that many compounds act on multiple channel family members. For example, ivermectin, picrotoxin, and propofol all act on multiple pLGICs[87,329–336]. It thus appears difficult to develop drugs that are specific to a pLGIC member or even to a specific subunit composition. Such compounds are, however, highly desirable to improve drug profiles and reduce off-target effects[337–341].

In this thesis, the activation cycle and drug interactions of pentameric ligand-gated ion channels (pLGICs) were investigated with molecular dynamics simulations, with a particular focus on the glycine receptor (GlyR). The identification of drug-permeable fenestrations in the glycine receptor in particular is of note in the context of subtype-specific modulator design.

7.2 Conclusions

In a collaborative effort, we were able to demonstrate the existence of intermediate, asymmetric functional states of the *Gloeobacter* ligand-gated ion channel (GLIC), a prototypical pLGIC. These intermediate states displayed stepwise, per-subunit activation reorientations. The main result of this study was the identification of these intermediate states through a refined cryo-EM analysis pipeline. My contribution was the assignment of functional states to the most open-like intermediates. In addition to the finding that some pre-open states already display partial conductivity, I reported statistically significant rectification behaviour in these simulations which is sure to prompt further investigation into the protonation states of key gating residues of this pH-gated ion channel.

A second collaboration was aimed at improving our understanding of the allosteric network in the glycine receptor. This network is responsible for the transmission of the structural rearrangements triggered by agonist binding in the extracellular domain (ECD) towards the transmembrane domain (TMD), where the activation gate of the channel lies and is opened in reaction to this signal. We reasoned that structures of GlyR with the TMD-binding allosteric activator ivermectin and the orthosteric inhibitor strychnine bound at the same time might show intermediate activation behaviour, with an activated-like TMD and an inactivated-like ECD. Our collaborators provided these structures, and I carried out MD simulations with a particular focus on how ivermectin disrupts strychnine binding. I was able to identify the precise elements of the strychnine binding site that react most strongly to the presence of ivermectin. In this study, we were able to demonstrate further that the conformational wave is able to communicate changes in both directions.

Faced with a simulation system in which a potential of mean force (PMF) calculation was too slow to converge to be feasible, I developed an improved umbrella sampling workflow by combining three existing techniques. Umbrella windows generated through alchemically growing the ligand in place along the reaction

coordinate instead of pulling it in steered MD simulations reduces the hysteresis artefacts seen in PMFs. As shown by its original authors, Simulated Tempering-enhanced Umbrella Sampling (STeUS) outperforms standard umbrella sampling simulations when slow orthogonal degrees of freedom hamper PMF convergence. However, a downside of STeUS in its original description is that a majority of simulation data is not used in PMF calculation as it is collected at temperatures other than the temperature of interest. I evaluated the effectiveness of temperature reweighting with the Multistate Bennet Acceptance Ratio (MBAR) and found that, for sufficiently small systems, this can significantly improve convergence speeds. I benchmarked the effects of these three components individually in a systematic study to identify the individual and synergistic improvements in convergence speed and found that the combination is a powerful toolkit to quickly obtain converged membrane permeation PMFs. Differences in PMFs between different membrane bilayers remain clearly visible. I also demonstrate the errors associated with commonly used boundary conditions applied to PMF calculations. In all but the simplest cases, these boundary conditions introduce errors into the PMF. The slower a PMF is to converge, the larger the error introduced by these constraints.

The key result of this thesis was obtained from the application of this umbrella sampling workflow to the study of putative fenestrations in the glycine receptor. To date, no drug-permeable fenestrations have been identified in the entire superfamily of pentameric ligand-gated ion channels. However, previous electrophysiological data appeared consistent with fenestration-based inhibition of GlyR through fipronil. I set out to determine energetics of fipronil pore access via membrane permeation and via fenestration-based access (Fig. 7.1).

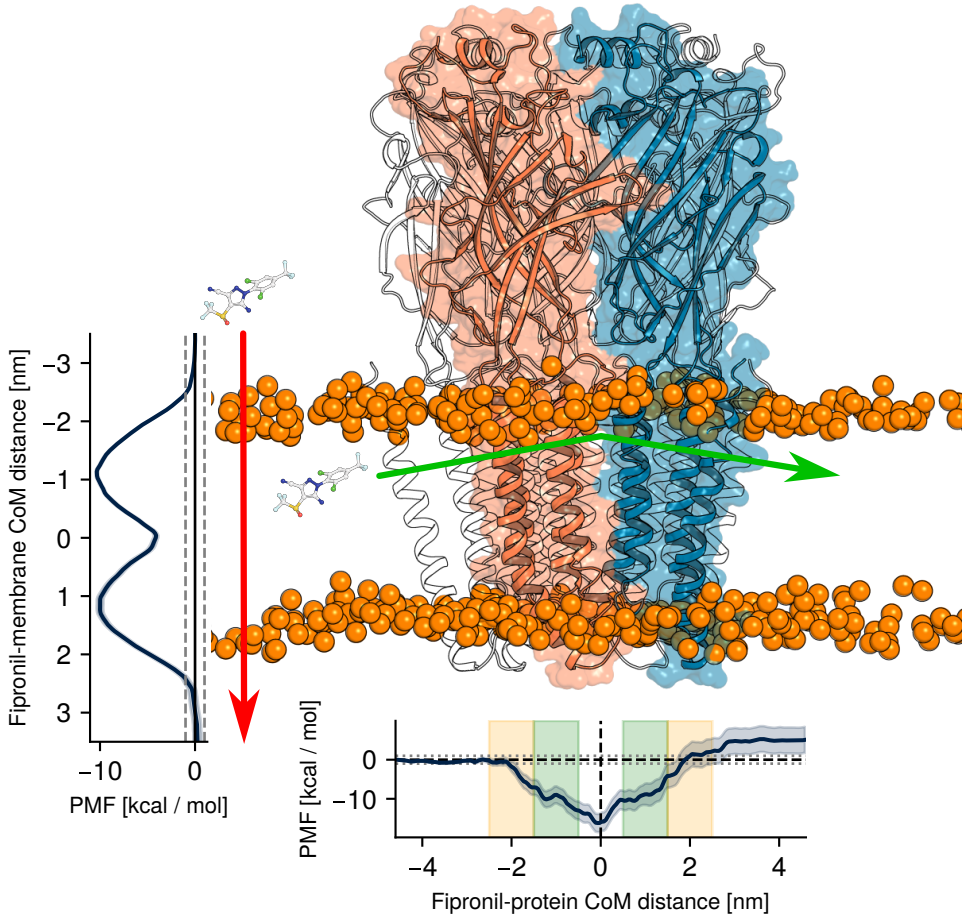

Figure 7.1: The key result of this thesis. Membrane permeation of fipronil (red arrow) has an energy barrier of 10 kcal mol^{-1}. A fipronil molecule permeating the fenestration (green arrow), on the other hand, does not encounter significant energy barriers on its way to the channel pore. Phosphorus atoms of lipids are shown as orange spheres. The protein structure is of a *Danio rerio* α1 homopentamer (PDB 3JAF[44]).

My simulations clearly show that fenestration-based access of fipronil into the GlyR pore is energetically preferable over membrane permeation. Combined with a clear state dependence of the fenestration radius I identified in the available static structures of the channel, my results provide the context needed for the original electrophysiology data. Taken together, this is strong evidence of the existence of a state-dependent and drug-permeable fenestration in the glycine receptor, the first of its kind in the pLGIC superfamily. This discovery opens the door to further directed study of this feature, which will improve our understanding

of lipid and drug interactions in pLGICs and might lead to the development of subtype-specific compounds that avoid or reduce off-target effects.

7.3 Future work

All studies reported on in this thesis would benefit from future work.

The main obstacle in our investigation of the *Gloeobacter* ligand-gated ion channel (GLIC) with molecular simulation was the assignment of suitable protonation states. Should this project be picked up again in the future, this assignment would need to be refined. While I chose protonation states in an all-or-nothing manner based on pK_a predictions, protonation of only a few of the five copies of each residue might be a more realistic representation. Alternatively, constant pH simulations could be used to dynamically protonate and deprotonate residues as appropriate.

Chapter 5 dealt with the changes in the orthosteric binding site of the glycine receptor that are allosterically induced by ivermectin. My work offers a clear description of the processes that occur in the orthosteric site to destabilise strychnine there. These results would benefit from further mechanistic study—an analysis of the dynamic network that connects the two binding sites would elucidate the precise pathways along which they communicate. This deeper level of understanding on the coupling between the extracellular and transmembrane domains could then be used to help inform further studies into GlyR activation pathways and gating.

The improved umbrella sampling workflow described in chapter 3 is robust, with the possible exception of the effect of Alchembed-based window generation. To understand why some of the tested compounds display worse convergence behaviour in Alchembed-based windows, the approach should be tested more thoroughly. One obvious difference of steered-MD windows is that the ligand orientation varies from window to window. This variance might be beneficial for the sampling of ligands that are slow to rotate within the membrane, such as

fluocinolone acetonide. As a next step, Alchembed-based window generation with randomised ligand orientation should be tested as a more direct comparison to sMD-based configurations.

The permeation of fipronil through fenestrations on the glycine receptor has been demonstrated for the first time in this thesis. While already an important result, the strength of the claim would be increased by fully converged potentials of mean force (PMF). To this end, more simulation data has to be collected. In the investigation of glycine receptor fenestrations in this thesis, I considered membrane permeation and membrane-to-channel inhibition (Fig. 4.4) to be the most likely access mechanisms and therefore only calculated PMF profiles along these two pathways (Fig. 7.1). The PMF of the membrane permeation path is as yet incomplete. While the energetics of the actual membrane permeation are now accounted for, simulations of fipronil entering its binding site in the channel pore from the intracellular aqueous compartment are still required. The resulting, complete PMF profile would fully describe the membrane permeation path (red arrow in Fig. 4.4), and would give access to cycle closure of the PMF as another convergence criterion. It should be noted, however, that as the energy barrier of membrane permeation by itself is already larger than that of fenestration permeation, additional findings here will not falsify our hypothesis that membrane-to-channel inhibition is the preferred pore access mechanism for fipronil at the glycine receptor.

A third pathway through which fipronil could reach its binding site is entering at the extracellular domain and permeation through the channel's ion-conducting pore. I did not investigate this possibility in this thesis as the electrophysiology data reported by Islam and Lynch makes it unlikely[90]. Due to its hydrophobic nature, fipronil is likely to quickly accumulate in the membrane. Fig. 4.2H shows that after a brief short-state application of fipronil, followed by wash-out, GlyR initially displays an uninhibited channel current. Inhibition then sets in rapidly, even though no ligand is left in the surrounding solution. This makes pore-based access from the extracellular domain unlikely. Additionally, the A288F mutation at the

fenestration site is unlikely to have an effect on pore-based access. Nevertheless, the available evidence does not rule out this access pathway. As such, further studies should calculate the PMF along this route, as well.

Appendices

A

Supplemental Information - Chapter 1

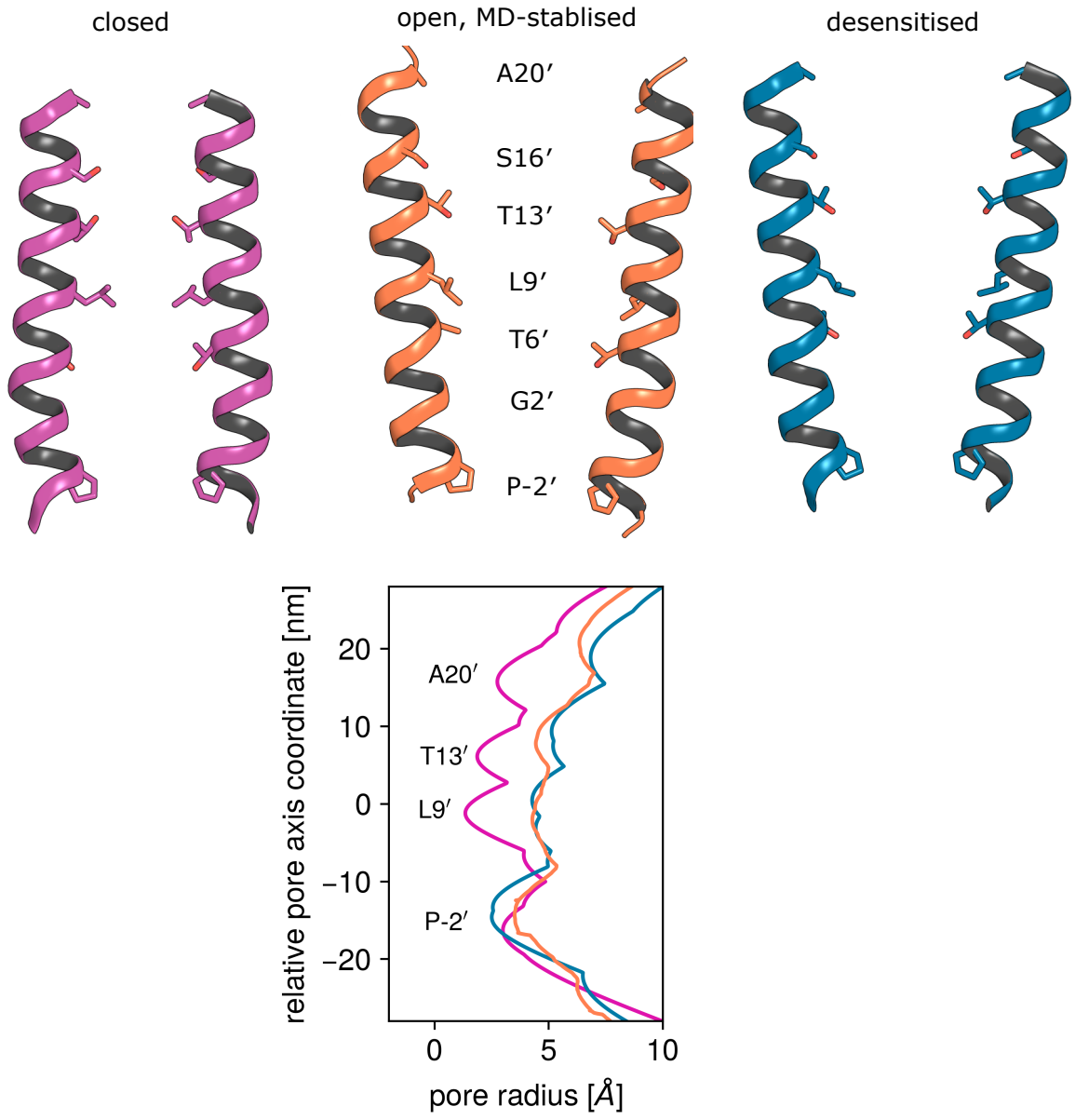

Figure A.1: Alternative version of Fig. 1.4. As the original open state structure (Du et al.[44], PDB 3JAE) is disputed, an MD-stabilised open state structure[59] is used here, instead. The key observations from the original figure about the M2 helix behaviour and gating at L9′ and P-2′ still hold.

B

Supplemental Information - Chapter 3

B.1 Convergence speed

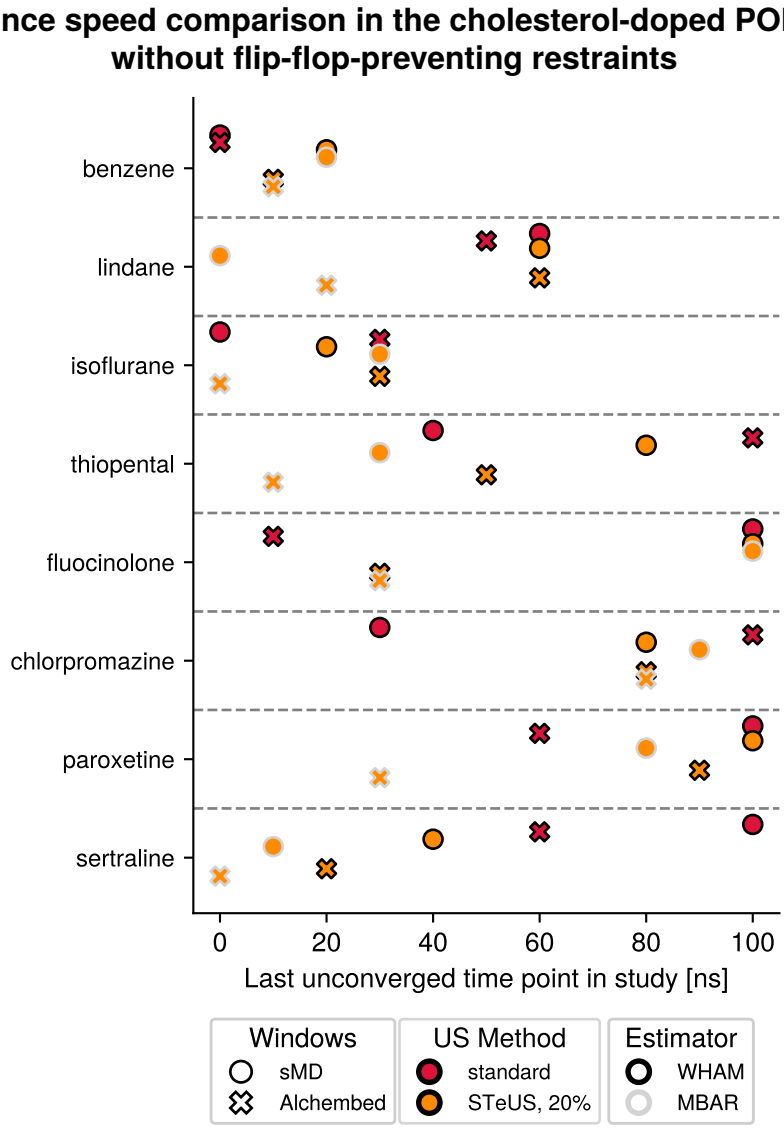

Figure B.1: Convergence speed comparison across all methods and compounds for data collected in the cholesterol-doped POPC bilayer without the application of flip-flop-preventing restraints.

B.2 Full, raw PMF convergence data

Full PMF convergence plots for all compounds and simulation conditions tested in this study are shown here. The same colour scale as in Fig. 3.17 is used throughout, with deep purple representing PMFs calculated from the first 10 ns of simulation data only and yellow representing the use of all 100 ns of data. Three membrane systems were investigated: POPC, POPC doped with 30 mol % cholesterol, and the cholesterol-doped bilayer with position restraints to prevent

cholesterol from flip-flopping. In each plot, 0 kcal mol^{-1} is indicated with a black line, and chemical accuracy of ±1 kcal mol^{-1} with dashed gray lines. The ten conditions are labelled A-J and were as follows:

- A: steered MD-based windows, standard US
- B: Alchembed-based windows, standard US
- C: steered MD-based windows, STeUS, 20 % ground state
- D: steered MD-based windows, STeUS, 20 % ground state, MBAR
- E: steered MD-based windows, STeUS, 40 % ground state
- F: steered MD-based windows, STeUS, 40 % ground state, MBAR
- G: Alchembed-based windows, STeUS, 20 % ground state
- H: Alchembed-based windows, STeUS, 20 % ground state, MBAR
- I: Alchembed-based windows, STeUS, 40 % ground state
- J: Alchembed-based windows, STeUS, 40 % ground state, MBAR

The cholesterol flip-flopping in the unrestrained cholesterol-doped POPC bilayer was noticed before STeUS simulations with 40 % ground state occupancy were started. As this bilayer was then considered incorrect, the new STeUS data was not collected, resulting in empty plots for conditions E, F, I, and J.

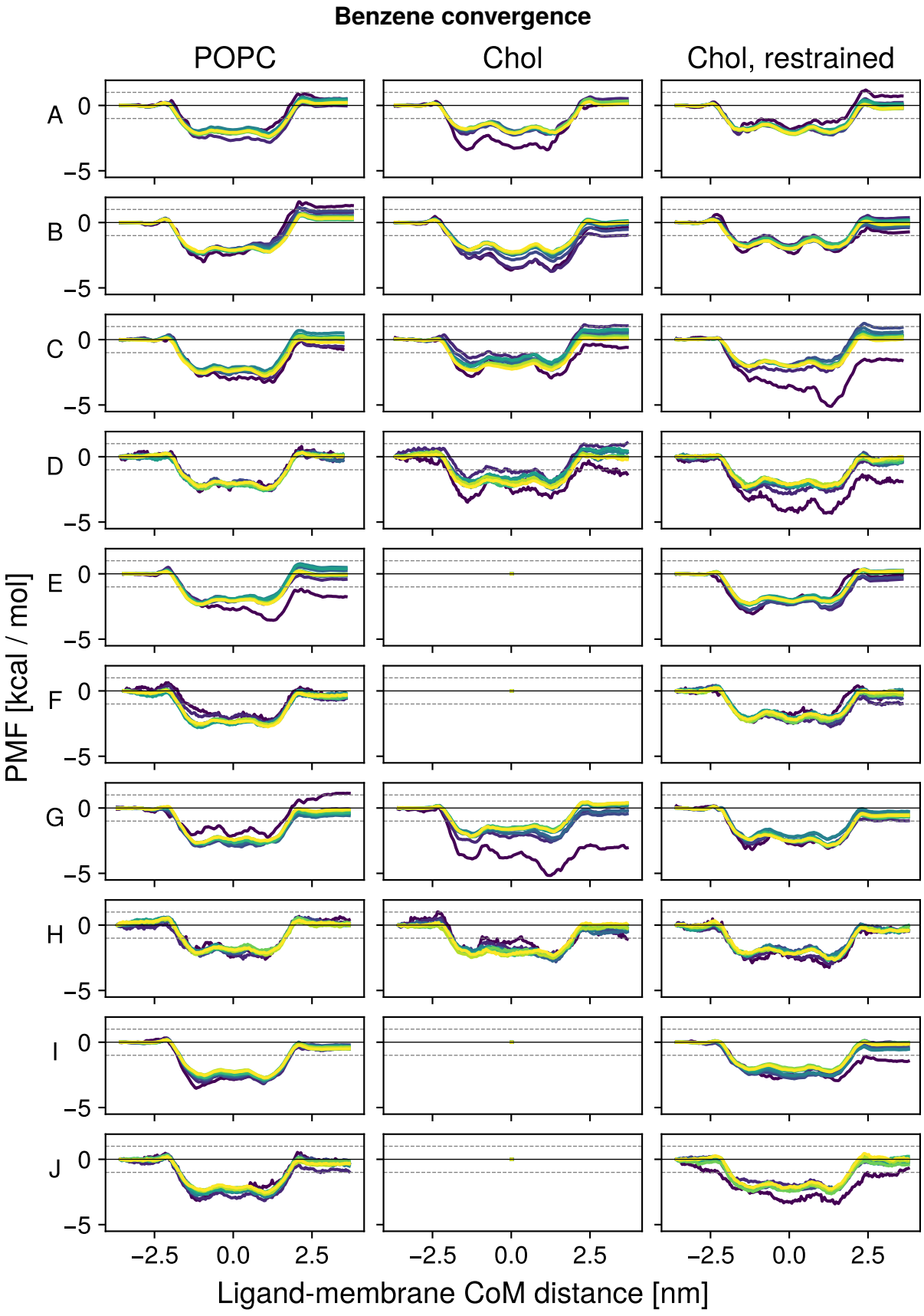

Figure B.2: Complete convergence data for benzene collected in this study.

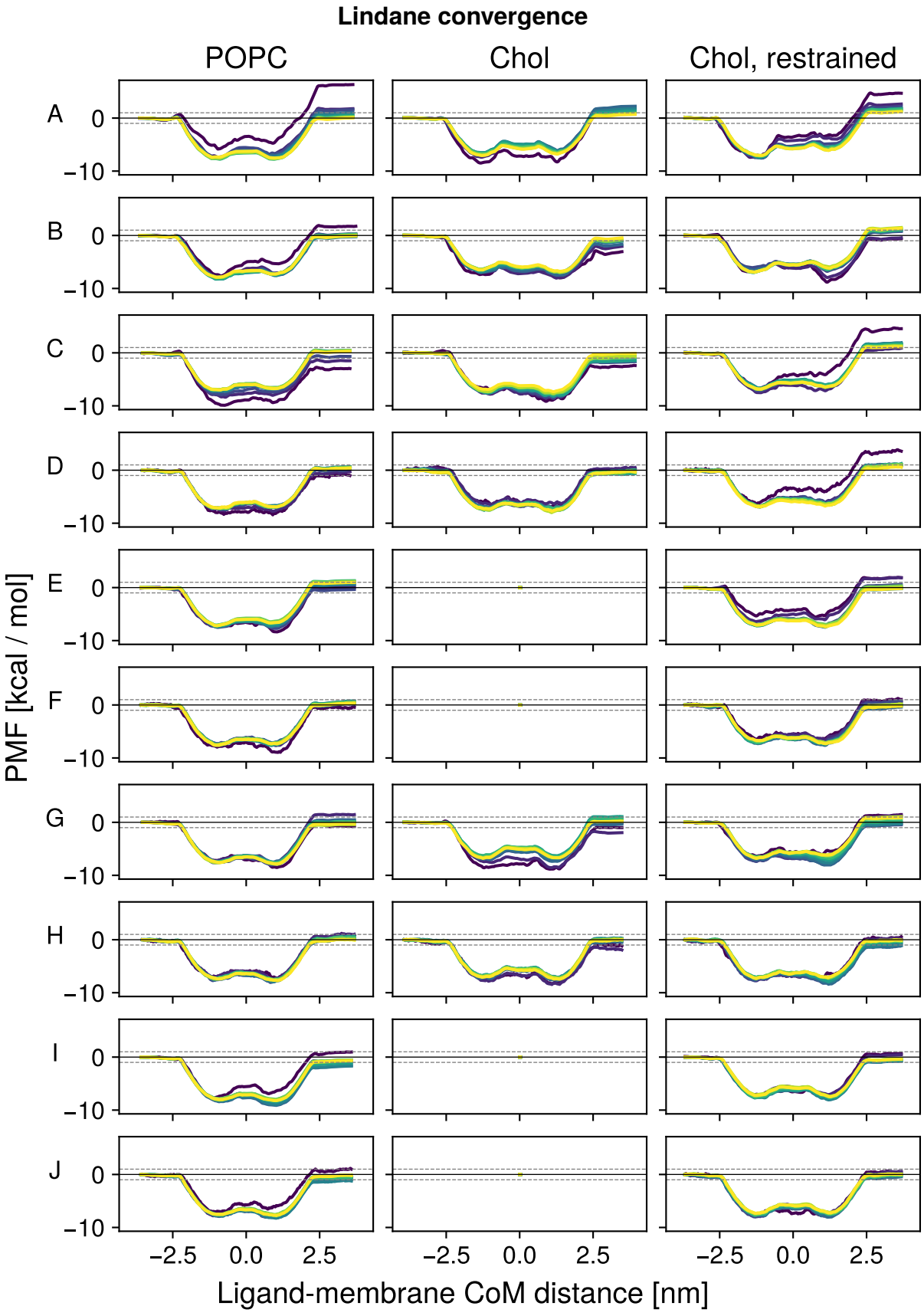

Figure B.3: Complete convergence data for lindane collected in this study.

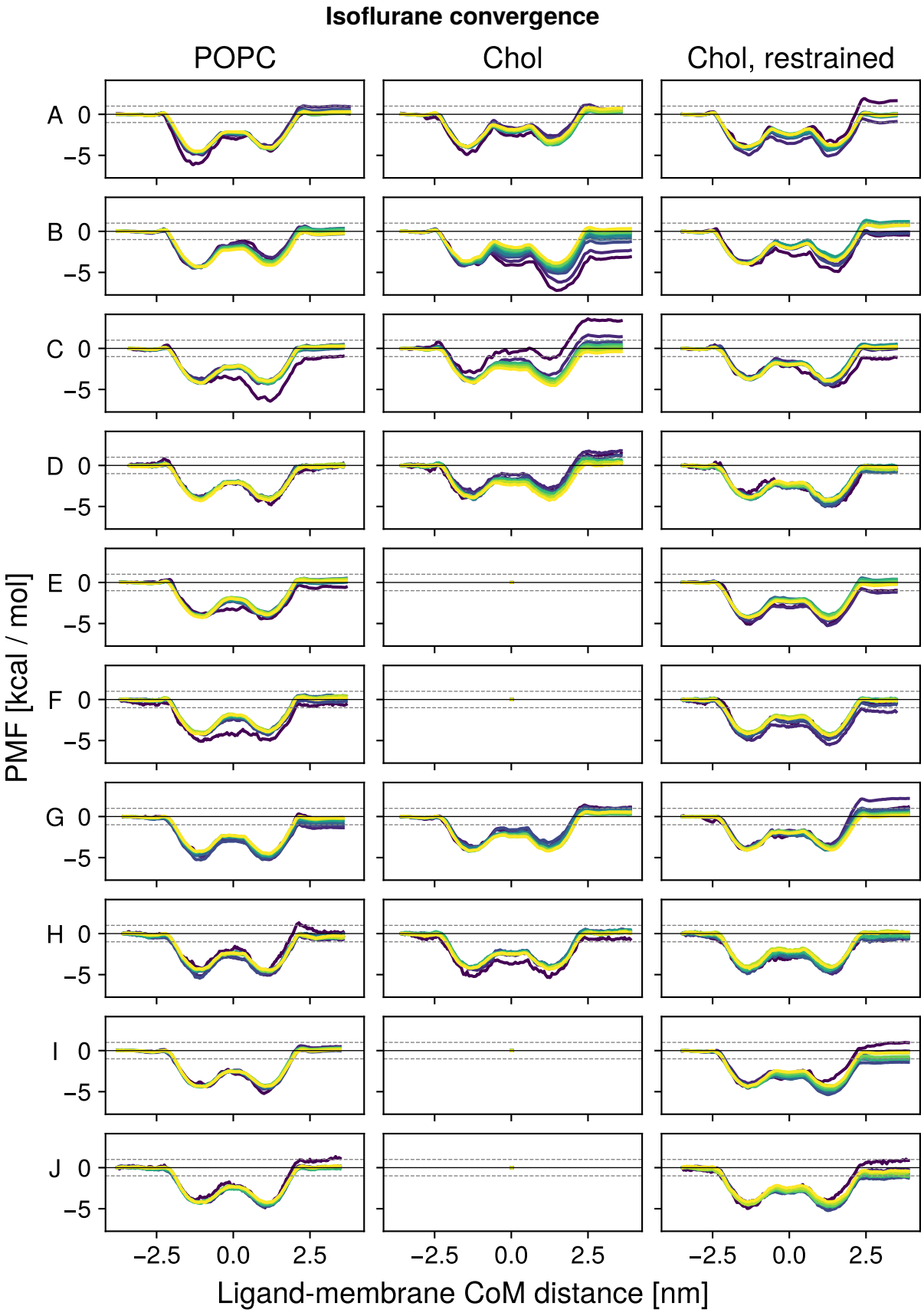

Figure B.4: Complete convergence data for isoflurane collected in this study.

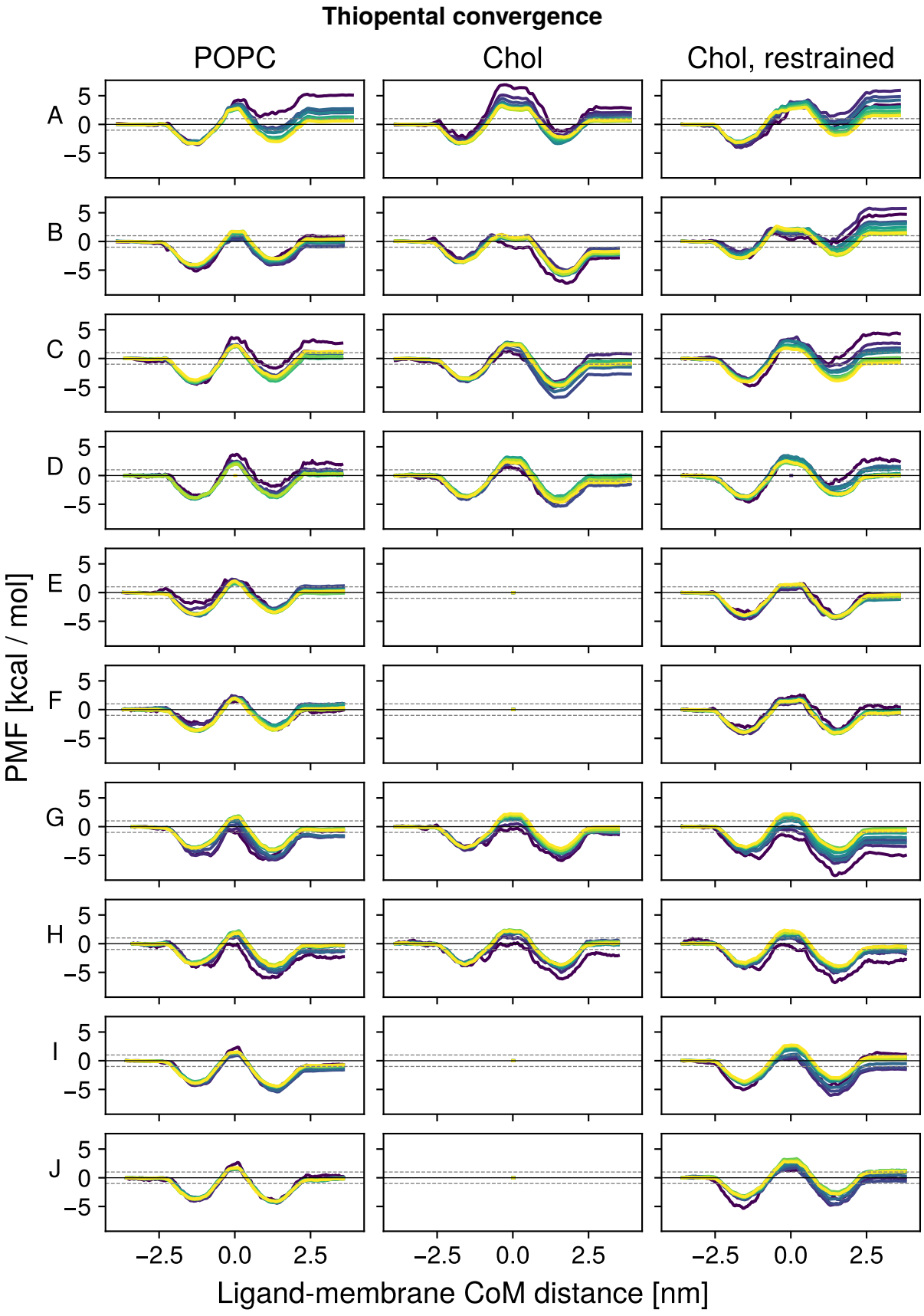

Figure B.5: Complete convergence data for thiopental collected in this study.

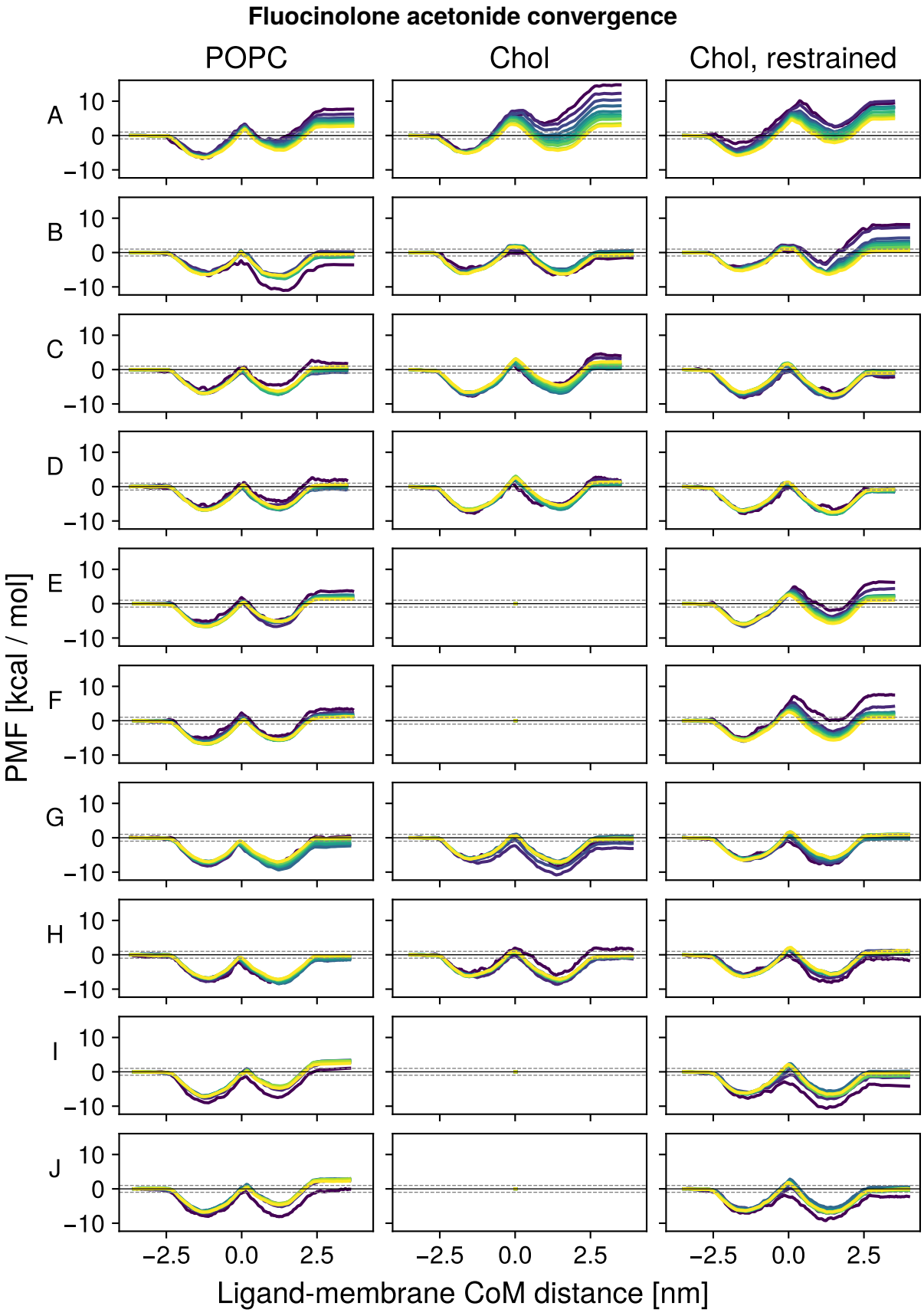

Figure B.6: Complete convergence data for fluocinolone acetonide collected in this study.

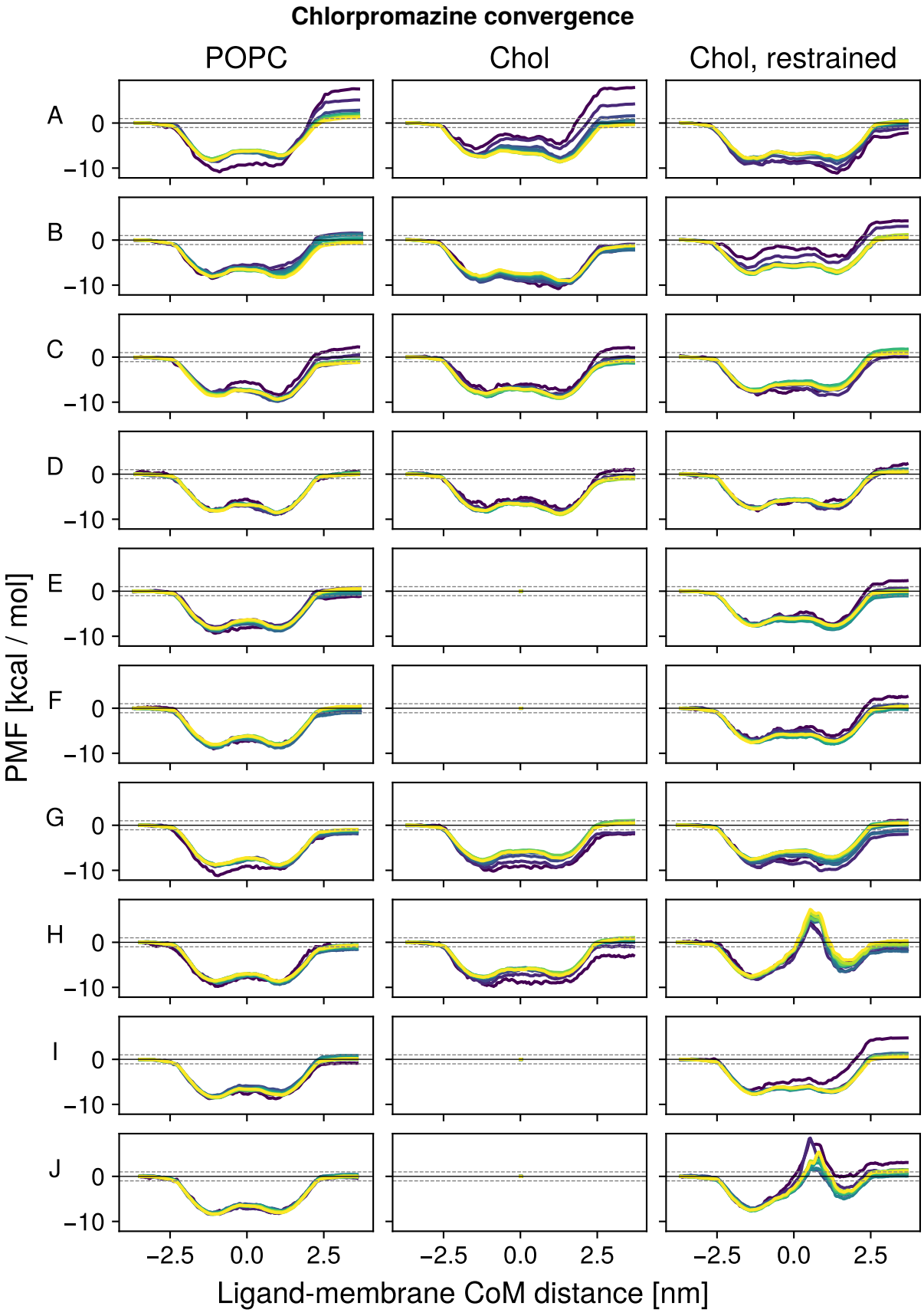

Figure B.7: Complete convergence data for chlorpromazine collected in this study.

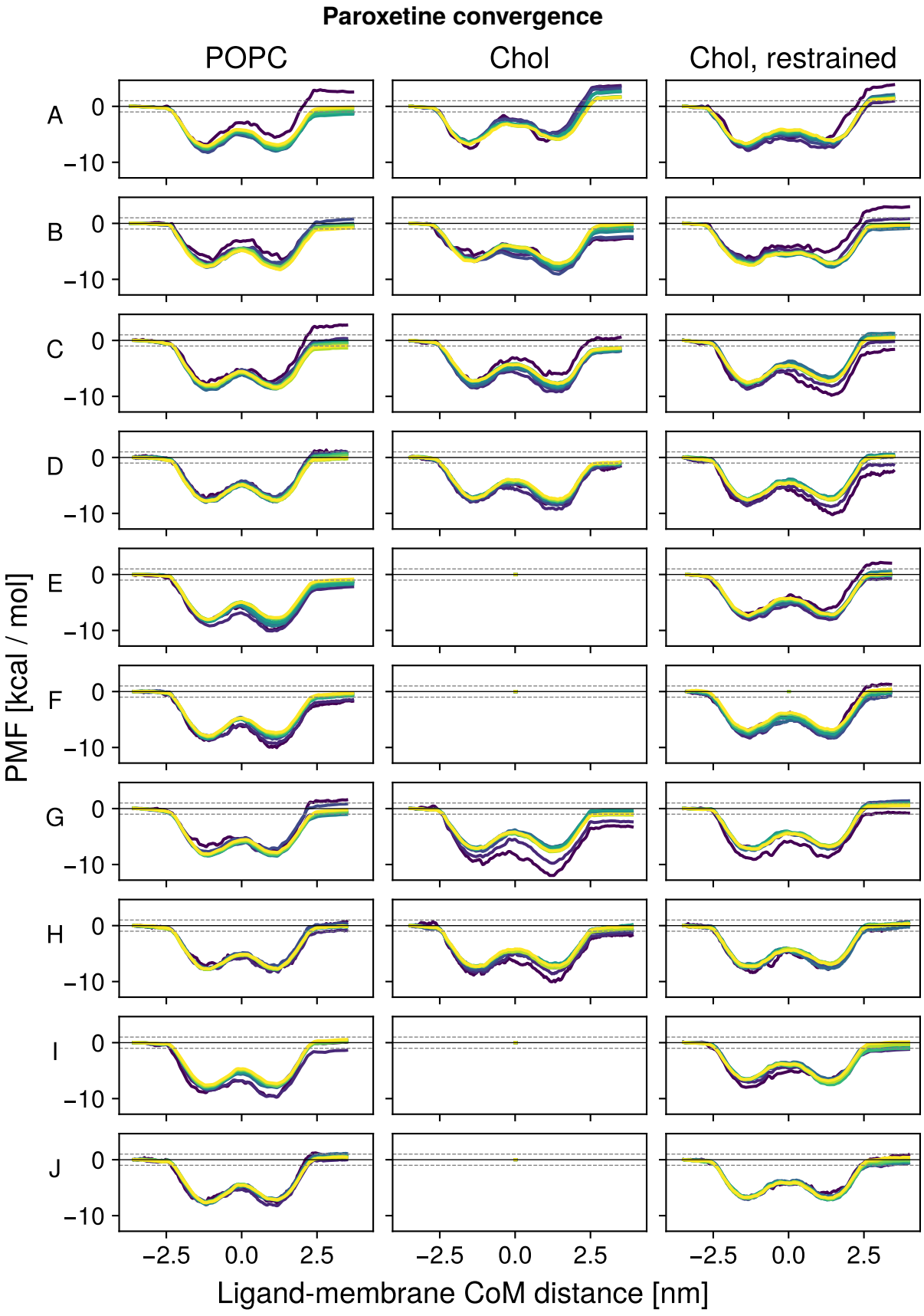

Figure B.8: Complete convergence data for paroxetine collected in this study.

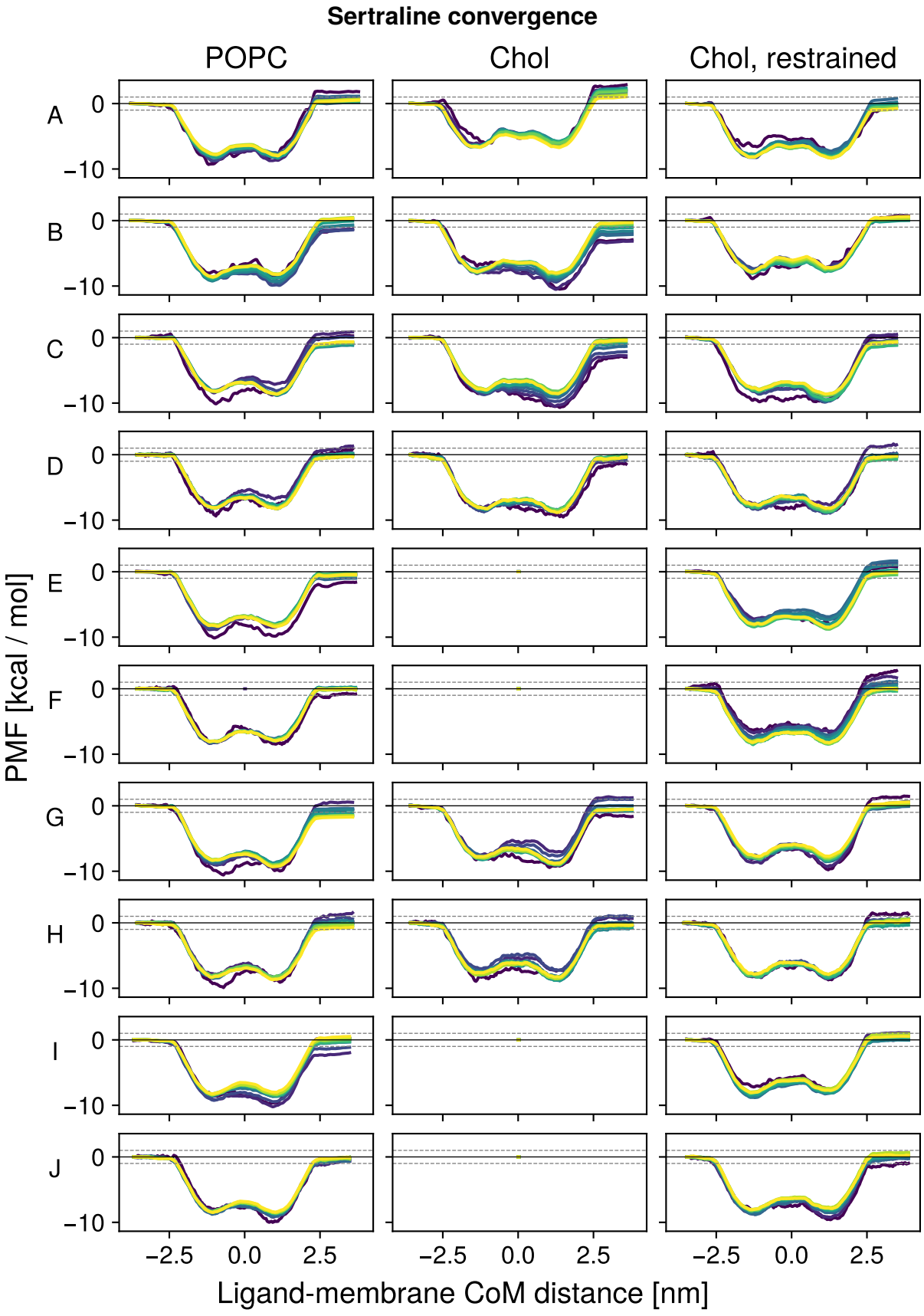

Figure B.9: Complete convergence data for sertraline collected in this study.

C

Supplemental Information - Chapter 4

Average fenestration bottleneck radii

Table C.1: Average bottleneck radii (BR) in Å as determined with CAVER for each of 55 structures, broken down by functional state.

index	PDB	state	BR	index	PDB	state	BR
1	3jad	closed	0.00	29	6pm6	open	1.72
2	3jae	open	1.72	30	6pxd	closed	0.00
3	3jaf	desensitized	1.50	31	6ubs	closed	0.00
4	5bkf	desensitized	1.38	32	6ubt	desensitized	1.66
5	5bkg	open	1.24	33	6ud3	open	1.45
6	5cfb	closed	0.24	34	6vm0	desensitized	1.73
7	5tin	desensitized	1.57	35	6vm2	desensitized	1.73
8	5tio	desensitized	1.31	36	6vm3	desensitized	1.83
9	5vdh	desensitized	1.94	37	7kuy	closed	0.24
10	5vdi	desensitized	1.97	38	7l31	closed	0.00
11	6plo	open	1.43	39	7m6m	closed	0.00
12	6plp	desensitized	1.47	40	7m6n	desensitized	1.64
13	6plq	wide-open	1.47	41	7m6o	open	2.05
14	6plr	desensitized	1.60	42	7m6p	desensitized	1.90
15	6pls	desensitized	1.53	43	7m6q	desensitized	1.88
16	6plt	closed	0.00	44	7m6r	desensitized	2.01
17	6plu	desensitized	1.56	45	7m6s	open	1.65
18	6plv	closed	0.00	46	7tu9	closed	0.26
19	6plw	wide-open	1.97	47	7tvi	desensitized	1.47
20	6plx	desensitized	1.31	48	7u2m	desensitized	1.41
21	6ply	open	1.51	49	7u2n	open	1.79
22	6plz	closed	0.00	50	7u2o	wide-open	1.82
23	6pm0	wide-open	1.76	51	8dn2	wide-open	1.75
24	6pm1	desensitized	1.50	52	8dn3	closed	0.59
25	6pm2	open	1.71	53	8dn4	desensitized	1.38
26	6pm3	closed	0.00	54	8dn5	open	1.21
27	6pm4	wide-open	1.56	55	8fe1	desensitized	1.89
28	6pm5	desensitized	1.27				

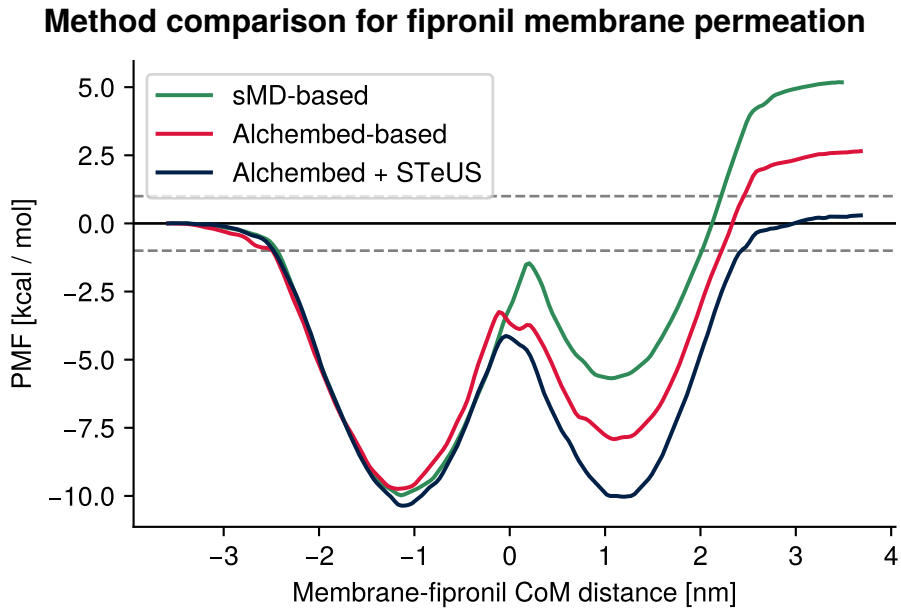

Figure C.1: PMFs of fipronil going through a POPC bilayer obtained with different methods. For each method, all windows were simulated for 100 ns. Umbrella sampling with windows extracted from a steered-MD trajectory and run without STeUS results in strong asymmetries (green line). Swapping steered MD-based for Alchembed-based window generation results in improvements (red line), but only the combination of STeUS and Alchembed-based window generation gives a symmetrical and well-converged PMF profile (blue line, bulk value closer than chemical accuracy of 1 kcal/mol, dashed grey lines).

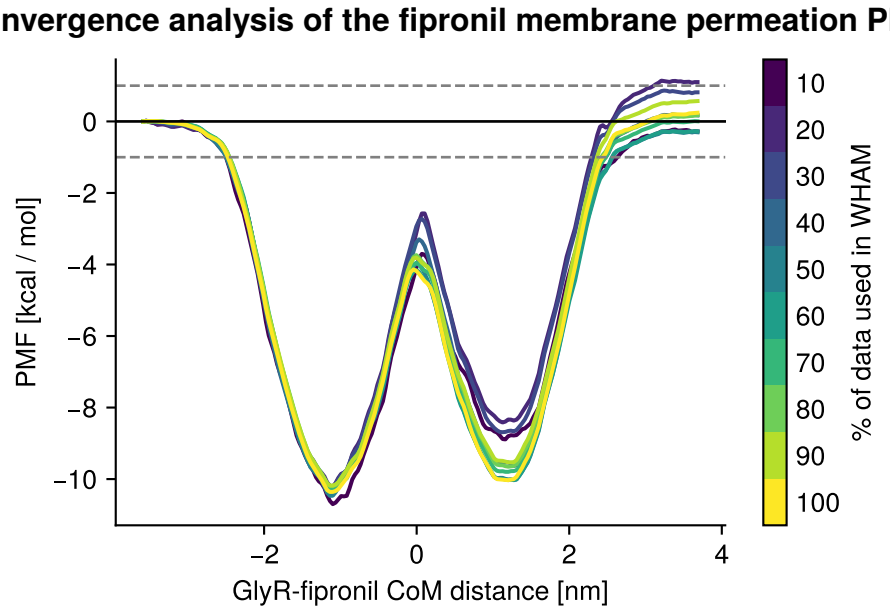

Figure C.2: Convergence analysis of the PMF of fipronil going through a POPC bilayer. The PMF is well converged. This can be seen through the symmetry of the profiles and through the value in bulk water sampled at positive reaction coordinates being very close to zero — closer than chemical accuracy (1 kcal/mol, dashed grey lines).

D

Supplemental Information - Chapter 5

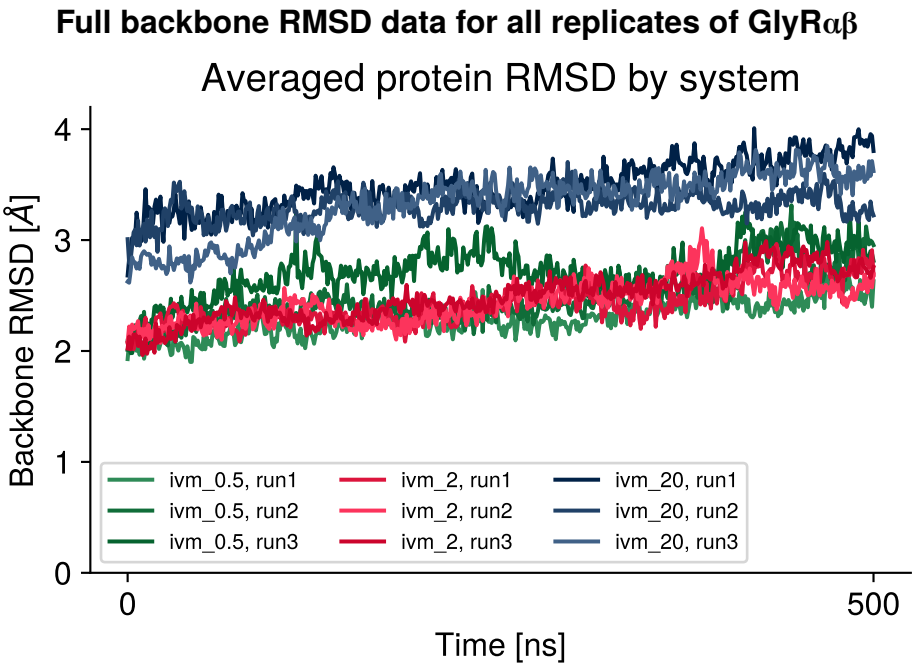

Figure D.1: Protein backbone RMSD data of all runs. Average values were reported in Fig. 5.8

E

Supplemental Information - Chapter 6

E.1 GLIC permeation plots with original protonation state

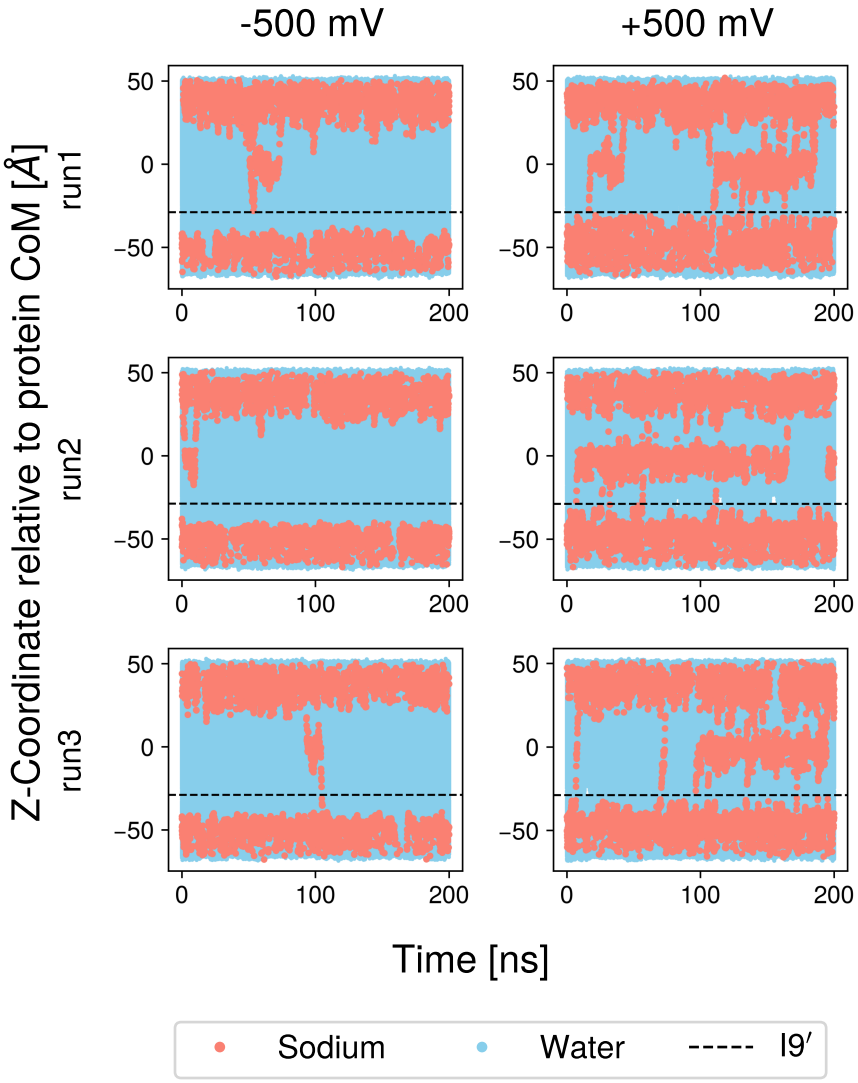

Figure E.1: Analysis of ion permeation events in *iiooo*. The z-coordinate of sodium ions and water molecules in the pore is plotted over time to visualise conduction events. More conduction events occur with a simulated positive membrane potential.

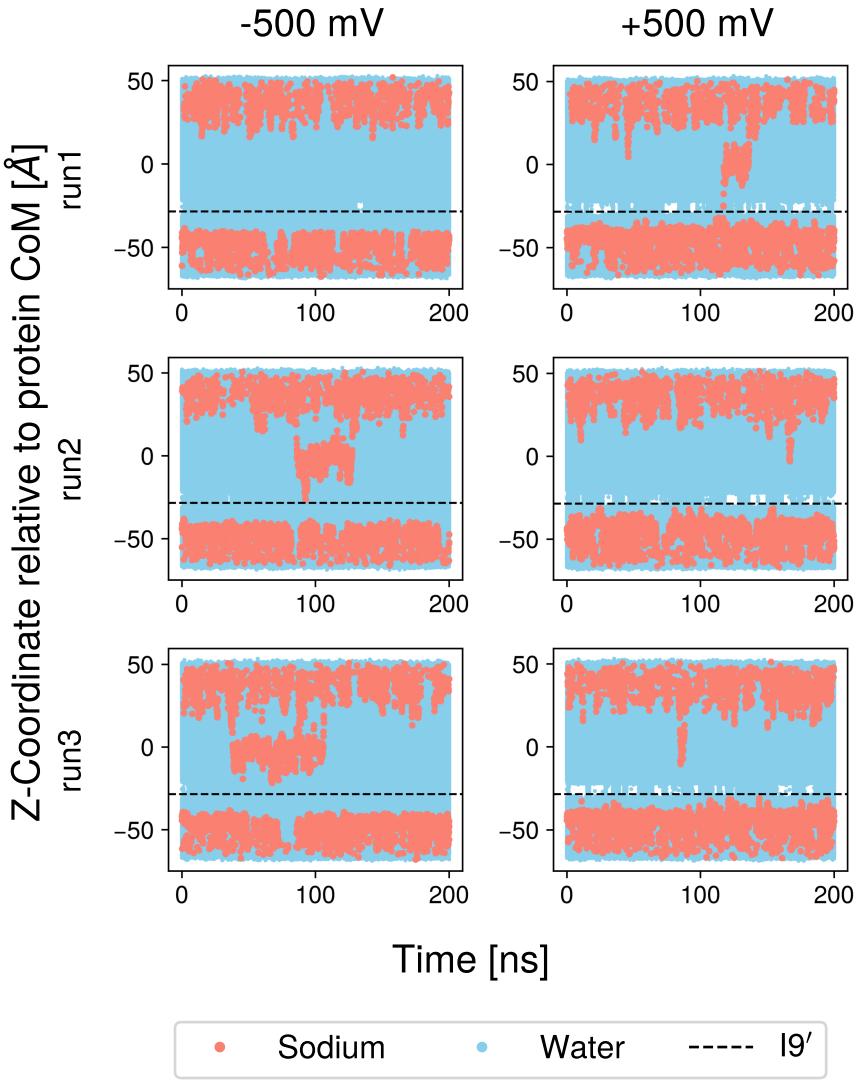

Figure E.2: Analysis of ion permeation events in *ioioo*. The z-coordinate of sodium ions and water molecules in the pore is plotted over time to visualise conduction events. More conduction events occur with a simulated positive membrane potential.

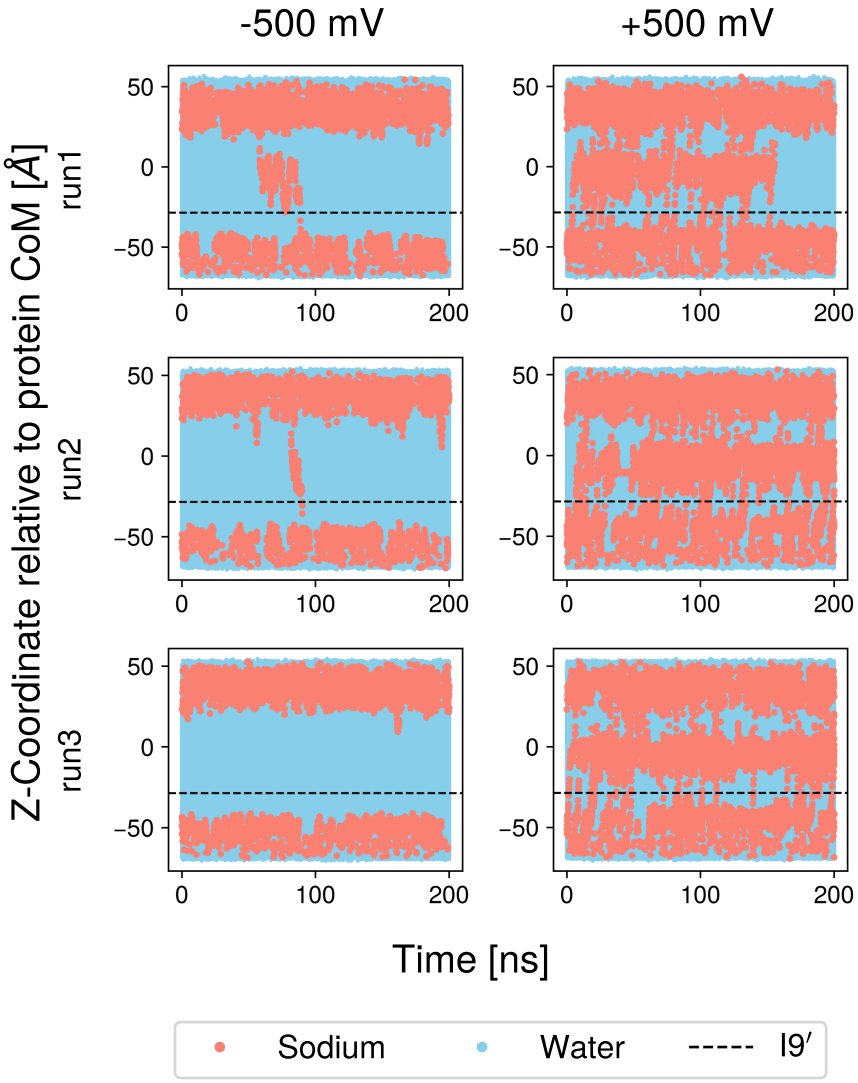

Figure E.3: Analysis of ion permeation events in *iooo*. The z-coordinate of sodium ions and water molecules in the pore is plotted over time to visualise conduction events. More conduction events occur with a simulated positive membrane potential.

E.2 GLIC permeation plots with new protonation state

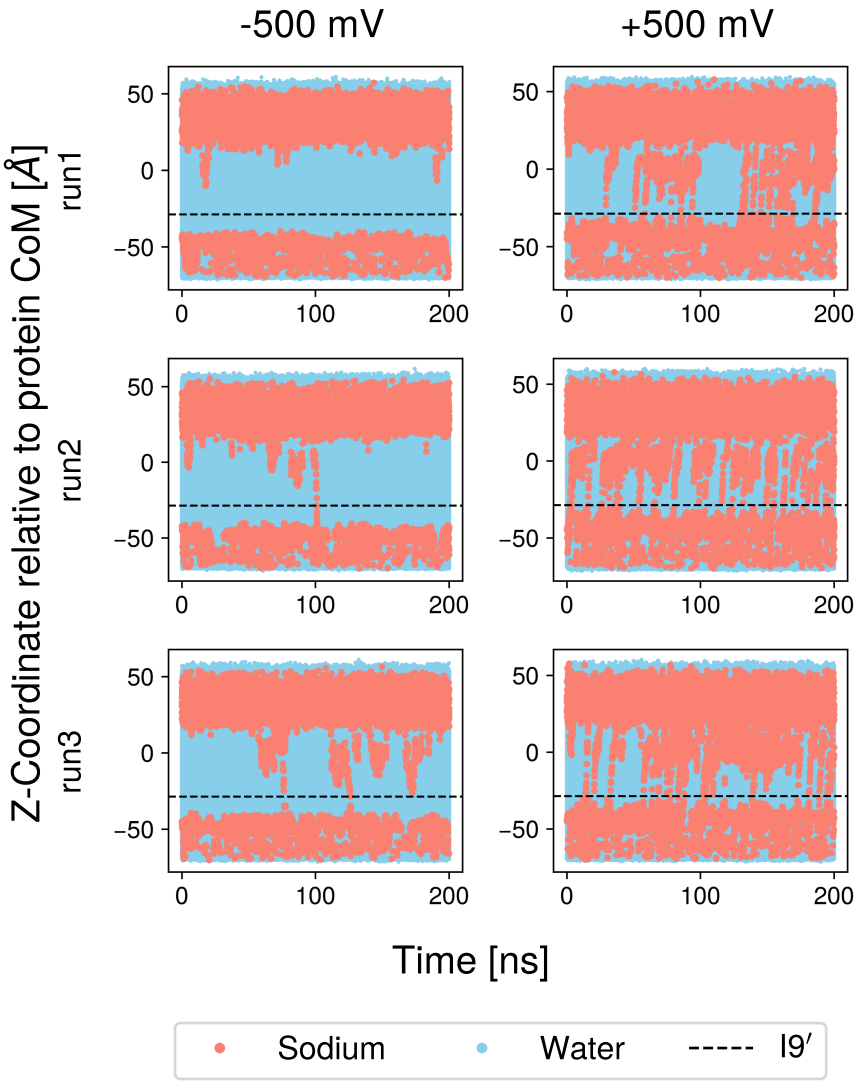

Figure E.4: Permeation plots for *iiooo* with the new protonation states. The rectification behaviour is still clearly visible. In these runs, *iiooo* is more conductive than in runs using the old protonation states (Fig. E.1).

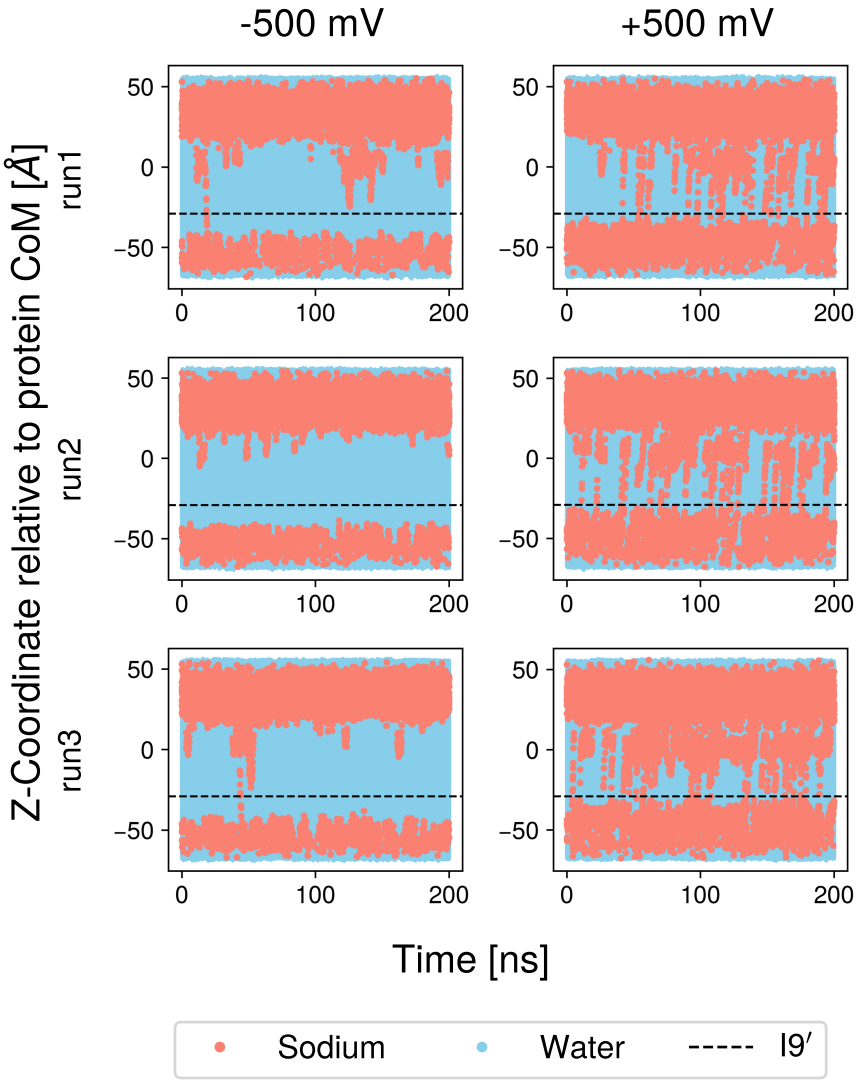

Figure E.5: Permeation plots for *iooоо* with the new protonation states. The rectification behaviour is still clearly visible. The channel conductance is similar to that determined with the old protonation states (Fig. E.3).

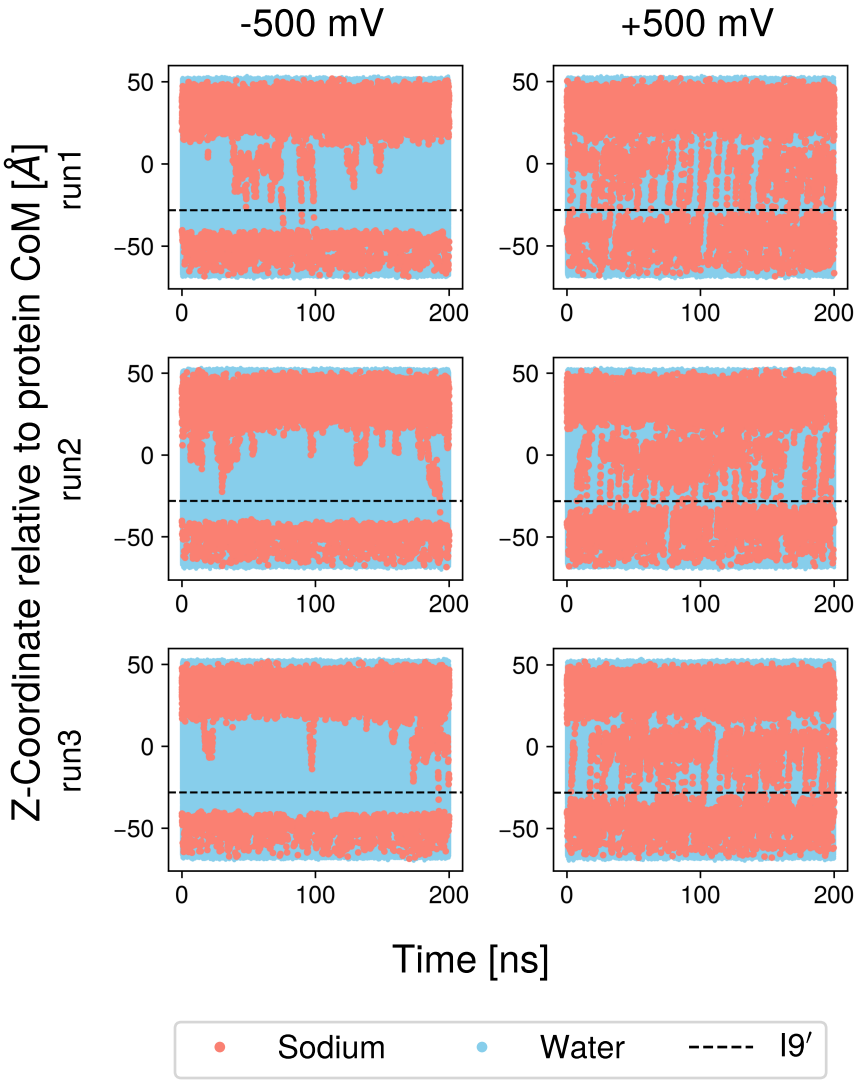

Figure E.6: Permeation data for an all-open *ooooo* state was collected to compare to the partially open *ioooo*. Also strongly rectifying, it is much more conductive than any other tested system under positive membrane potentials.

References

1. Schrödinger, E. *What is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, 1944).
2. Von Bertalanffy, L. The Theory of Open Systems in Physics and Biology. *Science* **111,** 23–29 (1950).
3. Alberts, B. *Molecular biology of the cell* 6th ed (Garland science, Taylor and Francis group, New York, 2015).
4. Hille, B. *Ion channels of excitable membranes* 3rd ed (Sinauer, Sunderland, Mass, 2001).
5. Chr. Skou, J. The identification of the sodium-pump as the membrane-boundNa+/K+-ATPase: a commentary The influence of some cations on an adenosine triphosphatase from peripheral nerves on 'The influence of some cations on an adenosine triphosphatase from peripheral nerves' by J.C. Skou Biochim. Biophys. Acta 23 (1957) 394–401. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1000,** 435–446 (1989).
6. Alexander, S. P. H., Mathie, A., Peters, J. A., Veale, E. L., Striessnig, J., Kelly, E., Armstrong, J. F., Faccenda, E., Harding, S. D., Pawson, A. J., Sharman, J. L., Southan, C., Davies, J. A. & CGTP Collaborators. THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: Ion channels. *British Journal of Pharmacology* **176 Suppl 1,** S142–S228 (Suppl 1 2019).
7. Kefauver, J. M., Ward, A. B. & Patapoutian, A. Discoveries in structure and physiology of mechanically activated ion channels. *Nature* **587,** 567–576 (2020).
8. Huang, J., Pan, X. & Yan, N. Structural biology and molecular pharmacology of voltage-gated ion channels. *Nature Reviews Molecular Cell Biology* **25,** 904–925 (2024).
9. Clapham, D. E. Calcium Signaling. *Cell* **131,** 1047–1058 (2007).
10. Clapham, D. E. TRP channels as cellular sensors. *Nature* **426,** 517–524 (2003).
11. Strange, K., Yamada, T. & Denton, J. S. A 30-year journey from volume-regulated anion currents to molecular structure of the LRRC8 channel. *Journal of General Physiology* **151,** 100–117 (2019).
12. Pedersen, S. F., Hoffmann, E. K. & Novak, I. Cell volume regulation in epithelial physiology and cancer. *Frontiers in Physiology* **4** (2013).
13. Wu, X., Li, Y., Maienschein-Cline, M., Feferman, L., Wu, L. & Hong, L. RNA-Seq Analyses Reveal Roles of the HVCN1 Proton Channel in Cardiac pH Homeostasis. *Frontiers in Cell and Developmental Biology* **10** (2022).

14. Taylor, A. R., Chrachri, A., Wheeler, G., Goddard, H. & Brownlee, C. A Voltage-Gated H+ Channel Underlying pH Homeostasis in Calcifying Coccolithophores. *PLOS Biology* **9,** e1001085 (2011).

15. Burotto, M., Chiou, V. L., Lee, J.-M. & Kohn, E. C. The MAPK pathway across different malignancies: A new perspective. *Cancer* **120,** 3446–3456 (2014).

16. Hager, G. L., Nagaich, A. K., Johnson, T. A., Walker, D. A. & John, S. Dynamics of nuclear receptor movement and transcription. *Biochimica Et Biophysica Acta* **1677,** 46–51 (2004).

17. Robertson, N. M., Schulman, G., Karnik, S., Alnemri, E. & Litwack, G. Demonstration of nuclear translocation of the mineralocorticoid receptor (MR) using an anti-MR antibody and confocal laser scanning microscopy. *Molecular Endocrinology* **7,** 1226–1239 (1993).

18. Thompson, A. J., Lester, H. A. & Lummis, S. C. R. The structural basis of function in Cys-loop receptors. *Quarterly Reviews of Biophysics* **43,** 449–499 (2010).

19. Hiroyoshi, M. *Encyclopedia of Neuroscience* 3952–3956 (Springer, Berlin, Heidelberg, 2009).

20. Karlin, A. & Akabas, M. H. Toward a structural basis for the function of nicotinic acetylcholine receptors and their cousins. *Neuron* **15,** 1231–1244 (1995).

21. Yakel, J. L., Shao, X. M. & Jackson, M. B. The selectivity of the channel coupled to the 5-HT3 receptor. *Brain Research* **533,** 46–52 (1990).

22. Fatima-Shad, K. & Barry, P. H. Anion permeation in GABA- and glycine-gated channels of mammalian cultured hippocampal neurons. *Proceedings. Biological Sciences* **253,** 69–75 (1993).

23. Keramidas, A., Moorhouse, A. J., Pierce, K. D., Schofield, P. R. & Barry, P. H. Cation-selective Mutations in the M2 Domain of the Inhibitory Glycine Receptor Channel Reveal Determinants of Ion-Charge Selectivity. *Journal of General Physiology* **119,** 393–410 (2002).

24. Cymes, G. D. & Grosman, C. Tunable pKa values and the basis of opposite charge selectivities in nicotinic-type receptors. *Nature* **474,** 526–530 (2011).

25. Unwin, N. Refined Structure of the Nicotinic Acetylcholine Receptor at 4 Å Resolution. *Journal of Molecular Biology* **346,** 967–989 (2005).

26. Hilf, R. J. C. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452,** 375–379 (2008).

27. Hilf, R. J. C. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457,** 115–118 (2009).

28. Cheng, M. H., Cascio, M. & Coalson, R. D. Homology modeling and molecular dynamics simulations of the alpha1 glycine receptor reveals different states of the channel. *Proteins* **68,** 581–593 (2007).

29. Murail, S., Wallner, B., Trudell, J. R., Bertaccini, E. & Lindahl, E. Microsecond Simulations Indicate that Ethanol Binds between Subunits and Could Stabilize an Open-State Model of a Glycine Receptor. *Biophysical Journal* **100,** 1642–1650 (2011).

30. Unwin, N. Nicotinic acetylcholine receptor and the structural basis of neuromuscular transmission: insights from Torpedo postsynaptic membranes. *Quarterly Reviews of Biophysics* **46,** 283–322 (2013).

31. Walstab, J., Rappold, G. & Niesler, B. 5-HT3 receptors: Role in disease and target of drugs. *Pharmacology & Therapeutics* **128,** 146–169 (2010).

32. Jiang, J., Li, X., Hu, A.-f., Zhou, G.-j., Gao, Y.-h., Xu, C., Wu, X.-m. & Wang, H.-J. Nicotine and neuronal nicotinic acetylcholine receptors: unraveling the mechanisms of nicotine addiction. *Frontiers in Neuroscience* **19** (2025).

33. Farrant, M. & Nusser, Z. Variations on an inhibitory theme: phasic and tonic activation of GABAA receptors. *Nature Reviews Neuroscience* **6,** 215–229 (2005).

34. Rudolph, U. & Möhler, H. Analysis of GABAA Receptor Function and Dissection of the Pharmacology of Benzodiazepines and General Anesthetics Through Mouse Genetics. *Annual Review of Pharmacology and Toxicology* **44,** 475–498 (Volume 44, 2004 2004).

35. Maillard, P.-Y., Baer, S., Schaefer, É., Desnous, B., Villeneuve, N., Lépine, A., Fabre, A., Lacoste, C., El Chehadeh, S., Piton, A., Porter, L. F., Perriard, C., Wardé, M.-T. A., Spitz, M.-A., Laugel, V., Lesca, G., Putoux, A., Ville, D., Mignot, C., Héron, D., Nabbout, R., Barcia, G., Rio, M., Roubertie, A., Meyer, P., Paquis-Flucklinger, V., Patat, O., Lefranc, J., Gerard, M., Consortium, E., de Bellescize, J., Villard, L., De Saint Martin, A. & Milh, M. Molecular and clinical descriptions of patients with GABAA receptor gene variants (GABRA1, GABRB2, GABRB3, GABRG2): A cohort study, review of literature, and genotype–phenotype correlation. *Epilepsia* **63,** 2519–2533 (2022).

36. Schaaf, C. P. Nicotinic acetylcholine receptors in human genetic disease. *Genetics in Medicine* **16,** 649–656 (2014).

37. Chung, S.-K., Vanbellinghen, J.-F., Mullins, J. G. L., Robinson, A., Hantke, J., Hammond, C. L., Gilbert, D. F., Freilinger, M., Ryan, M., Kruer, M. C., Masri, A., Gurses, C., Ferrie, C., Harvey, K., Shiang, R., Christodoulou, J., Andermann, F., Andermann, E., Thomas, R. H., Harvey, R. J., Lynch, J. W. & Rees, M. I. Pathophysiological Mechanisms of Dominant and Recessive GLRA1 Mutations in Hyperekplexia. *Journal of Neuroscience* **30,** 9612–9620 (2010).

38. Curtis, D. R., Hösli, L. & Johnston, G. a. R. Inhibition of Spinal Neurones by Glycine. *Nature* **215,** 1502–1503 (1967).

39. Lynch, J. W. Molecular Structure and Function of the Glycine Receptor Chloride Channel. *Physiological Reviews* (2004).

40. Mizzi, N. & Blundell, R. Glycine receptors: Structure, function, and therapeutic implications. *Molecular Aspects of Medicine* **103,** 101360 (2025).

41. Werman, R., Davidoff, R. A. & Aprison, M. H. Is Glycine a Neurotransmitter ?: Inhibition of Motoneurones by Iontophoresis of Glycine. *Nature* **214,** 681–683 (1967).

42. Pfeiffer, F., Graham, D. & Betz, H. Purification by affinity chromatography of the glycine receptor of rat spinal cord. *Journal of Biological Chemistry* **257,** 9389–9393 (1982).

43. Grenningloh, G., Rienitz, A., Schmitt, B., Methfessel, C., Zensen, M., Beyreuther, K., Gundelfinger, E. D. & Betz, H. The strychnine-binding subunit of the glycine receptor shows homology with nicotinic acetylcholine receptors. *Nature* **328,** 215–220 (1987).

44. Du, J., Lü, W., Wu, S., Cheng, Y. & Gouaux, E. Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526,** 224–229 (2015).

45. Huang, X., Chen, H., Michelsen, K., Schneider, S. & Shaffer, P. L. Crystal structure of human glycine receptor-α3 bound to antagonist strychnine. *Nature* **526,** 277–280 (2015).

46. Lester, H. A., Dibas, M. I., Dahan, D. S., Leite, J. F. & Dougherty, D. A. Cys-loop receptors: new twists and turns. *Trends in Neurosciences* **27,** 329–336 (2004).

47. Langlhofer, G. & Villmann, C. The Intracellular Loop of the Glycine Receptor: It's not all about the Size. *Frontiers in Molecular Neuroscience* **9** (2016).

48. Lynagh, T. & Pless, S. A. Principles of agonist recognition in Cys-loop receptors. *Frontiers in Physiology* **5** (2014).

49. Pless, S. A. & Lynch, J. W. Ligand-specific Conformational Changes in the α1 Glycine Receptor Ligand-binding Domain *. *Journal of Biological Chemistry* **284,** 15847–15856 (2009).

50. Borghese, C. M., Galpin, J. D., Eriksson Lidbrink, S., Zhuang, Y., Desai, N. G., Howard, R. J., Lindahl, E., Ahern, C. A. & Goldschen-Ohm, M. P. A single main-chain hydrogen bond required to keep GABAA receptors closed. *Nature Communications* **16,** 6107 (2025).

51. Grosman, C., Zhou, M. & Auerbach, A. Mapping the conformational wave of acetylcholine receptor channel gating. *Nature* **403,** 773–776 (2000).

52. Lee, W. Y. & Sine, S. M. Principal pathway coupling agonist binding to channel gating in nicotinic receptors. *Nature* **438,** 243–247 (2005).

53. Calimet, N., Simoes, M., Changeux, J.-P., Karplus, M., Taly, A. & Cecchini, M. A gating mechanism of pentameric ligand-gated ion channels. *Proceedings of the National Academy of Sciences* **110,** E3987–E3996 (2013).

54. Grewer, C. Investigation of the α1-Glycine Receptor Channel-Opening Kinetics in the Submillisecond Time Domain. *Biophysical Journal* **77,** 727–738 (1999).

55. Charnet, P., Labarca, C., Leonard, R. J., Vogelaar, N. J., Czyzyk, L., Gouin, A., Davidson, N. & Lester, H. A. An open-channel blocker interacts with adjacent turns of α-helices in the nicotinic acetylcholine receptor. *Neuron* **4,** 87–95 (1990).

56. Kumar, A., Basak, S., Rao, S., Gicheru, Y., Mayer, M. L., Sansom, M. S. P. & Chakrapani, S. Mechanisms of activation and desensitization of full-length glycine receptor in lipid nanodiscs. *Nature Communications* **11,** 3752 (2020).

57. Labarca, C., Nowak, M. W., Zhang, H., Tang, L., Deshpande, P. & Lester, H. A. Channel gating governed symmetrically by conserved leucine residues in the M2 domain of nicotinic receptors. *Nature* **376,** 514–516 (1995).

58. Miyazawa, A., Fujiyoshi, Y. & Unwin, N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature* **423,** 949–955 (2003).

59. Dämgen, M. A. & Biggin, P. C. A Refined Open State of the Glycine Receptor Obtained via Molecular Dynamics Simulations. *Structure* **28,** 130–139.e2 (2020).
60. Gielen, M., Thomas, P. & Smart, T. G. The desensitization gate of inhibitory Cys-loop receptors. *Nature Communications* **6,** 6829 (2015).
61. Beckstein, O. & Sansom, M. S. P. A hydrophobic gate in an ion channel: the closed state of the nicotinic acetylcholine receptor. *Physical Biology* **3,** 147 (2006).
62. Beckstein, O. & Sansom, M. S. P. The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Physical Biology* **1,** 42 (2004).
63. Trick, J. L., Aryal, P., Tucker, S. J. & Sansom, M. S. P. Molecular simulation studies of hydrophobic gating in nanopores and ion channels. *Biochemical Society Transactions* **43,** 146–150 (2015).
64. Rao, S., Lynch, C. I., Klesse, G., Oakley, G. E., Stansfeld, P. J., Tucker, S. J. & Sansom, M. S. P. Water and hydrophobic gates in ion channels and nanopores. *Faraday Discussions* **209,** 231–247 (2018).
65. Legendre, P., Muller, E., Badiu, C. I., Meier, J., Vannier, C. & Triller, A. Desensitization of Homomeric α1 Glycine Receptor Increases with Receptor Density. *Molecular Pharmacology* **62,** 817–827 (2002).
66. Liu, X. & Wang, W. Gating mechanism of the human $\alpha1\beta$ GlyR by glycine. *Structure* **32,** 1621–1631.e3 (2024).
67. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics* **14,** 354–360 (1996).
68. Pedersen, J. E., Bergqvist, C. A. & Larhammar, D. Evolution of vertebrate nicotinic acetylcholine receptors. *BMC Evolutionary Biology* **19,** 38 (2019).
69. Millar, N. S. & Gotti, C. Diversity of vertebrate nicotinic acetylcholine receptors. *Neuropharmacology* **56,** 237–246 (2009).
70. Olsen, R. W. & Sieghart, W. GABAA receptors: Subtypes provide diversity of function and pharmacology. *Neuropharmacology* **56,** 141–148 (2009).
71. Becker, C. M., Hoch, W. & Betz, H. Glycine receptor heterogeneity in rat spinal cord during postnatal development. *The EMBO journal* **7,** 3717–3726 (1988).
72. Meyer, G., Kirsch, J., Betz, H. & Langosch, D. Identification of a gephyrin binding motif on the glycine receptor beta subunit. *Neuron* **15,** 563–572 (1995).
73. Handford, C. A., Lynch, J. W., Baker, E., Webb, G. C., Ford, J. H., Sutherland, G. R. & Schofield, P. R. The human glycine receptor beta subunit: primary structure, functional characterisation and chromosomal localisation of the human and murine genes. *Brain Research. Molecular Brain Research* **35,** 211–219 (1996).
74. Cerdan, A. H., Sisquellas, M., Pereira, G., Barreto Gomes, D. E., Changeux, J.-P. & Cecchini, M. The Glycine Receptor Allosteric Ligands Library (GRALL). *Bioinformatics* **36,** 3379–3384 (2020).
75. Huang, X., Chen, H. & Shaffer, P. L. Crystal Structures of Human GlyRα3 Bound to Ivermectin. *Structure* **25,** 945–950.e2 (2017).

76. Masiulis, S., Desai, R., Uchański, T., Serna Martin, I., Laverty, D., Karia, D., Malinauskas, T., Zivanov, J., Pardon, E., Kotecha, A., Steyaert, J., Miller, K. W. & Aricescu, A. R. GABAA receptor signalling mechanisms revealed by structural pharmacology. *Nature* **565,** 454–459 (2019).

77. Basak, S., Gicheru, Y., Kapoor, A., Mayer, M. L., Filizola, M. & Chakrapani, S. Molecular mechanism of setron-mediated inhibition of full-length 5-HT3A receptor. *Nature Communications* **10,** 3225 (2019).

78. Laverty, D., Thomas, P., Field, M., Andersen, O. J., Gold, M. G., Biggin, P. C., Gielen, M. & Smart, T. G. Crystal structures of a GABAA-receptor chimera reveal new endogenous neurosteroid-binding sites. *Nature Structural & Molecular Biology* **24,** 977–985 (2017).

79. Mori, M., Gähwiler, B. H. & Gerber, U. β-Alanine and taurine as endogenous agonists at glycine receptors in rat hippocampus in vitro. *The Journal of Physiology* **539,** 191–200 (2002).

80. Curtis, D., Duggan, A. & Johnston, G. The specificity of strychnine as a glycine antagonist in the mammalian spinal cord. *Experimental Brain Research* **12** (1971).

81. Zarbin, M. A., Wamsley, J. K. & Kuhar, M. J. Glycine receptor: light microscopic autoradiographic localization with [3H]strychnine. *Journal of Neuroscience* **1,** 532–547 (1981).

82. Hansen, S. B., Sulzenbacher, G., Huxford, T., Marchot, P., Taylor, P. & Bourne, Y. Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *The EMBO Journal* **24,** 3635–3646 (2005).

83. Lobo, I. A. & Harris, R. A. *International Review of Neurobiology* 53–87 (Academic Press, 2005).

84. Mihic, S. J., Ye, Q., Wick, M. J., Koltchine, V. V., Krasowski, M. D., Finn, S. E., Mascia, M. P., Valenzuela, C. F., Hanson, K. K., Greenblatt, E. P., Harris, R. A. & Harrison, N. L. Sites of alcohol and volatile anaesthetic action on GABA(A) and glycine receptors. *Nature* **389,** 385–389 (1997).

85. McCracken, M. L., Gorini, G., McCracken, L. M., Mayfield, R. D., Harris, R. A. & Trudell, J. R. Inter- and Intra-Subunit Butanol/Isoflurane Sites of Action in the Human Glycine Receptor. *Frontiers in Molecular Neuroscience* **9** (2016).

86. Yamakura, T., Mihic, S. J. & Harris, R. A. Amino acid volume and hydropathy of a transmembrane site determine glycine and anesthetic sensitivity of glycine receptors. *The Journal of biological chemistry* **274,** 23006–23012 (1999).

87. Shan, Q., Haddrill, J. L. & Lynch, J. W. Ivermectin, an Unconventional Agonist of the Glycine Receptor Chloride Channel*. *Journal of Biological Chemistry* **276,** 12556–12564 (2001).

88. Althoff, T., Hibbs, R. E., Banerjee, S. & Gouaux, E. X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature* **512,** 333–337 (2014).

89. Dämgen, M. A. & Biggin, P. C. State-dependent protein-lipid interactions of a pentameric ligand-gated ion channel in a neuronal membrane. *PLOS Computational Biology* **17,** e1007856 (2021).

90. Islam, R. & Lynch, J. W. Mechanism of action of the insecticides, lindane and fipronil, on glycine receptor chloride channels. *British Journal of Pharmacology* **165,** 2707–2720 (2012).

91. Bennion, B. J., Be, N. A., McNerney, M. W., Lao, V., Carlson, E. M., Valdez, C. A., Malfatti, M. A., Enright, H. A., Nguyen, T. H., Lightstone, F. C. & Carpenter, T. S. Predicting a Drug's Membrane Permeability: A Computational Model Validated With in Vitro Permeability Assay Data. *The Journal of Physical Chemistry B* **121,** 5228–5237 (2017).

92. *Pharmacology for Anaesthesia and Intensive Care* (eds Harris, B. & Peck, T.) 5th ed., 1–7 (Cambridge University Press, Cambridge, 2021).

93. Nelson, E. Kinetics of Drug Absorption, Distribution, Metabolism, and Excretion. *Journal of Pharmaceutical Sciences* **50,** 181–192 (1961).

94. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews. In Vitro Models for Selection of Development Candidates* **23,** 3–25 (1997).

95. Kansy, M., Senner, F. & Gubernator, K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *Journal of Medicinal Chemistry* **41,** 1007–1010 (1998).

96. Hubatsch, I., Ragnarsson, E. G. E. & Artursson, P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nature Protocols* **2,** 2111–2119 (2007).

97. Awoonor-Williams, E. & Rowley, C. N. Molecular simulation of nonfacilitated membrane permeation. *Biochimica et Biophysica Acta (BBA) - Biomembranes. New approaches for bridging computation and experiment on membrane proteins* **1858,** 1672–1687 (2016).

98. Lee, C. T., Comer, J., Herndon, C., Leung, N., Pavlova, A., Swift, R. V., Tung, C., Rowley, C. N., Amaro, R. E., Chipot, C., Wang, Y. & Gumbart, J. C. Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds. *Journal of Chemical Information and Modeling* **56,** 721–733 (2016).

99. Bernardi, A., Bennett, W. F. D., He, S., Jones, D., Kirshner, D., Bennion, B. J. & Carpenter, T. S. Advances in Computational Approaches for Estimating Passive Permeability in Drug Discovery. *Membranes* **13,** 851 (2023).

100. Marrink, S.-J. & Berendsen, H. J. C. Simulation of water transport through a lipid membrane. *The Journal of Physical Chemistry* **98,** 4155–4168 (1994).

101. Lapierre, J. & Hub, J. S. Converging PMF Calculations of Antibiotic Permeation across an Outer Membrane Porin with Subkilocalorie per Mole Accuracy. *Journal of Chemical Information and Modeling* **63,** 5319–5330 (2023).

102. Nitschke, N., Atkovska, K. & Hub, J. S. Accelerating potential of mean force calculations for lipid membrane permeation: System size, reaction coordinate, solute-solute distance, and cutoffs. *The Journal of Chemical Physics* **145,** 125101 (2016).

103. Sun, R., Han, Y., Swanson, J. M. J., Tan, J. S., Rose, J. P. & Voth, G. A. Molecular transport through membranes: Accurate permeability coefficients from multidimensional potentials of mean force and local diffusion constants. *The Journal of Chemical Physics* **149,** 072310 (2018).

104. Hille, B. Local anesthetics: hydrophilic and hydrophobic pathways for the drug-receptor reaction. *The Journal of General Physiology* **69,** 497–515 (1977).

105. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **280,** 69–77 (1998).

106. Payandeh, J., Scheuer, T., Zheng, N. & Catterall, W. A. THE CRYSTAL STRUCTURE OF A VOLTAGE-GATED SODIUM CHANNEL. *Nature* **475,** 353–358 (2011).

107. Neureiter, E. G., Erickson-Oberg, M. Q., Nigam, A. & Johnson, J. W. Inhibition of NMDA receptors and other ion channel types by membrane-associated drugs. *Frontiers in Pharmacology* **16** (2025).

108. Gamal El-Din, T. M. & Lenaeus, M. J. Fenestropathy of Voltage-Gated Sodium Channels. *Frontiers in Pharmacology* **13** (2022).

109. Jiang, D., Banh, R., El-Din, T. M. G., Tonggu, L., Lenaeus, M. J., Pomès, R., Zheng, N. & Catterall, W. A. Open-state structure and pore gating mechanism of the cardiac sodium channel. *Cell* **184,** 5151–5162.e11 (2021).

110. Yan, Z., Zhou, Q., Wang, L., Wu, J., Zhao, Y., Huang, G., Peng, W., Shen, H., Lei, J. & Yan, N. Structure of the Nav1.4-β1 Complex from Electric Eel. *Cell* **170,** 470–482.e11 (2017).

111. Huang, J., Fan, X., Jin, X., Jo, S., Zhang, H. B., Fujita, A., Bean, B. P. & Yan, N. Cannabidiol inhibits Nav channels through two distinct binding sites. *Nature Communications* **14,** 3613 (2023).

112. Gamal El-Din, T. M., Lenaeus, M. J., Zheng, N. & Catterall, W. A. Fenestrations control resting-state block of a voltage-gated sodium channel. *Proceedings of the National Academy of Sciences of the United States of America* **115,** 13111–13116 (2018).

113. Dong, Y. Y., Pike, A. C. W., Mackenzie, A., McClenaghan, C., Aryal, P., Dong, L., Quigley, A., Grieben, M., Goubin, S., Mukhopadhyay, S., Ruda, G. F., Clausen, M. V., Cao, L., Brennan, P. E., Burgess-Brown, N. A., Sansom, M. S. P., Tucker, S. J. & Carpenter, E. P. K2P channel gating mechanisms revealed by structures of TREK-2 and a complex with Prozac. *Science* **347,** 1256–1259 (2015).

114. Schott, K., Usher, S. G., Serra, O., Carnevale, V., Pless, S. A. & Chua, H. C. Unplugging lateral fenestrations of NALCN reveals a hidden drug binding site within the pore region. *Proceedings of the National Academy of Sciences* **121,** e2401591121 (2024).

115. Tao, E. & Corry, B. Characterizing fenestration size in sodium channel subtypes and their accessibility to inhibitors. *Biophysical Journal* **121,** 193–206 (2022).

116. Tao, E. & Corry, B. Drugs exhibit diverse binding modes and access routes in the Nav1.5 cardiac sodium channel pore. *Journal of General Physiology* **157,** e202413658 (2025).

117. Nguyen, P. T., DeMarco, K. R., Vorobyov, I., Clancy, C. E. & Yarov-Yarovoy, V. Structural basis for antiarrhythmic drug interactions with the human cardiac sodium channel. *Proceedings of the National Academy of Sciences* **116,** 2945–2954 (2019).

118. Hollingworth, D., Thomas, F., Page, D. A., Fouda, M. A., De Castro, R. L.-R., Sula, A., Mykhaylyk, V. B., Kelly, G., Ulmschneider, M. B., Ruben, P. C. & Wallace, B. A. Structural basis for the rescue of hyperexcitable cells by the amyotrophic lateral sclerosis drug Riluzole. *Nature Communications* **15,** 8426 (2024).

119. Sula, A., Montini, G., Booker, J. & Wallace, B. A. Fenestration Differences in Open and Closed Gate Sodium Channels: A Molecular Basis for State-Dependent Drug Design. *Biophysical Journal* **116,** 311a (2019).

120. Barygin, O. I., Komarova, M. S., Tikhonova, T. B. & Tikhonov, D. B. Non-classical mechanism of α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor channel block by fluoxetine. *European Journal of Neuroscience* **41,** 869–877 (2015).

121. Aryal, P., Jarerattanachat, V., Clausen, M. V., Schewe, M., McClenaghan, C., Argent, L., Conrad, L. J., Dong, Y. Y., Pike, A. C., Carpenter, E. P., Baukrowitz, T., Sansom, M. S. & Tucker, S. J. Bilayer-Mediated Structural Transitions Control Mechanosensitivity of the TREK-2 K2P Channel. *Structure* **25,** 708–718.e2 (2017).

122. Proks, P., Schewe, M., Conrad, L. J., Rao, S., Rathje, K., Rödström, K. E., Carpenter, E. P., Baukrowitz, T. & Tucker, S. J. Norfluoxetine inhibits TREK-2 K2P channels by multiple mechanisms including state-independent effects on the selectivity filter gate. *Journal of General Physiology* **153,** e202012812 (2021).

123. Wilcox, M. R., Nigam, A., Glasgow, N. G., Narangoda, C., Phillips, M. B., Patel, D. S., Mesbahi-Vasey, S., Turcu, A. L., Vázquez, S., Kurnikova, M. G. & Johnson, J. W. Inhibition of NMDA receptors through a membrane-to-channel path. *Nature Communications* **13,** 4114 (2022).

124. Coetzee, W. A., Amarillo, Y., Chiu, J., Chow, A., Lau, D., McCormack, T., Moreno, H., Nadal, M. S., Ozaita, A., Pountney, D., Saganich, M., Vega-Saenz de Miera, E. & Rudy, B. Molecular diversity of K+ channels. *Annals of the New York Academy of Sciences* **868,** 233–285 (1999).

125. Kudaibergenova, M., Perissinotti, L. L. & Noskov, S. Y. Lipid roles in hERG function and interactions with drugs. *Neuroscience Letters. Molecular Modeling in Neuroscience: New Windows on Function of the Nervous System* **700,** 70–77 (2019).

126. Marzian, S., Stansfeld, P. J., Rapedius, M., Rinné, S., Nematian-Ardestani, E., Abbruzzese, J. L., Steinmeyer, K., Sansom, M. S. P., Sanguinetti, M. C., Baukrowitz, T. & Decher, N. Side pockets provide the basis for a new mechanism of Kv channel–specific inhibition. *Nature Chemical Biology* **9,** 507–513 (2013).

127. Zhong, L., Lin, X., Cheng, X., Wan, S., Hua, Y., Nan, W., Hu, B., Peng, X., Zhou, Z., Zhang, Q., Yang, H., Noé, F., Yan, Z., Jiang, D., Zhang, H., Liu, F., Xiao, C., Zhou, Z., Mou, Y., Yu, H., Ma, L., Huang, C., Wong, V. K. W., Chung, S. K., Shen, B., Jiang, Z.-H., Neher, E., Zhu, W., Zhang, J. & Hou, P. Secondary structure transitions and dual PIP2 binding define cardiac KCNQ1-KCNE1 channel gating. *Cell Research,* 1–13 (2025).

128. McClenaghan, C., Schewe, M., Aryal, P., Carpenter, E. P., Baukrowitz, T. & Tucker, S. J. Polymodal activation of the TREK-2 K2P channel produces structurally distinct open states. *Journal of General Physiology* **147,** 497–505 (2016).

129. Nilius, B. & Honoré, E. Sensing pressure with ion channels. *Trends in Neurosciences* **35,** 477–486 (2012).

130. Brohawn, S. G., Campbell, E. B. & MacKinnon, R. Physical mechanism for gating and mechanosensitivity of the human TRAAK K+ channel. *Nature* **516,** 126–130 (2014).

131. Aryal, P., Abd-Wahab, F., Bucci, G., Sansom, M. S. P. & Tucker, S. J. A hydrophobic barrier deep within the inner pore of the TWIK-1 K2P potassium channel. *Nature Communications* **5,** 4377 (2014).

132. Jorgensen, C., Darré, L., Oakes, V., Torella, R., Pryde, D. & Domene, C. Lateral Fenestrations in K+-Channels Explored Using Molecular Dynamics Simulations. *Molecular Pharmaceutics* **13,** 2263–2273 (2016).

133. Yao, X., Gao, S. & Yan, N. Structural biology of voltage-gated calcium channels. *Channels* **18,** 2290807 (2024).

134. Dong, Y., Gao, Y., Xu, S., Wang, Y., Yu, Z., Li, Y., Li, B., Yuan, T., Yang, B., Zhang, X. C., Jiang, D., Huang, Z. & Zhao, Y. Closed-state inactivation and pore-blocker modulation mechanisms of human CaV2.2. *Cell Reports* **37** (2021).

135. Tikhonov, D. B. & Zhorov, B. S. Mechanisms of dihydropyridine agonists and antagonists in view of cryo-EM structures of calcium and sodium channels. *Journal of General Physiology* **155,** e202313418 (2023).

136. Cerdan, A. H., Peverini, L., Changeux, J.-P., Corringer, P.-J. & Cecchini, M. Lateral fenestrations in the extracellular domain of the glycine receptor contribute to the main chloride permeation pathway. *Science Advances* **8,** eadc9340 (2022).

137. Leach, A. R. *Molecular Modelling: Principles and Applications* Second Edition (Pearson Education, 2001).

138. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* Second Edition (Oxford University Press, 2017).

139. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91,** 43–56 (1995).

140. Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Molecular modeling annual* **7,** 306–317 (2001).

141. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. GROMACS: fast, flexible, and free. *Journal of Computational Chemistry* **26,** 1701–1718 (2005).

142. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **4,** 435–447 (2008).

143. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B. & Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics (Oxford, England)* **29,** 845–854 (2013).

144. Páll, S., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. *Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS Solving Software Challenges for Exascale* (eds Markidis, S. & Laure, E.) (Springer International Publishing, Cham, 2015), 3–27.

145. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. & Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2,** 19–25 (2015).

146. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **78,** 1950–1958 (2010).

147. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18,** 1463–1472 (1997).

148. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **13,** 952–962 (1992).

149. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. & Pedersen, L. G. A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103,** 8577–8593 (1995).

150. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81,** 3684–3690 (1984).

151. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126,** 014101 (2007).

152. Bernetti, M. & Bussi, G. Pressure control using stochastic cell rescaling. *The Journal of Chemical Physics* **153,** 114107 (2020).

153. Zhang, S., Hahn, D. F., Shirts, M. R. & Voelz, V. A. Expanded Ensemble Methods Can be Used to Accurately Predict Protein-Ligand Relative Binding Free Energies. *Journal of Chemical Theory and Computation* **17,** 6536–6547 (2021).

154. Hsu, W.-T. & Shirts, M. R. Replica Exchange of Expanded Ensembles: A Generalized Ensemble Approach with Enhanced Flexibility and Parallelizability. *Journal of Chemical Theory and Computation* **20,** 6062–6081 (2024).

155. Invernizzi, M., Piaggi, P. M. & Parrinello, M. Unified Approach to Enhanced Sampling. *Physical Review X* **10,** 041034 (2020).

156. Marinari, E. & Parisi, G. Simulated Tempering: A New Monte Carlo Scheme. *Europhysics Letters* **19,** 451 (1992).

157. Park, S. & Pande, V. S. Choosing weights for simulated tempering. *Physical Review E* **76,** 016703 (2007).

158. Wang, F. & Landau, D. P. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Physical Review Letters* **86,** 2050–2053 (2001).

159. Lemkul, J. A. & Bevan, D. R. Assessing the Stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics. *The Journal of Physical Chemistry B* **114,** 1652–1660 (2010).

160. Bowman, J. D. & Lindert, S. Molecular Dynamics and Umbrella Sampling Simulations Elucidate Differences in Troponin C Isoform and Mutant Hydrophobic Patch Exposure. *The Journal of Physical Chemistry B* **122,** 7874–7883 (2018).

161. Cetin, H., Epstein, M., Liu, W. W., Maxwell, S., Rodriguez Cruz, P. M., Cossins, J., Vincent, A., Webster, R., Biggin, P. C. & Beeson, D. Muscle acetylcholine receptor conversion into chloride conductance at positive potentials by a single mutation. *Proceedings of the National Academy of Sciences* **116,** 21228–21235 (2019).

162. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **13,** 1011–1021 (1992).

163. Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics* **129,** 124105 (2008).

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

165. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics* **54,** 5.6.1–5.6.37 (2016).

166. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *Journal of computational chemistry* **32,** 2319–2327 (2011).

167. Gowers, R. J., Linke, M., Barnoud, J., Reddy, T. J. E., Melo, M. N., Seyler, S. L., Domański, J., Dotson, D. L., Buchoux, S., Kenney, I. M. & Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *scipy* (2016).

168. Klesse, G., Rao, S., Sansom, M. S. P. & Tucker, S. J. CHAP: A Versatile Tool for the Structural and Functional Annotation of Ion Channel Pores. *Journal of Molecular Biology* **431,** 3353–3365 (2019).

169. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J. & Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Computational Biology* **8** (ed Prlic, A.) e1002708 (2012).

170. Smith, R. H. B., Dar, A. C. & Schlessinger, A. *PyVOL: a PyMOL plugin for visualization, comparison, and volume calculation of drug-binding sites* 2019.

171. Camenisch, G., Folkers, G. & van de Waterbeemd, H. Review of theoretical passive drug absorption models: historical background, recent developments and limitations. *Pharmaceutica Acta Helvetiae* **71,** 309–327 (1996).

172. Mangas-Sanjuan, V., González-Alvarez, M., Gonzalez-Alvarez, I. & Bermejo, M. Drug penetration across the blood-brain barrier: an overview. *Therapeutic Delivery* **1,** 535–562 (2010).

173. Lundborg, M., Wennberg, C., Lidmar, J., Hess, B., Lindahl, E. & Norlén, L. Skin permeability prediction with MD simulation sampling spatial and alchemical reaction coordinates. *Biophysical Journal* **121,** 3837–3849 (2022).

174. Gleeson, M. P. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *Journal of Medicinal Chemistry* **51,** 817–834 (2008).

175. Wu, F., Zhou, Y., Li, L., Shen, X., Chen, G., Wang, X., Liang, X., Tan, M. & Huang, Z. Computational Approaches in Preclinical Studies on Drug Discovery and Development. *Frontiers in Chemistry* **8,** 726 (2020).

176. Ong, S., Liu, H., Qiu, X., Bhat, G. & Pidgeon, C. Membrane Partition Coefficients Chromatographically Measured Using Immobilized Artificial Membrane Surfaces. *Analytical Chemistry* **67,** 755–762 (1995).

177. Takegami, S., Kitamura, K., Funakoshi, T. & Kitade, T. Partitioning of anti-inflammatory steroid drugs into phosphatidylcholine and phosphatidylcholine-cholesterol small unilamellar vesicles as studied by second-derivative spectrophotometry. *Chemical & Pharmaceutical Bulletin* **56,** 663–667 (2008).

178. Symons, J., Cho, K.-J., Chang, J., Du, G., Waxham, M. N., Hancock, J. F., Levental, I. & Levental, K. R. Lipidomic atlas of mammalian cell membranes reveals hierarchical variation induced by culture conditions, subcellular membranes, and cell lineages. *Soft matter* **17,** 288–297 (2021).

179. Wennberg, C. L., van der Spoel, D. & Hub, J. S. Large Influence of Cholesterol on Solute Partitioning into Lipid Membranes. *Journal of the American Chemical Society* **134,** 5351–5361 (2012).

180. Hub, J. S., De Groot, B. L. & Van Der Spoel, D. g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *Journal of Chemical Theory and Computation* **6,** 3713–3720 (2010).

181. Markthaler, D., Jakobtorweihen, S. & Hansen, N. Lessons Learned from the Calculation of One-Dimensional Potentials of Mean Force [Article v1.0]. *Living Journal of Computational Molecular Science* **1,** 11073–11073 (2019).

182. Sousa, C. F., Becker, R. A., Lehr, C.-M., Kalinina, O. V. & Hub, J. S. Simulated Tempering-Enhanced Umbrella Sampling Improves Convergence of Free Energy Calculations of Drug Membrane Permeation. *Journal of Chemical Theory and Computation* **19,** 1898–1907 (2023).

183. Cockcroft, S. Mammalian lipids: structure, synthesis and function. *Essays in Biochemistry* **65,** 813–845 (2021).

184. Kučerka, N., Nieh, M.-P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1808,** 2761–2771 (2011).

185. Koynova, R. & Caffrey, M. Phases and phase transitions of the phosphatidylcholines. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1376,** 91–145 (1998).

186. Chapman, D. Phase transitions and fluidity characteristics of lipids and cell membranes. *Quarterly Reviews of Biophysics* **8,** 185–235 (1975).

187. Zhang, L., Zhang, Z., Jasa, J., Li, D., Cleveland, R. O., Negahban, M. & Jérusalem, A. Molecular dynamics simulations of heterogeneous cell membranes in response to uniaxial membrane stretches at high loading rates. *Scientific Reports* **7,** 8316 (2017).

188. Sonntag, Y., Musgaard, M., Olesen, C., Schiøtt, B., Møller, J. V., Nissen, P. & Thøgersen, L. Mutual adaptation of a membrane protein and its lipid bilayer during conformational changes. *Nature Communications* **2,** 304 (2011).

189. Itaya, H., Kasahara, K., Xie, Q., Yano, Y., Matsuzaki, K. & Takahashi, T. All-Atom Molecular Dynamics Elucidating Molecular Mechanisms of Single-Transmembrane Model Peptide Dimerization in a Lipid Bilayer. *ACS Omega* **6,** 11458–11465 (2021).

190. Marrink, S. J., Corradi, V., Souza, P. C., Ingólfsson, H. I., Tieleman, D. P. & Sansom, M. S. Computational Modeling of Realistic Cell Membranes. *Chemical Reviews* **119,** 6184–6226 (2019).

191. Mouritsen, O. G. & Zuckermann, M. J. What's so special about cholesterol? *Lipids* **39,** 1101–1113 (2004).

192. Sackmann, E. *Handbook of Biological Physics* (eds Lipowsky, R. & Sackmann, E.) 1–63 (North-Holland, 1995).

193. Ingólfsson, H. I., Melo, M. N., van Eerden, F. J., Arnarez, C., Lopez, C. A., Wassenaar, T. A., Periole, X., de Vries, A. H., Tieleman, D. P. & Marrink, S. J. Lipid Organization of the Plasma Membrane. *Journal of the American Chemical Society* **136,** 14554–14559 (2014).

194. Khodadadi, E., Khodadadi, E., Chaturvedi, P. & Moradi, M. Comprehensive Insights into the Cholesterol-Mediated Modulation of Membrane Function Through Molecular Dynamics Simulations. *Membranes* **15,** 173 (2025).

195. De Young, L. R. & Dill, K. A. Solute partitioning into lipid bilayer membranes. *Biochemistry* **27,** 5281–5289 (1988).

196. Programme, U. N. E. Stockholm Convention on Persistent Organic Pollutants (POPs) as amended in 2009: Text and Annexes (2009).

197. Wafford, K. A., Lummis, S. C. & Sattelle, D. B. Block of an insect central nervous system GABA receptor by cyclodiene and cyclohexane insecticides. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **237,** 53–61 (1989).

198. Raymond-Delpech, V., Matsuda, K., Sattelle, B. M., Rauh, J. J. & Sattelle, D. B. Ion channels: molecular targets of neuroactive insecticides. *Invertebrate Neuroscience* **5,** 119–133 (2005).

199. Antunes-Madeira, M. C. & Madeira, V. M. C. Partition of lindane in synthetic and native membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **820,** 165–172 (1985).

200. Grasshoff, C. & Antkowiak, B. Effects of isoflurane and enflurane on GABAA and glycine receptors contribute equally to depressant actions on spinal ventral horn neurones in rats. *British Journal of Anaesthesia* **97,** 687–694 (2006).

201. Matta, J. A., Cornett, P. M., Miyares, R. L., Abe, K., Sahibzada, N. & Ahern, G. P. General anesthetics activate a nociceptive ion channel to enhance pain and inflammation. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 8784–8789 (2008).

202. Kanaya, H. J., Kuwajima, K., Ito, Y., Shinohara, Y., Okubo, Y., Shiono, S., Tatsuki, F., Ohno, R.-i., Ukai, H., Ukai-Tadenuma, M., Sumiyama, K., Fujishima, H., Yamada, R. G., Tone, D., Kiyonari, H., Kikuchi, M., Umehara, T., Murayama, T., Kanemaru, K., Iino, M., Ode, K. L., Hirokawa, T. & Ueda, H. R. Isoflurane activates the type 1 ryanodine receptor to induce anesthesia in mice. *PLOS Biology* **23,** e3003172 (2025).

203. Chen, Q., Kinde, M. N., Arjunan, P., Wells, M. M., Cohen, A. E., Xu, Y. & Tang, P. Direct Pore Binding as a Mechanism for Isoflurane Inhibition of the Pentameric Ligand-gated Ion Channel ELIC. *Scientific Reports* **5,** 13833 (2015).

204. Qiu, J., Yang, Y., Liu, J., Zhao, W., Li, Q., Zhu, T., Liang, P. & Zhou, C. The volatile anesthetic isoflurane differentially inhibits voltage-gated sodium channel currents between pyramidal and parvalbumin neurons in the prefrontal cortex. *Frontiers in Neural Circuits* **17** (2023).

205. Dickinson, R., Franks, N. P. & Lieb, W. R. Can the stereoselective effects of the anesthetic isoflurane be accounted for by lipid solubility? *Biophysical Journal* **66,** 2019–2023 (1994).

206. Wasmuth, C. E. & Hale, D. E. THIOPENTAL SODIUM ANESTHESIA IN INFANTS AND CHILDREN. *Journal of the American Medical Association* **156,** 1321–1323 (1954).

207. Koniaris, L. G., Zimmers, T. A., Lubarsky, D. A. & Sheldon, J. P. Inadequate anaesthesia in lethal injection for execution. *Lancet (London, England)* **365,** 1412–1414 (2005).

208. Kala, A. K. Of ethically compromising positions and blatant lies about 'truth serum'. *Indian Journal of Psychiatry* **49,** 6–9 (2007).

209. Korten, K., Sommer, T. J. & Miller, K. W. Membrane composition modulates thiopental partitioning in bilayers and biomembranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **599,** 271–279 (1980).

210. Martynyuk, A. E., Morey, T. E., Raatikainen, P. M. J., Seubert, C. N. & Dennis, D. M. Ionic Mechanisms Mediating the Differential Effects of Methohexital and Thiopental on Action Potential Duration in Guinea Pig and Rabbit Isolated Ventricular Myocytes. *Anesthesiology* **90,** 156 (1999).

211. George, P., Jones, N., Goldman, H. & Rosenblatt, A. Cycles of reform in the history of psychosis treatment in the United States. *SSM - Mental Health* **3,** 100205 (2023).

212. Luxnat, M. & Galla, H.-J. Partition of chlorpromazine into lipid bilayer membranes: the effect of membrane structure and composition. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **856,** 274–282 (1986).

213. Avdeef, A. *Absorption and drug development: solubility, permeability, and charge state* 1 p. (Wiley-Interscience, Hoboken, N.J, 2003).

214. Chen, J. Y., Brunauer, L. S., Chu, F. C., Helsel, C. M., Gedde, M. M. & Huestis, W. H. Selective amphipathic nature of chlorpromazine binding to plasma membrane bilayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1616,** 95–105 (2003).

215. Neubert, R. Ion pair transport across membranes. *Pharmaceutical Research* **6,** 743–747 (1989).

216. Elferink, J. G. R. The asymmetric distribution of chlorpromazine and its quaternary analogue over the erythrocyte membrane. *Biochemical Pharmacology* **26,** 2411–2416 (1977).

217. PubChem. *Paroxetine* `https://pubchem.ncbi.nlm.nih.gov/compound/43815` (2025).

218. Ngo, D. T. N., Nguyen, T. Q., Huynh, H. K. & Nguyen, T. T. Thermodynamics of selective serotonin reuptake inhibitors partitioning into 1,2-dioleoyl-sn-glycero-3-phosphocholine bilayers. *RSC Advances* **10,** 39338–39347.

219. Ngo, D. T. N., Ho, T. H., Huynh, L. K. & Nguyen, T. T. The interplay of membrane fluidity, acyl chain order and area per lipid on the partitioning of two antidepressants paroxetine and sertraline. *Soft Matter* **19,** 5527–5537 (2023).

220. Deák, K., Takács-Novák, K., Tihanyi, K. & Noszál, B. Physico-chemical profiling of antidepressive sertraline: solubility, ionisation, lipophilicity. *Medicinal Chemistry (Shariqah (United Arab Emirates))* **2,** 385–389 (2006).

221. Jämbeck, J. P. M. & Lyubartsev, A. P. Another Piece of the Membrane Puzzle: Extending Slipids Further. *Journal of Chemical Theory and Computation* **9,** 774–784 (2013).

222. Jämbeck, J. P. M. & Lyubartsev, A. P. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *Journal of Chemical Theory and Computation* **8,** 2938–2948 (2012).

223. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79,** 926–935 (1983).

224. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **25,** 1157–1174 (2004).

225. Jo, S., Kim, T. & Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS ONE* **2** (ed Yuan, A.) e880 (2007).

226. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **29,** 1859–1865 (2008).

227. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophysical Journal* **97,** 50–58 (2009).

228. Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., MacKerell, A. D., Klauda, J. B. & Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **12,** 405–413 (2016).

229. Wu, E. L., Cheng, X., Jo, S., Rui, H., Song, K. C., Dávila-Contreras, E. M., Qi, Y., Lee, J., Monje-Galvan, V., Venable, R. M., Klauda, J. B. & Im, W. CHARMM-GUI *Membrane Builder* toward realistic biological membrane simulations. *Journal of Computational Chemistry* **35,** 1997–2004 (2014).

230. Jefferys, E., Sands, Z. A., Shi, J., Sansom, M. S. P. & Fowler, P. W. Alchembed: A Computational Method for Incorporating Multiple Proteins into Complex Lipid Geometries. *Journal of Chemical Theory and Computation* **11,** 2743–2754 (2015).

231. Lichtinger, S. M. & Biggin, P. C. Tackling Hysteresis in Conformational Sampling: How to Be Forgetful with MEMENTO. *Journal of Chemical Theory and Computation* **19,** 3705–3720 (2023).

232. Zeng, X., Hu, H., Zhou, H.-X., Marszalek, P. E. & Yang, W. Equilibrium Sampling for Biomolecules under Mechanical Tension. *Biophysical Journal* **98,** 733–740 (2010).

233. Oshima, H., Re, S. & Sugita, Y. Replica-Exchange Umbrella Sampling Combined with Gaussian Accelerated Molecular Dynamics for Free-Energy Calculation of Biomolecules. *Journal of Chemical Theory and Computation* **15,** 5199–5208 (2019).

234. Jo, S., Suh, D., He, Z., Chipot, C. & Roux, B. Leveraging the Information from Markov State Models To Improve the Convergence of Umbrella Sampling Simulations. *The Journal of Physical Chemistry. B* **120,** 8733–8742 (2016).

235. Bartosik, A., Wiśniewska, M. & Makowski, M. Potentials of mean force for hydrophobic interactions between hydrocarbons in water solution: dependence on temperature, solute shape, and solute size. *Journal of Physical Organic Chemistry* **28,** 10–16 (2015).

236. Massi, F. & Palmer, A. G. Temperature dependence of NMR order parameters and protein dynamics. *Journal of the American Chemical Society* **125,** 11158–11159 (2003).

237. Roux, B. The calculation of the potential of mean force using computer simulations. *Computer Physics Communications* **91,** 275–282 (1995).

238. Noe, F. *Reweighting distributions between thermodynamic states* `http://docs.markovmodel.org/reweighting.html` (2025).

239. Okamoto, Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *Journal of Molecular Graphics and Modelling. Conformational Sampling* **22,** 425–439 (2004).

240. Chodera, J. D., Swope, W. C., Pitera, J. W., Seok, C. & Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *Journal of Chemical Theory and Computation* **3,** 26–41 (2007).

241. Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. E. Array programming with NumPy. *Nature* **585,** 357–362 (2020).

242. Pople, J. A. Nobel Lecture: Quantum chemical models. *Reviews of Modern Physics* **71,** 1267–1274 (1999).

243. Doktorova, M., Symons, J. L., Zhang, X., Wang, H.-Y., Schlegel, J., Lorent, J. H., Heberle, F. A., Sezgin, E., Lyman, E., Levental, K. R. & Levental, I. Cell membranes sustain phospholipid imbalance via cholesterol asymmetry. *Cell* **188,** 2586–2602.e24 (2025).

244. Gu, R.-X., Baoukina, S. & Tieleman, D. P. Cholesterol Flip-Flop in Heterogeneous Membranes. *Journal of Chemical Theory and Computation* **15,** 2064–2070 (2019).

245. Choubey, A., Kalia, R. K., Malmstadt, N., Nakano, A. & Vashishta, P. Cholesterol Translocation in a Phospholipid Membrane. *Biophysical Journal* **104,** 2429–2436 (2013).

246. Zhu, Y., Porcar, L., Ravula, T., Batchu, K. C., Lavoie, T. L., Liu, Y. & Perez-Salas, U. Unexpected asymmetric distribution of cholesterol and phospholipids in equilibrium model membranes. *Biophysical Journal* **123,** 3923–3934 (2024).

247. Cheng, T., Zhao, Y., Li, X., Lin, F., Xu, Y., Zhang, X., Li, Y., Wang, R. & Lai, L. Computation of octanol-water partition coefficients by guiding an additive model with knowledge. *Journal of Chemical Information and Modeling* **47,** 2140–2148 (2007).

248. Allen, T. W., Andersen, O. S. & Roux, B. Ion Permeation through a Narrow Channel: Using Gramicidin to Ascertain All-Atom Molecular Dynamics Potential of Mean Force Methodology and Biomolecular Force Fields. *Biophysical Journal* **90,** 3447–3468 (2006).

249. Hanson, S. M., Newstead, S., Swartz, K. J. & Sansom, M. S. P. Capsaicin Interaction with TRPV1 Channels in a Lipid Bilayer: Molecular Dynamics Simulation. *Biophysical Journal* **108,** 1425–1434 (2015).

250. Ribeiro, R. P., Coimbra, J. T. S., Ramos, M. J. & Fernandes, P. A. Diffusion of the small, very polar, drug piracetam through a lipid bilayer: an MD simulation study. *Theoretical Chemistry Accounts* **136,** 46 (2017).

251. Liu, P., Kim, B., Friesner, R. A. & Berne, B. J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proceedings of the National Academy of Sciences* **102,** 13749–13754 (2005).

252. Wang, L., Friesner, R. A. & Berne, B. Replica Exchange with Solute Scaling: A more efficient version of Replica Exchange with Solute Tempering (REST2). *The journal of physical chemistry. B* **115,** 9431–9438 (2011).

253. Feddersen, B. & Lundborg, M. *Expanded ensemble: Performance impact of perturbing more atoms* GROMACS User Discussions. `https://gromacs.bioexcel.eu/t/expanded-ensemble-performance-impact-of-perturbing-more-atoms/11099/1` (2025).

254. Kutzner, C., Köpfer, D. A., Machtens, J.-P., de Groot, B. L., Song, C. & Zachariae, U. Insights into the function of ion channels by computational electrophysiology simulations. *Biochimica Et Biophysica Acta* **1858,** 1741–1752 (2016).

255. Pipatpolkai, T., Corey, R. A., Proks, P., Ashcroft, F. M. & Stansfeld, P. J. Evaluating inositol phospholipid interactions with inward rectifier potassium channels and characterising their role in disease. *Communications Chemistry* **3,** 147 (2020).

256. Pipatpolkai, T., Quetschlich, D. & Stansfeld, P. J. From Bench to Biomolecular Simulation: Phospholipid Modulation of Potassium Channels. *Journal of Molecular Biology. Ion channels : Intersection of Structure, Function, and Pharmacology* **433,** 167105 (2021).

257. Domański, J., Hedger, G., Best, R. B., Stansfeld, P. J. & Sansom, M. S. P. Convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association. *The Journal of Physical Chemistry B* **121,** 3364–3375 (2017).

258. Lundborg, M., Narangifard, A., Wennberg, C. L., Lindahl, E., Daneholt, B. & Norlén, L. Human skin barrier structure and function analyzed by cryo-EM and molecular dynamics simulation. *Journal of Structural Biology* **203,** 149–161 (2018).

259. Stehr, C. M., Linbo, T. L., Incardona, J. P. & Scholz, N. L. The Developmental Neurotoxicity of Fipronil: Notochord Degeneration and Locomotor Defects in Zebrafish Embryos and Larvae. *Toxicological Sciences* **92,** 270–278 (2006).

260. Sparks, T. C. & Nauen, R. IRAC: Mode of action classification and insecticide resistance management. *Pesticide Biochemistry and Physiology. Insecticide and Acaricide Modes of Action and their Role in Resistance and its Management* **121,** 122–128 (2015).

261. Taylor-Wells, J., Hawkins, J., Colombo, C., Bermudez, I. & Jones, A. K. Cloning and functional expression of intracellular loop variants of the honey bee (*Apis mellifera*) RDL GABA receptor. *NeuroToxicology* **60,** 207–213 (2017).

262. Li, X., Bao, C., Yang, D., Zheng, M., Li, X. & Tao, S. Toxicities of fipronil enantiomers to the honeybee Apis mellifera L. and enantiomeric compositions of fipronil in honey plant flowers. *Environmental Toxicology and Chemistry* **29,** 127–132 (2010).

263. Bonmatin, J.-M., Giorio, C., Girolami, V., Goulson, D., Kreutzweiser, D. P., Krupke, C., Liess, M., Long, E., Marzaro, M., Mitchell, E. A. D., Noome, D. A., Simon-Delso, N. & Tapparo, A. Environmental fate and exposure; neonicotinoids and fipronil. *Environmental Science and Pollution Research* **22,** 35–67 (2015).

264. Tomlin, C. D. S. *The Pesticide Manual, A World Compendium* 14th Edition. pp. 462-464 (British Crop Protection Council, Hampshire, England, 2006).

265. Qiao, X., Zhou, T., Zhang, J., Zhang, L., Lu, Y. & Huang, J. Functional validation of A2'N mutation of the RDL GABA receptor against fipronil via molecular modeling and genome engineering in drosophila. *Pest Management Science* **80,** 1924–1929 (2024).

266. Zheng, N., Cheng, J., Zhang, W., Li, W., Shao, X., Xu, Z., Xu, X. & Li, Z. Binding difference of fipronil with GABAARs in fruitfly and zebrafish: insights from homology modeling, docking, and molecular dynamics simulation studies. *Journal of Agricultural and Food Chemistry* **62,** 10646–10653 (2014).

267. Law, R. J. & Lightstone, F. C. Gaba Receptor Insecticide Non-Competitive Antagonists May Bind at Allosteric Modulator Sites. *International Journal of Neuroscience* **118,** 705–734 (2008).

268. Chen, X., Cromer, B., Webb, T. I., Yang, Z., Hantke, J., Harvey, R. J., Parker, M. W. & Lynch, J. W. Dihydropyridine inhibition of the glycine receptor: Subunit selectivity and a molecular determinant of inhibition. *Neuropharmacology. Ligand-Gated Ion Channels* **56,** 318–327 (2009).

269. Shindyalov, I. N. & Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Engineering* **11,** 739–747 (1998).

270. Eastman, P., Friedrichs, M. S., Chodera, J. D., Radmer, R. J., Bruns, C. M., Ku, J. P., Beauchamp, K. A., Lane, T. J., Wang, L.-P., Shukla, D., Tye, T., Houston, M., Stich, T., Klein, C., Shirts, M. R. & Pande, V. S. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *Journal of Chemical Theory and Computation* **9,** 461–469 (2013).

271. Kandt, C., Ash, W. L. & Tieleman, D. P. Setting up and running molecular dynamics simulations of membrane proteins. *Methods (San Diego, Calif.)* **41,** 475–488 (2007).

272. Owen, B. B. The Dissociation Constants of Glycine at Various Temperatures. *Journal of the American Chemical Society* **56,** 24–27 (1934).

273. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **185,** 604–613 (2014).

274. Bonomi, M., Bussi, G., Camilloni, C., Tribello, G. A., Banáš, P., Barducci, A., Bernetti, M., Bolhuis, P. G., Bottaro, S., Branduardi, D., Capelli, R., Carloni, P., Ceriotti, M., Cesari, A., Chen, H., Chen, W., Colizzi, F., De, S., De La Pierre, M., Donadio, D., Drobot, V., Ensing, B., Ferguson, A. L., Filizola, M., Fraser, J. S., Fu, H., Gasparotto, P., Gervasio, F. L., Giberti, F., Gil-Ley, A., Giorgino, T., Heller, G. T., Hocky, G. M., Iannuzzi, M., Invernizzi, M., Jelfs, K. E., Jussupow, A., Kirilin, E., Laio, A., Limongelli, V., Lindorff-Larsen, K., Löhr, T., Marinelli, F., Martin-Samos, L., Masetti, M., Meyer, R., Michaelides, A., Molteni, C., Morishita, T., Nava, M., Paissoni, C., Papaleo, E., Parrinello, M., Pfaendtner, J., Piaggi, P., Piccini, G., Pietropaolo, A., Pietrucci, F., Pipolo, S., Provasi, D., Quigley, D., Raiteri, P., Raniolo, S., Rydzewski, J., Salvalaglio, M., Sosso, G. C., Spiwok, V., Šponer, J., Swenson, D. W. H., Tiwary, P., Valsson, O., Vendruscolo, M., Voth, G. A., White, A. & The PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods* **16,** 670–673 (2019).

275. Tukey, J. W. Comparing Individual Means in the Analysis of Variance. *Biometrics* **5,** 99–114 (1949).

276. Vickery, O. *owenvickery/umbrella_sampling: pmf setup and analysis* version v1.0.1. 2020.

277. Starmer, C. F., Grant, A. O. & Strauss, H. C. Mechanisms of use-dependent block of sodium channels in excitable membranes by local anesthetics. *Biophysical Journal* **46,** 15–27 (1984).

278. Borghese, C. M. & Goldschen-Ohm, M. P. State-dependent energetics of GABAA receptor modulators. *Biophysical Journal* **123,** 1903–1906 (2024).

279. Ghosh, B., Satyshur, K. A. & Czajkowski, C. Propofol Binding to the Resting State of the *Gloeobacter violaceus* Ligand-gated Ion Channel (GLIC) Induces Structural Changes in the Inter- and Intrasubunit Transmembrane Domain (TMD) Cavities*. *Journal of Biological Chemistry* **288,** 17420–17431 (2013).

280. Emigh Cortez, A. M., DeMarco, K. R., Furutani, K., Bekker, S., Sack, J. T., Wulff, H., Clancy, C. E., Vorobyov, I. & Yarov-Yarovoy, V. Structural modeling of hERG channel–drug interactions using Rosetta. *Frontiers in Pharmacology* **14** (2023).

281. Prokop, Z., Gora, A., Brezovsky, J., Chaloupkova, R., Stepankova, V. & Damborsky, J. *Protein Engineering Handbook* (Wiley, 2013).

282. Kingsley, L. J. & Lill, M. A. Substrate tunnels in enzymes: Structure–function relationships and computational methodology. *Proteins: Structure, Function, and Bioinformatics* **83,** 599–611 (2015).

283. Leacock, S., Syed, P., James, V. M., Bode, A., Kawakami, K., Keramidas, A., Suster, M., Lynch, J. W. & Harvey, R. J. Structure/Function Studies of the α4 Subunit Reveal Evolutionary Loss of a GlyR Subtype Involved in Startle and Escape Responses. *Frontiers in Molecular Neuroscience* **11,** 23 (2018).

284. Zhu, H. & Gouaux, E. Architecture and assembly mechanism of native glycine receptors. *Nature* **599,** 513–517 (2021).

285. Schweizer, C., Balsiger, S., Bluethmann, H., Mansuy, I. M., Fritschy, J.-M., Mohler, H. & Lüscher, B. The gamma 2 subunit of GABA(A) receptors is required for maintenance of receptors at mature synapses. *Molecular and Cellular Neurosciences* **24,** 442–450 (2003).

286. Gibbs, E., Feddersen, B., Kindig, K., Seiferth, D., Biggin, P. C. & Chakrapani, S. Structural Basis for Domain Coupling in Heteromeric Glycine Receptors Revealed by an Atypical Allosteric Agonist. *Science Advances.*

287. Siebler, M., Pekel, M., Köller, H. & Müller, H.-W. Strychnine-sensitive glycine receptors in cultured primary neurons from rat neocortex. *Developmental Brain Research* **73,** 289–292 (1993).

288. Young, A. B. & Snyder, S. H. Strychnine Binding Associated with Glycine Receptors of the Central Nervous System. *Proceedings of the National Academy of Sciences of the United States of America* **70,** 2832–2836 (1973).

289. Cully, D. F., Vassilatis, D. K., Liu, K. K., Paress, P. S., Van der Ploeg, L. H. T., Schaeffer, J. M. & Arena, J. P. Cloning of an avermectin-sensitive glutamate-gated chloride channel from Caenorhabditis elegans. *Nature* **371,** 707–711 (1994).

290. Wolstenholme, A. J. Glutamate-gated Chloride Channels. *The Journal of Biological Chemistry* **287,** 40232–40238 (2012).

291. Lynagh, T., Webb, T. I., Dixon, C. L., Cromer, B. A. & Lynch, J. W. Molecular Determinants of Ivermectin Sensitivity at the Glycine Receptor Chloride Channel*. *Journal of Biological Chemistry* **286,** 43913–43924 (2011).

292. Maher, A., Radwan, R. & Breitinger, H.-G. In vivo protection against strychnine toxicity in mice by the glycine receptor agonist ivermectin. *BioMed Research International* **2014,** 640790 (2014).

293. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I. & Mackerell, A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **31,** 671–690 (2010).

294. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., De Groot, B. L., Grubmüller, H. & MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods* **14,** 71–73 (2017).

295. Vanommeslaeghe, K. & MacKerell, A. D. J. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *Journal of Chemical Information and Modeling* **52,** 3144–3154 (2012).

296. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F. & van Mulbregt, P. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17,** 261–272 (2020).

297. Bouysset, C. & Fiorucci, S. ProLIF: a library to encode molecular interactions as fingerprints. *Journal of Cheminformatics* **13,** 72 (2021).

298. DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA, USA, 2002).

299. Bowerman, S. & Wereszczynski, J. *Methods in Enzymology* (ed Voth, G. A.) 429–447 (Academic Press, 2016).

300. Meng, Y., Shukla, D., Pande, V. S. & Roux, B. Transition path theory analysis of c-Src kinase activation. *Proceedings of the National Academy of Sciences* **113,** 9193–9198 (2016).

301. Bowman, G. R. & Geissler, P. L. Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 11681–11686 (2012).

302. Sethi, A., Eargle, J., Black, A. A. & Luthey-Schulten, Z. Dynamical networks in tRNA:protein complexes. *Proceedings of the National Academy of Sciences* **106,** 6620–6625 (2009).

303. Li, Z., Bharambe, N., Lande, K., Feddersen, B., Balakrishna, A., Biggin, P. C., Sahu, G. & Basak, S. Asymmetric gating of the homopentameric ion channel GLIC revealed by cryo-EM. *Proceedings of the National Academy of Sciences* (2025).

304. Ghosh, B., Tsao, T.-W. & Czajkowski, C. A chimeric prokaryotic-eukaryotic pentameric ligand gated ion channel reveals interactions between the extracellular and transmembrane domains shape neurosteroid modulation. *Neuropharmacology* **125,** 343–352 (2017).

305. Velisetty, P. & Chakrapani, S. Desensitization Mechanism in Prokaryotic Ligand-gated Ion Channel. *The Journal of Biological Chemistry* **287,** 18467–18477 (2012).

306. Taly, A., Hénin, J., Changeux, J.-P. & Cecchini, M. Allosteric regulation of pentameric ligand-gated ion channels. *Channels* **8,** 350–360 (2014).

307. Salari, R., Murlidaran, S. & Brannigan, G. Pentameric Ligand-gated Ion Channels : Insights from Computation. *Molecular simulation* **40,** 821–829 (2014).

308. Song, C. & Corry, B. Ion conduction in ligand-gated ion channels: Brownian dynamics studies of four recent crystal structures. *Biophysical Journal* **98,** 404–411 (2010).

309. Zhang, Y., Dijkman, P. M., Zou, R., Zandl-Lang, M., Sanchez, R. M., Eckhardt-Strelau, L., Köfeler, H., Vogel, H., Yuan, S. & Kudryashev, M. Asymmetric opening of the homopentameric 5-HT3A serotonin receptor in lipid bilayers. *Nature Communications* **12,** 1074 (2021).

310. Lev, B., Murail, S., Poitevin, F., Cromer, B. A., Baaden, M., Delarue, M. & Allen, T. W. String method solution of the gating pathways for a pentameric ligand-gated ion channel. *Proceedings of the National Academy of Sciences* **114,** E4158–E4167 (2017).

311. Nury, H., Poitevin, F., Van Renterghem, C., Changeux, J.-P., Corringer, P.-J., Delarue, M. & Baaden, M. One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue. *Proceedings of the National Academy of Sciences* **107,** 6275–6280 (2010).

312. Ruan, Y., Kao, K., Lefebvre, S., Marchesi, A., Corringer, P.-J., Hite, R. K. & Scheuring, S. Structural titration of receptor ion channel GLIC gating by HS-AFM. *Proceedings of the National Academy of Sciences* **115,** 10333–10338 (2018).

313. Bharambe, N., Li, Z., Seiferth, D., Balakrishna, A. M., Biggin, P. C. & Basak, S. Cryo-EM structures of prokaryotic ligand-gated ion channel GLIC provide insights into gating in a lipid environment. *Nature Communications* **15,** 2967 (2024).

314. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation* **7,** 525–537 (2011).

315. Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *Journal of Chemical Theory and Computation* **7,** 2284–2295 (2011).

316. Yazdani, M., Jia, Z. & Chen, J. Hydrophobic dewetting in gating and regulation of transmembrane protein ion channels. *The Journal of Chemical Physics* **153,** 110901 (2020).

317. Bocquet, N., Prado de Carvalho, L., Cartaud, J., Neyton, J., Le Poupon, C., Taly, A., Grutter, T., Changeux, J.-P. & Corringer, P.-J. A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature* **445,** 116–119 (2007).

318. Welch, B. L. The Generalization of 'Student's' Problem when Several Different Population Variances are Involved. *Biometrika* **34,** 28–35 (1947).

319. Koenig, F. & Schmidt, M. Gloeobacter violaceus— investigation of an unusual photosynthetic apparatus. Absence of the long wavelength emission of photosystem I in 77 K fluorescence spectra. *Physiologia Plantarum* **94,** 621–628 (1995).

320. Aho, N., Buslaev, P., Jansen, A., Bauer, P., Groenhof, G. & Hess, B. Scalable Constant pH Molecular Dynamics in GROMACS. *Journal of Chemical Theory and Computation* **18,** 6148–6160 (2022).

321. Miller, P. S. & Smart, T. G. Binding, activation and modulation of Cys-loop receptors. *Trends in Pharmacological Sciences* **31,** 161–174 (2010).

322. Philip, A. B., Brohan, J. & Goudra, B. The Role of GABA Receptors in Anesthesia and Sedation: An Updated Review. *CNS Drugs* **39,** 39–54 (2025).

323. Gil, S. M. & Metherate, R. Enhanced Sensory–Cognitive Processing by Activation of Nicotinic Acetylcholine Receptors. *Nicotine & Tobacco Research* **21,** 377–382 (2019).

324. Koukouli, F. & Changeux, J.-P. Do Nicotinic Receptors Modulate High-Order Cognitive Processing? *Trends in Neurosciences* **43,** 550–564 (2020).

325. San Martín, V. P., Sazo, A., Utreras, E., Moraga-Cid, G. & Yévenes, G. E. Glycine Receptor Subtypes and Their Roles in Nociception and Chronic Pain. *Frontiers in Molecular Neuroscience* **15** (2022).

326. Fagerlund, M. J. & Eriksson, L. I. Current concepts in neuromuscular transmission. *British Journal of Anaesthesia* **103,** 108–114 (2009).

327. Richardson, R. J., Petrou, S. & Bryson, A. Established and emerging GABAA receptor pharmacotherapy for epilepsy. *Frontiers in Pharmacology* **15** (2024).

328. Bakker, M. J., Dijk, J. G. v., Maagdenberg, A. M. v. d. & Tijssen, M. A. Startle syndromes. *The Lancet Neurology* **5,** 513–524 (2006).

329. Hibbs, R. E. & Gouaux, E. Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474,** 54–60 (2011).

330. Adelsberger, H., Lepier, A. & Dudel, J. Activation of rat recombinant alpha(1)beta(2)gamma(2S) GABA(A) receptor by the insecticide ivermectin. *European journal of pharmacology* **394,** 163–170 (2000).

331. Krause, R. M., Buisson, B., Bertrand, S., Corringer, P.-J., Galzi, J.-L., Changeux, J.-P. & Bertrand, D. Ivermectin: A Positive Allosteric Effector of the α7 Neuronal Nicotinic Acetylcholine Receptor. *Molecular Pharmacology* **53,** 283–294 (1998).

332. Das, P., Bell-Horner, C. L., Machu, T. K. & Dillon, G. H. The GABAA receptor antagonist picrotoxin inhibits 5-hydroxytryptamine type 3A receptors. *Neuropharmacology* **44,** 431–438 (2003).

333. Wang, D.-S., Mangin, J.-M., Moonen, G., Rigo, J.-M. & Legendre, P. Mechanisms for Picrotoxin Block of α2 Homomeric Glycine Receptors *. *Journal of Biological Chemistry* **281,** 3841–3855 (2006).

334. Barann, M., Dilger, J. P., Bönisch, H., Göthert, M., Dybek, A. & Urban, B. W. Inhibition of 5-HT3 receptors by propofol: equilibrium and kinetic measurements. *Neuropharmacology* **39,** 1064–1074 (2000).

335. Barann, M., Linden, I., Witten, S. & Urban, B. W. Molecular actions of propofol on human 5-HT3A receptors: enhancement as well as inhibition by closely related phenol derivatives. *Anesthesia and analgesia* **106,** 846–57, table of contents (2008).

336. Fodale, V. & Santamaria, L. B. Different actions of sevoflurane and propofol on central nicotinic receptors may explain differences in hypnotic antagonism by cholinesterase inhibitors. *British Journal of Anaesthesia* **92,** 773–775 (2004).

337. Zeilhofer, H. U., Werynska, K., Gingras, J. & Yévenes, G. E. Glycine Receptors in Spinal Nociceptive Control—An Update. *Biomolecules* **11,** 846 (2021).

338. O'Brien, B. C. V., Weber, L. & Weltzin, M. M. Design and optimization of nicotinic acetylcholine receptor subtype selective peptides. *Biophysical Journal* **122,** 395a (2023).

339. Nemecz, Á., Prevost, M. S., Menny, A. & Corringer, P.-J. Emerging Molecular Mechanisms of Signal Transduction in Pentameric Ligand-Gated Ion Channels. *Neuron* **90,** 452–470 (2016).

340. Rudolph, U. & Knoflach, F. Beyond classical benzodiazepines: novel therapeutic potential of GABAA receptor subtypes. *Nature Reviews Drug Discovery* **10,** 685–697 (2011).

341. Sigel, E. & Ernst, M. The Benzodiazepine Binding Sites of GABAA Receptors. *Trends in Pharmacological Sciences* **39,** 659–671 (2018).