THE UNIVERSITY OF
SYDNEY

Computational Studies of Glutamate Transporters and Receptors

Germano Heinzelmann

THESIS SUBMITTED
IN FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

Doctor of Philosophy
IN
Physics

FACULTY OF SCIENCE

UNIVERSITY OF SYDNEY

2014

Germano Heinzelmann, Author — Date

Serdar Kuyucak, Supervisor — Date

Abstract

The creation and transmission of nerve impulses involve a variety of membrane proteins such as secondary transporters, receptors and ion channels. These structures are related to a number of neurodegenerative and auto-immune diseases, and the investigation of their properties may bring an important contribution to the treatment of such illnesses. In this thesis we use molecular dynamics simulations to study glutamate transporters and receptors, as well as toxin binding to human potassium channels. We apply a variety of techniques in computational biology such as ligand docking, parameterization of molecules, homology modeling, free energy perturbation and potential of mean force.

We start by studying in detail the archaeal aspartate transporter $\mathrm{Glt_{Ph}}$, which has crystal structures available and is a homolog of the human glutamate transporters. We first locate the third sodium ion binding site, which is not present in the crystal data, and show that this site is also conserved in the EAATs. We then perform molecular dynamics simulations in the outward and inward facing states of $\mathrm{Glt_{Ph}}$, calculating the ligand affinities in both conformations and the order of ligand binding/unbinding. We also investigate the nature of the aspartate-sodium coupling in these transporters, as well as the differences in the gating mechanism between the outward and inward states. From the $\mathrm{Glt_{Ph}}$ structure, we build models for the human glutamate transporters (EAATs), for which there are no crystal structures. Besides the ligands present in $\mathrm{Glt_{Ph}}$, the EAATs also bind a proton and a potassium during the transport cycle. We use our models to attempt to locate the potassium and proton binding sites, using free energy perturbation calculations to locate the potassium site and pK_a calculations to explain the mechanism of proton transport. We also use molecular dynamics simulations to investigate the gating differences between the archaeal and human transporters.

The glutamate receptor GluA2 binds to a variety of agonists and antagonists, with high-resolution structures of the complexes. It plays an essential role in the transmission of nerve impulses, also serving as a target to potential drugs. We calculate the standard binding free energy of five small molecules to the ligand binding domain of this receptor, using a free energy perturbation approach that produces results with very good agreement to experimental values. We show that this method is effective for polar and charged ligands, and can be applied to many biological systems, providing an useful tool in computer aided drug design.

Finally, we develop a method based on free energy perturbation that can be used to calculate the binding free energy differences of peptide toxin mutants to Kv potassium channels, producing very good agreement to potential of mean force and experimental results. We use it to obtain a ShK mutant that is selective for the Kv1.1 channel over Kv1.3, and therefore may be used in the treatment of autoimmune diseases.

Acknowledgments

I would like to give special thanks to my supervisor, Prof. Serdar Kuyucak, for giving me this great opportunity in Sydney and providing all the support I needed during my research. I would also like to thank Dr. Po-Chia Chen, for invaluable lessons on molecular dynamics simulations in the first years of my degree. Thanks to Renae Ryan and Robert Vandenberg, for their experimental collaboration and discussions regarding the glutamate transporters. I also extend my gratitude to my other co-workers Harunur Rashid, Turgut Bastug, Jeff Timko and Sahar Mahdavi.

Finally my parents, Ernesto and Elisabeth Heinzelmann, who always believed in me and supported me in all my endeavours.

Statement of originality

Chapters 1 is a review of the general aspects of the systems we investigate in this thesis, as well as their biological relevance. Chapter 2 explains the physical and mathematical background behind the methods we utilize. Chapters 3-9 are original works from a collection of published papers, except for chapter 7 which is an article currently under review. The conclusions and future work perspectives are outlined in Chapter 10. The bibliographical information of all the articles, as well as my contribution to each of them, is listed below:

Chapter 3: **Position of the Third Na^+ Site in the Aspartate Transporter Glt_{Ph} and the Human Glutamate Transporter, EAAT1**

Turgut Bastug[1,2], **Germano Heinzelmann**[1], Serdar Kuyucak[1], Marietta Salim[3], Robert J. Vandenberg [3], Renae M. Ryan[3]

[1] *School of Physics, The University of Sydney, Sydney, Australia*
[2] *Faculty of Arts and Sciences, TOBB University of Economy and Technology, Ankara, Turkey*
[3] *Transporter Biology Group, Discipline of Pharmacology, School of Medical Sciences and Bosch Institute, The University of Sydney, Sydney, Australia*

Publication Details: Published in *PLoS ONE*, 2012, Volume 7, Issue 3, e33058.

Contribution: I was one of the main contributors to this work, initially helping on the search for the Na3 ion using molecular dynamics. I pointed out the Serine-93 residue of Glt_{Ph} as coordinating Na3, which was later proved experimentally. I have also performed most of the MD simulations and FEP/TI calculations present in the paper, as well as being an integral part of the interpretation and discussion of the results.

Chapter 4: **Free Energy Simulations of Ligand Binding to the Aspartate Transporter Glt_{Ph}**

Germano Heinzelmann[1], Turgut Bastug[1,2], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*
[2] *Faculty of Arts and Sciences, TOBB University of Economy and Technology,*

Ankara, Turkey

Publication Details: Published in *Biophysical Journal*, 2011, Volume 101, 2380–2388

Contribution: I was primarily responsible for this work. I performed most of the MD simulations and FEP/TI calculations, as well as analysing and interpreting the results obtained.

Chapter 5: **Mechanism and Energetics of Ligand Release in the Aspartate Transporter Glt_{Ph}**

Germano Heinzelmann[1], Turgut Bastug[2], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*
[2] *Department of Materials Science and Nanotechnology Engineering, TOBB University of Economy and Technology, Ankara, Turkey*

Publication Details: Published in *Journal of Physical Chemistry B*, 2013, Volume 117, 5486–5496

Contribution: I was primarily responsible for this work. I created the systems, and performed all of the MD simulations together with the FEP/TI calculations. I also analysed and interpreted the results obtained.

Chapter 6: **Molecular Dynamics Simulations of the Mammalian Glutamate Transporter EAAT3**

Germano Heinzelmann[1], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*

Publication Details: Published in *PLoS ONE*, 2014, Volume 9, Issue 3, e92089.

Contribution: I was primarily responsible for this work. I reviewed the literature, created the models, performed all of the MD simulations and FEP/TI calculations. I also analysed and interpreted the results obtained. .

Chapter 7: **Molecular Dynamics Simulations Elucidate the Mechanism of Proton Transport in the Glutamate Transporter EAAT3**

Germano Heinzelmann[1], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*

Publication Details: Submitted to *Biophysical Journal*.

Contribution: I was primarily responsible for this work. I implemented the method for the pK_{a} calculations, using the thermodynamic cycle of Fig. 1. I also created the models, performed all of the MD simulations and FEP/TI calculations, analysed and interpreted the results obtained.

Chapter 8: **Computation of Standard Binding Free Energies of Polar and Charged Ligands to the Glutamate Receptor GluA2**

Germano Heinzelmann[1], Po-Chia Chen[2], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*
[2] *Computational Molecular Biophysics Group, Department of Molecular Structural Biology, Georg-August-University, 37073 Göttingen, Germany*

Publication Details: Published in *Journal of Physical Chemistry B*, 2014, Volume 118, 1813–1824

Contribution: I was primarily responsible for this work. I implemented the method for the standard binding free energy calculations, using the thermodynamic cycle of Fig. 1, as well the the confine-and-release approach. I parametrized the CNQX and AMPA molecules, performed all of the FEP/TI/PMF calculations, and did the analysis and interpretation of all the results.

Chapter 9: **A Potent and Selective Peptide Blocker of the Kv1.3 Channel: Prediction from Free-Energy Simulations and Experimental Confirmation**

M. Harunur Rashid[1], **Germano Heinzelmann**[1], Redwan Huq[2,3], Rajeev B. Tajhya[2,3], Shih Chieh Chang[4], Sandeep Chhabra[4], Michael W. Pennington[5], Christine Beeton[2], Raymond S. Norton[4], Serdar Kuyucak[1]

[1] *School of Physics, The University of Sydney, Sydney, Australia*
[2] *Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas, United States of America*
[3] *Graduate Program in Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas, United States of America*
[4] *Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria, Australia*
[5] *Peptides International, Louisville, Kentucky, United States of America*

Publication Details: Published in *PLoS ONE*, 2013, Volume 8, Issue 11, e78712

Contribution: I was one of the main contributors to this work. I devised and implemented the FEP/TI method to calculate the selectivity of ShK mutants, depicted in the thermodynamic cycle of Fig. 1. I performed all of the TI calculations, and was involved in the interpretation and discussion of the paper results. The electrophysiological experiments presented in this study were performed by Ray Norton's group at Monash University.

Contents

List of Tables

List of Figures

List of Abbreviations

Kv	**Voltage-gated potassium channel**
PMF	**Potential of mean force**
VMD	**Software for visualization of molecular dynamics.**
MD	**Molecular dynamics**
$NAMD$	**Software for nanoscale molecular dynamics**
$POPE$	**1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine**
$EAAT$	**Excitatory amino acid transporter**
$MODELLER$	**Software to build homology models**
NpT	**Constant number of particles, pressure and temperature**
NVT	**Constant number of particles, volume and temperature**
PDB	**Protein data bank**
$RMSD$	**Root mean square deviation**
PME	**Particle mesh Ewald, method to calculate electrostatics.**
$CMAP$	**Backbone correction terms in CHARMM**
$CHARMM$	**Chemistry at Harvard macromolecular mechanics**
$HADDOCK$	**A protein-protein docking program**
COM	**Center of mass**
FEP	**Free energy perturbation**
TI	**Thermodynamic Integration**
LJ	**Lenard-Jones potential**
$WHAM$	**Weighted histogram analysis method**

Chapter 1

Introduction

1.1 Nerve impulses

The creation and transmission of nerve impulses play a central role in neuron-to-neuron communication, allowing information to be extracted and processed within the nervous system. These impulses rely on the polarity of the lipid membrane, and its ability to generate and propagate action potentials[1]. A typical neuron has a resting potential of around -70 mV compared to the extracellular side. In an action potential, the voltage across the membrane suddenly reverses, making the inside positive relative to the outside, and then abruptly reverses again, after which the resting potential is restored. The changes in voltage that produce an action potential are caused by a brief, large influx of sodium ions and a brief, large efflux of potassium ions (Figure 1.1A). This flux is controlled by voltage-sensitive ion channels, which have the ability to selectively allow or block the passage of a particular ion through the cell membrane. They also rely on concentration gradients, which are maintained by active transporters such as the sodium and potassium pumps. The action potential allows a nerve impulse to travel through the neuron axon until it reaches another nerve cell, where there is a synapse.

A synapse is the interface between two neurons, where the nerve impulse can be transmitted electrically or chemically. Electrical synapses, altough a minority, are found in all nervous systems, including the human brain. Here two communicating neurons come extremely close and are linked together by an intercellular specialization called a gap junction[2]. These synapses are extraordinarily fast, since passive current flow across the gap junction is virtually instantaneous. Chemical synapses, on the other hand, are dependent on the release and recognition of neurotransmitters. Information is transmitted across a chemical synapse in four basic steps (Figure 1.1B):

1. Neurotransmitter molecules are synthesized and stored in vesicles in the axon terminal.
2. Neurotransmitter vesicles are transported to the presynaptic membrane and neu-

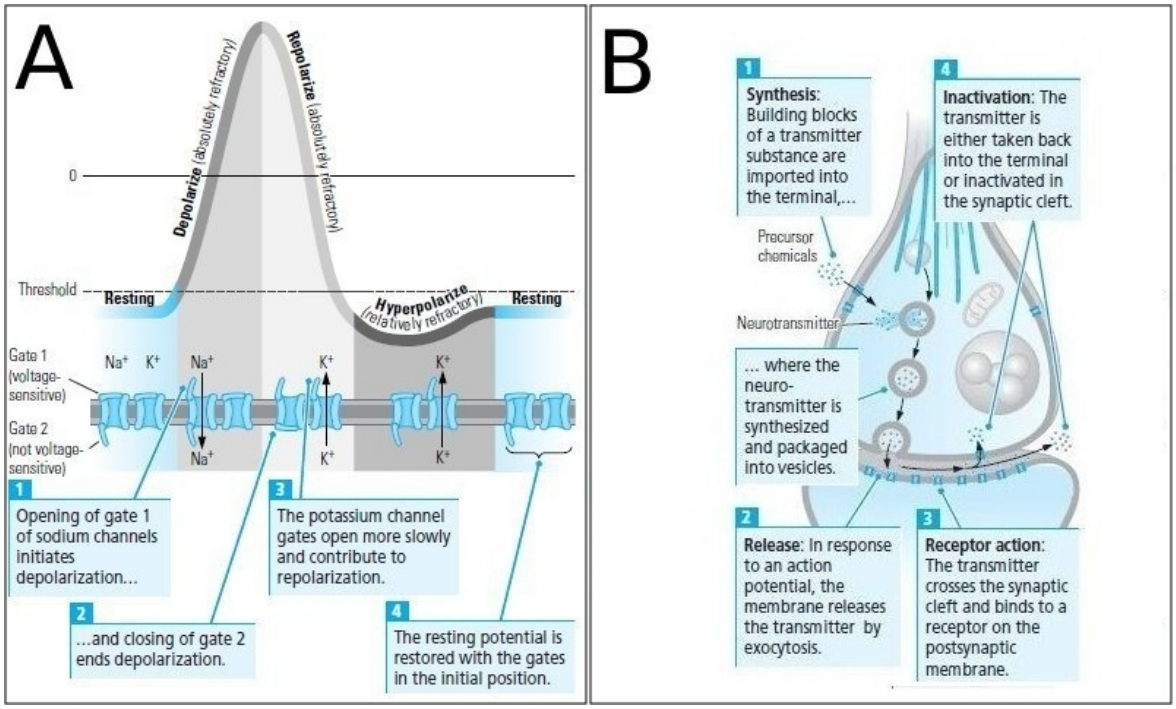

Figure 1.1: (A) The steps involved in the transmission of the action potential through the membrane, in which the potassium and sodium channels open in response to an applied voltage. (B) The four stages in the release of the neurotransmitter in the nerve synapses, from its synthesis in the presynaptic cell to its inactivation by enzymes and membrane transporters. Figure adapted from Ref.[1]

rotransmitters are released in response to an action potential.
3. Neurotransmitters interact with the receptors on the membrane of the target cell located on the other side of the synapse.
4. Neurotransmitters are inactivated (or they will continue to work indefinitely).

These steps depend on a number of cell structures like enzymes, RNA, lipid vesicles, calcium channels, transmitter receptors and transporters.

Despite the great variety of chemical synapses, neurotransmitters can be divided in two classes: excitatory or inhibitory. In other words, a neurotransmitter either increases or decreases the probability that the neuron with which it comes to contact will produce an action potential. The most important excitatory neurotransmitter in the human brain is the amino acid glutamate, and the glutamate receptor proteins play a major role in step 3 listed above[3,4]. They can be divided into metabotropic and ionotropic, depending on their mechanism[5]. Metabotropic glutamate receptors induce an indirect biochemical process, like the activation of biochemical cascades and the modification of other proteins. The ionotropic glutamate receptor, on the other hand, couple the binding of the glutamate molecule to the opening of an ion channel, which is fundamental for neurotransmission at the nerve synapses[6]. They

also serve as potential drug targets for the treatment of numerous neurodegenerative conditions. There are several ligands that are known to bind to these receptors, either as agonists or antagonists. In the past years, crystal structures of the ligand binding domains (LBD)[7,8] and amino terminal domains (ATD)[9,10] from different ionotropic receptors were determined. More recently, the crystal structure of the full-length eukaryotic AMPA receptor (iGluA2) was reported[11], shedding light into the mechanism of substrate recognition and the conformational changes happening upon substrate binding. In this thesis we study the free energy of binding of five ligands to the GluA2 receptor LBD, presenting a method to accurately calculate the binding free energy of polar and charged ligands using free energy perturbation (FEP). This protocol can be applied to many systems, and we present it as a good alternative for rational drug design.

Glutamate transporters, known as Excitatory Amino Acid Transporters (EAATs) in mammals, are membrane proteins responsible for clearing the excess glutamate released at the synapses after the transmission of the nerve impulse (step 4 above)[12]. Extracellular glutamate can be toxic to neurons in high concentrations, and malfunctioning of glutamate transporters has been implicated in many pathological conditions including cerebral ischemia, amyotrophic lateral sclerosis and Alzheimer's disease. The EAATs are sodium- and potassium-dependent secondary transporters, and they use the concentration gradient of these two ions to transport glutamate into the intracellular media. They cotransport 3 Na^+ ions and a proton with the substrate, and countertransport one K^+ ion at each transport cycle[13]. Five EAATs have been identified in humans, called EAAT1-EAAT5. Because they are embedded in a lipid environment, the EAATs are very difficult to crystallize, and therefore no crystal structures of these transporters are available. To overcome this limitation, the crystal structure of the homologous archaeal aspartate transporter from *Pyrococcus Horikoshii*, called Glt_{Ph}, was determined in 2004[14]. Further crystal structures of this transporter, identifying the sodium and substrate binding sites, as well as different conformations of the protein during the transport cycle, have been unveiled since then[15,16]. Even though Glt_{Ph} has around 36% homology to the EAATs and its transport mechanism does not involve potassium or a proton, the substrate binding site and the tertiary/quaternary structure of Glt_{Ph} shows very good agreement to previous experimental studies on the EAATs. This similarity opened up the possibility of producing EAATs homology models based on Glt_{Ph} to understand their transport mechanism. This is the main focus of this thesis — we perform computational studies on Glt_{Ph} and in EAAT3 homology models. Our goal is to identify the binding sites of the substrate, proton and ions; the mechanism of ligand binding and release; and the local conformational changes happening in the transporter upon ligand binding/unbinding.

The voltage-gated potassium channels (Kv), besides their importance for the transmission of the action potential through the neuron membrane, are also essential for maintaining cellular homeostasis and contribute to numerous biological process in non-excitable cells[17]. Peptide toxin blockers of Kv1 potassium channels, such as the

ones found in sea anemone and scorpion, may be good alternative for the treatment of auto-immune diseases because the blocking of potassium channel Kv1.3 inhibits T cell proliferation[18]. Since the Kv1.1 channel is present in the central nervous system, Kv1.3/Kv1.1 selectivity is necessary for the use these toxins as therapeutic drugs. Computational studies are able to find mutations in these toxins that can increase this selectivity, by performing docking and free energy calculations. The free energy methods we use to investigate the glutamate transporters and receptors were applied to these channel-toxin complexes, and the results are also presented in this thesis.

1.2 Glutamate transporters

L-glutamate is the major excitatory neurotransmitter in the human brain, and therefore this molecule is highly abundant in the central nervous system. Even though the concentration of glutamate in the brain is estimated to be around 5–15 mM per wet net weight, the extracellular concentration of this neurotransmitter is in the order of 3–4 μM[12]. This gives a several-thousand fold glutamate concentration gradient between the extracellular and intracellular media, and there is a number of reasons for this discrepancy. First, it is essential to keep a high signal-to-noise ratio for synaptic transmissions. If glutamate is not quickly removed from the synaptic cleft after receptor activation, the postsynaptic cell could not respond to subsequent messages that the presynaptic cell might send. Second, glutamate can be reused once it is transported back to the intracellular media, by accumulating into vesicles that are subsequently used for neurotransmitter release. Third, large extracellular concentrations of glutamate have been associated with neuronal death due to excessive activation of glutamate receptors. There is a scarcity of extracellular enzymes that can metabolize glutamate, leaving cellular uptake as the only quick way to remove this neurotransmitter from the nerve synapses. There are many different glutamate transporters characterized so far, having distinct mechanisms and locations within the cell tissue. Some are sodium and potassium-dependent, which is the case of the EAATs, the main focus of this thesis. There are also glutamate transporters that work by exchanging glutamate with another amino acid like cysteine or glycine, which are sodium independent. In mitochondria, glutamate can be taken up by proton-glutamate symporters or aspartate-glutamate antiporters. Finally, there are transporters that work inside neurons by carrying glutamate into the synaptic vesicles, using the membrane potential and an inward concentration gradient of protons as driving force.

Due to their high concentration on glial cells, the EAATs play an important role in glutamate uptake from synapses. Five subtypes have been identified in humans, called EAAT1-5. Within a single species, they share 50–60% amino acid sequence identity with each other, and also the same transport mechanism. The EAATs are secondary transporters, using as a driving force the concentration gradients of Na^+

and K^+, which are directed to the outside and inside the cell, respectively. Malfunctioning in the EAATs has been associated with numerous neuronal diseases, like ischemia, amyotrophic lateral sclerosis (ALS), Alzheimer's disease and epilepsy. In the case of ischemia, it has been observed that large amounts of glutamate are released into the extracelullar fluid when the energy supply of the brain is compromised. Studies have shown that this release might be related to the reversed operation of the glutamate transporters, which may cause the collapse of ion gradients due to overactivation of glutamate receptors[19]. Patients with ALS show an increased glutamate concentration in cerebrospinal fluid, as well as a reduced expression of EAAT2 transporters, which might be related to the excitotoxic mechanisms in ALS[20]. Reduced expression of the EAAT2 transporters, or the occurrence of splice variants, has also been linked to the onset of Alzheimer's disease, and the inhibition of glutamate transport might aggravate this condition[21]. Even though there is no direct evidence that defective glutamate uptake is related to epilepsy, increased levels of glutamate are observed in epileptic patients prior to the onset of seizures[22]. Furthermore, reduced expression of glutamate transporters in mice brings the development of epilepsy and neurodegeneration[23].

Many mutagenesis and functional experiments have been performed on human EAATs in the past years, as well as on its rodent homologs. These studies have revealed the main features of the tertiary and quaternary structures, the transport stoichiometry, and the residues involved in ligand binding and selectivity. Hydrophobicity profiles initially revealed that the EAATs consist of six transmembrane segments in the N-terminal, organized as α-helices[24]. Two additional transmembrane segments and two reentrant loops were identified by experiments with cysteine-substituted residues[25,26], based on their interaction with the biotinylated sulfhydryl reagent 3-*N*-maleimidyl(propionyl)biocytin (BM), both from the outside and inside of the cell (Figure 1.2). Furthermore, immunocytochemical data showed that both the C- and N-termini are located intracellularly[27]. Regarding the transport stoichiometry, functional studies show that, at each transport cycle, three sodium ions and one proton are co-transported with glutamate, and one potassium ion is countertransported[13] (Figure 1.3). The binding of the substrate is coupled to sodium, since glutamate cannot bind in the absence of this ion. At each half-cycle, the whole binding site is translocated across the membrane, with either $Glu/3Na^+/H^+$ or K^+ bound to the transporter. This process is electrogenic, with a net charge of +2 moving into the cell, and therefore can be increased by a negative membrane potential. In the absence of potassium, or if the binding of this ion is impaired, the EAATs work as exchangers, completing only the steps 1–6 in Figure 1.3. By increasing the potassium concentration in the extracellular media, these transporters can also work in reverse mode, thereby leaking glutamate from the cell[28].

In all EAATs, there is also an anion conductance that is thermodinamically independent from the transport process, but is coupled to substrate and sodium binding. This conductance is more pronounced for larger anions such as SCN^- and NO_3^-, and less for Cl^- and F^-[29]. This feature of the EAATs is very important

for functional experiments, since the SCN^- and NO_3^- conductances can be used to probe the binding of different ligands to the these transporters[30,31]. Regarding the residues involved in ligand binding, mutagenesis experiments point to an arginine and an aspartate residues interacting with the substrate α-amino and side-chain carboxyl groups, respectively[32,33]. A serine residue located in the HP2 segment is involved in the sodium/lithium selectivity in EAAT2, as well as the ability of lithium to support transport[34]. Aspartate residues close to the substrate binding site are shown to be involved in the binding of sodium and potassium to the EAATs, since their neutralization affects the binding properties of these ions[30].

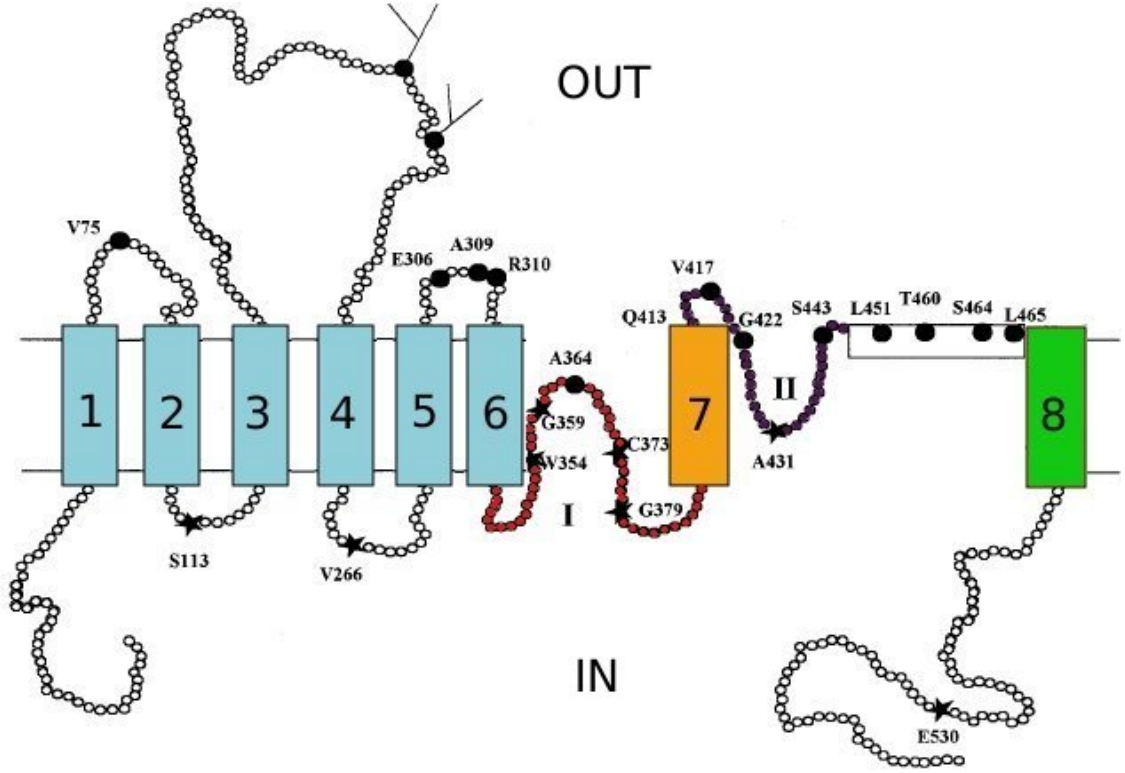

Figure 1.2: (A) The topology of the EAAT1 transporter obtained by hydropathy profiles and cysteine substitution experiments. The transmembrane domains are indicated by Arabic numerals, and the reentrant loops are indicated by roman numerals. Residues that react with 3-*N*-maleimidyl(propionyl)biocytin (BM) from the outside (filled circles) or from the inside (stars) are indicated by their position number.

An important step in our understanding of the EAATs came in 2004, with the first crystal structure of the sodium-coupled archaeal transporter Glt_{Ph}, which shares an overall 36% sequence similarity with the human glutamate transporters[14]. In contrast to the EAATs, Glt_{Ph} does not require the cotransport of a proton or the countertransport of a potassium ion for the completion of the transport cycle, and is also selective for aspartate[35]. Despite the relatively low homology, Glt_{Ph} shows structural features that are very similar to the predicted structure of the EAATs. The 2004

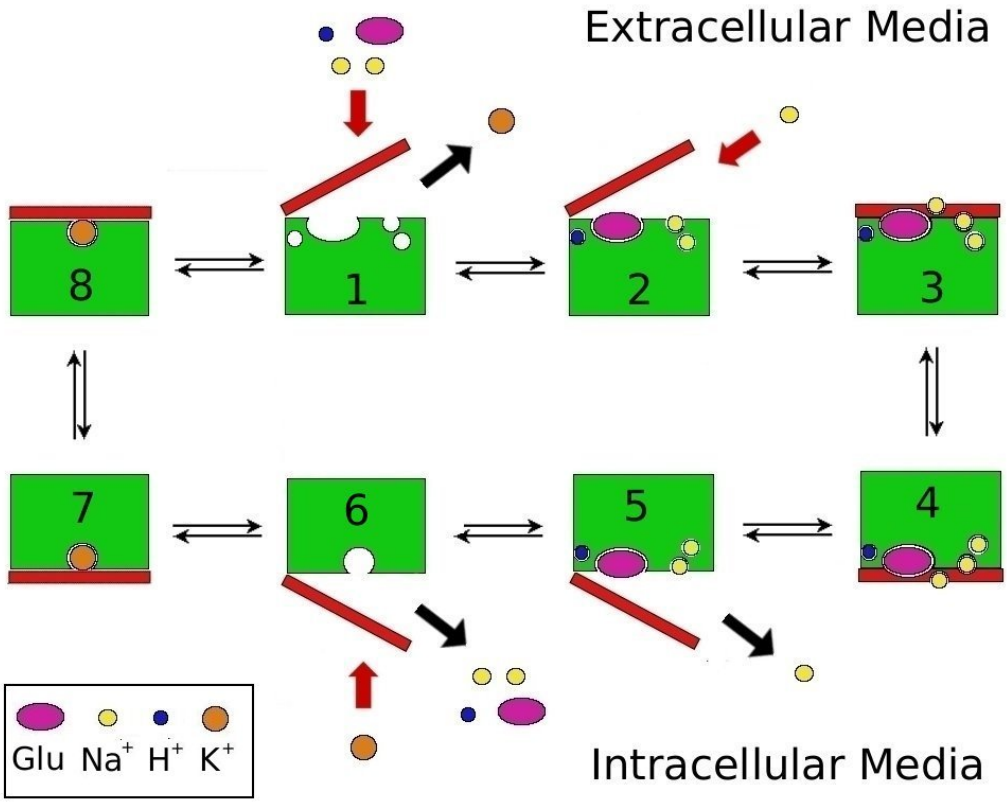

Figure 1.3: Cartoon showing the functional mechanism of the EAAT transporters. Steps 3-4 denote the translocation of the binding site across the membrane with $3Na^+$, H^+, and glutamate bound, while steps 7-8 denote the same with only K^+ bound. Step 2 shows the last ion to bind to the outward-facing transporter, which happens after the binding of substrate. Step 5 shows the opposite happening in the intracellular media, assuming that the binding and unbinding of ligands are symmetrical.

structure has the archaeal transporter in the extracellular(EC)-facing *holo* state of Glt_{Ph}, with the substrate and an unknown number of sodium ions bound. Glt_{Ph} is assembled as a trimer, with three identical subunits connected by non-covalent bonds (Figure 1.4A). Each monomer has 8 transmembrane segments, called TM1 to TM8, and 2 reentrant loops, or hairpins, called HP1 and HP2. HP1 is sequentially located between TM6 and TM7, and HP2 between TM7 and TM8 (Figure 1.4B). A subsequent crystal structure from 2007, also in the outward-facing state, identified the binding sites from the substrate and two Na^+ ions, called Na1 and Na2[15]. It revealed that these sites are formed mainly by residues located in segments HP1, TM7, HP2 and TM8, with this region having an homology of around 60% between Glt_{Ph} and the EAATs. Furthermore, almost all residues involved in ligand binding are similar in these two transporters. By using the non-transportable blocker TBOA as a substrate, the same study identified the HP2 segment as the gate that opens/closes the aspartate binding site to the solvent in the EC-facing state. Molecular dynamics studies have related the opening of the HP2 gate to substrate and Na2 binding, and proposed that the latter binds after the substrate and stabilizes the transporter in the closed conformation[36]. Even though only two sodium ions are present in the 2007 crystal structure, radioactive-labeled experiments have shown that a third sodium ion is cotransported with the substrate in Glt_{Ph}[37]. Molecular dynamics simulations and mutagenesis experiments have located the binding site for this third sodium, called Na3, which is also conserved in the EAATs[38]. They concluded that Na3 should be the first ligand to bind, followed by Na1, aspartate and Na2. The mechanism of substrate translocation was unknown until the release of a crystal structure in the closed intracellular(IC)-facing state of the Glt_{Ph} transporter, also with 2 sodium ions and the substrate bound[16]. Unfortunately, the gating mechanism for the IC-state could not be determined due to the lack of substrates that would stabilize this structure in the open conformation. It has been initially suggested that HP1 would work as the IC gate, but molecular dynamics studies have shown that both HP1 and HP2 are involved in the opening/closure of binding site in this case[39]. More recently, an asymmetric trimer of Glt_{Ph} with an intermediate state between the outward and inward conformations was also resolved, which will help solving the mechanism of substrate translocation across the membrane[40].

The main focus of this thesis is using computational techniques to investigate the properties of the Glt_{Ph} transporter, such as the location of the Na3 binding site, the mechanism of binding and release of the different ligands, as well as the gating properties in the outward and inward states of Glt_{Ph}. Once we characterize the archaeal transporter, we use it as a template to create homology models for the human glutamate transporter EAAT3. By performing molecular dynamics simulations on EAAT3, we investigate the coordination of the different ligands, and their conservation between Glt_{Ph} and EAAT3. We also attempt to locate the proton and the potassium binding sites, which are not present in Glt_{Ph}, as well as identifying the gating differences between the two transporters. Finally, we use rigorous pK_a calculations to unveil the mechanism of proton transport in the EAATs, and how

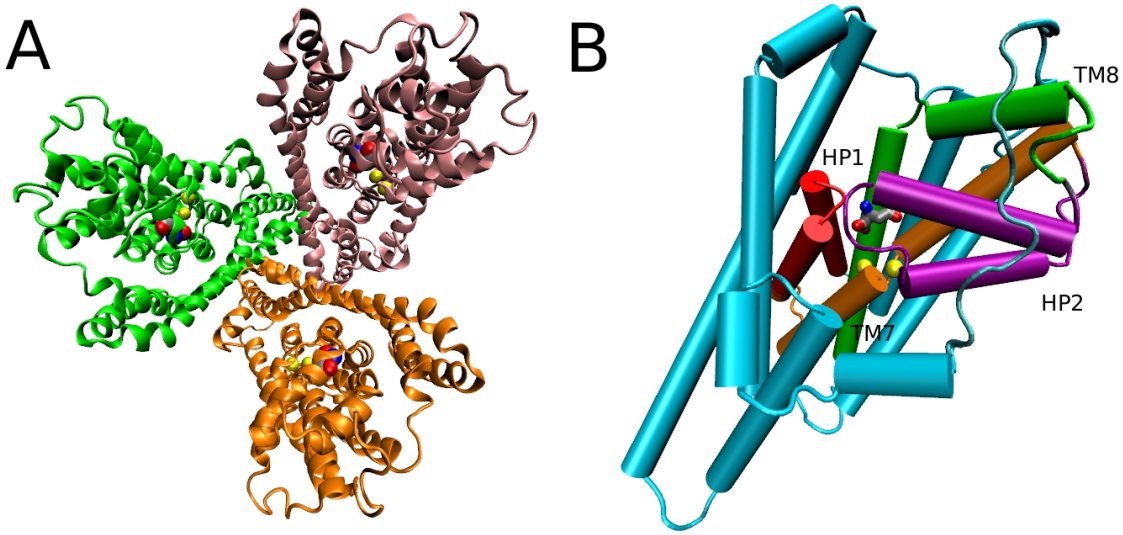

Figure 1.4: (A) The crystal structure of the Glt_{Ph} trimer, with the substrate and two sodium ions bound in each chain (PDB ID: 2NWX) (B) An individual monomer from the same structure, showing the transmembrane segments in the same coloring scheme of Fig. 1.2. HP1 is colored red, TM7 colored orange, HP2 purple and TM8 green.

that is related to the binding and unbinding of the other ligands to the transporter.

1.3 Ionotropic glutamate receptors

The ionotropic glutamate receptors are ligand-gated ion channels, coupling the interaction with glutamate to the opening of an ion channel in the plasma membrane[6]. These receptors can be divided in three major categories, according to their selectivity for the agonists AMPA (GluA1-GluA4), NMDA (GluN1, GluN2A–GluN2D,GluN3A–GluN3B) and kainate (GluK1-GluK5). As mentioned in the previous section, overactivation of the glutamate receptors can be related to several neurodegenerative disorders like ischemia and epilepsy. In addition, these proteins are potential targets for therapeutic drugs for neuropathic pain, Parkinson's disease, and schizophrenia. Due to the critical role that NMDA and AMPA receptors play on learning and memory, they can also serve as the base for cognitive enhancement treatments.

The AMPA receptors form hetero or homotetrameric structures, and each subunit can be divided into four domains: an amino-terminal domain (ATD) that participates in subtype-specific receptor assembly, trafficking and modulation; a ligand-binding domain (LBD) central to agonist/competitive antagonist binding and to activation gating; a transmembrane domain (TMD) that forms the membrane-spanning ion channel; and a cytoplasmic carboxy-terminal domain involved in receptor localization and regulation (Figure 1.5A). There are several crystal structures

for the LBD and the ATD domains of different AMPA receptors, and recently the three dimensional structure of a full rat GluA2 receptor determined their location in the protein and relative to the membrane[11]. The LBD is organized as a tetramer, located in the extracellular media, with each monomer assembled as a clamshell structure composed of two globular domains, called S1 and S2. Agonists like AMPA and glutamate bind strongly to the LBD, bringing the opening of the coupled ion channel (TMD) and the flow of cations through the membrane. Antagonists like CNQX and DNQX, in contrast, bind to the LBD with high-affinity but block the receptor activation. This discrepancy has been linked to the conformational changes happening in the receptor LBD upon ligand binding. The distance between the S1 and S2 lobes exhibits large fluctuations in the absence of substrates, and shows different degrees of opening depending on the ligand bound (Figure 1.5B and 1.5C). The mechanism of ligand binding and the associated conformational changes in the receptor for different substrates have been investigated using MD simulations, using a variety of methods to calculate the free energies involved in the process[41,42]. It was found that agonist binding changes the conformational landscape of the receptor LBD by strongly favouring the closed conformation, and this closure accounts for most of the free energy of binding of these molecules. Antagonist binding energetically favours the open state, and the full closure of the LBD is not permitted. This difference in the LBD opening for agonists and antagonists may be transmitted to the TMD, which would explain the different response of the receptor for these two classes of substrates.

In the present thesis we calculate the standard binding free energy of polar and charged ligands to the GluA2 ligand-binding domain (S1S2J), using a free energy perturbation (FEP) method developed initially to calculate the binding affinity of aspartate to the Glt_{Ph} transporter (Chapter 4). FEP calculations often have convergence problems for charged ligands like glutamate and aspartate[43], which motivated us to develop a detailed mathematical and statistical framework to compute the standard binding free energies of such molecules. This method has great potential on rational drug design, since it enables the calculation of the affinity of small molecules to a number of different proteins. GluA2 S1S2J has high resolution structures available in complex with a variety of agonists and antagonists, as well as in the apo state. We choose for our calculations the agonists AMPA, glutamate and ACPA; and the antagonists CNQX and DNQX. There are experimental values for the binding affinities of these ligands available in the literature, which can be directly compared to the results we obtained using FEP. Ligand affinities calculated by other computational methods like umbrella sampling and metadynamics have also been reported, and therefore we can check the consistency of different free energy methods used in molecular dynamics simulations. Due to the fact that FEP calculations allow a separation of electrostatic and Lennard-Jones contributions to the binding free energies, we were also able to discuss the nature of the ligand-receptor interactions, showing that they can be mostly hydrophobic or hydrophilic depending on the substrate.

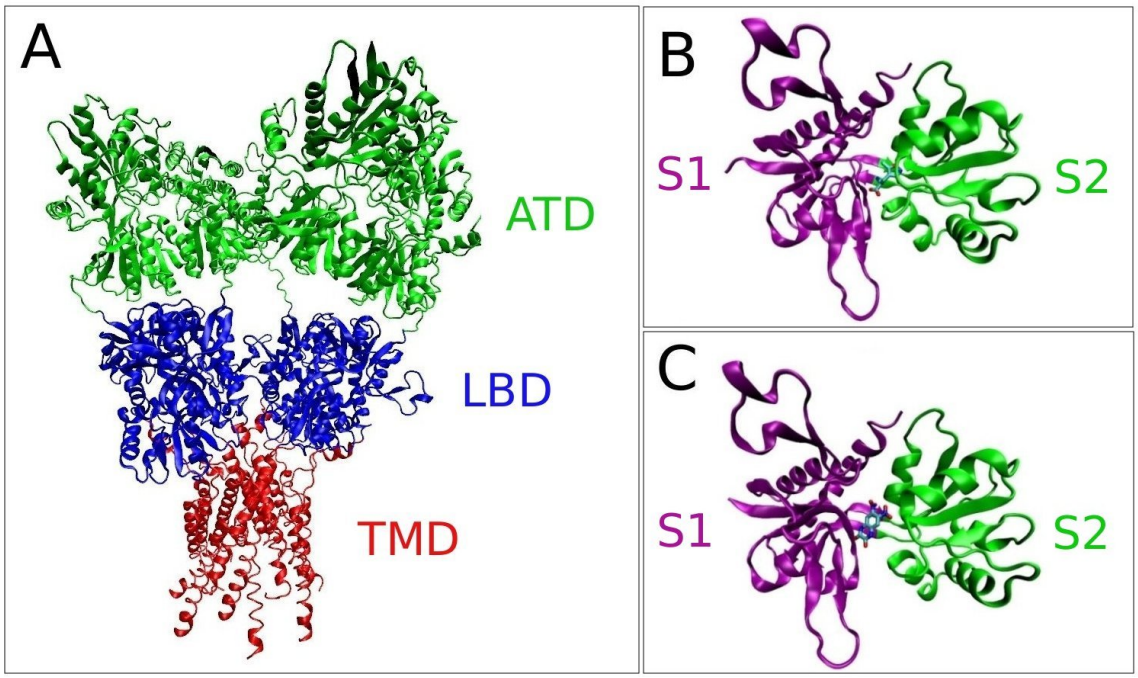

Figure 1.5: (A) The crystal structure of the full GluA2 receptor (PDB ID: 3KG2), showing the amino terminal domain (ATD), the ligand binding domain (LBD) and the transmembrane domain (TMD) (B) An individual monomer from the tetrameric LBD bound to agonist AMPA (PDB ID: 1FTM), with the S1 and S2 domains closer to each other (closed conformation). (C) The same structure now bound to the antagonist DNQX, with the S1 and S2 lobes exhibiting a larger separation between them (open conformation).

1.4 Voltage-gated potassium channels as therapeutic targets

Kv1 potassium channels, besides their importance in excitable cells such as nerve and muscle, are essential for maintaining homeostasis and controlling cell proliferation. For example, in cancer cells expression patterns of Kv1.3 and Kv1.5 are remodelled, and in some cases, a correlation has been established between protein abundance and grade of tumour malignancy[44]. The Kv1.3 potassium channels are also expressed in T and B lymphocytes. Blocking of Kv1.3 inhibits T cell proliferation, hence Kv1.3 blockers could be used in treatment of autoimmune diseases such as multiple sclerosis (MS), rheumatoid arthritis, and Type 1 and Type 2 diabetes mellitus[18]. The animal kingdom provides a great number of peptides that can target and effectively block the potassium Kv1 channels, in the form of toxins from cone snails, sea anemones and scorpions. A Kv1.3 blocker from sea anemone, called ShK, has a picomolar affinity for Kv1.3, and it has been used as a template for the development of immunosuppressant drugs in the past years[45,46]. It has also a similar affinity for Kv1.1, which is present in the central nervous system and should not be targeted. Increasing the Kv1.3/Kv1.1 selectivity, therefore, is essential in the development of new drug leads based on ShK. This can be achieved by mutating one or more residues in the toxin peptide, and more than 400 ShK analogs have been developed for this purpose.

Computational methods provide a powerful tool in this line of research, allowing identification of the binding modes and the calculation of the binding free energy for different toxin analogs. To identify the binding position of the toxin one can use a program like HADDOCK, followed by molecular dynamics simulations to refine the complex structure[47]. For the calculation of the binding free energy, the potential of mean force (PMF) is used, obtained by umbrella sampling and the weighted histogram analysis method (WHAM). This method is computationally costly, and therefore it would be interesting to develop a cheaper alternative. In this thesis we develop a method based on free energy perturbation (FEP), that can be used to calculate the binding free energy of toxin mutants in cases when the binding mode is not affected by the mutation (Chapter 9). We analyse the increase in Kv1.1/Kv1.3 selectivity when the K18A mutation is performed in ShK, and compare the results we obtain using FEP to the experimental and PMF ones.

Chapter 2

Theory and methods

Albert Einstein, in one of his famous 1905 papers, explained Brownian motion as the collision of small particles with surrounding atoms moving at high velocities, proving the atomic theory of matter[48]. It also corroborated the idea, introduced by Rudolf Clausius and James Clerk Maxwell, that macroscopic properties of a particular system like volume, temperature and pressure can be derived from the underlying mechanics of its fundamental particles. The development of quantum theory in the 20th century revealed the atomic structure, giving a mathematical description of the electron energy levels in terms of electromagnetic interactions and the discretization of energy. All these findings laid the ground for the development of Statistical Mechanics. Statistical Mechanics treats systems with a large number of particles statistically, making the connection between measurable quantities and the fundamental interactions between its constituents. It is the method of choice if we want to obtain a mathematical description of life and the processes involved in it. Even the most complex biochemical reactions can be seen as the movement and interaction of electron orbitals, defined by elements that have different properties according to the particles present in their nuclei. Mechanisms like self-assembly, essential to the formation of living cells, can be explained in terms of local electromagnetic interactions and the laws of diffusion. The intricate conformational changes observed in macromolecules, such as proteins and DNA, can be described by the free energies involved in the process, arising from electrostatic interactions and chemical bonds between individual atoms.

Ideally, one should use sophisticated quantum mechanical calculations to model all biological processes, in order to capture all the degrees of freedom and provide the most rigorous solutions. In practice this is not possible due to the complexity of systems with only a few thousand atoms, which are beyond the current capability of supercomputers if we want to achieve experimental time scales. Quantum mechanical simulations from first principles (*ab initio*), based on density-functional theory and the Car-Parrinello method[49,50], have been a large step in that direction. Nowadays it is possible to perform *ab initio* simulations on systems with a few thousand atoms in a picosecond time scale. Although suitable for small systems involving

chemical reactions, this time scale is still very short to describe the dynamics of macromolecules. Fortunately, we can approximate the quantum mechanical aspects of a system to a few classical parameters, greatly simplifying the calculations. Although this approach does not allow chemical reactions to take place, it has been widely used in the past years for investigation of a great number of systems, and is commonly known as all-atom classical molecular dynamics simulations (MD)[51]. It works by combining a force field, which assigns a few parameters for the properties of different atoms and molecules, to a numerical integration scheme, which describes the evolution of the system over time using Newton's equation of motion. The time scales that can be simulated today are in the nano and microsecond range, still far away from the time scales involved in most laboratory experiments. In many cases they do not allow for an efficient sampling of the system, and that severely limits the macroscopic properties that can be extracted from them. To overcome that limitation, there are several methods that can simulate only the relevant states for a given calculation, and that has greatly increased the success of MD calculations for phenomena like protein conformational changes and affinity for different ligands. Examples include free energy perturbation (FEP)[52], umbrella sampling[53] and metadynamics[54].

In this chapter we present a description of main aspects involved in molecular dynamics, like the force field parameters and the numerical integration method used on the time evolution of our model. We discuss the ensemble used in our simulations, how the atomic interactions are calculated, and the boundary conditions of the system. For proteins that do not have an available crystal structure, we explain how models can be built using the structure of homologous proteins. We also describe how we can use quantum mechanical calculations to obtain parameters for molecules that are not present in the force field database, as well as a docking method, which are very important for the investigation of possible drug leads. Finally, we show the physical-mathematical background of the enhanced sampling methods used in this thesis, and we apply them to our systems.

2.1 Molecular Dynamics Simulations

Classical MD simulations reduce the complexity of a quantum mechanical system to few simple parameters, like sizes and partial charges assigned to individual atoms and spring constants for chemical bonds, angles and dihedrals. They can also take into account polarization of charges, which we don't utilize in the present thesis. The time evolution of the system is done in small steps by using the Verlet algorithm, which solves Newton's equation of motion at each step using the initial coordinates and the force field parameters. Additional tools are used to calculate electrostatic interactions without truncation and to keep the system in a given temperature or pressure. The systems are built from the known crystal structures of proteins or homology models, which are inserted in a membrane or in solution using a visualiza-

tion software. In this section we discuss the physical background of our simulations, showing how the potential at each atom is calculated and how the equation of motion is solved at each time step.

2.1.1 Force Fields

There are plenty of classical force fields available in the literature, such as CHARMM[55], AMBER[56] and GROMOS[57]. We use CHARMM for all of our simulations, so this is the force field we are going to describe in detail here. CHARMM is a non-polarizable force field, which assigns for each atom a Lennard-Jones (LJ) potential and a charge. All particles are considered spherically symmetrical. The partial charge reflects the shifting of the electronic cloud in covalent bonds towards the more electronegative atoms, such as the oxygen atom in the water molecule, and it can have negative or positive values depending on the molecule properties. In the case of ions, the charge assigned is the electric charge of the cation or anion of interest. The electrostatic energy between two molecules is then calculated using the usual Coulomb interaction:

$$U_{\mathrm{Coul}}(r_{12}) = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{12}}, \tag{2.1}$$

where q_1 and q_2 are the charges or partial charges of atoms 1 and 2, and r_{12} is the distance between the point charges located at their centers of mass. This potential has an infinite range, which can be computationally very costly if we use periodic boundary conditions. To circumvent this problem we use the Ewald summation, which we explain in the next sections. For each atom we also assign a size, that is, two parameters that are used to calculate a Lennard Jones (LJ) potential between a pair of atoms[58]. This is also called a 12-6 potential, and it combines an attraction term for larger distances and a hard-core repulsion term for distances close to the atomic scale:

$$U_{\mathrm{LJ}}(r_{12}) = \varepsilon \left[\left(\frac{\sigma}{r_{12}} \right)^{12} - 2 \left(\frac{\sigma}{r_{12}} \right)^{6} \right], \tag{2.2}$$

where ε is the value of the potential between atoms 1 and 2 at its minimum, and σ the associated distance between their centers of mass. This equation is valid for atoms with the same ε and σ values. When the two atoms are different, the new values are calculated according to:

$$\varepsilon_{12} = \sqrt{\varepsilon_1 \varepsilon_2}, \quad \sigma_{12} = \frac{1}{2}(\sigma_1 + \sigma_2). \tag{2.3}$$

The r^{-6} term in equation 2.2 describes the attraction between atoms at larger distances, known as the Van der Waals interaction. The r^{-12} term represents the Pauli repulsion at short ranges due to overlapping electron orbitals. The LJ interactions have a very small value for distances much larger than the σ value, therefore this potential is truncated during the molecular dynamics simulations, as we are going to explain in section 2.1.3.

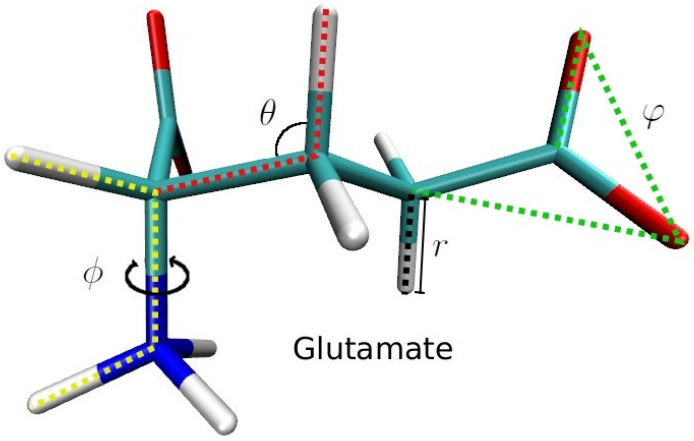

Figure 2.1: Examples of the four bonded parameters used in the CHARMM force field, shown in a glutamate molecule. The bond distance (r), angle (θ), dihedral (ϕ) and improper dihedral (φ) are shown in black, red, yellow and green, respectively.

Classical force fields do not allow for chemical reactions, and therefore all covalent bonds have to be assigned prior to the simulation. Most molecules are not rigid, being able to sample multiple conformations when in solution. An accurate description of the conformational transitions of macromolecules is essential for the understanding of how biological processes take place. This can be achieved by, in addition to non-bonded atomic interactions, also creating a set of parameters for bonded interactions in a given molecule (Figure 2.1). In CHARMM a covalent bond is created using an harmonic approximation, by assigning a minimum value and a spring constant for the distance between two atoms. In the same fashion, a given angle between three atoms in a molecule is restrained by an harmonic oscillator. Impropers also have harmonic potentials, involving four atoms at a time, and are used to maintain the planarity of molecular groups. Dihedrals, on the other hand, have a different expression that includes a dihedral force constant, the multiplicity of the function and a phase shift. This is necessary because torsion angles often have more than one minimum, as the case of a hydroxyl group attached to a aromatic ring. Putting all the bonded and non-bonded terms together, we have the complete

set of interactions in the CHARMM force-field:

$$U = \begin{cases} \sum_{\text{bonds}} \frac{1}{2} K_r (r - r_0)^2 \\ + \sum_{\text{angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \mid \\ + \sum_{\text{impropers}} \frac{1}{2} K_\varphi (\varphi - \varphi_0)^2 \\ + \sum_{\text{dihedrals}} \frac{1}{2} V_\phi (1 + cos(n\phi - \gamma) \end{cases} \quad \begin{cases} + \sum_{\text{pairs}} (U_{\text{Coul}} + U_{\text{LJ}}). \end{cases} \tag{2.4}$$

Here r, θ, φ and ϕ represent a given bond distance, angle, improper and dihedral, respectively. r_0, θ_0, and φ_0 are the positions where the potential energy is zero, with K_r, K_θ and K_φ as the associated spring constants. In the case of dihedrals, V_ϕ is the force constant, n is the multiplicity and γ is the phase shift. With all the terms involved in the calculations of the potential energies, we can now discuss the thermodynamic ensembles and the time evolution of our simulations.

2.1.2 Ensembles

An ensemble is a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state. There exist different ensembles with different characteristics. The microcanonical ensemble (NVE) is characterized by a fixed number of atoms, N, a fixed volume, V, and a fixed energy, E. This corresponds to an isolated system. The canonical ensemble (NVT) is a collection of all systems whose thermodynamic state is characterized by a fixed number of atoms, a fixed volume, and a fixed temperature, T. The isothermal-isobaric ensemble (NPT) has a fixed number of atoms, a fixed temperature, and a fixed pressure, P. The grand canonical ensemble (μVT) is defined by a fixed volume, a fixed temperature, and a fixed chemical potential, μ. For any ensemble, the average of any dynamic property can be obtained from:

$$\langle A \rangle = \sum_i A_i p_i \tag{2.5}$$

where the angled brackets denote an ensemble average, A_i is the value of A in the state i, and p_i represents the probability of observing the i-th state. The idea behind molecular dynamics simulations is that if one allows the system to evolve in time indefinitely, that system will eventually pass through all possible states. Even though we can use this to calculate time averages, experimental observables are assumed to be ensemble averages. This leads us to the ergodic hypothesis, which states that, if we sample a system for long enough, the time average equals the ensemble average. Therefore, in a molecular dynamics simulation, one should

generate enough representative conformations such that the following equality is satisfied:

$$\langle A \rangle_{time} = \langle A \rangle_{ensemble} \tag{2.6}$$

If this is the case, experimentally relevant information concerning structural, dynamic and thermodynamic properties may then be calculated using a feasible amount of computer resources. Because the simulations are of fixed duration, one must be certain to sample a sufficient amount of phase space.

In biological systems the temperature and pressure are usually kept constant, so our simulations are always carried out in the NPT ensemble. The partition function for the NPT ensemble is:

$$\Delta(N, P, T) = \sum_{V} e^{-\beta PV} \sum_{i} e^{-\beta E_i(N,V)}, \tag{2.7}$$

where $\beta = 1/k_B T$ and i are the possible states of the system, with an energy given by E_i. The isothermal-isobaric partition function is directly related to the Gibbs free energy by the expression:

$$G(N, P, T) = -k_B T \ln \Delta(N, P, T). \tag{2.8}$$

The Gibbs free energy is the amount of useful work that can be extracted from a given system at constant temperature and pressure. It plays a major role in chemistry because it describes the spontaneity of a process in these conditions, by combining enthalpic and entropic contributions. It is also related to the equilibrium constant between two states by the expression:

$$\Delta G = -k_B T \ln K_{eq} \tag{2.9}$$

where ΔG is the free energy difference between the two states and K_{eq} is the equilibrium constant. In order to accurately calculate the free energies of binding and conformational changes in proteins using molecular dynamics, we always use the Gibbs free energy.

2.1.3 Numerical Integration

We use the software NAMD[59] for all our simulations, so here we describe how this program solves the equation of motion at each time step. We also account for other features of NAMD that are used to keep the temperature and pressure constant, as well as calculate long-range electrostatic interactions.

Biological systems are chaotic, which means that they will exhibit a completely different outcome even if we make the slightest changes in the initial conditions. Therefore, rather than the actual time evolution of the model, MD simulations are useful to obtain a good ensemble of states and from these states extract relevant properties of the system. The chain of states created in a molecular dynamics

simulation, each of them defined by values for the position and momentum of all atoms, follows Newton's equation of Motion:

$$m_i \frac{d^2 r_i(t)}{dt^2} = -\nabla_i U(\{\mathbf{r}_i\}) \tag{2.10}$$

where $-\nabla_i U(\{\mathbf{r}_i\})$ is the force acting on the atom i as a function of all atomic coordinates and the force field parameters. $\mathbf{r}_i$ and m_i are the coordinates and mass of the atom i, respectively. The acceleration $d^2 r_i(t)/dt^2$ leads to a change in the atom's velocity and position within a discrete time step Δt. This time has to be sufficiently short to capture the fastest motions of the system, which usually come from the vibrations of chemical bonds and angles. The fastest vibrational modes come from bonds that involve hydrogens, due to the small mass of this atom, and they would require a time step of no more than 1 femtosecond. We always use a time step of 2 femtoseconds, so we have to use the SHAKE[60] algorithm in order to keep these bonds rigid during the simulation. To solve equation 2.10 in NVE ensemble, NAMD uses the Verlet method[61]. The velocity-Verlet method obtains the position and velocity for each atom at the next time step (r_{n+1}, v_{n+1}) from the current one (r_n, v_n), assuming the force $F_n = F(\{\mathbf{r}_n\})$ is already computed, in the following way:

$$v_{n+1/2} = v_n + \frac{F_n}{m}\frac{\Delta t}{2}, \tag{2.11}$$

$$r_{n+1} = r_n + v_{n+1/2}\,\Delta t, \tag{2.12}$$

$$F_{n+1} = F(\{\mathbf{r}_{n+1}\}), \tag{2.13}$$

$$v_{n+1} = v_{n+1/2} + \frac{F_{n+1}}{m}\frac{\Delta t}{2}, \tag{2.14}$$

where m is the mass of the particle. We can see that the forces and positions are calculated once and the velocities are calculated twice at each step. The NVE ensemble does not have a constant pressure or temperature, and therefore is not interesting for the systems we want to simulate.

In order to obtain an ensemble that has a constant temperature, one has to modify Eq. 2.10 in a way that does not affect the short-time trajectories. NAMD does that by coupling the system to a temperature reservoir, done in a stochastic way by using the Langevin equation[62]:

$$m\frac{dv}{dt} = F(r) - \gamma v + \sqrt{\frac{2\gamma k_B T}{m}} R(t) \tag{2.15}$$

where m is the mass, v is the velocity, F is the force, r is the position, γ is the friction coefficient, T is the temperature and $R(t)$ is a univariate Gaussian random process. Coupling to the reservoir is modeled by adding the fluctuating (the last term) and dissipative ($-\gamma v$ term) forces to the Newtonian equations of motion (Eq. 2.10). To integrate the Langevin equation, NAMD uses the Brünger–Brooks–Karplus (BBK)

method[63], a natural extension of the Verlet method for the Langevin equation. This adds a stochastic element to the simulations, which means that molecular dynamics simulations that use this method are no longer deterministic.

In order to work in the NPT ensemble, we also need to use a method to keep the pressure constant while varying the volume of the simulation cell. NAMD provides constant pressure simulation using a modified Nosé-Hoover method[65] in which Langevin dynamics is used to control fluctuations in the barostat. An additional variable, e, is used to represent a piston that can compress or expand the system volume. Langevin dynamics is then applied to this piston, giving the following equations:

$$\frac{de}{dt} = \frac{1}{3V}\frac{dV}{dt} \tag{2.16}$$

$$\frac{d^2e}{dt^2} = \frac{3V}{W}(p - p_0) - \gamma_e\frac{de}{dt} + R_e, \tag{2.17}$$

$$\langle R_e^2 \rangle = 2\gamma_e kT/h, \tag{2.18}$$

$$W = 3N\tau^2 kT, \tag{2.19}$$

where analogous values of the Langevin terms are denoted by the subscript e, and the weight factor W is the effective mass of the piston defined through the oscillation period, τ. The work done by this piston on the system volume affects the atomic velocities and accelerations as (de/dt)r and $-(de/dt)v$. In practice, the friction coefficients γ and γ_e determine the degree to which simulations are coupled to the thermostat and barostat. The decay constant τ is always scaled with $1/\gamma_e$ such that piston behaviour is close to critical damping. Larger decay values tend to prioritize equilibration, but affect system evolution and sampling rates. Thus, properties such as the self-diffusion rate of water are amongst those affected by these coupling parameters.

All simulations performed make use of periodic boundary conditions, which means that the simulation box is replicated infinitely on the three coordinate axes. This is necessary to avoid edge effects on the small systems we simulate, but it also brings problems to the calculation of the non-bonded interactions. Both the electrostatic and the LJ potentials have an infinite range, which becomes computationally costly since the number of interactions grows with the square of the number of atoms. If we include the interactions with all the periodic cells, such calculation becomes prohibitive. In the case of the Lennard-Jones potential (Eq. 2.2), the interaction energy becomes very small for distance values much larger than σ, so we can safely use a cut off distance. The σ_{12} values for pairs of atoms are typically smaller then 5 Å, so for distances greater then 12 Å the LJ interaction is set to zero. To avoid a discontinuity in the LJ potential, this potential is gradually reduced starting from 10 Å until 12 Å. The same approach cannot be used for the electrostatic interactions, since in this case the magnitude decreases with $1/r$. The solution in this case is to use the Ewald summation[66], which separates the electrostatic interactions in two parts. The first is a short-range potential, calculated using

the equation for the Coulomb interaction (Eq. 2.1), which converges quickly in the three spatial coordinates. The second is a long range potential, that is calculated by taking advantage of the periodicity of the system, and converges quickly in the Fourier space. By dividing the system into a grid, these calculations can be done even faster, and this method is known as the Particle Mesh Ewald (PME)[67]. We use PME for all of our systems, and to avoid artifacts from the electrostatic calculations the total net charge of the simulation cell must always be zero. This is done by adding or removing ions from the system, which is usually kept at a physiological concentration of NaCl.

Each model we investigate has different features, like being in solution or immersed in a lipid membrane. They also vary concerning the size of the simulation cell, the number of atoms and the simulation parameters. The way we build our systems, as well as the minimization and equilibration procedures, are also important. We discuss these and further details of our simulations in the respective chapters for each of the proteins we study, since they can be different depending on the model.

2.2 Docking of ligands to proteins

Although the binding free energy of ligands to proteins can be measured using a variety of experiments, often we don't have the crystal structure of the complex. This severely limits our molecular understanding of the binding mechanism, and how a particular ligand could be modified to increase or decrease the affinity for a particular channel or receptor. Docking programs like ZDock[68] and HADDOCK[69] aim to fill this gap, by using several steps to predict the binding modes using a scoring function based on the protein and ligand parameters. They can be applied to small ligands, like the ones we bind to the glutamate receptor, as well as large toxin peptides, such as ShK binding to Kv1.1 and Kv1.3. We use HADDOCK followed by MD refinement for all our docking purposes, so in this section we discuss the main aspects of this method.

HADDOCK has first been developed for protein-protein docking, but has since been applied to a variety of protein-ligands, protein-peptides and protein-DNA/RNA complexes. It is based on Ambiguous Interaction Restraints, in which a few residues that are experimentally shown to be involved in binding can be chosen to narrow the search for the possible binding poses. It combines three steps: 1) Randomization and rigid body energy minimization; 2) semi-flexible simulated annealing; 3) flexible explicit solvent refinement. In the first step the two molecules are separated in space and considered rigid, with their rotation and translation randomized and then minimized in turns. This step can be performed in vacuum or in a solvated environment, and generally 1000 or 2000 different complexes are obtained from this procedure. The best structures from step one, typically a few hundred, will be subjected to step two, a semi-flexible simulated annealing (SA) in torsion angle space. Here HADDOCK performs high temperature body searches followed by simulated

annealing, adding flexibility to the side-chains and backbones of chosen residues in the final stages. Step three is a further refinement of the complexes obtained from step 2, keeping the residues flexible and performed with a layer of 8 Å of explicit solvent around the complex.

The final structures are clustered using the pairwise backbone RMSD at the interface, then analyzed and ranked according to their average intermolecular energies calculated using the OPLS non-bonded parameters. The binding mode that is considered the most likely is based on the energy score and the formation of clusters. A large ensemble of low energy structures that have similar binding modes is a good indication of the correct complex formation, even if they don't provide the absolute lower energies. The complex chosen is further refined using molecular dynamics simulations, by slowly releasing restraints applied on both molecules and performing equilibrium simulations of several nanoseconds. During our simulations, this step was found to be especially important for the side chain interactions involving hydrophobic residues.

2.3 Parametrization of molecules

The CHARMM force field provides simulation parameters for proteins, nucleic acids, lipids and carbohydrates. This limits the number of structures that we can treat with MD, since often it is necessary to include an organic molecule in our simulations that is not fully parameterized in CHARMM. This could be the case of a possible drug lead or a non-natural amino acid. The CHARMM General Force Field (CGenFF)[70] allows one to obtain missing parameters for a given molecule that can be readily used in any NAMD simulation, assigned by analogy to existing molecules on the CGenFF database. Unfortunately, this analogy is not always satisfactory, and the parameters have to be validated by combining quantum mechanics (QM) and molecular mechanics (MM) simulations. This is the case of four molecules for which we measure the affinity to the glutamate receptor: AMPA, ACPA, DNQX and CNQX. In this section we describe how to include a new molecule in a MD simulation, starting with the download of the 3-dimensional structure from a chemical database, followed by determination of the CGenFF parameters and their validation using QM simulations. We show as examples some of the molecules listed above, for which we have successfully obtained CHARMM parameters to use in our simulations.

2.3.1 Obtaining the initial parameters

We first download the three dimensional structure of the molecule of interest, which can be done in a chemical database such as Chemspider (http://www.chemspider.com/) or PubChem (https://pubchem.ncbi.nlm.nih.gov/). There is a number of possible formats for this structure, like .mol, .mol2 or .sdf. The CGenFF online tool only

accepts molecules in .mol2 format, so we use the program OpenBabel[71] to convert files from any other format to .mol2. OpenBabel also provides additional tools, like defining the protonation states of titratable groups for a specified pH. Once we have the file in the desired format, we use the online tool from ParamChem (https://www.paramchem.org/) to obtain the initial parameters based on the CGenFF force field, which are assigned by analogy to molecules that have similar groups as the one we wish to parameterize. The output file is called a stream file, and it contains all information needed to include our molecule in a MD simulation. Each atom is assigned an atom type, which has associated LJ parameters to account for its size. These atom types are going to be used to decide which parameters are missing in CGenFF, and also to assign new ones by similarity to parameterized molecules. They involve the partial charges of each atom and the missing bond, angle, dihedral and improper parameters (Fig. 2.2). They also come with a penalty, an indication of how good the analogy is and therefore how reliable are the assigned parameters. Penalties larger than 10 indicate that some basic validation is recommended, and larger than 50 indicate that extensive validation is needed. We can see in the example that a few parameters have to be validated due the large penalties, so we have performed QM simulations using Gaussian09[72] to optimize them, which we describe next.

2.3.2 Optimizing the molecule geometry and calculating the partial charges

The first step in using QM simulations for parameter validation is to perform a geometry minimization. This is done to obtain the optimal structure of our target molecule, which is going to be used in the subsequent steps. For that purpose we use the program Gaussian09, using the 6-31G basis set including single first polarization functions. The geometry optimization combines an initial Hartree-Fock (HF) calculation followed by a Møller-Plesset correlation energy correction, truncated at second-order (MP2). We also obtain the MP2 Merz-Kollman charges for each atom, in order to compare them to the values assigned by ParamChem. There is usually good agreement between the two. In the case of a big discrepancy, it is important to remember that the CHARMM force field is empirical and it is fitted to the solvation energies of different molecules. That means that directly using the charges assigned by Gaussian might not always be the best alternative, due to effects like polarization, which are not taken into account in a classical MD simulation.

The next step is a further optimization of the partial charges with water molecules, using the minimized geometry from the previous step. This is done by performing QM simulations to compute the scaled HF minimum interaction energies and distances between the model compound and individual water molecules. In Fig. 2.3 we show an example of all the water molecules used in the case of the ligand DNQX, with each QM simulation containing only one water molecule at a time.

```
RESI ACPA      -1.000 ! param penalty= 130.000 ; charge penalty= 108.641
GROUP            ! CHARGE   CH_PENALTY
ATOM O1     OG2R50 -0.215 !   40.946
ATOM N2     NG2R50 -0.432 !   48.903
ATOM C3     CG2R52  0.285 !   83.968
ATOM C4     CG2R51 -0.138 !   61.497
ATOM C5     CG2R51  0.307 !   35.592
ATOM C31    CG2O3   0.587 !   85.040
(...)
BONDS
CG2O3  CG2R52  200.00     1.5000 ! penalty= 45
ANGLES
CG2R52 CG2O3  OG2D2    40.00    116.00   50.00   2.35300 ! penalty= 8.5
CG2R52 CG2R51 CG321    45.80    130.00 ! penalty= 2
CG331  CG2R51 OG2R50   45.80    124.00 ! penalty= 45
CG2O3  CG2R52 CG2R51   45.80    130.00 ! penalty= 80
CG2O3  CG2R52 NG2R50   45.80    120.00 ! penalty= 80
DIHEDRALS
OG2D2  CG2O3  CG2R52 CG2R51     3.1000  2   180.00 ! penalty= 93.5
OG2D2  CG2O3  CG2R52 NG2R50     3.1000  2   180.00 ! penalty= 130
CG2R52 CG2R51 CG2R51 CG331      6.6000  2   180.00 ! penalty= 31
CG321  CG2R51 CG2R51 CG331     12.0000  2   180.00 ! penalty= 62
(...)
IMPROPERS
CG2O3  OG2D2  OG2D2  CG2R52    96.0000  0     0.00 ! penalty= 4.5
```

Figure 2.2: Part of the output file from ParamChem for the ACPA molecule. This file contains the missing parameters, a value for each of them assigned by analogy, as well as the associated penalties. The partial charges section is in blue, and the others follow the coloring scheme from Fig. 2.1. We have several penalty values over 50, so this molecule requires further optimization using QM simulations.

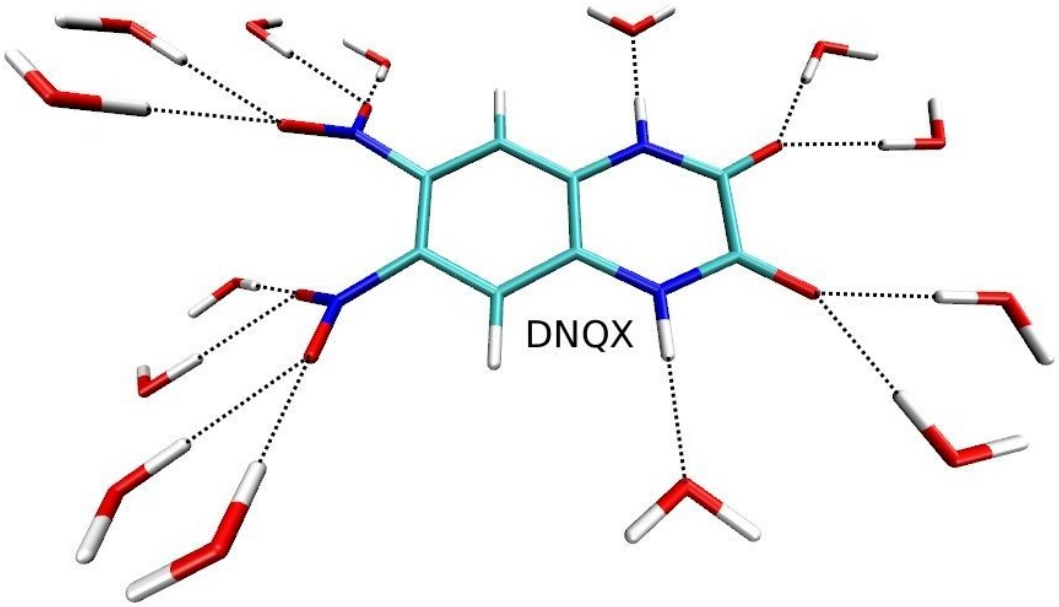

Figure 2.3: The minimized DNQX molecule with the interacting water molecules that are used in the QM and the MM simulations. These simulations are performed with one water molecule at a time, and the partial charges are adjusted to obtain good agreement between QM and MM.

We perform the same procedure in MM simulations by using the optimized DNQX geometry, the molecule parameters we have so far, and the TIP3 model for water. The goal is to obtain good agreement between the interaction energies from the QM and the MM simulations, by performing multiple MM simulations and adjusting the partial charges of each atom until satisfactory agreement between the two methods is obtained. When this objective is met, we turn to the bond and angle equilibrium values and force constants.

2.3.3 Bond and angle equilibrium values and force constants

The missing parameters may also include bonds and angles for which there are no values available in the CGenFF. If the penalties are high enough, we should also validate these parameters of our model compound. This is the case for the isoxazole ring of AMPA, with an hydroxyl group and a methyl group attached to it, which has bond and angle penalties with values over 40. We have parametrized this group separately, before attaching it to the rest of the molecule and calculating the additional parameters. To obtain the bond and angle equilibrium values we perform a geometry minimization in CHARMM, and compare it to the QM minimized structure. We adjust the bond and angle values in the stream file until there is good agreement between the bond distances and angle values obtained by QM and MM geometry minimization.

Symbolic PED matrix [%] (sorted)

1	78.6	dih-C5-C	96.					
2	178.4	rC5Hwag	46.	rngTor2	32.	rC3Owag	16.	
3	267.5	rC5Hrck	63.	rC3Orck	22.			
4	338.3	rC3Owag	31.	rC5Hwag	27.	rngTor1	22.	
5	402.9	dih-OH	87.					
6	413.5	rC3Orck	63.	rC5Hrck	24.			
7	562.2	rngDef1	45.	bC5-C51	25.	bC4-C5	18.	
8	567.5	rngTor1	71.	rC5Hwag	13.	rngTor2	11.	
9	619.7	rngTor2	54.	rC3Owag	45.			
10	708.3	rngDef2	49.	bC3-O31	18.	bC5-C51	15.	
11	708.5	rC4Hwag	106.					
12	895.5	bO1-N2	42.	C51rock	22.	bC5-O1	15.	
13	944.8	bC3-C4	37.	rngDef2	31.	bC5-O1	13.	
14	974.9	C51rock	40.	bC4-C5	26.	rngDef1	24.	
15	1000.5	bO1-N2	43.	bC5-O1	12.	bC3-C4	10.	
16	1016.2	C51rock2	87.					
17	1099.3	rC4Hrck	47.	bC3-O31	19.	bC5-C51	16.	
18	1168.0	aO31-H31	61.	bN2-C3	17.			
19	1249.7	bC5-O1	40.	rC4Hrck	11.	C51rock	10.	
20	1363.9	C51sym	95.					
21	1426.6	C51as2	93.					
22	1439.0	C51as1	82.					
23	1457.6	bN2-C3	51.	bC3-C4	15.	aO31-H31	13.	
24	1474.0	bC3-O31	40.	bC3-C4	24.			
25	1578.1	bC4-C5	65.	bC5-C51	13.	bC3-O31	-11.	
26	2895.8	bC51-H53	37.	bC51-H52	37.	bC51-H51	25.	
27	2972.8	bC51-H53	50.	bC51-H52	50.			
28	2992.4	bC51-H51	75.	bC51-H52	13.	bC51-H53	13.	
29	3101.4	bC4-H4	99.					
30	3455.0	bO31-H31	100.					

Symbolic PED matrix [%] (sorted)

1	94.6	dih-C5-C	99.					
2	171.6	rC5Hwag	43.	rC3Owag	38.	dih-OH	10.	
3	267.0	rC5Hrck	72.	rC3Orck	16.			
4	285.5	rC5Hwag	42.	dih-OH	36.	rC3Owag	21.	
5	417.7	dih-OH	44.	rngTor2	20.	rngTor1	20.	
6	449.6	rC3Orck	60.	rC5Hrck	19.			
7	489.0	bC5-C51	32.	rngDef1	28.	bC4-C5	26.	
8	600.5	rngTor2	56.	rngTor1	44.			
9	677.9	rngDef2	29.	bC5-C51	25.	bC3-O31	22.	
10	731.8	rngTor1	39.	rC3Owag	31.	rngTor2	22.	
11	850.3	rC4Hwag	107.					
12	885.6	bC3-C4	28.	rngDef1	19.	bN2-C3	16.	
13	913.1	rngDef2	26.	bC3-C4	21.	rngDef1	14.	
14	990.3	bO1-N2	62.					
15	1063.1	C51rock	64.					
16	1074.8	C51rock2	83.	C51as2	14.			
17	1099.1	rC4Hrck	55.	aO31-H31	12.			
18	1151.3	aO31-H31	57.	bC3-O31	11.	bN2-C3	10.	
19	1401.6	bC4-C5	28.	C51sym	22.	bC3-O31	18.	
20	1422.1	C51as1	83.					
21	1432.5	C51as2	86.	C51rock2	14.			
22	1508.9	bC5-O1	28.	bC3-O31	27.	bN2-C3	10.	
23	1514.7	C51sym	56.	bN2-C3	16.			
24	1591.8	rngDef1	21.	bN2-C3	20.	bC5-O1	12.	
25	1652.6	bC3-C4	40.	bC4-C5	27.	rngDef2	20.	
26	2848.0	bC51-H51	34.	bC51-H53	33.	bC51-H52	33.	
27	2913.8	bC51-H52	50.	bC51-H53	50.			
28	2914.6	bC51-H51	65.	bC51-H53	17.	bC51-H52	17.	
29	3062.8	bC4-H4	99.					
30	3691.5	bO31-H31	100.					

Figure 2.4: The vibrational modes of the isoxazole ring attached to a hydroxyl and a methyl group, which is part of the parametrization for the AMPA molecule. The QM result (left) and the MM result (right) show good agreement, since they reproduce similar vibrational modes. The different components that contribute to each mode are a combination of bonded interactions, which are assigned according to Pulay *et al.*[73].

The calculation for the force constants is based on the reproduction of the scaled MP2 vibrational spectrum. This is done by computing force constants, the resulting vibrational frequencies and the intensities in an MP2/6-31G* Gaussian simulation, using as an input the optimized geometry obtained previously. The same vibrational modes can be calculated in CHARMM, using the VIBRAN command and the MM minimized structure. This step also involves the creation of a U MATRIX (UMAT), that transforms the molecule's internal coordinates to the local internal valence coordinates as outlined in Pulay et al.[73]. This procedure gives as an output two vibrational spectra, one from the QM and one from the MM simulations. Here we change the bond and angle force constants and perform MM simulations until there is good agreement between the vibrational modes obtained by Gaussian and CHARMM. We give as an example in Fig. 2.4 the two spectra obtained for the AMPA isoxazole ring using the two methods, after we adjust the force constants in the stream file. We can see that the main vibrational modes, as well as the contribution from the different components, show good agreement. Therefore we can move to the next step in the parameterization, which is the calculation of the dihedral force constants, multiplicities and phase shifts.

2.3.4 Calculation of the dihedral parameters

To calculate the dihedral parameters we perform potential energy scans using QM simulations, by rotating the torsional angle of interest and minimizing the molecule conformation at each dihedral value. The initial structure is the optimized geometry of the target molecule, and here also the MP2 optimization, as well as the 6-31G basis set, is used. The potential energy scan is done in discrete steps, by choosing a range and an increment for the dihedral of interest. The range does not necessarily has to be 360 degrees, as it can be smaller in the case of rigid molecules, and the increment is usually around two degrees. At each dihedral value the minimum of the MP2 potential energy is calculated, together with the atomic coordinates for this minimum. These coordinates are going to be used as inputs for the MM calculations, where CHARMM calculates the same potential energies at each dihedral value, based on the parameters we have defined so far in our stream file. With the potential energies of both QM and MM calculations, we can now plot them as a function as the dihedral value, and find the optimal dihedral parameters that show the best agreement between the two methods. We show in Fig. 2.5 the potential energy scans for the missing dihedrals of the CNQX compound, after we adjust the dihedral parameters. In the graph we have both the QM and the MM results, and we can see the good agreement between them.

Correctly describing the torsional angles is very important for MD simulations, since these are the bonded parameters that have the largest variation for molecules in solution. This is because they often have more than one minima, and the energy barriers between these minima can be overcome in the time scales of our simulations. They play a great role in the conformational changes of proteins, so they have to be properly adjusted to correctly describe the time evolution of our system. We have not mentioned the impropers in our parameterization validation, since in all molecules we have parameterized they have large force constants around a single minima, as well as a penalty below 5. Therefore, they did not call for refined parameterization procedure as is the case for the bonds, angles and dihedrals.

2.3.5 Re-optimization of the other parameters

The partial charges were the first parameters we calculated, and after that we have made important structural changes in our target molecule. These changes can affect the interaction energy of water molecules, so we have to check if any further charge optimizations are needed. This is done by repeating the procedure on the second paragraph of section 2.3.2 with the new values for the bonded parameters. The other steps also might need further optimization if the charges are reassigned, so the parameterization process is iterative. Once the stream file provides satisfactory agreement between the QM and the MM interactions in all steps, it is ready to be used in a simulation. Often there is a trade-off between parameters, which imposes a limit to the accuracy on the description of a QM system using CHARMM. Never-

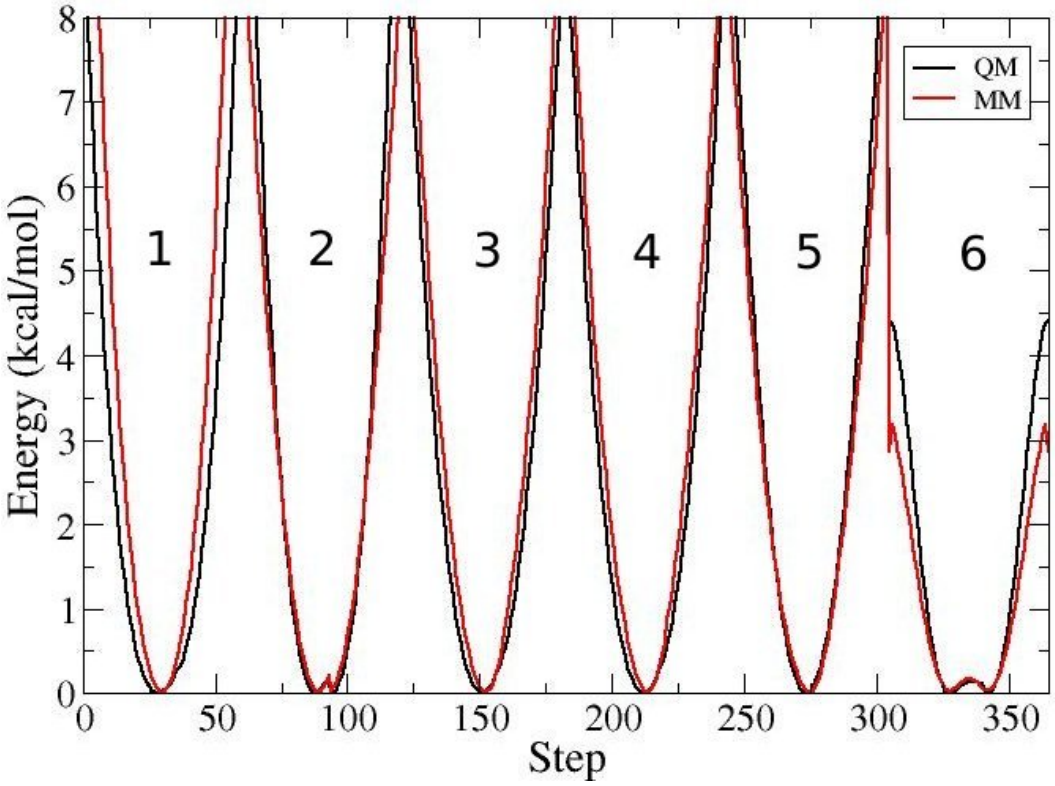

Figure 2.5: Potential energy scans for six different dihedrals (numbered one to six) belonging to the CNQX molecule, which required QM validation. The first five dihedral scans used 60 steps with 2 degrees increments, with a total of 120 degrees each, and the last scan used 60 steps with 3 degree increments, totaling 180 degrees. Each potential energy scan is calculated in a separate simulation. The agreement between the potential energy scan using QM (black) and MM (red) is very good, which means that the parameters have been successfully optimized.

theless, the procedure described here is very effective, as we show in the calculations of the binding free energies of AMPA, ACPA, CNQX and DNQX to the glutamate receptor in Chapter 8.

2.4 Homology Modeling

Membrane proteins may be very difficult to crystallize, limiting the number of systems we can build in order to perform MD simulations. To overcome this limitation, one can construct models based on the known structure of homologous proteins. This is the case for the human glutamate transporters (EAATs), for which no crystal structure has been resolved yet, even though the main elements of the tertiary structure were determined using other methods. The homology between two proteins is obtained by aligning their amino acid sequence, and checking for residue similarity in equivalent positions. The closest homolog of the EAATs whose crystal structure has been resolved is the archaeal transporter Glt_{Ph}, which shares an overall 36% sequence identity with the human transporters. It is also a homolog of the neutral amino acid transporters ASCT1 and ASCT2, with a 23% homology. As explained in the introduction, Glt_{Ph} also conserves the main structural elements present in the EAATs, as well as the majority of residues involved in ion and substrate binding. Therefore, we use Glt_{Ph} as a template for building EAAT and ASCT models using sequence alignment and the program MODELLER[74]. We also check for the quality of the structures we obtain using a variety of scoring methods. The alignment, model building and model evaluation are outlined in the next sections.

2.4.1 Sequence alignment

Obtaining the correct sequence alignment between the template and the target protein constitutes one of the most important steps in homology modeling. This alignment is done by comparing the amino acid sequence of two proteins, and aligning equivalent regions based on their sequence similarity. It can include gaps and insertions, which are caused by creation and deletion of residues during evolution. In the case of homology modelling, the alignment includes the sequences of the template and the target proteins. It can be performed using more than two amino acid sequences, as in the case of multiple alignment for a whole family of proteins, or for a target protein that has more than one template crystal structure. MODELLER has a built-in tool to perform sequence alignments, loosely based on the program COMPARER[75], which provides additional options like imposing gap penalties and using structural information from the template sequence. Another alignment tool can be found in the Swiss Model server, by using the program T-COFFEE[76]. Swiss Model also provides a template search using BLAST[77], which is useful if we have a given target sequence and we wish to find the closest homologs that can be used as template structures. We have used these tools to align the sequences of the human

transporters to the Glt_{Ph} transporter, and obtained very similar results to the alignment provided in Ref[14]. We chose to use the alignment provided in this reference to build our models, since it also includes the sequences from other archaeal and human neutral transporters. By converting it to FASTA format, we use it as the starting point for MODELLER to build our EAATs models, which we explain in the next section.

2.4.2 Building the models

MODELLER implements comparative protein structure modeling through satisfaction of spatial restraints[78]. It can perform many additional tasks, including *de novo* modeling of loops in protein structures and optimization of various models of protein structure. It also allows for the inclusion of restraints on distances, angles, dihedrals, as well as the secondary structure of a particular region. MODELLER works by imposing spatial restraints on the protein according to a probability distribution function (pdf) $p(x)$, which is integrated to give:

$$p(x_1 < x < x_2) = \int_{x_1}^{x_2} p(x)dx, \tag{2.20}$$

where the term on the left side is the probability of an event x to be found within the boundaries x_1 and x_2. The pdf is always a function of x and other specified variables:

$$p(x) = p(x, a, b, c...) \tag{2.21}$$

These variables are determined according to stereochemical restraints, the template structure and a database of known protein structures. The stereochemical restraints include the bond distances, angles, dihedrals, Coulomb's law and Van der Waals repulsion between pairs of atoms, as well as specified disulphide bonds. The values for these interactions are based on the CHARMM22 parameter set. The restraints based on the template structure are applied on the C^{α}–C^{α} distances, the main chain N-O distances, the main chain conformation, and the residue side chain conformations.

Additional terms can be added to the pdf above, if we wish to restrain a particular feature of the system such as an angle or a dihedral between atoms. They also allow for the creation of a secondary structure like an α-helix or a β-bridge in a given segment of the target protein, if this sequence has no equivalent on the template. This was the case for a 50 residue loop that is present in the EAATs but not in Glt_{Ph}, that was proposed to be an α-helix by FRET studies, and is discussed further in the results section. MODELLER is also capable of producing homology models that include bound ligands in the alignment, such as ions or amino acid substrates. This is useful if we want to model a given protein that has the same ligands bound as the template structure, since their presence is also taken into account in the calculation of the spatial restraints. That was the case for all of our models, since Glt_{Ph} and the EAATs share the same ligands and their binding sites are believed to be conserved.

We have produced homology models for EAAT1, EAAT2, EAAT3, ASCT1 and ASCT2, all using Glt_{Ph} as a template. In this thesis we focus on the results from the EAAT3 model, since this is the one we have investigated further in Chapters 6 and 7.

2.4.3 Evaluating the models

We have built 20 models for each transporter using the method explained above, and performed different model evaluation procedures to choose the best ones to use in our MD simulations. We have used three separate methods: the QMEAN (Qualitative Model Energy ANalysis), a composite scoring function ranging from 0 to 1, which describes the major geometrical aspects of protein structures[79]; the DFIRE (Distance-scaled, Finite Ideal-gas REference), a potential based on a database of non-homologous proteins[80]; and the DOPE (Discrete Optimized Protein Energy), an atomic distance-dependent statistical potential calculated from a sample of native structures[74]. The first two are available on the Swiss Model Server, and the last one is built into the program MODELLER. The model for each transporter that had the best quality, based on the methods above, was chosen for the MD simulations. The scores obtained for the models, as well as the system creation and equilibration procedures, are discussed in further detail in Chapters 6 and 7.

2.5 Free energy methods

Calculating the free energies involved in different biological processes is one of the main goals of MD simulations. If one could perform simulations for a sufficient amount of time, the partition function in Eq. 2.7 could be directly obtained, as well as the probability of the system to be in any given state. Since this is not feasible in all-atom simulations, we can associate the probability ratio of two macroscopic properties of our system to the free energy difference between them, according to the equation:

$$\Delta G_{12} = -k_B T \ln\left(\frac{P_1}{P_2}\right) = -k_B T \ln K_{eq} \tag{2.22}$$

This expression can be calculated using a reasonable amount of computational resources, as long as we find ways to sample the relevant states of the system and their associated free energies. This can be done in many ways, like performing an alchemical transformation or adding a biasing potential along a chosen reaction coordinate. In this section we discuss the free energy methods that we use in this thesis: the Free Energy Perturbation (FEP), Thermodynamic Integration (TI)[52] and the Potential of Mean Force (PMF)[53]. The first two are calculated by mutating or inserting a given molecule in the system in small steps, and the third by introducing biasing potentials along a given reaction coordinate and obtaining the PMF using the Weighted Histogram Analysis Method (WHAM)[81]. In section 2.5.1

we explain the Free Energy Perturbation (FEP) and the Thermodynamic Integration (TI) methods, which have a similar formalism. In section 2.5.2 we discuss the application of umbrella sampling to obtain a given Potential of Mean Force (PMF).

2.5.1 Free Energy Perturbation and Thermodynamic Integration

Molecular dynamics simulations sample a given state of the system with a probability proportional to $e^{\frac{-G}{k_B T}}$, and therefore they seek out preferentially the lowest free energy regions of configurational space. For densities typical of molecular liquids, the high- and low-energy regions of configurational space are sufficiently well separated that a conventional simulation of practical length will never adequately sample the high-energy regions that contribute significantly to the ensemble average expression for free energy. However, in molecular systems the free energy problem can be seen in terms of a free energy difference between two states, ΔG. If we associate the initial and final states to Hamiltonians given by H_0 and H_1, respectively, this free energy difference can be cast in terms of a coupling parameter. Let us assume that the Hamiltonian of the system, H, depends on a continuous parameter λ, such that as λ is varied from 0 to 1, $H(\lambda)$ passes smoothly from H_0 to H_1. In the perturbation method one introduces a hybrid Hamiltonian $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, where H_0 represents the Hamiltonian for the initial state and H_1 for the final state of the system. The interval between H_0 and H_1 can be divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average:

$$\Delta G_i = -k_{\mathrm{B}} T \ln\langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_{\mathrm{B}} T]\rangle_{\lambda_i}. \tag{2.23}$$

The subscript λ_i indicates that, even though we are using two different Hamiltonians, the averaging is over the ensemble of configurations representative of the initial state of the system. This means that there must be significant overlap between the configurational space of equilibrated systems using the $H(\lambda_{i+1})$ and $H(\lambda_i)$ sets of interactions. This can be achieved by choosing appropriate subintervals, or windows, for λ. Successful application of the perturbation method is commonly limited to free energy differences of < 2kT, or ~ 1.5 kcal mol. Consequently, the free energy difference between two adjacent lambda values should be less than $\sim$1.5 kcal/mol. The total free energy difference ΔG is then simply given by the sum:

$$\Delta G = \sum_i \Delta G_i \tag{2.24}$$

This calculation should also be done in the opposite direction, H_1 to H_0, which is commonly called the backward FEP calculation. The discrepancy or hysteresis in these two estimates of ΔG gives an estimate of the statistical uncertainty in the result, and it arises from configurations sampled in $H(\lambda_i)$ and not in $H(\lambda_{i+1})$, and vice versa.

The Thermodynamic Integration method is similar to the FEP method described above, but here we obtain the ensemble average of the derivative, $\frac{\partial H(\lambda)}{\partial \lambda}$, and integrate over λ to obtain the final value of the free energy difference:

$$\Delta G = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda. \tag{2.25}$$

The best way to solve this integral is to obtain $\frac{\partial H(\lambda)}{\partial \lambda}$ for a finite set of λ values, and use them to numerically solve the integral above. We always use the Gaussian quadrature, which is an integration method that uses a weighted sum of function values at specified points within the domain of integration:

$$\int_0^1 f(x) = \sum_{i=1}^{n} \omega_i f(x_i) \tag{2.26}$$

An n-point Gaussian quadrature rule yields an exact result for polynomials of degree $2n-1$ or less, and in our calculations we use a value for n between 7 and 12. As we explain in the next paragraph, most of our alchemical free energy calculations are divided into an electrostatic and a Lennard-Jones (LJ) term for quicker convergence. The calculations related to the electrostatic interactions can always be approximated by a 7-point quadrature, but the LJ interactions calculations often require 9 or 12 points due to the steepness of the ΔG curve close to the end points.

FEP/TI calculations can bring serious convergence problems in the end points when we have the insertion/deletion of particles in a given system. This problem is often related to the high values of the LJ potential when the distance between two atoms, r_{12}, approaches zero. Even if we use very small λ subintervals, there is always a singularity for the LJ potential at $r_{12} = 0$, and that causes instability for values of lambda close to 0 and 1. The solution is to add a term to the expression of the Lennard-Jones potential (Eq. 2.2)[82], where r_{12} is scaled with lambda according to the expression:

$$r_{12} = \sqrt{r_{12}^2 + c(1-\lambda)}, \tag{2.27}$$

where c is the value of the shift coefficient, which is specified in the NAMD input file. Even though this does not affect the initial and final states of the FEP calculations, it modifies the LJ interaction into a soft core potential for $0 < \lambda < 1$, so that particles can be smoothly added to the system. We find that c values between 4.0 and 7.0 are suitable for our calculations, so we always use shift coefficients in this range. The use of a soft core potential, on the other hand, is not compatible with an electrostatic potential that is $-\infty$ for opposite charges when $r_{12} = 0$. Therefore, we separate the Lennard-Jones and the electrostatic contributions during the FEP calculations, by first creating a neutral molecule and then charging it, or vice-versa.

Further details about our FEP/TI calculations, as well as how to use restraints to calculate the entropic contributions of ligand binding free energies, are explained in Chapter 8.

2.5.2 Potential of Mean Force

When a particular transformation can be described in terms of a physical coordinate, like a distance or a dihedral, the calculation of the potential of mean force (PMF) becomes a very useful tool to calculate the free energy difference between two states. There are many methods that can be used to calculate the PMF, like metadynamics[54], adaptative biasing force[83] and umbrella sampling[53]. The main idea behind these methods is to introduce a bias in the system, so that it visits states that are essential for the calculation of the free energy, but cannot be sampled during equilibrium simulations due to their low probability. This is the case for the unbound states of ligands that have a high affinity[84], or intermediate states in protein conformational transitions[41]. In this thesis we use umbrella sampling, which consists of harmonic potentials applied as a function of a reaction coordinate, that restrain the system in a particular region of the phase space. By using the biasing potential as an input, the PMF can be obtained using the Weighted Histogram Analysis Method (WHAM). The PMF is then numerically integrated to obtain the final value of the free energy. In the next paragraphs we discuss each of these steps, as well as how they apply to our calculations.

By choosing an appropriate reaction coordinate, we can apply harmonic potentials at different positions along the latter so that its whole extent is sampled at a relatively low computational cost. This method is called umbrella sampling, and we use it to calculate the binding affinity of different toxins to the Kv1.1 and Kv1.3 potassium channels, as well as to separately calculate the influence of a salt bridge in the binding free energy of AMPA and Glutamate to the GluA2 receptor. Umbrella sampling can be done in multiple dimensions, although they bring a larger computational cost, since the number of simulations varies exponentially with the number of dimensions we choose. We use a one dimensional PMF, which is enough for our purposes, as well as a reaction coordinate defined by a distance between two groups of atoms. To create the states along this distance, also called windows, we perform steered molecular dynamics (SMD) simulations starting from the structure with the lowest free energy. We then use this path to create the umbrella windows, which are simulated until good convergence is obtained. Generally the SMD windows are not yet equilibrated properly, so we have to run a few nanoseconds of equilibration and then production for each window. The distance between each window, as well as the force constants, can have different values depending on the system. The important thing is to have sufficient overlap between adjacent windows[84], so that the whole path is sampled properly. In Figure 2.6A we show the biased histogram as a function of the distance between the α-amino group of glutamate and the GluA2 E705 residue side chain carboxyl group, for each of the windows we simulate. The distance between windows is 0.5 Å, ranging from 2.5 Å to 9.0 Å, with an harmonic biasing potential of 8 (kcal/mol)/Å^2 centered at the distance value for each window. This histogram cannot be directly used to calculate the free energy, due to the strong bias applied on the system, which has to be removed in order to obtain the

unbiased PMF.

To remove the applied bias on the histogram of Fig. 2.6A, and obtain the unbiased PMF of the system, we use the WHAM method[81]. WHAM provides a scheme for obtaining the optimal estimate of the unbiased probability distribution from biased probability distributions in the presence of biasing potentials. For one reaction coordinate ξ, the WHAM equations are given by:

$$\rho(\xi) = \frac{\sum_{i=1}^{N_w} n_i \rho_i^{(b)}(\xi)}{\sum_{j=1}^{N_w} n_j e^{-\beta(V_j(\xi)-F_j)}} \tag{2.28}$$

$$e^{-\beta F_i} = \int e^{-\beta V_i(\xi)} \rho(\xi) d\xi, \tag{2.29}$$

where $\beta = 1/k_B T$, N_w is the number of windows, $V_i(\xi)$ is the biasing potential of the ith window, $\rho_i^{(b)}(\xi)$ is the biased probability of finding the system at ξ at simulation i, n_i is the number of samples used to construct $\rho_i^{(b)}(\xi)$ at simulation i, F_i is the free energy constant of window i, and $\rho(\xi)$ is the unbiased probability distribution. The last two terms are unknown, and WHAM solves these two equations self-consistently to obtain them. The potential of mean force is directly related to $\rho(\xi)$, by the equation:

$$PMF(\xi) = -k_B T ln(\rho(\xi)) + C \tag{2.30}$$

The constant C is obtained by setting a ξ value for which $PMF(\xi) = 0$, usually the value of the PMF minimum. In Figure 2.6B we show the unbiased PMF calculated from the distributions of Fig. 2.6A, after we have applied the WHAM method. Integrating this curve as a function of the chosen distance gives the free energy in this region of the phase space, that can be used to calculate the free energy of ligand binding or local conformational changes. The way we integrate the PMF, and how that applies to the systems in question, is further explained in Chapters 8 and 9.

2.5.3 Convergence

In both FEP and umbrella sampling, it is essential that the simulations are properly converged for the obtention of reliable results. This means that the majority of states involved in a given tranformation or translocation have to be appropriately sampled, which is harder to accomplish when dealing with large systems. In the case of FEP simulations, large conformational changes in the binding protein or the ligand, caused by the binding/unbinding of the substrate, can cause great problems when performing the transformation from $\lambda = 0$ to $\lambda = 1$ (or vice-versa). This issue can be identified by a large histeresis between the forward and backward calculations, or by significant differences between results using different simulation times. The use of TI is helpful in these situations since it allows for greater sampling of

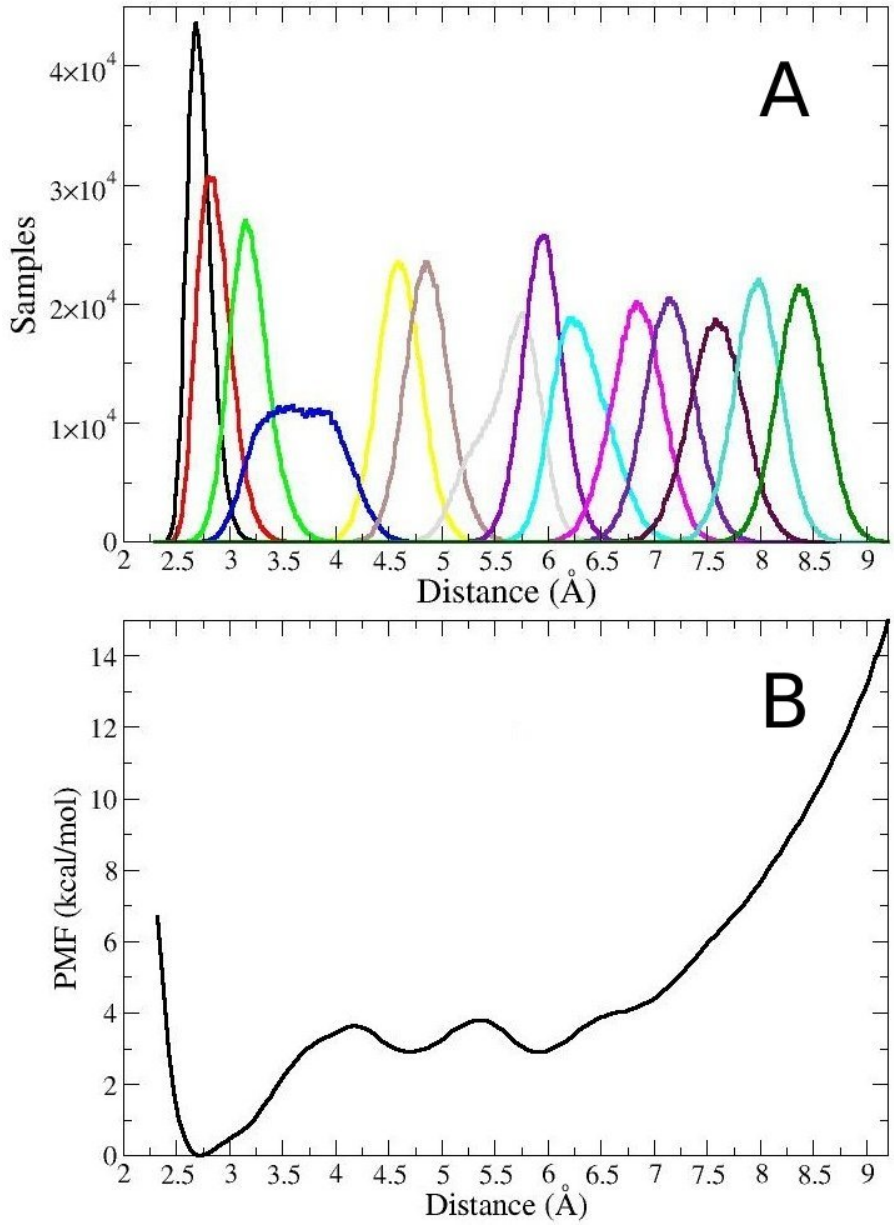

Figure 2.6: (A) Biased probability distribution for each of the 14 windows we simulate, when calculating the PMF of the distance between the α-amino group of glutamate and the E705 side chain carboxyl group of the GluA2 receptor. (B) The final PMF after we use WHAM and the biased probability of the previous item.

a smaller number of windows. Another way to solve convergence problems is to treat a given conformational change on the protein, like a dihedral rotation, separately and perform the FEP calculations in a confined state. Restraining the ligand conformation, translation and rotation during the FEP simulations, and calculating the free energy of applying these restraints separately, also allows for accurate FEP calculations with less computational effort. This is explained further in Chaper 8.

In the case of PMF calculations of toxin binding to ion channels, the toxin might experience intermediate conformational changes when it is pulled from the binding site during the steered molecular dynamics and umbrella sampling simulations. These transient states between the bound and unbound states of the toxin can cause convergence problems and therefore influence the results. Similarly to the case of FEP, we can perform restrained simulations and calculate the effect of these restraints on the toxin separately, in the binding site and in bulk. This method also greatly increases convergence at a relatively low computational cost[84].

Chapter 3

Position of the Third Na^+ Site in the Aspartate Transporter Glt_{Ph} and the Human Glutamate Transporter, EAAT1

ABSTRACT

Glutamate transport via the human excitatory amino acid transporters is coupled to the co-transport of three Na^+ ions, one H^+ and the counter-transport of one K^+ ion. Transport by an archaeal homologue of the human glutamate transporters, Glt_{Ph}, whose three dimensional structure is known, is also coupled to three Na^+ ions but only two Na^+ ion binding sites have been observed in the crystal structure of Glt_{Ph}. In order to fully utilize the Glt_{Ph} structure in functional studies of the human glutamate transporters, it is essential to understand the transport mechanism of Glt_{Ph} and accurately determine the number and location of Na^+ ions coupled to transport. Several sites have been proposed for the binding of a third Na^+ ion from electrostatic calculations and molecular dynamics simulations. In this study, we have performed detailed free energy simulations for Glt_{Ph} and reveal a new site for the third Na^+ ion involving the side chains of Threonine 92, Serine 93, Asparagine 310, Aspartate 312, and the backbone of Tyrosine 89. We have also studied the transport properties of alanine mutants of the coordinating residues Threonine 92 and Serine 93 in Glt_{Ph}, and the corresponding residues in a human glutamate transporter, EAAT1. The mutant transporters have reduced affinity for Na^+ compared to their wild type counterparts. These results confirm that Threonine 92 and Serine 93 are involved in the coordination of the third Na^+ ion in Glt_{Ph}.

3.1 Introduction

Glutamate is the major excitatory neurotransmitter in the mammalian central nervous system. The extracellular concentration of glutamate is predicted to be as low as 25 nM[86] and is maintained by specific transport proteins called excitatory amino acid transporters (EAATs). Excessive extracellular glutamate is toxic for neurons, and therefore its concentration needs to be strictly controlled. Loss of this control due to dysfunction of EAATs has been implicated in several neurological diseases such as Alzheimer's disease, motor neuron disease and amyotrophic lateral sclerosis[12].

Glutamate transport via the EAATs is coupled to the co-transport of three Na^+ ions and one H^+ ion followed by the counter-transport of one K^+ ion[13,87]. Binding of Na^+ and glutamate to the EAATs also activates an uncoupled Cl^- conductance[88–90]. A considerable amount of data has been gathered on the functional properties of the EAATs, but in the absence of any molecular structures it has been difficult to interpret this data and arrive at a molecular-level understanding of the transport mechanism of the EAATs. Determination of the crystal structure of a glutamate transporter homologue from *Pyrococcus horikoshii* (Glt_{Ph})[14] has therefore caused much excitement in the field. Glt_{Ph} is a Na^+-dependent aspartate transporter that also has an uncoupled Cl^- conductance and the binding sites for the substrate aspartate and two Na^+ ions have been identified[15,91] (Figure 3.1A). Whether two or three Na^+ ions are co-transported with each substrate molecule was not clear from the initial functional studies of Glt_{Ph}, where a Hill coefficient >2 was observed[15,35,91]. This issue has recently been resolved in an experiment using ^{22}Na and ^{3}H-aspartate, where it has clearly been shown that three Na^+ ions are co-transported with each substrate as in the EAATs[37]. In this study, we follow the nomenclature already established and call the Na^+ ion binding sites observed in the crystal structure Na1 and Na2[15] and the third Na^+ binding site Na3.

Glt_{Ph} shares about 36% amino acid sequence identity with the EAATs, which increases to above 60% for the residues forming the binding pocket[15]. Furthermore, many of the residues that have been implicated in substrate and ion binding, and also Cl^- permeation, in the EAATs are conserved in Glt_{Ph}[28,32,90,92–94]. Thus, the Glt_{Ph} structure offers a good starting point for constructing homology models of the EAATs which can be used to perform simulations to gain a detailed understanding of the transport mechanism. All three Na^+ ions need to be included in such studies, and this requires an accurate determination of the location of the Na3 site. Even a relatively small inaccuracy in the binding position of the ion could affect its coordination shell and lead to sizable inaccuracies in the calculated binding free energy. Several Na3 sites have been proposed from electrostatic calculations[31,95] and molecular dynamics (MD) simulations[96,97], which are compared with the available experimental data below.

The most important clue for the location of the Na3 site comes from the crystal structure of Glt_{Ph}[15], where the side chains of residues T92 and D312 are not in-

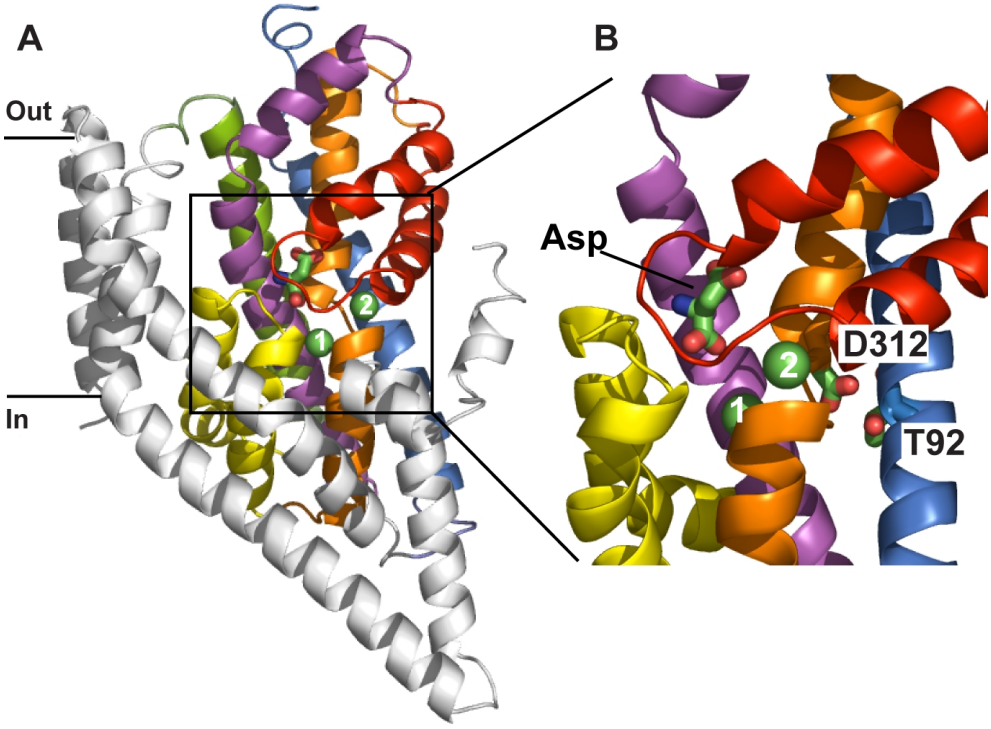

Figure 3.1: Structure of Glt_{Ph} with bound aspartate, Na1 and Na2 (PDB 2NWX). (A) One protomer in cartoon representation; transmembrane domain (TM)1, TM2, TM4 and TM5 (coloured in grey), TM3 (blue), TM6 (green), TM7 (orange), TM8 (magenta), hairpin (HP)1 (yellow) and HP2 (red). Bound aspartate is in stick representation and Na1(1) and Na2(2) are indicated as green spheres. (B) Close-up of the substrate binding site that has been rotated 90° around the vertical axis (coloured as in A). Aspartate 312 (D312) and threonine 92 (T92) are shown in stick representation. Figures were made using Pymol[98].

volved in coordination of Na1 or Na2 (Figure 3.1B). This is in contrast to evidence from mutagenesis experiments on EAAT3 where both of these residues have been proposed to coordinate one of the coupled Na^+ ions during glutamate transport[30,31]. This suggests that the vicinity of the T92 and D312 side chain oxygens is the most likely place for the Na3 site and MD simulations of the Glt_{Ph} structure provide further clues in this regard. In MD simulations lasting 50 ns we have observed that, in the absence of a Na^+ ion nearby, the D312 side chain swings ∼5 Å and coordinates Na1. Similar D312–Na1 coordination has been observed in other MD simulations of Glt_{Ph}[96]. The resulting coordination shell for Na1 is in substantial disagreement with that observed in the crystal structure, indicating that D312 should not be involved in the coordination of Na1. The only way the D312 side chain can be prevented from swinging towards Na1 in MD simulations is by allowing it to coordinate the third Na^+ ion. Such a Na3 binding site was first proposed from electrostatic calculations, where four other oxygens besides the two oxygens from D312 were identified as possible ligands, namely, the side chains of Y88, T92 and N310, and the backbone of G404[31]. This Na3 site was refined in subsequent MD simulations, where it was proposed to be formed by the side chains of T92, N310 and D312, and a water molecule[97].

The aim of this study was to carefully examine the proposed Na3 binding sites and find the best site that is consistent with all of the available data. Because the D312–Na3 coordination is fairly well established from both experiments[30] and simulations[97], we impose this condition which considerably simplifies the search. Our simulation results indicate that there is an alternate Na3 binding site near the one proposed in[97], which involves the side chains of T92, S93, N310, D312, and the backbone of Y89. This site provides a better coordination for the third Na^+ ion and hence a lower binding free energy. We have used site-directed mutagenesis to change the proposed ligand residues T92, S93, N310 and D312 to alanine in Glt_{Ph}. The N310A and D312A mutant transporters are non-functional, but functional analysis of T92A and S93A, and the equivalent residues in EAAT1, confirm their involvement in the coordination of Na3.

3.2 Results

3.2.1 MD simulations of Glt_{Ph} reveal a novel Na3 site

To date, all MD simulations of Glt_{Ph} and predictions for the Na3 site have used the aspartate–bound closed structure. In this study, we employ two simulation systems derived from the aspartate–bound 'closed' (PDB: 2NWX) and the TBOA–bound 'open' (PDB: 2NWW) structures. In the closed structure the helical hairpin 2 (HP2) gate is closed and aspartate and two Na^+ ions (Na1 and Na2) are bound (2NWX). While in the open structure the non-transportable inhibitor TBOA is bound in place of aspartate which prevents closing of the HP2 gate and only one Na^+ ion (Na1) is bound (2NWW). We have created an open structure from 2NWX by removing

the ligands and observing the opening of the HP2 gate in MD simulations[36,99]. A similar apo structure has been created from 2NWW by removing TBOA. To distinguish between these two structures, we will refer to them as "closed" and "open" following their origin, although in both structures the HP2 gate is open.

As T92 and D312 have been proposed to bind a Na^+ ion in EAAT3[30,31] and they are unlikely to interact with Na1 or Na2, we placed a Na^+ ion in the vicinity of the T92 and D312 side chains in the system obtained from the closed structure, and equilibrated this system in MD simulations. The resulting coordination shell is consistent with that found previously[97] and consists of the oxygens from the sidechains of T92, N310 and D312 (two), and a water molecule (this site will be called Na3' henceforth, see Figure 3.2A). We next considered the system obtained from the open structure of Glt_{Ph}. Surprisingly, when this system is equilibrated in MD simulations with a Na^+ ion placed near the T92 and D312 side chains, a rather different coordination shell emerges, consisting of the oxygens from the side chains of T92, S93, N310, D312 (one oxygen), and the backbone of Y89 (Figure 3.2B). In the new site (to be called Na3), water and one of the D312 carboxyl oxygens in the Na3' site are replaced with two new ligands, which leads to a tighter Na^+ coordination shell due to shorter Na–O distances.

An immediate question is why there are two different Na3 sites if the two structures are so similar. The answer becomes obvious when the open and closed crystal structures are superimposed – the O and N atoms in the N310 side chain are interchanged in the two structures (Figure 3.2A,B). This different position of the N310 side chain leads to a difference of about 2 Å between the position of Na3 and Na3'. Na3 is closer to transmembrane domain 3 (TM3), where the extra coordinating ligands reside. The aspartate and TBOA bound crystal structures were solved at ∼3 Å and ∼3.5 Å respectively. At this resolution it is difficult to assign the positions of side chains with accuracy so we have performed several tests to see which of the two positions of the N310 side chain is more realistic. To check the stability of the respective binding sites, we have equilibrated the open and closed structures without Na^+ ions and Asp. In 5 ns MD simulations, the N310 side chain has retained its original configuration in both the open and closed structures, indicating that there is no intrinsic preference for one of the configurations in the 'apo' state. Next, we have equilibrated the two structures in the presence of a Na^+ ion in the respective Na3 binding sites. Snapshots obtained from the MD simulations are compared to the respective crystal structures (Figure 3.3). Introduction of the Na^+ ion at Na3' is seen to cause a relatively large perturbation of the N310 side chain in the closed structure, e.g., the Cα atom is displaced by about 2 Å from its crystal structure position (Figure 3.3A). In contrast, placement of a Na^+ ion at Na3 induces very little change in the N310 side chain configuration in the open structure (Figure 3.3B). If there is an unobserved third Na^+ ion in these crystal structures, its introduction in the correct location in the MD simulations should cause minimal disruption to the coordinating ligands, which is the case for the Na3 site but not for Na3'.

To ensure that the N310 side chain conformation is the only difference between

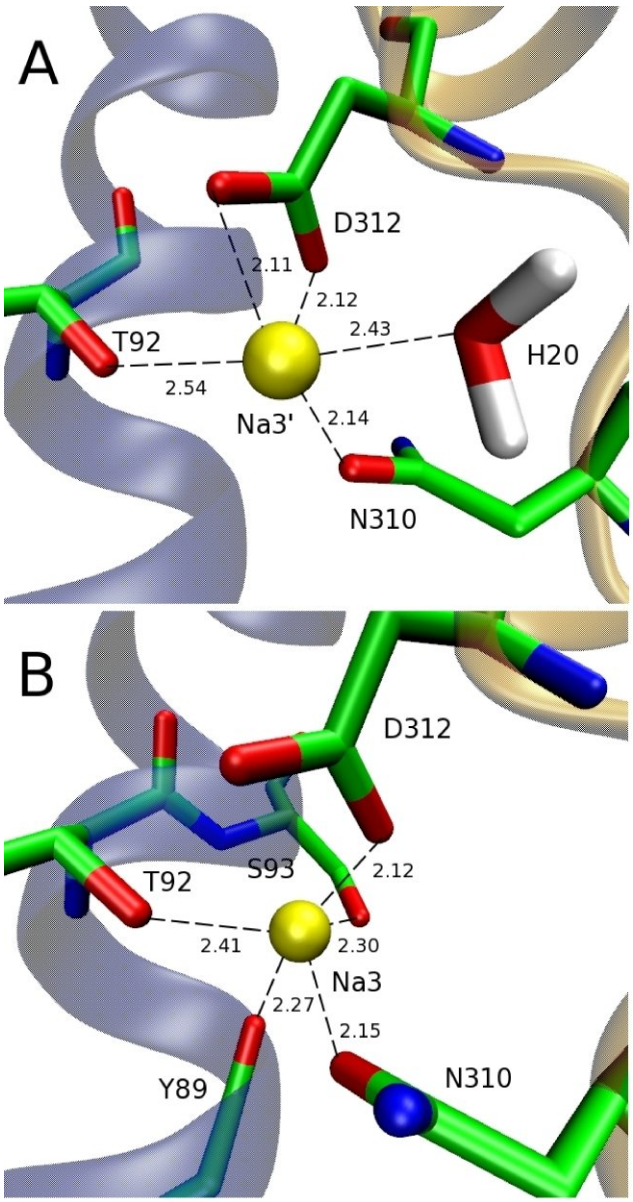

Figure 3.2: Comparison of the Na3 and Na3' sites. Snapshots of the Na3' (A) and Na3 (B) sites showing the oxygens (red) involved in the coordination of a Na^+ ion (yellow). Note the flipping of the N310 side chain between (A) and (B). Distances between the Na^+ ion and coordinating ligands are shown in angstroms.

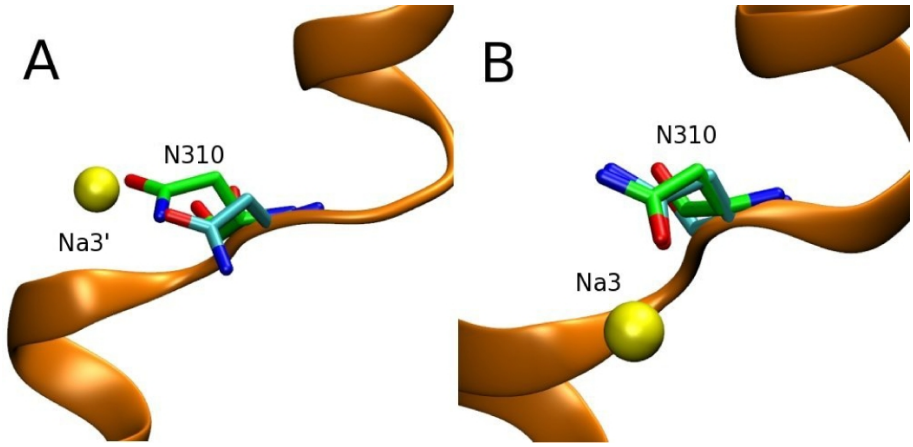

Figure 3.3: The N310 configurations in the closed and open structures contrasted with the MD results. Comparison of the N310 configuration from the crystal structure (green backbone) with that obtained from MD simulations (cyan backbone). Na^+ ions are shown as yellow spheres. (A) Closed structure (PDB 2NWX) vs. Na3' site, (B) Open structure (PDB 2NWW) vs. Na3 site.

the 2NWX and 2NWW structures with regard to the Na3 binding site, we have repeated the MD simulations with these two structures, but with the O and N atoms in the N310 side chain flipped in each case. The average Na^+–O distances listed in Table 3.1 essentially reveal an identical coordination shell structure in both cases, as long as the N310 side chains have the same conformation. Finally, we have calculated the free energy of binding of a Na^+ ion to the Na3 and Na3' sites in the absence of other ions and substrate. An example of the thermodynamic integration (TI) calculations illustrating the convergence of the binding free energy for the Na3 site is shown in Figure 3A.1. The binding free energy is found to be 4 kcal/mol lower for the Na3 site compared to that of Na3' (Table 3.2), which again supports Na3 as the correct binding site. Using the coordination for the Na3 and Na3' ions shown in Table 3.1 and the valence mapping algorithm used by Nayal et al.[100], we can also compare the valences of the Na3 and Na3' sites from the equation $v = (R/R_1)^{-N}$. In this expression, R_1 and N are empirical constants, and R is the distance of each coordinating atom from the sodium ion. To obtain the final value of v, we sum over all the coordinating oxygens for each case. We find a value of 1.18 for the valence of Na3' and 1.23 for the valence of Na3, making it difficult to assert which one is more likely based on valence calculations. On the other hand, both are similar to the values obtained for the valence of the Na1 site[95], which further indicates the validity of the D312 region on harbouring a sodium ion. To summarize the evidence presented, the Na3 site is more consistent with the crystal structure compared to the Na3' site and leads to a lower binding free energy. Therefore, the Na3 site identified in this study is the most likely binding site for the third Na^+ ion in Glt_{Ph}.

Table 3.1: Coordination of the third Na^+ ion showing the Na-O distances (in Å) for the closed (2NWX), and open (2NWW) structures, which give rise to the Na3' and Na3 binding sites, respectively. When the N310 side chain is flipped in the closed structure (2NWX*), the same coordination shell is observed as in the open structure (2NWW). Similarly, when the N310 side chain is flipped in the open structure (2NWW*), an identical Na^+ coordination to that of 2NWX is observed.

	Na3'		Na3	
	2NWX	2NWW*	2NWW	2NWX*
Y89 (O)	3.8 ± 0.3	3.9 ± 0.3	2.3 ± 0.1	2.3 ± 0.1
T92 (OH)	2.4 ± 0.1	2.4 ± 0.1	2.4 ± 0.1	2.4 ± 0.1
S93 (OH)	4.3 ± 0.2	4.1 ± 0.3	2.3 ± 0.1	2.4 ± 0.1
N310 (OD)	2.2 ± 0.1	2.2 ± 0.1	2.2 ± 0.1	2.2 ± 0.1
D312 (O1)	2.2 ± 0.1	2.2 ± 0.1	2.1 ± 0.1	2.1 ± 0.1
D312 (O2)	2.2 ± 0.1	2.2 ± 0.1	3.5 ± 0.3	3.6 ± 0.2
H20 (O)	2.4 ± 0.2	2.4 ± 0.2		

3.2.2 Order of binding of Na^+ and aspartate to Glt_{Ph}

The binding order of Na^+ ions and aspartate provides more evidence in this regard. It has been argued that, because access to the Na3 site is through the Na1 site and the path is very narrow, the Na3 site must be occupied first otherwise a Na^+ ion would not be able to access the Na3 site[97]. Here we provide evidence from the binding free energy calculations in favour of this proposal. As indicated in Table 3.2, the free energy of binding of a Na^+ ion is 7 kcal/mol lower for the Na3 site compared to the Na1 site. The large free energy difference between these two sites suggests that once a Na^+ ion binds to Na3 there is a higher probability it will remain bound than if it were to bind to Na3' where the free energy difference compared to Na1 is smaller (3 kcal/mol). The higher free energy of binding of Na^+ to Na3 suggests that this site is a more stable one compared to Na3'.

A related question is whether aspartate binds before or after Na1. As shown in Table 3.2, the second Na^+ ion binds to the Na1 site with a fairly large free energy (-7.1 kcal/mol) in the absence of aspartate. When aspartate is included in the MD simulations following the binding of two Na^+ ions at the Na3 and Na1 sites, the resulting system is extremely stable – all three aspartate molecules in the trimer have been observed to remain in their binding pockets for more than 50 ns. However, once the Na^+ ion is removed from the Na1 site, aspartate becomes unstable and dissociates – all three aspartate molecules in the trimer have been observed to move away from their binding pockets within 5 ns of MD simulations. These observations provide strong evidence that aspartate binds only after the binding of two Na^+ ions. Thus the proposed binding order of the ligands to the transporter is Na3, Na1,

Table 3.2: Binding free energies for Na^+ ions in various locations (in kcal/mol). The interaction energy (ΔG_{int}), entropic contributions (ΔG_{tr}) and the total binding energy (ΔG_b) are listed separately. The interaction free energy differences are obtained from the average of the forward and backward transitions. Presence of a second Na^+ ion is indicated in parentheses. Errors are estimated from block data analysis using 50 ps windows. The Na2 site is not considered because it is formed only after the HP2 gate shuts.

Ligand	ΔG_{int}	ΔG_{tr}	ΔG_b
Na3	-23.3 ± 1.2	4.6	-18.7 ± 1.2
Na3'	-19.3 ± 0.9	4.6	-14.7 ± 0.9
Na1	-16.2 ± 1.3	4.9	-11.3 ± 1.3
Na1 (Na3)	-11.9 ± 1.3	4.8	-7.1 ± 1.3

aspartate and Na2.

3.2.3 The Na3 site in the archaeal aspartate transporter Glt_{Ph}

To confirm the alternate Na3 binding site described above, site-directed mutagenesis was performed on the Na3 coordinating residues T92, S93, N310 and D312. Y89 contributes a backbone carbonyl oxygen to the Na3 site and therefore was not investigated by mutagenesis. All mutations were well tolerated by Glt_{Ph}; similar protein yields were obtained and all mutant transporters eluted as symmetrical peaks following size exclusion chromatography at the same point as wild type Glt_{Ph} (Figure 3.4A). Purified protein was reconstituted into liposomes and function was assayed by examining ^{3}H-L-aspartate uptake.

The N310A and D312A transporters did not exhibit ^{3}H-L-aspartate uptake over background levels and were not investigated further (Figure 3.4B). This is not surprising as mutation of either of these residues in EAAT2 results in a non-functional transporter[101]. In addition, these two residues are in the highly conserved NMDGT motif that contains residues important for substrate binding[15]. Both T92A and S93A are functional aspartate transporters (Figure 3.4B) but display altered aspartate affinity and changes in the Na^+ dependence of transport. T92A transports aspartate with a slightly higher affinity than wild type Glt_{Ph}; $K_{0.5}$ of 61 ± 10 nM compared to 100 ± 10 nM for wild type Glt_{Ph}. In contrast, S93A exhibits a reduction in aspartate affinity with $K_{0.5}$ of 370 ± 60 nM (Figure 3.4C). The Na^+ dependence of aspartate transport was measured by monitoring the uptake of ^{3}H-L-aspartate under a range of Na^+ concentrations, from 0.1 mM to 300 mM. Both T92A and S93A showed a reduction in Na^+ affinity compared to wild type Glt_{Ph}; T92A $K_{0.5}$ of 19 ± 1 mM and S93A $K_{0.5}$ of 44 ± 2 mM compared to Glt_{Ph} $K_{0.5}$ of 3.9 ± 0.4 mM (Figure 3.4D) but there was no change in the Hill coefficient (Glt_{Ph},

Table 3.3: Effect of the T92A and S93A mutations on the binding free energies of Na^+ ions at the Na3 and Na1 sites (in kcal/mol). Mutations significantly reduce the binding free energies of Na3 but not of Na1. Presence of a second Na^+ ion is indicated in parentheses.

Ligand	Glt_{Ph}	T92A	S93A
Na3	-18.7 ± 1.2	-11.2 ± 1.4	-12.8 ± 1.2
Na1 (Na3)	-7.1 ± 1.3	-6.7 ± 1.2	-6.4 ± 1.4

2.3 ± 0.5; T92A, 2.1 ± 0.2; S93A, 1.9 ± 0.2). As T92 and S93 are not close enough to coordinate Na1 or Na2, we propose that these mutations are affecting the Na3 site. Indeed, when the binding free energy of Na3 is calculated in the presence of the T92A and S93A mutations, it is reduced by 7.5 kcal/mol and 5.9 kcal/mol respectively, while the binding free energy of Na1 is unaltered (Table 3.3). Interestingly, the maximal velocity of ^{3}H-L-aspartate transport measured in the presence of 100 mM Na^+ and saturating substrate concentrations by T92A is 42 ± 1 nmol/mg protein/min and S93A is 58 ± 3 nmol/mg protein/min. These rates are increased compared to the maximal velocity of ^{3}H-L-aspartate transport by Glt_{Ph} (7.5 ± 0.5 nmol/mg protein/min) (Figure 3A.2). The increased rate of transport observed for the mutant transporters could be due to the reduced affinity of Na3 for its binding site. Na3 may unbind more easily from T92A and S93A than wild type Glt_{Ph} and thus facilitate the turnover of the transporter. Taken together, the simulation and experimental results strongly support a role for T92 and S93 in coordinating Na3 in Glt_{Ph}.

3.2.4 The Na3 site in the human glutamate transporter EAAT1

Glt_{Ph} shares ∼36% amino acid identity with the EAATs and this conservation is even higher in the transport domain consisting of hairpin (HP)1, transmembrane domain (TM) 7, HP2 and TM8 where amino acid identity increases to ∼ 60%. The recent crystal structure of the inward facing Na^+ and aspartate bound Glt_{Ph} revealed that TM3 and TM6 are also part of the transport domain core that is proposed to move over 18 Å during the transport cycle[16]. TM3 is not as highly conserved as the rest of the transport domain, Glt_{Ph} shares 40% amino acid identity with EAAT1 in this region, but there are several residues that are conserved throughout the Na^+ dependent transporters of this family including Glt_{Ph}, EAAT1-5 and the neutral amino acid transporter, ASCT1. T92 is highly conserved, while S93 is a threonine residue in the rest of the family (Figure 3.5A). Interestingly, T92 is replaced with an isoleucine residue in the H^+ dependent glutamate transporters from *Bacillus stearothermophilus* ($GltT_{Bs}$) and *Esherichia coli* ($GltP_{Ec}$). To investigate if the third Na^+ site predicted by simulation and experimental studies on Glt_{Ph} is

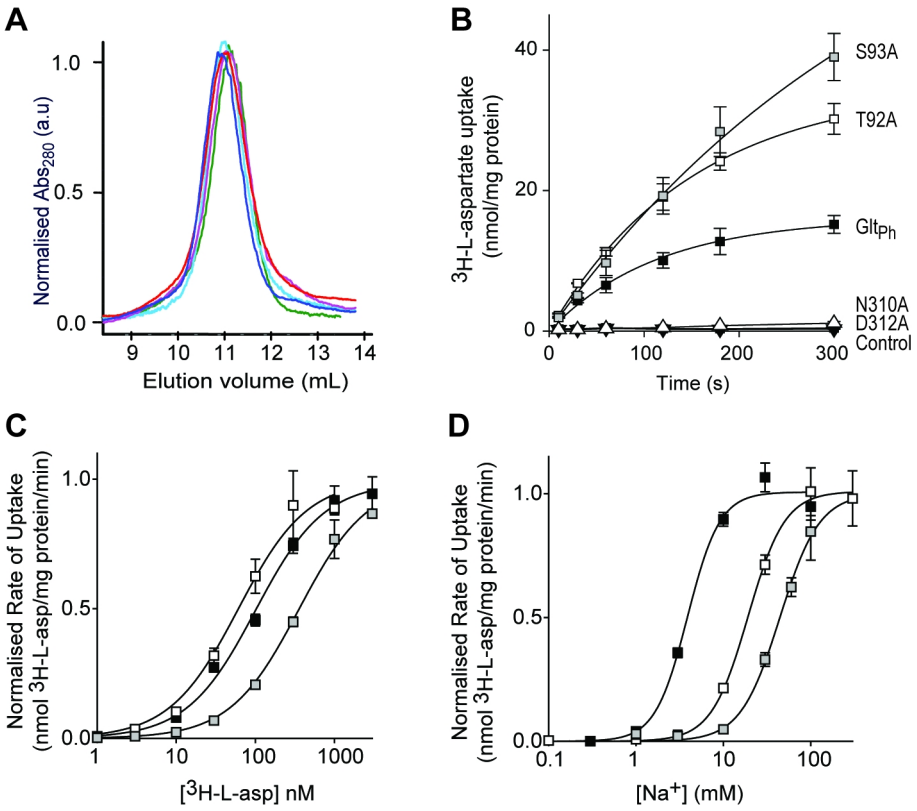

Figure 3.4: T92A and S93A mutations in Glt_{Ph} have reduced Na^+ affinity. (A) Size exclusion column profile for wild type Glt_{Ph} (dark blue), T92A (red), S93A (green), N310A (pink) and D312A (cyan). (B) Uptake of 100 nM ^{3}H-L-aspartate in the presence of 100 mM NaCl for Glt_{Ph} (black squares), T92A (white squares), S93A (grey squares), N310A (white triangles) and D312A (black triangles). Control levels are from uptake performed in the presence of internal buffer (100 mM KCl, 20 mM HEPES/Tris pH 7.5) (C) ^{3}H-L-aspartate concentration-dependent transport in the presence of 100 mM NaCl by Glt_{Ph} (black squares), T92A (white squares) and S93A (grey squares). (D) Na^+ concentration-dependent transport of 100 nM ^{3}H-L-aspartate for Glt_{Ph} (black squares), T92A (white squares) and 500 nM ^{3}H-L-aspartate for S93A (grey squares). Data in (C) and (D) are normalised to the maximal velocity of transport and all data are from the mean of at least 3 separate experiments ± s.e.m.

conserved in the human glutamate transporters the T130A (equivalent to T92) and T131A (equivalent to S93) mutations were introduced into EAAT1. The residues equivalent to N310 and D312 were not investigated as these transporters have been shown to be non-functional in EAAT2[101].

Application of 100 μM L-glutamate to oocytes expressing EAAT1 elicits a current of 395 ± 20 nA when the oocytes are clamped at -60 mV (Figure 3.5B). Application of 100 μM L-glutamate to oocytes expressing the mutant transporters T130A and T131A also activates a conductance at -60 mV but it is reduced in amplitude to 21 ± 2 nA and 25 ± 2 nA for T130A and T131A respectively (Figure 3.5B). To further investigate the impact of the T130A and T131A mutations in EAAT1 the $K_{0.5}$ for L-glutamate transport and the $K_{0.5}$ for the Na^+-dependence of L-glutamate transport were determined. Oocytes expressing EAAT1, T130A or T131A were clamped at -60 mV and increasing doses of L-glutamate were applied. For wild type EAAT1 the $K_{0.5}$ value for L-glutamate is 15 ± 1 M (Figure 3.5C). In contrast, the L-glutamate affinity for T130A and T131A is reduced with $K_{0.5}$ values of 87 ± 3 μM and 124 ± 11 μM, respectively (Figure 3.5C).

To investigate the Na^+-dependence of transport the current elicited by saturating L-glutamate concentrations (300 μM for EAAT1 and 1 mM for T130A and T131A) in the presence of increasing amounts of Na^+ was determined. For wild type EAAT1 the $K_{0.5}$ for Na^+ is 28 ± 4 mM and the Hill coefficient is 2.8 ± 0.4. The Na^+ dependence of L-glutamate transport for T130A and T131A was significantly altered compared to wild type EAAT1 (Figure 3.5D). For both mutant transporters, the highest Na^+ concentration tolerated by Xenopus laevis oocytes (150 mM) was not saturating and the data could not be fit to the Hill equation to determine a $K_{0.5}$ for Na^+ or a Hill coefficient. Nevertheless, it is evident that both of these mutations have significantly reduced the ability of Na^+ to support glutamate transport.

3.3 Discussion

Molecular dynamics is a powerful tool that can simulate movements and binding events from static snapshots of proteins obtained from X-ray crystallography, but functional studies are vital to confirm these predictions and maximise the information obtained from MD simulations. In this study, we have combined these techniques to explore the third Na^+ binding site in the glutamate transporter family using Glt_{Ph} and EAAT1 as representative transporters.

The crystal structure of the aspartate transporter Glt_{Ph} with bound substrate and two Na^+ ions was an important advance in our understanding of the transport mechanism of the glutamate transporter family[15]. Substrate transport by the human glutamate transporters EAAT2[87] and EAAT3[13], and also by Glt_{Ph}[37], is coupled to the co-transport of 3 Na^+ ions, yet only 2 Na^+ binding sites were identified in the Glt_{Ph} structure[15] leaving the location of the third Na^+ ion unknown. Previous studies using electrostatic calculations, MD simulations and mutagenesis studies

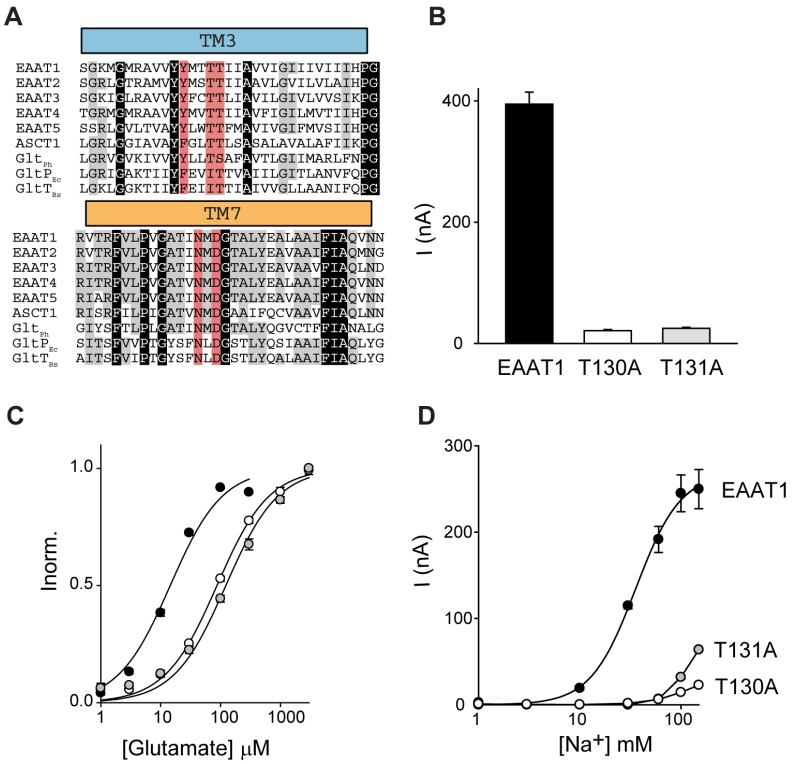

Figure 3.5: T130A and T131A mutations in EAAT1 have reduced Na^+ affinity. (A) Amino acid alignment of TM3 and TM7. Alignment was made using ClustalW2 [102] and adjusted manually. T130A and T131A mutations in EAAT1 have reduced Na^+ affinity. (A) Amino acid alignment of TM3 and TM7. Alignment was made using ClustalW2 [102] and adjusted manually. Amino acid sequences are; human (h)EAAT1 (NP_004163.3), hEAAT2 (NP_004162.2), hEAAT3 (NP_004161.4); hEAAT4 (NP_005062.1); hEAAT5 (NP_006662.3); hASCT1 (NP_003029.2); *Pyrococcus horikoshii* Glt_{Ph} (NP_143181); *Escherichia coli* $GltP_{Ec}$ (EGT70436.1); *Bacillus stearothermophilus* $GltT_{Bs}$ (P24943.1). Red shading indicates the residues that form the Na3 site, grey shading indicates conserved residues and black shading indicates residues absolutely conserved. (B) Average current activated by the application of 100 μM L-glutamate to oocytes clamped at -60 mV expressing EAAT1 (black), T130A (white) and T131A (grey). Red shading indicates the residues that form the Na3 site, grey shading indicates conserved residues and black shading indicates residues absolutely conserved. (B) Average current activated by the application of 100 μM L-glutamate to oocytes clamped at -60 mV expressing EAAT1 (black), T130A (white) and T131A (grey). (C) L-glutamate concentration-dependent currents for EAAT1 (black), T130A (white) and T131A (grey). (D) Na^+ concentration-dependent currents in the presence of 300 μM L-glutamate for EAAT1 (black circles), and 1 mM L-glutamate for T130A (white circles) and T131A (grey circles). Data in (C) are normalised to the current at saturating glutamate concentrations and all data are from the mean of at least 3 separate experiments $\pm$ s.e.m.

have proposed several different sites for Na3[30,31,95–97]. Here, we present a new location of the Na3 site as determined by MD simulations of both the open and closed Glt_{Ph} structures. The veracity of this site has been further tested by calculating the binding free energy of substrate and ions and site-directed mutagenesis experiments, which corroborate our proposed Na3 location.

The binding site for Na3 is made up of residues from TM3 and TM7 which are both in the transport domain that is proposed to undergo a large conformational change during substrate translocation[16]. T92 (TM3), S93 (TM3), N310 (TM7) and D312 (TM7) all donate side chain oxygens to the binding site while Y89 (TM3) contributes a backbone carbonyl oxygen (Figure 3.6A). To verify the contribution of the side chains of T92, S93, N310 and D312 to the Na3 site predicted by MD simulations we used site-directed mutagenesis to change these residues in Glt_{Ph} to alanine. The N310A and D312A transporters were non-functional. These residues reside in the highly conserved NMDGT motif that contains residues important for substrate binding[15] and mutations of the equivalent residues in EAAT2 leads to non-functional transporters[101]. In contrast, the T92A and S93A transporters were functional aspartate transporters but these mutations result in a reduced affinity for Na^+ compared to wild type Glt_{Ph}, while the Hill coefficient is unaltered. This data suggests that the ability of Na^+ to support aspartate transport is diminished but the number of Na^+ ions coupled to transport is unchanged. There is also some variation in the affinity of aspartate for these mutant transporters. T92A has a slightly higher aspartate affinity, while S93A displays a reduced aspartate affinity compared to wild type Glt_{Ph}. As aspartate binding and transport is coupled to Na^+, it is not surprising that disruption of the Na3 site also affects the $K_{0.5}$ for aspartate transport. Further evidence to support the location of Na3 comes from the binding free energies calculated when the T92A and S93A mutations were introduced into Glt_{Ph}. Both of these mutations reduced the binding free energy calculated for Na3, but did not affect the binding free energy calculated for Na^+ binding to the Na1 site (Table 3.3). To translate the information gained about the location of the Na3 site in Glt_{Ph} to the human glutamate transporters, parallel mutagenesis experiments were performed in EAAT1. The residues corresponding to T92 and S93 in EAAT1 were also changed to alanine and these mutant transporters displayed a similar phenotype to the mutant transporters in Glt_{Ph} in that the ability of Na^+ to support glutamate transport was substantially reduced.

Why hasn't Na3 been observed in any of the crystal structures of Glt_{Ph}? The resolution of the available Glt_{Ph} structures is not good enough to visualise Na^+ ions directly and the heavy atom thallium (Tl^+) was used to probe cation binding sites on Glt_{Ph}. Two Tl^+ sites were identified and ion competition studies confirmed that Na1 and Na2 were genuine Na^+ sites, but Tl^+ binding was not observed at the Na3 site under these conditions[15]. The Na^+ affinity for Na3 is considerably higher than for Na1 and Na2 as suggested by the binding free energies (Table 3.2) and as Tl^+ has a larger ionic radius than Na^+ (1.40 vs. 0.95 Å), it simply may not be able to bind to the Na3 site. In addition, the pathway between Na1 and Na3 is narrow[97] and as

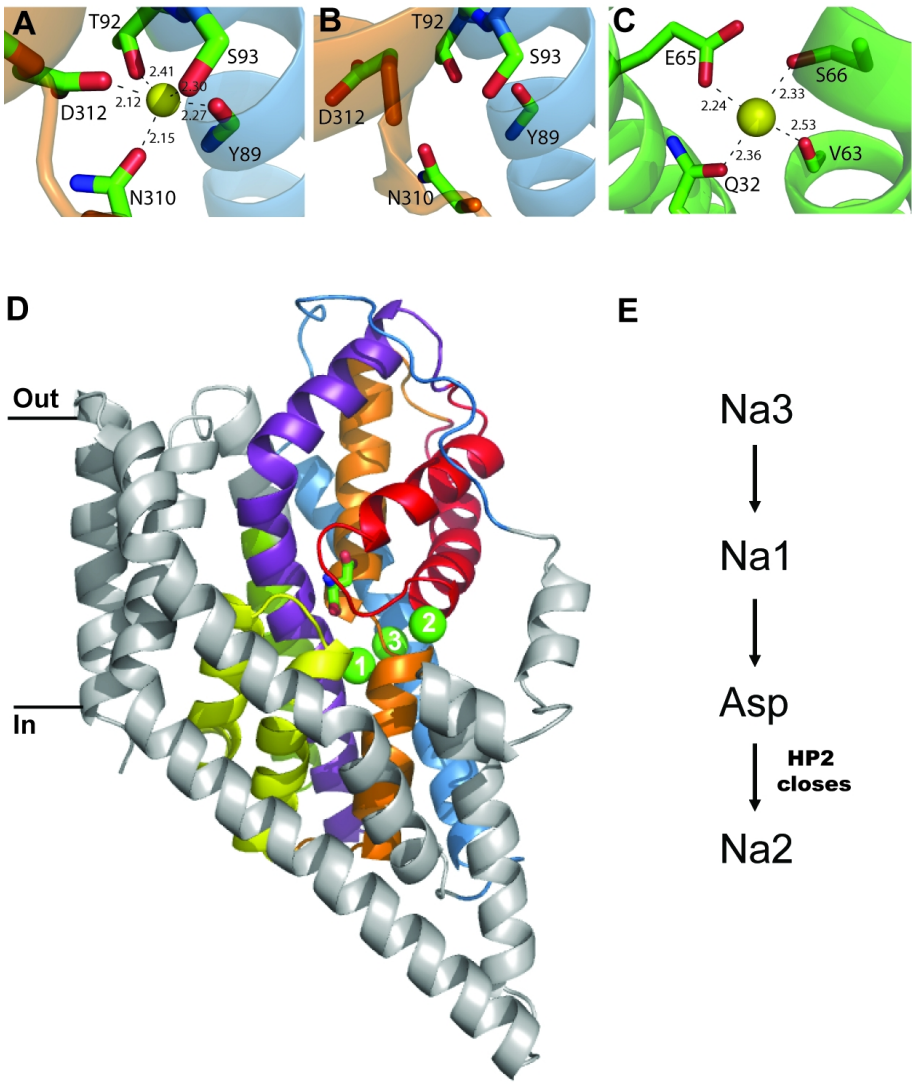

Figure 3.6: The third Na^+ binding site in Glt_{Ph}. (A) The coordination of the Na3 binding site identified in this study. TM3 (blue) and TM7 (orange) are shown in cartoon and the coordinating residues are shown in stick representation and labelled. Na3 is shown as a yellow sphere. (B) The area of the Na3 site in the inward occluded structure of Glt_{Ph} (PDB 3KBC); colouring as in A. (C) Na^+ binding site on the rotor ring (c ring) of the F-Type Na^+ ATPase from *Ilyobacter tartaricus* (PDB 1YCE). V63, S66 are from the A subunit and Q32, E65 are from the B subunit of the c ring. Na^+ ion is shown as a yellow sphere. (D) One protomer of Glt_{Ph} with bound aspartate, Na1, Na2 and Na3. The transport domain is shown in colour; TM3 (blue), TM6 (green), HP1 (yellow), TM7 (orange), HP2 (red) and TM8 (purple) and the 3 Na^+ ions are shown as green spheres and numbered. (E) The proposed order of binding for substrate and ions to Glt_{Ph}. All distances are in angstroms and structure figures were made using Pymol.[98]

Tl^+ was soaked into Glt_{Ph} crystals that already contained Na^+, exchange between Na^+ and Tl^+ may not have been possible at the Na3 site. We have compared the Na3 site of Glt_{Ph} identified in this study (Figure 3.6A) with the Na^+ binding site on an unrelated protein, the F-type Na^+-ATPase from *Ilyobacter tartaricus*[103]. The F-type Na^+-ATPase is a molecular motor that utilises a Na^+ gradient to synthesise ATP. The Na^+ binding sites (11 identical sites) are located on the rotor ring (or c ring) and investigation of these Na^+ sites (Figure 3.6C) reveals that they share a number of similarities with Na3 site in Glt_{Ph} (Figure 3.6A). Both sites are formed by the side chain oxygen of an acidic amino acid (D312 in Glt_{Ph}; E65 in c ring), a polar amino acid (N310 in Glt_{Ph}; Gln32 in c ring), a backbone carbonyl oxygen (Y89 in Glt_{Ph}; Val63 in c ring) and the hydroxyl oxygen of a serine residue (S93 in Glt_{Ph}; S66 in c ring). The Na3 binding site in Glt_{Ph} has an additional coordinating ligand from T92. The striking similarities between these two Na^+ binding sites on proteins unrelated in sequence or function further supports the correct identification of the third Na^+ binding site of Glt_{Ph} in this study.

Two different orders for Na^+ and aspartate binding to Glt_{Ph} have been proposed where aspartate binds either before or after Na1[97]. The binding free energy calculations for Na^+ and MD simulations of aspartate binding to Glt_{Ph} presented in this study suggest an order of binding where Na3 binds first, followed by Na1. The binding of these two Na^+ ions creates a favourable site for aspartate to bind. Finally, Na^+ binding to the Na2 site secures HP2 down over the substrate and results in the 'occluded' state observed in previous crystal structures (Figure 3.6D,E)[14,15]. The transport domain with aspartate and Na1-3 bound is then predicted to undergo a conformational change and move to the intracellular side of the membrane where it is ready to release its cargo into the cell[16]. Interestingly, examination of the Na3 site in the 'inward occluded' structure[16] reveals that two of the coordinating ligands for Na3 (the side chains of N310 and D312) appear to be facing away from this site (Figure 3.6B). The resolution of this structure is moderate (3.8 Å) and side chain placement may not be accurate, but if these movements do occur, the affinity of Na^+ at the Na3 site in the intracellular occluded state would be reduced and may facilitate unbinding of Na^+. Further molecular dynamics simulations of the 'inward occluded' structure and/or a structure of an inward facing 'apo' state of Glt_{Ph} are required to confirm these side chain movements and will shed light on the process of Na^+ and substrate unbinding.

Three dimensional crystal structures of prokaryotic homologues of membrane proteins have resulted in a major advance in our understanding of the structure and mechanism of their human counterparts. Here, we present the precise location of the third Na^+ site in Glt_{Ph} and show that this site is conserved in EAAT1. This information is required to develop accurate homology models of the EAATs which will provide useful starting points to investigate the transport mechanism of the human glutamate transporters. Such studies may help to explain the differences between Glt_{Ph} and the EAATs including differences in transport rates, substrate selectivity and K^+ dependence of transport.

3.4 Materials and Methods

3.4.1 Ethics statement

Frogs used in this study were anaesthetised to minimize suffering and all surgical procedures followed a protocol approved by The University of Sydney Animal Ethics Committee (protocol # K21/2-2010/3/5269) under the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes.

3.4.2 Model system and MD simulations

The simulation systems are prepared using the software VMD[104]. The crystal structure of the Glt_{Ph} trimer is embedded in a 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) phospholipid bilayer and solvated in a box of water molecules with physiological concentration of NaCl. The final system consists of the trimer, 239 lipid molecules, 15688 water molecules and 35 Na^+ and 35 Cl^- ions. Glt_{Ph} has a net charge of +6e, so to keep the system neutral 6 Cl^- ions are added in the apo system (more Cl^- ions are added when the bound Na^+ ions are included to preserve neutrality). After the system is built, it is equilibrated in two stages: first, the coordinates of the protein atoms are fixed and the system is equilibrated with 1 atm pressure coupling until the correct water and lipid densities are obtained. The x and y-dimensions of the simulation box are then fixed (at 115 and 113 Å, respectively), and pressure coupling is applied in the z-direction (the average z length is 71 Å). In the second stage, the protein is gradually relaxed in 2.4 ns MD simulations by reducing the restraints on the protein atoms in several steps. The system is further equilibrated for 5 ns with only a small (0.1 kcal/mol/Å^2) restraint left on the backbone atoms of the protein. This helps to preserve the structural integrity of the protein during long MD simulations. This procedure is repeated for both the closed and open structures, and the final states obtained have been used in all subsequent MD simulations and free energy perturbation (FEP) calculations.

MD simulations are performed using the NAMD package, version 2.7b2[59] with the CHARMM22 force field[55] including the CMAP corrections[105]. The temperature is kept at 300 K using the Langevin damping method, and the pressure is kept at 1atm using the Langevin piston method. The Lennard-Jones interactions are switched off for distances over 12 Å with a switching distance of 10 Å. Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. A time step of 2 fs is used in all MD simulations.

3.4.3 Free energy calculations

The binding free energies of Na^+ ions are calculated using the expression, $\Delta G_b = \Delta G_{int} + \Delta G_{tr}$, where the first term gives the free energy difference for the interactions of the ion in the binding site and bulk (i.e. translocation energy) and the second

term measures the free energy loss due to reduction in translational entropy upon binding[106]. The latter can be estimated from the rms fluctuations of the ion in the binding site (σ_x,σ_y,σ_z) as $\Delta G_{tr} = -k_B T \ln[(2\pi e)^{3/2}\sigma_x\sigma_y\sigma_x/V_0]$, where $V_0 = 1660 Å^3$ which is the reference volume for the standard concentration[107]. The interaction energy is calculated using both the free energy perturbation (FEP) and the thermodynamic integration (TI) methods[52,108]. Because the two methods have yielded essentially the same results, here we report only the TI calculations. Initially, the Na^+ ion is placed at the appropriate binding position in the protein and the system is equilibrated. In the forward calculation, the bound Na^+ ion is alchemically transformed to a water molecule while a water molecule in the bulk is transformed to a Na^+ ion simultaneously. After equilibrating the last window in the forward calculation, a backward calculation is performed, where the opposite transformations are performed bringing the system back to the initial state with a bound Na^+ ion. Any significant difference between the forward and backward calculations points to hysteresis effects in the TI calculations. Otherwise the binding free energy of the ion is determined from the average of the forward and backward calculations.

Integrals in the TI calculations are evaluated using a seven-point Gaussian quadrature, which has been shown to be sufficiently accurate for calculation of the binding free energies of ions[108]. The simulation systems for the seven windows are adapted from the parallel FEP calculations because the windows are more closely spaced in FEP and hence equilibrate faster. Each window is equilibrated for 0.5 ns followed by a 1 ns production run. Convergence of the free energy results are checked from the running averages which become flat once sufficient sampling is obtained.

The same procedure is used to calculate the binding free energy of the Na+ ions at the Na1 and Na3 sites in the T92A and S93A mutations of the transporter. The mutations are implemented using the MUTATOR plug-in from the VMD software, where the side chain of the chosen residue is mutated while the backbone coordinates remain the same. To make sure that the mutated residue is in a stable conformation, the resulting structure is energy minimized and equilibrated for 3 ns in MD simulations before starting the free energy calculations.

3.4.4 Site-directed mutagenesis

Site-directed mutagenesis was performed using a polymerase chain reaction (PCR) based method[109] and Velocity DNA polymerase (Bioline). All mutations were sequenced on both strands by Dye Terminator Cycle Sequencing (ABI PRISM, PerkinElmer Life Sciences). The wild type EAAT1 and mutant transporter cDNAs were linearized with SpeI and cRNA transcribed with T7 RNA polymerase using the mMessage mMachine T7 kit (Ambion Inc.).

3.4.5 Protein purification and reconstitution

Glt_{Ph} protein was purified as described previously[91]. Briefly, membranes were isolated, solubilized with (40 mM) n-dodecyl-β-D-maltopyranoside ($C_{12}M$, Anatrace), and protein was purified using Ni-NTA resin (Qiagen). The $C_{12}M$ concentration in the buffer was reduced to 2 mM before addition to Ni-NTA beads. The histidine tag was subsequently removed by digestion with thrombin (10 U/mg protein) and the protein further purified on a size exclusion column where the detergent was exchanged to 7 mM n-decyl-β-D-maltopyranoside ($C_{10}M$, Anatrace).

Pure protein was reconstituted into liposomes using a method modified from Gaillard et al., 1996[110]. *Escherichia coli* polar lipids and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids), at a ratio of 3:1, were mixed, dried under nitrogen, and resuspended in internal buffer (100 mM KCl, 20 mM HEPES pH 7.5 or the appropriate internal buffer as indicated below). The lipid mixture was briefly sonicated using a cylindrical sonicator (Labatory Supplies Co.) and the lipid suspension was frozen in liquid nitrogen and thawed at 45°C several times. Liposomes were formed by extrusion through 400 nm polycarbonate membranes (Avanti Polar Lipids) and were treated with Triton X-100 at a 0.5:1 (w/w) detergent to lipid ratio prior to the addition of protein at 0.25 μg protein/mg lipid[35]. The protein/lipid mixture was left at room temperature for 30 minutes before detergent was removed using SM2 Biobeads (Biorad). The protein/lipid mixture was incubated, with gentle agitation, with three consecutive batches of the Biobeads (80 mg/ml). Proteoliposomes were concentrated by centrifugation at 150,000g for 30 minutes in a Beckman Optima TLX centrifuge, resuspended at 100 mg lipid/mL and either used immediately or flash frozen in liquid nitrogen and stored at -80°C.

3.4.6 Proteoliposome transport assay

^{3}H-L-aspartate transport by wild type Glt_{Ph} and mutant Glt_{Ph} transporters was assayed using a protocol modified from Gaillard et al., 1996[110]. Proteoliposomes were loaded with internal buffer (100 mM KCl, 20 mM HEPES pH 7.5) by several freeze/thaw cycles followed by extrusion as described above. The uptake reaction was initiated by diluting the proteoliposomes (100 mg lipid/mL) 133 fold into reaction buffer pre-warmed to 30°C. The reaction buffer contained: 100 mM NaCl, 20 mM HEPES pH 7.5, 1 μM valinomycin and the indicated concentrations of ^{3}H-L-aspartate. At each time point, a 200 μL aliquot was removed and diluted 10 fold into ice cold quench buffer (100 mM LiCl, 20 mM HEPES pH 7.5), followed by immediate filtration over nitrocellulose filters (0.22 μm pore size, Millipore). The filters were washed once with 2 mL of ice cold quench buffer and assayed for radioactivity using a Trilux beta counter (Perkin Elmer). The Na^+ dependence of ^{3}H-L-aspartate transport was measured by varying extraliposomal Na^+ from 0.1 to 300 mM Na^+. Choline (Ch^+) was used to balance osmolarity and the intraliposomal K^+ concentration was adjusted to 300 mM. ^{3}H-L-aspartate concentrations used for the Na^+

titrations were 100 nM for Glt_{Ph} and T92A and 500 nM for S93A; concentrations close to the aspartate $K_{0.5}$ value for each transporter. Background levels of uptake were measured by diluting proteoliposomes into internal buffer (100 mM KCl, 20 mM HEPES pH 7.5) containing 1 μM valinomycin and the indicated concentrations of ^{3}H-L-aspartate.

3.4.7 Electrophysiology

All chemicals were obtained from Sigma unless otherwise stated. Stage V oocytes were harvested from Xenopus laevis as described previously[111]. 4 ng of cRNA was injected into oocytes and incubated in standard frog Ringers solution (ND96: 96 mM NaCl, 2 mM KCl, 1 mM MgCl2, 1.8 mM CaCl2, 5 mM HEPES, pH 7.5) supplemented with 50 μg/mL gentamycin, 2.5 mM sodium pyruvate, 0.5 mM theophylline at 16-18^{o}C.

Current recordings were made using the two-electrode voltage clamp technique with a Geneclamp 500 amplifier (Axon Instruments, Foster City, CA) interfaced with a PowerLab 2/20 chart recorder (ADI Instruments, Sydney, Australia) using the chart software and a Digidata 1322A (Axon Instruments) controlled by an IBM-compatible computer using the pClamp software (version 10, Molecular Devices, Union City, California). When measuring Na^+ dependence of substrate induced currents, saturating concentrations of L-glutamate were used (300 μM for EAAT1 and 1 mM for T130A and T131A) and the concentration of Na^+ varied from 1 to 150 mM. $NMDG^+$ was used as a Na^+ substitute to maintain ionic strength of the buffers.

3.4.8 Analysis of kinetic data

Substrate responses were fitted by least squares as a function of substrate concentration to $I/I_{max} = [S]/([S] + K_{0.5})$, where I is the current, I_{max} is the maximal current, $K_{0.5}$ is the concentration of substrate that generates half-maximal response, and [S] is the substrate concentration. Na^+ concentration responses were fit to the Hill equation $I/I_{max} = [S]^n/([S]^n + (K_{0.5})^n)$ where n is the Hill coefficient and all other terms are as described above. For ^{3}H-L-aspartate uptake by Glt_{Ph}, initial rates were calculated from the linear portion of the curve and all data represent the mean ± s.e.m. of at least 3 experiments.

3.5 Acknowledgements

We thank Cheryl Handford for expert technical assistance, all those who maintain the University of Sydney Xenopus laevis colony and members of the Transporter Biology Group for helpful discussions. Calculations were performed using the HPC facilities at the National Computational Infrastructure (Canberra) and ULAKBIM (Ankara).

3.A Appendix for Chapter 3

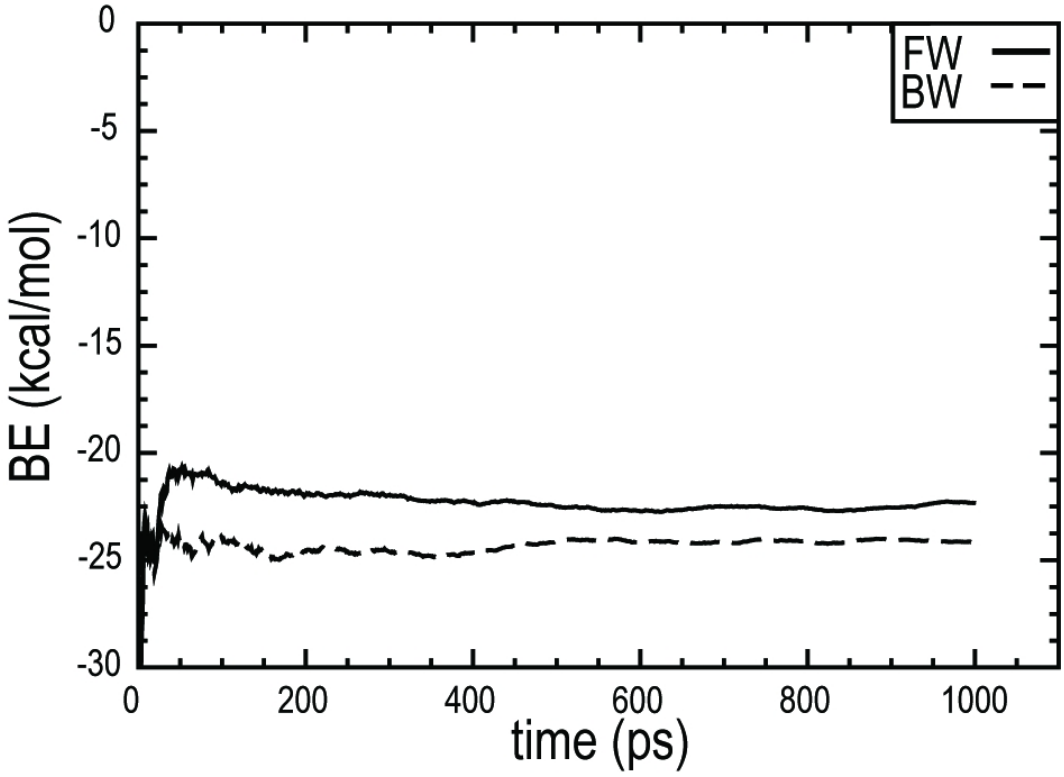

Figure 3A.1: TI calculations for Na^+. Convergence of the binding free energy of a Na^+ ion to the Na3 site is demonstrated using the running averages of the free energies, which flatten out as the data are accumulated. Binding free energies for the negative of the forward (binding site $\rightarrow$ bulk) and the backward transitions are shown with solid and dashed lines, respectively. The final results read from the end points of the curves, are -22.3 kcal/mol (forward) and -24.2 kcal/mol (backward), whose average gives the interaction energy value (ΔG_{int}) quoted in Table 3.2.

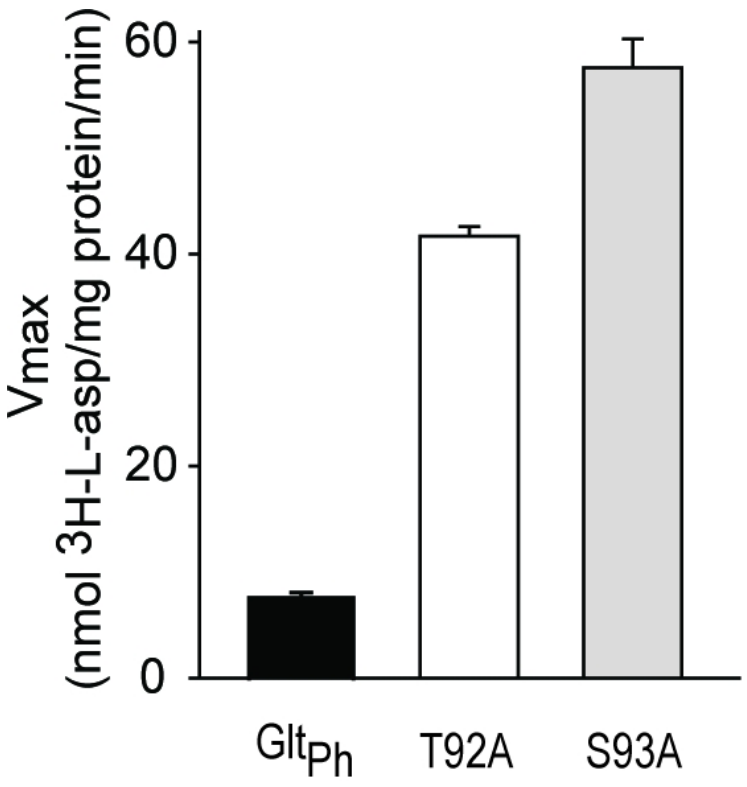

Figure 3A.2: Maximal velocity of ^{3}H-L-aspartate transport. The maximal rate of transport in the presence of saturating aspartate concentrations and 100 mM NaCl for Glt_{Ph} (black), T92A (white) and S93A (grey).

Chapter 4

Free Energy Simulations of Ligand Binding to the Aspartate Transporter $\mathrm{Glt_{Ph}}$

ABSTRACT
Glutamate/Aspartate transporters cotransport three Na^+ and one H^+ ions with the substrate and countertransport one K^+ ion. The binding sites for the substrate and two Na^+ ions have been observed in the crystal structure of the archaeal homologue $\mathrm{Glt_{Ph}}$, while the binding site for the third Na^+ ion has been proposed from computational studies and confirmed by experiments. Here we perform detailed free energy simulations of $\mathrm{Glt_{Ph}}$, giving a comprehensive characterization of the substrate and ion binding sites, and calculating their binding free energies in various configurations. Our results show unequivocally that the substrate binds after the binding of two Na^+ ions. They also shed light into Asp/Glu selectivity of $\mathrm{Glt_{Ph}}$, which is not observed in eukaryotic glutamate transporters.

4.1 Introduction

Glutamate transporters (GlTs) remove the excess extracellular glutamate using the electrochemical gradient of Na^+ ions[12]. In mammalian GlTs (called excitatory amino acid transporters, or EAATs for short), the transport mechanism involves cotransport of three Na^+ and one H^+ ions, and countertransport of one K^+ ion[13,85]. A large amount of functional data has been accumulated on GlTs in the last two decades[12] but their interpretation has proved difficult in the absence of any structures. As in ion channels, the first crystal structure of a GlT was determined in prokaryotes, namely that of Glt_{Ph} from *Pyrococcus horikoshii*[14]. The binding sites for the substrate and two ions were identified in a subsequent crystal structure of Glt_{Ph}[15]. Initial functional studies of Glt_{Ph} revealed that the transport mechanism was independent of H^+ and K^+, but they were inconclusive on whether two or three Na^+ ions were cotransported in each cycle[15,35]. This issue was settled in a recent experiment with radioactive $^{22}Na^+$ ions and [^{14}C] aspartate, which showed conclusively that three Na^+ ions are cotransported with the substrate as in EAATs[37].

Glt_{Ph} has about 36% amino acid sequence identity with the EAATs, which is somewhat low. However, the sequence identity is much higher for the residues forming the binding pocket ($\sim 65\%$), and most of the residues implicated in substrate and ion binding in EAATs are conserved[28,32,90,94]. Thus, in the absence of any EAAT structures, one may use the Glt_{Ph} structure to interpret the functional data from EAATs and thereby gain insights for the transport mechanism in GlTs[112,113]. This can be achieved by performing detailed molecular dynamics (MD) simulations of Glt_{Ph}, where one characterizes the binding sites for the substrate and the ions and calculates their binding free energies in various configurations to determine the binding order. But before one can proceed with such a program, the binding site for the third Na^+ ion (to be called Na3) needs to be located accurately. Otherwise a reliable prediction of the binding free energies would not be possible.

So far several MD simulations of the Glt_{Ph} structure have been carried out focusing on the opening of the extracellular gate and substrate binding[36,99], location of the Na3 site[96,97], and substrate translocation and release[39,114]. The Na3 site was also searched via electrostatic calculations and valence mapping[31,95]. Of the predicted Na3 sites, only those involving D312[31,97] are completely consistent with the functional[30,31] and structural[15] data. Our initial free energy calculations indicated that even those Na3 sites did not yield binding energies that were consistent with the transport mechanism, which prompted us to further refine the Na3 site proposed in Ref.[97]. The proposed new site for Na3 involves, besides the already identified T92, N310 and D312 side chains[31,97], two other residues, Y89 backbone and S93 side chain. The presence of the S93 side chain in the coordination shell of the third Na^+ ion has been confirmed by S93A mutation experiments both in Glt_{Ph} and EAAT1, where a large reduction in Na^+ affinity has been observed[38].

It appears that all the binding sites in Glt_{Ph} have been well identified so the system is ripe for a detailed MD study. Our aim here is first to show that the MD

results for the binding sites and coordination of the substrate and ions are in good agreement with the experimental data. This is important to build confidence in the model and dispel any doubts about the ion binding sites in the crystal structure where Tl^+ ions were employed instead of Na^+. We then perform free energy calculations for the substrate and ions in various configurations to determine their binding free energies, which are used in predicting their binding order. In particular, a strong dependence of Asp binding to Na1 binding is found, which is explained by a H-bond network linking the two ligands. Finally we examine the Asp/Glu selectivity in Glt_{Ph} and discuss its implications for constructing homology models for EAATs, where such an Asp/Glu selectivity is not observed.

4.2 Model system and simulation details

Several crystal structures of the Glt_{Ph} protein are available from the Protein Data Bank. The most revealing structure is the closed structure (PDB ID: 2NWX), where an Asp and two Na^+ ions are bound and the helical hairpin 2 (HP2) gate is closed[15]. Following previous MD studies[36,99], we have created an open structure from 2NWX by removing the bound Asp and Na^+ ions and observing opening of the HP2 gate in MD simulations. In a second crystal structure (PDB ID: 2NWW), the transporter inhibitor DL-threo-β-Benzyloxyaspartate (TBOA) is bound in place of Asp, which keeps the HP2 gate open. We have created a similar apo structure of Glt_{Ph} from 2NWW by removing the bound TBOA and equilibrating the system in MD simulations. The two open structures are very similar in all respects except for the orientation of the N310 side chain, which is flipped 180° between the two structures. Because the N310 side chain is directly involved in the coordination of the Na3 site, it is important to choose the correct conformation for this side chain. The Na3 site predicted using the 2NWW structure has been shown to be more consistent with experiments[38]. Therefore we have adopted the conformation of the N310 side chain from the 2NWW structure in the 2NWX structure as well. With this change, the two open structures yield essentially identical results in free energy simulations.

The simulation systems are prepared using the software VMD[104]. The crystal structure of the Glt_{Ph} trimer is embedded in a 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) phospholipid bilayer and then solvated in a box of water molecules with 125 mM of NaCl. The system contains the trimer, 239 lipid molecules, 15688 water molecules and 35 NaCl ions. The Glt_{Ph} trimer has a net charge of $+6e$ in the apo state. Therefore, 6 extra Cl^- ions are included in the apo system to maintain charge neutrality. When the bound ligands are included in simulations, neutrality of the system is preserved by adding more counter ions. After the initial preparation, the system is equilibrated in two stages.

In the first stage, the coordinates of the protein atoms are fixed and the system is equilibrated with 1 atm pressure coupling in all directions until the correct water and

lipid densities are obtained. At this point, the x and y-dimensions of the simulation box are fixed, and pressure coupling is applied only in the z direction henceforth (typical dimensions of the simulation box are $115 \times 111 \times 71$ Å^3). In the second stage, the protein is gradually relaxed in 2.4 ns MD simulations by reducing the restraints on the protein atoms in several steps. The system is further equilibrated for 5 ns with only a small (0.1 kcal/mol/Å^2) restraint left on the backbone atoms of the protein. This helps to preserve the structural integrity of the protein during long MD simulations. The above procedure is repeated for all structures employed in MD simulations.

MD simulation are performed using the NAMD package (version 2.7b2)[59] with the CHARMM22 force field[55] including the CMAP dihedral corrections[105]. The temperature is maintained at 300 K using Langevin damping with a coefficient of 5 ps^{-1}, and the pressure is kept at 1 atm using the Langevin piston method with a damping coefficient of 20 ps^{-1}[64]. Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. The Lennard-Jones (LJ) interactions are switched off between 10 and 12 Å using a smooth switching function. A time step of 2 fs is used in all MD simulations.

4.3 Free Energy calculations

The standard binding free energy of a ligand can be expressed as[106,115,116]

$$\Delta G_{\mathrm{b}} = \Delta G_{\mathrm{tr}} + \Delta G_{\mathrm{rot}} + \Delta G_{\mathrm{con}} + \Delta G_{\mathrm{int}}, \tag{4.1}$$

where the first two terms represent the translational and rotational free energy differences of the ligand between the binding site and the bulk, the third term measures the free energy cost of applying conformational restraints on the ligand in the binding site relative to bulk, and the last term represents the free energy of translocating the restrained ligand from the bulk to the binding site. The translational and rotational free energy differences can be evaluated assuming Gaussian distributions for the ligand configurations in the binding site, which yields[107,117–119]

$$\Delta G_{\mathrm{tr}} = -k_{\mathrm{B}}T \ln \left[\frac{(2\pi e)^{3/2}\sigma_x\sigma_y\sigma_z}{V_0} \right], \Delta G_{\mathrm{rot}} = -k_{\mathrm{B}}T \ln \left[\frac{(2\pi e)^{3/2}\sigma_{\phi 1}\sigma_{\phi 2}\sigma_{\phi 3}}{8\pi^2} \right], \tag{4.2}$$

where $V_0 = 1660$ Å^3 is the reference volume for the standard concentration, σ_x, σ_y, and σ_z, are the principal rms fluctuations of the center of mass of the ligand in the binding site, and $\sigma_{\phi 1}$, $\sigma_{\phi 2}$, and $\sigma_{\phi 3}$, are the rotational rms fluctuations of the ligand calculated using the quaternion representation[119]. The various σ's are estimated from MD simulations of the bound ligands without using any restraints because they are stably bound.

The free energy difference for applying conformational restraints can be written as

$$\Delta G_{\mathrm{con}} = \Delta G_{\mathrm{con}}^{\mathrm{bulk}} - \Delta G_{\mathrm{con}}^{\mathrm{site}}, \tag{4.3}$$

where ΔG_{con}^{bulk} and ΔG_{con}^{site} are the free energies of restraining the ligand in the bulk and the binding site, respectively. To calculate ΔG_{con}^{site}, the equilibrium position of the ligand is determined as accurately as possible, and the heavy atoms on the ligand are restrained using $k_f = 4.5$ kcal/mol/Å^2. Following Cecchini et al.[120], these restraints are gradually relaxed and the free energy change is calculated from

$$\Delta G_{con}^{site} = \frac{1}{2}\int_0^{k_f} \langle |\mathbf{X} - \mathbf{X}_0|^2 \rangle \, dk, \tag{4.4}$$

where $|\mathbf{X} - \mathbf{X}_0|$ represents the difference of the restrained coordinates from the reference values at a given k. The integral is evaluated using 10 windows with k = 4.5, 2.0, 1.0, 0.5, 0.2, 0.1, 0.05, 0.02, 0.01 and 0 kcal/mol/Å^2. Each window is simulated for 600 ps, and the ensemble average of the coordinate difference is calculated from the last 300 ps. A similar calculation is performed in the bulk for ΔG_{con}^{bulk} using a box containing the ligand and 1193 water molecules (plus counterions to keep the system neutral).

The free energy of translocation, ΔG_{int}, is calculated using the free energy perturbation (FEP) and thermodynamic integration (TI) methods[52]. In both methods, one introduces a hybrid Hamiltonian $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, where H_0 represents the Hamiltonian for the initial (apo) state with the ligand in the bulk and H_1 for the final (ligand-bound) state. In the FEP method, the interval between $\lambda = 0$ and 1 is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average

$$\Delta G_i = -k_B T \ln\langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_B T] \rangle_{\lambda_i}. \tag{4.5}$$

The free energy of translocation is obtained from the sum, $\Delta G_{int} = \sum_i \Delta G_i$. A criterion for successful application of FEP is that $\Delta G_i \lesssim 2$ kcal/mol for each window. Because charges are involved in the FEP calculations, satisfaction of this criterion would require over 130 windows if uniform subintervals are used. We have used instead exponentially spaced subintervals, which halves the required number of windows to 66. In a typical FEP calculation, each window is equilibrated for 40 ps followed by a 40-ps production run. We have used shorter and longer simulation times to check the convergence of the results in sample calculations and found that 40 + 40 ps sampling is adequate.

In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, is obtained at several λ values, and the free energy of translocation is calculated from the integral

$$\Delta G_{int} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda. \tag{4.6}$$

Again because charges are involved, Gaussian quadrature provides a better method for the evaluation of this integral. In previous applications of the TI method to ion binding free energies[108], a seven-point quadrature was found to be adequate, which is also adapted here. Each window is typically equilibrated for 0.5 ns and sampled

for 1 ns. Convergence of the results are checked from the running averages of the free energy differences, and the sampling runs are extended if necessary.

For Na^+ ions, application of these methods is straightforward. It is observed that the Na1 and Na3 sites in the apo structure are occupied by water molecules within a few ns of MD simulations. A similar thing happens when the Na^+ ion is removed from the Na2 site in the closed structure. This justifies alchemical transformation of a Na^+ ion in one of the binding sites to a water molecule while a water molecule in the bulk is simultaneously transformed to a Na^+ ion. To check against the hysteresis effects, we also perform the backward calculation, where a Na^+ ion in the bulk is translocated to the binding site. Thus TI calculations for each Na^+ ion took 11.5 x 2 = 23 ns. Convergence of the TI results are checked from the running averages (see Fig. 4A.1 in the Appendix). Denoting the free energy differences obtained from the forward and backward calculations as $\Delta G_{\rm int}^{\rm for}$ and $\Delta G_{\rm int}^{\rm back}$, the free energy of translocation is obtained from their average as

$$\Delta G_{\rm int} = \frac{1}{2}[-\Delta G_{\rm int}^{\rm for} + \Delta G_{\rm int}^{\rm back}], \tag{4.7}$$

and the absolute binding free energy follows upon adding the translational free energy difference

$$\Delta G_{\rm b} = \Delta G_{\rm int} + \Delta G_{\rm tr}. \tag{4.8}$$

When Asp is removed from the open structure, the binding site is similarly flooded with water molecules in a few ps. However, a direct application of the FEP method runs into convergence issues for molecules, which necessitates adoption of a staged approach to free energy calculations[116]. Accordingly, Asp is restrained in the binding site and the electrostatic and LJ interactions are switched off in turn, which results in annihilation of Asp in the binding site. A reverse process is performed in the bulk simultaneously, resulting in creation of a restrained Asp in the bulk, which is then unrestrained. The reference positions for the restrained atoms are chosen so that there is a maximal overlap between the restrained and unrestrained Asp. This is demonstrated in Table 4.2, where the restrained Asp (last column) is seen to have an almost identical coordination geometry as the unrestrained Asp (fifth column).

To improve convergence and prevent instabilities in the FEP calculations, we use a soft-core LJ potential with a shift coefficient of 5.0[82]. This avoids the necessity of splitting the LJ interaction into short-range and dispersion parts. We have discovered that during the backward calculation—while an uncharged Asp is created in the binding site—a water molecule gets trapped under Asp, leading to a final state inconsistent with the crystal structure. To prevent this from happening, we have inserted Asp to the binding site in two stages, first the backbone and then the side chain. A similar splitting is employed in the forward calculation—the side chain of Asp is destroyed first, followed by the backbone.

The free energy difference at each stage is calculated separately, and the forward free energy of translocation is obtained from the sum

$$\Delta G_{\rm int}^{\rm for} = \Delta G_{\rm elec} + \Delta G_{\rm LJ-bb} + \Delta G_{\rm LJ-sc}, \tag{4.9}$$

where the three terms represent the contributions from the electrostatic interactions, LJ-backbone, and LJ-side chain. To calculate ΔG_{elec}, Asp in the binding site is discharged while an uncharged Asp in the bulk (30 Å away from the binding site) is charged simultaneously. We have used both FEP and TI methods for this purpose employing the window parameters specified above. In total four FEP and two TI simulations are performed. Evidence for convergence of the FEP results are presented in the Appendix (see Fig. 4A.2). In the second and third stages, the LJ interactions for the side chain and backbone of the uncharged Asp in the binding site are switched off while they are switched on for an uncharged Asp in the bulk. The FEP method is employed in these calculations with 52 windows for $\Delta G_{\text{LJ}-\text{sc}}$ and 44 windows for $\Delta G_{\text{LJ}-\text{bb}}$. In both cases, a shorter spacing ($\Delta\lambda = 0.01$) is used closer to the end points to facilitate convergence. For each of the forward calculations mentioned above, a complementary backward calculation is performed to check against hysteresis. Because of the larger relative fluctuations observed, ten independent FEP calculations are run for the LJ interactions.

We have also investigated the selectivity of Glt_{Ph} for Asp over Glu. Since all the residues that coordinate Asp in the open Glt_{Ph} are conserved in the EAATs (R276 is an exception but it coordinates with the backbone oxygen), it is assumed that Glu binds to the transporter in the same position as Asp. The free energy of mutating Asp to Glu in the binding site and bulk are determined via TI calculations as described above (except using 1 ns equilibration time for each window). The selectivity free energy is determined from the difference

$$\Delta\Delta G(\text{Asp} \rightarrow \text{Glu}) = \Delta G^{\text{site}}_{\text{Asp}\rightarrow\text{Glu}} - \Delta G^{\text{bulk}}_{\text{Asp}\rightarrow\text{Glu}}. \tag{4.10}$$

Note that no restraints are needed in this case as both Asp and Glu are bound to Glt_{Ph}.

We have further simulated the Glu-bound system for 10 ns starting from the last window of the Asp → Glu transformation. Differences in the coordination of Asp and Glu in the binding pocket are examined to find out which changes are required in EAATs in order to accommodate Glu as well as Asp and transport it with equal efficiency.

4.4 Results and discussion

4.4.1 Characterization of the binding sites

We start with a comparison of the binding sites for the three Na^+ ions and Asp obtained from the MD simulations of Glt_{Ph} with those observed in the crystal structure 2NWX. The residues involved in the coordination of the Na^+ ions are listed in Table 4.1 together with the average Na^+–O distances obtained from 2-ns MD simulations of the closed and open structures. The main difference between the two is that the HP2 gate is open in the latter which abolishes the Na2 site. From Table

4.1 it is seen that the Na3 site is tightly coordinated by one backbone and four side chain oxygens. Note that only one of the D312 oxygens is in the first hydration shell and there are no water molecules. A similarly tight binding site is found for the Na1 site. There is a one-to-one correspondence between the coordinating residues observed in the crystal structure and those found from the MD simulations. The larger ion-oxygen distances in the crystal structure is presumably due to the larger size of the Tl^+ ion employed in the experiments. Almost identical coordinations are obtained in the closed and open structures for the Na3 and Na1 sites, which indicates that the opening of the HP2 gate, while abolishing the Na2 site, has a minimal effect on the Na3 and Na1 sites.

Table 4.1: The Glt_{Ph} residues coordinating the three Na^+ ions, listed according to their binding order. The average Na^+–O distances obtained from 2-ns MD simulations of the closed and open structures are compared to those of the crystal structure (2NWX). Nature of the coordinating oxygens are indicated in parenthesis, i.e., bare O denotes backbone and the others are side chain. The Na3 site is not seen in the crystal structure. The open state has Asp and two Na^+ ions bound because the Na2 site is not formed. All distances are in Å.

Ion	Helix - Residue	Cry. str.	Closed st.	Open st.
Na3	TM3 - Y89 (O)		2.3 ± 0.1	2.3 ± 0.1
	TM3 - T92 (OH)		2.4 ± 0.1	2.4 ± 0.1
	TM3 - S93 (OH)		2.4 ± 0.1	2.3 ± 0.1
	TM7 - N310 (OD)		2.2 ± 0.1	2.2 ± 0.1
	TM7 - D312 (O_1)		2.1 ± 0.1	2.1 ± 0.1
	TM7 - D312 (O_2)		3.6 ± 0.2	3.5 ± 0.3
Na1	TM7 - G306 (O)	2.8	2.4 ± 0.2	2.4 ± 0.2
	TM7 - N310 (O)	2.7	2.3 ± 0.1	2.4 ± 0.2
	TM8 - N401 (O)	2.7	2.4 ± 0.2	2.5 ± 0.2
	TM8 - D405 (O_1)	3.0	2.2 ± 0.1	2.2 ± 0.1
	TM8 - D405 (O_2)	2.8	2.2 ± 0.1	2.3 ± 0.1
	H_2O	-	2.3 ± 0.1	2.3 ± 0.1
Na2	TM7 - T308 (O)	2.6	2.3 ± 0.1	
	TM7 - T308 (OH)	5.5	2.4 ± 0.1	
	HP2 - S349 (O)	2.1	4.5 ± 0.3	
	HP2 - I350 (O)	3.2	2.3 ± 0.1	
	HP2 - T352 (O)	2.2	2.3 ± 0.1	
	H_2O	-	2.3 ± 0.1	

Inspection of the Na2 coordination in Table 4.1 shows that there are some differences between the MD results and the crystal structure, which we again attribute

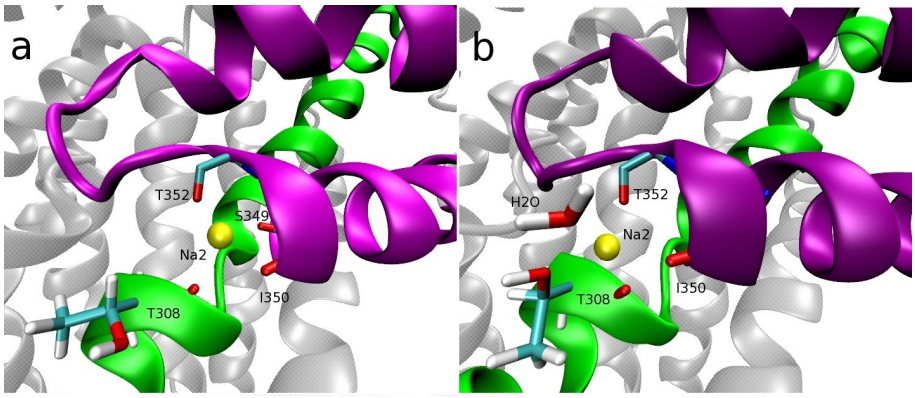

Figure 4.1: The coordination of Na2 (a) in the closed crystal structure of Glt_{Ph} and (b) in our MD simulations. See Table 4.1 for a list of the coordinating residues and the Na^+–O distances.

to differences between the Tl^+ and Na^+ ions. When a Na^+ ion is placed in the experimentally observed position (to be called Na2′), it immediately shifts 2-3 Å away from the coordinating S349 backbone oxygen. With further equilibration, the side chain of T308 enters the coordination shell of Na^+, forming a very stable binding site to be called Na2 (see Fig. 4.1). This is observed in all three monomers suggesting that the Na2 site is more likely to be the binding site for a Na^+ ion. The function of the Na2 site is to lock the HP2 gate in the closed conformation by binding to the residues in HP2 and TM7. In the crystal structure, this lock (the Na2′ site) has three hinges on HP2 and only one on TM7, whereas the Na2 site has two hinges on each, which provides a more reliable lock. More quantitative evidence for the proposed Na2 site will presented below from the binding free energies.

A similar comparison is presented for the Asp binding site in Table 4.2. Again there is a one-to-one correspondence between the coordinating residues seen in the crystal structure and those found from the MD simulation of the closed state. Most of the contact distances match very well except the side chain of S278 which is found to coordinate the α-carboxyl of Asp. This has been observed in other MD simulations of Glt_{Ph} as well[99,114]. Below we will show that S278 side chain plays a key role in binding of Asp following Na1. A water molecule coordinating Na1 also makes a H-bond with the S278 side chain. We speculate that absence of this water molecule in the crystal structure may have caused the shift in the position of the S278 side chain. Opening of the HP2 gate leads to the removal of the V355 and G359 residues from the coordination shell of Asp, which is partly compensated by the entry of the N_1 atom of the R397 side chain. The remaining eight contacts are preserved between the closed and open states and continue to provide a stable binding site for Asp.

Table 4.2: List of the Glt_{Ph} residues (first column) coordinating Asp atoms (second column). The average N–O and O–O distances obtained from the MD simulations are compared to those of the crystal structure 2NWX (third column). Again bare N and O in parenthesis refer to the backbone atoms and the others refer to the side chain atoms. The last column shows the effect of the conformational restraints on the coordination of Asp. All distances are in Å.

Helix - Residue	Asp	Cry. str.	Closed st.	Open st.	Open $st._{res}$
HP1 - R276 (O)	α-N	2.4	3.0 ± 0.2	3.0 ± 0.2	3.0 ± 0.2
HP1 - S278 (N)	α-O_1	2.8	2.8 ± 0.1	2.8 ± 0.1	2.8 ± 0.1
HP1 - S278 (OH)	α-O_2	3.8	2.7 ± 0.1	2.8 ± 0.2	2.8 ± 0.1
TM7 - T314 (OH)	β-O_2	2.7	2.7 ± 0.1	2.8 ± 0.1	2.8 ± 0.1
HP2 - V355 (O)	α-N	2.9	2.9 ± 0.2	11.9 ± 0.4	11.9 ± 0.3
HP2 - G359 (N)	β-O_2	2.8	3.1 ± 0.2	6.1 ± 0.4	6.3 ± 0.3
TM8 - D394 (O_1)	α-N	2.6	2.7 ± 0.1	2.7 ± 0.1	2.7 ± 0.1
TM8 - R397 (N_1)	β-O_2	4.6	4.2 ± 0.2	2.7 ± 0.1	2.7 ± 0.1
TM8 - R397 (N_2)	β-O_1	2.5	2.9 ± 0.2	2.9 ± 0.2	2.9 ± 0.2
TM8 - T398 (OH)	α-N	3.2	3.2 ± 0.2	3.0 ± 0.2	3.0 ± 0.2
TM8 - N401 (ND)	α-O_2	2.8	2.8 ± 0.1	3.0 ± 0.2	2.9 ± 0.2

4.4.2 Binding free energies

We first consider the binding free energies of the Na^+ ions to the sites Na1, Na2 and Na3 (Table 4.3). Test calculations with the FEP and TI methods have yielded essentially identical results, therefore only the TI method is used in the calculations for convenience. The forward and backward results are seen mostly to agree within 1 kcal/mol indicating minimal hysteresis effects. The free energy difference due to loss of translational entropy remains around 4.5–5 kcal/mol for all sites. The absolute binding free energies listed in the last column reveal that the Na3 site provides by far the largest affinity for Na^+ ions and hence it will be occupied first, followed by Na1 and Na2. It has been shown that the path to the Na3 site goes through the Na1 site[97]. The 7 kcal/mol energy gradient found between the Na3 and Na1 sites here provides sufficient incentive for a Na^+ ion bound to the Na1 site to move quickly to the empty Na3 site. As shown below, the Na1 site is occupied before Asp. We have therefore calculated the binding free energy for the Na1 ion in the presence of Na3. Coulomb repulsion due to the Na3 ion results in about 4 kcal/mol increase in the binding free energy for Na1 compared to the apo state. Nevertheless this is still sufficient to provide a μM affinity for Na1.

The Na2 site is occupied last after Na3, Na1 and Asp bind, and the HP2 gate is closed. The calculated binding free energy for Na2 is higher than all the other ligands, consistent with its binding last. Because Na2 binds last, its binding free

Table 4.3: Binding free energies for Na^+ ions in the open state (for Na3 and Na1) and closed state (for Na2) of Glt_{Ph} (in kcal/mol). Various contributions to the binding free energy as described in Eqs. 4.7 and 4.8 are listed separately. Other ligands present in the transporter are indicated in parenthesis. Errors are estimated from block data analysis using 100 ps windows.

Ion	$-\Delta G_{\text{int}}^{\text{for}}$	$\Delta G_{\text{int}}^{\text{back}}$	ΔG_{int}	ΔG_{tr}	ΔG_{b}
Na3	-22.3 ± 1.2	-24.2 ± 0.9	-23.3 ± 1.1	4.6	-18.7 ± 1.1
Na1	-15.9 ± 0.9	-16.5 ± 1.6	-16.2 ± 1.3	4.9	-11.3 ± 1.3
Na1 (3)	-11.8 ± 1.6	-11.9 ± 0.9	-11.9 ± 1.3	4.8	-7.1 ± 1.3
Na2 (1, 3, Asp)	-6.6 ± 1.3	-7.7 ± 1.3	-7.1 ± 1.3	4.4	-2.7 ± 1.3
Na2′ (1, 3, Asp)	-1.4 ± 0.9	-2.1 ± 0.8	-1.7 ± 0.8	4.4	$+2.7 \pm 0.8$

energy can be estimated from the observed Michaelis–Menten constant, $K_m = 3.9$ mM as -3.3 kcal/mol[35], which is in good agreement with the calculated value of -2.7 kcal/mol. As argued above, the Na2 binding site found from the MD simulations is slightly different from the Na2′ site observed in the crystal structure, and the former provides a better site for binding of a Na^+ ion. Indeed the binding free energy calculated for the Na2′ site is about 5 kcal/mol higher than that of Na2, and it is positive. This provides quantitative evidence that the Na2 site is more likely to be the actual binding site. We note that the T308 side chain—proposed to be involved in coordination of the Na2 site here—is highly conserved in the EAATs and may not have been resolved in the crystal structure due to the replacement of Na^+ by the larger Tl^+ ion. This hypothesis could be confirmed in future experiments by mutating T308 to a non-polar residue and observing how the Na^+ affinity and the Asp transport rate are affected by this mutation.

The results of the binding free energy calculations for Asp are presented in Table 4.4. As discussed below, Asp binds only in the presence of Na1 and Na3, therefore the calculations are performed only for this configuration. To ensure the reliability of the electrostatic calculations, we have determined it using both the FEP and TI methods. As seen from Table 4.4, the two methods yield consistent results. Electrostatic interactions dominate the binding of Asp resulting in a tightly bound ligand. This happens at the expense of the LJ interactions, which, therefore make a positive contribution (4.6 kcal/mol) to the binding energy. The translational and rotational free energy losses are even more significant, reducing the binding free energy by a further 7.2 kcal/mol. Comparable results have been obtained for the entropic free energy losses for similar-sized ligands[116]. Finally, the conformational restraints make a relatively small contribution to the binding free energy, reflecting the fact that Asp is slightly more restrained in the binding site compared to the bulk. We note that the measured binding free energy of Asp, -12 kcal/mol[15], refers to the closed state of Glt_{Ph}, which includes the free energies of gate-closure

Table 4.4: Binding free energy of aspartate in the open state of Glt_{Ph} with the Na1 and Na3 sites occupied (in kcal/mol). The forward and backward calculations differed less than 1 kcal/mol in all cases, therefore only their average is listed. The electrostatic contribution is taken as the average of the FEP and TI results given in the first two columns. The absolute binding free energy of Asp is given in the last column.

ΔG_{elec} (FEP)	ΔG_{elec} (TI)	ΔG_{LJ-bb}	ΔG_{LJ-sc}	ΔG_{tr}	ΔG_{rot}	ΔG_{con}	ΔG_{b}
-15.8 ± 0.6	-16.4 ± 0.9	3.8 ± 0.5	0.8 ± 0.5	3.3	3.9	0.5	-3.8 ± 1.0

and Na2 binding. The calculated free energy is for the binding of Asp to the open state of Glt_{Ph}. Therefore, a direct comparison with experiment is not possible at this stage. Binding free energy calculations that include the gating free energy have been performed for a glutamate receptor recently[41]. We are planning similar calculations for the HP2 gate in Glt_{Ph}, which will enable comparison with experiments.

4.4.3 Role of Na1 in facilitating Asp binding

We have also attempted to calculate the binding free energy of Asp in the presence of Na3 only, but failed to stabilize Asp in the binding site in the absence of Na1. Na1 is not directly involved in the coordination of Asp so the cause of this instability is not immediately clear. A careful study of this phenomenon has revealed a close relationship between the binding of Asp and the occupation of the Na1 site, namely, a network of hydrogen bonds that pin down Asp in the binding site is only stable in the presence of Na1. This H-bond network is indicated by a dashed line in Fig. 4.2, which shows the binding site of Asp obtained from the equilibrium simulations of the open structure with Asp, Na1 and Na3 bound. It is seen from Fig. 4.2 that the water molecule that coordinates the Na1 ion also makes a H-bond with the side-chain oxygen of S278, which, in turn, initiates the following chain of H-bonds: S278 (HG) – Asp (α-O_2) – N401 (HD_2); N401 (HD_1) – T314 (OG); T314 (HG) – Asp (β-O_2) – R397 (HN). The net effect of the H-bond network is to firmly anchor the backbone and side chain oxygens of Asp to the protein via four H-bonds. This network of H-bonds remains very stable throughout the simulations of the open structure as indicated by the average O–H distances (see Table 4A.1 in the Appendix).

To see how removal of the Na1 ion destabilizes and unzips the H-bond network, we have performed MD simulations of the system after removing the Na1 ion with only Na3 and Asp present. Initially 4 kcal/mol/Å^2 restraints are applied to all the C and N atoms of Asp, which are gradually relaxed during 2 ns of MD simulations. The system is then run for a further 3 ns with no restraints applied, during which time Asp is observed to unbind in all three monomers. The sequence of events leading to the collapse of the H-bond network and release of Asp are as follows (see Fig. 4A.3 in the Appendix for the snapshots).

1. After 1 ns of simulations when the restraints on Asp are 1 kcal/mol/Å^2

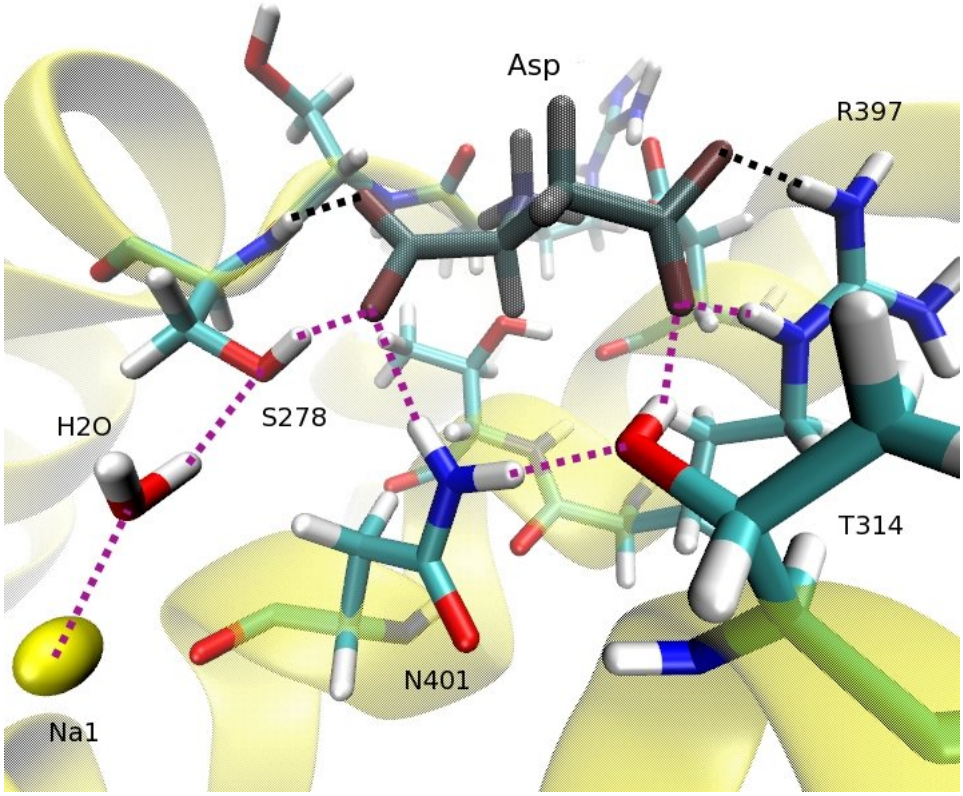

Figure 4.2: The binding site of Glt_{Ph} in the open state, with aspartate (dark-shaded) bound. The hydrogen bonds are denoted by dotted lines, with the purple ones showing the chain of H-bonds that stabilizes Asp in the binding site in the presence of Na1. The three contacts to the substrate α-amino group are not shown. This structure is stable throughout the simulations performed in the open state, as shown in Table 4A.1 in the Appendix.

(Fig. 4A.3a): a water molecule is near the Na1 site now with its hydrogens facing the Na1 site and its oxygen directed towards the S278 side chain. This causes the S278 side chain break the H-bond with Asp (α-O_2) and make a H-bond with the water oxygen instead.

2. At 2 ns with no restraints (Fig. 4A.3b): Sideway motion of S278 allows entry of a water molecule to the binding site. This water molecule breaks the Asp (α-O_2) – N401 (HD_2) H-bond, and starts making H-bonds with these two atoms and S278 (OG). Freeing of the N401 side chain disrupts the N401 (HD_1) – T314 (OG) H-bond, which, in turn, destabilizes the T314 side chain. Entry of another water molecule (not shown) helps to break the T314 (HG) – Asp (β-O_2) H-bond, with T314 side chain now making a H-bond with this water molecule. At this stage, α-O_2 of Asp is free, and mobility of Asp has increased considerably.

3. At 2.5 ns (Fig. 4A.3c): The two contacts from the β-carboxyl group of Asp to the R397 side chain are lost, leaving it completely free. Only four out of the nine contacts are still intact, namely, S278 (N) that coordinates α-O_1, and R276 (O), D394 (O_1) and T398 (OH) that coordinate α-N. More water molecules enter the binding site, further destabilizing Asp.

4. At 3.6 ns (Fig. 4A.3d): The remaining contacts are also lost. The binding site is now full of water molecules, and Asp is released to the solvent.

The above mechanism could play a role in the binding/unbinding of Asp. To appreciate its potential significance, we note that none of the Na^+ ions in Glt_{Ph} is in direct contact with Asp, which makes it difficult to understand the physical basis of the Asp–Na^+ coupling in glutamate transporters. For example, in LeuT transporters a Na^+ ion is in direct contact with the substrate, which makes their coupling obvious[121]. Here the Na1 ion is separated from Asp by a water molecule and a hydroxyl side chain but they are nevertheless coupled through the H-bond network depicted in Fig. 4.2. Removal of the Na1 ion triggers the unzipping of the H-bond network and simultaneous entry of water molecules to the binding site, which facilitate the unbinding of Asp. A likely scenario for binding of Asp is to follow this script in reverse, that is, Asp binds to Glt_{Ph} after Na1 following the steps 1–4 in the reverse direction. Experimental evidence for the key role played by the S278 side chain in the binding of Asp comes from cysteine scanning experiments in EAAT2[26]. Mutation of the residue corresponding to S278 in EAAT2 (S363) to cysteine results in 100-fold reduction in the transport activity. Due to the strong electronegative potential in the D405 region in the presence of the substrate and in the absence of Na1, it is also possible that this residue becomes protonated at some point during the transport cycle. Since there is no proton transport in Glt_{Ph}[35] and, as we shown in Chapter 7, the protonation of D405 is not favourable in the fully bound state, the presence of the proton could only be possible as a transient state in either the outward or inward conformation of Glt_{Ph}.

In a recent crystal structure of Glt_{Ph}, the transporter was captured in the inward-facing conformation. This structure has indicated that HP1 is most likely to be the intracellular gate, functioning in a similar fashion to the extracellular HP2 gate[16].

A similar conclusion has also been reached from modeling studies[39,122]. Inspection of Table 4.2 shows that three residues in HP1 coordinate Asp, including the S278 side chain that plays a key role in the H-bond network. Thus opening of HP1 will not only remove these three contacts from Asp but will also lead to the unzipping of the H-bond network, further destabilizing Asp. In the light of the above discussions, we suggest that Asp should unbind quite fast and may even be the first ligand to vacate the binding pocket.

4.4.4 Aspartate/Glutamate selectivity

The bacterial transporter Glt_{Ph} does not transport Glu, which is the main substrate for the mammalian GlTs (EAATs). In order to build a homology model of EAATs based on Glt_{Ph}, one needs to understand why Glu is not transported by Glt_{Ph}, and find the differences in structures that enable EAATs to transport both substrates. Fig. 4.3 shows the results obtained from 10 ns MD simulations with Glu bound to the open and closed states of Glt_{Ph}. In the open state, the three contacts that coordinate the Asp side chain, two from R397 and one from T314, do not coordinate Glu, even though Glu stays bound to the transporter. The longer side chain of Glu does not quite fit in the binding pocket, resulting in the loss of these contacts. Mutagenesis experiments on the R447 residue in EAAT3 (equivalent to R397 in Glt_{Ph}) have shown that interaction of this residue with the side-chain carboxyl of Glu is of critical importance for transport[32], which is prevented in the open state of Glt_{Ph} due to space constraints. Closing of the HP2 gate (Fig. 4.3b) induces a conformational change in the Glu side chain such that the contacts with the R397 and T314 side chains are restored.

When this conformation of Glu is placed in the open state and equilibrated, it goes back to the original open-state conformation (Fig. 4.3a) within few ps of releasing the restraints. This indicates that Glu is in a higher energy conformation in the closed state. However, even in this conformation, the Glu side chain does not fit in the binding pocket as well as Asp, resulting in displacement of the HP2 gate by about 1.5 Å relative to the Asp-bound case (Fig. 4.3b). These observations suggest that the binding pocket in EAATs must be larger or more flexible to accommodate Glu. Sequence alignment indicates that majority of the residues forming the binding pocket are conserved, and the only residue that may have an effect on the volume is R276 in Glt_{Ph}, which is replaced by A or S in EAATs. The R276 side chain is located between HP1 and TM8 and its replacement by a smaller side chain may help to increase the volume of the binding pocket.

For a quantitative assessment of the perturbations caused by the binding of Glu, we have calculated the free-energy required to mutate Asp to Glu in the open and closed structures of Glt_{Ph} (listed in Table 4.5). In the open state, the binding free energy difference between Asp and Glu is $\Delta G^{op}_{Asp\rightarrow Glu} = 5.2$ kcal/mol. Using the binding free energy of Asp (−3.8 kcal/mol), we can estimate the binding free energy of Glu as +1.4 kcal/mol. However, for the reasons stated above for Asp, this result

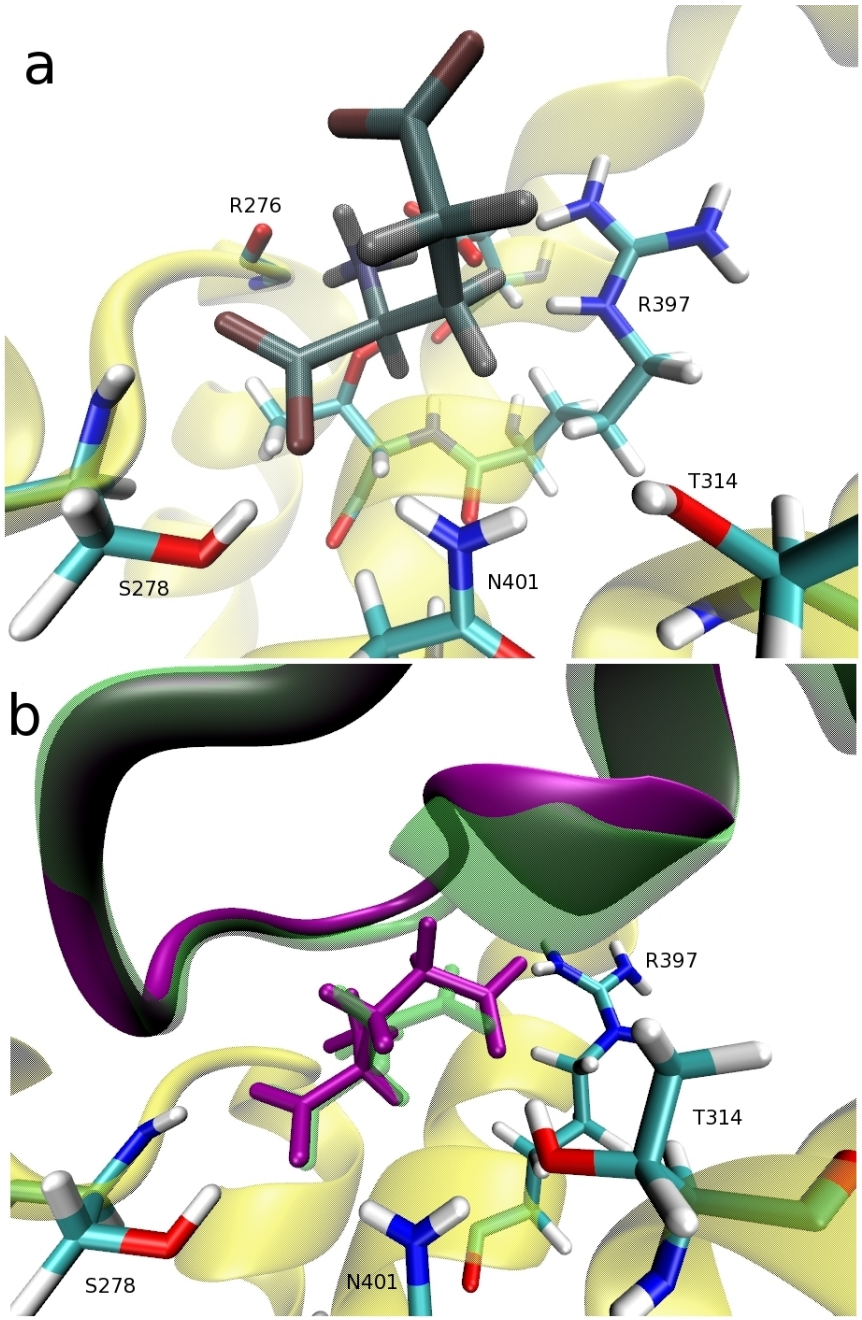

Figure 4.3: The binding site of Glt_{Ph} with a bound Glu in the open and closed states. (a) Open state: residues R397 and T314, which coordinate the β-carboxyl group of Asp, do not coordinate Glu (dark shade), resulting in a lower binding affinity for Glu. (b) Closed state: The side chain of Glu (dark shade) sticks out further relative to Asp (transparent light shade), resulting in displacement of the HP2 residues (dark shade) with respect to the Asp-bound case (transparent light shade).

Table 4.5: Asp → Glu selectivity in the binding site for the open state of Glt_{Ph} with Na1 and Na3 bound, and the closed state with all the ions bound to the transporter. The second column gives the free energy of mutating Asp to Glu in the binding site, minus the free energy of doing the same in the bulk. The third column gives the same for the backward transition. The average of the two is listed in the fourth column. The next two columns give the translational and rotational free energy differences between Asp and Glu, and the last column gives the total selectivity free energy. Errors are estimated from block data analysis using 100 ps windows.

Glt_{Ph} state	ΔG_{int}^{for}	$-\Delta G_{int}^{back}$	ΔG_{int}^{av}	ΔG_{tr}	ΔG_{rot}	$\Delta G_{Asp\to Glu}$
open	5.3 ± 1.1	5.3 ± 1.6	5.3 ± 1.3	0.0	-0.1	5.2 ± 1.3
closed	6.3 ± 1.4	5.7 ± 1.7	6.0 ± 1.5	-0.1	-0.5	5.4 ± 1.5

cannot be directly compared to the experimental value for Glu binding. In the closed state, the binding free energy difference is given by $\Delta G_{Asp\to Glu}^{cl} = 5.4$ kcal/mol. The experimental value obtained from the ratio of the dissociation constants of Asp and Glu is 6.6 kcal/mol[15], which is close to the calculated binding free energy difference.

Since we know the free-energy of mutating Asp to Glu in the closed and open states, we can estimate the difference in free energy of gating between the Asp and Glu-bound configurations. Using the thermodynamic cycle, this quantity can be written as

$$\Delta G_{Glu}^{op\to cl} - \Delta G_{Asp}^{op\to cl} = \Delta G_{Asp\to Glu}^{cl} - \Delta G_{Asp\to Glu}^{op} \tag{4.11}$$

From the values given in Table 4.5, it is seen that the gating energy remains about the same when Glu is bound compared to the Asp-bound case. It will be interesting to calculate the individual gating energies for the Asp-bound and Glu-bound cases from umbrella sampling MD simulations[41] and compare the substrate binding free energies directly to the experimental results.

4.5 Conclusions

We have performed free energy simulations for the Asp transporter Glt_{Ph}, a homolog of the mammalian glutamate transporters (EAATs). We have investigated in detail the coordination of all the ligands and calculated their binding free energies. The coordination shells found for Asp and Na1 in the closed state are in good agreement with the crystal structure. A small displacement of the experimental Na2 site is observed, which results in replacement of the S349 carbonyl in the coordination shell with the T308 hydroxyl. The new Na2 site is shown to be more stable and also more consistent with the gate-locking function of Na2.

Our simulations show that Asp is firmly bound to the open structure in the presence of Na3 and Na1 but becomes unstable when Na1 is removed, unbinding in

less that 5 ns in all three monomers. A close scrutiny of the unbinding process has revealed that Na1 and Asp are coupled via a hydrogen-bond network, and removal of Na1 leads to the unzipping of this network destabilizing Asp. The binding sequence of the ligands is determined from their binding free energies as: Na3, Na1, Asp (HP2 gate closes), Na2. The binding order of the second Na^+ ion and Asp has not been resolved from experimental observations and previous MD simulations[97]. Our results clearly show that Asp binding is contingent upon the presence of Na1 and thus happens after the binding of the second Na^+ ion. Finally our MD simulations with bound Glu and FEP calculations for transforming Asp $\rightarrow$ Glu in the open and closed states provide a simple mechanistic explanation for why Glt_{Ph} transports Asp but not Glu—the binding pocket is too small to accommodate the larger Glu, which remains in a higher energy state compared to Asp both in the open and closed states. The insights provided for Asp/Glu selectivity here will provide valuable guidance in constructing homology models for EAATs based on the Glt_{Ph} structure.

4.6 Acknowledgments

This work was supported by grants from the Australian Research Council and Turkish Scientific and Technical Research Council. Calculations were performed using the HPC facilities at the National Computational Infrastructure (Canberra). We thank Rob Vandenberg and Renae Ryan for discussions on the structure and function of Glt_{Ph}.

4.A Appendix for Chapter 4

Table 4A.1: List of the residues involved in the H-bond network that stabilizes Asp in the binding site when the Na1 site is occupied, and the average distances between the contact atoms (indicated in parentheses). Oxygens are defined as in Tables 4.1 and 4.2. HG refers to the hydroxyl hydrogen, HD_1 and HD_2 refer to the two hydrogens of the β-amide group of N401, and HN refers to the closest hydrogen in the guanidium group of R397.

Residue	Residue	Distance
Na1	H_2O (O)	2.3 ± 0.1
H_2O (H)	S278 (OG)	2.1 ± 0.3
S278 (HG)	Asp (α-O_2)	1.9 ± 0.2
Asp (α-O_2)	N401 (HD_2)	2.1 ± 0.3
N401 (HD_1)	T314 (OG)	2.0 ± 0.1
T314 (HG)	Asp (β-O_2)	1.8 ± 0.2
Asp (β-O_2)	R397 (HN)	1.8 ± 0.2

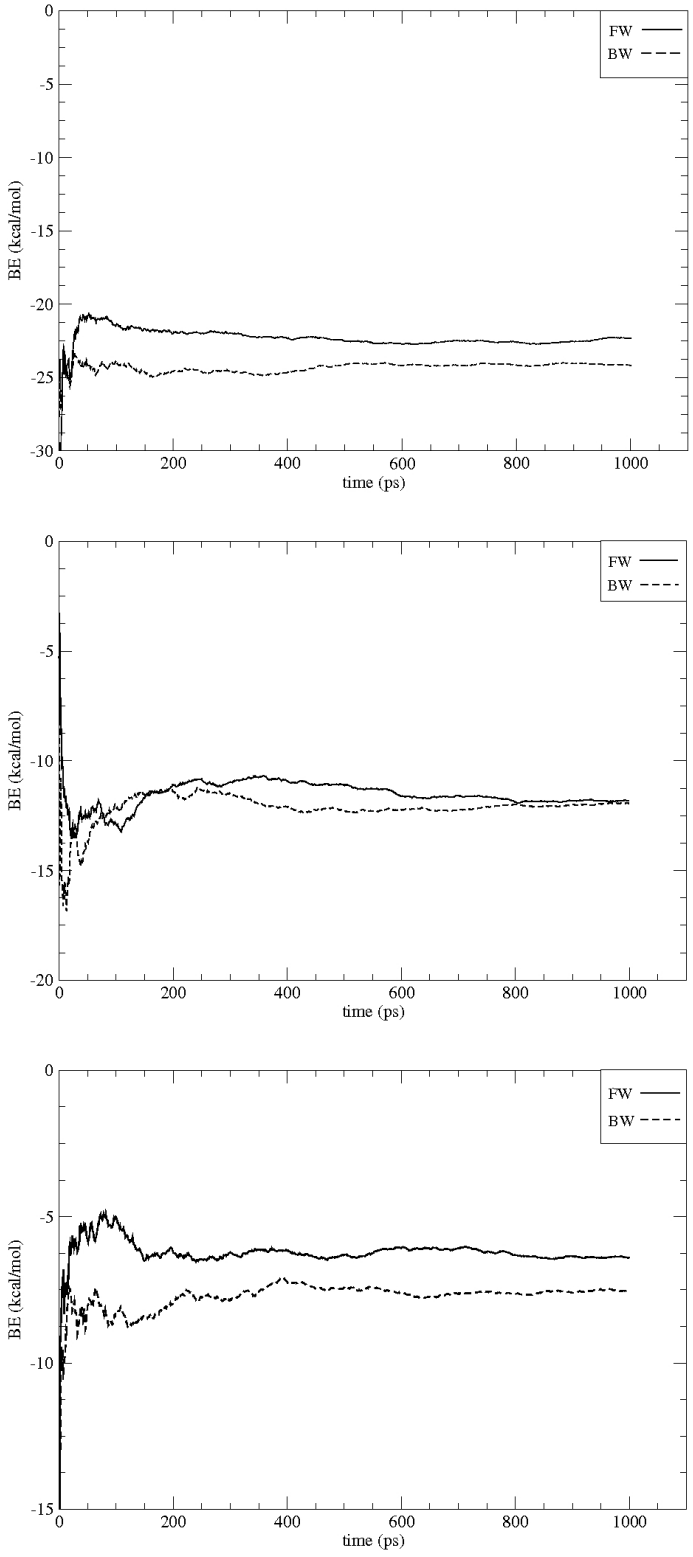

Figure 4A.1: Convergence of the TI results for the translocation free energy of the Na3 (top) and Na1 (middle) and Na2 (bottom) ions from running averages. The Na3 ion is present in the Na1 calculation, and the Na2 calculation is for the closed state with all the ligands present. The negative of the forward (binding site $\rightarrow$ bulk) and the backward transitions of Na^+ are shown with solid and dashed lines, respectively.

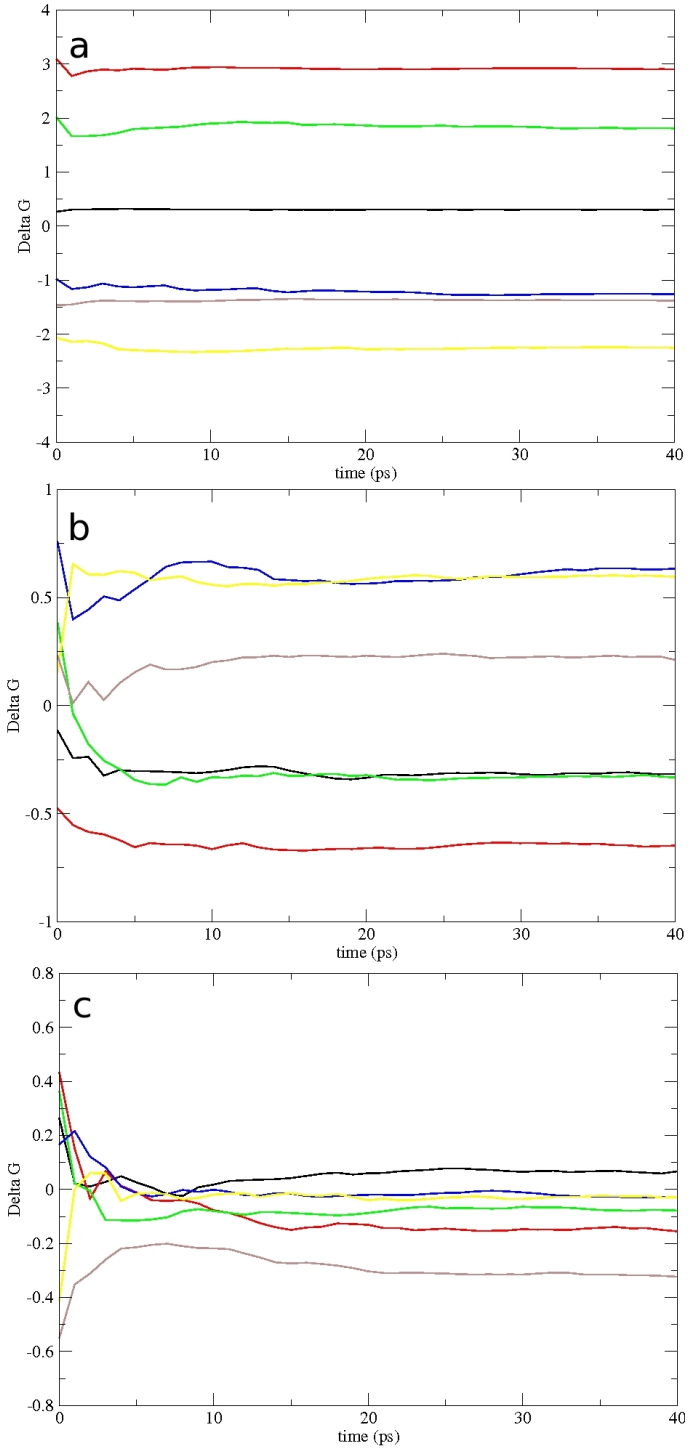

Figure 4A.2: Convergence of the FEP simulations for the binding free energy of Asp. a) The variation on the free energy during the FEP simulations of the discharging of aspartate inside the binding pocket and charging it in the bulk. The graph shows the free energies for the windows located at $\lambda = 0.0$ (black), $\lambda = 0.2$ (red), $\lambda = 0.4$ (green), $\lambda = 0.6$ (blue), $\lambda = 0.8$ (yellow) and $\lambda = 0.99$ (purple). b, c) The same for the destruction/creation in the binding pocket/bulk of the side chain (b) and backbone (c) of the discharged substrate.

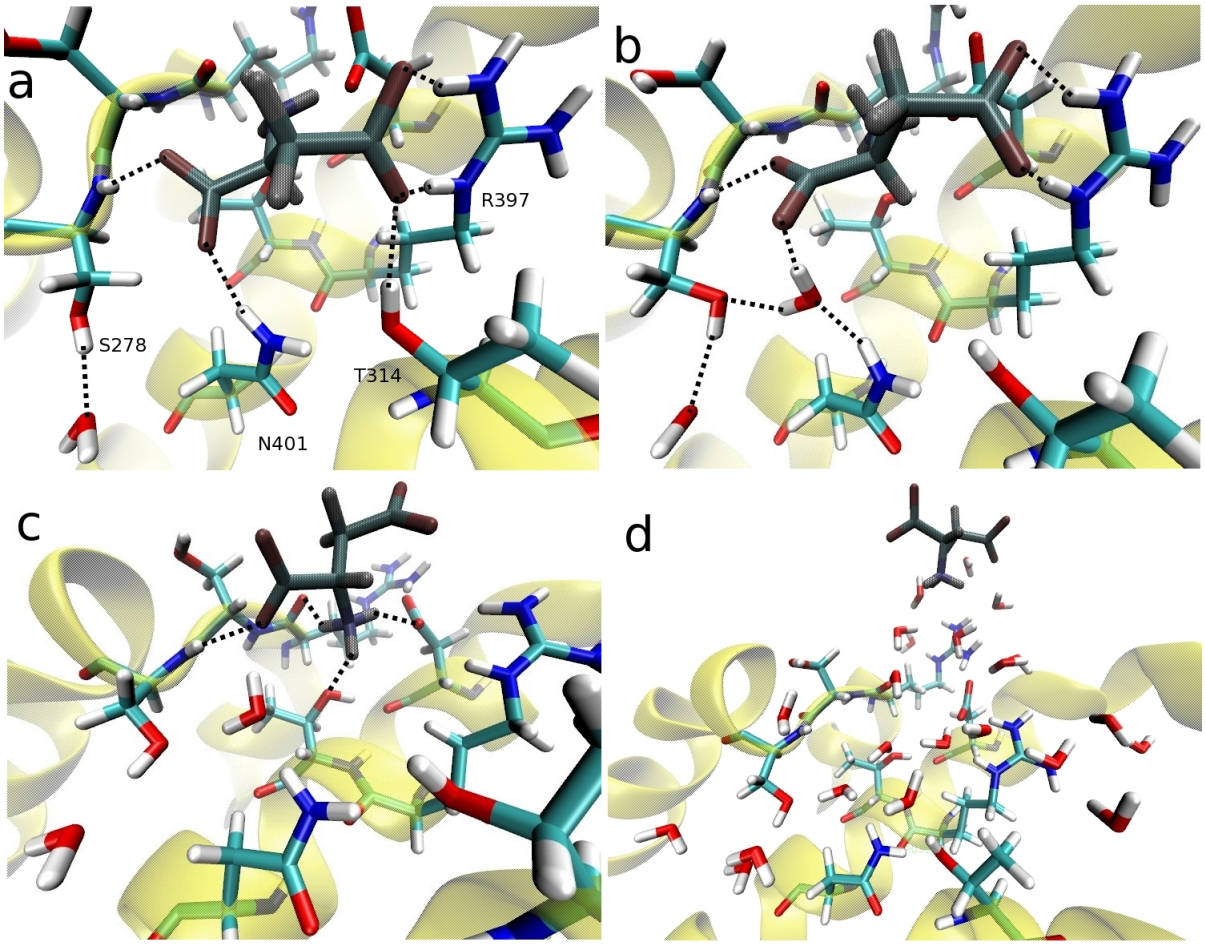

Figure 4A.3: The sequence of events that leads to the unbinding of Asp after the removal of the Na1 ion are summarized in the the snapshots a-d (see the main chapter for further explanations). The H-bonds forming the network are indicated by dotted lines.

Chapter 5

Mechanism and Energetics of Ligand Release in the Aspartate Transporter $\mathrm{Glt_{Ph}}$

ABSTRACT

The bacterial aspartate transporter $\mathrm{Glt_{Ph}}$ cotransports three $\mathrm{Na^+}$ ions with the substrate. Mechanism and energetics of ligand binding have previously been studied using molecular dynamics simulations on the crystal structure of $\mathrm{Glt_{Ph}}$ captured in the outward-facing state. Here we use the recent crystal structure of the inward-facing state of $\mathrm{Glt_{Ph}}$ to study the reverse process of unbinding of ligands. Gating behavior is studied in the presence of different ligands. A detailed characterization of the intracellular gate is given, pointing out the differences from the extracellular gate. We then perform free energy simulations to calculate the binding affinities of all the ligands in different combinations, from which the unbinding order is determined as Na2, (gate opens), Asp, Na1 and Na3. The strong coupling between Asp and Na1 is quantified from several free energy calculations. Na3 has the largest affinity to $\mathrm{Glt_{Ph}}$, and therefore its unbinding is proposed as the rate-limiting step in the transport cycle. The release time of Na3, estimated from Kramers' rate theory, is shown to be consistent with the experimental turnover rate of the transporter.

5.1 Introduction

Mammalian glutamate transporters, known as excitatory amino acid transporters (EAATs), are membrane proteins responsible for clearing the excess glutamate released at the nerve synapses. Malfunctioning of glutamate transporters has been implicated in many pathological conditions including cerebral ischemia, amyotrophic lateral sclerosis and Alzheimer's disease.[12] The transport mechanism of EAATs involves the cotransport of three Na^+ and one H^+ ions with the substrate, and the countertransport of one K^+ ion.[13,85] Glutamate transporters function by cycling between two states: one facing the extracellular (EC) medium where the ligands bind to the transporter, and the other facing the intracellular (IC) medium where the ligands are released to the cytoplasm. The first crystal structure of an archaeal homolog of EAATs—the aspartate transporter Glt_{Ph} from *Pyrococcus horikoshii*—was determined in the outward-facing state.[14] In subsequent crystal structures of this state, Glt_{Ph} was captured both in the open and closed conformations, and the binding sites for the substrate and two ions were identified.[15] Recently, the crystal structure of Glt_{Ph} was reported in the inward-facing closed state, also complexed with the substrate and two ions.[16] So far, there is no crystal structure available for the inward-facing open state of Glt_{Ph}.

Glt_{Ph} shares about 37% sequence identity with the EAATs, but the homology is much larger for the residues involved in the binding of the ligands, with almost of all of them being conserved.[28,32,90,94] Experiments have shown that, while the transport mechanism in Glt_{Ph} is independent of H^+ and K^+ ions,[35] it is still coupled to the cotransport of three Na^+ ions as in EAATs.[37] The third Na^+ site could not be identified in the crystal structures of Glt_{Ph}, but a binding site determined from molecular dynamics (MD) simulations was confirmed from mutagenesis experiments and shown to be conserved in Glt_{Ph} and the EAAT family.[38] Due to the structural and functional similarities between these proteins, and in the absence of any crystal structures for the EAATs, Glt_{Ph} provides a good model to interpret functional data on EAATs and gain new insights into the transport mechanism.[112,113]

Several MD simulations have been performed on the outward-facing state of Glt_{Ph}, to study the opening of the HP2 segment which works as the EC-gate,[36,99] to find the location of the Na3 site,[38,96,97] and to calculate the binding free energies of the ligands.[96,123] MD simulations and elastic network models have also been used to study substrate translocation,[114,122,124,125] and release.[39] In an interesting application of metadynamics simulations, Grazioso et al.[126] investigated the mechanism of substrate uptake and release, and calculated the free energy surfaces for opening of the EC and IC-gates with different combinations of bound ligands. This study revealed that, i) contrary to the outward-facing state where only the HP2 segment works as the EC-gate, opening of the IC-gate involves the movement of both HP1 and HP2, and ii) there is an energy barrier between the open and closed conformations of the IC-gate, which is not seen in the case of the EC-gate. We note that the metadynamics simulations were carried out using a single monomer in solution

instead of a trimer embedded in membrane and in the absence of the Na3 ion. It is important to confirm that the gating results obtained in Ref.[126] are retained in a more realistic environment and in the presence of the Na3 ion.

In a previous paper,[123] we performed detailed free energy simulations in the outward-facing state of Glt_{Ph} with an Asp and three Na^+ ions as ligands. From the binding free energies of ligands in different configurations, we determined their binding order as, Na3, Na1, Asp, (EC-gate closes), and Na2. A key finding of this work is that Asp and Na1 are coupled through a hydrogen-bond network despite being separated by about 7 Å. Here we perform a similar study for the inward-facing state of Glt_{Ph}. We first investigate the opening of the IC-gate by performing unbiased MD simulations with different combinations of the bound ligands. This is followed by free energy calculations of the binding affinities for Na2 in the closed configuration, and for other ligands in the open configuration. As in the outward-open case, we find a close link between binding of Asp and Na1, namely, Asp can only bind in the presence of Na1 and it is released to the cytoplasm when Na1 is removed from the binding site. The binding free energy results indicate that the ligands are released in a reverse order compared to their binding, that is, Na2, (IC-gate opens), Asp, Na1, and Na3. The Na3 ion is released last with a binding affinity much higher than the other ligands, which suggests that this step is the rate-limiting step of the transport cycle.

5.2 Model system and simulation details

. There are several crystal structures of Glt_{Ph} available from the Protein Data Bank, most of them in the outward-facing state.[14,15] In this study we use the crystal structure of Glt_{Ph} K55C–A364C mutant crosslinked with divalent mercury, which is in the inward-facing, closed conformation with two Na^+ ions (Na1 and Na2) and Asp bound (PDB ID: 3KBC).[16] We choose this structure here because our goal is to study the mechanism and energetics of release of the ligands to the cytoplasm. As in the other crystal structures of Glt_{Ph}, the Na3 site was not resolved in 3KBC. Therefore, we add a third Na^+ ion at the Na3 site identified in Ref.[38], where it is coordinated by the side chains of D312, N310, T92 and S93, and the backbone carbonyl of Y89. Justification for this site comes from mutagenesis experiments,[38] which show that the T92A and S93A mutations significantly reduce Na^+ affinity to the transporter. Also when an alternative Na3 site is used in MD simulations, the D312 side chain swings away from the crystal structure position by about 6 Å and starts coordinating Na1, which is in conflict with the Na1 coordination shell found in the crystal structure.[38]

The simulation system is prepared using the software VMD.[104] We first embed the crystal structure of the Glt_{Ph} trimer in a 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) phospholipid bilayer containing 251 POPE molecules. Each subunit of Glt_{Ph} has four ligands bound: three Na^+ ions, namely Na1, Na2 and Na3, and Asp

as the substrate. We then solvate the protein-membrane complex in a box of 16204 water molecules with 34 Na^+ ions and 43 Cl^- ions. The extra Cl^- ions are required to keep the simulation box neutral, and their number changes depending on the ligands bound to the transporter. The system is then equilibrated in two stages. In the first stage, the coordinates of the ligands and protein atoms are fixed and the system is equilibrated with 1 atm pressure coupling in all directions until the correct water and lipid densities are obtained. At this point, the x and y-dimensions of the simulation box are fixed, and pressure coupling is applied only in the z direction thereafter (typical dimensions of the simulation box are $114 \times 112 \times 75$ Å^3). In the second stage, the protein is gradually relaxed by reducing the restraints on the protein atoms in several steps during MD simulation lasting 2.4 ns. The system is then further equilibrated for 10 ns without restraints before we start collecting data.

All of our MD simulations are performed using the NAMD package (version 2.8)[59] with the CHARMM22 force field[55] including the CMAP dihedral corrections.[105] We note that the latest version (CHARMM36) has an updated force field for lipids but not for proteins. We have checked that using CHARMM36 instead of CHARMM22 yields essentially the same results for the binding free energies of ligands. The temperature is maintained at 300 K using Langevin damping with a coefficient of 5 ps^{-1}, and the pressure is kept at 1 atm using the Langevin piston method with a damping coefficient of 20 ps^{-1}.[64] Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. The Lennard-Jones (LJ) interactions are switched off between 10 and 12 Å using a smooth switching function. A time step of 2 fs is used in all MD simulations.

5.3 Free energy calculations

. We use essentially the same methods as in our previous work[123] in calculation of the binding free energies of ligands. Therefore, we briefly review the free energy methods here and refer to Ref.[123] for details. The standard binding free energy of a ligand can be expressed as[106,115,116]

$$\Delta G_{\mathrm{b}} = \Delta G_{\mathrm{tr}} + \Delta G_{\mathrm{rot}} + \Delta G_{\mathrm{con}} + \Delta G_{\mathrm{int}}. \tag{5.1}$$

Here the first two terms give the free energy change due to translational and rotational entropy loss upon binding, the third term measures the free energy cost of applying conformational restraints on the ligand in the binding site relative to bulk, and the last term represents the free energy of translocating the restrained ligand from the bulk to the binding site. The translational and rotational free energy differences are determined using[107,117–119]

$$\Delta G_{\mathrm{tr}} = -k_{\mathrm{B}}T \ln\left[\frac{(2\pi e)^{3/2}\sigma_x\sigma_y\sigma_z}{V_0}\right],\ \Delta G_{\mathrm{rot}} = -k_{\mathrm{B}}T \ln\left[\frac{(2\pi e)^{3/2}\sigma_{\phi 1}\sigma_{\phi 2}\sigma_{\phi 3}}{8\pi^2}\right], \tag{5.2}$$

where $V_0 = 1660$ Å^3, σ_x, σ_y, and σ_z, are the principle rms fluctuations of the center of mass of the ligand in the binding site, and $\sigma_{\phi 1}$, $\sigma_{\phi 2}$, and $\sigma_{\phi 3}$, are the rotational rms fluctuations of the ligand calculated using the quaternion representation.[119] The various σ values are estimated from 8 ns of MD simulations of the bound ligands with no restraints applied.

The free energy cost of applying conformational restraints is determined using[120]

$$\Delta G_{\text{con}} = \frac{1}{2} \int_0^{k_f} \langle |\mathbf{X} - \mathbf{X}_0|^2 \rangle \, dk, \tag{5.3}$$

where $|\mathbf{X} - \mathbf{X}_0|$ represents the difference of the restrained coordinates from the reference values at a given k. The integral is evaluated using 10 windows with k = 4.0, 2.0, 1.0, 0.5, 0.2, 0.1, 0.05, 0.02, 0.01 and 0 kcal/mol/Å^2. Each window is simulated for 600 ps, and the ensemble average of the coordinate difference is calculated from the last 300 ps. For the bulk simulations, we use a box containing the ligand and 1193 water molecules (plus counterions to keep the system neutral). The value of ΔG_{con} is then obtained from $\Delta G_{\text{con}} = \Delta G_{\text{con}}^{\text{bulk}} - \Delta G_{\text{con}}^{\text{site}}$.

The free energy of translocation, ΔG_{int}, is calculated using the free energy perturbation (FEP) and thermodynamic integration (TI) methods,[52] In the FEP method, the interval between $\lambda = 0$ and 1 is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average

$$\Delta G_i = -k_{\text{B}} T \ln \langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_{\text{B}}T] \rangle_{\lambda_i}, \tag{5.4}$$

where $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, with H_0 and H_1 representing the free-ligand and bound-ligand states. The free energy of translocation is obtained from the sum $\Delta G_{\text{int}} = \sum_i \Delta G_i$. For electrostatic interactions, we use exponentially spaced subintervals which are smaller near the end points, with a total of 66 windows.[123] In a typical FEP calculation, each window is equilibrated for 40 ps followed by 40 ps of production. In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, is obtained at several λ values, and the free energy of translocation is calculated from the integral

$$\Delta G_{\text{int}} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda. \tag{5.5}$$

Provided the integrand can be fitted well with a polynomial, Gaussian quadrature provides an efficient and accurate method for evaluation of such integrals because it allows for longer sampling of a smaller number of windows. This is the case for all of our calculations, therefore we use a seven-point quadrature for the TI simulations that involve electrostatic interactions.[108] Seven windows are found to be insufficient for the LJ interactions so a nine or twelve-point quadrature has been used in this case. All the TI windows are created from the previous FEP runs, which have already been equilibrated. Each window is sampled for 0.8 ns, of which the last 0.7 ns is used for production. Adequacy of this protocol is demonstrated from the

running averages of the free energy differences which converge to a flat line, and from lack of hysteresis effects between the forward and backward calculations. In addition, the free energies calculated using the FEP and TI methods agree within statistical fluctuations. To prevent duplication, we report here only the TI results, which are expected to be slightly more accurate due to longer sampling of each window.

For the Na^+ ions, only two of the terms in Eq. 5.1 contribute to the binding free energy, $\Delta G_{\rm tr}$ and $\Delta G_{\rm int}$. All the Na sites are observed to be occupied by a water molecule in the absence of the ion. Thus $\Delta G_{\rm int}$ is calculated by alchemically transforming a Na^+ ion in one of the binding sites to a water molecule while a water molecule in the bulk is simultaneously transformed to a Na^+ ion. The backward calculation for the reverse process is also performed, and $\Delta G_{\rm int}$ is determined from their average as, $\Delta G_{\rm int} = (-\Delta G_{\rm int}^{\rm for} + \Delta G_{\rm int}^{\rm back})/2$.

Due to its more complex nature, we adopt a staged approach for free energy calculations of Asp.[116] Accordingly, Asp is restrained in the binding site and the electrostatic and LJ interactions are switched off one at a time, which results in annihilation of Asp in the binding site. A reverse process is performed in the bulk simultaneously, resulting in creation of a restrained Asp in the bulk, which is then unrestrained. The reference positions for the restrained atoms are chosen so that there is a maximal overlap between the restrained and unrestrained Asp. The LJ interactions are switched off using a soft-core LJ potential with a shift coefficient of 5.0.[82] To avoid trapping of water molecules under the substrate during the backward LJ calculation, it is performed in two stages: first the neutral backbone is created/destroyed and then the neutral side chain is created/destroyed. This approach is also used in the forward LJ calculation to preserve symmetry. Thus the forward free energy of translocation is given by, $\Delta G_{\rm int}^{\rm for} = \Delta G_{\rm elec} + \Delta G_{\rm LJ-bb} + \Delta G_{\rm LJ-sc}$, where the three terms represent the contributions from the electrostatic interactions, LJ-backbone, and LJ-side chain. $\Delta G_{\rm elec}$ is calculated by discharging Asp in the binding site while an uncharged Asp in the bulk (30 Å away from the binding site) is charged simultaneously. In the second and third stages, the LJ interactions for the side chain and backbone of the uncharged Asp in the binding site are switched off while they are switched on for an uncharged Asp in the bulk (see Ref.[123] for details).

5.4 Kramers' rate theory

. The binding free energies of ions, especially that of Na3, are quite large which may delay their release to bulk considerably. Therefore, it is important to show that the estimated release times are consistent with the observed turnover rates in $Glt_{\rm Ph}$. For this purpose, we use Kramers' rate theory,[127,128] which provides an estimate for the rate of escape of a particle from a binding pocket over a barrier described by a

1-dimensional external potential energy function $U(x)$

$$k = \frac{\omega_a}{2\pi}\left\{\left[1+\left(\frac{\zeta}{2\omega_b}\right)^2\right]^{1/2} - \frac{\zeta}{2\omega_b}\right\}\exp(-U_b/k_BT). \tag{5.6}$$

Here U_b is the barrier height, ω_a and ω_b are the frequencies of the harmonic potentials at the bottom of the well and top of the barrier, and ζ is the friction coefficient which is related to the diffusion coefficient via the Einstein relation, $\zeta = k_BT/mD$. We assume an inverse harmonic potential for the barrier with height U_b and range x_b. Because $U(x)$ is an external potential energy, U_b is taken from the enthalpic contribution to the binding free energy of the ion. The frequency, ω_b, can be determined from $U_b = m\omega_b^2 x_b^2/2$. The frequency in the binding pocket follows from the equipartition theorem, $m\omega_a^2\sigma_x^2 = k_BT$, where σ_x is the RMS fluctuations of the particle along the reaction coordinate which is calculated from the MD simulations. We note that, in a low friction environment ($\zeta \ll \omega_b$), the rate coefficient has the maximum value of

$$k_{\max} = (\omega_a/2\pi)\exp(-U_b/k_BT). \tag{5.7}$$

With increasing friction, the rate coefficient steadily decreases, and in the high friction limit ($\zeta \gg \omega_b$), it takes the form, $k = k_{\max}\omega_b/\zeta$.

5.5 Results and discussion

5.5.1 Opening of the IC-gate

. In the 3KBC crystal structure of Glt_{Ph}, the IC-gate is closed, and only the Na2 ion has access to the solvent. Thus using this structure, only the binding free energy of the Na2 ion can be calculated. In order to calculate the binding free energies of the other ligands, we need the open conformation of Glt_{Ph} where the IC-gate is open and the ligands are exposed to the solvent. To find the open conformation, we have performed several 30-ns MD simulations of the 3KBC structure with different ligand configurations. The opening of the IC-gate is monitored using the S277(Cα)–V355(Cα) distance, which has been used in studies of the opening of the EC-gate previously.[36] The S277 and V355 residues are located at the tips of the HP1 and HP2 hairpins, respectively, and hence provide a good measure for the gate opening. The results of these simulations are summarized below.

i) All the ligands bound: No perceptible changes are observed in the transporter structure with regard to the opening of the IC-gate. Na2 ion remains bound in all subunits for the entire 30 ns of simulations. Note that the Na2 site is different from the Na2′ site obtained from the crystal structure, which is found to be unstable.[123]

ii) Asp, Na1 and Na3 bound: Na2 ions are removed from all subunits in the last frame of (i) before running the simulations. In Fig. 5.1A, we show the time evolution of the S277–V355 distance for the three subunits. The subunits A and B remain

in the same state as in (i) and do not show any sign of gate opening. In contrast, subunit C displays a spontaneous opening of the IC-gate at 5 ns with a temporary intermediate state between the open and closed configurations as can be seen from the histogram of distance distributions (Fig. 5.1B). We note that the EC-gate has not been observed to open in the presence of Asp in MD simulations of the outward structure. The rare event observed in subunit C here is traced to entry of water molecules to the binding pocket through the HP2 segment, which facilitate breaking of the H-bonds that hold the IC-gate. Inspection of the H-bonds during the gate opening shows that the gate is stabilized by two H-bonds HP2 makes with Asp and HP1, namely, Asp(N)–V355(O) and S279(N)–G354(O) (see Table 5.1). Breaking of the first H-bond results in the half-open intermediate state and breaking of the second H-bond leads to the fully open state. In the open state, Asp and Na1 are exposed to the solvent, so, in principle, both ligands could unbind at this point.

iii) Asp and Na3 bound: Simulations are run after removing Na1 ions from all subunits in the last frame of (ii). As shown in Fig. 5.2A, the S277–V355 distances do not show any gating transitions, with subunits A and B remaining in the closed state and C in the open state. Interestingly, Asp in subunit C remains in the binding site for the whole 30 ns, indicating a semi-stable binding of Asp in the absence of Na1 (see Fig. 5.3A). In contrast, in similar simulations of the outward structure, Asp is released within 5 ns in all subunits. We have, therefore, extended this simulation for a further 20 ns. Asp is eventually observed to lose most of its contacts at 45 ns (Fig. 5.3B).

iv) Na1 and Na3 bound: Removing Asp instead of Na1 in the last frame of (ii) leads to a very different result. The IC-gates in subunits A and B also open but the gates are extremely unstable, fluctuating freely between the open and closed conformations in all subunits (Fig. 5.2B). This further confirms the role of Asp in stabilizing the IC-gate.

v) Na3 bound: Removing Na1 from the last frame of (iv) yields very similar results to that observed in (iv) (not shown).

vi) No ligands: Simulations are run after removing all the ligands from the last frame of (i). All the IC-gates open within 6 ns (Fig. 5.2C). Although the gates still exhibit large fluctuations, they don't visit the fully closed state anymore, indicating that it is less favored in the apo state.

The above results demonstrate that the bound Asp plays an important role in stabilizing the IC-gate via the V355(O)–Asp(N) H-bond. The H-bonds noted in simulation (ii) result in a sizable energy barrier between the open and closed conformations, which suppresses random gating transitions. To confirm the role of these H-bonds in gating and also to create open-states in all subunits, we attempt to break them in simulation (ii) where Asp, Na1 and Na3 are bound. For this purpose, we introduce a repulsive half-harmonic potential defined by $U = 0.5k(r - r_0)^2$ for $r < r_0$ and $U = 0$ for $r > r_0$, with parameters $k = 25$ kcal/mol/Å^2 and $r_0 = 4.5$ Å. This potential acts on the H-bond partners only when the distance between the atoms (r) is less than 4.5 Å. It speeds up breaking of the specified H-bonds and

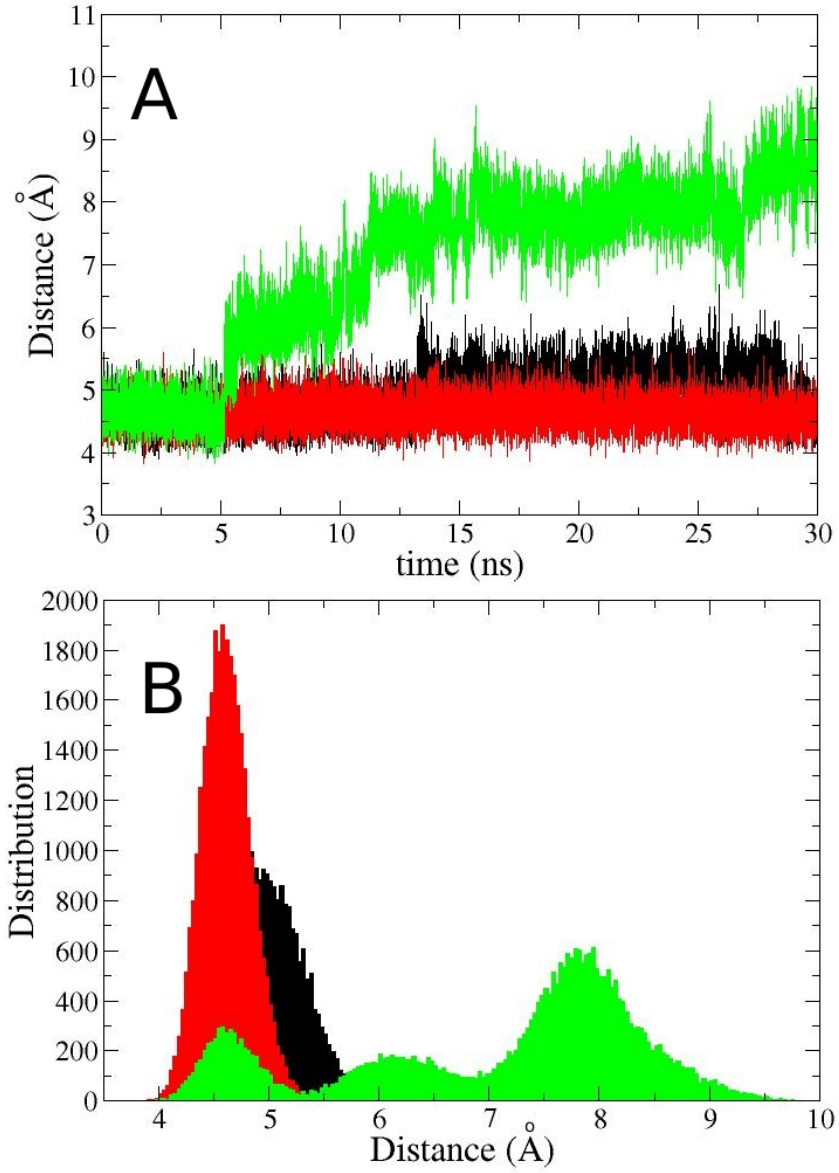

Figure 5.1: (A) Time evolution of the S277(Cα)–V355(Cα) distance in subunits A (black), B (red), and C (green) when all the ligands are bound except Na2. There is a spontaneous opening of the IC-gate in subunit C starting at 5 ns. (B) The histogram analysis of the same graph, showing the unstable intermediate state in subunit C during the opening of the IC-gate.

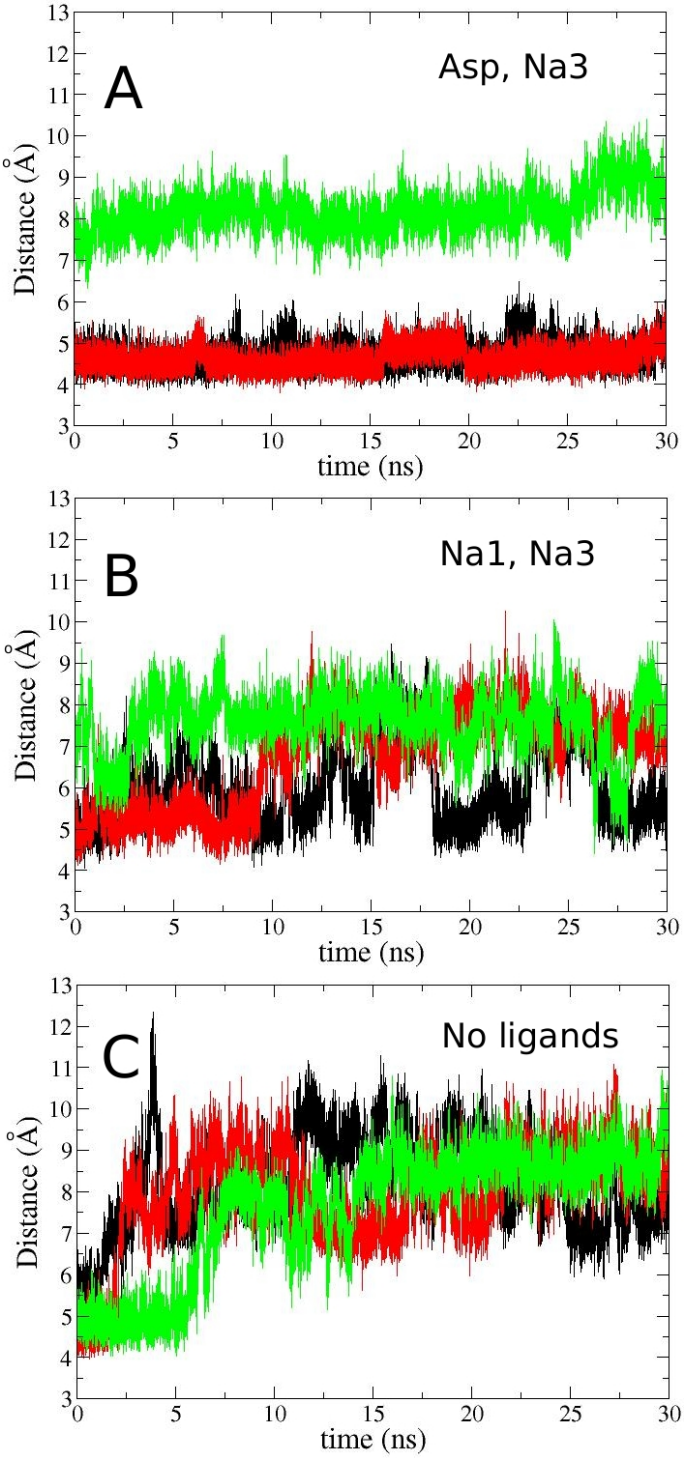

Figure 5.2: Similar to Fig. 5.1A but for different combinations of bound ligands. (A) Only Asp and Na3 are bound. (B) Only Na1 and Na3 are bound. (C) No ligands are bound. The gate distances are stable in case A but exhibit large fluctuations in B and C. This indicates that presence of Asp results in an energy barrier between the open and closed conformations.

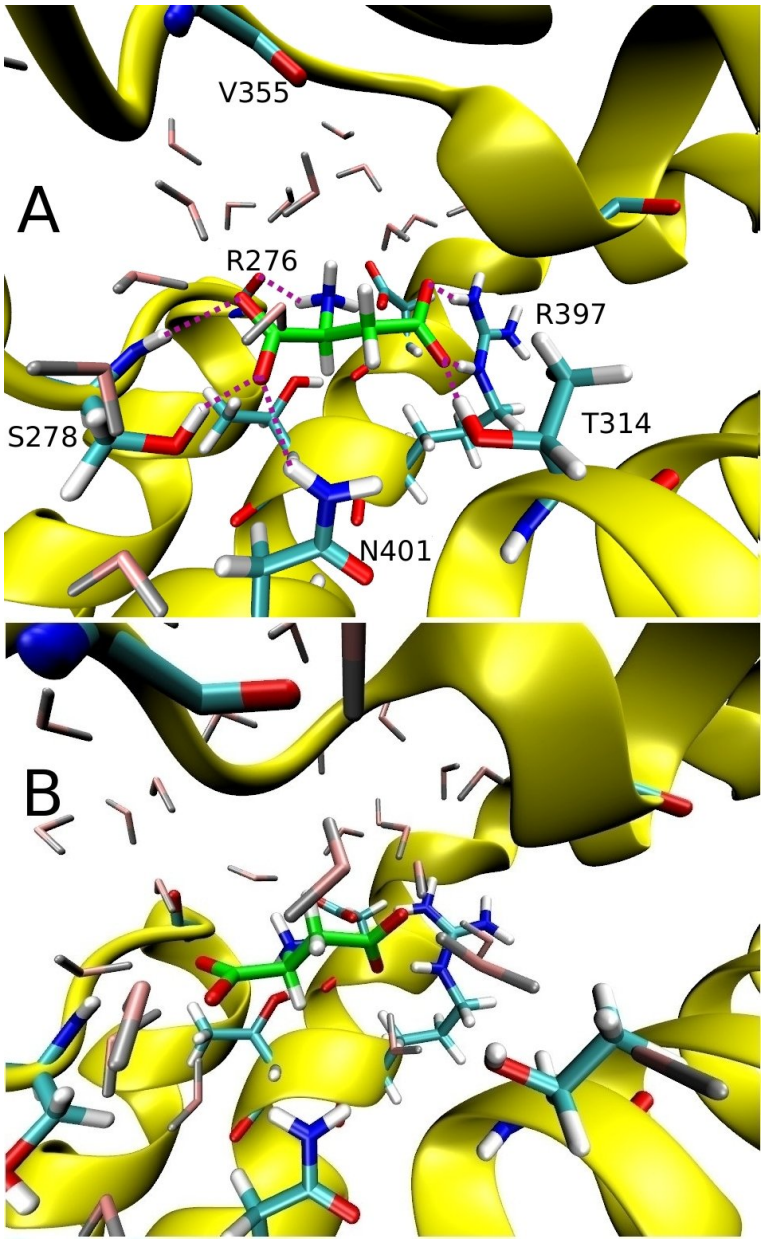

Figure 5.3: (A) Position of Asp in the binding site in the absence of Na1 after 10 ns of MD simulations. Contacts are indicated with purple dashed lines. (B) The same after 45 ns of simulations, where Asp has lost most of the contacts indicated in A.

Table 5.1: List of the Glt_{Ph} residues (first column) coordinating Asp atoms (second column). The average N–O and O–O distances (in Å) obtained from the MD simulations in the closed and open states are compared to those of the crystal structure 3KBC (third column). Bare O and N atoms in parenthesis refer to the backbone atoms and the others refer to the side chain atoms. The first row shows the S279(N)–G354(O) H-bond distance involved in the IC-gate.

Glt_{Ph} residue	Asp residue	Crystal structure	Closed state	Open state
HP1 - S279 (N)	HP2 - G354 (O)	3.3	3.0 ± 0.2	6.0 ± 0.8
HP1 - R276 (O)	Asp (α-N)	3.1	2.9 ± 0.1	3.0 ± 0.2
HP1 - S278 (N)	Asp (α-O_1)	2.5	2.8 ± 0.1	2.9 ± 0.1
HP1 - S278 (OH)	Asp (α-O_2)	3.7	2.7 ± 0.1	2.8 ± 0.1
TM7 - T314 (OH)	Asp (β-O_2)	2.6	2.7 ± 0.1	2.7 ± 0.1
HP2 - V355 (O)	Asp (α-N)	2.1	3.0 ± 0.2	5.5 ± 0.3
HP2 - G359 (N)	Asp (β-O_2)	2.7	3.0 ± 0.2	3.2 ± 0.4
TM8 - D394 (O_1)	Asp (α-N)	3.4	2.7 ± 0.1	2.7 ± 0.1
TM8 - R397 (N_1)	Asp (β-O_2)	3.9	4.3 ± 0.2	2.7 ± 0.1
TM8 - R397 (N_2)	Asp (β-O_1)	2.5	2.9 ± 0.2	2.7 ± 0.1
TM8 - T398 (OH)	Asp (α-N)	3.6	3.2 ± 0.2	3.0 ± 0.2
TM8 - N401 (N_δ)	Asp (α-O_2)	2.9	2.9 ± 0.1	2.9 ± 0.2

thereby facilitates the gate opening process observed in subunit C. Applying this potential to the Asp(N)–V355(O) and S279(N)–G354(O) distances in subunits A and B, we observe that the IC-gate in subunit B quickly opens and remains in this state for almost 10 ns with no bias forces acting as they are outside the range (Fig. 5.4, red line). In subunit A, upon breaking of the S279(N)–G354(O) H-bond, G354(O) forms two intermediate H-bonds with the hydroxyl groups of S277 and S279, keeping the IC-gate closed. These H-bonds are also broken by applying the half-harmonic potentials at 5 ns, after which the IC-gate in subunit A also opens and remains in this state with no bias forces acting (Fig. 5.4, black line). Despite the differences in the opening mechanisms, all three states have very similar final conformations as demonstrated in the alignment diagram in Fig. 5.5. We stress, in particular, the similarity between the open states obtained using the bias forces (subunits A and B) and the one obtained from MD simulations (subunit C). This is further quantified by calculating the RSMDs between the subunits A-C and B-C after aligning the segments HP1, HP2, TM7 and TM8, which yield 0.9 Å in both cases.

The conformational changes occurring during the opening of the IC-gate is demonstrated in Fig. 5.6A, where the open apo state obtained from the MD simulations is superposed with the closed state obtained from the crystal structure.

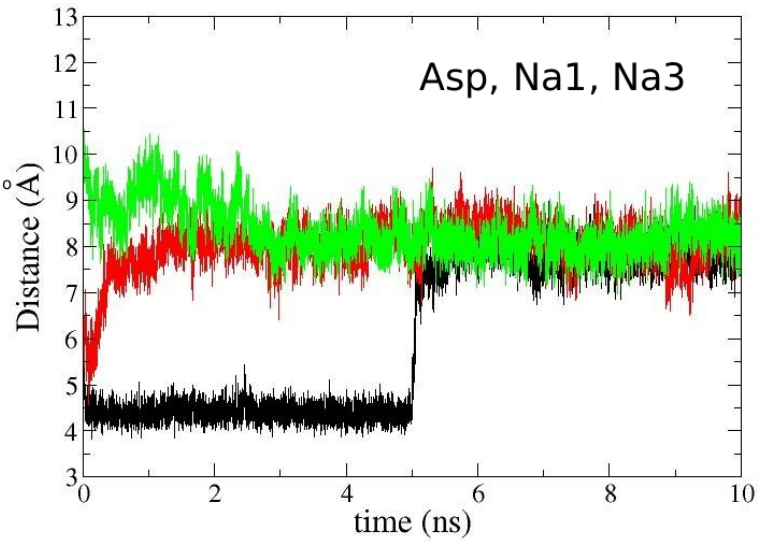

Figure 5.4: Time evolution of the distance in Fig. 5.1A in a 10 ns simulation with Asp, Na1 and Na3 bound. Here a bias force is introduced in subunits A (black) and B (red) to observe the opening of the IC-gate in these subunits. Even though the bias force does not act after the IC-gate opens, the transporter is seen to remain in the open state, indicating presence of an energy barrier between the two states when there is a bound Asp.

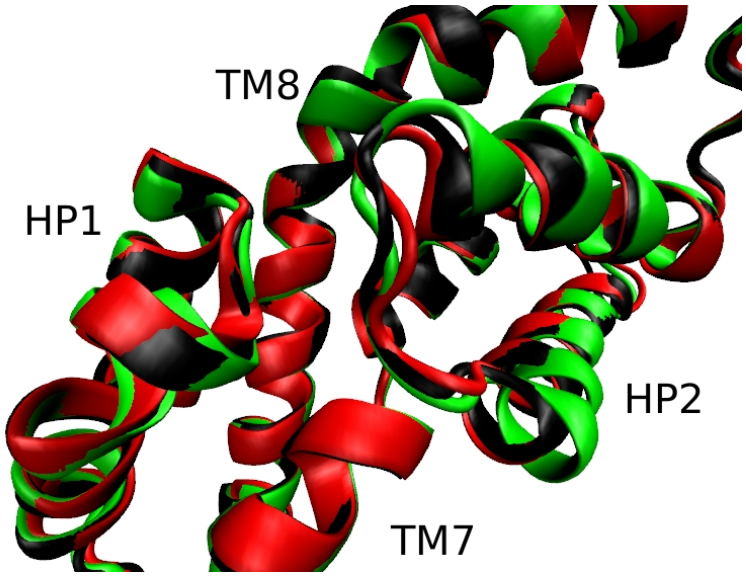

Figure 5.5: Comparison of the open states obtained using bias forces in subunits A (black) and B (red) with the one obtained from MD simulations in subunit C (green).

Opening of the IC-gate involves a partial movement of HP2 towards TM2 and a similarly small movement of HP1 away from HP2, e.g., the S277–V355 distance increases by about 4 Å after the IC-gate opens while the same distance increases by 10 Å in the EC-gate. As shown in Fig. 5.6B, essentially the same open conformation is obtained from the MD simulations with bound Asp, Na1 and Na3, which is used in the free energy calculations. To show the effect of gate opening on the bound Asp, we present in Table 5.1 the residues involved in the coordination of Asp, and how the contact distances change when the IC-gate opens. The S279(N)–G354(O) distance is also given in Table 5.1 because it is involved in the gating. Comparing the results in Table 5.1 with those obtained for the outward-facing conformation (Table 4.2[123]), we find a rather different mechanism. Opening of the EC-gate involves breaking of the V355(O)–Asp(N) and G359(N)–Asp(βO_2) bonds, which allows HP2 to move away from a stationary HP1. In opening of the IC-gate, only the first of these HP2–Asp bonds is broken while the second remains intact and performs a hinge function (Table 5.1). More importantly, none of the HP1–Asp bonds are broken.

In order to quantify and contrast the changes occurring in the IC and EC-gates further, we show the RMSD of the Cα atoms at the tip of HP1 and HP2 using the closed crystal structures as reference (Fig. 5.7). Here it is clearly seen that HP2 is the only moving part in the EC-gate whereas both HP1 and HP2 move in the IC-gate but to a much lesser degree. The movement is also less uniform across the tip with the maximum occurring in the residues S279 in HP1 and A353 in HP2 (indicated in Fig. 5.6B). As noted above, the movement of the HP2 tip is restricted by the unbroken G359–Asp H-bond, which results in partial opening of HP2 around the G359 hinge. We also see that the HP1–Asp bonds are maintained during gating because HP1 moves only a small distance. Thus the IC-gate involves a smaller opening and preserves the Asp contacts better relative to the EC-gate. As a result we expect Asp to have a higher affinity on the IC side compared to the EC side, which will affect its release to the cytoplasm. This has already been noted in case (iii) above where Asp is observed to remain in the binding site much longer in the absence of Na1. In order to make sure that the smaller opening of the IC-gate does not prevent the release of Asp, we have facilitated its release by applying repulsive harmonic potentials to the N–O contacts between the R397 side chain and Asp(β-carboxyl), and between D394 side chain and Asp(α-amino). The same values of potential parameters are employed as in breaking of the H-bonds in the IC-gate. The key events leading to the release of Asp are captured in four snapshots in Fig. 5.8:

Frame A: This is the same snapshot as in Fig. 5.3A (MD simulations with only Asp and Na3 bound after 45 ns). Most of the contacts with the Glt_{Ph} residues have already been lost, and Asp is relatively free to move and rotate within the binding pocket. The half-harmonic potentials are applied at this point and the system is simulated until Asp is released.

Frame B: In this position, the α-carboxyl group of Asp still makes contacts with S278(N) and S279(N), and the β-carboxyl group of Asp with G357(N). There are

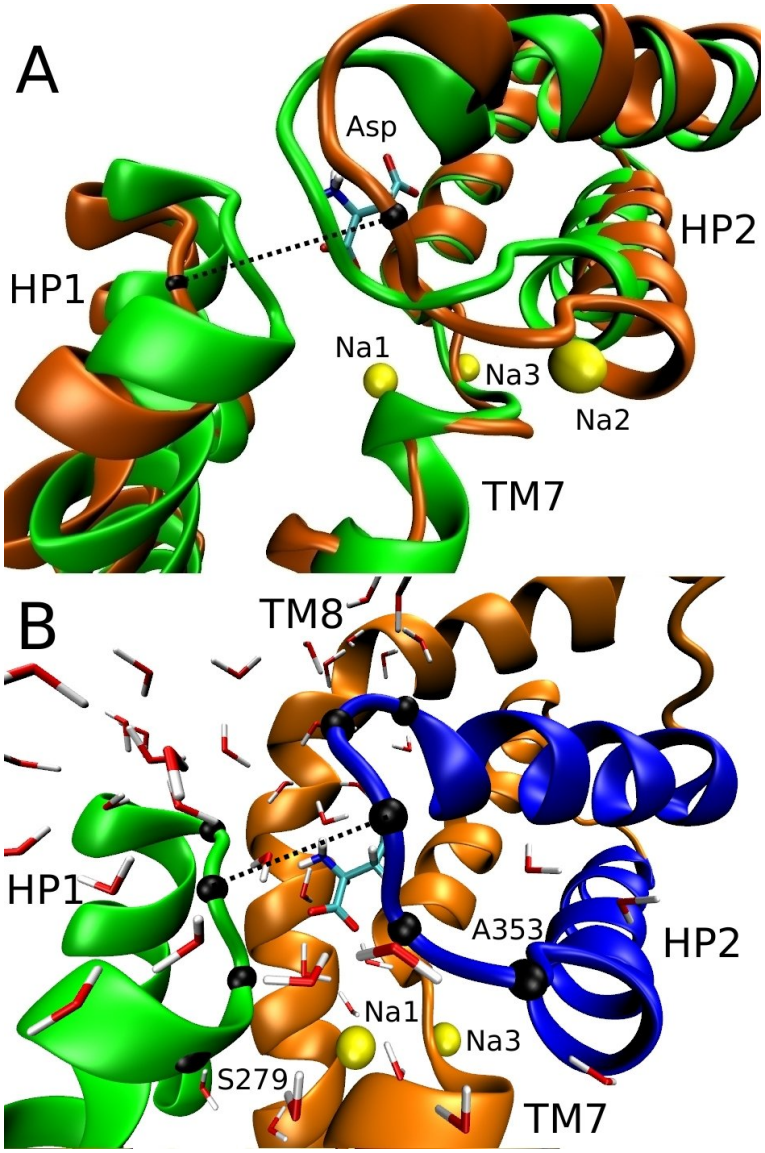

Figure 5.6: (A) Comparison of the closed state taken from the crystal structure (green) and the open state obtained from the MD simulations with no bound ligands (orange). The black dotted line indicates the S277(Cα)–V355(Cα) distance used in measuring the opening of the IC-gate. (B) The open state obtained from the simulations with Na1, Na3 and Asp bound. In this conformation both Na1 and Asp have access to the solvent. The Cα atoms used in RMSD measurements in Fig. 5.7 (R276–S279 and A353–G357) are indicated with black spheres.

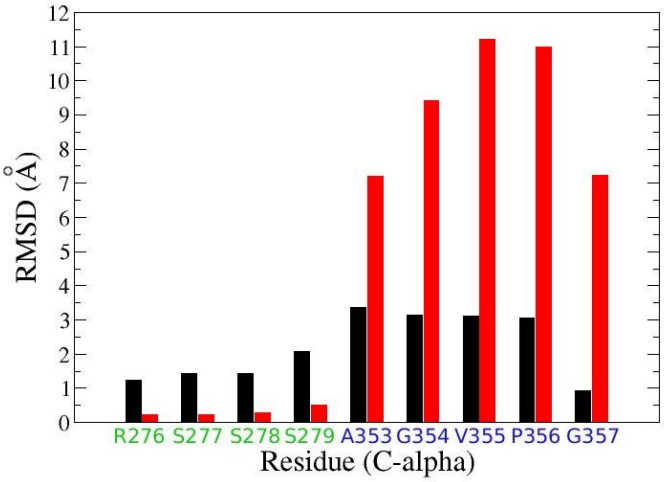

Figure 5.7: Comparison of the IC-gate opening (black) with the EC-gate opening (red). The RMSD of the open-state Cα atoms at the tips of HP1 and HP2 are measured with respect to the closed crystal structure after aligning the segments HP1, HP2, TM7 and TM8.

no other contacts made by Asp.

Frame C: At this point the G357(N) contact is lost, and the Asp side-chain moves up towards the solution and away from the binding site. The S278(N) and S279(N) contacts are still maintained.

Frame D: Finally the S278(N) and S279(N) contacts are also lost, and Asp is released to the cytoplasm.

A similar picture is obtained for the opening of the IC-gate from metadynamics simulations, e.g., the IC-gate remains closed in the presence of Na2, opening of the IC-gate involves movement of both HP1 and HP2, and there is an energy barrier between the open and closed conformations.[126] The metadynamics simulations were performed for a monomer in solution with two bound Na^+ ions. Our results confirm that the basic features of the gating dynamics are not influenced by embedding the Glt_{Ph} trimer in membrane and including the third Na^+ ion. This indicates that Na3 is not involved in gating or coupling to Asp, rather it provides the electrostatic energy needed to move the binding pocket from the EC to IC side. This energy comes from the membrane associated voltage of around -80 mV, which produces an inward-directed potential for positive charges.

5.5.2 Binding free energies

. The binding free energies of Na^+ ions are calculated using the formalism described in methods and the results are presented in Table 5.2. Convergence of the translo-

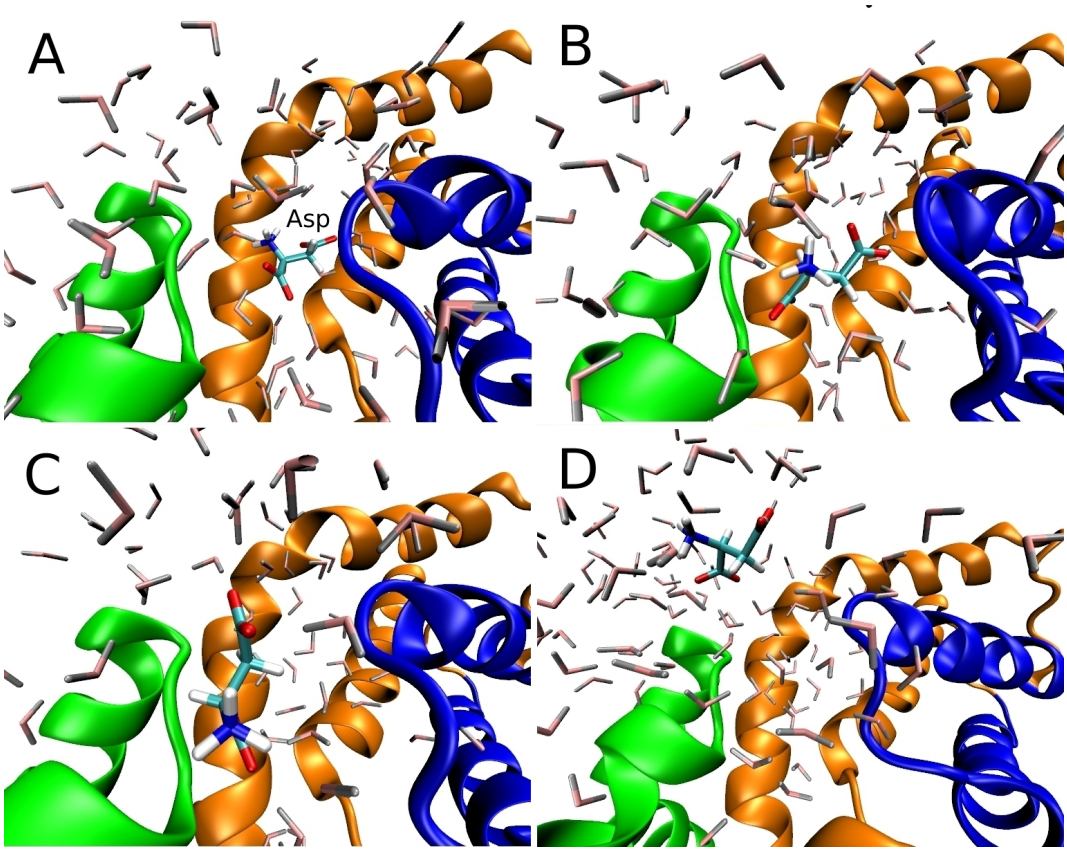

Figure 5.8: Four snapshots depicting the release of Asp. Explanations are given in the text.

Table 5.2: Binding free energies for Na^+ ions in the closed state (for Na2 and Na2′) and in the open state (for Na1 and Na3) of Glt_{Ph}. The forward and backward calculations of the interaction energy are listed in the second and third columns, and their average in the fourth column. The total binding free energy that includes the translational free energy difference (fifth column) is given in the last column. Other ligands present in the transporter are indicated in parenthesis. Errors are estimated from block data analysis using 100 ps windows. All energies are in kcal/mol.

Ion	$-\Delta G_{\text{int}}^{\text{for}}$	$\Delta G_{\text{int}}^{\text{back}}$	ΔG_{int}	ΔG_{tr}	ΔG_{b}
Na2 (1, 3, Asp)	-6.8 ± 1.2	-6.7 ± 1.1	-6.8 ± 1.2	4.4	-2.4 ± 1.2
Na2′ (1, 3, Asp)	-2.1 ± 0.8	-2.1 ± 1.0	-2.1 ± 0.9	4.3	$+2.2 \pm 0.9$
Na1 (3, Asp)	-19.3 ± 1.2	-20.0 ± 1.3	-19.6 ± 1.3	4.7	-14.9 ± 1.3
Na1 (3)	-12.3 ± 1.6	-11.8 ± 1.8	-12.1 ± 1.3	4.8	-7.3 ± 1.3
Na3	-21.5 ± 1.5	-19.9 ± 1.2	-20.7 ± 1.1	4.4	-16.3 ± 1.1

cation free energies and lack of hysteresis effects are demonstrated in Fig. 5.9. We first discuss binding of Na2, which is bound to the closed state and already exposed to the solvent, so it should be the first ligand to be released. The free energy calculation is performed using the closed crystal structure 3KBC. In the outward crystal structure, Na2 is coordinated by the backbone oxygens of T308, S349, I350 and T352. But in MD simulations, Na2 is observed to shift away from the S349 backbone oxygen immediately, which we call the Na2′ site.[123] With further equilibration, the S349 oxygen is replaced by the side chain oxygen of T308 in all subunits and Na2 remains stably bound in this position.[123] The same thing is observed to happen in the inward structure simulations albeit more slowly, resulting in an identical coordination shell for Na2 as in the inward structure. The calculated Na2 binding free energy is −2.4 kcal/mol, which is also similar to the value obtained for the outward structure, −2.7 kcal/mol.[123] The binding free energy for the Na2′ site is +2.2 kcal/mol, indicating its instability. We note that the Na2 site is not as well located as the Na1 and Na3 sites due to Tl^+ substitution and intrinsic flexibility of the binding site, so there is still some uncertainty about its exact location.

Once the IC-gate is open, the binding pocket is exposed to the solvent, and either Asp or Na1 can unbind. To cover both possibilities, we calculate the Na1 binding free energy in two cases: with only Na3 bound, and with Na3 and Asp bound. These and the remaining free energy calculations are performed using the open structure obtained from the MD simulations (subunit C in Fig. 5.1). The Na1 binding free energy in the presence of Na3 is found to be −7.3 kcal/mol, which is very close to the value obtained for the outward structure under the same conditions (−7.1 kcal/mol).[123] This is expected because the same residues coordinate Na1 in the inward and outward structures. From Table 5.2, presence of Asp is seen to have a big influence on the binding affinity of Na1, boosting its binding free energy

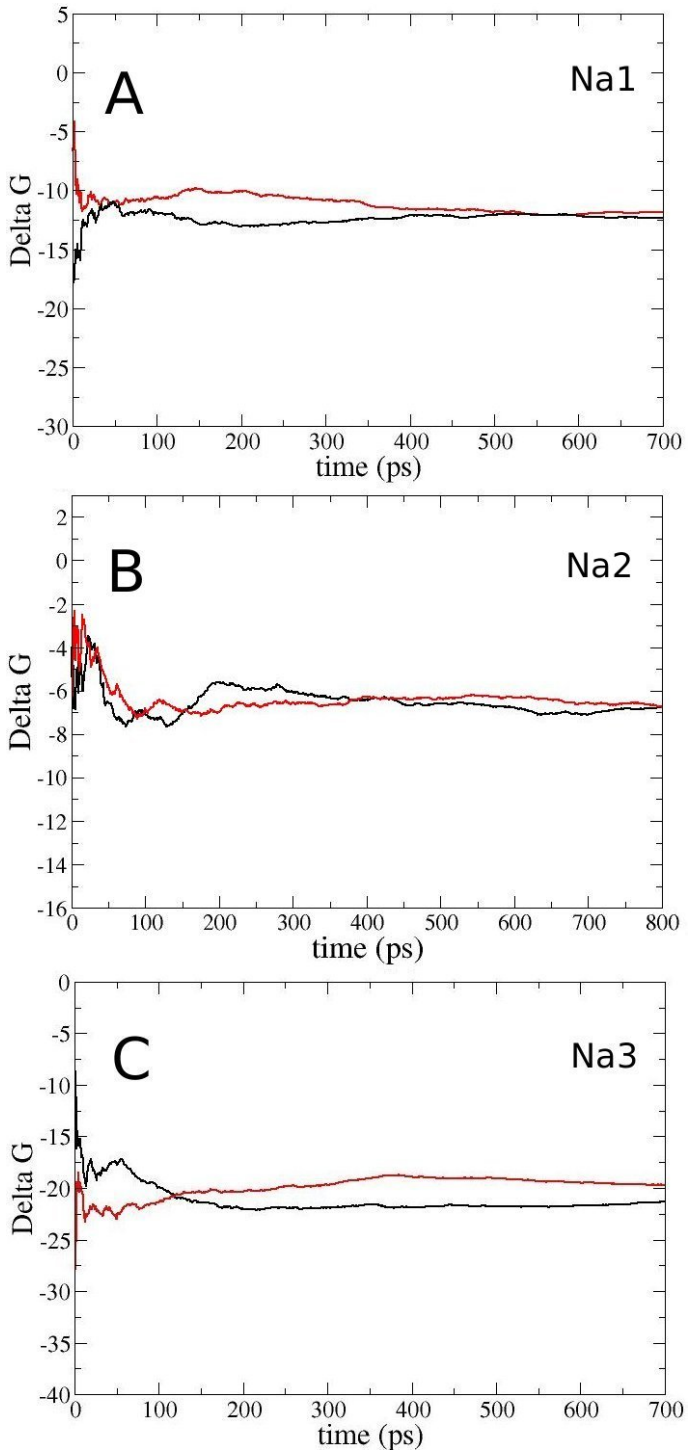

Figure 5.9: Convergence of the TI results for the translocation free energy of the (A) Na1, (B) Na2 and (C) Na3 ions from running averages. For the Na1 ion we show the case with only Na3 bound, and the Na2 calculation is for the closed state with all the ligands present. The negative of the forward (binding site → bulk) and the backward transitions of Na^+ are shown with black and red lines, respectively.

by −7.6 kcal/mol to −14.9 kcal/mol. In the outward structure calculations,[123] we provided a qualitative demonstration of the Asp–Na1 coupling via a hydrogen-bond network, which linked the α-carboxyl group of Asp with Na1. Because Asp(α-carboxyl) and Na1 are separated by 7 Å with a water molecule and the S278-OH group in between them (see Fig. 5.10), the strength of this coupling is not clear. In the outward structure, Asp is not stable in the absence of Na1, and therefore the Na1 binding free energy could not be calculated in the presence of Asp. In the inward structure, Asp remains stable in the absence of Na1 long enough to allow such a calculation. The above estimate of 7.6 kcal/mol difference in Na1 binding free energies with and without Asp indicates that the Asp–Na1 coupling is fairly strong despite their sizable separation. We emphasize this point because, unlike in LeuT transporters where a Na^+ ion is in direct contact with the substrate,[121] the nature and strength of the substrate-ion coupling is not obvious in Glt_{Ph}.

In order to investigate the nature of the Asp–Na1 coupling and examine the role of the electrostatic interactions between Asp and Na1 in this coupling, we repeat the Na1 binding free energy calculation in the presence of Na3 and Asp but with a neutralized α-carboxyl group. This is expected to abolish the electrostatic interaction between Asp and Na1 because the β-carboxyl and α-amino groups of Asp are at similar distances from Na1 (10 Å) so their effects should mostly cancel each other (Fig. 5.10). The calculated binding free energy (−4.2 kcal/mol) is about 3.1 kcal/mol larger than the value obtained in the absence of Asp (−7.3 kcal/mol). Thus other factors also contribute to the Asp–Na1 coupling besides the direct electrostatic interactions. For example, the α-carboxyl group of Asp is involved in the H-bond network responsible for binding of Asp,[123] and its neutralization will affect this H-bond network.

Na3 is buried in the protein and can only be released after Na1 and Asp. We calculate the binding free energy of Na3 using the last frame from the 20 ns MD simulations with only Na3 bound. The calculated value (−16.3 kcal/mol) is again comparable to the value obtained for the outward-facing state (−18.7 kcal/mol), reflecting the fact that Na3 coordination is also similar in the inward and outward structures. The Na3 ion has by far the largest affinity of all the ligands, with a high enthalpic barrier (20.7 kcal/mol) that has to be crossed for this ion to be released. We use Kramers' rate theory discussed in Methods to provide a rough estimate for the release time of Na3. The barrier potential is specified with height $U_b = 20.7$ kcal/mol and range $x_b \sim 4$ Å (half of the distance from the Na3 site to solvent which is 8 Å), from which we obtain $\omega_b = 7 \times 10^{12}$ s^{-1}. Similarly, from the RMSD fluctuations of the Na3 ion in the binding pocket ($\sigma_x = 0.2$ Å), we obtain $\omega_a = 16 \times 10^{12}$ s^{-1}. From Eq. 5.7, these values correspond to a maximum rate of $k_{\max} = 2.9 \times 10^{-3}$ s^{-1} at $T = 30$ °C, or average release time of 6 min. Including friction in the rate equation (Eq. 5.6) will lower the rate and increase the release time. The friction coefficient of Na^+ ions in bulk water is $\zeta = 8 \times 10^{13}$ s^{-1}, which is larger than ω_b. However, there are no water molecules in the Na3 binding pocket, and hence we expect ζ to be much smaller than the bulk value. As an example, for

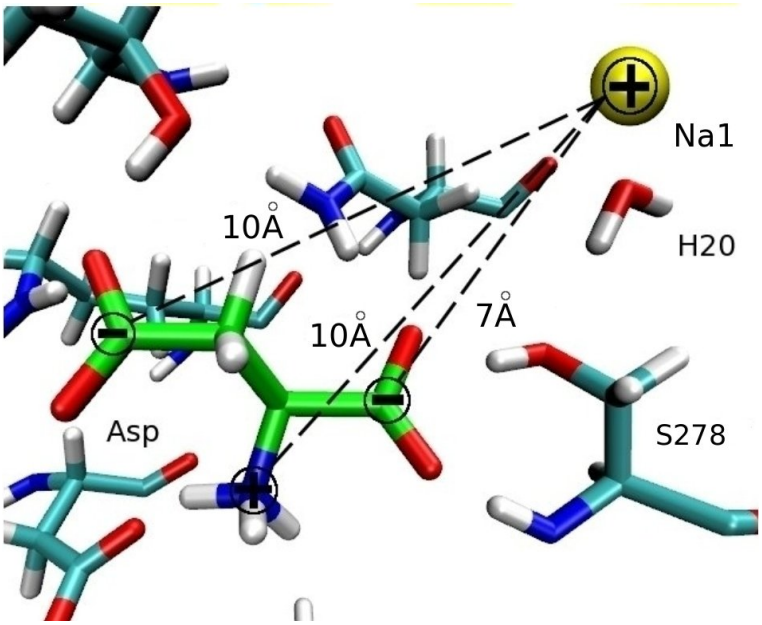

Figure 5.10: Picture depicting the strong coupling between Asp and Na1. The distances between the charges on Asp and Na1 are indicated with dashed lines. Asp α-carboxyl is closer to Na1 which is the source of strong coupling while the other positive and negative charges on Asp are at equal distance to Na1 and hence cancel each other.

$\zeta \sim \omega_b$, the rate would be $k = 0.6k_{\max} = 1.7 \times 10^{-3}$ s^{-1} corresponding to an average release time of 10 min.

The experimental turnover time of Glt_{Ph} is 3.5 min at $T = 30$ °C.[35] There is strong support from mutagenesis experiments that unbinding of Na3 is the rate-limiting step in the transport cycle. Mutation of the residues coordinating Na3, which reduce its binding affinity and facilitate its release, result in almost 20-fold increase in the transport rate.[38] Experiments also indicate that the outward-inward transition is relatively rapid (seconds) and does not limit the rate of transport.[16] Thus, assuming that the rate-limiting step in the transport cycle is the release of Na3, the calculated release time is consistent with the experimental turnover time (note that the computational accuracy of free energies is about 1 kcal/mol which translates to a factor of five in the rate coefficients or release times).

While the release time of the Na^+ ions is quite long, their electrostatic potential energy provides sufficient incentive for keeping the transporter in the inward conformation until they are released. For a membrane potential of -80 mV, transport of each Na^+ ion against this potential gradient costs $80/25 = 3.2$ kT. For two Na^+ ions, the cost is 6.4 kT. Thus compared to the apo state, transition of the Na1-Na3 bound transporter to the outward-facing state is suppressed by a factor of $\exp(-6.4) \sim 1/600$.

Temperature dependence of Asp uptake provides another test for the proposed model. The Q_{10} value estimated from the Na3 release rates at $T = 25$ and 35 °C is

Table 5.3: Various terms contributing to the binding free energy of Asp in the open state of Glt_{Ph}. The forward and backward calculations differed by less than 1 kcal/mol in all cases, therefore only their average is listed. Other ligands present in the transporter are indicated in parenthesis. The absolute binding free energy of Asp is given in the last column. Errors are estimated from block data analysis using 100 ps windows. All energies are in kcal/mol.

Substrate	ΔG_{elec}	$\Delta G_{\mathrm{LJ-bb}}$	$\Delta G_{\mathrm{LJ-sc}}$	ΔG_{tr}	ΔG_{rot}	ΔG_{con}	ΔG_{b}
Asp (1,3)	-17.0 ± 1.0	2.6 ± 0.6	0.3 ± 0.6	3.9	4.6	0.7	-4.9 ± 1.1
Asp (3)	-9.3 ± 1.2	1.8 ± 0.7	0.4 ± 0.5	3.9	4.4	0.6	$+1.8 \pm 1.3$

3.1, which is in reasonable agreement with the measured Q_{10} value of 3.7.[35]

We next consider the binding free energy of Asp in the presence of Na1 and Na3, which is calculated using the open subunit from the last frame of the MD simulations in Fig. 5.1A (Table 5.3). Convergence of the Asp translocation free energies for various contributions and lack of hysteresis effects are demonstrated in Fig. 5.11. The calculated value (−4.9 kcal/mol) is 10 kcal/mol higher than the binding free energy of Na1 in the presence of Asp and Na3 (−14.9 kcal/mol), which assures that Na1 is released only after Asp. Thus the ligands are released into the cytoplasm in the order of Na2, Asp, Na1 and Na3, which is the reverse of their binding order in the outward-facing state.[123] As noted in the discussion of the IC-gate, Asp is better coordinated in the inward structure compared to the outward structure, and this is corroborated by its binding free energy which is 1.1 kcal/mol lower in the inward state relative to the outward state. In fact, the enthalpic contribution to the binding free energy is 2.6 kcal/mol lower in the inward state but this is partly offset by the entropic contributions which are also higher. All of these point to a tighter binding of Asp in the inward state.

As pointed out above, Asp remains stable in the absence of Na1 long enough (> 30 ns) to allow free energy calculations. This means that we can also calculate the binding free energy of Asp in the presence of only Na3 (Table 5.3). The calculation is performed using the open subunit of the MD simulations in Fig. 5.2A. The calculated value is 1.8 kcal/mol, showing that Asp is indeed unstable in this configuration, but the relatively large enthalpic contribution to the binding free energy (−7.1 kcal/mol) delays its release. Comparing the Asp binding free energies with and without Na1 (Table 5.3) gives another estimate for the Asp–Na1 coupling. The binding free energy difference is 6.7 kcal/mol, which is consistent with the value obtained from the Na1 binding free energies with and without Asp (7.6 kcal/mol).

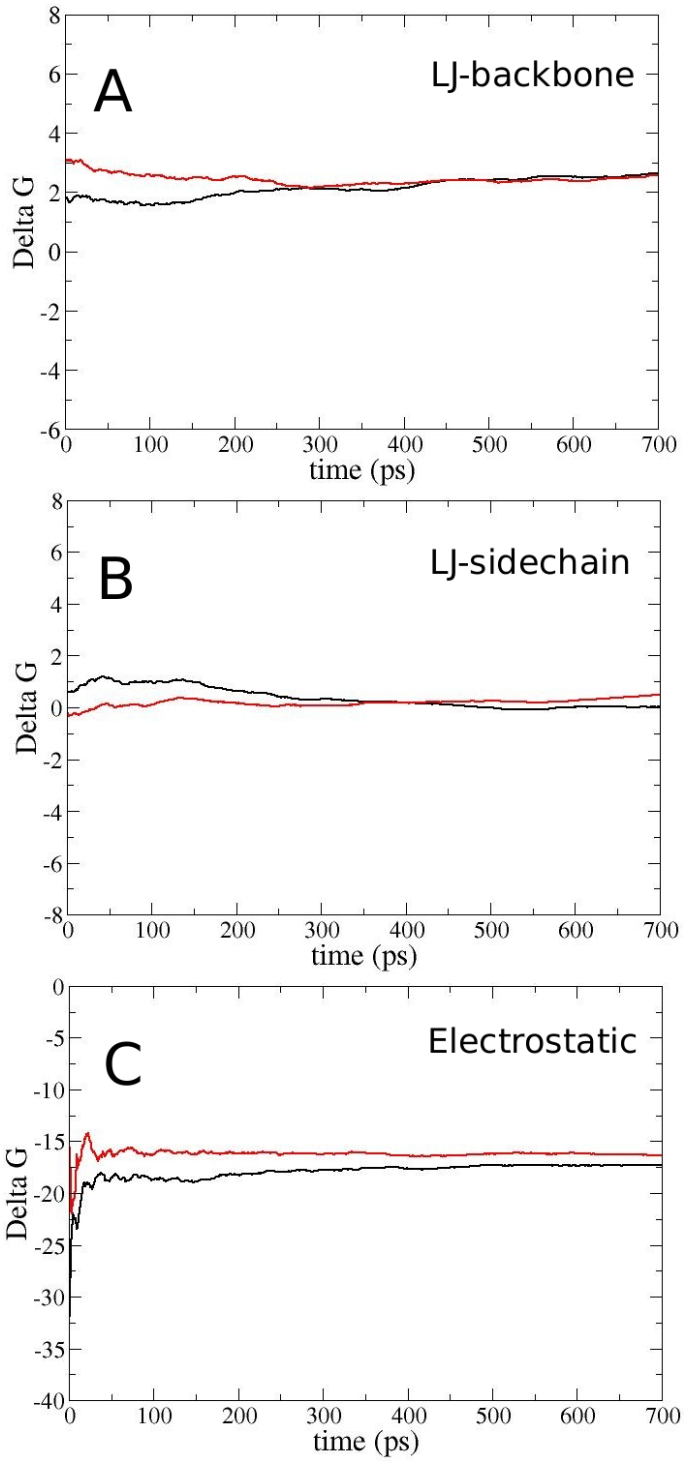

Figure 5.11: Convergence of the TI results for the three different components of the translocation free energy of aspartate from running averages: (A) the creation/destruction of the neutral backbone, (B) the same for the side-chain, and (C) the charging/discharging of the whole aspartate molecule. Here we show the case in which we have both Na1 and Na3 bound to the transporter. The negative of the forward (binding site → bulk) and the backward transitions of Na^+ are shown with black and red lines, respectively.

5.6 Conclusions

We have performed MD simulations and free energy calculations in the inward-facing state of the Asp transporter Glt_{Ph}. The gating mechanism is studied in equilibrium simulations with different ligands bound to the transporter. Spontaneous opening of the IC-gate with Na1, Na3 and Asp bound is observed in one of the MD simulations, which happens through the breaking of two hydrogen bonds. The IC-gate is found to be quite different from the EC-gate as it involves movement of both HP1 and HP2 but in much smaller amounts. Our results indicate that presence of Asp leads to an energy barrier between the open and closed states, consistent with the metadynamics simulations.[126] This mechanism presumably facilitates the unbinding of Asp by keeping the transporter in an open state until it is released.

The unbinding order predicted from the binding free energies is, Na2, (IC-gate opens), Asp, Na1 and Na3, which is the reverse of the order of binding found in the outward structure.[123] Another difference between the inward and and outward structure calculations is the semi-stable binding of Asp in the inward case when only Na3 is present. This has allowed us to give a quantitative estimate for the strength of the coupling between Asp and Na1, which was only qualitatively described in the previous work.[123] The binding free energy of Na3 is found to be substantially lower than the other ligands, suggesting that its unbinding is the the rate-limiting step in the transport process. Estimates of the release time obtained from Kramers' rate theory is consistent with the experimental turnover rate of Glt_{Ph}, which provides further support for this proposal.

The free energy calculations show that the binding free energies of Na^+ ions and Asp in the inward structure are very similar to those obtained from the outward structure. This is expected from the similarity of the binding sites in the inward and outward open state structures. The faster turnover rates observed in mammalian glutamate transporters could result from the extra proton transported and/or changes in the binding sites during the transition, which could facilitate the release of ligands. It will be interesting to see if the homology models of mammalian glutamate transporters could provide such mechanistic explanations for the differences observed between them and Glt_{Ph}. Although the results showed in this work and our previous study of Glt_{Ph}[123] focus mainly on thermodynamic properties of ligand binding, the kinetics of transport also plays an important role in the understanding of Glt_{Ph} and the EAATs. These properties are harder to obtain due to the different time scales of laboratory experiments and MD simulations. Simulations that can reproduce the unbinding of the substrate as shown in Ref.[123], or the calculation of transitions rates using Kramers' theory and equilibrium properties, are an effort in that direction. Unfortunately, we are still limited to the calculation of the free energies between bound or unbound states using FEP, which usually do not bring insights into the further kinetic properties of our systems.

Chapter 6

Molecular Dynamics Simulations of the Mammalian Glutamate Transporter EAAT3

ABSTRACT
Excitatory amino acid transporters (EAATs) are membrane proteins that enable sodium-coupled uptake of glutamate and other amino acids into neurons. Crystal structures of the archaeal homolog Glt_{Ph} have been recently determined both in the inward- and outward-facing conformations. Here we construct homology models for the mammalian glutamate transporter EAAT3 in both conformations and perform molecular dynamics simulations to investigate its similarities and differences from Glt_{Ph}. In particular, we study the coordination of the different ligands, the gating mechanism and the location of the proton and potassium binding sites in EAAT3. We show that the protonation of the E374 residue is essential for binding of glutamate to EAAT3, otherwise glutamate becomes unstable in the binding site. The gating mechanism in the inward-facing state of EAAT3 is found to be different from that of Glt_{Ph}, which is traced to the relocation of an arginine residue from the HP1 segment in Glt_{Ph} to the TM8 segment in EAAT3. Finally, we perform free energy calculations to locate the potassium binding site in EAAT3, and find a high-affinity site that overlaps with the Na1 and Na3 sites in Glt_{Ph}.

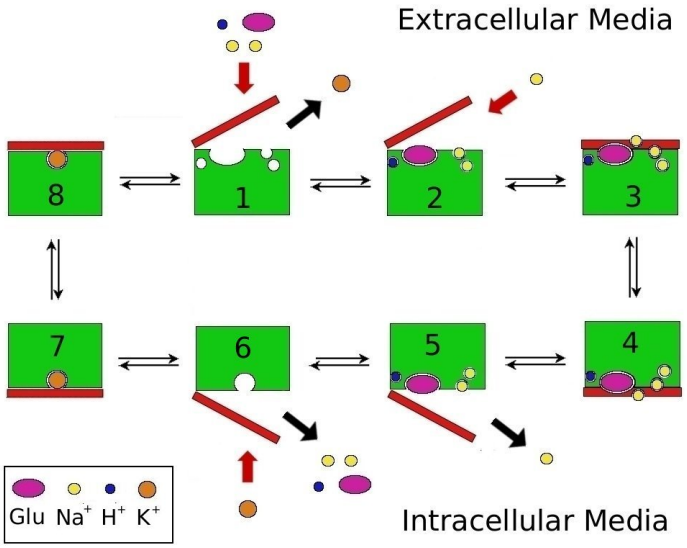

Figure 6.1: Cartoon showing the functional mechanism of the EAAT transporters. Steps 3-4 denote the translocation of the binding site across the membrane with 3 Na^+, H^+, and glutamate bound to EAAT3, while steps 7-8 denote the same with only K^+ bound. Step 2 shows the binding of the Na2 ion, which binds after the substrate and the closure of the HP2 gate. Step 5 shows the opposite happening in the intracellular media, assuming that the binding and unbinding of ligands are symmetrical.

6.1 Introduction

Excitatory Amino Acid Transporters (EAATs) are responsible for clearing the excess glutamate released at the nerve synapses. There are five subtypes identified in humans, called EAAT1 to EAAT5. Malfunctioning of these proteins has been implicated in many pathological conditions including cerebral ischemia, amyotrophic lateral sclerosis and Alzheimer's disease[12]. The transport mechanism of EAATs involves two half cycles, which enables transport of Glu or Asp against their concentration gradients. The first part involves the binding and translocation of three Na^+ and one H^+ ions with the substrate across the membrane while the second half cycle brings the transporter back to its initial conformation by binding and counter-transporting one K^+ ion[13,85] (Figure 6.1). Many mutagenesis and functional experiments have been performed on EAATs in the past, especially in the rodent homologs of EAAT2 and EAAT3, also known as GLT-1 and EAAC1, respectively. We have summarized the results of several studies concerning the residues close to the binding site in Table 6.1, pointing out the effects of different mutations on the main functional properties of the EAATs. These include transporter/exchanger functionality, the cation and substrate affinities, the Na^+/Li^+ selectivity, and the ability to interact with K^+ ions[28,30–34,38,92,96,129–142]. We will refer to this table to motivate the choices made in modeling of EAAT3 and justify the simulation results.

A recent breakthrough in our understanding of the EAATs came with the crystal structure of the archaeal homolog Glt_{Ph}[14], and the identification of the Asp substrate and two of the Na^+ binding sites (Na1 and Na2)[15]. These initial structures captured Glt_{Ph} in the outward-facing state. The inward-facing state was similarly characterized in a subsequent crystal structure[16]. Glt_{Ph} shares about 37% sequence identity with the EAATs, but the homology is much higher for residues near the binding pocket, reaching around 60% in this region. Furthermore, almost all of the residues shown to be involved in the binding of the ions and aspartate in Glt_{Ph} are conserved in the EAATs[14–16,32,90,94]. Experiments have shown that, while the transport mechanism in Glt_{Ph} is independent of H^+ and K^+ ions[35], it is still coupled to the co-transport of three Na^+ ions as in EAATs[37]. The third Na^+ site (Na3) could not be identified in the crystal structures of Glt_{Ph}, but a binding site determined from molecular dynamics (MD) simulations was confirmed from mutagenesis experiments and shown to be conserved in Glt_{Ph} and the EAAT family[38]. Because of its structural and functional similarity to EAATs, Glt_{Ph} provides a good model for studying the transport mechanism in EAATs[112,113]. This has prompted several MD simulations of Glt_{Ph}, investigating the gating mechanism in the outward[36,99,126] and inward-facing states[126,143,144], location of the Na3 site[38,96,97], binding free energies of the ligands[96,123,144], and substrate translocation[114,124,125].

In order to facilitate homology modeling of EAATs from Glt_{Ph}, we list in Table 6.2 all the functionally important residues close to the binding site, which are identified from the crystal structures and the mutation experiments in Table 6.1. Considering the conserved residues first, there seems to be good agreement between the crystal structure data and mutagenesis experiments, e.g., identification of D444 and R447 in EAAT3 as being involved in the coordination of glutamate α-amino and side-chain carboxyl groups, respectively[32,33]. There is, however a discrepancy related to D455 in EAAT3, which corresponds to D405 in Glt_{Ph}. In Glt_{Ph}, D405 coordinates the Na1 ion with both side-chain carboxyl oxygens. But the D455N mutation in EAAT3 has shown little effect on the Na^+ affinity both in the presence and absence of the substrate (Table 6.1). To resolve this issue, it has been proposed that D455 might be protonated during the whole transport cycle, which would explain why neutralizing this residue does not affect the Na^+ affinity[141]. Another important difference between EAATs and Glt_{Ph} is the K^+ binding site in the inward-facing state of EAATs, which is not present in Glt_{Ph}. Different sites have been suggested for K^+ binding, overlapping with either the proposed protonation site[28], or the substrate α-amino group[95], or the Na1 site[140,141].

There are only six non-conserved mutations between Glt_{Ph} and EAATs in Table 6.2. Of these, R276 and T352 contribute backbone carbonyls to the coordination of the Glu substrate and the Na2 ion, respectively, while the others are not directly involved in ligand coordination. Mutation of R276 and M395 in Glt_{Ph} to S331 and R445 in EAAT3 switches the position of the arginine residue from the HP1 segment in Glt_{Ph} to the TM8 segment in EAATs, which will be discussed later in the paper. The Q242 residue in Glt_{Ph} is a histidine in EAATs, which has been proposed as the

protonation site in EAATs. But this was contradicted by the H296Q mutation in EAAT3, which had no effect on the transport mechanism (Table 6.1). The Q318 residue in Glt_{Ph} is a glutamate in EAATs, suggesting that this might be the protonation site. There is strong experimental evidence for this hypothesis, e.g., the E374Q mutation does not affect the Glu affinity of EAAT3 (Table 6.1), but it abolishes the pH dependence of Glu translocation seen in the wild-type transporter[137,145]. Another piece of evidence comes from the neutral amino acid transporters, where the corresponding residue is also a glutamine like in Glt_{Ph}, and the pH sensitivity can be engineered into these transporters by mutating this glutamine to glutamate[137]. The G354 residue in Glt_{Ph} is conserved in EAAT1 and EAAT3, but it is a serine in EAAT2. Among EAATs, EAAT2 is the only one where Li^+ cannot replace Na^+ in the transport cycle. Either the S440G or Y403F mutation enables Li^+ to support transport in EAAT2, indicating that the S440 and Y403 residues are involved in the Na^+/Li^+ selectivity of EAAT2[34].

Here we attempt to address the issues raised above by performing MD simulations on two homology models for EAAT3 based on the Glt_{Ph} transporter in the inward- and outward-facing states. We analyze several properties of EAAT3 such as the coordination of the ligands and the stability of the substrate with a protonated/deprotonated E374 side chain, the effect of the D455N mutation (or protonation of D455) on Na^+ affinity, and the extracellular (EC) and intracellular (IC) gating mechanisms. We also discuss binding of a K^+ ion in the inward-facing state and propose a binding site based on free energy calculations.

Table 6.1: Effect of different mutations on the properties of EAAT1, EAAT2 (GLT-1) and EAAT3 (EAAC1) transporters. Transporter/Exchanger functionality is given in the third and fourth columns. Columns 5 and 6 gives the Na^+ affinity in the presence and absence of glutamate, respectively. Columns 7 and 8 specify K^+ and Li^+ interactions with the transporter. Glu and Asp affinities are given in the last two columns. For the affinities, N/S means no saturation of the binding site in the range considered, and nRed/nInc means reduction/increase in the ligand affinity by n orders of magnitude (0, same order). Human sequences are used for EAAT3 which have one extra residue than the rat sequences employed in some experiments.

Transporter		Trans.	Exch.	Na^+ aff.(s)	Na^+ aff.	K^+ int.	Li^+ int.	Glu aff.*	Asp aff.*
EAAT1	Wild-Type[129,130]	Yes	Yes	20 mM		Yes	Yes	24 μM	16 μM
	T131A[38]		Yes	N/S				1Red	
	S363R/R477M[129]	No	Yes	0Inc		No		2Inc	2Inc
EAAT2	Wild-Type[28,34,131,132]	Yes	Yes	24 mM		Yes	No	5.3 μM	2.8 μM†
	Y403F[132]	No	Yes	1Inc		No	Yes	WT	WT†
	E404D[28]	No	Yes	WT		No			
	S440G[34]	Yes	Yes			Yes	Yes	WT	WT†
EAAT3	Wild-Type[30,133,134]	Yes	Yes	8 mM	120 mM	Yes	Yes	14 μM	14 μM
	T101A[31]	Yes	Yes	0Red	N/S	Yes		2Red	
	H296Q[135]	Yes	Yes	WT		Yes		WT	
	N366Q[92]	Yes	Yes	N/S		Yes	No	3Red	2Red
	D368E[92]	Yes	Yes	WT		Yes	No	None	1Red
	D368N[30]	No	Yes	0Red	N/S	No		2Red	
	N366D/D368N[92]	No	Yes			No	No	1Red	0Red
	T370S[96,134,136]	Yes	Yes	N/S	0Red	Yes	No	2Red	WT
	E374Q[133,137]	Yes	Yes	WT	WT	Yes		WT	
	D440N[133]	No	No	1Red	WT			2Red	1Red
	D440E[138]	Yes	Yes	1Inc		Yes	Yes	1Inc	1Inc
	D444C[33]	No	Yes‡					WT	WT
	D444E[33]	No	Yes	0Inc				WT	WT
	R445E[33]	Yes	Yes						
	R445S[139]	Yes	Yes	WT			Yes	1Inc	1Inc†
	R447C[32]	No	No	Becomes a Na^+ dependent serine/cysteine exchanger					
	N451S[136]	Yes	Yes	N/S		Yes	No	2Red	1Red
	D455N[30,133,140,141]	No	Yes	WT	WT	Yes	Yes	0Inc	0Inc
	D455S[140]	No	Yes	0Red		No	Yes	WT	0Inc
	D455E[140]	No	No						
	D455A[140–142]	No	No		WT				

*All glutamate affinities measured at [Na^+] between 100–150 mM.

†Result for D-aspartate.

‡Enables succinate exchange, unlike wild-type.

Table 6.2: Glt_{Ph} residues involved in the coordination of the ligands and their equivalents in EAAT1, EAAT2 and EAAT3. Residues from the mutagenesis experiments in Table 6.1 are also included in the table. Those residues that are not conserved between Glt_{Ph} and EAATs are indicated with boldface. Human sequences are used for EAAT3.

Glt_{Ph}	Y89	T92	S93	**Q242**	**R276**	S277	S278	G306	T308
EAAT1	Y127	T130	T131	**H328**	**S363**	S364	S365	G394	T396
EAAT2	Y124	T127	T128	**H326**	**A361**	S362	S363	G392	T394
EAAT3	Y98	T101	T102	**H296**	**S331**	S332	S333	G362	T364
Glt_{Ph}	N310	D312	T314	Y317	**Q318**	S349	I350	**T352**	**G354**
EAAT1	N398	D400	T402	Y405	**E406**	S437	I438	**A440**	**G442**
EAAT2	N396	D398	T400	Y403	**E404**	S435	I436	**A438**	**S440**
EAAT3	N366	D368	T370	Y373	**E374**	S405	I406	**A408**	**G410**
Glt_{Ph}	V355	G359	D390	D394	**M395**	R397	T398	N401	D405
EAAT1	I443	G447	D472	D476	**R477**	R479	T480	N483	D487
EAAT2	I441	G445	D470	D474	**R475**	R477	T478	N481	D485
EAAT3	V411	G415	D440	D444	**R445**	R447	T448	N451	D455

Table 6.3: QMEAN scores and DFIRE energy values for the Glt_{Ph} crystal structures, and the EAAT3 models constructed in this work. The QMEAN score goes from 0 to 1, 1 being the best possible model. Lower DFIRE energy values indicate better quality structures.

Structure	QMEAN score	DFIRE Energy
Glt_{Ph} (2NWX)	0.524	−629.6
EAAT3 (outward)	0.489	−605.0
Glt_{Ph} (3KBC)	0.557	−628.3
EAAT3 (inward)	0.540	−605.6

6.2 Results and Discussion

6.2.1 Model Description

We have used the crystal structures of Glt_{Ph} in the outward- (PDB ID 2NWX) and inward-facing states (PDB ID 3KBC) to construct two homology models for the EAAT3 transporter. Both conformations have the substrate binding site occluded from the solvent, and have an Asp and two Na^+ ions bound (Na1 and Na2). To build the models, we have used the program MODELLER[74] and the alignment from Ref.[14] (Figure 6A.1 in the Appendix). The templates had Asp, Na1, Na2 and Na3 bound, with this last Na^+ added at the site as described in Ref.[38]. We have created the models including the ligands in the alignment in order to obtain EAAT3 structures in the fully-bound states. We have created 20 models for each conformation and assessed the quality of the structures by using the Swiss Model Server[79,80]. After choosing the best models, we have constructed a trimer for each conformation by superposing the EAAT3 monomer to chains A, B and C of Glt_{Ph}. These two models have been employed in the MD simulations.

EAAT3 has a sequence of around 50 residues close to the trimerization domain, called 4B-4C loop, which is not present in Glt_{Ph}. In order to ensure robustness of EAAT3 trimers, we had to exclude this region from the models. Experiments have shown that cysteine substitution in the residues from this loop has little effect on the transporter properties[146]. However, studies of the LacY transporter indicate that this does not necessarily mean that this segment is not important for transport, as it can be involved in conformational changes happening during the transport cycle[147]. To make sure that the truncation of this segment has little effect on our results, we have also built outward and inward models with the inclusion of the 4B-4C loop. This is achieved by placing the 4B-4C loop in the position suggested from FRET studies[146], and adding restraints in MODELLER (see Figure 6A.2 in the Appendix). It can be seen from Figure 6A.2 that this loop is not close to the conformational changes we investigate in this study or to the bound ligands, which corroborates our approach. While this full model of EAAT3 is sufficient to give an idea about the distance between the 4B-4C loop and the binding pocket, there are large uncertainties in the secondary structure of the 4B-4C loop, which may cause simulation artifacts. Therefore, we prefer to use the truncated model without the 4B-4C loop in the MD simulations. Regarding the other segments of EAAT3, a previous FRET study has shown that their relative positions have a strong correlation with the Glt_{Ph} crystal structure, which indicates that the structures are well conserved between the two transporters.[148]

6.2.2 Evaluation of the Models

We have assessed the quality of our EAAT3 homology models using three different methods. QMEAN (Qualitative Model Energy ANalysis) is a composite scoring

function ranging from 0 to 1, which describes the major geometrical aspects of protein structures[79]. DFIRE (Distance-scaled, Finite Ideal-gas REference) is a potential based on a database of non-homologous proteins[80]. DOPE (Discrete Optimized Protein Energy) is an atomic distance-dependent statistical potential calculated from a sample of native structures[74]. The first two are available on the Swiss Model Server, and the last one is built into the program MODELLER. In Table 6.3 we show the QMEAN scores and DFIRE energies for the original Glt_{Ph} crystal structures, as well as our EAAT3 models in the outward and inward-facing states. It is seen that the QMEAN and DFIRE values are comparable between the crystal structures and the EAAT3 models, which shows that the EAAT3 models are reliable, and in principle provide a good option to investigate EAATs. The DOPE energy profile for the outward-facing EAAT3 model is compared with that of the 2NWX crystal structure in Figure 6A.3A in the Appendix, and the same is done for the inward-facing EAAT3 model and the 3KBC crystal structure in Figure 6A.3B in the Appendix. The profiles are very similar, especially in the regions that form the glutamate binding site, which range from residues 260 to 410. We have also compared the Ramachandran plots for the templates and the models, which agree very well with few residues in the forbidden regions in all cases (Figure 6A.4 in the Appendix). These results confirm that, even though the homology between EAAT3 and Glt_{Ph} is not very high, we can still build good models and hopefully gain insights into the mammalian glutamate transporters using the structural information from the archaeal homolog.

6.2.3 Fully Bound EAAT3

We first discuss the stability of the ligands in the outward and inward fully-bound models of EAAT3 with the EC-gate and IC-gates closed. To check the hypothesis that E374 is the proton carrier in EAAT3, we have prepared two systems for the outward-facing state of EAAT3: one with the side chain of E374 protonated and the other with a deprotonated E374 side chain in all three monomers. All of these systems have been equilibrated for 20 ns initially. In order to show that our models are stable for longer periods, we have performed an additional 40 ns of MD simulations for the outward and inward fully bound states with protonated E374. The RMSDs of the protein backbone for these models, obtained from the 60 ns MD data, are shown in Figure 6A.5 in the Appendix. Both RMSDs show no major changes after the first 10 ns. More importantly, the coordination of the ligands remain intact, exhibiting no perceptible changes between 10–20 ns and 20–60 ns of data collection. These results indicate that our systems are stable in the vicinity of the outward and inward states, and 20 ns appears to be sufficient for their equilibration.

Our simulations with the protonated E374 residue show that Glu is stable, remaining bound to the transporter with the EC-gate closed and the coordination shown in Table 6.4 throughout the 60 ns of MD simulations (Figure 6.2A). We have applied the same procedure in the inward-facing state of EAAT3, and ob-

Table 6.4: List of the EAAT3 residues coordinating the Glu substrate atoms. The average N–O and O–O distances (in Å) are obtained from 5 ns of MD simulations in the closed conformation of the outward and inward-facing states of the transporter. Bare O and N atoms in parenthesis refer to the backbone atoms and the others with subscripts refer to the side chain atoms.

Helix-Residue	Glu-atom	Outward EAAT3	Inward EAAT3
HP1 - S331 (O)	α-N	2.8 ± 0.2	3.0 ± 0.2
HP1 - S333 (N)	α-O_1	2.7 ± 0.1	2.7 ± 0.1
HP1 - S333 (OH)	α-O_2	2.8 ± 0.1	2.7 ± 0.1
TM7 - T370 (OH)	β-O_2	2.7 ± 0.1	2.7 ± 0.1
HP2 - V411 (O)	α-N	2.9 ± 0.2	3.0 ± 0.3
HP2 - G415 (N)	β-O_1	2.7 ± 0.1	2.8 ± 0.1
TM8 - D444 (O_1)	α-N	2.7 ± 0.1	2.7 ± 0.1
TM8 - R447 (N_1)	β-O_1	2.7 ± 0.1	2.7 ± 0.1
TM8 - R447 (N_2)	β-O_2	2.8 ± 0.2	2.7 ± 0.1
TM8 - T448 (OH)	α-N	3.0 ± 0.2	3.1 ± 0.2
TM8 - N451 (N_δ)	α-O_2	2.8 ± 0.1	2.8 ± 0.1

tained very similar results for the coordination of Glu when the E374 side-chain is protonated (Table 6.4). In contrast, if we have a deprotonated E374 side-chain, the substrate shows instability and loses most of its contacts with the transporter (Figure 6.2B). In the closed state, the substrate and the E374 side-chain carboxyl groups are separated by around 7 Å, and there are no water molecules between them so the dielectric screening is expected to be small in this region. Thus the electrostatic repulsion between these two groups is likely to destabilize Glu in the binding site, making the protonation of E374 necessary for Glu binding. This is consistent with the experimental observations indicating that protonation of E374 precedes Glu binding[137,145]. We have further tested this possibility by performing 20 ns of simulations of the Glt_{Ph} Q318E mutant in two systems previously equilibrated in a recent study[123]. One of them had Asp, Na1, Na2 and Na3 bound in the outward-facing closed state, and the other one had Asp, Na1 and Na3 bound in the outward-facing open state. In the first case we don't see any changes, except for the partial hydration of the E318 residue that takes place in one of the chains after a few nanoseconds. In the second, however, we see that the Asp substrate becomes unstable and loses most of its contacts with the protein, and Asp is released back to the solvent in two of the three chains of Glt_{Ph}. A movie is included in the Supporting Info of Ref.[161], showing the release of the Asp substrate in one of the chains that occurs between 5 and 15 ns.

Regarding the ions in our EAAT3 models, Na1 and Na3 are very stable and

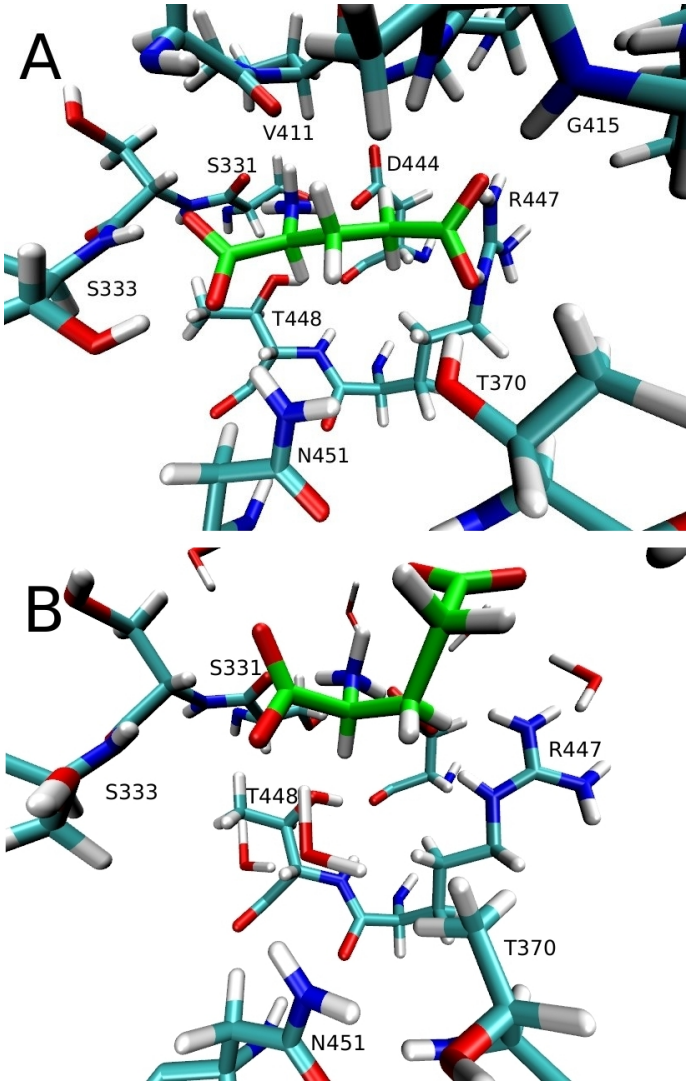

Figure 6.2: Binding of glutamate to EAAT3 with a protonated vs deprotonated E374 side chain. (A) Glutamate substrate (in green) bound to EAAT3 in the closed outward state with a protonated E374 side chain. (B) The same but with a deprotonated E374 side chain. Glutamate is stable in A for 20 ns, but becomes unstable in B, losing most of the contacts after 10 ns of simulations.

have very similar coordinations compared to those in Glt_{Ph}[123,144] (Table 6.5). This is expected because all of the residues in the coordination shell of both ligands are conserved between EAAT3 and Glt_{Ph}, except for T102 which is a serine (S93) in Glt_{Ph}. The Na2 ion, on the the other hand, is unstable and spontaneously unbinds in two of the chains in the outward-facing state and one of the chains in the inward-facing state. The ones that remain bound are coordinated by the residues shown in Table 6.5. Because the Na2 site in the crystal structure has contacts with the backbone carbonyl oxygens, it is difficult to assess it through mutagenesis experiments. Previous MD studies of Glt_{Ph} have shown that this site is unstable[149], which is consistent with our observations. It is possible that the Na2 site is not conserved in EAATs. Mutagenesis experiments have shown that the T370S[96,136] and the N451S[136] mutations seem to affect the Na^+ affinity and selectivity in EAAT3 (Table 6.1), and a site having the coordination of these residues and the substrate has been suggested[96]. Since Na2 is likely to be the only Na^+ ion to bind to EAAT3 after the substrate, this site would make a good candidate for the Na2 site if the original site from the Glt_{Ph} crystal structure is shown not to be conserved in EAATs. More experiments and simulations are needed to clarify this issue.

6.2.4 Opening of the EC- and IC-Gates

In order to observe the gating mechanism in EAAT3, we have performed two 60 ns simulations for each model: one of them had no ligands bound, and the other had only the Na1 and Na3 ions bound (Figure 6.3). The opening of the EC- and IC-gates is monitored using the S332(Cα)–V411(Cα) distance, which has been used in previous studies of the Glt_{Ph} transporter[36,144]. These two residues are located at the tips of the HP1 and HP2 hairpins, respectively, and hence provide a good measure for the opening of the EC-gate. In the simulations of the fully-bound state this distance remains around 5 Å, meaning that the gates are closed in this configuration. Once the Glu substrate is removed, both the EC- and the IC-gates show opening in all of the chains regardless of whether the Na1 and Na3 ions are present or not (Figure 6.3). Presence of the Na^+ ions appears to reduce the fluctuations in the gate motions but does not have much effect on the opening of the gates. In both Figures 6.3B and 6.3D, the Na1 and Na3 ions are not affected by the gate opening despite the flooding of the binding site by water molecules, and remain bound with the coordination shown in Table 6.5. This shows that these sites are very stable in EAAT3 even after 60 ns of simulations in the absence of the substrate, so they are likely to be the two Na^+ ions that are proposed to bind to EAATs before glutamate binds[31].

To assess in more detail the conformational changes brought by the gate opening in the outward and inward-facing states, we choose the chains that show the largest opening in the absence of ligands in both cases, which are chain A and chain B, respectively, in Figures 6.3A and 6.3C. In Figures 6.4A and 6.4B, we show the overlap of these two structures with their respective closed states, and in Figure 6.4C

Table 6.5: List of the EAAT3 residues coordinating the three Na^+ ions. The average Na^+–O distances (in Å) are obtained from 5 ns MD simulations of the closed state in the outward and inward-facing conformations of EAAT3. Nature of the coordinating oxygens are indicated in parenthesis.

	Helix-Residue	Outward EAAT3	Inward EAAT3
Na1	TM7 - G362 (O)	2.4 ± 0.1	2.7 ± 0.3
	TM7 - N366 (O)	2.3 ± 0.1	2.3 ± 0.1
	TM8 - N451 (O)	3.4 ± 0.5	2.5 ± 0.2
	TM8 - D455 (O_1)	2.2 ± 0.1	2.2 ± 0.1
	TM8 - D455 (O_2)	2.2 ± 0.1	2.2 ± 0.1
	H_2O	2.3 ± 0.1	2.3 ± 0.1
Na3	TM3 - Y98 (O)	2.3 ± 0.1	2.3 ± 0.1
	TM3 - T101 (OH)	2.4 ± 0.1	2.4 ± 0.1
	TM3 - T102 (OH)	2.4 ± 0.2	2.4 ± 0.2
	TM7 - N366 (O_δ)	2.2 ± 0.1	2.2 ± 0.1
	TM7 - D368 (O_1)	2.1 ± 0.1	2.2 ± 0.1
	TM7 - D368 (O_2)	2.7 ± 0.4	2.6 ± 0.3
Na2	TM7 - T364 (OH)	2.3 ± 0.1	2.6 ± 0.4
	HP2 - I406 (O)	4.2 ± 0.4	2.4 ± 0.3
	HP2 - A408 (O)	2.3 ± 0.1	2.4 ± 0.3
	H_2O	2.4 ± 0.2	2.3 ± 0.2

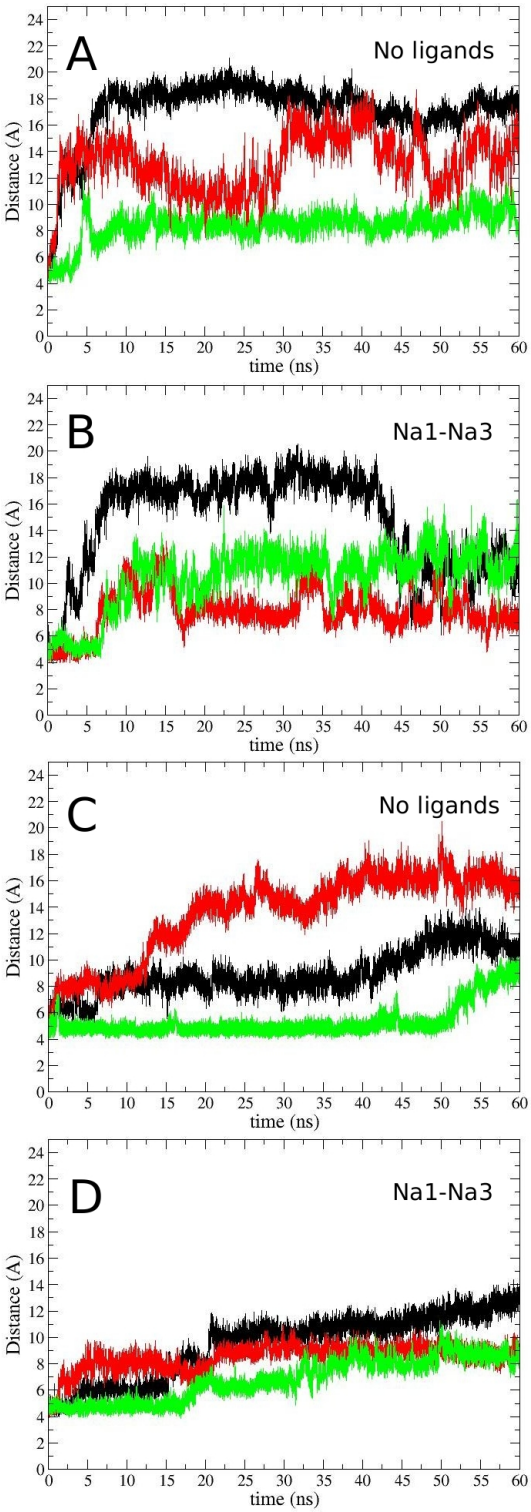

Figure 6.3: Gating motions in the outward- and inward-facing conformations of EAAT3. Time evolution of the S332(Cα)–V411(Cα) distance in the outward-facing conformation with (A) no ligands bound to the transporter, and (B) only the Na1 and Na3 ions bound to the transporter. The same for the inward-facing conformation with (C) no ligands bound to the transporter, and (D) only the Na1 and Na3 ions bound to the transporter. The color code for the subunits are A (black), B (red), and C (green).

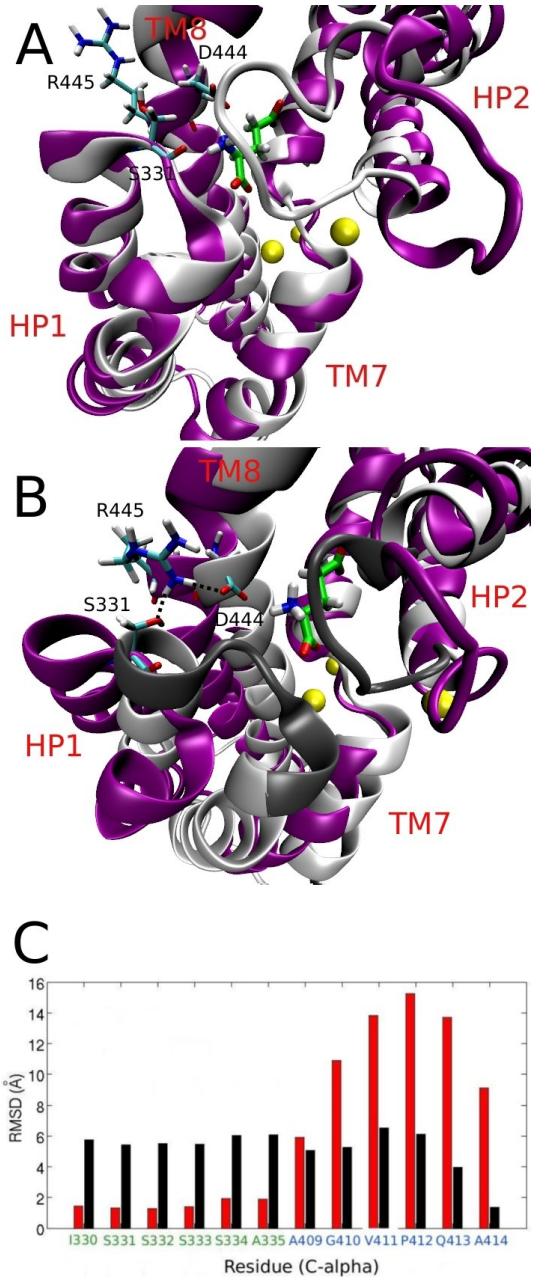

Figure 6.4: Comparison of gate opening in the outward- and inward-facing conformations of EAAT3. (A) Overlap between the closed (grey) and the open (purple) states in the outward EAAT3. Glutamate and three Na^+ ions bound to the transporter, as well as the S331, D444 and R445 residues are indicated. (B) The same for the inward-facing state, showing a large movement of HP1, HP2 and TM8. The contacts made by the insertion of the R445 side-chain between HP1 and TM8 are also shown. The dark shade of grey in the closed state indicates the position of the residues from part C below. (C) Comparison of the IC-gate opening (black) with the EC-gate opening (red). The RMSD of the Cα atoms at the tips of HP1 and HP2 is obtained by aligning the open and closed states and calculating the RMSD from 2 ns of MD simulations.

we show the average RMSD displacement of the Cα atoms of six residues located at the tips of each of the two hairpins. These are calculated by aligning each frame and calculating the RMSD of each atom between the closed and open states during 2 ns of MD simulations. The opening observed in the outward facing EAAT3 looks very similar to Glt_{Ph}, with a large movement of HP2 and a small displacement of the HP1, TM7 and TM8 segments. The inward state, on the other hand, shows movements of HP1, HP2 and TM8, with a much larger opening of HP1 compared to our previous study of Glt_{Ph}[144]. From Figure 6.4B, we see that this is caused by the insertion of the R445 side-chain in between the residues S331 and D444, making contacts with them. This shifts the HP1 and TM8 segments apart, causing a larger exposure of the binding site. This event is observed at around 40 ns in the simulation of Figure 6.3C, but the contacts are kept and the transporter maintains the large opening in the remainder of that simulation. Interestingly, Glt_{Ph} has a methionine in the position equivalent to R445, and an arginine replaces the serine in position 331.

It is instructive to compare our results to those obtained from metadynamics simulations on Glt_{Ph} by Grazioso et al.[126]. In both cases, opening of the gate in the outward-facing state shows a large movement of HP2, and a small displacement of the neighbouring segments. Similarly, during the opening of the gate in the inward-facing state, both studies show a movement of not only HP1 and HP2, but also of the nearby segments TM7 and TM8. A salt bridge between R276 and D394 is observed in the inward open state of Glt_{Ph}, which is similar to the R445–D444 salt bridge we see in our simulations. We also compare our results for EAAT3 to the microsecond MD simulations of Zomot and Bahar[143] in the inward-facing state of Glt_{Ph}. In Glt_{Ph}, a salt bridge between D394 and R276 prevents HP1 from opening further, but it also attracts the Glu substrate away from the binding site, facilitating its release. In EAAT3 there is still a salt bridge in this position, but it has the opposite effect on the gating; it allows the transporter to open further and stabilizes an open conformation through the contact between R445 and D444. Therefore, the switching of this arginine from position 331 to 445 in EAAT3—which is conserved in all EAATs—has a two-fold effect: first, it creates a more stable and open conformation of the transporter in the inward state, which facilitates the release of the larger Glu substrate; and second, it conserves the salt bridge (R445–D444) in the vicinity of the binding site that has been shown to help the release of the substrate. Even though this is an intermediate state followed by further conformational changes in the transporter, these are significant features of inward gating that are worth emphasizing.

6.2.5 The D455N Mutation

As mentioned in the Introduction, there is an apparent disagreement between the crystal structure and the mutagenesis experiments concerning the Na1 binding site, raising the possibility that the D455 residue is protonated throughout the trans-

port cycle. To address this issue, we have prepared four simulation systems in the outward-facing state of EAAT3: two systems with the D455N mutation and with either Na1 and Na3 bound or with only Na1 bound, and another two with a protonated D455 residue and the same Na1/Na3 occupancy as above. For the Na1 and Na3 bound systems, we use the last frame of the simulation in Figure 6.3B. The Na1 bound system is prepared by removing Na3 from this system and equilibrating it for a further 10 ns. We perform a 2 ns alchemical transformation in the D455 side-chain to an asparagine or to a protonated aspartate, and at the same time gradually release restraints applied to the Na1 ion. After the mutation/protonation step, we equilibrate the system for a further 8 ns, monitoring the distance between the side-chain carboxyl/carboxamide carbon atom and the Na1 ion.

The results for the D455N mutation and the D455 protonation in the presence of Na3 are displayed in Figure 6.5A. It is seen that Na1 is released to the solvent in all of the chains after the protonation of D455, and in two out of three chains after the D455N mutation. Therefore, it is unlikely that EAAT3 can support the binding of both Na1 and Na3 if the D455 side-chain is protonated. When we perform the same simulations in the absence of the Na3 ion, Na1 is found to be stable in all the chains for the whole 10 ns (not shown). In this case, we observe that the Na1 ion finds a new coordination shell involving the carboxyl oxygens of D368, one carboxamide/carboxyl oxygen of N455/D455 and the carbonyl oxygens of residues N366 and N451 (Figure 6.5B). This site involves some of the residues that coordinate the original Na1 ion (Table 6.5), and is located about 3 Å from the Na1 site. In the absence of the Glu substrate, the Na^+ affinity is measured via the [Na^+]-dependent anion leak current, which is unaffected by the D455N mutation in EAAT3. This may be explained by assuming that this current is coupled to the binding of two Na^+ ions in the wild-type transporter, but only to one Na^+ ion after the D455N mutation. It is important to mention that, even though the D455N mutation does not affect the Na^+ or Glu affinity in EAAT3, it does lock this transporter in the exchange mode (Table 6.1), meaning that one of the steps in the transport cycle is impaired by this mutation.

6.2.6 The Potassium Binding Site

In contrast to Glt_{Ph}, EAATs counter-transport one K^+ ion back to the extracellular medium at each cycle. Binding of K^+ is essential for the reorientation of the transporter back to the outward-facing state after the release of all the ligands into the intracellular medium. It has been suggested that mutants that are locked in the exchange mode cannot interact with K^+, and therefore are not able to carry out the second half of the transport cycle. This is the case for a few mutations close to the Na1 and Na3 binding sites, such as the D368N, D455A, D455S mutations and the N366D/D368N double mutation in EAAT3 (Table 6.1). The D455N mutation is an exception where the transporter is still locked in the exchange mode, but it could interact with K^+ albeit with a reduced affinity compared to wild type[141]. Therefore,

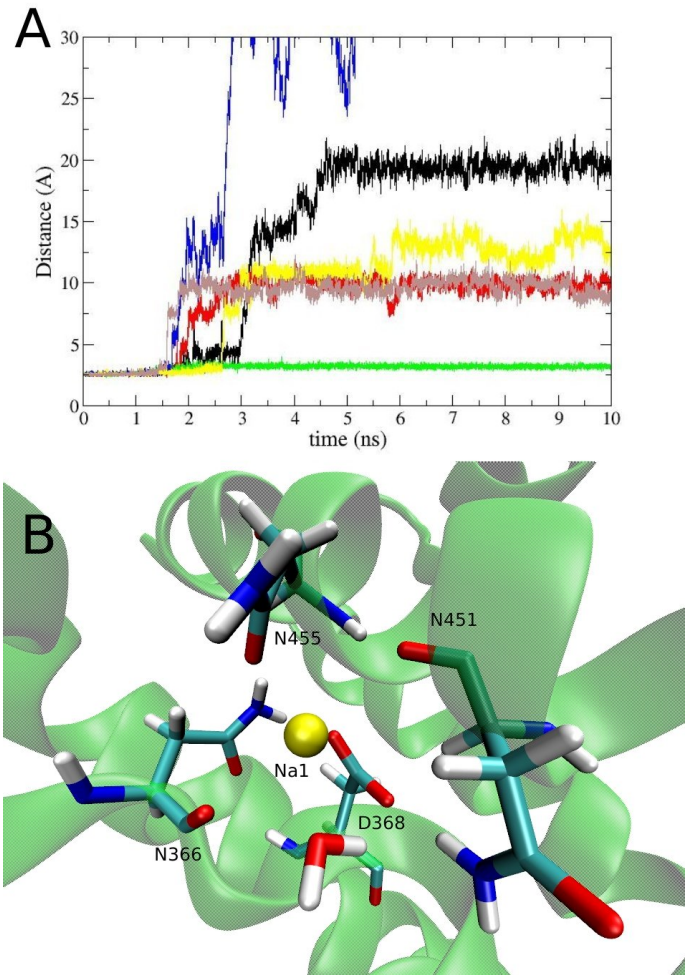

Figure 6.5: Effect of the D455N mutation on sodium binding. (A) The distance between the Na1 ion and the side-chain carboxyl/carboxamide carbon atom of the residue D455 during its mutation from aspartate to an asparagine (black, red and green), and to a protonated aspartate (blue, yellow and brown) in the presence of the Na3 ion. In the first case, Na1 leaves the binding site in 2 out of 3 chains, and in the second case, Na1 leaves the binding site in all chains. (B) The new coordination of Na1 when the D455N mutation is performed in the absence of the Na3 ion. This site has the coordination of the carbonyl oxygens of N366 and N451, as well as the side chain oxygens of D368 and N455.

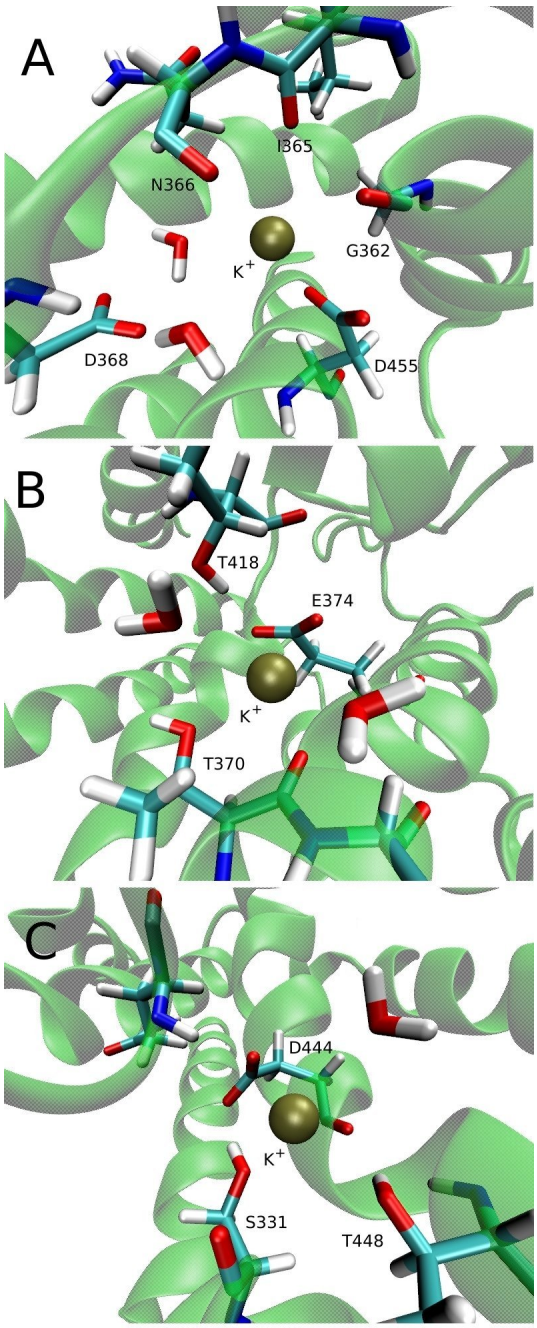

Figure 6.6: Three potential sites for binding of a K^+ ion. (A) Site 1 is obtained by replacing Na1 by a K^+ ion and equilibrating the system for 10 ns in the absence of other ligands (B) Site 2 is obtained by restraining K^+ close to the E374 carboxyl group and equilibrating the system for 10 ns with no other ligands bound. (C) Site 3 is obtained by placing a K^+ ion near glutamate α-N atom and equilibrating the system for 10 ns in the absence of other ligands.

this region is a good candidate for the K^+ binding site in EAAT3, which we call site 1. To obtain this site, we remove the Na3 ion from the last state of the simulation in Figure 6.3D—where EAAT3 is in the Na1-Na3 bound inward-facing state—replace the Na^+ ion at Na1 to K^+ in all of the three chains and equilibrate the system for a further 50 ns. The final coordination of K^+ is shown in Figure 6.6A, which is similar in all of the three subunits. Site 1 makes contacts with both side-chain oxygens of residue D455, the backbone atoms of residues G362, I365, and N366, and two water molecules (see Table 6.6 for the contact distances). The two water molecules are very stable as they are kept in place by the D368 side-chain, which is also close to the K^+ ion. This region has a strong negative potential due to the presence of two negatively-charged side chains, and therefore is likely to harbour a cation binding site.

Table 6.6: EAAT3 residues coordinating the K^+ ion at sites 1, 2, and 3. Notation is the same as in Table 6.5.

	Helix-Residue	Distance
site 1	TM7 - G362 (O)	2.8 ± 0.2
	TM7 - I365 (O)	2.9 ± 0.2
	TM7 - N366 (O)	2.7 ± 0.1
	TM8 - D455 (O_1)	2.6 ± 0.1
	TM8 - D455 (O_2)	2.7 ± 0.2
	H_2O (1)	2.7 ± 0.1
	H_2O (2)	2.9 ± 0.3
site 2	TM7 - T370 (OH)	3.0 ± 0.3
	TM7 - T370 (O)	2.7 ± 0.2
	TM7 - E374 (O_1)	2.8 ± 0.2
	TM7 - E374 (O_2)	2.7 ± 0.2
	H_2O (1)	2.8 ± 0.2
	H_2O (2)	2.9 ± 0.3
site 3	HP1 - S331 (O)	2.7 ± 0.2
	HP1 - S331 (OH)	3.1 ± 0.4
	TM8 - D444 (O)	3.2 ± 0.4
	TM8 - D444 (O_1)	2.8 ± 0.3
	TM8 - D444 (O_2)	3.4 ± 0.4
	TM8 - T448 (OH)	2.8 ± 0.1
	H_2O	2.9 ± 0.3

In EAAT2, the Y403F and E404D mutations also lock this transporter in the exchange mode (Table 6.1). The same happens for the E374Q mutation in EAAT3,

even though in this case the transporter is still able to interact with K^+ in the homoexchange mode[137]. This raises the possibility that the K^+ binding site might overlap with the proton binding site, which we call site 2. To obtain this site, we restrain a K^+ ion close to the deprotonated E374 side-chain and equilibrate the system in the absence of other ligands for 50 ns, gradually releasing the restraints applied to the K^+ ion. The K^+ ion remained in a similar position in two of the three chains and was unstable . The one that has the tightest coordination is chosen for site 2 (Figure 6.6B), and will be used in free energy calculations. The contacts in site 2 are the two side-chain oxygens of E374, and both the backbone oxygen and the side-chain hydroxyl group of T370 (Table 6.6). While the T418 residue does not directly interact with K^+, it remains close to this ion and might have an indirect effect on the stability of K^+ in site 2.

The last site we consider (to be called site 3) overlaps with the positively-charged substrate α-amino group and was proposed from electrostatic mapping calculations of a homology model of EAAT3[95]. This site has also been suggested in a recent study of the novel mosquito dicarboxylate transporter CuqDCT, which has an asparagine at the position equivalent to the D444 residue in EAAT3[150]. Mutations of this residue or its equivalent affect K^+ counter-transport in both EAAT3 and CuqDCT, so this site is also a good candidate for binding of a K^+ ion. Here we obtain this site by removing the substrate and placing a K^+ ion in the same position as its α-amino N atom. The system is then equilibrated for 50 ns in the absence of other ligands. The same coordination is obtained for site 3 in all chains of EAAT3 (Figure 6.6C), which is also very similar to the one suggested in Ref. [95]. Site 3 has contacts with an oxygen of the D444 side-chain, the hydroxyl groups of S331 and T448, and the backbone atoms of D444 and S331 (Table 6.6). Even though site 3 has many coordinating residues, it is also close to the positively-charged R445 and R447 residues, which will reduce its affinity for a K^+ ion.

To find out which site is more likely to be the binding site for K^+, we calculate the affinity and the selectivity of the three sites for K^+, using FEP and TI calculations as described in Methods. The FEP values show up to 3 kcal/mol difference between the forward and backward calculations whereas hysteresis in the TI values remains within 1 kcal/mol. This is because TI allows for longer simulations of fewer windows, which increases sampling and reduces the observed hysteresis. Therefore, we report only the TI results here but stress that very similar results are obtained from the averages of the forward and backward FEP values. The results are summarized in Table 6.7 and evidence for the convergence of the TI calculations is provided in Figure 6A.6 in the Appendix. Site 1 has by far the highest affinity of all sites (-20.5 kcal/mol), with a low selectivity margin for Na^+; the binding free energy decreases by only 0.5 kcal/mol when we go from K^+ to Na^+. Site 2 has an affinity of -9.5 kcal/mol and is highly selective for Na^+ with a selectivity free energy of 3.9 kcal/mol. Site 3 has the lowest affinity (-6.5 kcal/mol) but this site is highly selective for K^+, having -3.1 kcal/mol lower binding free energy compared to Na^+ in agreement with previous valence studies in an EAAT3 model[95].

Table 6.7: Binding free energies for K^+ and K^+/Na^+ selectivity free energies at the three sites proposed for the inward-facing EAAT3. The forward and the backward calculations of the interaction energy are listed in the second and third columns, respectively, and their average in the fourth column. The total binding free energy that includes the translational free energy difference (fifth column) is given in the last column. Errors are estimated from block data analysis using 100 ps windows. All energies are in kcal/mol.

Ion	$-\Delta G_{\text{int}}^{\text{for}}$	$\Delta G_{\text{int}}^{\text{back}}$	ΔG_{int}	ΔG_{tr}	ΔG_{b}
K^+-site 1	-25.6 ± 1.2	-24.2 ± 1.0	-24.9 ± 1.1	4.4	-20.5 ± 1.1
K^+-site 2	-13.7 ± 1.1	-14.0 ± 1.3	-13.8 ± 1.2	4.3	-9.5 ± 1.2
K^+-site 3	-10.6 ± 0.8	-11.0 ± 0.8	-10.8 ± 0.8	4.3	-6.5 ± 0.8
K^+/Na^+-site 1	0.4 ± 0.3	0.3 ± 0.5	0.3 ± 0.4	0.2	0.5 ± 0.4
K^+/Na^+-site 2	3.7 ± 0.5	4.0 ± 0.4	3.8 ± 0.4	0.1	3.9 ± 0.4
K^+/Na^+-site 3	-3.1 ± 0.7	-3.5 ± 0.5	-3.3 ± 0.6	0.2	-3.1 ± 0.4

The three K^+ sites tested here are all based on a model which was initially built from the Glt_{Ph} structure with all the ligands bound. In a recent study, the crystal structure of Glt_{Tk}—a close homolog of Glt_{Ph}—was resolved in a substrate-free state and in the absence of Na^+ ions[151]. This structure shows a few conformational changes compared to the holo state of Glt_{Ph}, in particular, the region around TM7 and TM8 which is close to our K^+ site 1. Comparison of this region in Glt_{Tk} and our apo state of EAAT3 shows that such conformational changes are well reproduced in our model (Figure 6A.7 in the Appendix). We stress that Glt_{Tk}, like Glt_{Ph}, does not interact with K^+ ions, so its crystal structure cannot be used directly in search of K^+ binding sites. For example, the K^+ binding site 2 is not formed because Glt_{Tk} has a glutamine at the position equivalent to E374 in EAAT3. Similarly, formation of the K^+ site 3 is compromised in Glt_{Tk} because D398 (D444 in EAAT3) interacts with R401.

The turnover rate of Glt_{Ph} is 0.29 min^{-1} at 30° C[35], and we have suggested in a previous study of Glt_{Ph} that the release of the Na3 ion is the rate-limiting step in the transport cycle due to its high affinity[144]. The speed of transport in EAATs is much higher, with turnover rates of EAAT2 and EAAT3 at room temperature being 15 s^{-1} and 90 s^{-1}, respectively. In this case, experimental evidence from EAAT2 and EAAT3 points to the K^+-dependent relocation step as rate-limiting in the transport cycle[131,152]. A K^+ ion at site 1, which is the site with the highest affinity, would be consistent with a model in which this ion would replace the Na^+ ions during the release of the ligands, therefore making this process faster than in Glt_{Ph}. This K^+ ion could again be replaced by Na^+ ions once the transporter goes back to the outward facing state, restarting the cycle. This last step is consistent with the MD simulations of Glt_{Ph}, which show that the first Na^+ ion to bind to the transporter occupies a transient state, equivalent to the K^+-site 1, before moving to the Na3

site[97]. It is important to note that this model is based only on the thermodynamics and ion binding, and does not take into account conformational changes of the transporter, which may indicate a different picture. Such conformational changes happen in time scales that are not accessible by all-atom MD simulations, and therefore, are difficult to analyze.

After the completion of this work, two FRET studies were published on the outward to inward transitions of Glt_{Ph} and the time scales involved in them. In one of them, the translocation of the binding pocket from the outward state to the inward state is observed in the order of minutes, suggesting that this transition is the rate-limiting step in the transport cycle[153]. The second study asserts that the transporter is always in a very dynamic state with the transitions between the outward and inward states occurring in the order of seconds regardless of which ligands are bound to it, and furthermore no long dwell times are observed[154]. Interestingly, this last study also shows that Glt_{Ph} cannot perform the full outward to inward translocation (or vice-versa) in the presence of the bound Na^+ ions—it can only alternate between an intermediate and an inward/outward conformation. The results of Ref.[154] are more consistent with our model, but further experiments and simulations are needed to clarify this issue.

6.3 Conclusions

In this work we have performed MD simulations on two homology models of the mammalian glutamate transporter EAAT3—one in the outward- and the other in the inward-facing conformation—based on the crystal structures of the archaeal homolog Glt_{Ph}. We have evaluated the quality of the EAAT3 models using three different methods and shown the reliability of the structures especially in the region of the substrate binding. We have examined the coordination of the ligands in EAAT3, and observed that Glu is stable in the binding site only if the residue E374 is protonated. This is in good agreement with numerous experiments that point to this residue as the proton carrier in EAATs. Regarding the opening of the EC- and IC-gates, our simulations show that the EC-gating in EAAT3 is very similar to that in Glt_{Ph} whereas the IC-gate in EAAT3 exhibits a larger opening compared to Glt_{Ph}. This difference can be traced to the change in the position of an arginine residue from the HP1 segment in Glt_{Ph} to the TM8 segment in EAATs, which creates a salt bridge that stabilizes the open state and also helps the release of the Glu substrate to the intracellular media. We have also investigated an apparent disagreement between the crystal structure and the mutagenesis experiments regarding the role of the D455 residue in Na^+ binding and the possible protonation of this residue. According to our simulations, when D455 is protonated or mutated to an asparagine, the substrate-free EAAT3 cannot support the binding of two Na^+ ions at the Na1 and Na3 positions. This disagreement can be explained if we assume that only one Na^+ ion binds to the D455N mutant before the substrate in contrast to two in the

wild-type transporter.

We have considered three potential sites for the counter-transported K^+ ion. Free energy calculations of K^+ binding at these sites reveal that a site which overlaps with the Na1 and Na3 binding sites has by far the highest affinity of the three. We suggest that binding of K^+ to this site, which does not happen in Glt_{Ph}, could be responsible for the much faster turnover rate in EAAT3 compared to the archaeal transporter. By looking at these results, we can speculate that the proton co-transport and the K^+ counter-transport might have evolved together and are closely related. As pointed out above, to speed up the transport rate, one needs to exchange the last Na^+ ion with K^+ in the inward-facing EAAT3 and counter-transport this K^+ ion. This extra positive charge in EAAT3 relative to Glt_{Ph} is neutralized by the residue E374 which is a glutamine in Glt_{Ph}. In the fully bound, outward-facing EAAT3, protonation of E374 results in the same charge content as in the fully bound, outward-facing Glt_{Ph}. Thus the charge content in the binding pocket of Glt_{Ph} is preserved in the EAAT3 transport cycle despite the co- and counter-transport of extra ions. To summarize, proton co-transport via E374 and its subsequent deprotonation enables K^+ counter-transport. This mechanism might also explain why mutations in the vicinity of the proposed protonation site, like E404D and Y403F in EAAT2 and E374Q in EAAT3, either lock the transporter in the exchange mode by not allowing the counter-transport of K^+ or significantly reduce this step in the transport cycle (Table 6.1).

6.4 Methods

6.4.1 Simulation Details

The simulation system is prepared using the software VMD[104]. We embed the EAAT3 trimer in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) phospholipid bilayer. Each monomer has four ligands bound: three Na^+ ions, namely Na1, Na2 and Na3, and Asp as the substrate. We then solvate the protein-membrane complex in a box of water molecules with 48 Na^+ ions and 45 Cl^- ions. The extra Na^+ ions are required to keep the system neutral, and their number changes depending on the ligands bound to the transporter. For the outward state there are a total of 244 lipid molecules and 15,953 water molecules, and for the inward state 239 lipid molecules and 15,871 water molecules, with a total of $\sim 100,000$ atoms in each simulation box. The two systems are equilibrated in three stages. In the first stage, the coordinates of the ligands and protein atoms are fixed and the system is equilibrated with 1 atm pressure coupling in all directions until the correct water and lipid densities are obtained. At this point, the x and y-dimensions of the simulation box are fixed, and pressure coupling is applied only in the z direction thereafter (typical dimensions of the simulation box for both cases are $112 \times 113 \times 76$ Å^3). In the second stage, the protein is gradually relaxed by reducing the restraints on the protein and ligand atoms in several steps during MD simulations lasting 2.4 ns.

At the end of this stage we are left only with a restraint of 0.1 kcal/mol/Å^2 in all the backbone atoms of the trimer. Following this, we perform a 2.5 ns alchemical transformation from the Asp substrate to Glu in the three chains, and equilibrate the system with Glu for another 5 ns by gradually releasing the backbone restraints.

All the MD simulations are performed using the NAMD package (version 2.9)[59] with the CHARMM36 force field[155]. The temperature is maintained at 300 K using Langevin damping with a coefficient of 5 ps^{-1}, and the pressure is kept at 1 atm using the Langevin piston method with a damping coefficient of 20 ps^{-1}[64]. Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. The Lennard-Jones (LJ) interactions are switched off between 10 and 12 Å using a smooth switching function. A time step of 2 fs is used in all MD simulations. Details of all the MD simulations we have performed in the outward and inward models of EAAT3 are summarized in Tables 6A.1 and 6A.2, respectively, giving the conditions and the total running time for each simulation.

6.4.2 Free Energy Calculations

We perform free energy calculations to determine the standard binding free energy of a K^+ ion at different sites in the protein, as well as the K^+/Na^+ selectivity. For this purpose, we use the free energy perturbation (FEP) and thermodynamic integration (TI) methods[52], as in our previous studies of Na^+ binding to Glt_{Ph}[123,144]. We note that the binding free energies calculated using the crystal structures cannot be compared directly to the experimental values due to the large conformational changes occurring in Glt_{Ph} following the binding/unbinding of the Na^+ ions[156]. Nevertheless they are still a useful tool for finding relative affinities of different sites, as we have done in locating the Na3 site Glt_{Ph}[38].

In the FEP method, the interval between $\lambda = 0$ and 1 is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average

$$\Delta G_i = -k_B T \ln\langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_B T]\rangle_{\lambda_i}. \tag{6.1}$$

where $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, with H_0 and H_1 representing the free-ligand and bound-ligand states. The free energy difference is obtained from the sum, $\Delta G_{\text{int}} = \sum_i \Delta G_i$. In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, is obtained at several λ values, and the free energy difference is calculated from the integral:

$$\Delta G_{\text{int}} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda. \tag{6.2}$$

Provided the integrand can be fitted well with a polynomial, Gaussian quadrature provides an efficient and accurate method for evaluation of such integrals because it allows for longer sampling of a smaller number of windows. The binding free

energy of a K^+ ion is calculated by alchemically transforming a a K^+ ion to a water molecule in the binding site, while at the same time the opposite transformation is performed in bulk. This approach is justified by the observation that, in the absence of a K^+ ion, a water molecule always occupies the binding site. For the Na^+/K^+ selectivity, we replace a K^+ ion with a Na^+ ion in the binding site, and vice-versa in bulk. In the FEP simulations we use 66 exponentially spaced windows with 20 ps of equilibration plus 40 ps of production for each window. In the TI method, we use seven windows starting from the FEP runs, and run 0.4 ns of equilibration plus 0.8 ns of data collection for each window. These parameters have been found to be optimal in previous studies of Glt_{Ph} [123,144]. We report here only the TI results because they show better convergence, but we always check for consistency between the two methods, as well as performing both the forward and backward transformations in each case. When calculating the standard binding free energies, we also have to account for the loss of translational entropy of the ligand upon binding. The translational free energy difference is determined using [118,119]

$$\Delta G_{\mathrm{tr}} = -k_{\mathrm{B}}T \ln \left[\frac{(2\pi e)^{3/2}\sigma_x\sigma_y\sigma_z}{V_0} \right], \tag{6.3}$$

where $V_0 = 1660$ Å^3, σ_x, σ_y, and σ_z, are the principal rms fluctuations of the center of mass of the ligand in the binding site, estimated from 5 ns of MD simulations of the bound ion with no restraints applied. The standard binding free energy of the ion can then be expressed as [106,115]:

$$\Delta G_{\mathrm{b}} = \Delta G_{\mathrm{tr}} + \Delta G_{\mathrm{int}}. \tag{6.4}$$

In this approach, the force field contribution to the free energy difference between the bound and unbound states (ΔG_{int}) is calculated using the alchemical FEP and TI methods, with a restrained ion/water molecule in the binding site/bulk solvent. To make sure that these restraints have little effect on ΔG_{int}, values of the spring constant are determined from the expression $k_{\mathrm{tr}} \sim 3k_BT\langle r^2\rangle$, where $\langle r^2\rangle$ is the mean-square fluctuations of the unrestrained ion in the binding site, which gives $k_{\mathrm{tr}} = 6$ kcal/mol/Å^2. The same goes for the replacement of a water molecule, which also samples a small volume in the ion binding site. In the bulk solvent, the restraints have no effect in the force field contribution to the free energy because bulk water is an isotropic medium. The free energy change due to loss of the translational entropy upon binding of the ion (ΔG_{tr}) is calculated separately using Eq. 6.3, where no restraints are employed. A more detailed validation of the method is given in a recent study of ligand binding to the ionotropic glutamate receptor GluA2 [157]. We note that the alchemical transformation approach adapted here is independent of path, and differs from the path-dependent methods that use a physical reaction coordinate. In the latter case, the ligand needs to be restrained within a cylinder or a cone while moving along the reaction coordinate, and this restriction has to be taken into account in calculating the binding affinity at the standard concentration [158,159].

6.A Appendix for Chapter 6

Table 6A.1: Equilibrium simulations performed in this study for the outward-facing model of EAAT3.

Mut.	Time	Glu	Na1	Na2	Na3	Comments
E374p	60ns	+	+	+	+	Transporter remained stable with gate closed, Na2 was released in two of the chains.
None	20ns	+	+	+	+	Instability in the substrate and in the HP2 gate, Na2 was released in all of the chains.
None	60ns	-	+	-	+	Opening and large fluctuations of the HP2 gate in all of the chains, Na^+ ions remained bound.
None	60ns	-	-	-	-	Opening and large fluctuations of the HP2 gate observed in all of the chains.
D455N	10ns	-	+	-	+	Na1 is released back to the solvent in two of the chains of the transporter.
D455N	10ns	-	+	-	-	Na1 remains bound coordinated by D368 in all of the chains of the transporter.
D455p	10ns	-	+	-	+	Na1 is released back to the solvent in all of the chains of the transporter.
D455p	10ns	-	+	-	-	Na1 remains bound coordinated by D368 in all of the chains of the transporter.

Table 6A.2: Equilibrium and FEP simulations performed in this study for the inward-facing model of EAAT3.

Mut.	Time	Glu	Na1	Na2	Na3	K^+	Comments
E374p	60ns	+	+	+	+	-	Transporter remained stable with gate closed, Na2 was released in one of the chains.
None	20ns	+	+	+	+	-	Instability in the substrate and in the HP1-HP2 gate, Na2 was released in all chains.
None	60ns	-	+	-	+	-	Opening of the HP1-HP2 gates in all of the chains, Na^+ ions remained bound.
None	60ns	-	-	-	-	-	Opening of the HP1-HP2 gates in all of the chains, with a large displacement of HP1.
None	50ns	-	-	-	-	s-1	K^+ remains bound at site 1 during the whole simulation.
None	50ns	-	-	-	-	s-2	K^+ remains bound at site 2 in two of the chains during the whole simulation.
None	50ns	-	-	-	-	s-3	K^+ remains bound at site 3 during the whole simulation.
FEP simulations							
None	16.8ns	-	-	-	-	s-1	K^+ binding affinity of -20.5 kcal/mol at site 1.
None	16.8ns	-	-	-	-	s-1	K^+ selectivity of 0.5 kcal/mol at site 1.
None	16.8ns	-	-	-	-	s-2	K^+ binding affinity of -9.5 kcal/mol at site 2.
None	16.8ns	-	-	-	-	s-2	K^+ selectivity of 3.9 kcal/mol at site 2.
None	16.8ns	-	-	-	-	s-3	K^+ binding affinity of -6.5 kcal/mol at site 3.
None	16.8ns	-	-	-	-	s-3	K^+ selectivity of -3.1 kcal/mol at site 3.

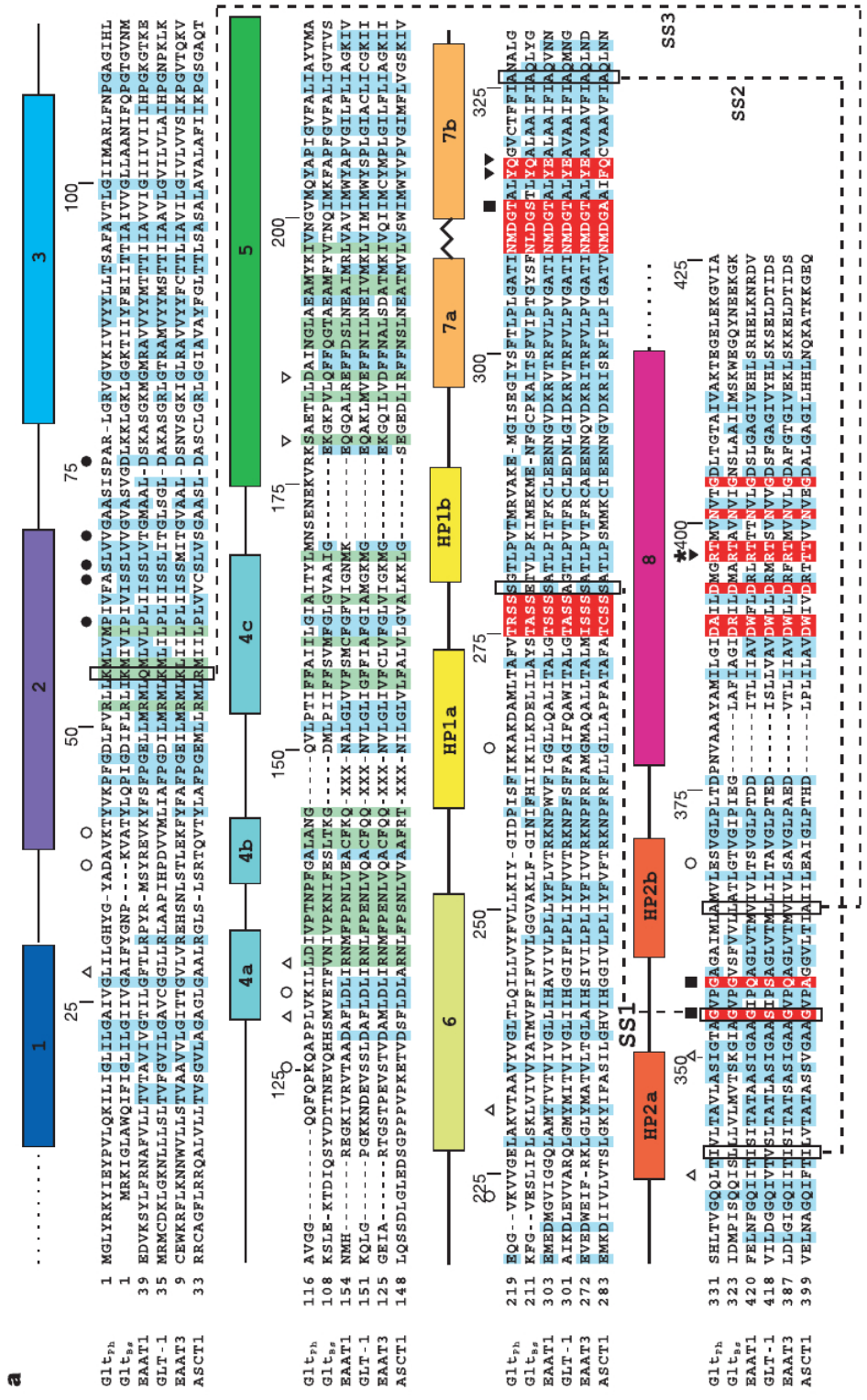

Figure 6A.1: The aligment used to build our models, from Ref.[14], showing the sequences of archaeal aspartate transporters, and human glutamate and neutral amino acid transporters.

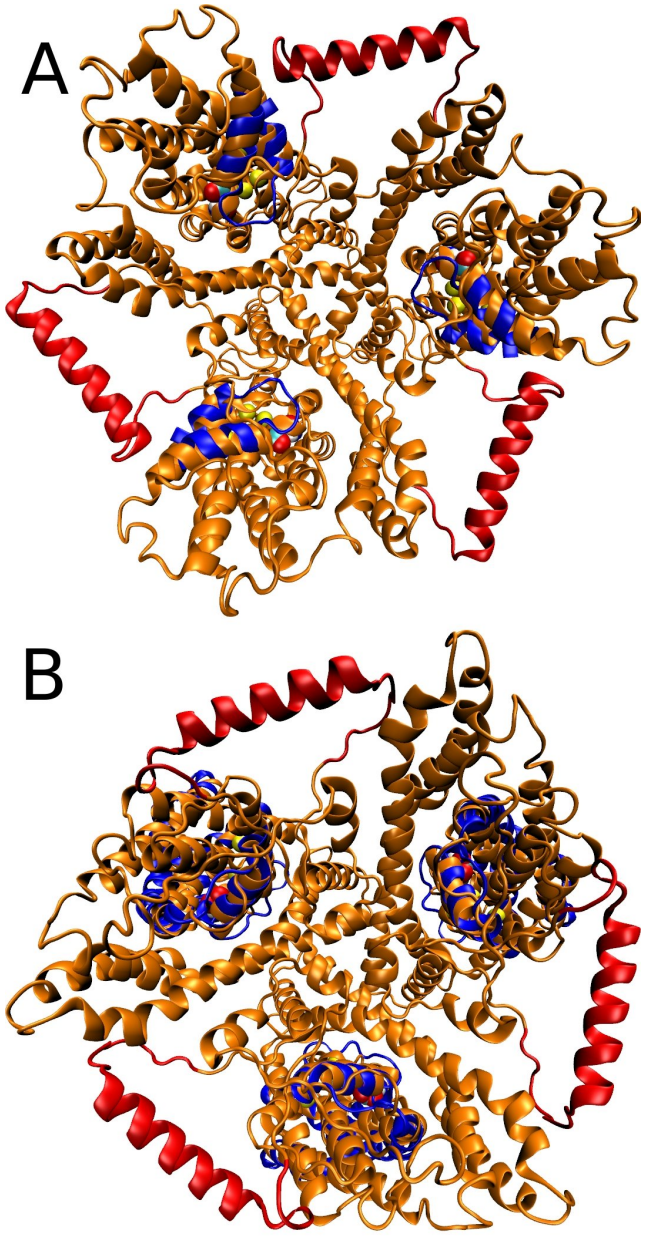

Figure 6A.2: (A) The model for the outward-facing state of EAAT3 including the 4B-4C loop 50-residue segment—which is not present in our model–placed at the position proposed in Ref.[146] (shown in red). The position of the HP2 segment in the open state, which is the segment that goes through conformational changes during the opening, is shown in blue. The substrate and sodium ions in the binding site are also shown. (B) The same for the inward-facing model of EAAT3, showing the segments that undergo conformational changes during the inward opening in blue, namely, HP1, HP2 and TM8. In both cases, the bound ligands and the conformational changes that happen during the opening of EAAT3 are not in proximity to the 4B-4C loop.

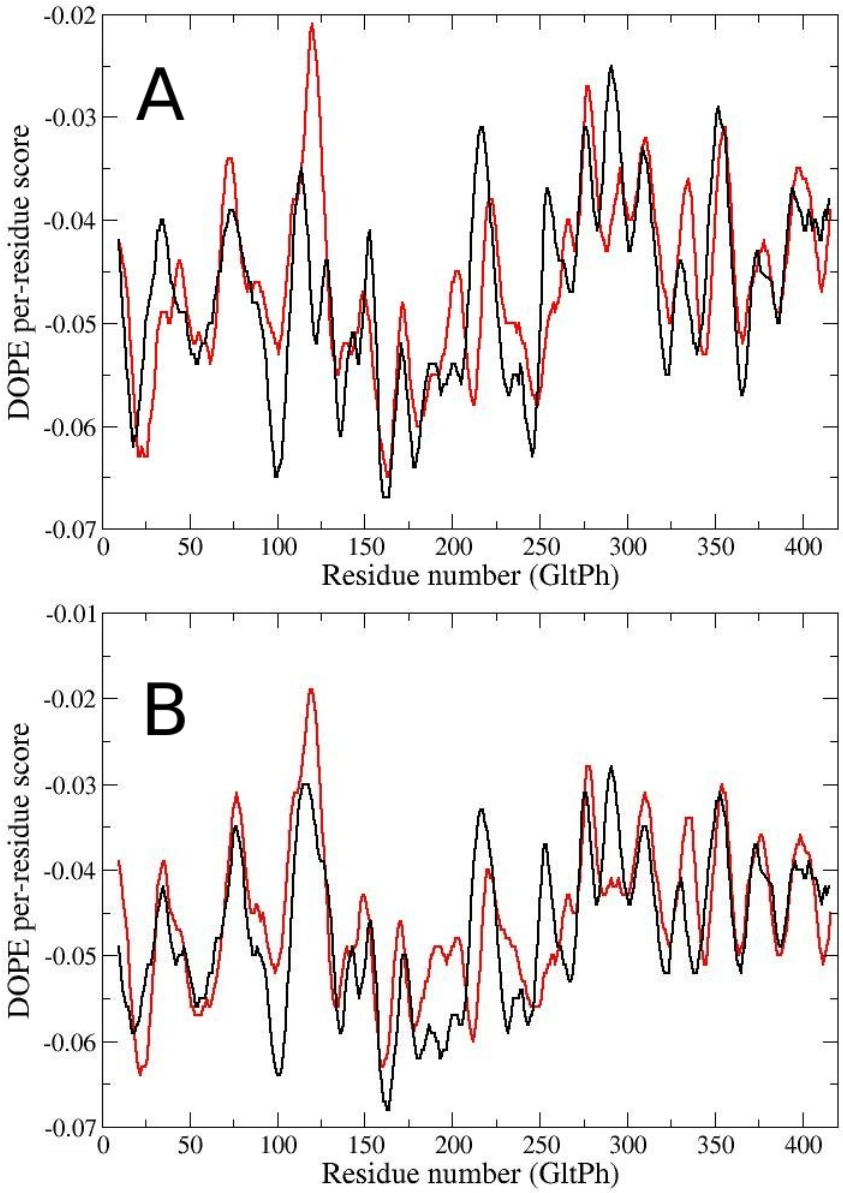

Figure 6A.3: (A) DOPE per-residue score for the 2NWX crystal structure (red) and the EAAT3 model in the outward-facing state (black). The residues are numbered according to the Glt_{Ph} sequence. (B) The same for the 3KBC crystal structure (red) and the EAAT3 model in the inward-facing state (black).

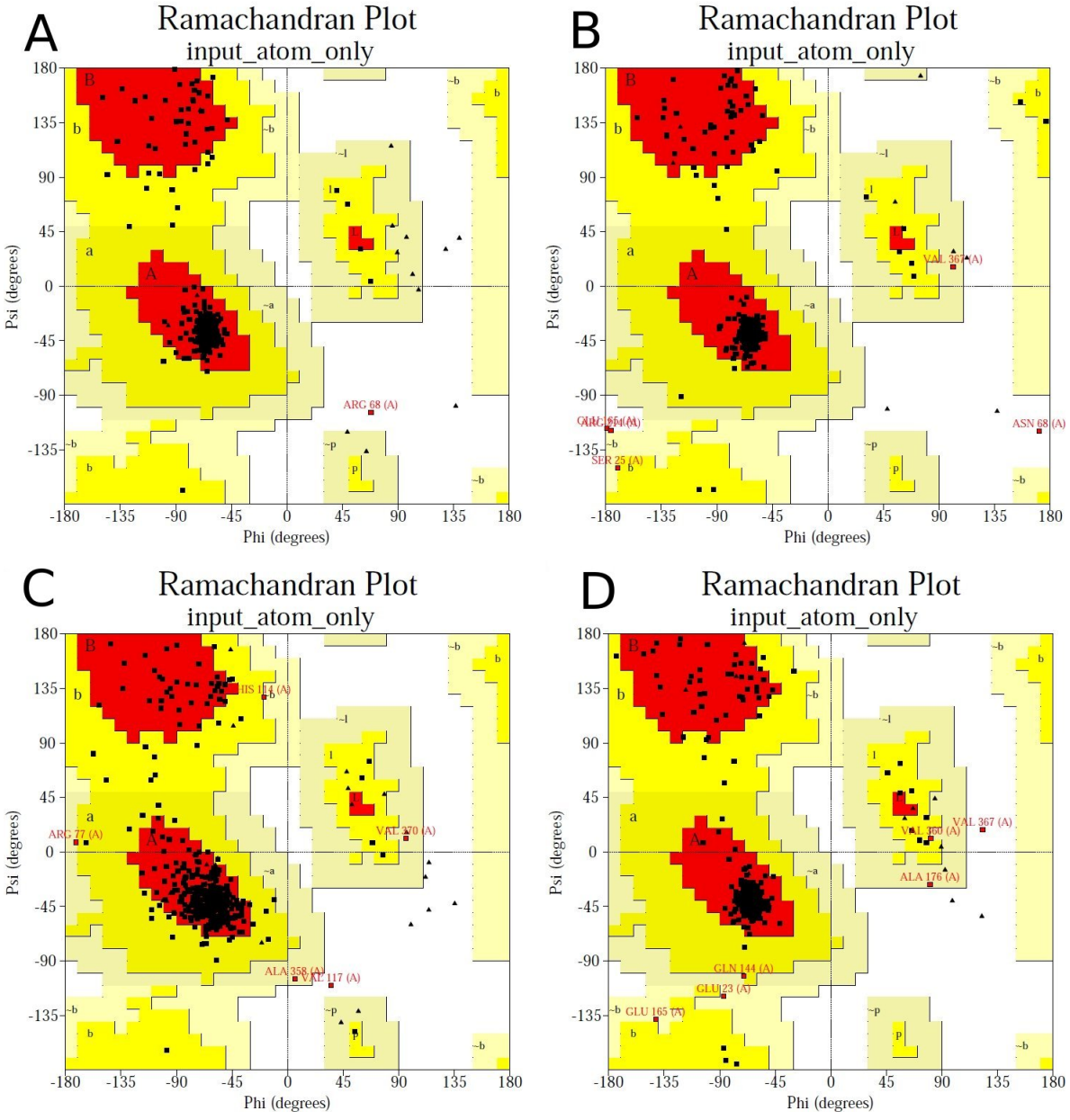

Figure 6A.4: Ramachandran plots for the templates and models used in this study.(A) 2NWX crystal structure.(B) EAAT3 model in the outward-facing conformation.(C) 3KBC crystal structure.(D) EAAT3 model in the inward-facing conformation.

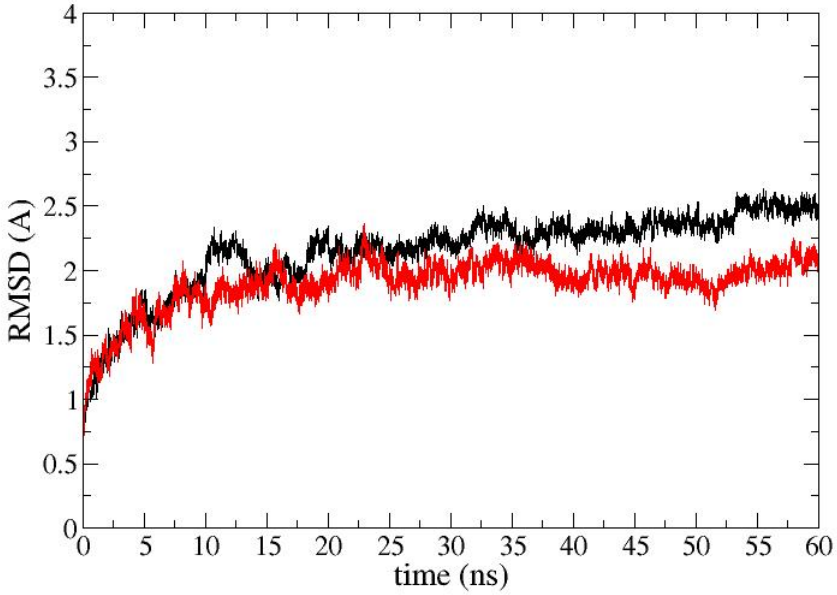

Figure 6A.5: RMSD of the protein backbone during 60 ns of simulations with all ligands bound and the protonated E374 side-chain, in the outward (black) and inward (red) states of EAAT3. The system shows no large conformational changes after 10 ns, which indicates that the models are stable.

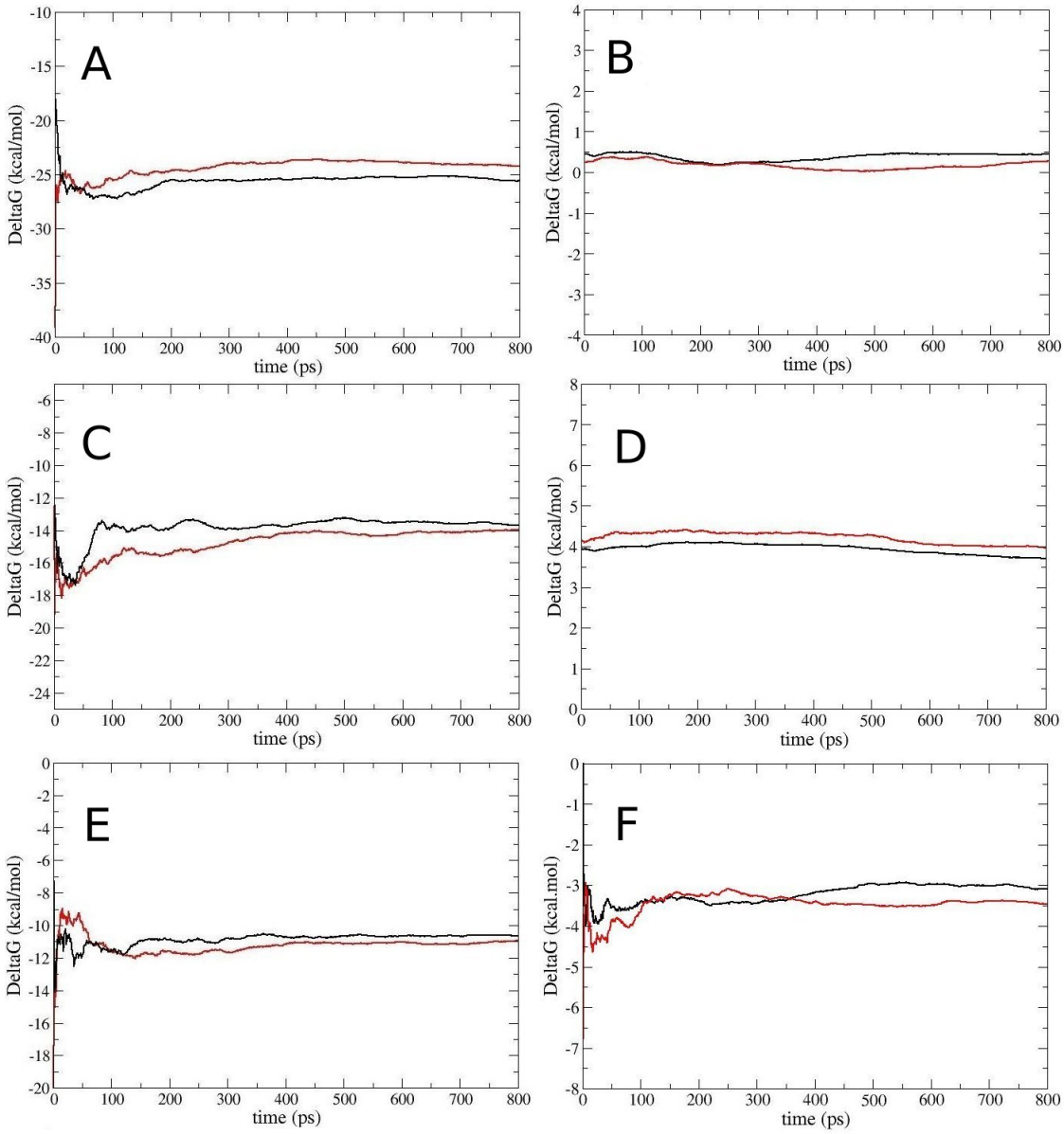

Figure 6A.6: Convergence of the TI results in the calculation of the binding free energies of K^+, and also the K^+/Na^+ selectivity. In all graphs the negative of the forward and the backward transformations results are shown with black and red lines, respectively. (A) Interaction free energy and (B) selectivity of the K^+ ion at site 1. (C,D) The same for the site 2. (E,F) The same for site 3.

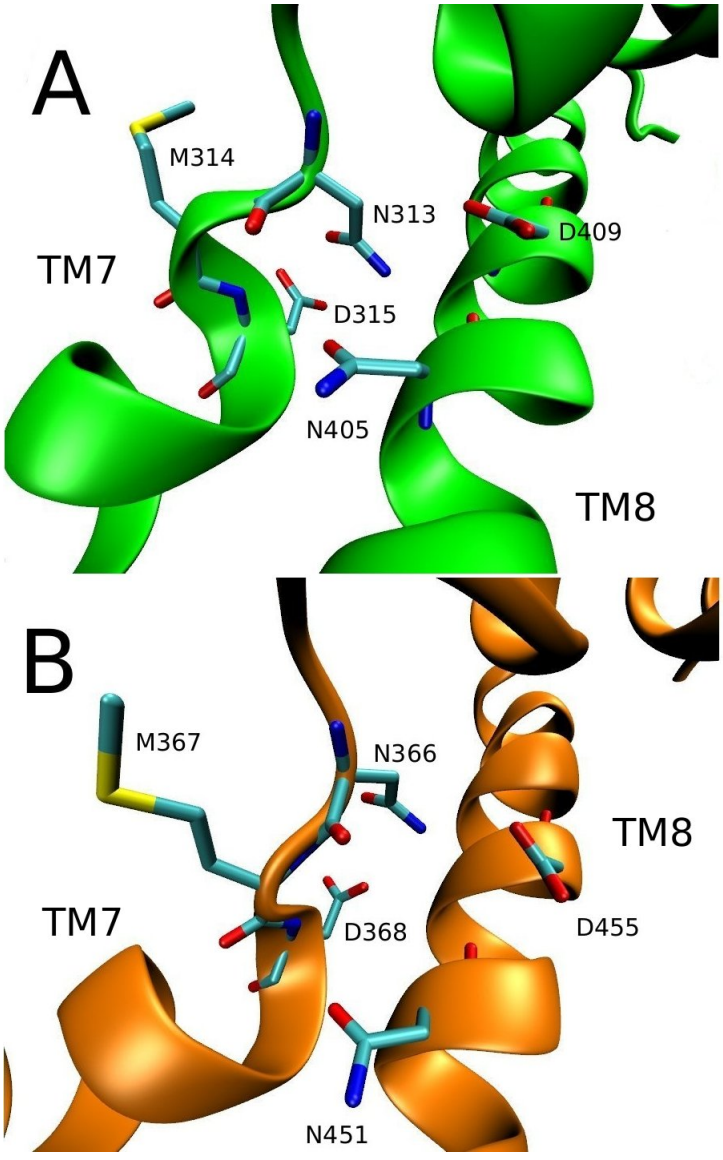

Figure 6A.7: Comparison of the conformations around the TM7–TM8 region in the Glt_{Tk} crystal structure (A) and in our simulations of the apo state of EAAT3 (B). Conformational changes that are observed between the loaded and apo states of the crystal structures are captured in our apo state model of EAAT3. Only the functionally important residues in this region are indicated explicitly.

Chapter 7

Molecular Dynamics Simulations Elucidate the Mechanism of Proton Transport in the Glutamate Transporter EAAT3

ABSTRACT

The uptake of glutamate in nerve synapses is carried out by the excitatory amino acid transporters (EAATs), involving the cotransport of a proton and three Na^+ ions and the countertransport of a K^+ ion. In this study we use an EAAT3 homology model to calculate the pK_a of several titratable residues around the glutamate binding site to locate the proton carrier site involved in the translocation of the substrate. After identifying E374 as the main candidate for carrying the proton, we calculate the protonation state of this residue in different conformations of EAAT3 and with different ligands bound. We find that E374 is protonated in the fully bound state, but removing the Na2 ion and the substrate reduces the pK_a of this residue and favors the release of the proton to solution. Removing the remaining Na^+ ions again favors the protonation of E374 in both the outward and inward-facing states, hence the proton is not released in the empty transporter. By calculating the pK_a of E374 with a K^+ ion bound in three possible sites, we show that binding of the K^+ ion is necessary for the release of the proton in the inward-facing state. This suggests a mechanism in which a K^+ ion replaces one of the ligands bound to the transporter, which may explain the faster transport rates of the EAATs compared to its archaeal homologs.

7.1 Introduction

Excitatory amino acid transporters (EAATs) are essential for proper functioning of the nervous system because they clear the excess glutamate released at the nerve synapses. Glutamate can be toxic to neurons in large concentrations, so problems in these transporters have been implicated in many pathological conditions including cerebral ischemia, amyotrophic lateral sclerosis and Alzheimer's disease[12]. EAATs transport the substrate across the membrane by coupling to three Na^+ ions and a proton, and a K^+ ion is countertransported at every cycle. There is a great amount of experimental data on EAATs, mostly through mutagenesis experiments[28,30–34,38,92,96,129–142], but also via structural information[146,160] and the pH dependence of transport[137,145]. A significant step in our understanding of the EAATs came with the solution of the crystal structures of the archaeal homolog Glt_{Ph} both in the outward- and inward-facing conformations[14–16]. Glt_{Ph} has aspartate as the main substrate and the transport mechanism is not coupled to a proton or K^+, but it still involves the cotransport of three Na^+ ions. Only two of them could be resolved in the crystal structure, but radio-labeled Na^+ experiments showed that three Na^+ ions were coupled to the substrate transport[37]. The third Na^+ site (Na3) was determined from molecular dynamics (MD) simulations, and mutagenesis experiments confirmed that this site was present in Glt_{Ph} and conserved in the EAAT family[38]. In Glt_{Ph} two Na^+ ions bind before the substrate, called Na1 and Na3, and the last Na^+ ion (Na2) binds after the substrate and the closure of the HP2 gate that forms the binding site[123]. Experimental studies show that this binding sequence is probably conserved in EAAT3, with the binding of two Na^+ ions in the absence of glutamate and the third Na^+ ion after the substrate[31,133].

Recently, we have created homology models for the outward- and inward-facing states of EAAT3 based on the Glt_{Ph} structures in the same conformations. Glt_{Ph} shares about 37% sequence identity with the EAATs, but the homology is much higher for residues near the binding pocket, reaching around 60% in this region. Furthermore, almost all of the residues shown to be involved in the binding of Na^+ ions and the substrate in Glt_{Ph} are conserved in the EAATs[14–16,32,90,94]. We have shown that the model is stable with glutamate as a ligand, and that the coordination of the ligands can be reproduced in good agreement with mutagenesis experiments[161]. A major result from this work is that the substrate is not stably bound unless the E374 side chain of EAAT3 is protonated, making this residue a natural candidate for the proton carrier, which also agrees well with experiments[137]. Using the inward-facing state of EAAT3 from our model, we have also investigated three possible binding sites for the K^+ ion, and found a high-affinity site that overlaps with the Na1 and Na3 binding sites.

Even though there is strong theoretical and experimental evidence pointing to the residue E374 as the proton carrier, the mechanism of proton association and dissociation from this residue in the outward and inward states of EAAT3 remains unknown. Also the protonation states of other titratable residues close to the bind-

ing site has not been thoroughly investigated, and in MD simulations, it is usually assumed that they retain the same protonation states as in solution. Here we use homology models of EAAT3 to address these issues. We perform free energy perturbation (FEP) calculations to obtain the pK_a values of different residues close to the substrate binding site in the fully bound outward conformation. After identifying the residue E374 as the only one being protonated under these conditions, we calculate the pK_a value of this residue in different states of the transporter and with different ligands bound. We show that the protonation/deprotonation of E374 is coupled to the binding/unbinding of the substrate, as well as the closure/opening of the HP1 and HP2 gates that give the substrate access to the solution. We also show that the binding of a K^+ ion in the inward state of EAAT3 is necessary for the deprotonation of E374, so that only the K^+ ion is transported back to the extracellular media in the second half of the transport cycle.

7.2 Model system and simulation details

For the outward and inward-facing conformations of EAAT3 we use the same models created in our previous work[161], built from their respective Glt_{Ph} templates and equilibrated in the presence and absence of the various ligands in both states. In addition to these two, we have also created a new model for the outward-apo state of EAAT3 from a recent crystal structure of Glt_{Tk} (PDB ID: 4KY0), another archaeal homolog of the EAATs and a close homolog of Glt_{Ph}[151]. This Glt_{Tk} structure is in the outward-apo state with the HP2 gate open, for which there were no crystal structures of EAAT3 homologs available until now. To build this model we have used the program MODELLER[74], following the same protocol as described in[161]. We have produced 20 initial structures and choose the one with the best scores from the Swiss Model Structure Assessment Tool[79,80]. We then construct a trimer by superposing the EAAT3 monomer to chains A, B and C of Glt_{Tk}.

The simulation systems for the models built from the Glt_{Ph} templates are prepared and equilibrated as described in[161]. The EAAT3 system built from the Glt_{Tk} template is prepared in the same way, using the software VMD[104]. First the EAAT3 trimer in the apo-state is embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) phospholipid bilayer. We then solvate the protein-membrane complex in a box of water molecules with 54 Na^+ ions and 45 Cl^- ions. The extra Na^+ ions are included to keep the system neutral. The simulation box has 237 lipid molecules and 15,851 water molecules, with a total of 98,124 atoms. The system is then equilibrated in two stages. First, the coordinates of the ligand and protein atoms are fixed and the system is equilibrated with 1 atm pressure coupling in all directions until the correct water and lipid densities are obtained. At this point, the x and y-dimensions of the simulation box are fixed, and pressure coupling is applied only in the z direction (typical dimensions of the simulation are $113 \times 113 \times 74$ Å^3). In the second stage, the protein is gradually relaxed by reducing the restraints in

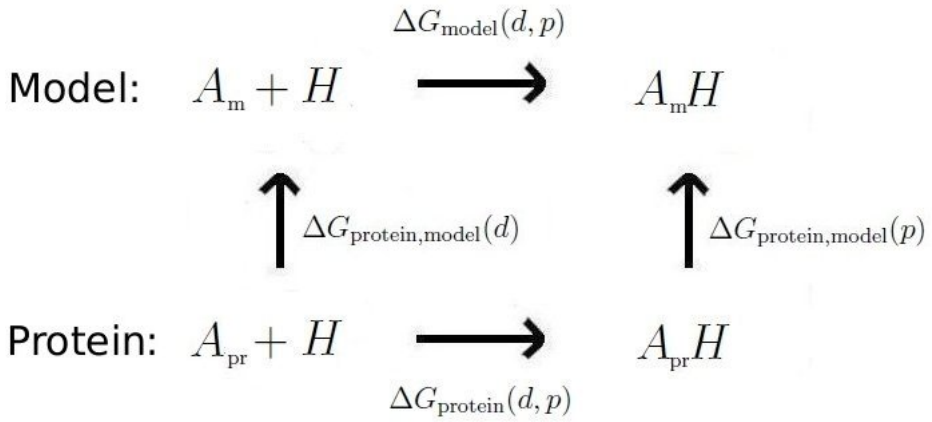

Figure 7.1: The thermodynamic cycle used in the FEP/TI calculations: A represents the side-chain of interest, H is the proton, the indices p and d refer to the protonated and deprotonated states, and the subscripts m and pr indicate that A is in the bulk solvent (model compound) and in the protein, respectively.

several steps during MD simulations lasting 2.4 ns. We then perform a further 15 ns of simulations with no restraints applied.

MD simulations are performed using the NAMD program[59] with the CHARMM36 force field[155]. The temperature is maintained at 300 K using Langevin damping with a coefficient of 5 ps^{-1}, and the pressure is kept at 1 atm using the Langevin piston method with a damping coefficient of 20 ps^{-1}[64]. Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. The Lennard-Jones (LJ) interactions are switched off between 10 and 12 Å using a smooth switching function. A time step of 2 fs is used in all MD simulations.

7.3 Free energy calculations

To calculate the pK_{a} values of the aspartate and glutamate side-chains in the vicinity of the glutamate binding site, we use the alchemical free energy perturbation (FEP) and thermodynamic integration (TI) methods in molecular dynamics (MD) simulations[52]. These methods use an explicit solvent and have advantages over continuum electrostatic models. Continuum models are very sensitive to the dielectric constant assigned to the protein interior, and the value used is usually an empirical parameter with a value ranging from 8 to 20 to account for different effects such as water penetration, protein flexibility, and conformational rearrangements in response to the ionization state[162–164]. These models are also influenced by the method used

in constructing the dielectric boundary between the protein and the surrounding environment[165]. All these effects can be taken into account rigorously by computing the free energy difference between the protonated and deprotonated forms using the MD/FEP method with an explicit solvent[166]. This is carried out by alchemically transforming a deprotonated carboxyl group to a protonated one in the protein while simultaneously performing the opposite transformation in a reference residue located in bulk solvent, according to the thermodynamic cycle shown in Fig. 7.1. The reference residues, or model compounds, that we use are *N-acetyl-L-isoasparagine* for the aspartate calculations and the *N-acetyl-L-isoglutamine* for the glutamate calculations (Fig. 7.2). These two compounds reproduce the side-chains of aspartate and glutamate and also the backbone atoms, but do not have the zwitterion charges that could affect the electrostatic component of the results. The experimental p$K_{\rm a}$ values of *N-acetyl-L-isoasparagine* and *N-acetyl-L-isoglutamine* are 4.0 and 4.3, which compare well with the p$K_{\rm a}$ values of aspartate (3.7) and glutamate (4.3).[167]. The CHARMM parameters are readily available for these compounds[55], which further facilitate their use in FEP calculations. The p$K_{\rm a}$ value of a residue in a protein is given by[168]:

$$\mathrm{p}K_{\rm a} = \mathrm{p}K_{\rm a(model)} - \frac{\Delta\Delta G_{\rm pr}}{2.3k_BT} \tag{7.1}$$

where $\Delta\Delta G_{\rm pr}$ can be calculated from the thermodynamic cycle in Fig. 7.1

$$\Delta\Delta G_{\rm pr} = \Delta G_{\rm pr\to m}(d) - \Delta G_{\rm pr\to m}(p) = \Delta G_{\rm pr}(d\to p) - \Delta G_{\rm m}(d\to p) \tag{7.2}$$

MD/FEP calculations are performed to determine the free energy difference on the right hand side of Eq. 7.2, corresponding to the horizontal legs in Fig. 7.1.

We use both the MD/FEP and MD/TI methods to calculate $\Delta\Delta G_{\rm pr}$ by alchemically transforming a deprotonated glutamate residue to a protonated one in the protein while doing the opposite transformation in the model compound in bulk. This process is carried out as a function of the reaction coordinate λ, in which the end points are at $\lambda = 0$ and $\lambda = 1$. In the FEP method, the interval between $\lambda = 0$ and 1 is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average

$$\Delta G_i = -k_{\rm B}T \ln\langle\exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_{\rm B}T]\rangle_{\lambda_i}. \tag{7.3}$$

where $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, with H_0 and H_1 representing the states with the proton in the model compound and in the protein residue, respectively. The free energy difference is obtained from the sum, $\Delta\Delta G_{\rm pr} = \sum_i \Delta G_i$. In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, is obtained at several λ values, and the free energy difference is calculated from the integral:

$$\Delta\Delta G_{\rm pr} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial\lambda} \right\rangle_\lambda d\lambda. \tag{7.4}$$

Provided the integrand can be fitted well with a polynomial, Gaussian quadrature provides an efficient and accurate method for evaluation of such integrals because it

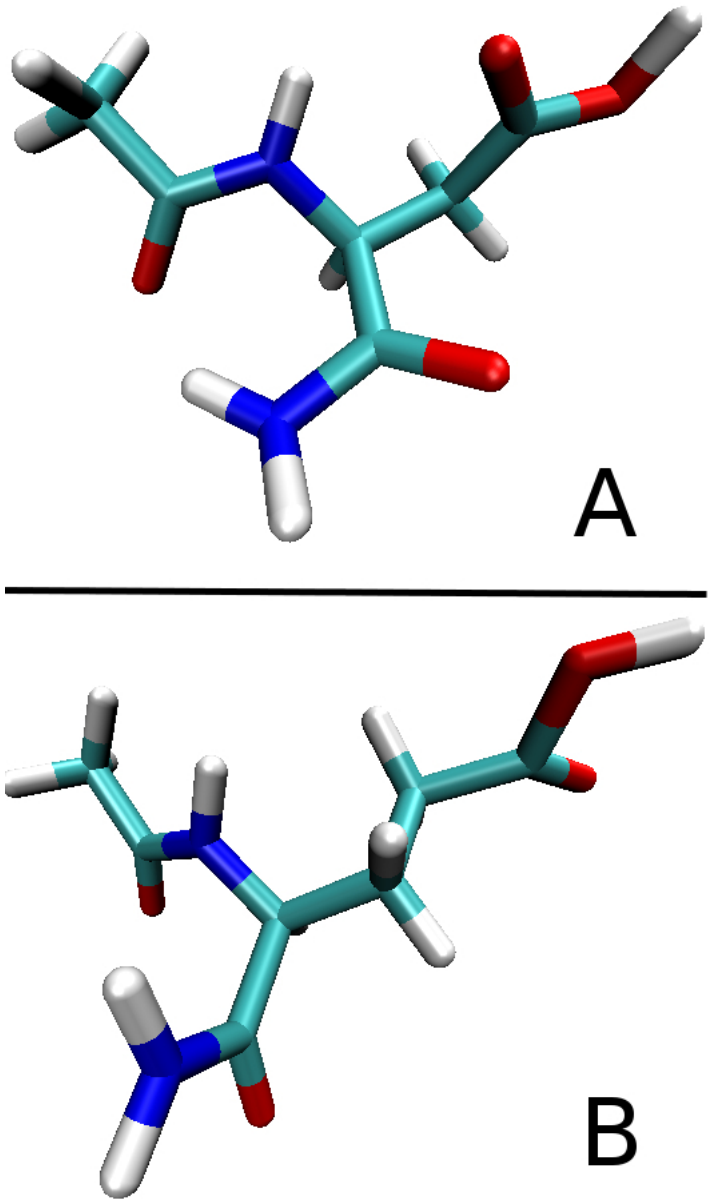

Figure 7.2: The model compounds used in our free energy calculations in bulk. (A) *N-acetyl-L-isoasparagine* is used for aspartate, (B) *N-acetyl-L-isoglutamine* is used for glutamate. The colors are cyan for carbon atoms, blue for nitrogen atoms, red for oxygen and white for hydrogens.

allows for longer sampling of a smaller number of windows. In the FEP simulations we use 66 exponentially spaced windows with 40 ps of equilibration plus 40 ps of production for each window. In the TI method, we use seven windows starting from the nearest FEP windows, and run 0.4 ns of equilibration plus 0.8 ns of data collection for each window. These parameters have been found to be optimal in previous studies of Glt_{Ph}[123,144]. We report here only the TI results because they show better convergence, but we always check for consistency between the two methods, as well as performing both the forward and backward transformations in each case to check against any hysteresis effects.

7.4 Results and discussion

7.4.1 Evaluating the model

We have assessed the quality the open-state EAAT3 model obtained from Glt_{Tk} using three different methods, similar to the case of the EAAT3 models obtained from the closed states of Glt_{Ph}[161]. QMEAN (Qualitative Model Energy ANalysis) is a composite scoring function ranging from 0 to 1, which describes the major geometrical aspects of protein structures[79]. DFIRE (Distance-scaled, Finite Ideal-gas REference) is a potential based on a database of non-homologous proteins[80]. DOPE (Discrete Optimized Protein Energy) is an atomic distance-dependent statistical potential calculated from a sample of native structures[74]. The first two are available on the Swiss Model Server, and the last one is built into the program MODELLER. In Table 7.1 we show the QMEAN scores and DFIRE energies for the original Glt_{Tk} crystal structure, as well as our EAAT3 model in the outward apo state. As in the case of Glt_{Ph}[161], the QMEAN scores and DFIRE energies are very similar, as well as the energy profiles obtained using DOPE (Fig. 7.3). This indicates that, even though the homology is not high between the archaeal and the human transporters, we can still built high quality models using the former as a template. Furthermore, we are looking at ligand binding and small conformational changes instead of large conformational transitions. Therefore, the models built for the EAATs in the vicinity of the outward- and inward-facing states are likely to retain the same features of the archaeal transporters in the same states.

7.4.2 pK_a values in the glutamate-bound outward EAAT3

The first step in identifying the proton carrier is to calculate the pK_a values of all titratable residues in the vicinity of the binding pocket. The fully-bound outward closed state precedes the transition to the inward-facing transporter, and therefore should have a proton bound. We choose this equilibrated model to look for the proton carrier in the EAATs, by performing pK_a calculations as described in methods. We have shown previously that the substrate is not stable unless the E374 residue, which is a glutamine in Glt_{Ph} and Glt_{Tk}, is protonated[161]. We cannot discard the

Table 7.1: Comparison of the QMEAN scores and DFIRE energies between the Glt_{Tk} crystal structure and the EAAT3 homology model. The QMEAN score goes from 0 to 1, 1 being the best possible model. Lower DFIRE energy values indicate better quality structures.

Structure	QMEAN score	DFIRE Energy
Glt_{Tk} (4KY0)	0.515	−679.6
EAAT3 (outward apo)	0.484	−608.3

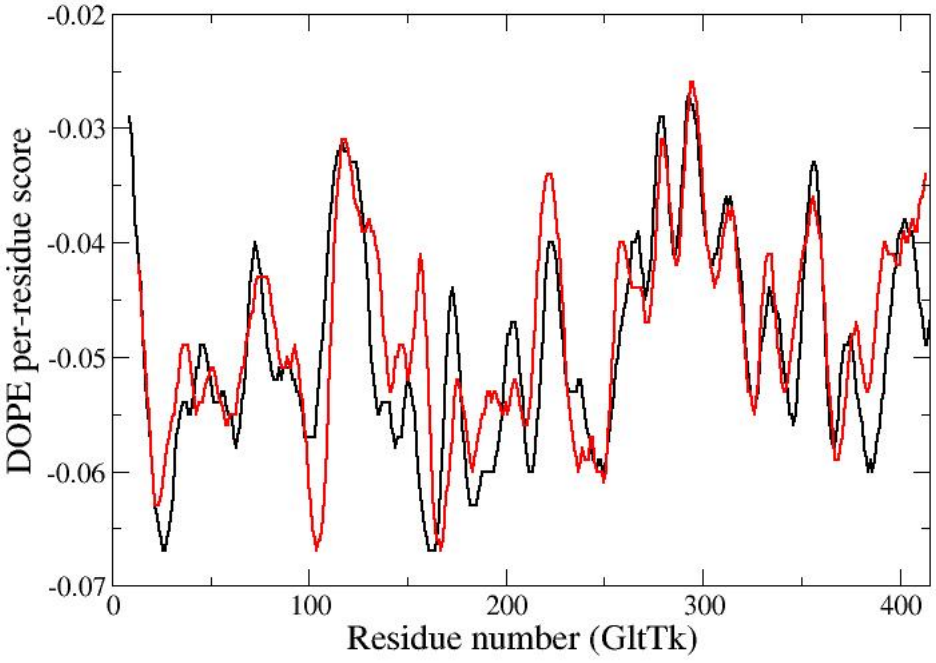

Figure 7.3: DOPE per-residue score for the 4KY0 crystal structure (black) and our EAAT3 model in the apo outward-facing state (red). The residues are numbered according to the Glt_{Tk} sequence.

possibility that other residues in the binding pocket could also be protonated, so, in addition to E374, we have also calculated the pK_a values of the residues D368, D440, D444 and D455. The side chains of D455 and D368 are involved in Na^+ binding at sites Na1 and Na3, the side chain of D444 coordinates the substrate α-amino group, and the side chain of D440 makes a salt bridge with the R447 residue (Fig. 7.4).

This arginine also coordinates the glutamate substrate side-chain carboxyl group. In order to test the suitability of our reference residues in bulk, we have also calculated the pK_a values of an aspartate (D83) and a glutamate (E135) residue in the protein that are fully exposed to the solvent and are not close to any charged groups. These residues are expected to retain the same pK_a values as in bulk. To avoid the unbinding of ligands during the protonation process, all the ligands are restrained during the pK_a calculations at the binding sites found in Ref.[161]. This procedure allows us to calculate the pK_a values in a given state, so we can determine which residues are protonated under these conditions. The Na1 and Na2 ions share the same binding sites as the Glt_{Ph} 2NWX crystal structure, and the Na3 site binding site was determined from MD simulations and mutagenesis experiments[38]. The coordination of the glutamate substrate is very similar to that of aspartate in Glt_{Ph}, which is stable with a protonated E374, and is in good agreement with the mutagenesis experiments[32,33,161].

Table 7.2: Calculated pK_a values of various aspartate and glutamate residues in EAAT3. The forward and the backward calculations of the protonation free energies are listed in the second and third columns, respectively, and their average in the fourth column. The shift ΔpK_a is given in the fifth column, and the final value of the pK_a is given in the last column. Errors are estimated from block data analysis using 100 ps windows. All free energies are in kcal/mol.

Residue	$\Delta\Delta G_{pr}^{for}$	$-\Delta\Delta G_{pr}^{back}$	$\Delta\Delta G_{pr}$	ΔpK_a	pK_a
D83	0.4 ± 0.8	1.1 ± 0.9	0.7 ± 0.8	-0.5 ± 0.6	3.5 ± 0.6
E135	0.2 ± 0.9	0.2 ± 0.8	0.2 ± 0.8	-0.1 ± 0.6	4.2 ± 0.6
E374	-20.1 ± 0.7	-20.3 ± 1.1	-20.2 ± 0.9	14.7 ± 0.7	19.1 ± 0.7
D368	18.8 ± 0.9	17.0 ± 0.9	17.9 ± 0.9	-13.1 ± 0.7	-9.1 ± 0.7
D440	6.8 ± 1.7	6.0 ± 1.1	6.4 ± 1.4	-4.7 ± 1.0	-0.7 ± 1.0
D444	-0.4 ± 0.9	0.8 ± 1.1	0.2 ± 1.0	-0.1 ± 0.7	3.9 ± 0.7
D455	16.3 ± 2.3	14.9 ± 1.5	15.6 ± 1.9	-11.4 ± 1.4	-7.4 ± 1.4

In Table 7.2, we present the calculated pK_a values of the glutamate and aspartate residues mentioned above. The pK_a values for the test residues D83 and E135 are 3.5 and 4.2, respectively, which are in good agreement with the experimentally determined pK_a values for the side chains of aspartate and glutamate (3.7 and 4.3). Thus our models can accurately reproduce the properties of the aspartate and glutamate side chains in bulk water. Returning to the EAAT3 residues, we find a pK_a value of 19.1 for E374 (Table 7.2), indicating that this residue is undoubtedly protonated in the fully-bound outward state of EAAT3. We believe that this large pK_a value is related to both the presence of the glutamate substrate and the closure of the HP2 gate after this ligand binds. We have shown previously that glutamate substrate is not stable in the binding site if E374 is deprotonated due to the repul-

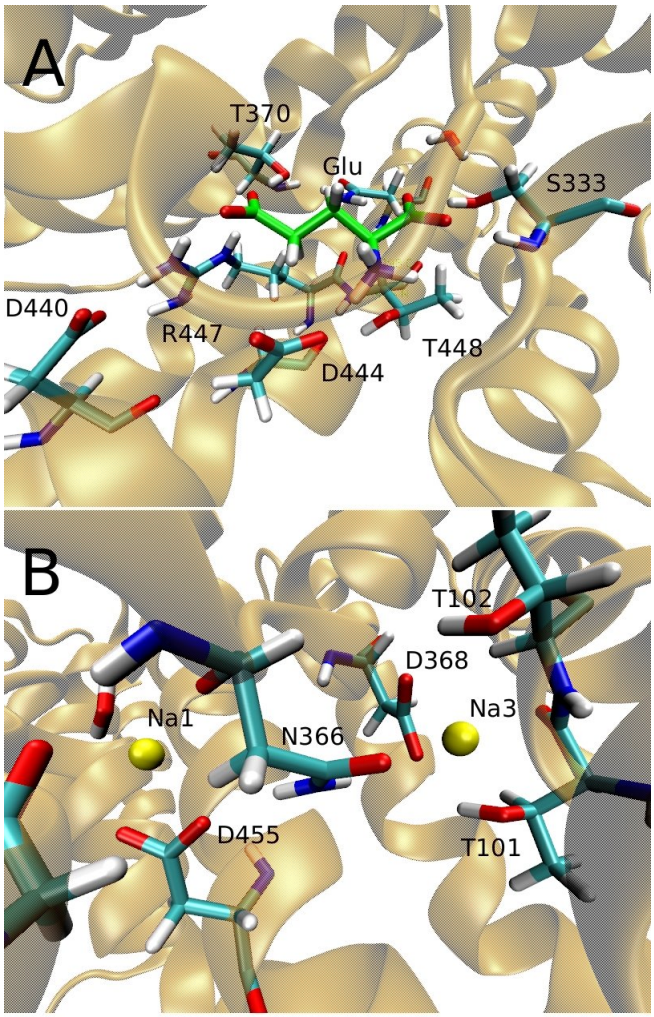

Figure 7.4: (A) The substrate binding site in the outward-facing fully-bound state, showing the bound glutamate (in green) and the titratable residues in this region. (B) The binding sites of the Na1 and Na3 ions in the same state of EAAT3, coordinated by the D455 and D368 side chains, respectively.

sion between the negatively charged carboxyl groups[161]. Therefore, protonation of E374 would be energetically favourable in the presence of the substrate. Furthermore, E374 is isolated from the solvent by a nonpolar region formed by a series of hydrophobic residues at the TM7 (M367 and T370) and HP2 (A404, A408, A414, T418 and V422) segments (Fig. 7.5A). All of them are strictly conserved throughout the EAAT family except for V422, which is an isoleucine in EAAT2. Conservation of this hydrophobic motif in all EAATs is indicative of its functional importance in stabilizing the protonated state of E374.

The experimental pK_a value of the proton carrier in EAAT3 is about 8[145]. This value is determined from the pH-dependence of the uncoupled and coupled currents associated with glutamate transport, thus reflecting the free energy difference between the fully bound protonated state and the ensemble of deprotonated conformations where the substrate translocation cannot happen. We believe that the deprotonated state of E374 is likely to be solvent exposed and without a bound substrate, given that the substrate is unstable under these conditions. Because accounting for the conformational changes and the unbinding of the substrate is not feasible in our FEP calculations, we have calculated the pK_a for the fully-bound closed state, where the E374 side-chain is isolated from the solvent and glutamate is bound. Hence the calculated pK_a value of 19.1 cannot be directly compared to the experimental value, but it is sufficient to ascertain that E374 is protonated in the fully-bound transporter. After 60 ns of MD simulations of EAAT3 in the absence of the substrate and the Na2 ion[161], we have observed the opening of the HP2 gate and the hydration of E374 (Fig. 7.5B). This will result in a change in the pK_a value of E374, which may trigger release of the proton in the inward state. This scenario will be explored further in the next section.

We next consider the pK_a values of other titratable residues. The side chains of D368 and D455, which coordinate Na3 and Na1, respectively, give very large values of $\Delta\Delta G_{pr}$, and consequently very low pK_a values. This is a direct consequence of the proximity of these two side-chains to Na1 and Na3 and the absence of water molecules in this region— there is only one water molecule in the binding sites of Na1, Na3 and the substrate (Fig. 7.4). Protonating a negatively charged aspartate side-chain that coordinates a positively charged Na^+ ion results in a very high energetic penalty, which is reflected in the calculated pK_a values. This indicates that the residues D368 and D455 have to be deprotonated in order to coordinate the Na1 and Na3 ions. We have shown previously that Na1 leaves the binding site shortly after D455 is protonated[161], which is consistent with the results obtained here. The D440 and D444 residues also give low pK_a values, most likely because they are also involved in charge interactions with the R447 side-chain and the substrate α-amino group, respectively (Fig. 7.4A).

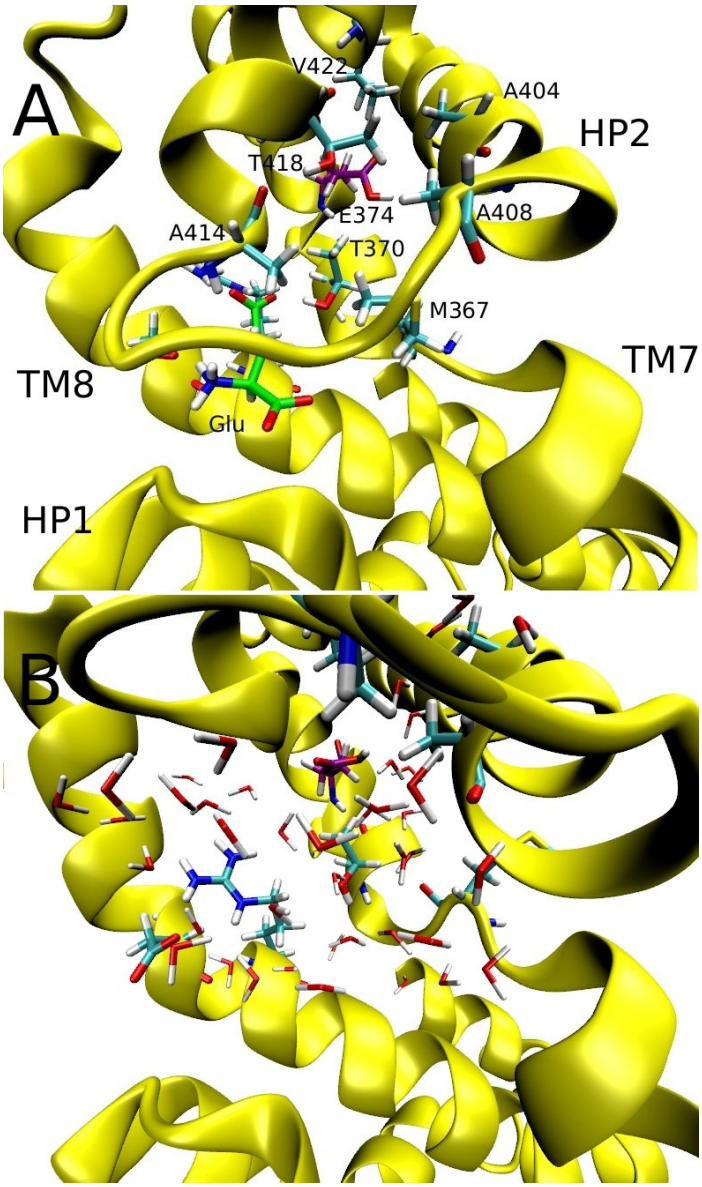

Figure 7.5: (A) The region around E374 (purple) in the fully bound state, where this residue is buried under a series of hydrophobic residues from the HP2 and TM7 segments. (B) After the removal of the substrate and the Na2 ion, the HP2 gate opens and the E374 side-chain is hydrated.

7.4.3 pK_a values of E374 in different states of EAAT3

In our previous study of EAAT3[161], we obtained a series of states by running simulations with different combinations of bound ligands. Besides the fully bound states in the outward- and inward-facing conformations, we also simulated these two conformations in the apo state and with only Na1 and Na3 bound to the transporter. Further, we investigated the binding sites of the K^+ ion in the inward model of EAAT3 and calculated the binding free energies of the K^+ ion in three sites suggested by the mutagenesis experiments[161]. Here we calculate the pK_a value of E374 in all these states, as well as in the EAAT3 model built from the apo Glt_{Tk} (Table 7.3). Our aim is to gain insights into the mechanism of proton binding and release and how that relates to the binding/unbinding of the other ligands. Because all our results are obtained using the EAAT3 model created from the fully bound Glt_{Ph}, it is important to check its consistency by comparing the apo state result with that obtained from the EAAT3 model based on the apo Glt_{Tk} template. As seen from the top two rows in Table 7.3, there is excellent agreement between the two results, confirming the relevance of the apo-state EAAT3 model used in the calculations. Comparison of the pK_a values between the outward and inward states for the apo, Na1-Na3 bound, and fully bound states (Table 7.3) shows that they are very similar. Even though this symmetry demonstrates the consistency of our results, experiments based on the pH dependence of the forward and reverse transport modes reveal an asymmetry of 1.5 pK_a units between the outward- and inward-facing states of EAAT3 in the presence of the substrate[145]. This discrepancy might arise from the free energies involved in the gating process of the substrate-bound EAAT3, which we do not consider in this study. This step isolates E374 from the solvent, directly affecting its pK_a, and is likely to be different in the inward and outward states, as shown in the metadynamics studies of gate opening in Glt_{Ph}[126].

We see from Table 7.3 that absence of the substrate and the Na2 ion (Na1, Na3 rows) has a great effect on the pK_a of this residue, which can be traced to hydration of E374 and the absence of the negatively charged substrate (Fig. 7.5). The pK_a value is reduced to 6.6 and 6.7 for the outward- and inward-states, meaning that around 85% of the E374 side-chains are deprotonated when only Na1 and Na3 are bound. This shows that the binding of the substrate and the closure of the HP2 gate in the presence of the these two Na^+ ions strongly favour proton binding to E374. To determine the relative contribution of these two processes, we have removed the Na2 ion from the fully bound outward state and simulated the system with a deprotonated E374 and a restrained substrate for 50 ns. In one of the chains of the transporter, we obtain a state in which Na1, Na3 and glutamate are bound, but the HP2 gate is fully open and E374 is hydrated. The pK_a value of E374 under these conditions is 10.4 (Table 7.3), which indicates that binding of glutamate is sufficient for the protonation of E374, and the closure of the gate further stabilizes it. As in Glt_{Ph}, Na1 and Na3 are necessary for substrate binding to EAAT3, but these ions also favor the deprotonation of E374 and the consequent instability of the

Table 7.3: Calculated pK_a values of E374 with different ligands bound to the outward and inward states of EAAT3. The forward and the backward calculations of the protonation free energies are listed in the second and third columns, respectively, and their average in the fourth column. The shift ΔpK_a is given in the fifth column, and the final value of the pK_a is given in the last column. Errors are estimated from block data analysis using 100 ps windows.All free energies are in kcal/mol.

Outward EAAT3					
Ligands	$\Delta\Delta G_{pr}^{for}$	-$\Delta\Delta G_{pr}^{back}$	$\Delta\Delta G_{pr}$	ΔpK_a	pK_a
apo	-10.4 ± 1.0	-10.1 ± 0.9	-10.2 ± 1.0	7.4 ± 0.7	11.7 ± 0.7
apo (Glt_{Tk})	-9.2 ± 1.5	-11.5 ± 1.2	-10.3 ± 1.4	7.5 ± 1.0	11.8 ± 1.0
Na1, Na3	-4.0 ± 0.9	-2.1 ± 1.1	-3.0 ± 1.0	2.2 ± 0.7	6.6 ± 0.7
Na1, Na3, Glu[a]	-8.9 ± 0.9	-7.7 ± 1.4	-8.3 ± 1.2	6.1 ± 0.9	10.4 ± 0.9
fully bound	-20.1 ± 0.7	-20.3 ± 1.1	-20.2 ± 0.9	14.7 ± 0.7	19.1 ± 0.7
Inward EAAT3					
Ligands	$\Delta\Delta G_{pr}^{for}$	-$\Delta\Delta G_{pr}^{back}$	$\Delta\Delta G_{pr}$	ΔpK_a	pK_a
apo	-10.2 ± 1.6	-10.1 ± 1.2	-10.2 ± 1.4	7.4 ± 1.0	11.7 ± 1.0
Na1, Na3	-3.3 ± 1.5	-3.4 ± 1.2	-3.3 ± 1.4	2.4 ± 1.0	6.7 ± 1.0
fully bound	-22.7 ± 1.1	-22.2 ± 0.9	-22.4 ± 1.0	16.3 ± 0.7	20.6 ± 0.7
K-site 1	-0.6 ± 1.0	-0.5 ± 1.1	-0.5 ± 1.0	0.4 ± 0.7	4.7 ± 0.7
K-site 3	-4.0 ± 1.0	-3.3 ± 1.4	-3.6 ± 1.2	2.6 ± 0.9	6.9 ± 0.9

[a] With the HP2 gate open and the hydrated E374 side-chain.

substrate in the binding site. On the other hand, the system with glutamate bound and a protonated E374 is very stable, and also produces a high value for the pK_a of this residue. Therefore, we propose that proton and substrate binding are mutually coupled, that is, substrate binding brings about the protonation of E374 and vice-versa. This mechanism is corroborated by experiments, which show that in the pH range from 6 to 10, saturation of glutamate binding can be achieved by either increasing the glutamate concentration or decreasing the pH in the extracellular media[145].

An interesting observation is that the pK_a value is 11.7 for both apo states, indicating a protonated E374 residue. We attribute this to the negative potential induced by the negative charges of the D368 and the D455 residues, which are 10 and 15 Å away from E374, respectively. Even though there are water molecules between these residues and E374, they are mostly limited to the surface of the protein, and hence provide a limited dielectric screening compared to bulk water (Fig. 7.6A). Presence of the Na1 and Na3 ions neutralizes these negative charges, thereby explaining the reduction in the pK_a value of E374 when the Na1 and Na3 ions are bound.

We next look at the release of the proton into the intracellular media, and how this is related to K^+ binding. Similar to the outward state, this proton is released when we remove the substrate and the Na2 ion, but it is bound if the transporter is empty because the pK_a of E374 is 11.7 in the inward-apo state (Table 7.3). To investigate the influence of K^+ binding on the proton release, we have also calculated the protonation state of E374 in the presence of K^+ in each of the three sites considered in our previous study[161]. For K^+ site number 2, which is coordinated by the same E374 side-chain, K^+ is not stable in the binding site and leaves it shortly after the protonation of this residue (not shown). Therefore, the deprotonation of E374 is absolutely necessary for K^+ binding at this site. In the case of K^+ sites 1 and 3, we obtain pK_a = 4.7 when K^+ is bound in site 1 and pK_a = 6.9 when K^+ is bound in site 3 (Table 7.3). This is a direct consequence of introduction of a K^+ ion in the apo state, which either replaces the Na1 and Na3 sodium ions (site 1) or the substrate α-amino group (site 3). Interestingly, introduction of one K^+ ion at site 1 causes a larger shift in the pK_a value compared to the apo state than the presence of both Na1 and Na3 (4.7 vs. 6.7). This can be explained by the fact that the binding of K^+ at site 1 not only reduces the negative potential in that region, but it also helps to hydrate the D368 side-chain, which does not happen in the apo transporter (Fig. 7.6A and 7.6B). In the case of site 3, the D444 side-chain is part of the substrate binding site (coordinates its α-amino group), and is separated from E374 by the salt bridge between R447 and D440 (Fig. 7.6C). The presence of a positive charge in this region, instead of the negatively charged substrate or an empty binding site, is responsible for reducing the pK_a value of E374, deprotonating this residue when K^+ is bound at site 3. These results show that binding of a K^+ ion is necessary for the deprotonation of the E374 side-chain and the completion of the transport cycle.

7.5 Conclusions

In conclusion, we have performed rigorous pK_a calculations using the MD/FEP method to identify protonated residues in the human glutamate transporter EAAT3, and also to elucidate the mechanism of proton transport. Our calculations show that E374 is the only protonated residue in the fully bound outward-facing state, which agrees with experiments that point to this residue as the proton carrier in EAATs. We find that both the presence of the substrate and the closure of the HP2 gate contribute to the high pK_a value obtained in our calculations. A hydrophobic region formed by conserved residues in the TM7 and HP2 segments isolates E374 from the solvent, favoring the presence of a neutral side-chain as opposed to a negatively charged one. When we remove the substrate and the Na2 ion, we see the opening of the HP2 gate and the hydration of the E374 side-chain, which now has a pK_a value below the physiological pH and therefore should be mostly deprotonated. If we remove the remaining Na^+ ions, the negative charges from the D368 and D455

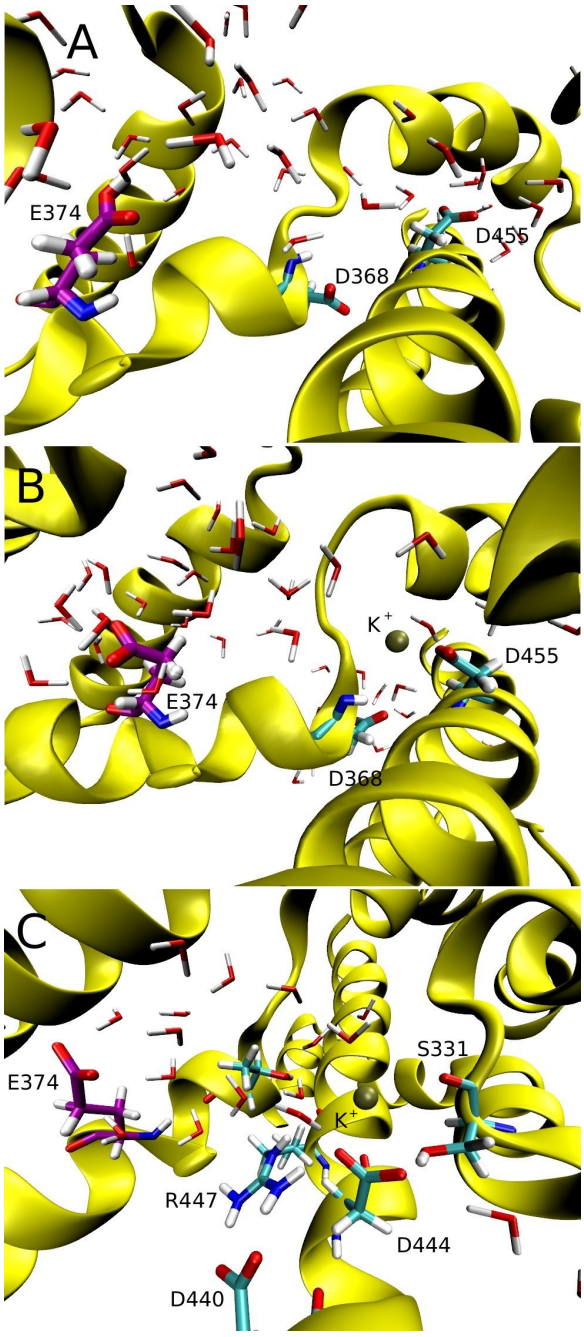

Figure 7.6: (A) The region between E374 (purple) and the D368 and D455 side-chains, which has water molecules located mostly on the surface of the protein, resulting in reduced dielectric screening. (B) The same area but now in the presence of the bound K^+ at site 1. We see that, in addition to the presence of the positively charged K^+ ion, the side chain of D368 is also hydrated. (C) The substrate binding site with a K^+ ion bound at site 3, which now has a positive charge instead of the negatively charged substrate, reducing the pK_a of E374.

side-chains induce the protonation of E374, which now has a pK_a of ∼12. There is good agreement between the results from the apo states created from the fully bound Glt_{Ph} and the apo Glt_{Tk} templates, confirming the relevance of the former in apo state calculations.

To understand the mechanism of proton release, we perform the pK_a calculations in the inward-facing state of EAAT3 in the fully-bound state, with only Na1 and Na3 bound, in the apo state, and also with K^+ bound in three possible sites previously determined. There is good symmetry between the values in the inward and the outward conformations of EAAT3, with the inward apo state also having a protonated E374 carboxyl group. This result shows that the release of all the ligands and the hydration of E374 is not enough to deprotonate this residue, and therefore the proton release should depend on K^+ binding to the inward transporter. Indeed, the binding of the K^+ ion in any of the sites tested induces the deprotonation of E374, by either reducing its pK_a or directly coordinating this ion. This is consistent with a mechanism in which the K^+ ion replaces either the Na1 and Na3 ions, the substrate, or the proton itself, so the transporter never needs to reach the inward apo state. This is in contrast to Glt_{Ph}, in which all the ligands have to unbind for the relocation of the transporter to the outward state. This difference might explain the much faster transport rates of the human transporters compared to the archaeal ones, and also the need for the countertransport of a K^+ ion in the EAATs.

7.6 Acknowledgments

This work was supported by grants from the Australian Research Council. Calculations were performed using the HPC facilities at the National Computational Infrastructure (Canberra), and the Victorian Life Sciences Computation Initiative (Melbourne). We thank Rob Vandenberg and Renae Ryan for discussions on the structure and function of Glt_{Ph} and the mammalian glutamate transporters.

Chapter 8

Computation of Standard Binding Free Energies of Polar and Charged Ligands to the Glutamate Receptor GluA2

ABSTRACT
Accurate calculation of the binding affinity of small molecules to proteins has the potential to become an important tool in rational drug design. In this study, we use the free energy perturbation (FEP) method with restraints to calculate the standard binding free energy of five ligands (ACPA, AMPA, CNQX, DNQX and glutamate) to the glutamate receptor GluA2, which plays an essential role in synaptic transmission. To deal with the convergence problem in FEP calculations with charged ligands, we use a protocol where the ligand is coupled in the binding site while it is decoupled in bulk solution simultaneously. The contributions from the conformational, rotational and translational entropies to the standard binding free energy are determined by applying/releasing respective restraints to the ligand in bulk/binding site. We also employ the confine-and-release approach, which helps to resolve convergence problems in FEP calculations. Our results are in good agreement with the experimental values for all five ligands, including the charged ones which are often problematic in FEP calculations. We also analyze the different contributions to the binding free energy of each ligand to GluA2 and discuss the nature of these interactions.

8.1 Introduction

The mechanism of action of many pharmaceutical drugs rely on binding of small molecules to receptor sites with high affinity and specificity.[169,170] If we can find efficient ways to calculate the protein-ligand binding affinities accurately, computational methods could become a powerful tool in the discovery and optimization of lead molecules.[171] To achieve that goal, one needs both a well-parametrized force field and a suitable method for calculating binding free energies. Software packages such as Antechamber in AMBER[172] and CGenFF in CHARMM[173,174] provide missing force-field parameters for drug-like molecules so that they can be incorporated in simulations of biomolecules. These parameters can be further optimized by performing quantum mechanical calculations using programs such as Gaussian 09.[72] Binding free energies of ligands can be calculated in a number of ways, and generally, greater accuracy of the results comes at a larger computational cost. Simple empirical scoring functions can sieve through millions of compounds, which may be desirable in an industrial environment.[175] However, these simplifications often compromise the accuracy of the predictions made.[176] At the other end of the spectrum, we have the free energy perturbation (FEP) and potential of mean force (PMF) methods that use all-atom molecular dynamics (MD) simulations.[106,115,177–182] They provide the most accurate approach for the prediction of ligand binding affinities, but they are also very time consuming and computationally costly. Nevertheless, thanks to the consistent, rapid increase in computing power and improvements in computational methods, these MD methods are fast becoming viable alternatives for lead selection.

Depending on the ligand-receptor system, the FEP or the PMF method may be more appropriate for the calculation of binding affinities. An important consideration in this regard is the size and charge state of the ligand, which determine its hydration energy. For small, neutral molecules, the hydration energy can be calculated accurately using the FEP method.[183,184] Hence FEP has become the standard method for such molecules.[116,185–191] Here the ligand interactions are decoupled from the rest of the system in both the binding site and bulk solvent, and the free energy of binding is obtained from the difference of these transformations. The standard binding free energies—needed for comparison of the results to experimental binding constants—are obtained by including entropy terms that take into account the loss of translational, rotational and conformational freedoms of the ligand upon binding.[107,117–119,192–194]

For large and especially charged molecules, the FEP method has rarely been used due to convergence problems, and therefore, the PMF method has been the method of choice.[43,47,84,195–200] The PMF is determined along one or more reaction coordinates of the ligand with respect to the receptor, and the free energy of binding is obtained by integrating the PMF along these coordinates. Even though in PMF calculations the ligand does not need a clear pathway to the binding site,[194] these calculations become very difficult for ligands that are buried inside proteins. Such

a situation occurs in the closed state of ionotropic glutamate receptors, where the ligand is partially isolated from the solvent. One way to deal with this problem is to create an open state so as to establish a clear pathway for the ligand. This method was recently applied to binding of neurotransmitters to the ionotropic glutamate receptor GluA2[41]. While good results were obtained for the binding free energies of several ligands, in general, it would be desirable to perform the free energy calculations using directly the crystal structures. This is feasible using the FEP method, where only the end points are needed in the calculations. Therefore, it is worthwhile to examine the convergence problem in FEP, and see if it can be resolved using a different protocol than the one commonly used. In most FEP calculations, the decoupling of the ligand from the binding site and bulk solution are performed in separate simulations. By increasing the system size slightly, it is possible to perform both transformations simultaneously, that is, coupling the ligand in the binding site while decoupling it in bulk at the same time. For charged ligands, this approach also solves the neutrality problem so that one does not need to annihilate ions together with the ligand in order to keep the system neutral during the FEP calculations,[201] or include correction terms for the inserted charges.[202] This method was recently used in FEP calculations of aspartate binding to the transporter protein Glt_{Ph},[123,144] and free energy change due to a charge mutation in a bound toxin,[203] and was shown to yield consistent results with experiments.

The aim of this work is to perform extensive FEP calculations using the above protocol to determine the binding free energies of five ligands (ACPA, AMPA, CNQX, DNQX and glutamate) to the S1S2 ligand binding domain (LBD) of GluA2. Glutamate is the predominant excitatory neurotransmitter in the brain, and ionotropic glutamate receptors are essential for the fast synaptic transmission between nerve cells.[3] They work by coupling the binding of a ligand to the opening of an ion channel, which allows flow of cations through the membrane and transmission of a nerve impulse. This coupling is related to the closure of the clamshell-shaped LBD in the presence of agonists such as glutamate and AMPA. Antagonists such as CNQX and DNQX, on the other hand, bind to GluA2 but do not allow the closure of the LBD, thereby blocking receptor activation. From a biomedical point of view, it has been shown that malfunctioning of the GluA2 receptors is implicated in a range of diseases, including stroke, depression and Parkinson's disease.[204,205] Accurate calculation of the free energies involved in ligand binding and the closure of the LBD will help to understand the mechanism of synaptic transmission,[41] which will facilitate finding treatments for such diseases.

There are several other reasons for choosing GluA2 in this study. First, there are high-resolution crystal structures of GluA2 in the apo state and in complex with all the ligands tested,[206–209] so that any conformational changes in GluA2 due to their presence can be easily identified. Next, the experimental binding free energies of these ligands are available in the literature,[206–209] so we can test the accuracy of our calculations. Finally, there are computational studies that use umbrella sampling[41] and metadynamics[42] to obtain the ligand affinities, which can be directly compared

to our FEP results. In umbrella sampling simulations,[41] ∼13 ns was used to calculate the binding affinity of each ligand from a 1-dimensional PMF, plus ∼40 ns to calculate the conformational PMF of the receptor in each case. Much longer simulation times were used in Ref.[42] with 0.5 μs of metadynamics simulations for the opening of the receptor LBD and 13.5 μs for calculating the binding affinity each ligand to the open state.

In the following sections, we first discuss the theoretical background and the methodology used in our calculations. We then present the results and compare the binding free energies obtained using the FEP method to the experimental/computational values available in the literature. For all five ligands considered, the FEP results are found to be in good agreement with the ones obtained experimentally and through umbrella sampling simulations.

8.2 Standard Binding Free Energy

The standard binding free energy for binding of a ligand to a receptor is defined from the equilibrium constant K_{b} as

$$\Delta G^0_{\mathrm{b}} = -k_{\mathrm{B}}T \ln(C^0 K_{\mathrm{b}}), \tag{8.1}$$

where C^0 is the standard concentration of 1 mol/L = 1/(1661 Å^3). For computational convenience, ΔG^0_{b} can be decomposed as follows[106,177,178]

$$\Delta G^0_{\mathrm{b}} = \Delta\Delta G_{\mathrm{cf}} + \Delta\Delta G_{\mathrm{rot}} + \Delta\Delta G^0_{\mathrm{tr}} + \Delta\Delta G_{\mathrm{int}} + \Delta G_{\mathrm{close}}, \tag{8.2}$$

where the first three terms represent the free energy contributions due to the changes in the conformational, rotational and translational entropies of the ligand upon binding and the fourth term gives the interaction free energy difference for translocating the ligand from the bulk to the binding site. The last term in Eq. 8.2 represents the free energy change due to conformational rearrangements in the receptor upon ligand binding such as the closure of the glutamate receptor. Straightforward calculation of the interaction energy term without any restraints on the ligand runs into convergence problem because it is difficult to sample all possible states of the ligand in both the binding site and bulk solvent. This problem is resolved in the double decoupling method (DDM),[106,177,178] where restraints are imposed on the ligand while its interactions with the solvent/receptor are decoupled/coupled. The contributions from these restraints are taken into account separately by means of simulations or analytical methods.

The thermodynamic cycle we use in the FEP calculations is illustrated in Figure 8.1, which describes the binding of a ligand to the receptor site. The ligand is restrained in bulk using conformational, rotational and translational potentials denoted by u_c, u_r and u_t, respectively. The restrained ligand is then translocated from bulk to the binding site, followed by removal of the restraints in reverse order. Below we first describe the free energy changes associated with the application/release of the restraints followed by the interaction energy.

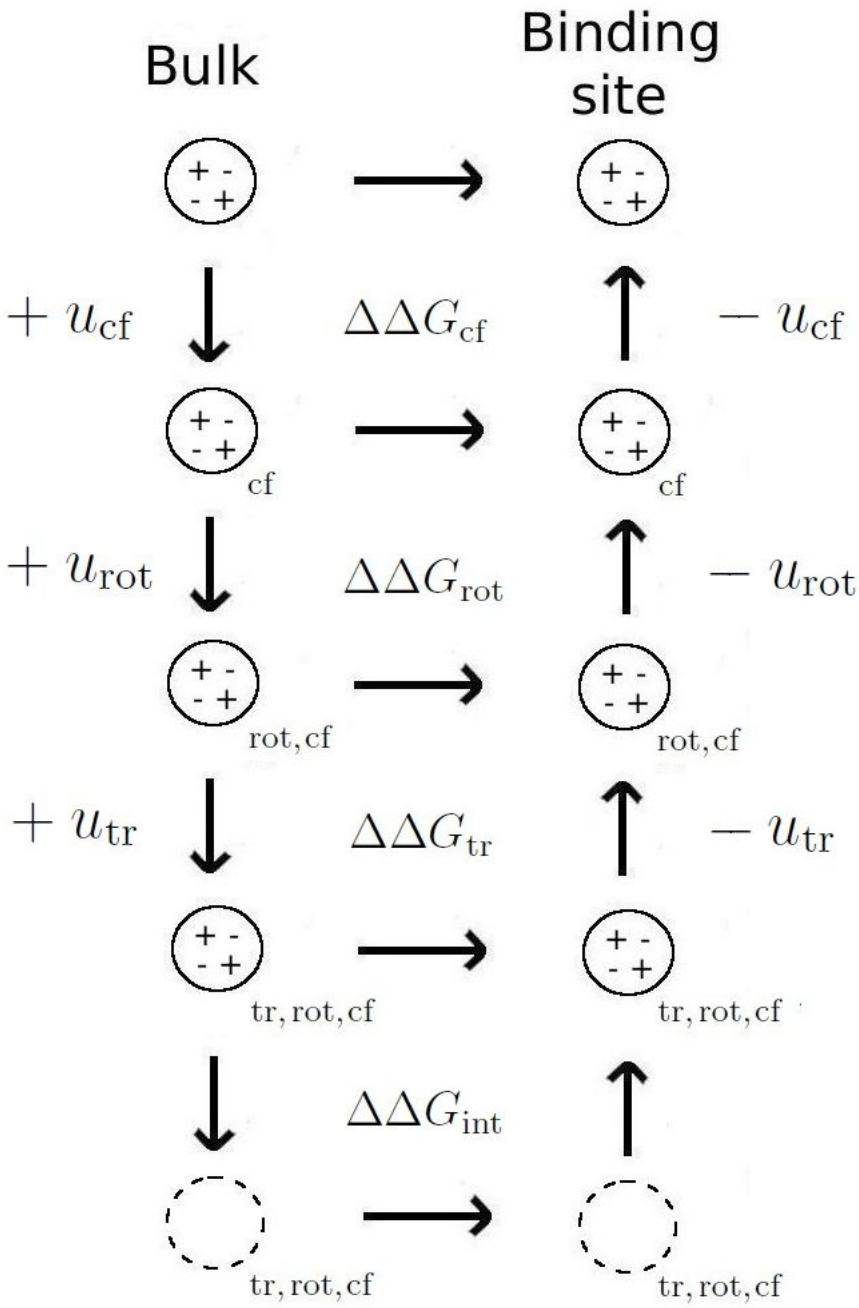

Figure 8.1: The thermodynamic cycle used in the FEP calculations, where the restraints are applied on the ligand in bulk solution and released while it is at the binding site.

8.2.1 Conformational Free Energy.

Conformational restraining of a ligand in bulk/binding site is achieved using the harmonic potential

$$u_c = \frac{k_c}{2n} \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{x}_{0i})^2, \tag{8.3}$$

with $k_c = 50$ kcal/mol/Å^2. Here $\{\mathbf{x}_i, i = 1, ..., n\}$ give the current positions of the ligand atoms after its position and orientation have been aligned with the reference state described by $\{\mathbf{x}_{0i}, i = 1, ..., n\}$. The reference state is chosen as the average conformation of the fully-interacting ligand in the bulk/binding site when no restraints are applied. The conformational free energy difference for applying the restraint in bulk and releasing it in the binding site is given by (see Figure 8.1)

$$\Delta\Delta G_{\rm cf} = \Delta G_{\rm cf}^{\rm bulk} - \Delta G_{\rm cf}^{\rm site} \tag{8.4}$$

The bulk/site contributions to $\Delta G_{\rm cf}$ are determined in separate FEP calculations where k_c is relaxed from 50 kcal/mol/Å^2 to 0 in ten steps given by {50, 30, 15, 8, 4, 2, 1, 0.5, 0.2, 0.1, 0}. Each FEP window is equilibrated for 300 ps followed by 800 ps of production data collection, and the free energy is obtained via numerical integration of the window values.

8.2.2 Rotational and Translational Free Energies.

In recent calculations of the standard free energy of ligands, the translational and rotational freedom of the ligand are restrained relative to the protein using anchoring atoms in the ligand and the protein.[116,188–191] This can be done in a number of ways, but it is essential that these anchors are chosen properly.[210] Here we use instead the Body Restraint Algorithm,[192] where the absolute coordinates are employed for both the position and orientation of the ligand, even though these quantities must be determined relative to the protein in order to obtain accurate results. We circumvent this problem by strongly restraining the orientation and the position of the center of mass (COM) of our receptor, such that these restraints are much larger than the restraints applied on the ligand. We stress that the strong restraints applied on the receptor affect only its position of the COM and overall orientation but not its internal degrees of freedom.

The orientational restraints are applied using the quaternion formalism: for a given ligand orientation $\mathbf{q}$ and its average orientation $\mathbf{q_0}$, the restraining potential is given by $u_r = (k_r/2)\Omega^2$, where $\Omega = \cos^{-1}(\mathbf{q} \cdot \mathbf{q}_0)$ and $k_r = 300$ kcal/mol/rad^2 is employed. Similarly, a harmonic potential of the form $u_t = (k_t/2)(\mathbf{r} - \mathbf{r}_0)^2$ is used to restrain the COM of the ligand in the binding site with $k_t = 10$ kcal/mol/Å^2. The average orientation $\mathbf{q_0}$ and the average position $\mathbf{r_0}$ are determined from restraint-free MD simulations of the ligand in the binding site. The magnitude of the spring constants k_r and k_t are chosen according to the fluctuations of the associated coordinates, namely, $k_t \sim 3k_{\rm B}T/\langle(\mathbf{r} - \mathbf{r}_0)^2\rangle$ and $k_r \sim 3k_{\rm B}T/\langle\Omega^2\rangle$. The force constants

applied to the position and orientation of the protein are about a hundred times larger than the ones applied on the ligand. This ensures that the protein can be considered fixed in space compared to the fluctuations of the ligand, and hence we can safely use the absolute coordinates of the ligand in the calculations.

The free energy differences associated with the application/release of the rotational and translational restraints in bulk/binding site are given by

$$\Delta\Delta G_{\text{rot}} = \Delta G_{\text{rot}}^{\text{bulk}} - \Delta G_{\text{rot}}^{\text{site}}, \tag{8.5}$$

$$\Delta\Delta G_{\text{tr}}^{0} = \Delta G_{\text{tr}}^{\text{bulk}} - \Delta G_{\text{tr}}^{\text{site}} \tag{8.6}$$

The binding site contributions to both terms are determined by performing FEP calculations similar to those of the conformational case. For $\Delta G_{\text{rot}}^{\text{site}}$, k_r is relaxed from 300 kcal/mol/rad^2 to 0 in eleven steps given by {300, 200, 150, 80, 50, 30, 15, 8, 4, 2, 1, 0}. Similarly for $\Delta G_{\text{tr}}^{\text{site}}$, k_t is relaxed from 10 kcal/mol/Å^2 to 0 in eight steps given by {10, 8, 5, 2, 1, 0.5, 0.2, 0.1, 0}. In both FEP calculations, each window is equilibrated for 300 ps followed by 800 ps of data collection. The bulk contributions can be determined analytically using the rigid-rotor approximation, which works well for conformationally restrained ligands[116]

$$\begin{aligned}\Delta G_{\text{rot}}^{\text{bulk}} &= -k_{\text{B}}T \ln\left[\frac{1}{8\pi^2}\int e^{-u_r(\phi_1,\phi_2,\phi_3)/k_{\text{B}}T} d\phi_1 d\phi_2 d\phi_3\right] \\ &= -k_{\text{B}}T \ln\left[\frac{1}{8\pi^2}\left(\frac{2\pi k_{\text{B}}T}{k_r'}\right)^{3/2}\right],\end{aligned} \tag{8.7}$$

$$\begin{aligned}\Delta G_{\text{tr}}^{\text{bulk}} &= -k_{\text{B}}T \ln\left[\frac{1}{V_0}\int e^{-u_t(\mathbf{r})/k_{\text{B}}T} d^3r\right] \\ &= -k_{\text{B}}T \ln\left[\frac{1}{V_0}\left(\frac{2\pi k_{\text{B}}T}{k_t}\right)^{3/2}\right],\end{aligned} \tag{8.8}$$

where $V_0 = 1661$ Å^3 is the reference volume for the standard concentration. In obtaining Eq. 8.7, we have converted the quaternion representation to the three axes of rotation assuming that they are not coupled, which is valid for small rotations. Therefore k_r' in Eq. 8.7 represents the spring constant applied to the three rotation axes.

It is of interest to compare the rotational and translational free energy differences obtained from the application of restraints with the corresponding ones obtained directly from the rotational and translational entropy differences for binding of a free ligand to a receptor site[107,117–119,192]

$$\Delta\Delta G_{\text{rot-ent}} = -k_{\text{B}}T \ln\left[\frac{(2\pi e)^{3/2}\sigma_{\phi 1}\sigma_{\phi 2}\sigma_{\phi 3}}{8\pi^2}\right], \tag{8.9}$$

$$\Delta\Delta G_{\text{tr-ent}}^{0} = -k_{\text{B}}T \ln\left[\frac{(2\pi e)^{3/2}\sigma_x\sigma_y\sigma_z}{V_0}\right]. \tag{8.10}$$

Here σ_x, σ_y, and σ_z, are the principal rms fluctuations of the COM of the ligand in the binding site, and $\sigma_{\phi 1}$, $\sigma_{\phi 2}$, and $\sigma_{\phi 3}$, are the principal rotational rms fluctuations of the ligand calculated using the quaternion representation.[119] The various σ values are calculated by performing 6 ns of MD simulations of the unrestrained ligand in the binding site and diagonalizing the translational and rotational covariance matrices obtained from the trajectory data.[118,119] The MD simulations show that the six degrees of freedom associated with the positions and orientations of the ligands in the binding site are well-described by normal distributions so that the entropic contributions to the binding free energy can be determined accurately using Eqs. 8.9 and 8.10. Although Eqs. 8.5 and 8.6 are formally different from Eqs. 8.9 and 8.10, we will show in the Results that they lead to very similar results when the restraining potentials are chosen appropriately.

8.2.3 Interaction Free Energy.

The free energy of translocation, that is, decoupling the restrained ligand in bulk and coupling it in the binding site, is given by

$$\Delta\Delta G_{\rm int} = \Delta G_{\rm int}^{\rm site} - \Delta G_{\rm int}^{\rm bulk}. \tag{8.11}$$

Here we perform the binding site and bulk calculations simultaneously and in the same system, which helps to maintain neutrality of the system for charged ligands.[123,144] Performing these calculations separately could lead to substantial errors due to the large solvation free energies associated with charged ligands.[43] In the present approach, the electrostatic part of the alchemical transformation can be seen as a transfer of charge between two regions of the same system. This reduces the errors arising from the large solvation energies of charged ligands as well as potential artifacts from use of the periodic boundary conditions.

The free energy of translocation in Eq. 8.11 is calculated using the FEP and thermodynamic integration (TI) methods.[180] In both methods, one introduces a hybrid Hamiltonian $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, where H_0 represents the Hamiltonian for the initial state and H_1 for the final state of the system. In the FEP method, the interval between $\lambda = 0$ and 1 is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average

$$\Delta G_i = -k_{\rm B}T \ln\langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_{\rm B}T]\rangle_{\lambda_i}. \tag{8.12}$$

The free energy of translocation is then obtained from the sum, $\Delta G_{\rm int} = \sum_i \Delta G_i$. In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, is obtained at several λ values, and the free energy of translocation is calculated from the integral

$$\Delta G_{\rm int} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda. \tag{8.13}$$

This integral is evaluated using a seven-point or nine-point Gaussian quadrature, which has been shown to provide a good method for free energy calculations involving

translocation of charged ligands.[123,144] In order to improve convergence and prevent instabilities in the FEP and TI calculations, decoupling/coupling of the ligand in bulk/site is performed in two stages by separating the Coulomb and Lennard-Jones (LJ) interactions

$$\Delta\Delta G_{\text{int}} = \Delta\Delta G_{\text{elec}} + \Delta\Delta G_{\text{LJ}}. \tag{8.14}$$

In the calculations, first the electrostatic interactions of the ligand are switched off/on in bulk/site followed by decoupling/coupling of the LJ interactions in bulk/site. We also use a soft-core potential for the LJ interactions to avoid problems at the end-points of the simulations.[82]

In the FEP calculations, it is important to choose the number of windows such that the ΔG_i value in each window is not much larger than 1.5 kT.[180] For the electrostatic terms in CNQX and DNQX, 66 windows are found to be sufficient while for ACPA, AMPA and glutamate—which carry charged groups—100 windows are required in order to obtain good convergence. Each window is equilibrated for 80 ps followed by a 120 ps of production run. In the case of the LJ interactions, 60 λ values are used for all ligands with smaller intervals close to the end-points. Each window is equilibrated for 80 ps followed by 80 ps of data collection. The windows for the TI calculations are adapted from those created during the FEP simulations. For the electrostatic terms, seven windows are used with 300 ps of equilibration and either 1 ns (CNQX and DNQX) or 1.5 ns (ACPA, AMPA and glutamate) of data collection. For the LJ terms, seven windows are used for CNQX and DNQX and nine windows for ACPA, AMPA and glutamate. In this case each window is equilibrated for 200 ps followed by 1 ns of production. The statistical errors in the TI calculations are estimated by performing block data analysis, using blocks of 100 ps. This does not take into account any systematic errors arising from the force field used or the asymmetry between the insertion and deletion of particles.[211] An estimate for the latter can be obtained from hysteresis between the forward and backward calculations.

To assess the convergence of the FEP and TI calculations, it is important to perform both the forward and the backward transformations, and check for consistency between the two results. A large hysteresis may be an indication of an energy barrier that cannot be crossed within the time scales of individual windows, and the results will not be reliable. In such situations, TI may provide a better option than FEP due to longer sampling of a smaller number of windows. Here we look for agreement between the forward and backward transformations in both FEP and TI calculations, and use it as evidence that our calculations are properly converged. The interaction binding free energy is determined from the average of the forward and the negative of the backward transformation results. In many studies, the Bennett acceptance ratio method[212,213] is used to obtain the optimal value of ΔG_i. This method uses an ensemble of forward and backward transformations at each window, in order to get the maximum likelihood estimator for the free energy difference. Here we perform the whole transformation in both directions, and the sigma values are based only on the TI calculations. Nevertheless, we can still obtain an estimate

of the boundary values for the FEP calculations from the forward and backward results. According to the Gibbs-Bogoliubov inequality,[212] the optimal value of the free energy difference lies between the values of the forward and the backward calculations. The values we obtain using TI are mostly within the range specified by FEP, which indicates consistency between the FEP and TI methods.

8.2.4 Confine-and-Release

Conformational changes in a protein that take place upon binding/unbinding of a ligand can sometimes affect convergence and prevent obtaining accurate results using FEP. This has been the case in the FEP calculations of the electrostatic terms for AMPA and glutamate, where a large hysteresis is observed between the forward and backward transformations. The problem is traced to the formation of a salt bridge between the ligand's α-amino group and the E705 side chain in GluA2. To resolve this problem we use the confine-and-release method,[214] which is a FEP/PMF hybrid—FEP is used to calculate the binding free energy of the ligand to the confined protein, and the free energy of confinement is obtained from the PMF of the confining coordinate. The interaction free energy can be written as

$$\Delta\Delta G_{\mathrm{int}} = \Delta\Delta G_{\mathrm{int,C}} + \Delta G_{\mathrm{con}} + \Delta G_{\mathrm{rel}}. \tag{8.15}$$

where $\Delta\Delta G_{\mathrm{int,C}}$ represents the free energy of translocation of the restrained ligand to the confined protein, and ΔG_{con} and ΔG_{rel} are, respectively, the free energies of confining the protein in the apo state and removing these restraints in the bound state. These are calculated from

$$\Delta G_{\mathrm{con}} = -\Delta G_{\mathrm{rel}} = -k_{\mathrm{B}}T \ln \left(\frac{\sum\limits_{i \in S_\alpha} e^{-F_i/k_{\mathrm{B}}T}}{\sum\limits_{all} e^{-F_i/k_{\mathrm{B}}T}} \right), \tag{8.16}$$

where S_α are the bins in the confined region, and F_i are the free energies for each bin in the apo states for ΔG_{con} and the bound states for ΔG_{rel}. These are computed by performing umbrella sampling simulations for the chosen confinement coordinate(s) and constructing the associated PMF using the weighted histogram analysis method.[81]

The confine-and-release approach can be related to the method used in Refs.[43,194], where the conformation of the ligand is confined and then released in the unbound and bound states. Here we are interested only in the conformational changes in the receptor that are related to the salt bridge in question. Using an overall conformational ensemble as in Refs.[43,194] would take much longer to converge, which rationalizes our choice for a simpler reaction coordinate that is sufficient to resolve the large hysteresis observed in the FEP calculations. The simplest choice for this confinement coordinate is the distance between the COM of the α-amino group of the ligand and the COM of the two oxygens in the E705 side chain. This requires

construction of a 1-D PMF, and therefore computationally less costly than restraining the two dihedrals of E705, which would require construction of a 2-D PMF. For both the apo and the bound states, we use 14 windows between the distances 2.5 Å and 9.0 Å spaced 0.5 Å apart. Note that the restrained, non-interacting ligand is still present in the apo state though it has no effect on the protein. The chosen range is sufficient to cover all the possible distances between these two groups as values outside this range are forbidden due to the limited flexibility of the E705 side-chain or steric clashes with the neighbouring groups. The initial windows for umbrella sampling simulations are created from 2 ns of steered MD. A spring constant of 8 kcal/mol/Å^2 is employed, and each umbrella window is equilibrated for 2 ns followed by 3 ns of data production.

8.3 Model System and Simulation Details.

The simulation systems are prepared using the crystal structures of GluA2 S1S2 ligand binding domain in complex with five ligands, namely, glutamate (PDB ID:1FTJ), AMPA (1FTM), ACPA (1M5E), CNQX (3B7D), DNQX (1FTL), and the apo state of GluA2 (1FTO). The structures of the five ligands are shown in Figure 8.2. In all cases only one monomer (chain A) is used. All simulation boxes are created using the software VMD.[104] The apo protein or complex are solvated in a box of dimensions $80 \times 80 \times 70$ Å^3 with about 12,000 water molecules. This size is chosen to ensure that during the decoupling of the ligands in bulk, they do not interact with the protein (the ligands are more than 25 Å away from the binding site). To create a 0.15 M NaCl solution and maintain the charge neutrality, we add to the solution 34 Na^+ and 38 or 39 Cl^- ions, depending on the ligand bound. The total number of atoms in each system is about 40,000.

After energy minimization of each system, the restraints on the non-hydrogen atoms of the protein and ligand are gradually released during 5 ns of MD simulations. Each system is further simulated for 4 ns with no restraints applied. In order to compare our results to experimental values and to the PMF results,[41] we perform the FEP simulations in the same open state of the receptor as employed in Ref.[41]. This state is defined by introducing two distance restraints between groups of atoms in lobes S1 and S2, namely, ξ_1 which measures the distance between the COM of residues the 479–481 in S1 and the COM of the residues 654–655 in S2, and ξ_2 which measures the distance between the COM of the residues 401–403 in S1 and the COM of the residues 686–687 in S2. These distances are restrained at $(\xi_1, \xi_2) =$ (14.4 Å, 13.7 Å). We perform 2 ns steered MD to achieve this state from the initial equilibrated complex, while keeping the ligands bound to the R485 residue on S1 and restraining the protein RMSD separately on the S1 and S2 lobes, as described in Ref.[41]. After the steered MD is completed, we release the ligand and the RMSD restraints and simulate the resulting system for a further 5 ns. Throughout the following MD simulations and in all the FEP calculations, restraints are applied

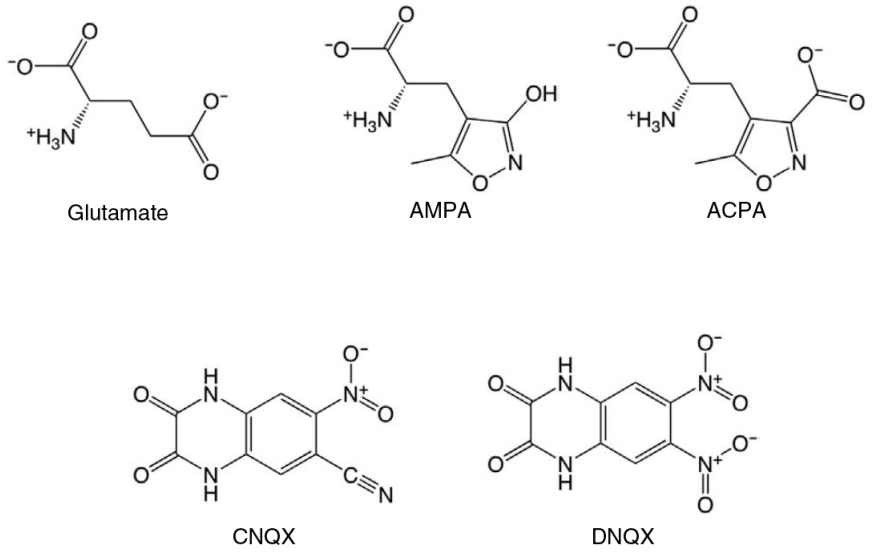

Figure 8.2: Structures of the five ligands considered in this study.

to the ξ_1 and ξ_2 distances keeping them at (ξ_1, ξ_2) = (14.4 Å, 13.7 Å). In order to compare the values obtained for the ligand affinities in the open state to the experimental values, we add the free energy of LBD closure obtained from Ref.[41] to the final value of the binding free energy of each ligand.

All of our MD simulations are performed using the NAMD package (version 2.9)[59] with the CHARMM36 force field.[70] Initial Parameters for ACPA, AMPA, CNQX and DNQX were obtained from the ParamChem server,[173,174] and parameters with penalties higher than 10 were optimized using Gaussian 09[72]. The parameters files for ACPA, AMPA, CNQX and DNQX are shown in the Appendix as Figures 8A.1, 8A.2, 8A.3 and 8A.4, respectively. The temperature of the simulation is maintained at 300 K using Langevin damping with a coefficient of 5 ps^{-1}, and the pressure is kept at 1 atm using the Langevin piston method with a damping coefficient of 20 ps^{-1}.[64] Periodic boundary conditions with the particle-mesh Ewald method are employed to calculate the electrostatic interactions without truncation. The Lennard-Jones (LJ) interactions are switched off between 10 and 12 Å using a smooth switching function. A time step of 2 fs is used in all MD simulations.

8.4 Results and discussion

The results of the FEP/TI calculations are summarized in Table 8.1, and the standard binding free energies obtained are compared to those determined from umbrella sampling[41] and experiments.[206–208] The electrostatic and LJ contributions to the interaction free energy have been calculated using both the FEP and TI methods, and also in both the forward and backward directions to check for any hysteresis effects. The two methods yield very similar results, therefore only the TI results are shown

in Table 8.1 as they allow for longer sampling and also estimation of the errors by block data analysis. As mentioned in the Methods section, all the calculations are performed in the same open state of the receptor as in Ref.[41]. This enables a direct comparison of the computational results obtained using the FEP/TI methods here with those determined from the PMF method. Also using the free energy of receptor closure from Ref.[41] allows us to determine the final standard binding free energy of the ligands, which can be compared with experiments. We first discuss CNQX and DNQX, which have similar structures and give very similar results, followed by discussions of the more challenging cases of ACPA, glutamate and AMPA.

8.4.1 CNQX and DNQX

CNQX and DNQX are antagonists whose binding to the receptor prevents closure of the LBD. Consistent with these properties, both ligands have substantially larger

Table 8.1: The Standard Binding Free Energies of the Five Ligands Considered in This Study. The free energy changes due to conformational, rotational and translational restraints are given in rows 1,2 and 3, respectively. The confine-and-release results (Eq. 8.16) for glutamate and AMPA are given in rows 4-5. The electrostatic and LJ contributions to the interaction free energy differences obtained from TI calculations are listed in rows 6-7. The binding free energies of the ligands to the open state of the receptor (sum of the rows 1–7) are given in row 8. The standard binding free energies in row 10 are obtained by adding the free energies of closure from Ref.[41] (row 9) to the open state results in row 8. The standard binding free energies obtained in this work (row 10) are compared to those determined from umbrella sampling (row 11) and experiments (row 12). The rotational free energies of DNQX (marked with a *) include a factor of $-k_{\mathrm{B}}T\ln(2) = -0.41$, due to the two-fold symmetry of this ligand. All energies are in kcal/mol.

	CNQX	DNQX	ACPA	Glu	AMPA
$\Delta\Delta G_{\mathrm{cf}}$	0.2	0.6	0.9	0.3	0.7
$\Delta\Delta G_{\mathrm{rot}}$	4.4	3.2*	4.5	3.9	4.6
$\Delta\Delta G^0_{\mathrm{tr}}$	3.8	3.6	4.1	3.8	4.0
ΔG_{con}	-	-	-	0.5	1.1
ΔG_{rel}	-	-	-	-2.3	-7.9
$\Delta\Delta G_{\mathrm{elec}}$	-7.3 ± 0.6	-6.9 ± 0.6	-10.6 ± 1.1	-5.5 ± 1.6	-4.8 ± 1.1
$\Delta\Delta G_{\mathrm{LJ}}$	-7.0 ± 1.3	-7.8 ± 1.4	-0.8 ± 0.7	0.9 ± 0.9	-1.0 ± 1.3
$\Delta G^0_{\mathrm{b}} - \Delta G_{\mathrm{close}}$	-5.9 ± 1.4	-7.3 ± 1.5	-1.9 ± 1.3	1.6 ± 1.8	-3.3 ± 1.7
$\Delta G_{\mathrm{close}}$[41]	-1.7	-1.6	-10.4	-10.1	-8.3
ΔG^0_{b}	-7.6 ± 1.4	-8.9 ± 1.5	-12.3 ± 1.3	-8.5 ± 1.8	-11.6 ± 1.7
ΔG^0_{b} (US)[41]	-7.7	-7.9	-11.2	-9.1	-10.1
ΔG^0_{b} (exp)	-8.5[206]	-8.2[207]	-10.6[208]	-8.4[207]	-10.8[208]

affinities for the open state of the receptor compared to the agonists ACPA, glutamate and AMPA (row 8 in Table 8.1), and much smaller free energies of closure (row 9). The main aim of this work is to demonstrate the feasibility of using the FEP/TI methods in calculation of the translocation free energies in ligand binding. To this end, we study the quality and convergence of the free energy calculations for the electrostatic and LJ terms. In Figure 8.3, we present the results of the FEP calculations as a function of the parameter λ, and the results of the TI calculations as running averages plotted as a function of the production time. In both cases, the free energy difference $\Delta\Delta G$ is given by the end points in the plots. Comparison of the forward and backward transformation results shows that there is a minimal hysteresis effect in all cases. For example, in TI calculations the difference between the two sets is seen to be less than 1 kcal/mol in all cases except for $\Delta\Delta G_{\mathrm{LJ}}$ in CNQX where it is 1.5 kcal/mol. Also the FEP and TI results obtained from the average of the forward and backward transformations are consistent with each other. Comparison of the standard binding free energies with those obtained from umbrella sampling and experiments (Table 8.1) provides further support for the feasibility of the FEP/TI approach—for both CNQX and DNQX, the calculated values are within 1 kcal/mol of the umbrella sampling and experimental results.

Table 8.2: Comparison of the Entropic Contributions to the Free Energy Obtained With and Without Restraints. The free energy changes due to rotational and translational restraints from Table 8.1 are reproduced in rows 1 and 3. The free energy changes calculated from the rotational and translational entropies (Eqs. 8.9 and 8.10) are listed in rows 2 and 4 for comparison. All energies are in kcal/mol.

	CNQX	DNQX	ACPA	Glu	AMPA
$\Delta\Delta G_{\mathrm{rot}}$	4.4	3.2*	4.5	3.9	4.6
$\Delta\Delta G_{\mathrm{rot\text{-}ent}}$	4.0	3.6*	4.6	3.4	4.3
$\Delta\Delta G^0_{\mathrm{tr}}$	3.8	3.6	4.1	3.8	4.0
$\Delta\Delta G^0_{\mathrm{tr\text{-}ent}}$	3.4	3.4	3.9	3.6	4.0

In Table 8.2, we compare the results of free energy changes calculated from the rotational and translational entropies of an unrestrained ligand using Eqs. 8.9 and 8.10 with those obtained from application of the rotational and translational restraints as presented in Table 8.1. The two methods are seen to yield very close results. In fact, using the values obtained from Eqs. 8.9 and 8.10 in Table 8.1, one would get very similar results for the binding free energies. The entropy calculations for an unrestrained ligand using Eqs. 8.9 and 8.10 are much easier and computationally less demanding than the free energy calculations involving application of rotational and translational restraints to a ligand. Therefore, it is worthwhile to check whether this alternative approach can be justified. Because restraints have to be used in calculation of the translocation free energy of a ligand in Eq. 8.2, we

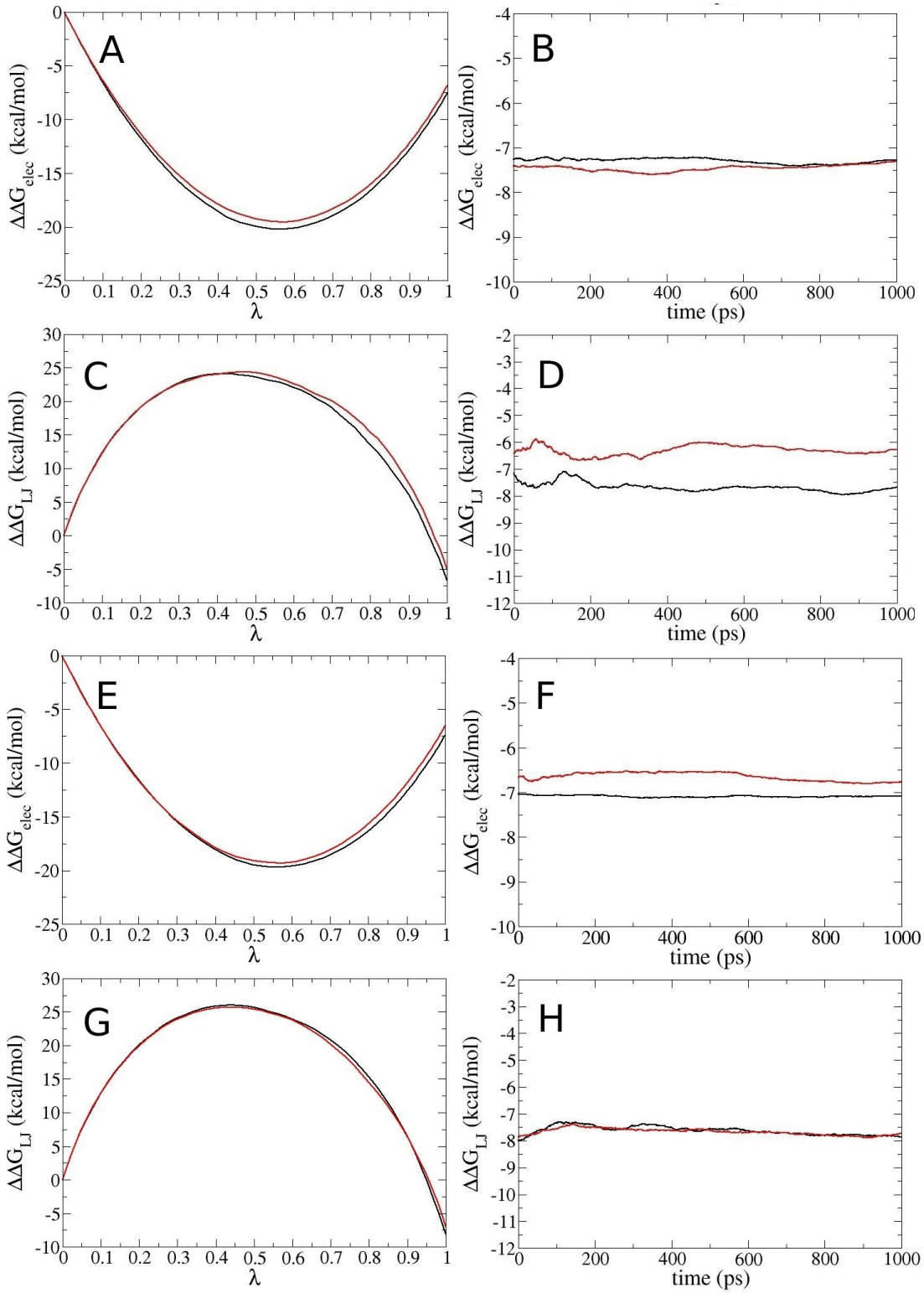

Figure 8.3: Plots of the FEP (first column) and TI (second column) calculations for the electrostatic (A, B) and LJ (C, D) interactions in CNQX, respectively. The graphs E-H show the same for DNQX. The black lines indicate the forward transformations and the red lines, the negative of the backward transformations.

need to show that the force-field interactions of the ligand in the binding site/bulk are not affected by restraining its position and orientation. We test this conjecture by performing FEP calculations for the electrostatic and LJ parts of DNQX using different sets of restraints. The results shown in Table 8.3 are all very similar indicating that for reasonable choices, these restraints have indeed little effect on the force field contributions to the binding free energy.

A distinguishing feature of the CNQX and DNQX results in Table 8.1 is the relatively large contribution of the LJ term to the binding free energy compared to the other ligands. This is directly responsible for the much higher affinities of CNQX and DNQX for the open state of the receptor, and therefore warrants a better understanding. The LJ term arises mostly from the hydrophobic interactions of the ligands with the receptor. Inspection of the binding modes of CNQX and DNQX shows that such an interaction occurs between the double-ring of CNQX/DNQX and the ring of the Y450 residue of the receptor as shown in Figure 8.4. These two structures are observed to be stacked stably on top of each other during the entire MD simulations, suggesting that this interaction is a major contributor to the high affinity of these two antagonists in the open state of the receptor.

8.4.2 ACPA

ACPA differs from CNQX and DNQX in that it has a net charge $-e$ but lacks the double ring structure. These differences help to explain the stronger electrostatic and weaker LJ interactions in ACPA relative to CNQX and DNQX (Table 8.1). Quality and convergence of the FEP and TI calculations for the electrostatic and LJ terms are studied in Figure 8.5. In all cases, the forward and backward results differ by less than 1 kcal/mol, indicating a minimal hysteresis effect. Also there is an almost perfect agreement between the FEP and TI results. Yet the calculated standard binding free energy deviates from the experimental value by 1.7 kcal/mol—the only case among the five ligands where this deviation is larger than 1 kcal/mol.

Table 8.3: Restraint Dependence of the Translocation Free Energy. Results of the FEP calculations for the electrostatic and LJ terms and their sum for DNQX using different translational and orientational restraints for the ligand. The free energy is in kcal/mol, k_t is in kcal/mol/Å^2 and k_r is in kcal/mol/rad^2.

Restraints	$\Delta\Delta G_{\text{elec}}$	$\Delta\Delta G_{\text{LJ}}$	$\Delta\Delta G_{\text{int}}$
$k_t = 5$, $k_r = 300$	-7.2	-7.2	-14.4
$k_t = 10$, $k_r = 100$	-6.6	-7.6	-14.2
$k_t = 10$, $k_r = 300$	-7.1	-7.3	-14.4
$k_t = 10$, $k_r = 500$	-6.7	-8.0	-14.7
$k_t = 25$, $k_r = 300$	-6.4	-7.0	-13.4

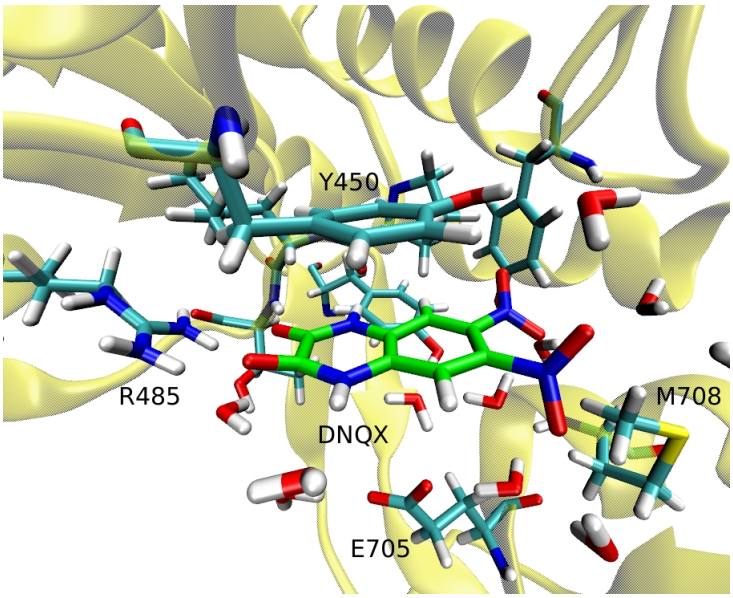

Figure 8.4: A snapshot of the DNQX complex in the open state of the receptor showing the hydrophobic interaction between the double ring of DNQX and the ring of the Y450 residue. The same pattern is also observed in the CNQX complex.

To see if this is due to some fluctuation, we have repeated the TI calculations but obtained essentially the same results. Another possibility is that the free energy of the receptor closure ΔG_{close} might be slightly overestimated, leading to the lower free energy value.

8.4.3 Glutamate and AMPA

As mentioned in Theory and Methods, large hysteresis effects have been observed in the FEP calculations of the electrostatic terms for glutamate and AMPA. Despite its similarity with AMPA, this does not happen for ACPA. To explain why the hysteresis effect occurs for glutamate and AMPA and not for ACPA, we show snapshots of these ligands in complex with the open state of the receptor (Figure 8.6). It is seen that, in the case of glutamate and AMPA, there is a salt bridge between the ligand's α-amino group and the E705 side chain of the receptor, but this does not occur in ACPA. Thus the problem is caused by this salt bridge and to circumvent it, we use the confine-and-release method as outlined in Theory and Methods.

For glutamate, we confine the distance between the ligand α-amino and the receptor E705 side-chain carboxyl groups to the range 4.5–6.5 Å by adding high energy barriers at these two points, which keeps the distance unbiased but confined within the specified range. This procedure breaks the salt bridge during the FEP simulations, and the contribution of this interaction is added later through the PMF term. The results of the FEP and TI calculations for the electrostatic and LJ terms,

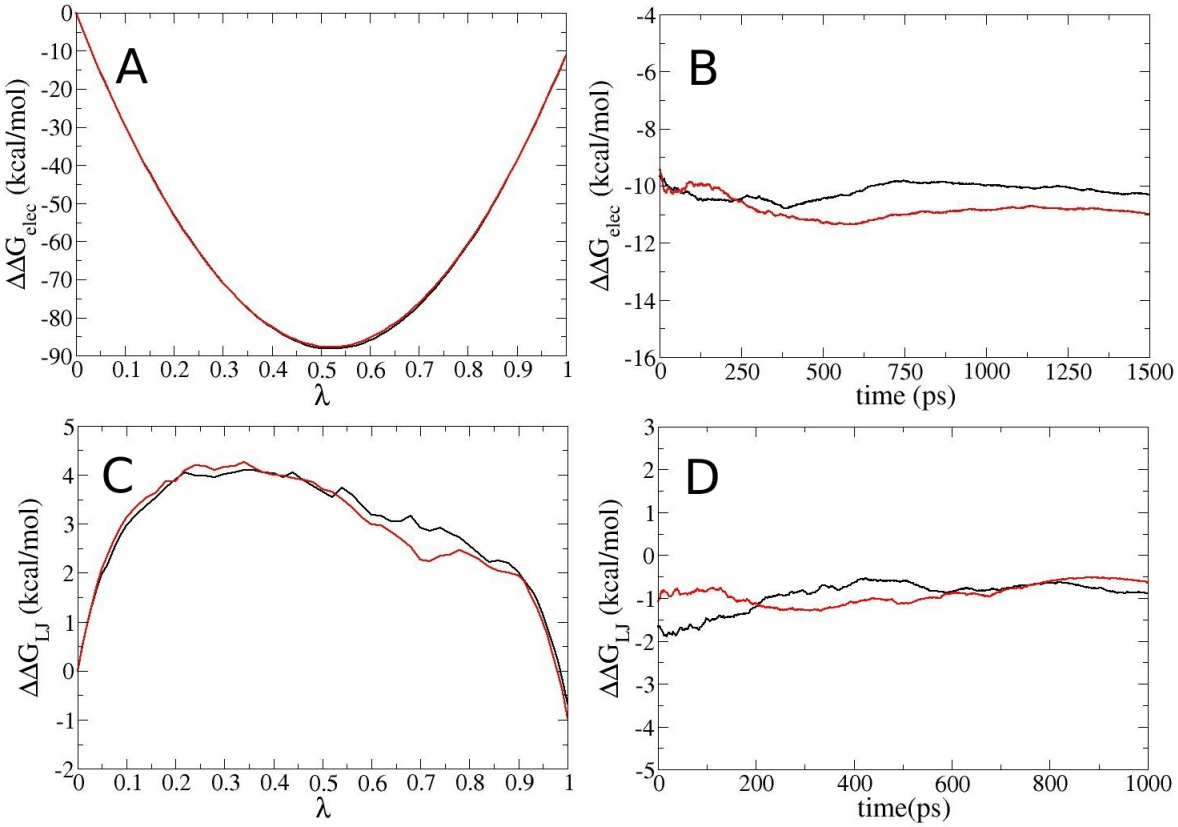

Figure 8.5: Same as Fig. 8.3 but for ACPA. Plots of the FEP (first column) and TI (second column) calculations for the electrostatic (A, B) and LJ (C, D) interactions.

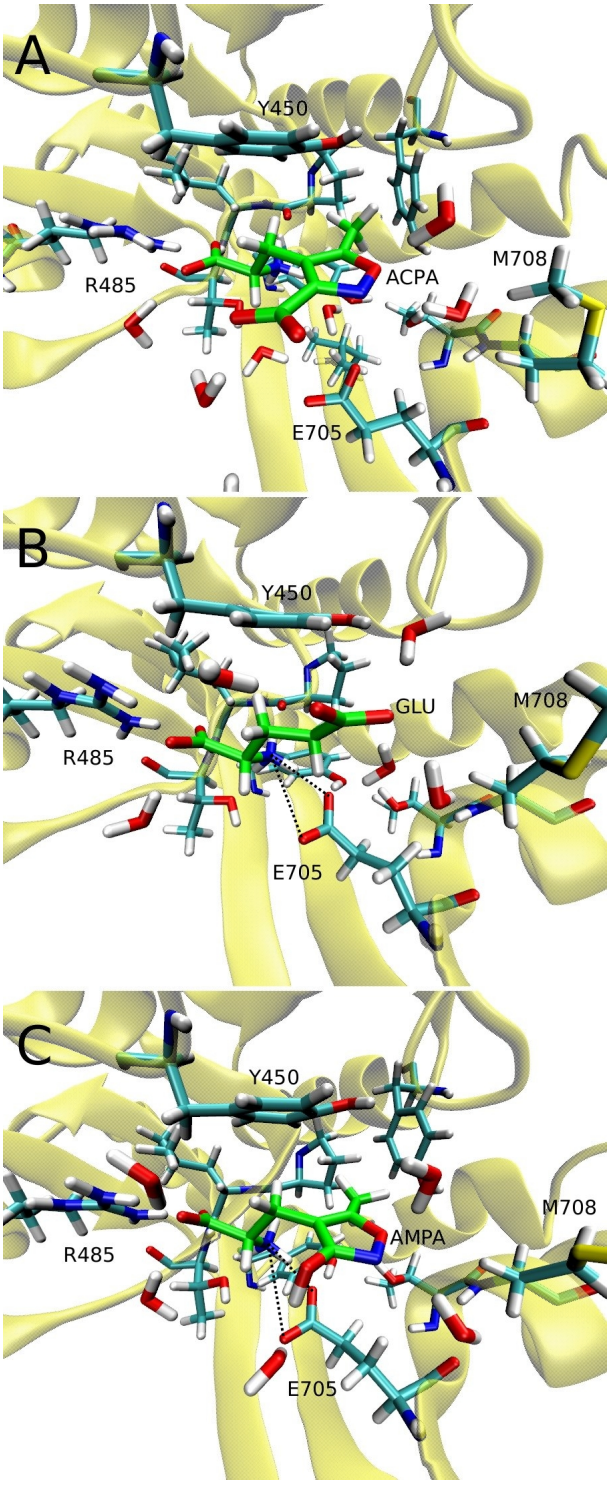

Figure 8.6: Snapshots of the ACPA, glutamate and AMPA complexes in the open state of the receptor. The critical E705 side chain has no interactions with ACPA but makes salt bridges with the α-amino group of glutamate and AMPA (indicated by dotted lines).

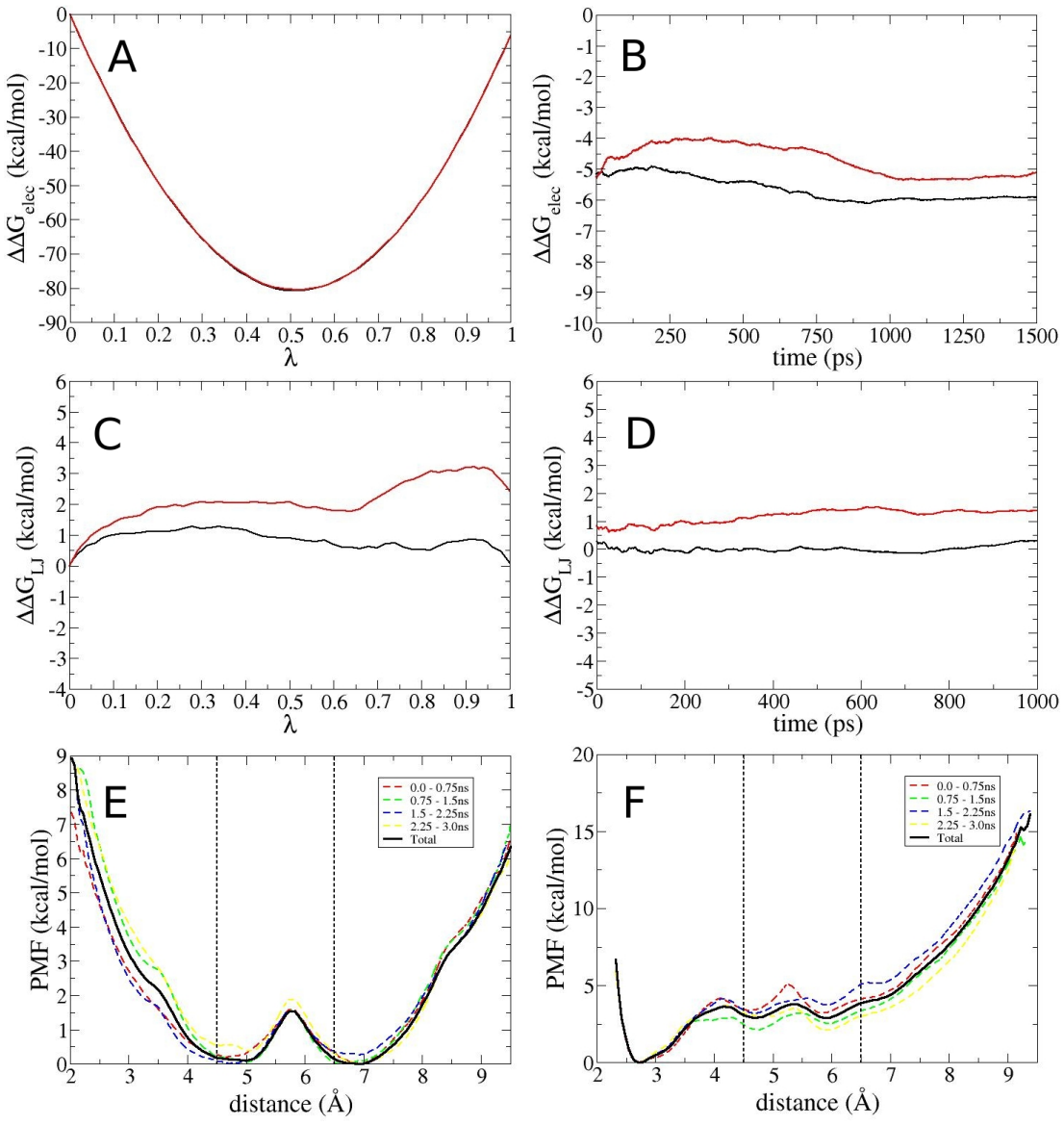

Figure 8.7: Plots of the FEP (first column) and TI (second column) calculations for the electrostatic (A, B) and LJ (C, D) interactions in glutamate. We also show the PMF as a function of the distance between the ligand α-amino and the receptor E705 side-chain carboxyl group in the apo (E) and bound (F) states of the receptor. The range of the confinement potential is indicated by the vertical dotted lines.

as well as the PMFs for the apo and bound states, are shown in Figure 8.7. Apart from the FEP result for the LJ term, there is minimal hysteresis between the forward and backward results. Because there is good agreement between the average FEP and TI results for the LJ term, we have not pursued this issue further. Convergence of the PMF results are demonstrated via block data analysis of the trajectory data (Figure 8.7E-F). The four PMF's constructed using 0.75 ns of data for each exhibit small fluctuations around the full PMF, which is taken as a sign of convergence. The standard binding free energy is in good agreement with the umbrella sampling and experimental results (Table 8.1), which shows that hysteresis problems in the FEP calculations can be overcome using the confine-and-release approach (Table 8.1). This is especially useful for calculations involving charged ligands because they often produce the largest conformational changes upon binding to a protein, and these changes can compromise the reliability of FEP calculations. By identifying the interactions causing problems and analyzing them separately, we can avoid such problems at the cost of some additional computations.

In the case of AMPA, using the same confinement range as in glutamate have still resulted in a large hysteresis in the FEP calculations. Therefore we have reduced the confinement range from 4.5–6.5 Å to 5.5–6.5 Å, further hindering the interaction between these two charged groups. As shown in Figure 8.8, this smaller range solved the hysteresis problem in the FEP calculations, yielding well-converged results in all FEP, TI, and PMF calculations. The standard binding free energy is within 1 kcal/mol of the experimental value but 1.5 kcal/mol smaller than the umbrella sampling result (Table 8.1). We note that this is the case among the five ligands where the difference between the FEP/TI and umbrella sampling results is largest. Another point to note is the relatively large contribution of the confine-and-release terms to the binding free energy of AMPA (rows 4-5 in Table 8.1). This is partly caused by the confinement potential which keeps the charged groups in the salt bridge further apart in AMPA compared to glutamate. Another reason can be gleaned from Figure 8.6. In the case of glutamate, there is some repulsive interaction between the β-carboxyl group and the E705 side-chain carboxyl, which reduces the effectiveness of the salt bridge. In AMPA, the carboxyl group is replaced by hydroxyl attached to the isoxazole ring, which does not hinder the formation of the salt bridge, rather it makes a hydrogen bond with E705. We can also see why the salt bridge does not form in ACPA—it is prevented by attachment of a carboxyl group instead of hydroxyl to the isoxazole ring. Thus the carboxyl group of E705 plays a critical role in binding of agonist ligands to the open state of GluA2.

8.5 Conclusions

In this study we have performed FEP/TI calculations of five polar and charged ligands using different techniques to obtain their standard binding free energies to the glutamate receptor GluA2. The FEP/TI calculations are performed by restraining

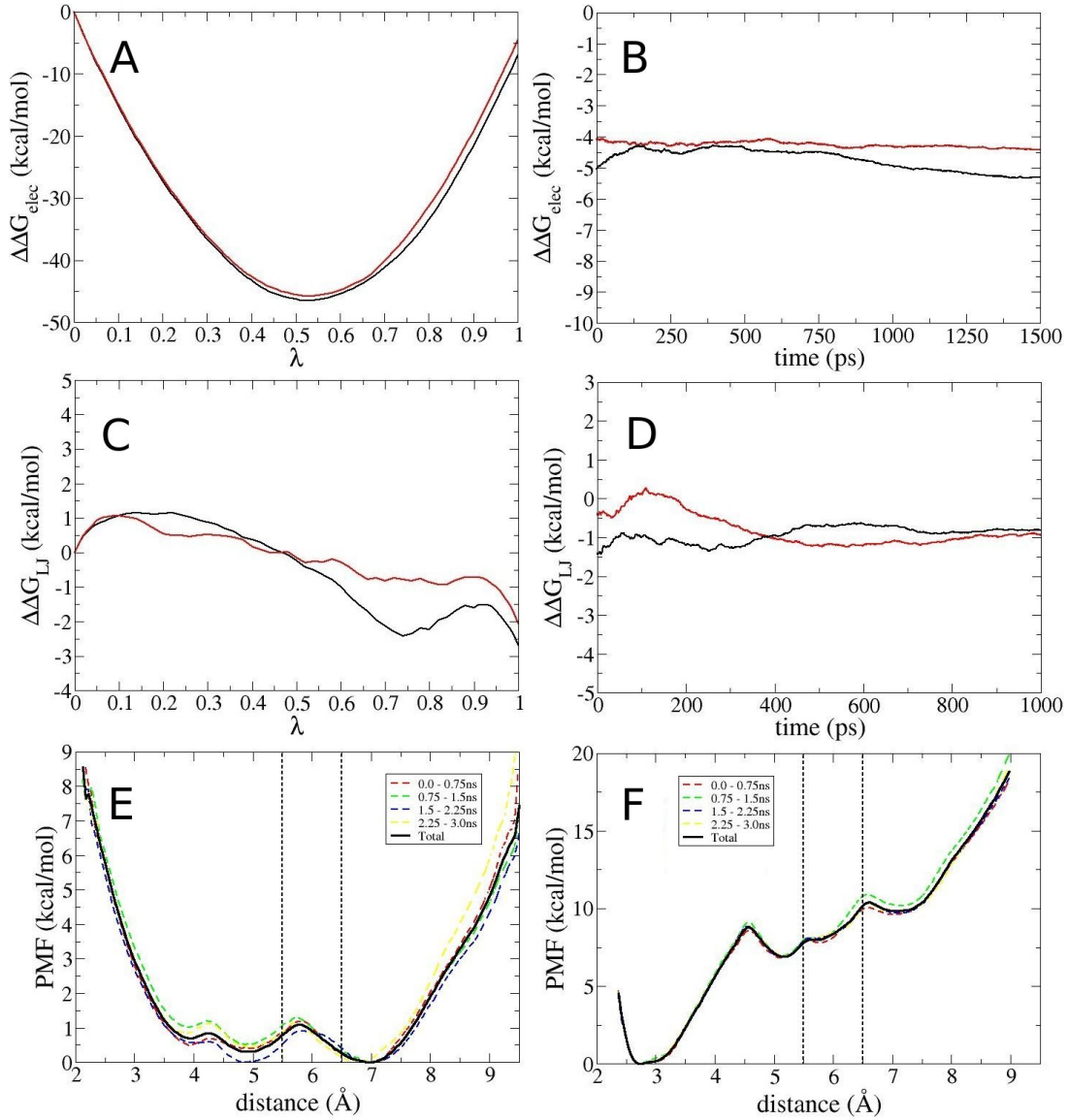

Figure 8.8: Same as Figure 8.7 but for AMPA. In this case, the range of the confinement potential is 5.5–6.5Å.

the position and orientation of ligands using the absolute coordinates, which do not require anchoring atoms in the protein and ligand, and therefore are easier to implement. It also allows a direct calculation of the entropy differences between the free and bound states of the ligand. We have shown that this approach produces very similar results to the double decoupling method, and considering its simplicity, may offer an alternative method for the calculation of the standard binding free energies. For glutamate and AMPA, for which the straightforward FEP calculations lacked good convergence, we have used the confine-and-release method, and obtained very good agreement with the experimental values. The results obtained in this study show that it is possible to calculate the binding free energies accurately using the FEP/TI methods even for ligands with net charges as long as the bulk and binding site transformations are performed in the same system. The FEP/TI methods, even though not suitable for protein-protein binding, are advantageous over the PMF method for binding sites that are deeply buried inside proteins. The techniques used here can be applied to many protein-ligand systems. The accuracy of the results means that they could, in particular, become very useful tools in rational drug design.

8.6 Acknowledgements

This work was supported by grants from the Australian Research Council. Calculations were performed using the HPC facilities at the National Computational Infrastructure (Canberra), and the Victorian Life Sciences Computation Initiative (Melbourne).

8.A Appendix for Chapter 8

```
RESI AM1           -1.000
GROUP         !    CHARGE
ATOM O1     OG2R50 -0.34
ATOM N2     NG2R50 -0.51
ATOM C3     CG2R52  0.30
ATOM C4     CG2R51 -0.21
ATOM C5     CG2R51  0.47
ATOM C31    CG2O3   0.67
ATOM O31    OG2D2  -0.74
ATOM O32    OG2D2  -0.74
ATOM C41    CG321  -0.08
ATOM CA     CG314   0.21
ATOM N      NG3P3  -0.30
ATOM CT     CG2O3   0.34
ATOM OT1    OG2D2  -0.67
ATOM OT2    OG2D2  -0.67
ATOM C51    CG331  -0.27
ATOM HT1    HGP2    0.33
ATOM HT2    HGP2    0.33
ATOM HT3    HGP2    0.33
ATOM H41    HGA2    0.09
ATOM H42    HGA2    0.09
ATOM HA     HGA1    0.10
ATOM H51    HGA3    0.09
ATOM H52    HGA3    0.09
ATOM H53    HGA3    0.09

BONDS
CG2O3  CG2R52  200.00     1.4600

ANGLES
CG2R52 CG2R51 CG321    45.80    130.00
CG331  CG2R51 OG2R50   25.00    112.00
CG2R52 CG2O3  OG2D2    40.00    116.00  50.0  2.3530
CG2O3  CG2R52 CG2R51   40.00    128.50
CG2O3  CG2R52 NG2R50   40.00    121.00


DIHEDRALS
OG2D2  CG2O3  CG2R52 CG2R51     0.8000  2   180.00
OG2D2  CG2O3  CG2R52 NG2R50     0.8000  2   180.00
CG2R51 CG2R51 CG2R52 CG2O3      2.3000  2   180.00
CG2O3  CG2R52 NG2R50 OG2R50     2.3000  2   180.00
CG2R52 CG2R51 CG2R51 CG331      2.6000  2   180.00
OG2R50 CG2R51 CG331  HGA3       0.0840  3     0.00
CG331  CG2R51 OG2R50 NG2R50     8.5000  2   180.00
CG321  CG2R51 CG2R51 CG331      0.6000  2   180.00
CG321  CG2R51 CG2R51 OG2R50     1.7000  2   180.00
CG321  CG2R51 CG2R52 CG2O3      2.0000  2   180.00
CG2R52 CG2R51 CG321  CG314      0.0000  1     0.00
CG2R52 CG2R51 CG321  CG314      0.7500  2     0.00
CG2R52 CG2R51 CG321  CG314      0.0000  3     0.00
CG2R52 CG2R51 CG321  HGA2       0.0000  3     0.00
CG321  CG2R51 CG2R52 NG2R50     4.0000  2   180.00

IMPROPERS
CG2O3  OG2D2  OG2D2  CG2R52    96.0000  0     0.00
CG2O3  X      X      CG314     96.0000  0     0.00

END
RETURN
```

```
          O31
           \\
            C31-O32(-)
           /
      N2==C3     HT1 HT2(+)
     /    |       \ /
    O1    | H41    N--HT3
     \    |  |    /
      C5==C4-C41--CA-HA
     /       |    \
H51--C51     H42   CT=OT1
    / \             |
  H52 H53           OT2(-)
```

Figure 8A.1: CHARMM parameters for ACPA.

```
RESI AMQ           0.000
GROUP           !  CHARGE
ATOM O1      OG2R50 -0.18
ATOM N2      NG2R50 -0.48
ATOM C3      CG2R52  0.49
ATOM C4      CG2R51 -0.19
ATOM C5      CG2R51  0.34
ATOM O31     OG311  -0.55
ATOM C41     CG321  -0.08
ATOM CA      CG314   0.21
ATOM N       NG3P3  -0.30
ATOM CT      CG2O3   0.3
ATOM OT1     OG2D2  -0.67
ATOM OT2     OG2D2  -0.67
ATOM C51     CG331  -0.27
ATOM HT1     HGP2    0.33
ATOM HT2     HGP2    0.33
ATOM HT3     HGP2    0.33
ATOM H41     HGA2    0.09
ATOM H42     HGA2    0.09
ATOM HA      HGA1    0.10
ATOM H31     HGP1    0.47
ATOM H51     HGA3    0.09
ATOM H52     HGA3    0.09
ATOM H53     HGA3    0.09

BONDS
CG2R52 OG311  400.00     1.37

ANGLES
CG331  CG2R51 OG2R50   25.00    112.00
CG2R51 CG2R52 OG311    88.00    130.00
NG2R50 CG2R52 OG311    55.00    133.00
CG2R52 OG311  HGP1     50.00    114.00
CG2R52 CG2R51 CG321    45.80    130.00

DIHEDRALS
CG2R52 CG2R51 CG2R51 CG331      2.6000   2  180.00
CG331  CG2R51 CG2R51 HGR51      0.3500   2  180.00
OG2R50 CG2R51 CG331  HGA3       0.0840   3    0.00
CG331  CG2R51 OG2R50 NG2R50     8.5000   2  180.00
CG2R51 CG2R51 CG2R52 OG311      20.0000  1    0.00
HGR51  CG2R51 CG2R52 OG311      2.0000   1  180.00
OG311  CG2R52 NG2R50 OG2R50     17.0000  1    0.00
CG2R51 CG2R52 OG311  HGP1       1.0000   2    0.00
NG2R50 CG2R52 OG311  HGP1       2.200    2  180.00
NG2R50 CG2R52 OG311  HGP1       1.850    1    0.00
CG321  CG2R51 CG2R51 CG331      0.6000   2  180.00
CG321  CG2R51 CG2R51 OG2R50     1.7000   2  180.00
CG321  CG2R51 CG2R52 OG311      2.0000   2  180.00
CG2R52 CG2R51 CG321  CG314      0.0000   1    0.00
CG2R52 CG2R51 CG321  CG314      0.7500   2    0.00
CG2R52 CG2R51 CG321  CG314      0.0000   3    0.00
CG2R52 CG2R51 CG321  HGA2       0.0000   3    0.00
CG321  CG2R51 CG2R52 NG2R50     4.0000   2  180.00

IMPROPERS
CG2O3  X      X      CG314     96.0000  0     0.00

END
RETURN
```

```
        H31--O31
             |
          N2==C3     HT1 HT2(+)
          /   |       \ /
         O1   |   H41   N--HT3
          \   |    |    |
           C5==C4--C41--CA-HA
          /        |    |
   H51--C51      H42   C=OT1
        / \            |
     H52 H53         OT2(-)
```

Figure 8A.2: CHARMM parameters for AMPA.

```
RESI CNI            0.000
GROUP           !  CHARGE
ATOM N1      NG2R61 -0.64
ATOM C2      CG2R63  0.54
ATOM C3      CG2R63  0.44
ATOM N4      NG2R61 -0.49
ATOM C5      CG2R61 -0.34
ATOM C6      CG2R61  0.15
ATOM C7      CG2R61 -0.23
ATOM C8      CG2R61 -0.21
ATOM C9      CG2R62  0.29
ATOM C10     CG2R62  0.16
ATOM O21     OG2D4  -0.43
ATOM O31     OG2D4  -0.42
ATOM N61     NG2O1   0.57
ATOM O61     OG2N1  -0.32
ATOM O62     OG2N1  -0.34
ATOM C71     CG1N1   0.40
ATOM N71     NG1T1  -0.37
ATOM H1      HGP1    0.42
ATOM H4      HGP1    0.40
ATOM H5      HGR61   0.22
ATOM H8      HGR61   0.20

BONDS
CG2R61 CG2R62  305.00     1.375
CG2R63 CG2R63  302.00     1.480

ANGLES
CG2R61 CG2R61 CG2R62   40.00    108.00   35.00   2.42
CG2R62 CG2R61 HGR61    30.00    126.00   22.00   2.15
CG2R61 CG2R62 CG2R62   40.00    116.00
CG2R61 CG2R62 NG2R61   85.00    132.90
CG2R63 CG2R63 NG2R61   70.00    105.50
CG2R63 CG2R63 OG2D4   100.00    110.60

DIHEDRALS
CG2R61 CG2R61 NG2O1  OG2N1      1.2000  2   180.00
CG1N1  CG2R61 CG2R61 CG2R62     2.0000  2   180.00
CG1N1  CG2R61 CG2R61 NG2O1      3.0000  2   180.00
CG2R61 CG2R61 CG2R61 CG2R62     3.1000  2   180.00
CG2R62 CG2R61 CG2R61 NG2O1      2.0000  2   180.00
CG2R61 CG2R61 CG2R62 CG2R62     3.1000  2   180.00
CG2R61 CG2R61 CG2R62 NG2R61     7.0000  2   180.00
HGR61  CG2R61 CG2R62 CG2R62     4.2000  2   180.00
HGR61  CG2R61 CG2R62 NG2R61     3.4000  2   180.00
CG2R61 CG2R62 CG2R62 CG2R61     6.0000  2   180.00
CG2R61 CG2R62 CG2R62 NG2R61     7.0000  2   180.00
NG2R61 CG2R62 CG2R62 NG2R61     0.0000  2   180.00
CG2R61 CG2R62 NG2R61 CG2R63     0.6000  2   180.00
CG2R61 CG2R62 NG2R61 HGP1       1.0000  2   180.00
NG2R61 CG2R63 CG2R63 NG2R61     0.0000  2   180.00
NG2R61 CG2R63 CG2R63 OG2D4      0.0000  2   180.00
OG2D4  CG2R63 CG2R63 OG2D4      0.7000  2   180.00
CG2R63 CG2R63 NG2R61 CG2R62     1.5000  2   180.00
CG2R63 CG2R63 NG2R61 HGP1       4.8000  2   180.00

IMPROPERS
CG2R63 CG2R63 NG2R61 OG2D4     90.0000  0     0.00

END
RETURN
```

```
        O21     O31
         \\   //
          C2-C3
         /     \
    H1-N1      N4-H4
        \     /
         C9=C10
        /     \
    H8-C8      C5-
        \\   //
         C7-C6
        /     \
      C71      N61=O62
      |||       |
      N71      O61
```

Figure 8A.3: CHARMM parameters for CNQX.

```
RESI DNQ           0.000
GROUP            ! CHARGE
ATOM N1     NG2R61 -0.55         O21     O31
ATOM C2     CG2R63  0.485         \\   //
ATOM C3     CG2R63  0.485          C2-C3
ATOM N4     NG2R61 -0.55          /     \
ATOM C5     CG2R61 -0.29       H1-N1     N4-H4
ATOM C6     CG2R61 -0.03          \     /
ATOM C7     CG2R61 -0.03          C9=C10
ATOM C8     CG2R61 -0.28          /     \
ATOM C9     CG2R62  0.23       H8-C8     C5-H5
ATOM C10    CG2R62  0.23          \\   //
ATOM O21    OG2D4  -0.42           C7-C6
ATOM O31    OG2D4  -0.42          /     \
ATOM N61    NG2O1   0.60    O71=N71    N61=O61
ATOM O61    OG2N1  -0.34         |      |
ATOM O62    OG2N1  -0.31       O72    O62
ATOM N71    NG2O1   0.60
ATOM O71    OG2N1  -0.34
ATOM O72    OG2N1  -0.31
ATOM H1     HGP1    0.40
ATOM H4     HGP1    0.40
ATOM H5     HGR61   0.22
ATOM H8     HGR61   0.22

BONDS
CG2R61 CG2R62  305.00     1.375
CG2R63 CG2R63  302.00     1.480

ANGLES
CG2R61 CG2R61 CG2R62   40.00    108.00   35.00   2.42
CG2R62 CG2R61 HGR61    30.00    126.00   22.00   2.15
CG2R61 CG2R62 CG2R62   40.00    116.00
CG2R61 CG2R62 NG2R61   85.00    132.90
CG2R63 CG2R63 NG2R61   70.00    105.50
CG2R63 CG2R63 OG2D4   100.00    110.60

DIHEDRALS
CG2R61 CG2R61 CG2R61 CG2R62     3.1000  2   180.00
CG2R62 CG2R61 CG2R61 NG2O1      1.8000  2   180.00
NG2O1  CG2R61 CG2R61 NG2O1      1.8000  2   180.00
CG2R61 CG2R61 CG2R62 CG2R62     3.1000  2   180.00
CG2R61 CG2R61 CG2R62 NG2R61     7.0000  2   180.00
HGR61  CG2R61 CG2R62 CG2R62     4.2000  2   180.00
HGR61  CG2R61 CG2R62 NG2R61     3.4000  2   180.00
CG2R61 CG2R62 CG2R62 CG2R61     6.0000  2   180.00
CG2R61 CG2R62 CG2R62 NG2R61     7.0000  2   180.00
NG2R61 CG2R62 CG2R62 NG2R61     0.0000  2   180.00
CG2R61 CG2R62 NG2R61 CG2R63     0.6000  2   180.00
CG2R61 CG2R62 NG2R61 HGP1       1.0000  2   180.00
NG2R61 CG2R63 CG2R63 NG2R61     0.0000  2   180.00
NG2R61 CG2R63 CG2R63 OG2D4      0.0000  2   180.00
OG2D4  CG2R63 CG2R63 OG2D4      0.7000  2   180.00
CG2R63 CG2R63 NG2R61 CG2R62     1.5000  2   180.00
CG2R63 CG2R63 NG2R61 HGP1       4.8000  2   180.00

IMPROPERS
CG2R63 CG2R63 NG2R61 OG2D4     90.0000  0     0.00

END
RETURN
```

Figure 8A.4: CHARMM parameters for DNQX.

Chapter 9

A Potent and Selective Peptide Blocker of the Kv1.3 Channel: Prediction from Free-Energy Simulations and Experimental Confirmation

ABSTRACT

The voltage-gated potassium channel Kv1.3 is a well-established target for treatment of autoimmune diseases. ShK peptide from a sea anemone is one of the most potent blockers of Kv1.3 but its application as a therapeutic agent for autoimmune diseases is limited by its lack of selectivity against other Kv channels, in particular Kv1.1. Accurate models of Kv1.x-ShK complexes suggest that specific charge mutations on ShK could considerably enhance its specificity for Kv1.3. Here we evaluate the K18A mutation on ShK, and calculate the change in binding free energy associated with this mutation using the path-independent free energy perturbation and thermodynamic integration methods, with a novel implementation that avoids convergence problems. To check the accuracy of the results, the binding free energy differences were also determined from path-dependent potential of mean force calculations. The two methods yield consistent results for the K18A mutation in ShK and predict a 2 kcal/mol gain in Kv1.3/Kv1.1 selectivity free energy relative to wild-type peptide. Functional assays confirm the predicted selectivity gain for ShK[K18A] and suggest that it will be a valuable lead in the development of therapeutics for autoimmune diseases.

9.1 Introduction

High-affinity binding to the target protein is a key criterion in the search for drug leads, and several methods, from high-throughput screening to computational docking, have been used for this purpose. The selectivity of a drug lead for a given target is also important as high affinity for an unintended target could lead to unacceptable side effects. In principle, experimental or computational screening methods developed for lead searching can be used to address the selectivity problem, although the complexity of proteins compared to lead ligands makes such brute force methods less likely to be successful. Methods based on structure-based drug design provide a promising alternative for solving selectivity problems, as we illustrate in this study, which also identifies a potent analog of ShK peptide that is selective for the voltage-gated potassium channel Kv1.3.

Kv1.3 is a well-established target for the treatment of autoimmune diseases mediated by effector memory T_{EM} lymphocytes, such as multiple sclerosis and rheumatoid arthritis[18,215]. ShK peptide, from the sea anemone *Stichodactyla helianthus*[45,46,216], is a potent blocker of the Kv1.3 channel with an IC_{50} of around 10 pM[217], and thus an excellent lead for development of an immunomodulatory drug[218]. However, ShK also blocks other Kv channels with high affinity, in particular Kv1.1 (IC_{50} = 16 pM)[217]. This lack of specificity for Kv1.3 is an impediment for the use of ShK as a therapeutic agent, and considerable effort has gone into developing ShK analogs with improved selectivity for Kv1.3 over Kv1.1 and other potassium channels[46,218]. Several analogs with the desired selectivity have been developed, e.g. ShK-Dap22[217], ShK-F6CA[219], ShK-186[220], and ShK-192[221], but all of them contain non-natural amino acids and/or adducts, and, in the case of ShK-186, the phosphorylated Tyr residue is susceptible to hydrolysis. For this reason, it is desirable to develop selective ShK analogs for Kv1.3 containing only natural amino acids.

An essential first step in searching for selective analogs using structure-based drug design methods is to construct accurate models of the relevant protein-ligand complexes. This has been achieved recently for Kv1.x-ShK complexes using docking and molecular dynamics (MD) simulations[47,222]. The validity of the complex models was established by comparing the predicted binding modes with alanine scanning mutagenesis data[223,224], and showing that the binding free energies of ShK obtained from the potential of mean force (PMF) calculations reproduce the experimental values within chemical accuracy[47]. Comparison of the binding modes in the Kv1.1-ShK and Kv1.3-ShK complexes reveals that the side chain of K18 in ShK is strongly coupled to a glutamate side chain in Kv1.1 via electrostatic interactions but it does not exhibit any interactions with the Kv1.3 residues[47]. Thus, mutation of this residue in ShK to alanine could lead to a substantial selectivity gain for Kv1.3 over Kv1.1.

Here we perform free energy simulations to calculate the change in the binding free energy of the Kv1.1-ShK and Kv1.3-ShK complexes associated with the K18A mutation. Two independent methods are used to check the veracity of the calcula-

tions. First we perform free energy perturbation (FEP) and thermodynamic integration (TI) calculations where ShK in the binding site is transformed to ShK[K18A] while the reverse transformation is carried out in the bulk for ShK[K18A], which gives directly the change in the binding free energy due to the mutation. We avoid convergence problems by using uncharged side chains as intermediate states. Convergence of the results was checked using the TI method, which also demonstrates the robustness of the calculations. Secondly, the PMFs for the dissociation of ShK[K18A] from Kv1.x are constructed from umbrella sampling MD simulations, from which the relative and absolute binding free energies of ShK[K18A] are determined. The change in the binding free energy is obtained by subtracting the relative binding free energy of the mutant from that of wild-type peptide[47]. The FEP and TI methods are computationally much cheaper but their reliability is not very well established for charge mutations. Thus, one of our aims was to demonstrate that the free energy change due to a charge mutation can be calculated accurately using the FEP and TI methods. The predictions of the free energy calculations were confirmed by measuring the binding constants of ShK[K18A] for Kv1.1 and Kv1.3 by patch-clamp electrophysiology.

9.2 Materials and methods

9.2.1 Ethics statement

All procedures involving vertebrate animals were conducted after approval by the Institutional Animal Care and Use Committee at Baylor College of Medicine. Baylor College of Medicine follows the requirements of the Guide for the Care and Use of Laboratory Animals (National Research Council, 8th edition). Female Lewis rats were purchased from Taconic (Germantown, NY, USA) and housed in autoclaved setup with food and water ad libitum in an Association for Assessment and Accreditation of Laboratory Animal Care International (ALAAAC)-approved facility. Every effort was made to minimize animal discomfort and to keep the number of animals used to a minimum. Rats were euthanized by deep isoflurane anesthesia (slow breathing and lack of reactivity to toe pinching), followed by cardiac puncture. Death was ensured by decapitation.

9.2.2 Peptide synthesis

The synthesis of the peptide has been described before[224]. Briefly, Fmoc-amino acid derivatives were obtained from Bachem A.G. (CH-4416 Bubendorf, Switzerland). Solid-phase assembly was initiated with an Fmoc-Cys(Trt)-2-chlorotrityl resin to minimize potential racemization of the C-terminal Cys residue. Automated stepwise assembly was carried out entirely on an ABI-431A peptide synthesizer (Applied Biosystems, Foster City, CA). ShK[K18A] was solubilized, oxidized, and purified by

reverse phase-high pressure liquid chromatography using the method described previously[217,223], and high pressure liquid chromatography-pure fractions were pooled and lyophilized. The structure and purity of the peptides were confirmed by reverse phase-high pressure liquid chromatography, amino acid analysis, and electrospray ionization-mass spectroscopy analysis. Samples were weighed and adjusted to account for peptide content before bioassay. The purity and mass were determined using LC and ESI-MS analyses (Figures 9A.1 and 9A.2 in the Appendix).

9.2.3 NMR spectroscopy and data analysis

Samples were prepared by dissolving freeze-dried ShK[K18A] in 90% H2O/10% ^{2}H2O, pH 4.8, to a concentration of 870 μM. One-dimensional ^{1}H spectra and two dimensional homonuclear TOCSY spectra with a spin lock time of 80 ms were acquired at 20°C on a Bruker Avance 600 MHz spectrometer. A NOESY spectrum (mixing times 200 ms) was also acquired at 20°C, pH 4.8. All spectra were processed in TOPSPIN (version 3.0, Bruker Biospin) and analysed using CcpNmr-Analysis (version 2.1.5). ^{1}H chemical shifts were referenced to the 1,4-dioxane signal at 3.75 ppm. Chemical shift assignments for backbone and side chain protons of ShK[K18A] were made by conventional analysis of TOCSY and NOESY spectra. A complete assignment of the proton NMR signals of ShK[K18A] was obtained.

The one-dimensional ^{1}H NMR spectrum (Figures 9A.3 and 9A.4 in the Appendix) of ShK[K18A] showed sharp and well-dispersed resonances similar to that of wild-type ShK, indicating that the K18A mutation did not cause any significant perturbation of the native structure. To further compare the structures of wild-type ShK and the K18A mutant, chemical shift differences from random coil values for amide, H^{α} and H^{β} resonances were plotted (Figure 9A.5 in the Appendix). This shows that the patterns of chemical shift deviations from random coil for ShK[K18A] closely resemble those for wild-type ShK, and confirms that the three-dimensional structures are basically the same. Small differences in chemical shifts between wild-type and mutant were observed only for resonances around the mutated residue K18 (Figure 9A.5 in the Appendix).

9.2.4 Cells

Mouse L929 fibroblasts stably expressing mKv1.1, mKv1.3, hKv1.5 channels[225] were gifts from Dr. K. George Chandy (University of California, Irvine). Human embryonic kidney 293 cells stably expressing hKv11.1 (hERG) were a gift from Dr. Craig January (University of Wisconsin, Madison). CCR7$^-$ Ova-GFP T_{EM} cells (gift from Dr. Flgel, Munich, Germany) were derived from Lewis rats and react to ovalbumin[226]; they were maintained in culture through alternating rounds of antigen-dependent stimulation and of expansion in a cytokine-rich medium[227,228]. As antigen-presenting cells for T_{EM} cell stimulation, we used irradiated (30 Gy) thymocytes harvested from humanely euthanized Lewis rats[227]. Histopaque-1077

gradients (Sigma-Aldrich) were used to isolate splenocytes from the same Lewis rats[220,229].

9.2.5 Electrophysiology

Electrophysiology experiments were conducted in the whole-cell configuration of the patch-clamp technique on a Port-a-Patch (Nanion Technologies, North Brunswick, NJ) connected to an EPC10-USB amplifier (HEKA Instruments, Bellmore, NY)[228,230]. A holding potential of -80 mV was used for all recordings. Pipette resistances averaged 2.0 MΩ and series resistance compensation of 80% was employed when currents exceeded 2 nA. Kv1.3 currents were elicited by repeated 200-ms pulses from -80 mV to 40 mV, applied at intervals of 30 or 60 s. Kv1.3 currents were recorded in normal Ringer solution with a Ca^{2+}-free pipette solution containing (in mM): 145 KF, 10 HEPES, 10 EGTA, 2 $MgCl_2$, pH 7.2, 300 mOsm. IC_{50}-values were determined by fitting the Hill equation to the reduction of area under the current curve measured at 40 mV. Kv1.1 and Kv1.5 currents were recorded with 200-ms depolarizing pulses to 40 mV applied every 10 sec. HERG (Kv11.1) currents were recorded with a 2-step pulse from -80 mV first to 20 mV for 2 sec and then to -50 mV for 2 sec and the reduction of both peak and tail current by the drug was determined.

9.2.6 T Lymphocyte proliferation assays

The T_{EM} cells were used after 4 days of IL-2-dependent expansion. The splenic T cells were used immediately after collection from Lewis rats. Cells were seeded in flat-bottom 96-well plates in 200 μL of culture medium supplemented with 1% homologous rat serum. ShK or ShK[K18A] were added to the cells 30 min before the antigen or mitogen. T_{EM} cells were stimulated with 10 μg/ml ovalbumin in the presence of irradiated thymocytes as antigen-presenting cells and splenocytes were stimulated with 1 μg/mL concanavalin A. The cells were cultured for 3 days and were pulsed with [^{3}H] thymidine (1 4μCi/well) 16-18 h before harvest. The proliferative response of T cells was assessed in a β scintillation counter measuring the [^{3}H] thymidine incorporated into their DNA[228,230].

9.2.7 Delayed type hypersensitivity

Lewis rats were immunized against ovalbumin[227] and a week later challenged in the pinna of one ear with ovalbumin or saline[227,231]. They received a subcutaneous injection of 100 μg/kg ShK[K18A] at time of challenge and ear thickness was determined 24 h later as a measure of inflammation[220,227,228].

9.2.8 Modeling of Kv1.x-ShK[K18A] complexes

The solution structure of ShK was obtained from the Protein Data Bank (PDB ID, 1ROO)[232]. Homology models of Kv1.1 and Kv1.3 were taken from previous work[47]. The structure of ShK[K18A] was generated from that of ShK using the VMD plugin, Mutator[104]. Two-dimensional homonuclear NMR spectroscopy was used to confirm that the solution structure of this analog was essentially identical with that of ShK[232,233] (Figures 9A.3-5 in the Appendix). The Kv1.x-ShK[K18A] complex models were generated via docking and refinement with MD following the same protocols used for Kv1.x-ShK complexes[47]. Briefly, the structure ShK[K18A] was equilibrated in bulk MD simulations for 5 ns, and ten frames were chosen for use in docking calculations. The initial Kv1.x-ShK[K18A] complexes were obtained using HADDOCK[69]. For both Kv1.1 and Kv1.3, the top ten poses with the highest scores were found to be very similar, and the one with the highest score was chosen for refinement with MD. Each complex structure was embedded in a lipid bilayer consisting of 125 POPC molecules and solvated with 10,525 water molecules and 0.1 M KCl. After the correct water and lipid densities were obtained with pressure coupling, the protein-toxin complex was gradually relaxed in 5 ns MD simulations. The equilibrated system was run for another 5 ns for analysis of the complex structure, where convergence of the channel protein and toxin peptide was monitored using the RMSD of the backbone atoms. MD simulations were performed using the NAMD program[59] with the CHARMM27 force field[55], which includes the CMAP correction[105]. An NpT ensemble was used with the temperature and pressure maintained at 300 K and 1 atm, respectively, via Langevin coupling with damping coefficients of 5 ps^{-1} and 10 ps^{-1}. Periodic boundary conditions were employed together with particle-mesh Ewald algorithm to compute the long-range electrostatic interactions. Lennard-Jones (LJ) interactions were switched off within a distance of 10–12 Å, and the list of non-bonded interactions was truncated at 13.5 Å. A time step of 2 fs was used, and the trajectory data were written at 1 ps intervals.

9.2.9 FEP and TI calculations

The free energy change due to a mutation is most efficiently calculated using the alchemical transformation methods such as FEP and TI[52], where the side chain of an amino acid is gradually transformed into another side chain and the free energy change is determined from the free energies of the end-points. This also applies to a protein-ligand complex, provided the binding mode is not altered by the mutation, which is the case for the K18A mutation in the Kv1.x-ShK complexes considered here. To avoid the problems posed by the change in the charge state, we implement the thermodynamic cycle in Figure 9.1. Accordingly, while the K18A mutation was applied to ShK in the binding site, the reverse transformation was applied to the toxin in the bulk simultaneously and in the same system, which ensures that the net charge in the system did not change during the calculations. To deal with

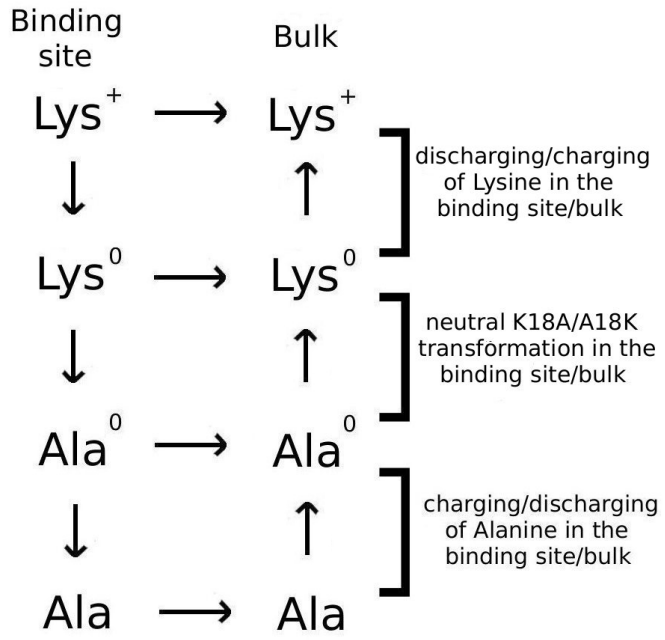

Figure 9.1: The thermodynamic cycle used in the free energy calculations. The superscript 0 denotes amino acids with no charges on the side chain atoms.

convergence issues, we separated the Coulomb and LJ interactions and performed the K18A mutation in three stages as depicted in Figure 9.1. First, the K18 side chain on the bound ShK was discharged, while the reverse process was performed on ShK in the bulk with an uncharged K18 side chain. Secondly, the uncharged K18 side chain on the bound ShK was transformed to an uncharged A18 side chain, while the reverse was performed on the bulk ShK. Finally, the A18 side chain in the binding site was charged, while the corresponding side chain in the bulk was discharged (note that the Ala side chain is neutral overall but has small partial charges on C and H atoms). The change in the binding free energy due to the K18A mutation can, thus, be expressed as:

$$\Delta\Delta G_b = \Delta G_b(ShK[K18A]) - \Delta G_b(ShK) = \Delta\Delta G(K \to K^0) + \Delta\Delta G(K^0 \to A^0) + \Delta\Delta G(A^0 \to A) \tag{9.1}$$

The first and last terms on the right hand side give the Coulomb contributions to the free energy change and the middle term gives the LJ contribution. Each term is calculated using the FEP and/or TI methods. In both methods, one introduces a hybrid Hamiltonian $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$, where H_0 represents the Hamiltonian for the initial state and H_1 for the final state. In the FEP method, the interval [0,1] is divided into n subintervals with $\{\lambda_i, i = 1, \ldots, n-1\}$, and for each subinterval the free energy difference is calculated from the ensemble average,

$\Delta G_i = -k_B T \ln\langle \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_B T]\rangle_{\lambda_i}$. The free energy difference between the initial and final states is obtained from the sum, $\Delta G = \sum_i \Delta G_i$. Provided $\Delta G_i \leq 2kT$ for each window[52], the FEP sum should converge with reasonable sampling. For charge mutations, satisfaction of this condition would require over 130 windows if uniform subintervals were used. Using exponentially spaced subintervals instead halves the required number of windows to 66 (the λ values are provided in the Appendix). Each window was equilibrated for 80 ps followed by a 120 ps production run. These were found to be optimal times after several tests with different times. We also performed the backward calculation to check for hysteresis effects. The system was equilibrated for 5 ns after the forward transformation before starting the backward transformation. For the LJ part, we used 30 equally-spaced windows at the end points, i.e., for λ between 0–0.3 and 0.7–1, and 20 equally-spaced windows in between. Using 20 ps equilibration and 30 ps production run for each window was found to be adequate. In order to improve convergence and prevent instabilities, we used a soft-core LJ potential with a shift coefficient of 7.0[82].

In the TI method, the ensemble average of the derivative, $\partial H(\lambda)/\partial\lambda$, was obtained at several values, and the free energy difference was calculated from the integral, $\Delta G = \int_0^1 \langle \partial H(\lambda)/\partial\lambda \rangle_\lambda \, d\lambda$. The TI method was especially advantageous for charge interactions because Gaussian quadrature allows evaluation of the integral using a small number of windows, which can be sampled for longer times to check convergence of the results. A seven-point quadrature was found to be adequate in previous applications of the TI method[123], which was also adopted here (the λ values are provided in the Appendix). The initial TI windows were taken from the nearest FEP window. Each window was typically equilibrated for 1 ns and sampled for 1 ns. The TI method has not been used in the calculation of the LJ interactions because it offers no advantages over FEP (it requires as many windows due to the unsuitability of Gaussian quadrature and takes longer time), and good results have already been obtained using FEP.

9.2.10 PMF Calculations

PMF provides the most reliable method for calculation of the absolute binding free energies of charged peptides, which can be directly compared to the experimental values. Unlike the FEP and TI methods, which give the free energy difference between two states, PMF provides a continuous free energy profile between the initial and final states. This makes PMF computationally more expensive than the FEP and TI methods. Nevertheless, it is very useful for testing of alternative methods. The PMFs for unbinding of ShK from Kv1.x are available from earlier work[47], and those for ShK[K18A] were constructed here using the umbrella sampling MD simulations employing the protocols developed in earlier work[47,84]. Briefly, umbrella windows are generated at 0.5 Å intervals along the channel axis using steered MD simulations with k = 30 kcal/mol/Å^2 and v = 5 Å/ns. After each pulling step, the toxin was equilibrated for 0.4 ns with the same restraining force to relax the effect of

steering on the environment. The reaction coordinate was measured from the center of mass of the channel protein to the center of mass of the toxin along the z axis, which was aligned with the channel axis. A minimum of 5% overlap between the sampling of neighboring windows was required for an accurate construction of the PMF[84]. When the overlap was below this limit, e.g. due to a sharp rise in the PMF, an extra window was included between the two windows. The same umbrella force constant (30 kcal/mol/Å^2) was used in all windows, which was found to be optimal for this size toxin[84]. The PMF was constructed from the reaction coordinates of ShK[K18A] collected during the umbrella sampling simulations via the weighted histogram analysis method[81]. Simulations were continued until the convergence of each PMF is assured. The binding constant was determined from the integral of the PMF, $W(z)$, along the reaction coordinate, and the absolute binding energy from the log of the binding constant:

$$K_{eq} = \pi R^2 \int_{z_1}^{z_2} e^{-W(z)/k_B T} dz \tag{9.2}$$

$$G_b = -k_B T ln(K_{eq} C_0) \tag{9.3}$$

Here z_1 is in the binding pocket and z_2 is in the bulk where W vanishes. The factor πR^2 measures the average cross-sectional area of the binding pocket as explored by the center of mass (COM) of the toxin. The values of R were determined from the widths of the transverse fluctuations of the toxin COM. They were obtained from restraint-free MD simulations of the complexes as 0.71 and 0.70 Å, respectively, for Kv1.1-ShK[K18A] and Kv1.3-ShK[K18A]. C_0 is the standard concentration of 1 M.

9.3 Results and discussion

9.3.1 ShK[K18A] complexes with Kv1.1 and Kv1.3

Equilibrated structures of the Kv1.1-ShK[K18A] and Kv1.3-ShK[K18A] complexes were obtained as described in Methods. To visualize the effect of the K18A mutation on the binding modes, we superimpose the snapshots of the Kv1.x-ShK[K18A] complexes with those of ShK (Figure 9.2). In both complex structures, there is a very good overlap between the backbones of the mutant and wild-type ShK, indicating that the K18A mutation causes minimal perturbation of the binding mode. This observation is made more quantitative in Table 9.1, which compares the average interatomic distances between the strongly interacting residues in the mutant and wild-type complexes. All the contact pairs are seen to be preserved after the mutation in both complexes, and the changes in the pair distances are well within the fluctuations for all the strongly coupled pairs (distance < 4 Å).

Table 9.1: Comparison of the strongly interacting pair distances in the ShK-Kv1.x complexes with those in the ShK[K18A]-Kv1.x complexes. The average atom-atom distances obtained from 5 ns MD simulations are listed in column 3 for the wild-type complexes and in column 6 for the mutant complexes (in units of Å). Subscripts refer to the side chain atoms and the monomer identity in K^+ channels is given in parentheses.

ShK	Kv1.1	MD average	ShK[K18A]	Kv1.1	MD average
R11-N_2	D361-O_2(B)	5.5 ± 0.5	R11-N_2	D361-O_1(B)	4.5 ± 0.3
K18-N_1	E353-O_2(C)	2.7 ± 0.2			
S20-O_H	Y379-O_H(B)	3.0 ± 0.3	S20-O_H	Y379-O_H(B)	3.0 ± 0.3
K22-N_1	Y375-O(ABC)	2.7 ± 0.2	K22-N_1	Y375-O(BCD)	2.7 ± 0.5
F27-$C_{\varepsilon 2}$	Y379-$C_{\varepsilon 1}$(A)	3.6 ± 0.2	F27-$C_{\varepsilon 2}$	Y379$C_{\varepsilon 1}$(A)	4.0 ± 0.4
R29-N_2	E353-O_2(D)	2.5 ± 0.3	R29-N_2	E353-O_2(D)	2.8 ± 0.4
ShK	Kv1.3	MD average	ShK[K18A]	Kv1.3	MD average
R1-N_1	D376-O_1(C)	4.5 ± 0.4	R1-N_1	D376-O_1(C)	3.5 ± 0.3
R11-N_2	D402-O_1(A)	3.5 ± 0.3	R11-N_2	D402-O_1(A)	4.0 ± 0.3
H19-N_H	S378-O(B)	3.0 ± 0.5	H19-N_H	S378-O(B)	3.5 ± 0.5
S20-O_H	G401-O(B)	2.7 ± 0.2	S20-O_H	G401-O(B)	2.7 ± 0.2
K22-N_1	Y400-O(ABD)	2.7 ± 0.2	K22-N_1	Y400-O(ABD)	2.7 ± 0.2
Y23-O_H	G401-O(A)	3.5 ± 0.3	Y23-O_H	G401-O(A)	3.5 ± 0.3
F27-$C_{\varepsilon 1}$	H404-C_γ(C)	3.6 ± 0.2	F27-$C_{\varepsilon 1}$	H404-C_γ(C)	3.7 ± 0.4

9.3.2 Free energy changes from FEP and TI calculations

Provided the binding mode is preserved, the most straightforward way to calculate the binding free energy change due to a mutation on a ligand is to use the FEP and/or TI methods. To facilitate convergence, we adopted a staged approach, as described in Figure 9.1. We used both methods for the charging/discharging steps where convergence issues are critical for the long-range Coulomb interactions. Evidence for the convergence of the FEP and TI calculations for charging/discharging of Lys side chain is provided in Figure 9.3. Mutation of the uncharged Lys to Ala involves short-range LJ interactions, where convergence is not an issue provided the windows at the end points are chosen at small intervals. Because the TI method offers no advantages, we have performed three independent FEP calculations for the LJ step. Convergence of the FEP calculations for the LJ step is demonstrated in Figure 9.3B. Detailed results of the FEP and TI calculations showing the individual contributions from the three steps in both forward and backward directions for the Kv1.1 and Kv1.3 channels are presented in Table 9.2. We note that there is negligible hysteresis between the forward and backward calculations in all cases, and

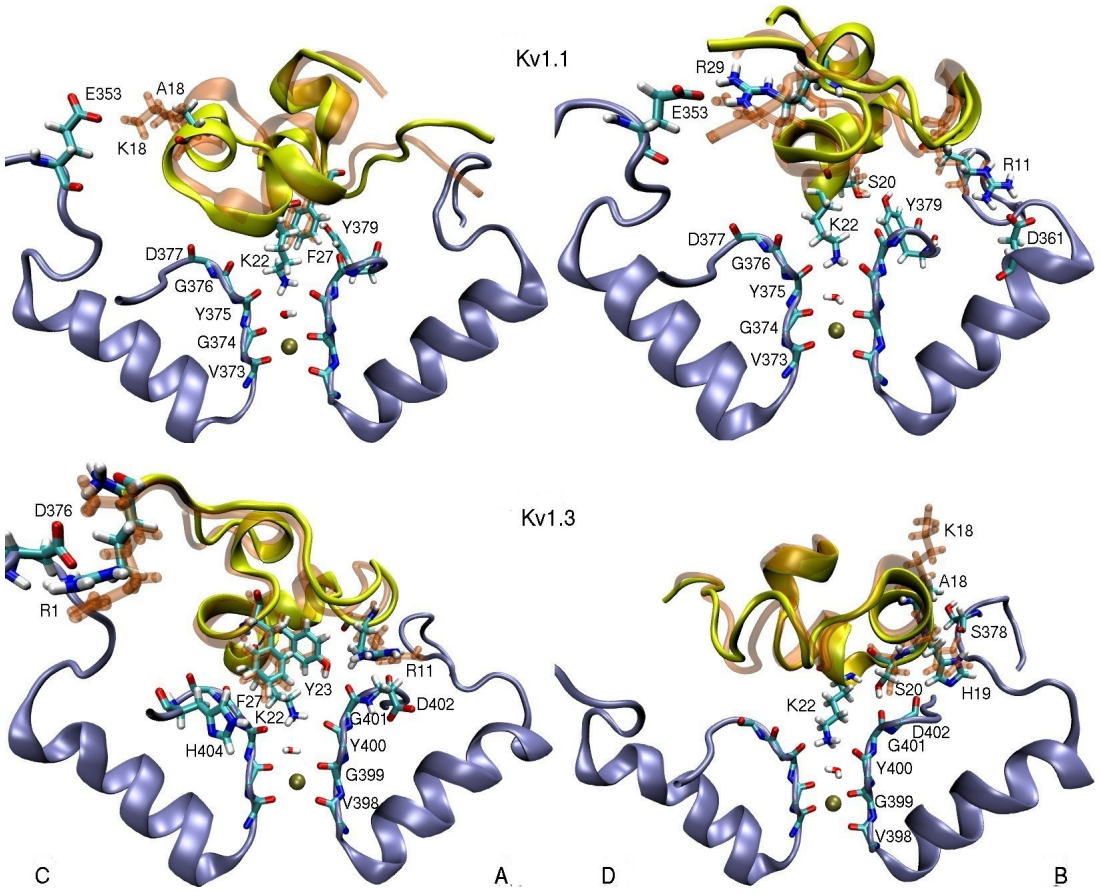

Figure 9.2: The K18A mutation does not change the binding modes of ShK to Kv1.1 and Kv1.3 channels. Snapshots of ShK[K18A] (yellow with colored side chains) and ShK (transparent orange) in complex with Kv1.1 and Kv1.3 are superposed to show that there is a good overlap between the toxin backbones, and all the important interactions identified in the ShK complexes are preserved in the ShK[K18A] complexes. To give the full picture, two views of the cross sections of the complex, depicting the monomers C and A (left panel) and D and B (right panel) are presented. Only the residues involved in binding are indicated explicitly.

the FEP and TI calculations yield consistent results for the charging/discharging steps. As expected, the main contribution to the free energy difference comes from the charging/discharging of the Lys side chain, while the charging/discharging of the Ala side chain makes a negligible contribution.

Table 9.2: Differences in the binding free energy of ShK to Kv1.1 and Kv1.3 due to the K18A mutation, calculated with the FEP and TI methods. The three contributions to the binding free energy difference and their sum (Eq. 9.1) are listed in columns 2-4 (in units of kcal/mol). In each case, the results of the forward and backward calculations are given separately, followed by their average. The TI calculations are performed for the Coulomb parts only; the LJ contribution is taken from the FEP calculations. The uncertainties are calculated from block data analysis of the data.

Kv1.1/FEP	$\Delta\Delta G(K \to K^0)$	$\Delta\Delta G(K^0 \to A^0)$	$\Delta\Delta G(A^0 \to A)$	$\Delta\Delta G_b$
FORWARD	2.4 ± 0.4	-0.3 ± 0.1	-0.1 ± 0.1	2.0 ± 0.4
BACKWARD	-2.5 ± 0.5	0.3 ± 0.1	0.1 ± 0.1	-2.1 ± 0.5
AVERAGE	2.5 ± 0.5	-0.3 ± 0.1	-0.1 ± 0.1	2.1 ± 0.5
Kv1.3/FEP				
FORWARD	0.9 ± 0.3	-0.2 ± 0.1	0.0 ± 0.1	0.7 ± 0.2
BACKWARD	-0.4 ± 0.3	0.2 ± 0.1	0.0 ± 0.1	-0.2 ± 0.2
AVERAGE	0.7 ± 0.3	-0.2 ± 0.1	0.0 ± 0.1	0.5 ± 0.2
Kv1.1/TI				
FORWARD	2.7 ± 0.5	-0.3 ± 0.1	-0.1 ± 0.1	2.3 ± 0.5
BACKWARD	-3.0 ± 0.6	0.3 ± 0.1	0.1 ± 0.1	-2.6 ± 0.6
AVERAGE	2.8 ± 0.6	-0.3 ± 0.1	-0.1 ± 0.1	2.4 ± 0.6
Kv1.3/TI				
FORWARD	0.6 ± 0.5	-0.2 ± 0.1	0.0 ± 0.1	0.4 ± 0.5
BACKWARD	-0.1 ± 0.5	0.2 ± 0.1	-0.1 ± 0.1	0.0 ± 0.5
AVERAGE	0.4 ± 0.5	-0.2 ± 0.1	0.0 ± 0.1	0.2 ± 0.5

9.3.3 Free energy changes from PMF calculations

PMF calculations were then performed to check the accuracy of the FEP and TI calculations. The binding constants and absolute binding free energies obtained from PMFs also allow a direct comparison with the experimental values. The PMFs for the unbinding of ShK from Kv1.1 and Kv1.3 were constructed in our earlier work[47]. Here we repeat these PMF calculations for ShK[K18A] using the same protocols. Convergence studies of the PMFs from block data analysis (Figures 9A.6

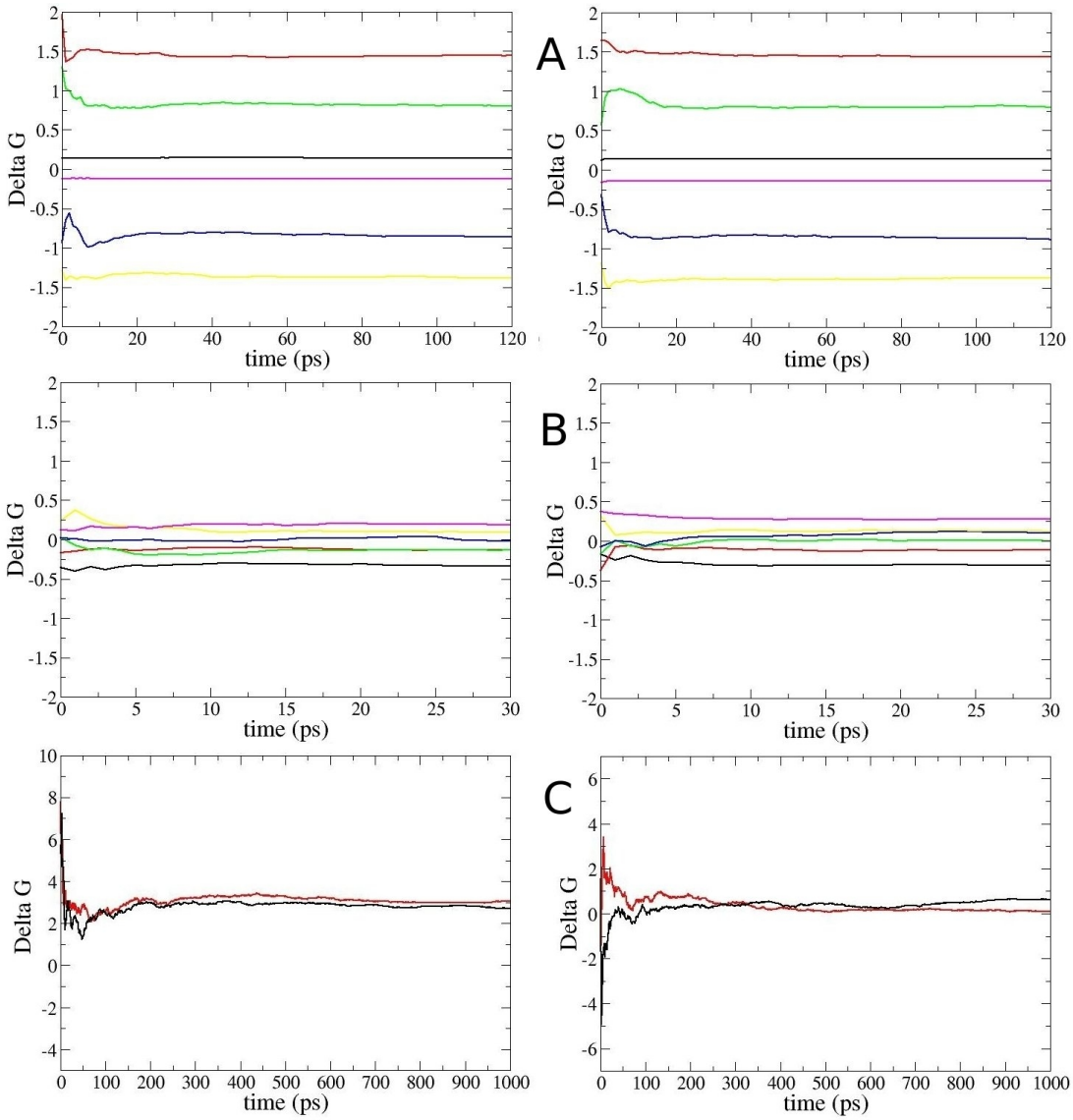

Figure 9.3: Convergence of the FEP and TI calculations. (A) Convergence of the FEP calculations for discharging/charging of the K18 in ShK in the binding site/bulk. Running averages of ΔG_i for windows at $\lambda = 0$ (black), $\lambda = 0.2$ (red), $\lambda = 0.4$ (green), $\lambda = 0.6$ (blue), $\lambda = 0.8$ (yellow) and $\lambda = 0.999$ (magenta) are plotted as a function of the production time for Kv1.1 (left) and Kv1.3 (right). (B) Similar to A but showing the convergence of the FEP calculations for transformation of the uncharged Lys side chain to that of Ala. (C) Convergence of the TI calculations for discharging/charging of K18 in ShK in the binding site/bulk. Running averages of the ΔG values obtained from the TI calculations are plotted as a function of the production time for Kv1.1 (left) and Kv1.3 (right). Both the forward (black) and the negative of the backward (red) results are shown to check for hysteresis effects, which remain well below 1 kcal/mol for both channels.

and 9A.7 in the Appendix) show that the PMF in the Kv1.3-ShK[K18A] system converges relatively quickly, just as in the case of the Kv1.3-ShK system[47], and the final PMFs for ShK and ShK[K18A] are very similar (Figure 9.4). Convergence of the PMF in the Kv1.1-ShK[K18A] system is delayed somewhat (Figures 9A.6 and 9A.7 in the Appendix), which can be traced to the slow equilibration of the E353-R29 side chain interactions. As expected, abolishing a charge interaction in the Kv1.1-ShK[K18A] system leads to a shallower PMF compared with that for Kv1.1-ShK. An interesting feature of the ShK[K18A] PMF is the absence of the shoulder region observed in the ShK PMF, which is caused by the long-range Coulomb interactions, in particular by the E353-R29 pair[47]. To understand this feature, we compare the E353(O2)-R29(N2) pair distances in ShK and ShK[K18A] PMFs as a function of the channel-toxin distance, z (Figure 9A.8 in the Appendix). The pair distance remains around 7-8 Å for $z > 35$Å in the ShK PMF but becomes double that in the ShK[K18A] PMF. Thus, there are no discernible charge interactions left after $z = 35$Å in the ShK[K18A] PMF, which explains its earlier flattening compared to the ShK PMF. Similarly the bump in the Kv1.1-ShK PMF around $z = 43$Å can be explained by the re-engagement of the E353-R29 pair due to rotation of the toxin and its subsequent dissociation (Figure 9A.8 in the Appendix). Absolute binding free energies obtained from the integration of the PMFs (Eq. 9.2) are compared to the experimental values in Table 9.3. In all cases, the calculated binding free energies agree with the experimental values within the chemical accuracy of 1 kcal/mol, which provides strong support for the accuracy of the complex models.

Table 9.3: The relative and absolute binding free energies obtained from the PMF calculations for the Kv1.x-ShK and Kv1.x-ShK[K18A] complexes. The relative binding energy, ΔW_{well}, is given by the well depth in the PMF and the absolute binding free energy, G_b, is obtained from the integration of the PMF via Eq. 9.2 (in units of kcal/mol). Experimental G_b is determined from the binding constants using Eq. 9.3.

Complex	ΔW_{well}	G_b (PMF)	G_b (exp)
Kv1.1-ShK	-18.0 ± 0.7	-14.3 ± 0.6	-14.4 ± 0.1
Kv1.1-ShK[K18A]	-15.3 ± 0.7	-11.7 ± 0.7	-11.3 ± 0.1
Kv1.3-ShK	-17.8 ± 0.8	-14.2 ± 0.7	-15.0 ± 0.1
Kv1.3-ShK[K18A]	-17.4 ± 0.7	-13.9 ± 0.6	-14.2 ± 0.1

9.3.4 Selectivity free energies

The results of the binding free energy differences obtained using various approaches are summarized in Table 9.4 together with the experimental values. The FEP and TI results are in good agreement with those of the PMF, which indicates that the

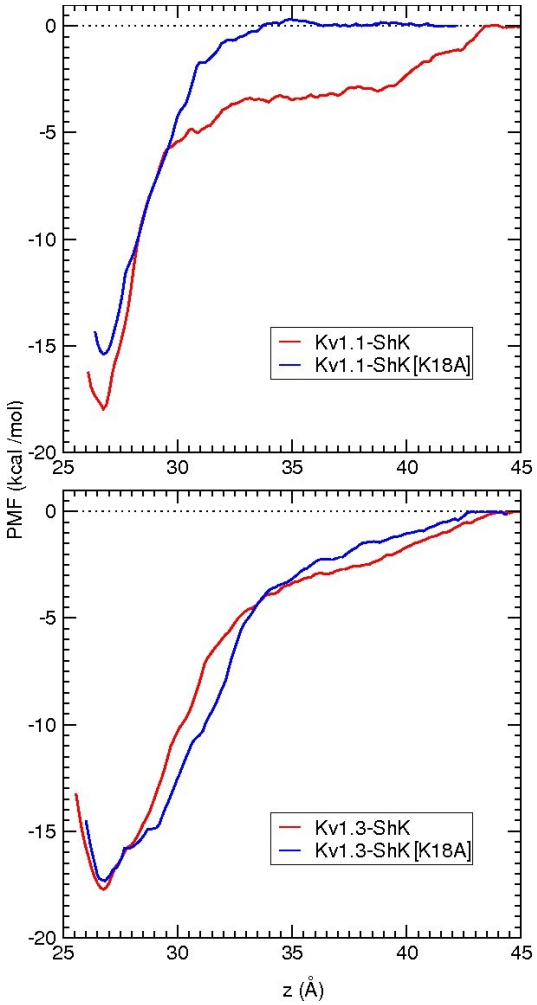

Figure 9.4: Comparison of the PMFs for the unbinding of ShK and ShK[K18A] from the Kv1.1 and Kv1.3 channels.

former methods can be used reliably in calculation of the free energy differences associated with mutation of amino acids. The TI method gives slightly better results compared with FEP for charge interactions, presumably because of longer sampling of windows. The crucial question of the selectivity free energy gain for Kv1.3/Kv1.1 associated with the K18A mutation is addressed in the last column of Table 9.4. All methods predict a gain in selectivity free energy of about 2 kcal/mol, which is confirmed with the binding constant measurements of ShK[K18A] for Kv1.1 and Kv1.3.

Table 9.4: Comparison of the binding free energy differences for Kv1.1 and Kv1.3, and the selectivity free energy for Kv1.3/Kv1.1, obtained using the FEP, TI and PMF methods, with the experimental results. The FEP and TI results are taken from Table 9.2, and the PMF results are obtained from Table 9.3 using $\Delta\Delta G_b = G_b$ (ShK[K18A]) - G_b (ShK) (a similar result is obtained if ΔW_{well} is used instead of G_b). The change in selectivity free energy due to the mutation in column 4 follows from $\Delta\Delta G_{sel} = \Delta\Delta G_b$ (Kv1.1) - $\Delta\Delta G_b$ (Kv1.3). All energies are in units of kcal/mol.

	$\Delta\Delta G_b$ (Kv1.1)	$\Delta\Delta G_b$ (Kv1.3)	$\Delta\Delta G_{sel}$
FEP	2.1 ± 0.5	0.5 ± 0.2	1.6 ± 0.6
TI	2.4 ± 0.6	0.2 ± 0.5	2.2 ± 0.8
PMF	2.6 ± 1.0	0.3 ± 1.0	2.3 ± 1.4
Exp.	3.1 ± 0.1	0.8 ± 0.2	2.3 ± 0.2

9.3.5 Experimental validation of modeling results

In order to validate the modeling results by experimentation, we used the well-established technique of whole-cell patch-clamp. Using this technique, mouse fibroblasts stably expressing either Kv1.1 or Kv1.3 channels were patch-clamped to measure total potassium currents through these channels. The perfusion of various dilutions of ShK or ShK[K18A] induced Kv1.1 or Kv1.3 channel block, measured by a steady-state reduction in potassium currents through the Kv1 channels, allowing for the determination of the half-inhibitory concentration (IC_{50}) for ShK and its analog. ShK[K18A] inhibited Kv1.3 channels with an IC_{50} of 39.6 ± 3.8 pM, displaying a modest decrease in potency compared to ShK (IC_{50} on Kv1.3 = 9.3 ± 2.0 pM) (Figure 9.5A)[230]. ShK also inhibited Kv1.1 channels with a high affinity (IC_{50} = 25.6 ± 2.8 pM) (Figure 9.5B). In contrast, ShK[K18A] exhibited a 124-fold selectivity for Kv1.3 over Kv1.1 channels (IC_{50} on Kv1.1 = 4,900 ± 200 pM).

A potential concern when developing potassium channel blockers is off-target toxic effects on cardiac channels. We therefore used the whole-cell patch-clamp assay to determine the effects of ShK[K18A] on the two cardiac potassium channels Kv1.5 and Kv11.1 (hERG). ShK[K18A] had no effects on either cardiac potassium channels when tested at concentrations of up to 100 nM (results not shown).

Since selective Kv1.3 channel blockers are known to preferentially inhibit the proliferation of CCR7$^-$ T_{EM} lymphocyte with little to no effect on the proliferation of CCR7$^+$ naïve and T_{CM} lymphocytes, we next tested the effects of ShK and ShK[K18A] on the proliferation of rat T lymphocytes. Both peptides displayed a marked preferential inhibition of the proliferation of T_{EM} lymphocytes (Figure 9.5C) with little effect on naïve/T_{CM} cells (Figure 9.5D). These results suggest that ShK[K18A] is selective for Kv1.3 over KCa3.1 channels as naïve/T_{CM} cells rely on KCa3.1 channels for their proliferation[215].

Finally, we used an active DTH reaction to ovalbumin to determine the *in vivo* ability of ShK[K18A] to inhibit a T_{EM} lymphocyte-mediated inflammatory reaction. ShK[K18A] reduced this active DTH reaction by *ca* 52% (Figure 9.5E), exhibiting similar activity to ShK-186 and ShK-192 in this model[221].

These results demonstrate that ShK[K18A] is bioavailable and can inhibit T_{EM} cells both in vitro and in vivo, making it a compound of considerable interest for immunomodulation and treatment of T_{EM}-mediated diseases.

9.4 Conclusions

Increasing the selectivity of a drug lead for a specified target is a design problem that in principle could be greatly facilitated using computational methods. Here we have described the case of ShK peptide, which binds with high affinity to the Kv1.3 channel and is a potential drug lead for the treatment of autoimmune diseases. We have analyzed the K18A mutation in ShK, which was suggested by accurate models of Kv1.x-ShK complexes, and shown via free energy calculations, to produce a 2 kcal/mol gain in selectivity free energy for Kv1.3/Kv1.1. Experimental confirmation of this result increases confidence in the ability of the free energy methods, in particular computationally cheaper FEP and TI methods, to predict the effects of mutations accurately. Moreover, the potency of ShK[K18A] against Kv1.3, its selectivity for this channel over Kv1.1, its preferential inhibition of T_{EM} cells over naïve and T_{CM} cells, its ability to significantly reduce a DTH reaction, and the fact that it is composed entirely of protein amino acids, make it an attractive candidate for further evaluation as a therapeutic for autoimmune diseases.

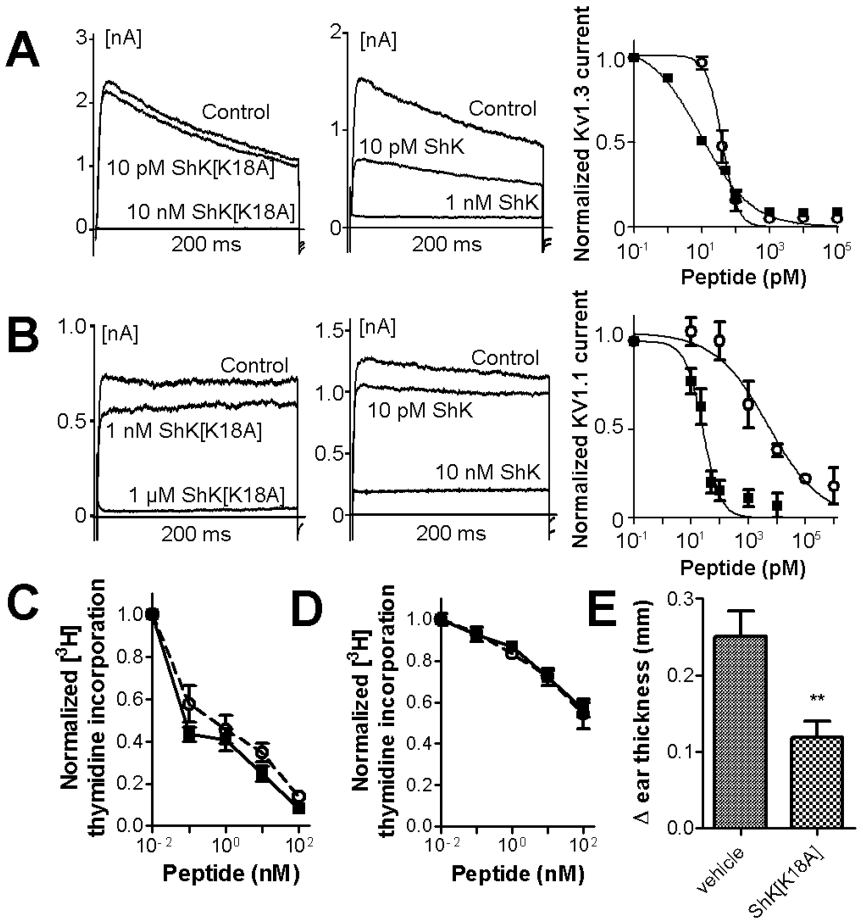

Figure 9.5: ShK[K18A] is selective for Kv1.3 over Kv1.1 channels and preferentially targets T_{EM} lymphocytes *in vitro* and *in vivo*. (A) Effects of ShK (black squares) and ShK[K18A] (open circles) on Kv1.3 currents measured by whole-cell patch-clamp on L929 fibroblasts stably transfected with mKv1.3. The two left panels show whole-cell Kv1.3 currents before (control) and after perfusion of ShK[K18A] (left panel) or ShK (middle panel). The panel on the right shows the Kv1.3 currents remaining after steady-state block is reached with the different concentrations of ShK and its analog, fitted to a Hill equation (N = 5-6 cells per concentration).(B) Effects of ShK (black squares) and ShK[K18A] (open circles) on Kv1.1 currents measured by whole-cell patch-clamp on L929 fibroblasts stably transfected with mKv1.1. The two left panels show whole-cell Kv1.1 currents before (control) and after perfusion of ShK[K18A] (left panel) or ShK (middle panel). The panel on the right shows the Kv1.1 currents remaining after steady-state block is reached with the different concentrations of ShK and its analog, fitted to a Hill equation (N = 6-7 cells per concentration).(C) Effects of ShK (black squares) and ShK[K18A] (open circles and dashed line) on the proliferation of rat T_{EM} cells measured *in vitro* by the incorporation of $[^3H]$ thymidine into the DNA of dividing cells (N = 3). (D) Effects of ShK (black squares) and ShK[K18A] (open circles and dashed line) on the proliferation of rat splenocytes (mainly naïve/T_{CM} cells; N = 3).(E) Effects of the subcutaneous administration of 100 μg/kg ShK[K18A] on an active DTH reaction elicited against ovalbumin. Data show the difference in challenged and non-challenged ear thickness in vehicle-treated rats and ShK[K18A]-treated rats (N = 6/group).

9.A Appendix for Chapter 9

Lambda values used in charge/discharge FEP calculations. 67 exponentially spaced lambda values were used:

0.000000000 1.0000004e-03 5.9452732e-03 1.1209489e-02 1.6813217e-02 2.2778355e-02 2.9128211e-02 3.5887597e-02 4.3082927e-02 5.0742315e-02 5.8895691e-02 6.7574915e-02 7.6813901e-02 8.6648751e-02 9.7117893e-02 1.0826224e-01 1.2012533e-01 1.3275353e-01 1.4619617e-01 1.6050579e-01 1.7573831e-01 1.9195323e-01 2.0921393e-01 2.2758785e-01 2.4714678e-01 2.6796716e-01 2.9013033e-01 3.1372291e-01 3.3883708e-01 3.6557097e-01 3.9402905e-01 4.2432252e-01 4.5656975e-01 0.5000000000 5.4343025e-01 5.7567748e-01 6.0597095e-01 6.3442903e-01 6.6116292e-01 6.8627709e-01 7.0986967e-01 7.3203284e-01 7.5285322e-01 7.7241215e-01 7.9078607e-01 8.0804677e-01 8.2426169e-01 8.3949421e-01 8.5380383e-01 8.6724647e-01 8.7987467e-01 8.9173776e-01 9.0288211e-01 9.1335125e-01 9.2318610e-01 9.3242508e-01 9.4110431e-01 9.4925769e-01 9.5691707e-01 9.6411240e-01 9.7087179e-01 9.7722165e-01 9.8318678e-01 9.8879051e-01 9.9405473e-01 9.9900000e-01 1.0000000000

Lambda values used in charge/discharge TI calculations (seven point Gaussian quadrature):

0.02544 0.12923 0.29707 0.50000 0.70292 0.87076 0.97455

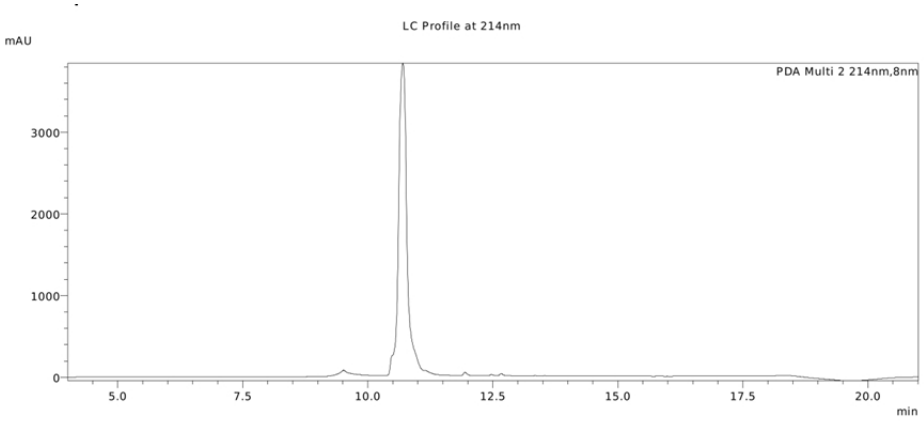

Figure 9A.1: The LC profile of ShK[K18A].

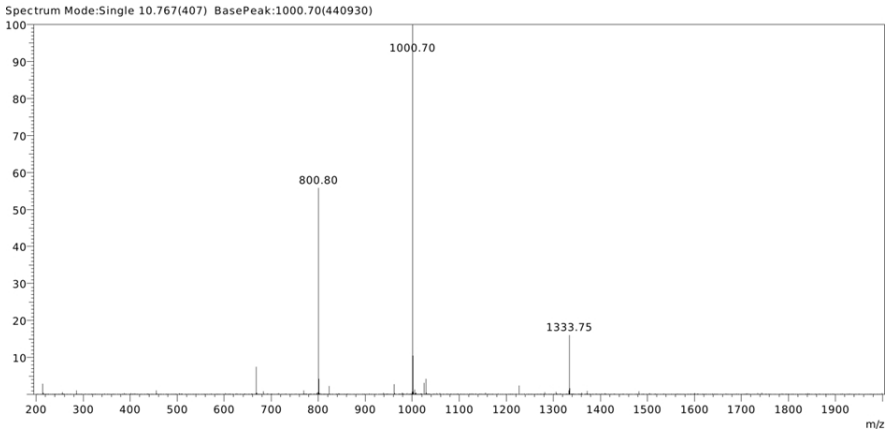

Figure 9A.2: ESI-MS analysis of ShK[K18A]. The mass of the peptide calculated from the predicted amino acid sequence (3998 Da), compared well to the electrospray mass spectrometry (ESI-MS, 3998 Da).

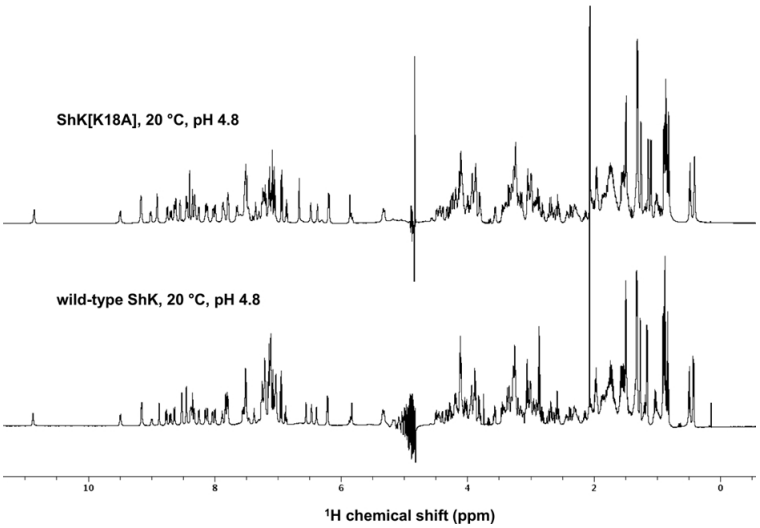

Figure 9A.3: The 1D ^{1}H NMR spectrum of wild-type ShK and ShK[K18A] at pH 4.8, acquired on a Bruker Avance 600 MHz spectrometer at 20°C.

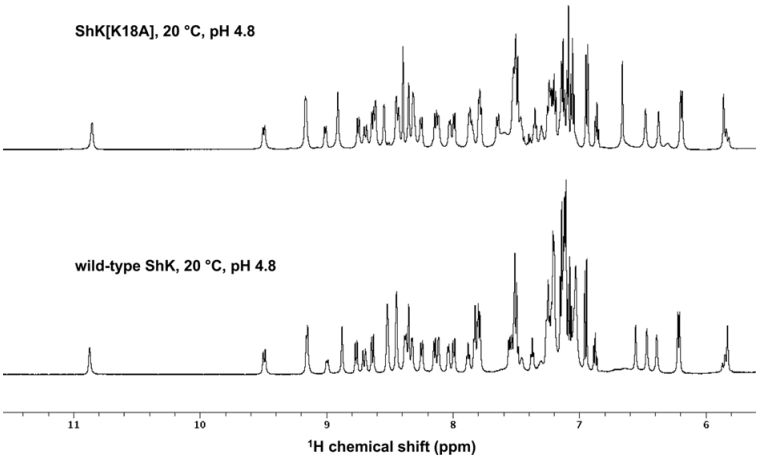

Figure 9A.4: The amide and aromatic region of ^{1}H NMR spectra of wild-type ShK and ShK[K18A] at pH 4.8, acquired on a Bruker Avance 600 MHz spectrometer at 20°C.

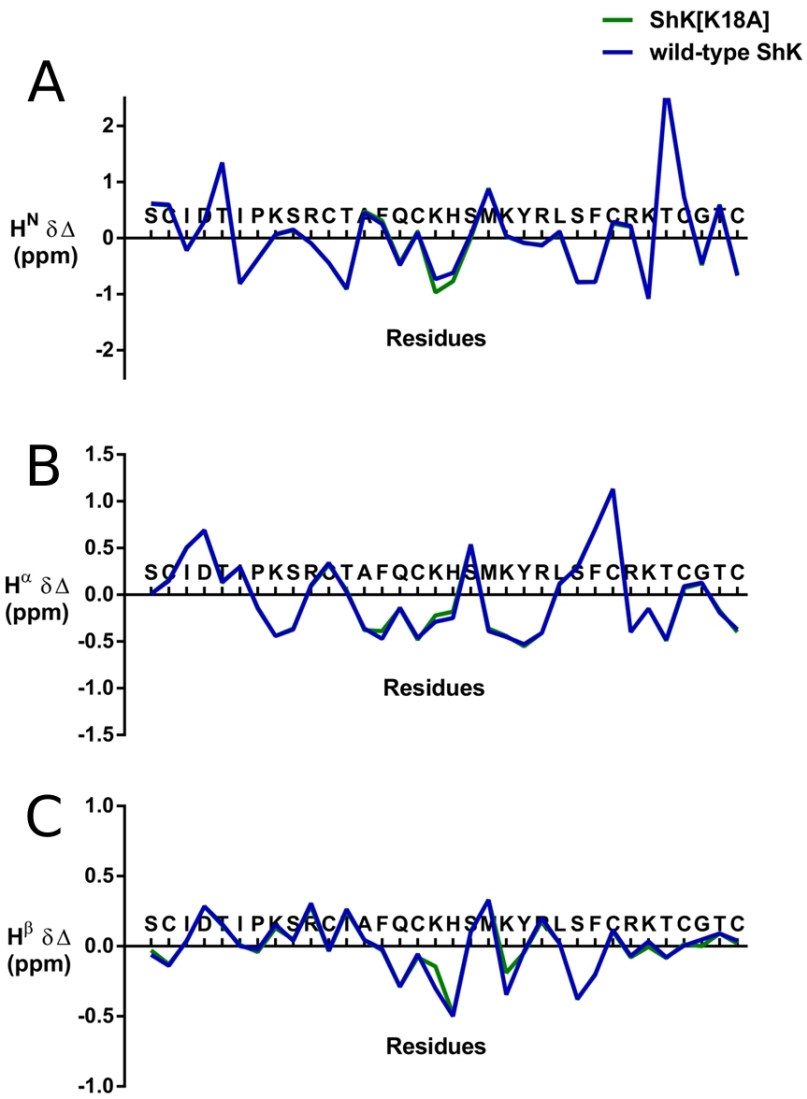

Figure 9A.5: Deviation from random coil chemical shifts (4) of the H^N (A), H^α (B), and H^β (C) resonances of wild-type ShK and ShK[K18A] at 20°C and pH 4.8.

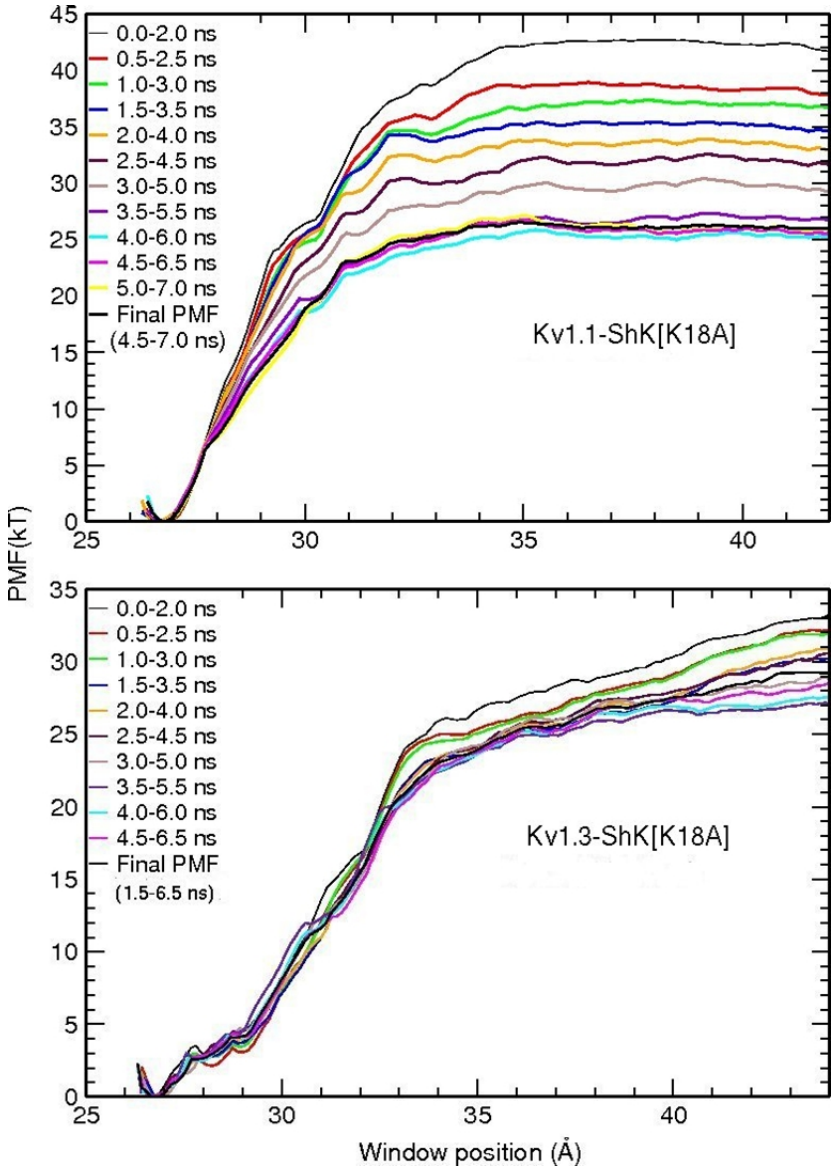

Figure 9A.6: Convergence of the Kv1.x-ShK[K18A] PMFs from 2 ns block data analysis. To reduce fluctuations, we use a large sampling size (2 ns), which is slid in 0.5 ns steps over the range of the data. In Kv1.1, the PMFs drop monotonically until 3.5-5.5 ns, after which they fluctuate around a base line. In Kv1.3, equilibration occurs earlier, presumably due to the stronger binding. To decide more precisely which part of the data should be discarded for equilibration, we have constructed PMFs from 0.5 ns blocks in Figure 9A.7.

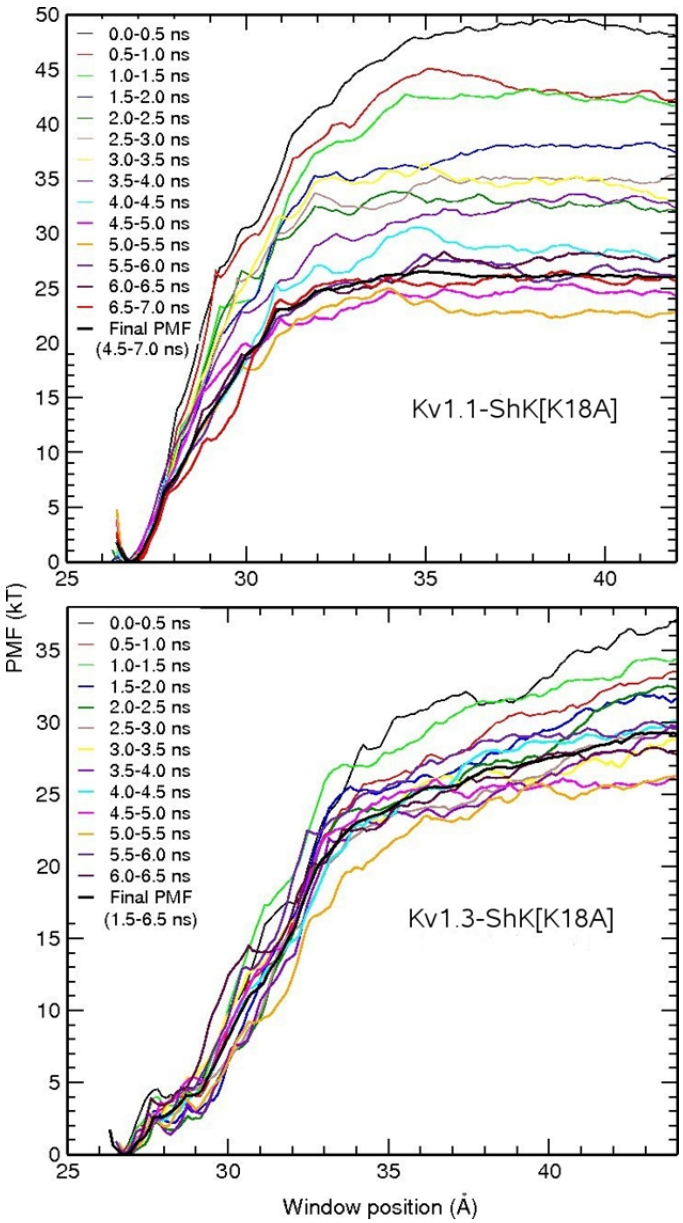

Figure 9A.7: Convergence of the Kv1.x-ShK[K18A] PMFs from 0.5 ns block data analysis. In Kv1.1, the first 4.5 ns of data are discarded as equilibration, and the final PMF is determined from the last 2.5 ns of data. In Kv1.3, the first 1.5 ns of data are discarded, and the final PMF is determined from the last 5 ns of data.

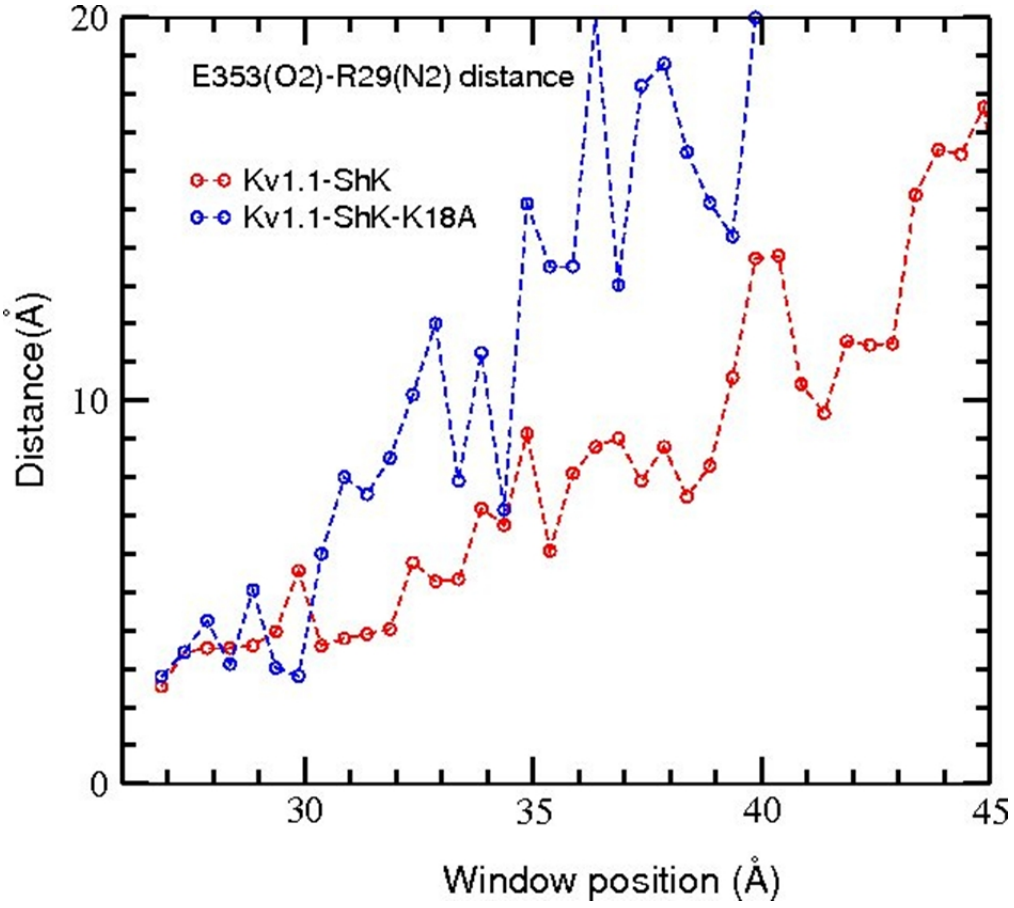

Figure 9A.8: The E353(O_2)R29(N_2) pair distances in ShK and ShK[K18A] PMFs as a function of the channel-toxin distance in Kv1.1. In the case of ShK[K18A], the pair decouple for $z > 35$ Å, which explains the earlier flattening of the corresponding PMF in Figure 9.4. The bump in the Kv1.1-ShK PMF around $z = 43$ Å (Figure 9.4) can be traced back to the temporary re-engagement of R29 with E353 facilitated by rotation of the toxin.

Chapter 10

Conclusions and future work

In this thesis we have performed all-atom molecular dynamics simulations on different proteins related to the process of neurotransmission in the brain. We focus on glutamate transporters and receptors, as well as potassium channels. These structures are related to a number of neurodegenerative conditions and auto-immune diseases, and therefore serve as potential drug targets. In order to investigate these proteins, we combine classical molecular dynamics with a variety of methods in computational biology. We parameterize drug-like molecules for use with the CHARMM force-field using *ab initio* calculations, and use archaeal homologs to build protein models that do not have crystal structures available. We also dock ligands to proteins to find the correct binding modes, which are used in binding free energy calculations. In order to calculate the ligand binding free energies and the protein conformational changes, we perform free energy perturbation calculations with restraints, as well as potential of mean force calculations.

The first structure we investigate is an archaeal homolog of the human glutamate transporters, called Glt_{Ph}, which is mainly an aspartate transporter. Glt_{Ph} displays many similarities with the EAATs, and therefore can provide important insights into the mechanism of the human transporters. Like the EAATs, Glt_{Ph} co-transports the substrate with three sodium ions, even though only two of them could be identified in the Glt_{Ph} crystal structure. We locate the binding site of the missing sodium ion (Na3), providing experimental evidence that this binding site is present in both Glt_{Ph} and the EAATs. We then perform MD simulations in the outward-facing Glt_{Ph} including the Na3 ion, showing the coordination of all the ligands, their order of binding, the nature of the sodium-aspartate coupling, and the reasons behind the aspartate/glutamate selectivity. We follow by performing simulations in the inward-facing structure of Glt_{Ph}, searching for an understanding of the mechanism of ions and substrate release. Our results demonstrate that the binding affinity of the ligands is symmetrical between the inward and outward states, as well as the order of binding/unbinding, with the ligands that bind last being released first. In the absence of any crystal structures of Glt_{Ph} in the inward open state, we obtain this conformation by simulating the inward Glt_{Ph} with different combinations of

ligands bound. The simulations show the inward gating involving both the HP1 and HP2 segments, as opposed to the large HP2 movement observed in the outward state. Free energy calculations also show that the sodium-aspartate coupling is mostly electrostatic, and that the large affinity of the Na3 ion might be behind the long turnover rates of $\mathrm{Glt_{Ph}}$. Our thorough investigation of the $\mathrm{Glt_{Ph}}$ properties enabled us to build homology models for the human transporter EAAT3 based on this archaeal homolog, which we studied next.

The EAATs and $\mathrm{Glt_{Ph}}$, although having similar binding sites for the sodium ions and the substrate, as well as tertiary structures, have a few important differences. One of them is that, in addition to the co-transport of three sodium ions, the EAATs also transport one proton with glutamate and counter-transport one potassium ion at each cycle. The EAATs also have a much faster turnover rate compared to $\mathrm{Glt_{Ph}}$, which might be related to these two additional ligands. This has prompted us to create models for the EAAT3 transporter based on $\mathrm{Glt_{Ph}}$, and study the properties of the former by using MD simulations and FEP/TI calculations. The fully-bound EAAT3 models we produce are very stable, and so is the coordination of the ligands, even after several nanoseconds of simulations. By removing the substrate, we find that the gating motions of EAAT3 and $\mathrm{Glt_{Ph}}$ in the outward-facing states are very similar, but there is a larger opening in the HP1 segment in the inward state of EAAT3 compared to $\mathrm{Glt_{Ph}}$. We trace that to the relocation of an arginine residue from the HP1 segment in $\mathrm{Glt_{Ph}}$ to the TM8 segment in the EAATs. To locate the potassium binding site in the EAATs, we test three potential sites for this ion as suggested by experiments; either overlapping with the substrate, the proton, or two sodium ions. We find that this last site has the lowest binding free energy, and that sodium/potassium exchange could explain the faster transport of the EAATs compared to $\mathrm{Glt_{Ph}}$.

In our EAAT3 model glutamate is not stably bound with a deprotonated E374 side-chain, making this residue a good candidate for the proton carrier, which is also corroborated by experiments. We perform rigorous $\mathrm{p}K_\mathrm{a}$ calculations using FEP/TI on all the titratable residues close to the substrate binding site in the fully bound EAAT3, and find that E374 is the only one protonated in these conditions. We then calculate the $\mathrm{p}K_\mathrm{a}$ of this residue in different states of EAAT3 and with different ligands bound, to elucidate the mechanism of proton transport. We find that the binding and release of the proton is dependent on the presence of the substrate and the closure of the HP1/HP2 gates, and that potassium binding in the inward facing state is necessary for proton release.

The next protein we investigate is the GluA2 receptor, which has high-resolution crystal structures available in complex with many ligands, and also the experimental binding affinities for each of them. We calculate the binding free energies of five ligands to GluA2, including two ligands with a net charge, using alchemical FEP/TI with restraints combined with PMF calculations. Alchemical FEP/TI methods are advantageous over PMF methods for binding sites that are deeply buried inside proteins, but they also present problems for ligands that have charges. Here we

show that these calculations are possible using FEP, as long as we perform the decoupling/coupling of charges in the bulk/binding site in the same system, which preserves the neutrality of the periodic cell. We also apply the confine-and-release method, which is a useful tool when there are large hysteresis effects in the forward and backward FEP calculations. We obtain results close to the experimental values for all the ligands, and therefore our method provides an efficient and accurate way of computing the binding free energies of small molecules. These calculations have the potential to become a useful tool in the final stages of rational drug design, since they can be applied to a number of different systems. In the case of the GluA2 receptor, we also gained insights into the components of the binding free energy of each ligand, since our FEP/TI approach separates the free energies arising from Lennard-Jones and electrostatic interactions.

In addition to the binding affinity of small molecules, alchemical FEP/TI also provides a good solution to calculate the binding free energy differences between wild-type and mutant toxin peptides to ion channels. In this case, to obtain good convergence and accurate results, the toxin binding modes should not be affected by the mutation. Here we apply alchemical FEP/TI to increase the selectivity of the sea-anemone ShK toxin for the Kv1.3 channel over Kv1.1, which is a problem for the wild-type toxin. We successfully create a ShK mutant that is a selective blocker of the Kv1.3 channel over Kv1.1, and therefore may provide a good alternative in the treatment of auto-immune diseases. The results we obtain are also in good agreement with PMF calculations and experiments, even though FEP/TI is less computationally costly compared to PMF methods. Therefore, this approach also provides a good alternative for rational drug design, but for small peptides instead of drug-like molecules.

There are still many unanswered questions related to the aspartate/glutamate transporters. Even though there are crystal structures in the inward and outward facing states of Glt_{Ph}, with many related MD studies, the outward to inward transition mechanism remains unknown. We plan as future work to perform multidimensional potential of mean force calculations on this transition, with different ligands bound to the transporter, to understand the mechanism of substrate translocation in Glt_{Ph} and the EAATs. Glt_{Ph} is also a homolog of the human neutral amino acid transporters (ASCTs), and we have recently used a ASCT1 model created by us to help interpret mutagenesis experiments in this transporter. We also plan on employing the FEP/TI calculations we did on the glutamate receptor to other proteins that are potential drug targets, turning this method into an automated protocol that could be quickly applied to any system. By combining this protocol with homology modeling, parameterization of molecules and ligand docking, we aim to have a complete set of tools that can be used on computer-aided drug design. In the case of peptide toxins, there is a large database of potassium channel blockers that can be engineered to bind to a particular channel based on MD simulations, opening up the possibility of new drugs developed from toxins found in the animal kingdom.

Chapter 11

Bibliography

1. Kolb, B., and I. Q. Whishaw. 2003. *Fundamentals of Human Neuropsychology.* Worth Publishers, London, UK.

2. Purves, D, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A. LaMantia, J. O. McNamara, and S. M. Williams. 2001. *Neuroscience.* Sinauer Associates, Sunderland, MA, USA.

3. Hayashi, T. 1954. Effects of sodium glutamate on the nervous system *Keio J. Med.* 3:183–192.

4. Curtis, D. R., J. W. Phillis, and J. C. Watkins. 1959. Chemical excitation of spinal neurons. *Nature* 183:611–612.

5. Sugiyama, H., I. Ito, and M. Watanabe. 1989. Glutamate receptor subtypes may be classified into two major categories: A study on Xenopus oocytes injected with rat brain mRNA. *Neuron* 3:129–132.

6. Dingledine, R., K. Borges, D. Bowie, and S. F. Traynelis. 1999. The glutamate receptor ion channels. *Pharmacol. Rev.* 51:7–61.

7. Armstrong, N., and E. Gouaux. 2000. Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core. *Neuron* 28:165–181.

8. Chaudhry, C., M. C. Weston, P. Schuck, C. Rosenmund, and M. L. Mayer. 2009. Stability of ligand-binding domain dimer assembly controls kainate receptor desensitization. *Embo J.* 28:1518–1530.

9. Jin, R., S. K. Singh, S. Gu, H. Furukawa, A. I. Sobolevsky, J. Zhou, Y. Jin, and E. Gouaux. 2009. Crystal structure and association behaviour of the GluR2 amino-terminal domain. *Embo J.* 28: 1812–1823.

10. Kumar, J., and M. L. Mayer. 2010. Crystal structures of the glutamate receptor ion channel GluK3 and GluK5 amino-terminal domains. *J. Mol. Biol.* 404:680–696.

11. Sobolevsky, A. I., M. P. Rosconi, and E. Gouaux. 2009. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* 462:745–756.

12. Danbolt, N. C. 2001. Glutamate uptake. *Prog. Neurobiol.* 65:1–105.

13. Zerangue, N. and M. P. Kavanaugh. 1996. Flux coupling in a neuronal glutamate transporter. *Nature.* 383:634–637.

14. Yernool, D., O. Boudker, Y. Jin, and E. Gouaux. 2004. Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature* 431:811–818.

15. Boudker, O., R. M. Ryan, D. Yernool, K. Shimamoto, and E. Gouaux. 2007. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature.* 445:387–393.

16. Reyes, N., C. Ginter, and O. Boudker. 2009. Transport mechanism of a bacterial homologue of glutamate transporters. *Nature.* 462:880–885.

17. Hille, B. 2001. *Ionic Channels of Excitable Membranes.* Sinauer Associates, Sunderland, MA, USA.

18. Beeton, C., H. Wulff, N. E. Standifer, P. Azam, K. M. Mullen, M. W. Pennington, A. Kolski-Andreaco, E. Wei, A. Grino, D. R. Counts *et al.*. 2006. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proc. Natl. Acad. Sci. USA.* 103:17414–17419.

19. Rossi, D.J., T. Oshima, and D. Attwell. 2000. Glutamate release in severe brain ischaemia is mainly by reversed uptake. *Nature* 403:316–321.

20. Rothstein, J.D., M. Van Kammen, A. I. Levey, L. J. Martin, and R. W. Kuncl. 1995. Selective loss of glial glutamate transporter GLT-1 in amyotrophic lateral sclerosis. *Ann. Neurol.* 38:73–84.

21. Li, S., M. Mallory, M. Alford, S. Tanaka, and E. Masliah. 1997. Glutamate transporter alterations in Alzheimer disease are possibly associated with abnormal APP expression. *J. Neuropathol. Exp. Neurol.* 56:901–911.

22. During, M.J., and D. D. Spencer. 1993. Extracellular hippocampal glutamate and spontaneous seizure in the conscious human brain. *Lancet* 341:1607–1610.

23. Tanaka, K., K. Watase, T. Manabe, K. Yamada, M. Watanabe, K. Takahashi, H. Iwama, T. Nishikawa, N. Ichihara, S. Hori, M. Takimoto, and K. Wada. 1997. Epilepsy and exacerbation of brain injury in mice lacking the glutamate transporter GLT-1. *Science* 276:1699–1702.

24. Kanai, Y., C. P. Smith, and M. A. Hediger. 1993. A new family of neurotransmitter transporters the high-affinity glutamate transporters. *FASEB J.* 7:1450–1459.

25. Grunewald, M., A. Bendahan, and B. I. Kanner. 1998. Biotinylation of single cysteine mutants of the glutamate transporter GLT-1 from rat brain reveals its unusual topology. *Neuron* 21:623–632.

26. Grunewald, M., and B. I. Kanner. 2000. The accessibility of a novel reentrant loop of the glutamate transporter GLT-1 is restricted by its substrate. *J. Biol. Chem.* 275:9684–9689.

27. Danbolt, N. C., K. P. Lehre, Y. Dehnes, F. A. Chaudhry, L. M. Levy. 1998. Localization of transporters using transporter-specific antibodies. *Methods Enzymol.* 296:388–407.

28. Kavanaugh, M. P., A. Bendahan, N. Zerangue, Y. Zhang, and B. I. Kanner. 1997. Mutation of an amino acid residue influencing potassium coupling in the glutamate transporter GLT-1 induces obligate exchange. *J. Biol. Chem.* 272:1703–1708.

29. Wadiche, J. I., S. G. Amara, and M. P. Kavanaugh. 1995. Ion fluxes associated with excitatory amino acid transport. *Neuron* 15:721–728.

30. Tao, Z., Z. Zhang, and C. Grewer. 2006. Neutralization of the aspartic acid residue Asp-367, but not Asp-454, inhibits binding of Na^+ to the glutamate-free form and cycling of the glutamate transporter EAAC1. *J. Biol. Chem.* 281:10263–10272.

31. Tao, Z., N. Rosental, B. I. Kanner, A. Gameiro, J. Mwaura, and C. Grewer. 2010. Mechanism of cation binding to the glutamate transporter EAAC1 probed with mutation of the conserved amino acid residue Thr^{101}. *J. Biol. Chem.* 285:17725–17733.

32. Bendahan, A., A. Armon, N. Madani, M. P. Kavanaugh, and B. I. Kanner. 2000. Arginine 447 plays a pivotal role in substrate interactions in a neuronal glutamate transporter. *J. Biol. Chem.* 275:37436–37442.

33. Teichman, S., and B. I. Kanner. 2007. Aspartate-444 is essential for productive substrate interactions in a neuronal glutamate transporter. *J. Gen. Physiol.* 129:527–539.

34. Zhang, Y. and B. I. Kanner. 1999. Two serine residues of the glutamate transporter GLT-1 are crucial for coupling the fluxes of sodium and the neurotransmitter. *Proc. Natl. Acad. Sci. USA.* 96:1710–1715.

35. Ryan, R. M., E. L. Compton, and J. A. Mindell. 2009. Functional characterization of a Na^+-dependent aspartate transporter from Pyrococcus horikoshii. *J. Biol. Chem.* 284:17540–17548.

36. Huang, Z., and E. Tajkhorshid. 2008. Dynamics of the extracellular gate and ion-substrate coupling in the glutamate transporter. *Biophys. J.* 95:2292–2300.

37. Groeneveld, M., and D. J. Slotboom. 2010. Na^+:aspartate coupling stoichiometry in the glutamate transporter homologue Glt(Ph). *Biochemistry.* 49:3511–3513.

38. Bastug, T., G. Heinzelmann, S. Kuyucak, M. Salim, R. J. Vandenberg, and R. M. Ryan. 2012. Position of the Third Na Site in the Aspartate Transporter Glt_{Ph} and the Human Glutamate Transporter, EAAT1. *PLoS ONE* 7:e33058.

39. DeChancie, J., I. H. Shrivastava, and I. Bahar. 2011. The mechanism of sustrate release by the aspartate transporter Glt_{Ph}: insights from simulations. *Mol. BioSyst.* 7:832–842.

40. Verdon, G., and O. Boudker. 2012. Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog. *Nat.Struct.Mol.Biol.* 19:355–357.

41. Lau, A. Y., and B. Roux. 2011. The Hidden Energetics of Ligand Binding and Activation in a Glutamate Receptor. *Nat. Struc. Mol. Biol.* 18:283–287.

42. Wolter, T., T. Steinbrecher, and M. Elstner. 2013. Computational Study of Synthetic Agonist Ligands of Ionotropic Glutamate Receptors. *PLoS ONE* 8:e58774.

43. Woo, H., and B. Roux. 2005. Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. *Proc. Nat. Acad. Sci. U.S.A.* 102:6825–6830.

44. Felipe, A., J. Bielanska, N. Comes, A. Vallejo, S. Roig, S. Ramn Y Cajal, E. Condom, J. Hernndez-Losa, and J. C. Ferreres. 2012. Targeting the voltage-dependent K^+ channels Kv1.3 and Kv1.5 as tumor biomarkers for cancer detection and prevention. *Curr. Med. Chem.* 19:661–674.

45. Norton, R.S., M. W. Pennington, and H. Wulff. 2004. Potassium channel blockade by the sea anemone toxin ShK for the treatment of multiple sclerosis and other autoimmune diseases. *Curr. Med. Chem.* 11:3041–3052.

46. Beeton, C., M. W. Pennington, and R. S. Norton. 2011. Analogs of the sea anemone potassium channel blocker ShK for the treatment of autoimmune diseases. *Inflamm. Allergy Drug Targets* 10:313–321.

47. Rashid, M. H., and S. Kuyucak. 2012. Afinity and selectivity of ShK toxin for the Kv1 potassium channels from free energy simulations. *J. Phys. Chem. B* 116:4812–4822.

48. Einstein, A. 1905. Über die von der molekularkinetischen Theorie der Wrme geforderte Bewegung von in ruhenden Flssigkeiten suspendierten Teilchen. *Annalen der Physik* 17:549–560.

49. Kohn, W., and L. J. Sham. 1965. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* 140:1133–1138.

50. Car, R., and M. Parrinello. 1985. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55:2471–2474.

51. Karplus, M., and J. A. McCammon. 2002. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9:646–652.

52. Beveridge, D. L., and F. M. DiCapua. 1989. Free energy via molecular simulation: applications to chemical and biomolecular systems. *Ann. Rev. Biophys. Biophys. Chem.* 18:431–492.

53. Torrie, G.M., and J. P. Valleau. 1997. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comp. Phys.* 23:187–199.

54. Laio, A., and M. Parrinello. 2002. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA.* 99:12562–12566.

55. MacKerell, A. D., D. Bashford, R. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, *et al.* 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586–3616.

56. Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25:1157–1174.

57. Christen, M., P. H. Hunenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Krutler, C. Oostenbrink, C. Peter, D. Trzesniak, and W. F. van Gunsteren. 2005. The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem.* 26:1719–1751.

58. Lennard-Jones, J.E. 1924. On the determination of molecular fields. II. The equation of state of a gas. *Proc. Roy. Soc. A.* 106:463–477.

59. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.

60. Ryckaert, J., G. Ciccotti, and H. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints:molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.

61. Frenkel, D., and B. Smit. 2002. *Understanding Molecular Simulation From Algorithms to Applications.* Academic Press, California, USA.

62. Kubo, R., M. Toda, and N. Hashitsume. 1991. *Statistical Physics II: Nonequilibrium Statistical Mechanics.* Springer.

63. Brnger, A., C. B. Brooks, and M. Karplus. 1984. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.* 105:495–500.

64. Feller, S. E., Y. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant-pressure molecular-dynamics simulation–The Langevin piston method. *J. Chem. Phys.* 103:4613–4621.

65. Martyna, G. J., D. J. Tobias, and M. L. Klein. 1994. Constant-pressure molecular-dynamics algorithm. *J. Chem.Phys.* 101:4177–4189.

66. Ewald, P. 1921. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys. Berlin.* 369:253–287.

67. Darden, T. A., D. M. York, and L. G. Pedersen. 1993. Particle mesh Ewald: an N log (N) method for Ewald sums in large systems. *J. Chem.Phys.* 98:10089–10092.

68. Chen, R., and Z. Weng. 2002. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins* 47:281–294.

69. Dominguez, C., R. Boelens, and A. M. Bonvin. 2003. HADDOCK: a protein-protein docking approach based on biochemical and/or biophysical information. *J. Am. Chem. Soc.* 125:1731–1737.

70. Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell. 2010. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Field. *J. Comput. Chem.* 31:671–690.

71. O'Boyle, N. M., M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison. 2011. Open Babel: An open chemical toolbox. *J. Cheminform.* 3:33.

72. Frisch, M. J. et al. 2009. Gaussian 09, Revision D.01.

73. Pulay, P., G. Fogarasi, F. Pang, and J. E. Boggs. 1979. *J. Am. Chem. Soc.* 101:2550–2560.

74. Eswar, N., D. Eramian, B. Webb, M. Shen, and A. Sali. 2008. Protein structure modeling with MODELLER. *Methods Mol. Biol.* 426:145–159.

75. Sali, A., and T. L. Blundell. 1990. Definition of general topological equivalence in protein structures: A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J. Mol. Biol.* 212:403-428.

76. Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302:205–217.

77. Altschul, S. F., T. L. Madden, A. A. Schffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

78. Sali, A., and T. L. Blundell. 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* 234:779-815.

79. Benkert, P., M. Biasini, and T. Schwede. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27:343–350.

80. Zhou, H., and Y. Zhou. 2002. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Sci.* 11:2714–2726.

81. Kumar, S. J., M. Rosenberg, M., D. Bouzida, R. H. Swendsen, and P. A. Kollman. 1992. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* 13:1011–1021.

82. Zacharias, M., T. P. Straatsma, and J. A. McCammon. 1994. Separation-Shifted Scaling, a New Scaling Method for Lennard-Jones Interactions in Thermodynamic Integration. *J. Chem. Phys.* 100:9025–9031.

83. Darve, E., D. Rodríguez-Gómez, and A. Pohorille. 2008. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* 128:144120.

84. Chen, P.C., and S. Kuyucak. 2011. Accurate determination of the binding free energy for KcsA-Charybdotoxin complex from the potential of mean force calculations with restraints. *Biophys. J.* 100:2466–2474.

85. Kanner, B. I. and A. Bendahan. 1982. Binding order of substrates to the sodium and potassium-ion coupled L-glutamic acid transporter from rat-brain. *Biochemistry.* 21:6327–6330.

86. Herman, M. A., and C. E. Jahr. 2007. Extracellular glutamate concentration in hippocampal slice. *J. Neurosci.* 27:9736–9741.

87. Levy, L. M., O. Warr, and D. Atwell. 1998. Stoichiometry of the glial glutamate transporter GLT-1 expressed inducibly in a chinese hamster ovary cell line selected for low endogenous Na^+-dependent glutamate uptake. *J. Neurosci.* 18:9620–9628.

88. Fairman, W. A., R. J. Vandenberg, J. L. Arriza, M. P. Kavanaugh, and S. G. Amara. 1995. An excitatory amino-acid transporter with properties of a ligand-gated chloride channel. *Nature* 375: 599–603.

89. Wadiche, J. I., and M. P. Kavanaugh. 1998. Macroscopic and microscopic properties of a cloned glutamate transporter/chloride channel. *J. of Neurosc.* 18:7650–7661.

90. Vandenberg, R. J., J. L. Arriza, S. G. Amara, and M. P. Kavanaugh. 1995. Constitutive ion fluxes and substrate binding domains of human glutamate transporters. *J. Biol. Chem.* 270:17668–17671.

91. Ryan, R. M., and J. A. Mindell. 2007. The uncoupled chloride conductance of a bacterial glutamate transporter homolog. *Nat. Struct. Mol. Biol.* 14:365–371.

92. Rosental, N., A. Bendahan, and B. I. Kanner. 2006. Multiple consequences of mutating two conserved beta-bridge forming residues in the translocation cycle of a neuronal glutamate transporter. *J. Biol. Chem.* 281:27905–27915.

93. Ryan, R. M., A. D. Mitrovic, and R. J. Vandenberg. 2004. The chloride permeation pathway of a glutamate transporter and its proximity to the glutamate translocation pathway. *J. Biol. Chem.* 279:20742–20751.

94. Seal, R. P., B. H. Leighton, and S. G. Amara. 2000. A model for the topology of excitatory amino acid transporters determined by the extracellular accessibility of substituted cysteines. *Neuron.* 25:695–706.

95. Holley, D. C., and M. P. Kavanaugh. 2009. Interactions of alkali cations with glutamate transporters. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364:155–161.

96. Larsson, H. P., X. Y. Wang, B. Lev, I. Baconguis, D. A. Caplan, N. P. Vyleta, H. P. Koch, A. Diez-Sampedro, and S. Y. Noskov. 2010. Evidence for a third sodium-binding site in glutamate transporters suggests an ion/substrate coupling model. *Proc. Natl. Acad. Sci. USA.* 107:13912–13917.

97. Huang, Z. and E. Tajkhorshid. 2010. Identification of the third Na^+ site and the sequence of extracellular binding events in the glutamate transporter. *Biophys. J.* 99:1416–1425.

98. Schrodinger, L. L. C. 2010. The PyMOL Molecular Graphics System, Version 1.3r1.

99. Shrivastava, I. H., J. Jiang, S. G. Amara, and I. Bahar. 2008. Time-resolved mechanism of extracellular gate opening and substrate binding in a glutamate transporter. *J. Biol. Chem.* 283:28680–28690.

100. Nayal, M. and E. di Cera. 1994. Predicting Ca^{2+}-binding sites in proteins. *Proc. Natl. Acad. Sci. USA.* 91:817–821.

101. Zarbiv, R., M. Grunewald, M. P. Kavanaugh, and B. I. Kanner. 1998. Cysteine scanning of the surroundings of an alkali-ion binding site of the glutamate transporter GLT-1 reveals a conformationally sensitive residue. *J. Biol. Chem.* 273:14231–14237.

102. Higgins D. G., M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, *et al.* 2007. Clustal W and clustal X version 2.0. *Bioinformatics* 23:2947–2948.

103. Meier, T., P. Polzer, K. Diederichs, W. Welte, and P. Dimroth. 2005. Structure of the rotor ring of F-Type Na^+-ATPase from Ilyobacter tartaricus. *Science* 308:659–662.

104. Humphrey, W., A. Dalke and K. Schulten. 1996. VMD - Visual Molecular Dynamics. *J. Molec. Graphics* 14:33–38.

105. MacKerell, A. D., M. Feig, and C. L. Brooks. 2004. Extending the treatment of backbone energetics in protein force fields. *J. Comput. Chem.* 25:1400–1415.

106. Boresch, S., F. Tettinger, M. Leitgeb, and M. Karplus. 2003. Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B.* 107:9535–9551.

107. Karplus, M., and J. N. Kushick. 1981. Methods for estimating the configuratonal entropy of macromolecules. *Macromolecules.* 14:325–332.

108. Bastug, T., and S. Kuyucak. 2006. Energetics of ion permeation, rejection, binding and block in gramicidin A from free energy simulations. *Biophys. J.* 90:3941–3950.

109. McKinzie, A. A., R. M. Ryan, and R. J. Vandenberg. 2010. Site-directed mutagenesis in the study of membrane transporters. *Methods Mol. Biol.* 637:277–293.

110. Gaillard, I., D. Slotboom, J. Knol, J. S. Lolkema, and W. N. Konings. 1996. Purification and reconstitution of the glutamate carrier GltT of the thermophillic bacterium Bacillus stearothermophillus. *Biochemistry* 35:6150–6156.

111. Vandenberg, R. J., A. D. Mitrovic, G. A. Johnston. 1998. Serine-O-sulphate transport by the human glutamate transporter, EAAT2. Br. J. Pharmacol. 123:1593–1600.

112. Forrest, L. R., R. Krämer, and C. Ziegler. 2011. The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta.* 1807:167–188.

113. Jiang, J., and S. G. Amara. 2011. New views of glutamate transporter structure and function: Advances and challenges. *Neuropharmacology.* 60:172–181.

114. Gu, Y., I. H. Shrivastava, S. G. Amara, and I. Bahar. 2009. Molecular simulations elucidate the substrate translocation pathway in a glutamate transporter. *Proc. Natl. Acad. Sci. USA.* 106:2589–2594.

115. Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon. 1997. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* 72:1047–1069.

116. Deng, Y., and B. Roux. 2006. Calculation of standard binding free energies : Aromatic molecules in T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* 2:1255–1273.

117. Luo, H, and K. Sharp. 2002. On the calculation of absolute macromolecular binding energies. *Proc. Natl. Acad. Sci. USA.* 99:10399–10404.

118. Carlson, J., and J. Aqvist. 2005. Absolute and relative entropies from computer simulation with applications to ligand binding. *J. Phys. Chem. B.* 109:6448–6456.

119. Minh, D. D. L., J. M. Bui, C. Chang, T. Jain, J. M. J. Swanson, and J. A. McCammon. 2005. The entropic cost of protein-protein association. *Biophys. J.* 89:L25–L27.

120. Cecchini, M., S. V. Krivov, M. Spichty, and M. Karplus. 2009. Calculation of free-energy differences by confinement simulations. Application to peptide conformers. *J. Phys. Chem. B.* 113: 9728–9740.

121. Yamashita, A., S. K. Singh, T. Kawate, Y. Jin, and E. Gouaux. 2005. Crystal Structure of LEUTAA, a bacterial homolog of Na^+/Cl^- dependent neurotransmitter transporters. *Nature.* 437: 215–223

122. Crisman, T. J., S. Qu, B. I. Kanner, and L. R. Forrest. 2009. Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. *Proc. Natl. Acad. Sci. USA.* 106:20752–20757.

123. Heinzelmann, G., T. Bastug, and S. Kuyucak. 2011. Free Energy Simulations of Ligand Binding to the Aspartate Transporter Glt_{Ph}. *Biophys. J.* 101:2380–2388.

124. Jiang, J., I. H. Shrivastava, S. D. Watts, I. Bahar, and S. G. Amara. 2011. Large collective motions regulate the functional properties of glutamate transporter trimers. *Proc. Natl. Acad. Sci. USA* 108:15141–15146.

125. Stolzenberg, S., G. Khelashvili, and H. Weinstein. 2012. Structural Intermediates in a Model of the Substrate Translocation Path of the Bacterial Glutamate Transporter Glt_{Ph}. *J. Phys. Chem. B.* 116:5372–5383.

126. Grazioso, G., V. Limongelli, D. Branduardi, E. Novellino, C. De Micheli, A. Cavalli, and M. Parrinello. 2012. Investigating the Mechanism of Substrate Uptake and Release in the Glutamate Transporter Homologue Glt_{Ph} Through Metadynamics Simulations. *J. Am. Chem. Soc.* 134:453–463.

127. Kramers, H. A. 1940. Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions. *Physica* 7:284–304.

128. Berry, R. S., S. A. Rice, and J. Ross. *Physical Chemistry*, Oxford University Press, Oxford, 2000.

129. Ryan, R. M., N. C. Kortt, T. Sirivanta, and R. J. Vandenberg. 2010. The position of an arginine residue influences substrate affinity and K^+ coupling in the human glutamate transporter, EAAT1. *J. Neurochem.* 114:565–575.

130. Huang, S., R. M. Ryan, and R. J. Vandenberg RJ. 2009. The role of cation binding in determining substrate selectivity of glutamate transporters. *J. Biol. Chem.* 284:4510–4515.

131. Bergles, D. E., A. V. Tzingounis, and C. E. Jahr CE. 2002. Comparison of coupled and uncoupled currents during glutamate uptake by GLT-1 transporters. *J. Neurosci.* 22:10153–10162.

132. Zhang, Y., A. Bendahan, R. Zarbiv, M. P. Kavanaugh, and B. I. Kanner. 1998. Molecular determinant of ion selectivity of an (Na^+ + K^+)-coupled rat brain glutamate transporter. *Proc. Natl. Acad. Sci. USA* 95:751–755.

133. Tao, Z., and C. Grewer. 2007. Cooperation of the conserved aspartate 439 and bound amino acid substrate is important for high-affinity Na^+ binding to the glutamate transporter EAAC1. *J. Gen. Physiol.* 129:331–344.

134. Borre, L., and B. I. Kanner BI. 2001. Coupled, but not uncoupled, fluxes in a neuronal glutamate transporter can be activated by lithium ions. *J. Biol. Chem.* 276:40396–40401.

135. Tao, Z., and C. Grewer. 2005. The conserved histidine 295 does not contribute to proton cotransport by the glutamate transporter EAAC1. *Biochemistry* 44:3466–3476.

136. Teichman, S., S. Qu, and B. I. Kanner. 2012, Conserved asparagine residue located in binding pocket controls cation selectivity and substrate interactions in neuronal glutamate transporters. *J. Biol. Chem.* 287:17198–17205.

137. Grewer, C., N. Watzke, T. Rauen, and A. Bicho. 2003. Is the glutamate residue glu-373 the proton acceptor of the excitatory amino acid carrier 1? *J. Biol. Chem.* 278:2585–2592.

138. Rosental, N., A. Gameiro, C. Grewer, and B. I. Kanner. 2011. A conserved aspartate residue located at the end of the binding pocket controls cation interactions in brain glutamate transporters. *J. Biol. Chem.* 286:41381–41390.

139. Borre, L., and B. I. Kanner. 2004. Arginine 445 controls the coupling between glutamate and cations in the neuronal transporter EAAC-1. *J. Biol. Chem.* 279:2513–2519.

140. Teichman, S., S. Qu, and B. I. Kanner. 2009. The equivalent of a thallium binding residue from an archaeal homolog controls cation interactions in brain glutamate transporters. *Proc. Natl. Acad. Sci. USA* 106:14297–14302.

141. Mwaura, J., Z. Tao, H. James, T. Albers, A. Schwartz *et al.*. 2012. Protonation state of a conserved acidic amino acid involved in Na^+ binding to the glutamate transporter EAAC1. *ACS Chem. Neurosci.* 3:1073–1083.

142. Tao, Z., A. Gameiro, and C. Grewer. 2008. Thallium ions can replace both sodium and potassium ions in the glutamate transporter excitatory amino acid carrier 1. *Biochemistry* 47:12923–12930.

143. Zomot, E., and I. Bahar. 2013. Intracellular gating in an inward-facing state of aspartate transporter Glt(Ph) is regulated by the movements of the helical hairpin HP2. *J. Biol. Chem.* 288:8231–8237.

144. Heinzelmann, G., T. Bastug, and S. Kuyucak. 2013. Mechanism and energetics of ligand release in the aspartate transporter Glt_{Ph}. *J. Phys. Chem. B* 117:5486–5496.

145. Watzke, N., T. Rauen, E. Bamberg, and C. Grewer. 2000. On the mechanism of proton transport by the neuronal excitatory amino acid carrier 1. *J. Gen. Physiol.* 116:609–621.

146. Koch, H. P., J. M. Hubbard, and H. P. Larsson. 2007. Voltage-independent sodium-binding events reported by the 4B-4C loop in the human glutamate transporter excitatory amino acid transporter 3. *J. Biol. Chem.* 282:24547–24553.

147. Frillingos, S., M. L. Ujwal, J. Z. Sun, H. R. Kaback. 1997. The role of helix VIII in the lactose permease of Escherichia coli. 1. Cys-scanning mutagenesis. *Protein Sci.* 6:431–437.

148. Koch, H. P., H. P. Larsson. 2005. Small-scale molecular motions accomplish glutamate uptake in human glutamate transporters. *J. Neurosci.* 25:1730–1736.

149. Lev, B, and S. Y. Noskov. 2013. Role of protein matrix rigidity and local polarization effects in the monovalent cation selectivity of crystallographic sites in the Na-coupled aspartate transporter Glt_{Ph}. *Phys. Chem. Chem. Phys.* 15:2397–2404.

150. Wang, H., A. M. Rascoe, D. C. Holley, E. Gouaux, and M. P. Kavanaugh. 2013. Novel dicarboxylate selectivity in an insect glutamate transporter homolog. *PLoS ONE* 8:e70947.

151. Jensen, S., A. Guskov, S. Rempel, I. Hanelt, and D. J. Slotboom DJ. 2013. Crystal structure of a substrate-free aspartate transporter. *Nat. Struct. Mol. Biol.* 20:1224–1226.

152. Grewer, C., N. Watzke, M. Wiessner, and T. Rauen T. 2000. Glutamate translocation of the neuronal glutamate transporter EAAC1 occurs within milliseconds. *Proc. Natl. Acad. Sci. USA* 97:9706–9711.

153. Akyuz, N., R. B. Altman, S. C. Blanchard, and O. Boudker. 2013. Transport dynamics in a glutamate transporter homologue. *Nature* 502:114–118.

154. Erkens, G. B., I. Hänelt, J. M. H. Goudsmits, D. J. Slotboom, and A. M. van Oijen. 2013. Unsynchronised subunit motion in single trimeric sodium-coupled aspartate transporters. *Nature* 502:119–123.

155. Klauda, J. B., R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, *et al.* 2012. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* 114:7830–7843.

156. Reyes, N., S. Oh, and O. Boudker. 2013. Binding thermodynamics of a glutamate transporter homolog. *Nat. Struct. Mol. Biol.* 20: 634–641.

157. Heinzelmann, G., P. C. Chen, and S. Kuyucak. 2014. Computation of standard binding free energies of polar and charged ligands to the glutamate receptor GluA2. *J. Phys. Chem. B* 118:1813–1824.

158. Limongelli, V., M. Bonomi, and M. Parrinello. 2013. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. USA* 110:6358–6363.

159. Roux, B., O. S. Andersen, and T. A. Allen. 2008. Comment on Free energy simulations of single and double ion occupancy in gramicidin A *J. Chem. Phys* 128:227101.

160. Mim, C., Z. Tao, and C. Grewer. 2007. Two conformational changes are associated with glutamate translocation by the glutamate transporter EAAC1. *Biochemistry* 46:9007–9018.

161. Heinzelmann, G., and S. Kuyucak. 2014. Molecular dynamics simulations of the mammalian glutamate transporter EAAT3. *PLoS One.* In press.

162. Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1996. The determinants of pKa's in proteins. *Biochemistry.* 35:7819–7833.

163. Luzhkov, V. B., and J. Aqvist. 2000. A computational study of ion binding and protonation states in the KcsA potassiun channel. *Biochim. Biophys. Acta.* 1481:360–370.

164. Berneche, S., and B. Roux. 2002. The ionization state and the conformation of Glu-71 in the KcsA K^+ channel. *Biophys. J.* 82:772–780.

165. Nina, M., D. Beglov, and B. Roux. 1997. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem.* 101:5239–5248.

166. Kollman, P. A. 1993. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* 93:2395–2417.

167. Nozaki, Y., and C. Tanford. Intrinsic Dissociation Constants of Aspartyl and Glutamyl Carboxyl Groups. 1967. *J. Biol. Chem.* 242:4731–4735.

168. Bashford, D., and M. Karplus. 1990. pKas of ionizable groups in proteins: atomic detail from a continuum, electrostatic model. *Biochemistry.* 29:10219–10225.

169. Vindigni, A. 1999. Energetic Dissection of Specificity in Serine Proteases. *Comb. Chem. High Throughput Screening* 2:139–153.

170. Cheng, A. C., V. Calabro, and A. D. Frankel. 2001. Design of RNA-Binding Proteins and Ligands. *Curr. Opin. Struct. Biol.* 11:478–484.

171. Mobley, D. L., and P. V. Klimovich. 2012. Perspective: Alchemical Free Energy Calculations for Drug Discovery. *J. Chem. Phys.* 137:230901.

172. Wang, J., W. Wang, P. A. Kollman, and D. A. Case. 2006. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* 25:247–260.

173. Vanommeslaeghe, K., and A. D. MacKerell. 2012. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* 52:3144–3154

174. Vanommeslaeghe, K., E. P. Raman, and A. D. MacKerell. 2012. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* 52:3155–3168.

175. Gohlke, H., and K. Klebe. 2002. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem. Int. Ed.* 41:2644–2676.

176. Warren, G. L., C. W. Andrews, A. M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. Semis, S. Senger, *et al.*. 2006. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* 49:5912–5931.

177. Gilson, M. K., and H. X. Zhou. 2007. Calculation of Protein-Ligand Binding Energies. *Annu. Rev. Biophys. Biomol. Struct.* 36:21–42.

178. Deng, Y., and B. Roux. 2009. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B* 113:2234–2246.

179. Christ, C. D., A. E. Mark, and W. F. van Gunsteren. 2010. Basic Ingredients of Free Energy Calculations. *J. Comput. Chem.* 31:1569–1582.

180. Chipot, C., and A. Pohorille. 2007. *Free Energy Calculations. Theory and Applications in Chemistry and Biology*; Springer Verlag, Berlin.

181. Lelivre, T., M. Rousset, and G. Stoltz. 2010. *Free Energy Computations: A Mathematical Perspective*; Imperial College Press, London.

182. Chipot, C. 2013. Frontiers in Free-Energy Calculations of Biological Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4:71–89.

183. Shirts, M. R., J. Pitera, W. C. Swope, and V. S. Pande. 2003. Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins. *J. Chem. Phys.* 119:5740–5761.

184. Deng, Y., and B. Roux. 2004. Hydration of Amino Acid Side Chains: Nonpolar and Electrostatic Contributions Calculated from Staged Molecular Dynamics Free Energy Simulations with Explicit Water Molecules. *J. Phys. Chem. B* 108:16567–16576.

185. Mann, G., and J. Hermans. 2000. Modeling Protein-Small Molecule Interactions: Structure and Thermodynamics of Noble Gases Binding in a Cavity in Mutant Phage T4 Lysozyme L99A. *J. Mol. Biol.* 302:979–989.

186. Fujitani, H., Y. Tanida, M. Ito, G. Jayachandran, C. D. Snow, M. R. Shirts, E. J. Sorin, and V. S. Pande. 2005. Direct Calculation of the Binding Free Energies of FKBP Ligands. *J. Chem. Phys.* 123:084108.

187. Oostenbrink, C., and W. F. van Gunsteren. 2005. Free Energies of Ligand Binding for Structurally Diverse Compounds. *Proc. Natl. Acad. Sci. USA* 102:6750–6754.

188. Mobley, D. L., J. D. Chodera, and K. A. Dill. 2006. On the Use of Orientational Restraints and Symmetry Corrections in Alchemical Free Energy Calculations. *J. Chem. Phys.* 125:084902.

189. Wang, J., Y. Deng, and B. Roux. 2006. Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophys. J.* 91:2798–2814.

190. Mobley, D., E. Dumont, J. Chodera, and K. Dill. 2007. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* 111:2242–2254.

191. Mobley, D. L., A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet, and K. A. Dill. 2007. Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *J. Mol. Biol.* 371:1118–1134.

192. Hermans, J., and L. Wang. 1997. Inclusion of Loss of Translational and Rotational Freedom in Theoretical Estimates of Free Energies of Binding. Application

to a Complex of Benzene and Mutant T4 Lysozyme. *J. Am. Chem. Soc.* 119:2707–2714.

193. Chang, C. A., W. Chen, and M. K. Gilson. 2007. Ligand Configurational Entropy and Protein Binding. *Proc. Natl. Acad. Sci. USA* 104:1534–1539.

194. Gumbart, J. C., B. Roux, and C. Chipot. 2013. Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *J. Chem. Theory Comput.* 9:794–802.

195. Bastug, T., P. C. Chen, S. M. Patra, and S. Kuyucak. 2008. Potential of Mean Force Calculations of Ligand Binding to Ion Channels from Jarzynski's Equality and Umbrella Sampling. *J. Chem. Phys.* 128:155104.

196. Chen, P. C., and S. Kuyucak. 2009. Mechanism and Energetics of Charybdotoxin Unbinding from a Potassium Channel from Molecular Dynamics Simulations. *Biophys. J.* 96:2577–2588.

197. Doudou, S., N. A. Burton, and R. H. Henchman. 2009. Standard Free Energy of Binding form a One-Dimensional Potential of Mean Force. *J. Chem. Theory Comput.* 5:909–918.

198. Buch, I., S. K. Sadiq, and G. De Fabritiis. 2011. Optimized Potential of Mean Force Calculations for Standard Binding Free Energies. *J. Chem. Theory Comput.* 7:1765–1772.

199. Chen, R., and S. H. Chung. 2012. Structural Basis of the Selective Block of Kv1.2 by Maurotoxin from Computer Simulations. *PLoS One* 7:e47253.

200. Gumbart, J. C., B. Roux, and C. Chipot. 2013. Efficient Determination of Protein-Protein Standard Binding Free Energies from First Principles. *J. Chem. Theory Comput.* 9:3789–3798.

201. Dixit, S. B, and C. Chipot. 2001. Can Absolute Free Energies of Association Be Estimated from Molecular Mechanical Simulations? The Biotin–Streptavidin System Revisited. *J. Phys. Chem. A* 105:9795–9799.

202. Rocklin, G. J., S. E. Boyce, M. Fischer, I. Fish, D. L. Mobley, B. K. Shoichet, and K. A. Dill. 2013. Blind Prediction of Charged Ligand Binding Affinities in a Model Binding Site. *J. Mol. Biol.* 425:4569–4583.

203. Rashid, M. H., G. Heinzelmann, R. Huq, R. B. Tajhya, S. C. Chang, S. Chhabra, M. W. Pennington, C. Beeton, R. S. Norton, and S. Kuyucak. 2013. A Potent and Selective Peptide Blocker of the Kv1.3 Channel: Prediction from Free-Energy Simulations and Experimental Confirmation. *PLoS One* 8:e78712.

204. Brauner-Osborne, H., J. Egebjerg, E. O. Nielsen, U. Madsen, and P. Krogsgaard-Larsen. 2000. Ligands for Glutamate Receptors: Design and Therapeutic Prospects. *J. Med. Chem.* 43:2609–2645.

205. O'Neill, M. J., and J. M. Witkin. 2007. AMPA Receptor Potentiators: Application for Depression and Parkinson's Disease. *Curr. Drug Targets* 8:603–620.

206. Ahmed, A. H., M. D. Thompson, M. K. Fenwick, B. Romero, A. P. Loh, D. E. Jane, H. Sondermann, and R. E. Oswald. 2009. Mechanisms of Antagonism of the GluR2 AMPA Receptor: Structure and Dynamics of the Complex of two Willardiine Antagonists with the Glutamate Binding Domain. *Biochemistry* 48:3894–3903.

207. Armstrong, N., and E. Gouaux. 2000. Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core. *Neuron* 28:165-181.

208. Hogner, A., J. S. Kastrup, R. Jin, T. Liljefors, M. L. Mayer, J. Egebjerg, I. K. Larsen, and E. Gouaux. 2002. Structural Basis for AMPA Receptor Activation and Ligand Selectivity: Crystal Structures of Five Agonist Complexes with the GluR2 Ligand-Binding Core. *J. Mol. Biol.* 322:93–109.

209. Menuz, K., R. M. Stroud, R. A. Nicoll, and F. A. Hays. 2007. TARP Auxiliary Subunits Switch AMPA Receptor Antagonists into Partial Agonists. *Science* 318:815–817.

210. Jo, S., W. Jiang, H. S. Lee, B. Roux, and W. Im. 2012. CHARMM-GUI Ligand Binder for Absolute Binding Free Energy Calculations and Its Application. *J. Chem. Inf. Model.* 53:267–277.

211. Kofke, D. A., and P. T. Cummings. 1998. Precision and Accuracy of Staged Free-Energy Perturbation Methods for Computing the Chemical Potential by Molecular Simulation. *Fluid Phase Equilib.* 150:41–49.

212. Bennett, C. H. 1976. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* 22:245–268.

213. Pohorille, A., C. Jarzynski, and C. Chipot. 2010. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* 114:10235–10253.

214. Mobley, D. L., J. D. Chodera, and K. A. Dill. 2007. Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. *J. Chem. Theory Comput.* 3:1231–1235.

215. Wulff, H., P. A. Calabresi, R. Allie, S. Yun, M. Pennington *et al.* 2003. The voltage-gated Kv1.3 K+ channel in effector memory T cells as new target for MS. *J. Clin. Invest.* 111:1703–1713.

216. Castaneda, O., V. Sotolongo, A. M. Amor, R. Stocklin R, A. J. Anderson, *et al.*. 1995. Characterization of a potassium channel toxin from the Caribbean Sea anemone Stichodactyla helianthus. *Toxicon* 33:603–613.

217. Kalman, K., M. W. Pennington, M. D. Lanigan, A. Nguyen, H. Rauer, *et al.* 1998. ShK-Dap22, a potent Kv1.3-specific immunosuppressive polypeptide. *J. Biol. Chem.* 273: 32697–32707.

218. Chi, V., M. W. Pennington, R. S. Norton, E. J. Tarcha, L. M. Londono, *et al.* 2012. Development of a sea anemone toxin as an immunomodulator for therapy of autoimmune diseases. *Toxicon* 59:529–546.

219. Beeton, C, H. Wulff, S. Singh, S. Botsko, G. Crossley, *et al.* 2003. A novel fluorescent toxin to detect and investigate Kv1.3 channel up-regulation in chronically activated T lymphocytes. *J. Biol. Chem.* 278:9928–9937.

220. Beeton, C., M. W. Pennington, H. Wulff, S. Singh, D. Nugent, *et al.* 2005. Targeting effector memory T cells with a selective peptide inhibitor of Kv1.3 channels for therapy of autoimmune diseases. *Molec. Pharmacol.* 67:1369–1381.

221. Pennington, M. W., C. Beeton, C. A. Galea, B. J. Smith, V. Chi, *et al.* 2009. Engineering a stable and selective peptide blocker of the Kv1.3 channel in T lymphocytes. *Molec. Pharmacol.* 75:762–773.

222. Pennington, M. W., M. H. Rashid, R. B. Tajhya, C. Beeton, S. Kuyucak, and R. S. Norton. 2012. A C-terminally amidated analogue of ShK is a potent and selective blocker of the voltage-gated potassium channel Kv1.3. *FEBS Letters* 586:3996–4001.

223. Pennington, M. W., V. M. Mahnir, I. Khaytin, I. Zaydenberg, M. E. Byrnes, and W. R. Kem. 1996. An essential binding surface for ShK toxin interaction with rat brain potassium channels. *Biochemistry* 35:16407–16411.

224. Rauer, H., M. W. Pennington, M. Cahalan, and K. G. Chandy. 1999. Structural conservation of the pores of calcium-activated and voltage-gated potassium channels determined by a sea anemone toxin. *J. Biol. Chem.* 274:21885–21892.

225. Grissmer, S., A. N. Nguyen, J. Aiyar, D. C. Hanson, R. J. Mather, *et al.* 1994. Pharmacological characterization of five cloned voltage-gated K+ channels, types Kv1.1, 1.2, 1.3, 1.5, and 3.1, stably expressed in mammalian cell lines. *Mol. Pharmacol.* 45:1227–1234.

226. Flugel, A., M. Willem, T. Berkowicz, and H. Wekerle. 1999. Gene transfer into CD4+ T lymphocytes: green fluorescent protein-engineered, encephalitogenic T cells illuminate brain autoimmune responses. *Nat. Medicine.* 5:843–847.

227. Beeton, C., and K. G. Chandy. 2007. Induction and monitoring of adoptive delayed-type hypersensitivity in rats. *J. Vis. Exp.* 8:325.

228. Matheu, M. P., C. Beeton, A. Garcia, V. Chi, S. Rangaraju, *et al.* 2008. Imaging of effector memory T cells during a delayed-type hypersensitivity reaction and suppression by Kv1.3 channel block. *Immunity* 29:602–614.

229. Beeton, C., H. Wulff, J. Barbaria, O. Clot-Faybesse, M. W. Pennington, *et al.* 2001. Selective blockade of T lymphocyte K+ channels ameliorates experimental autoimmune encephalomyelitis, a model for multiple sclerosis. *Proc. Natl. Acad. Sci. USA* 98:13942–13947

230. Chang, S. C., C. A. Galea, E. W. Leung, R. B. Tajhya, C. Beeton, *et al.* 2012. Expression and isotopic labelling of the potassium channel blocker ShK toxin as a thioredoxin fusion protein in bacteria. *Toxicon* 60:840–850.

231. Beeton, C., and K. G. Chandy. 2007. Induction and monitoring of active delayed type hypersensitivity (DTH) in rats. *J. Vis. Exp.* 6:237.

232. Tudor, J. E., P. K. Pallaghy, M. W. Pennington, and R. S. Norto. 1996. Solution structure of ShK toxin, a novel potassium channel inhibitor from a sea anemone. *Nat. Struct. Biol.* 3:317–320.

233. Tudor, J. E., M. W. Pennington, and R. S. Norton. 1998. Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin. *Eur. J. Biochem.* 251:133–141.

Author's Biography

Germano Heinzelmann was born in Joinville, Brazil, in June 28th, 1982. He obtained his Bachelor's degree and Master's degree in Physics from Universidade Federal de Santa catarina (UFSC), located in Florianópolis, Brazil. He was then offered an international Ph.D. scholarship at the University of Sydney, starting his studies in 2010. Since May, 2010, he has been doing research under his supervisor Prof. Serdar Kuyucak.