Role of extracellular gate in cation coupling in the glutamate transporter family

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Claudia Alleva

Jülich, August 2020

aus dem Institute of Complex Systems – Zelluläre Biophysik
des Forschungszentrums Jülich

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Christoph Fahlke
Korreferent: Prof. Dr. Holger Gohlke
Tag der mündlichen Prüfung:

Contents

List of Figures

List of Tables

1. Abstract

Excitatory amino acid transporters (EAATs) terminate synaptic transmission by taking up glutamate from the synaptic cleft and moving it back into nearby neurons and glial cells. Thereby, EAATs ensure low extracellular glutamate concentrations in the brain and permit reliable high frequency signal transduction in excitatory synapses. Under pathological conditions, e.g. ischemia, impaired EAAT function leads to accumulation of glutamate in the synaptic cleft, causing glutamate excitotoxic effects.

Concentrative glutamate uptake is achieved through a secondary active transport cycle, which couples the uptake of one glutamate molecule to the co-transport of three Na^+, and one H^+, with the counter-transport of one K^+. This coupling stoichiometry permits glutamate accumulation up to 10^6-fold gradients. X-ray crystal structures of the prokaryotic EAAT homologs Glt_{Ph}, Glt_{Tk}, and the human glutamate transporter EAAT1 have revealed the position of the substrate- and the Na^+-binding sites, and defined two hairpin-like structures that serve as gates to regulate access to the substrate-binding pocket.

Glutamate transport involves at least two major conformational changes: extracellular gate opening and closing, and transmembrane translocation of the transport domain harboring the substrate and ion-binding sites. It has recently been established that substrate binding causes subsequent extracellular gate closure, and this induced-fit binding mechanism has been shown to confer substrate selectivity. In contrast, the coupling between cation binding and conformational changes is still unclear.

Here, we perform extensive all-atom molecular dynamics simulations to study the coupling mechanisms between Na^+, K^+, and H^+ binding and extracellular gate dynamics. Our molecular dynamics simulations suggest that the binding of two Na^+ prior to substrate binding induces gate opening via a conformational selection-like mechanism, and we define the allosteric interaction network underlying this Na^+–gate coupling. We confirmed these results through stopped flow fluorescence spectroscopy experiments, and we identified two point mutations, M311A and R397A, in Glt_{Ph} that uncouple Na^+ binding from gate dynamics by selectively altering Na^+ binding or the Na^+-induced gate opening. The coupling mechanism involves Na^+ binding to the so-called Na1 and Na3 sites: Na^+ bind-

ing at Na1 is preceded by a binding pocket rearrangement, while Na3 binding involves passage through the Na1 site. We show that only occupation of both sites ensures complete gate opening to permit subsequent substrate binding.

Moreover, we show that K^+ binding is also coupled to extracellular gate dynamics, but exerting an opposite effect to Na^+ binding and causing hairpin closure in EAAT1. Similarly to Na^+, the coupling to K^+ is driven by allosteric interactions, which we demonstrate by disrupting it with mutations that destabilize the open state (E386Q, E386D, and R459A EAAT1) or increase the flexibility of the hairpin (L428A EAAT1). Finally, we show the physiological role of the putative H^+-acceptor site, demonstrating that protonating it stabilized the closed state of the extracellular gate.

In summary, Na^+ binding causes extracellular gate opening via a conformational selection-like mechanism, which permits induced-fit substrate binding causing gate re-closure, which finally permits transmembrane translocation. Our analysis reveals how the energetic coupling of the extracellular gate to ion and substrate binding forms the molecular basis for secondary active glutamate transport.

2. Introduction

2.1 Physiological role of Excitatory Amino Acid transporters

Glutamate is the predominant neurotransmitter used in excitatory synapses in the mammalian nervous system. In order to achieve the highly sensitive and specific signal transmission typical of excitatory synapses, glutamate needs to be eliminated from the synaptic cleft through a fast and specific uptake. This need is a product of the fundamental role that glutamate plays in excitatory signaling, and the neurotoxicity for which high glutamate concentrations are responsible. Example of this double requirement for keeping low synaptic glutamate concentration are the physiological functions, i.e. learning and memory [1], and the multiple pathologies in which this neurotransmitter is involved, such as schizophrenia, Alzheimer's disease, multiple sclerosis, traumatic brain injury, epilepsy and ataxia [2–7].

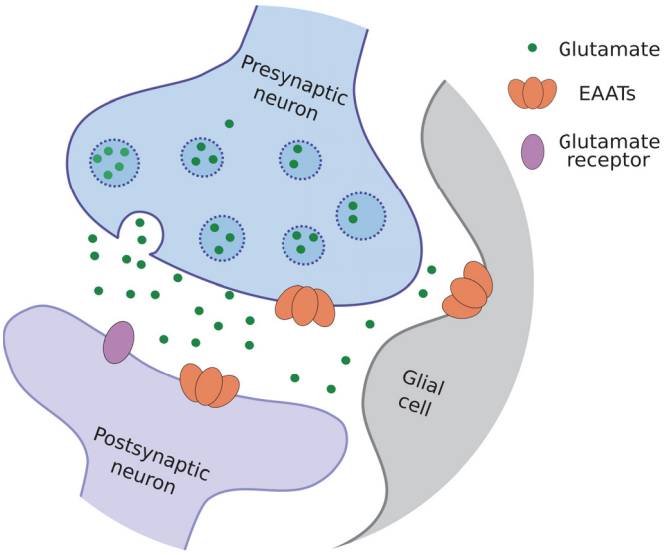

Figure 2.1: Schematic representation of an excitatory synapse, formed by the presynaptic and postsynaptic neurons and a glial cell. The position of glutamate receptors and the excitatory amino acid transporters are shown. The role of EAATs is to take up the released glutamate and thus to keep a glutamate gradient across the cell membrane.

This crucial role of maintaining up to 10^6-fold gradients between intracellular ($\sim$10 mM) and extracellular space (in the nM range), is played out by excitatory amino acid transporters (EAATs). These are secondary active transporters, which exploit the electrochemical potential difference maintained across biological membrane, in order to transport species against their concentration gradient by thermodynamically coupling cotransport and counter transport of different species; EAATs stoichiometry, in particular, includes co-transport of 3 Na^+, one H^+ (which move down their gradient) and counter transport of one K^+ [8–10]. The stoichiometry of glutamate uptake is thus electrogenic, with the net transport of two charges for each complete cycle, in contrast to the electroneutral neutral amino acid uptake of the closely related ASCTs.

An added layer of complexity in EAATs role, is their ability to conduct Cl^- currents, functioning as chloride channels thermodynamically uncoupled to glutamate uptake and yet activated upon glutamate binding [11].

2.2 Glutamate transporter family structures

A plethora of Glutamate transporter family structures have been resolved: the *Pyroccocus horikoshii* glutamate transporter (Glt_{Ph}) was the first resolved structure in the family in 2007 [12]. Since then Glt_{Ph} has been crystallized in different occupation states and conformations, often with the aid of specific mutations, so that most steps of the transport cycle have been elucidated [12–18]. Following high resolution Glt_{Tk} structures [19] have been released and finally by the EAAT1 [20] and ASCT [21] structures. The first EAAT crystal structure [20], in particular, became available only recently, so most of the structural and computational studies carried out have used the archaeal homologs crystal structures, in particular Glt_{Ph}.

Table 2.1 presents a list of the structures currently available.

PDB ID	Resolution (Å)	Conformation	Mutation	Ligands	Citation
Glt_{Ph}					
2NWW	3.2	OF,open	7H	Na^+ + TBOA	
2NWX	3.29	OF, closed	7H	$2Na^+$ + Aspartate	[12]
2NWL	2.96	OF, closed	7H	Aspartate	
3KBC	3.51	IF, closed	7H, V216C, M385C	$2Na^+$ + Aspartate	[13]
3V8F	3.8	IF, closed	7H, V216C, M385C	$2Na^+$ + Aspartate	[14]
3V8G	4.66	OF, closed	7H, V216C, M385C	$2Na^+$ + Aspartate	
4OYE	4	OF, closed	7H, R397A	apo	[15]
4IZM	4.5	OF, closed	7H, L66C, S300C	$2Na^+$ + Aspartate	
4OYF	3.41	OF, closed	7H, R397A	Na^+	
5CFY	3.5	OF, closed	7H, R397A	$2Na^+$ + Aspartate	
4P3J	3.5	OF, closed	7H, R397A, V216C, M385C	apo	[16]
4P19	3.25	IF, closed	7H, R397A, V216C, M385C	apo	
4P1A	3.75	IF, closed	7H, V216C, M385C	Tl^+	
4P6H	4.08	IF, closed	7H	Tl^+	
4X2S	4.21	IF, closed	7H, R276S, M395R, L66C, S300C	$2Na^+$ + Aspartate	[17]
6BAU	3.8	OF, closed	7H, R397C	$2Na^+$ + Cysteine	
6BMI	3.9	OF, closed	7H, R397C	$2Na^+$ + Serine	[18]
6BAT	3.4	OF, closed	WT	$2Na^+$ + Aspartate	
6BAV	3.7	OF, closed	R397C	Na^+ + Benzylcystein	
-	2.6	OF, closed	-	$3Na^+$ + Aspartate	[22]
Glt_{Tk}					
5DWY	2.7	OF, closed	-	apo	[19]
5E9S	2.8	OF, closed	-	$3Na^+$ + Aspartate	
EAAT1					
5LLU	3.32	OF, closed	mutant II	Na^+ + Glutamate	
5MJU	3.71	OF, open	mutant II	TFB-TBOA + UCPH101	[20]
5LM4	3.1	OF, open	mutant II	Na^+ + Aspartate + UCPH101	
5LLM	3.25	OF, open	mutant II	Na^+ + Aspartate + UCPH101	
ASCT					
6GCT	3.85	OF, closed	no	Glutamine	[21]

Table 2.1: The available glutamate transporter family structures in this table are divided by protein homologs; OF stands for outward-facing, IF for inward facing; open and closed refer to the state of the hairpin; the structures used for this study are shown in boldface.

In light of the sequence identity of ∼30% between different structures (i.e. 37% between Glt_{Ph} and EAAT1), the similarity in structure between homologs is staggering. Not only the general division in domains and the respective positions of monomers is perfectly overlapping, but even the majority of the residues in the functional sites are oriented in a comparable manner. Due to this high similarity, we will present Glt_{Ph} structure as the archetype of the family, commenting on important differences when necessary.

Glt_{Ph} has a trimeric structure, composed of identical monomers (Fig. 2.2), arranged in order to form a large hydrophilic basin (diameter of ∼50 Å) on the extracellular side, which encompass about half of the membrane plane (with a total depth of 30 Å and ∼15 Å of depth across the membrane) [12].

Each monomer is composed of eight transmembrane helical domains and two helix-turn-helix hairpins (shown in red in Fig. 2.2 and named HP1 and HP2). The two helical hairpins form the structural basis for the extracellular gate. The three monomers interact through their trimerization domains (shown in blue in Fig. 2.2), which form the base of the bowl-like soluble basin, while the transport domains (shown in yellow in Fig. 2.2) line the sides of the bowl. The trimerization domain is composed of TM1, TM2, TM4 and TM5, while the transport domain contains the TM3, TM6, TM7, TM8 transmembrane helices, and the two hairpins (HP1 between TM6 and TM7, HP2 between TM7 and TM8).

The tips of the HP1 and HP2 directly interact in the structures denominated as closed (i.e. in $Glt_{Ph,2Na,substrate}$ in [12] and $Glt_{Ph,apo}$ in [16]); just behind the two hairpin lays the binding pocket for the negatively charged amino acid (which we will call from now on substrate) and the Na^+, lined in the back by TM7 and TM8, the first of which is interrupted by an unstructured region corresponding to a five-residue highly conserved motif (NMDGT). In the apo crystal structures [16], the NMDGT motif undergoes a rearrangement, with M311 moving away from the substrate binding site and pulling the motif away from TM4, while at the same time TM3 retrieves away from TM7, with N310 moving away from N401, which rearranges in order to occupy part of the Na1 site. HP2 is generally completely wound, irrespectively of its position, but it partially unwinds in $Glt_{Ph,apo}$, with the formation of a longer loop that includes the last residues of the first helix, HP2a. All of these rearrangements in the structure describe the collapsing of Na1, Na2 and Na3 sites, so that $Glt_{Ph,apo}$ has a different binding pocket arrangement than the $Glt_{Ph,2Na,substrate}$ structure, but the same compactness, which would allow for the rigid body movement across the membrane described in [13]. Interestingly, most of these changes can be reverted by the presence of Na^+, as we can observe in $Glt_{Ph,Na1}$ [16]), where the structure has been obtained in presence of Na^+ and Na1 occupation has been resolved. These observations demonstrate how Na^+ occupancy might be enough to bring the binding pocket to the fully-bound state. Similar results are observed for the apo inward structures obtained in [16].

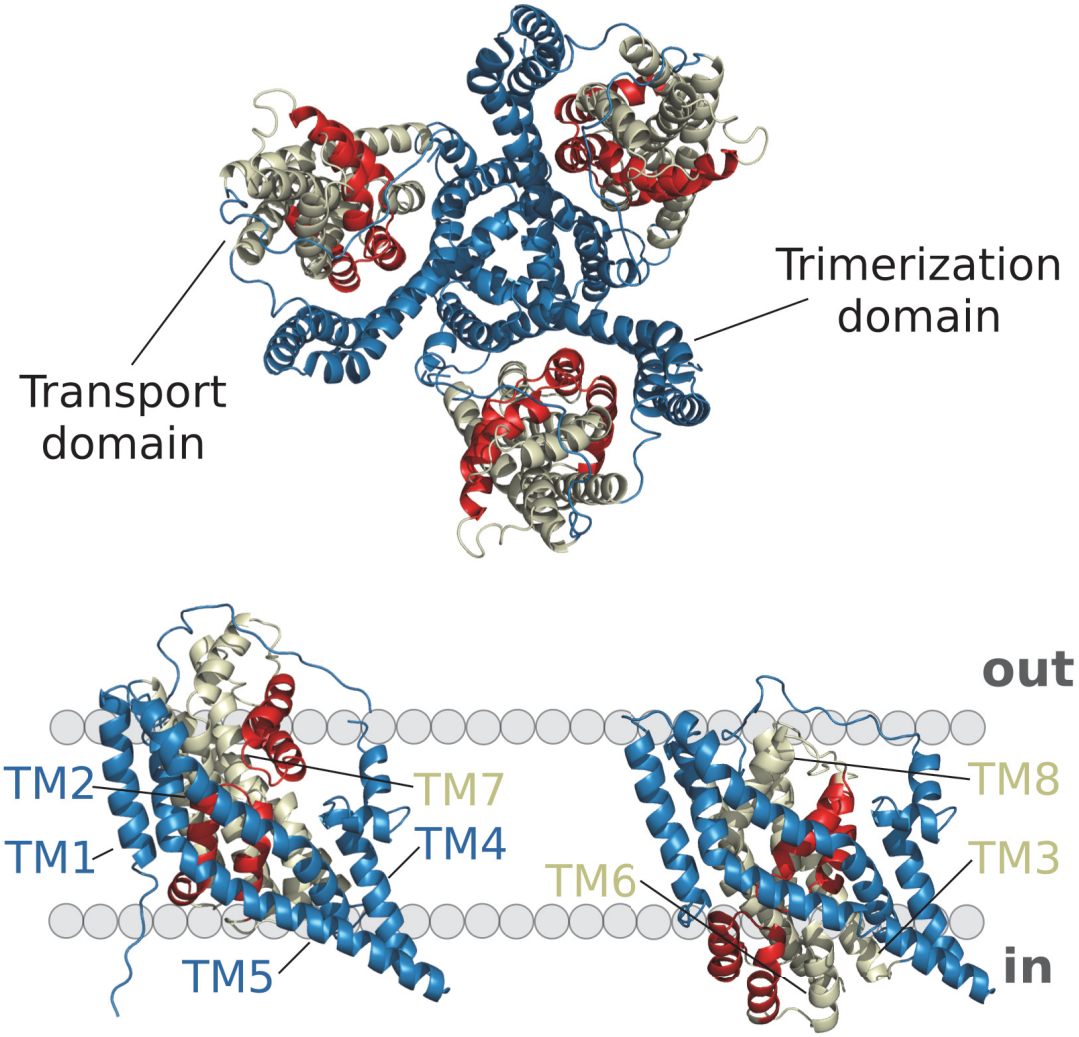

Figure 2.2: Glt_{Ph} trimer is shown on the left as a topview from the extracellular side of the membrane plane; the trimerization domain, transport domain and the hairpins are indicated by arrows, as are the transmembrane domains. On the bottom the monomeric structure is shown in the outward-facing (on the left) and the inward (on the right) conformations, in order to emphasize the movement of the transport domain along the trimerization domain and across the membrane.

The inward structure, obtained in [13], helps describing the alternating access mechanism that characterizes the transport cycle of glutamate transporters. In order to obtain this structure, a double cysteine mutation K55C/A364C was used. While the internal structure of the transport and trimerization domains is

unaffected, the relative position of the two is modified: the two loops between TM2-TM3 and TM5-TM6 work as hinges, while the transport domain slides along the trimerization domain, which anchors the monomer in the membrane (Fig. 2.2). The movement consists of a 16 Å translation across the membrane, towards the intracellular side, and a 37° rotation. More recently, the structure of the Thermococcus kodakarensis glutamate transporter (Glt_{Tk}) has been resolved with different ligand occupations in the outward-facing state, adding important informations on key residues previously mutated (i.e. R397A)and the position of the Na3 site [19, 23]. Despite the coupling to three Na^+, until the Glt_{Tk} crystal structure, the position of the third Na^+ was only speculated upon, with different sites being proposed. It is important to note here that the Glt_{Tk} structures were the first ones to have a resolution high enough to directly position Na^+ ions, in contrast to the two sites previously found in Glt_{Ph}, which where interpreted as Na^+ sites but were resolved with the aid of Tl^+ soaking. Despite the strong similarity to Na^+, Tl^+ could not sustain the uptake of the substrate, thus being unable to effectively substitute Na^+. One could speculate that Tl^+ has chemical characteristics closer to K^+, but the direct resolution of the Na1 and Na2 sites in Glt_{Tk}, confirms the site modeled previously in Glt_{Ph}. These structures also directly show the third Na^+ position, which, interestingly were previously postulated with the aid of molecular dynamics simulations. Another important point in the new crystals, is the resolution of R399 (R397 in Glt_{Ph}) in $Glt_{Ph,apo}$; this residue was known to directly stabilize the negatively charged substrate in the binding pocket and to have an impact on Na^+-substrate coupling. Despite its importance, $Glt_{Ph,apo}$ could not be resolved in earlier crystals without mutating this residue to an alanine, and thus decreasing dramatically the transported affinity for the substrate. The position of this arginine reveals to have a characteristic distribution in the $Glt_{Ph,apo}$ and $Glt_{Ph,2Na,substrate}$ structures, similarly to what previously observed for other residues in [16]. Despite the resolution of the third Na^+ site and the presence of R397, the structural characteristic of the crystal are rather similar to $Glt_{Ph,apo}$, with a partially unwound HP2a, the extracellular gate closed and a rearrangement of residues in the binding pocket that allows for a very compact transport domain. The position of the residues coordinating the resolved Na3 site in Glt_{Tk}, is similar to the homologous positions in Glt_{Ph}, hinting at the occupation of this site in the previously obtained crystals.

Recently the first mammalian EAAT structure was resolved [20]. EAATs have a similar overall structure, with the addition of an extracellular domain which is believed to be site of a post translational N-glycosylation [20]. A thermostable version of the transporter, with 75% of identity with the wild type, was used to obtain the crystal structure. Despite the mutation, the protein still showed substrate uptake and it maintained the stoichiometry and the inhibition profile of the wild type. In this study, we use the residue numbering described for the crystal structure in [20]. Addition of two isoleucine residues (M231I and F235I) was needed in order to crystallize the transporter in the inhibitor free form. Overall, the EAAT1 structure is very similar to the Glt_{Ph} and Glt_{Tk} structures, with a trimeric organization and the clear distinction of a transport domain and a trimerization domain, but with the addition of a bulkier extracellular side. The position of the substrate binding site is conserved and so is the one of Na2, which could be resolved in the crystal structure, while Na1 and Na3 were not resolved, even though they are generally thought to be present due to the coordinating residues position.

In the transport domain, the transmembrane helix TM8 can here be divided in three parts, namely TM8a, TM8b, and TM8c, the last of which protrudes in the cytoplasmic side, while TM8b in particular directly interacts with HP1 and HP2.

2.3 Mechanism of transport in the glutamate transporters family

EAATs are remarkable molecular machines, undergoing a transport cycle consisting of multiple steps, at each of which some degree of control can be applied.

The general model used to describe the mechanism exploited by transporters to move solutes across the membrane is the alternating access model, which calls for a differential access to the binding pocket from the two sides of the membrane, controlled by a conformational change. The transport of solutes across the membrane can be described by the alternating access mechanism [24]; this model calls for a conformational change that allow for differential access from the two sides of the membrane. Transporters can be divided in classes due to the specific molecular mechanism by which the differential access is guaranteed [25];

glutamate transporters are the prototypical example of one of these classes, the one characterized by the elevator-like mechanism. This mechanism is characterized by the relative rigid-body movement of one domain, containing the binding pocket, across the membrane and along a second domain, which plays a scaffolding role and is responsible for the oligomerization. Glt_{Ph} is the prototypical example of the elevator-like alternating access mechanism, with the trimerization domain (Fig. 2.2) of each monomer interacting with the same domains of the other monomers, keeping the oligomer together, while also fixing it in the membrane; the transport domain, containing the substrate and Na^+ binding sites, undergoes a translational and rotational movement along the trimerization domain, responsible for a movement of $\sim$18 Å of the binding pocket (see movement of transport domain across the membrane in Fig. 2.2).

Interestingly, the translocation mechanism is controlled by the state of the extracellular gate, formed by HP1 and HP2; HP2 can move away from HP1, so that the tips of the two hairpins are at a distance of $\sim$12 Å , which characterizes the open state and does not allow free access to the binding pocket. This open state is incompatible with translocation, due to the position of HP2, which would clash with the trimerization domain during translocation. HP2 acquires then the role of switch for the translocation process. Conversely, HP2 state depends directly on the association of Na^+ (which induces opening) and the substrate (which induces closure) [26].

These major structural and stoichiometric points of control are combined in order to obtain a highly regulated process, in which the transporter can be in the inward or outward-facing conformation, the gate might be open or closed and a different set of ligand can bind at any time. At the same time, the different regulation hubs are also interdependent: an open gate is incompatible with re-translocation [16]; a stable intermediate state allows the uncoupled Cl^- permeation while shielding access to any other ligand from both sides of the membrane [11]; Na^+ and glutamate have opposite effects on the state of the gate [26]. Using these observations and the stoichiometry information described in literature [10, 27] we can describe the entire transport cycle.

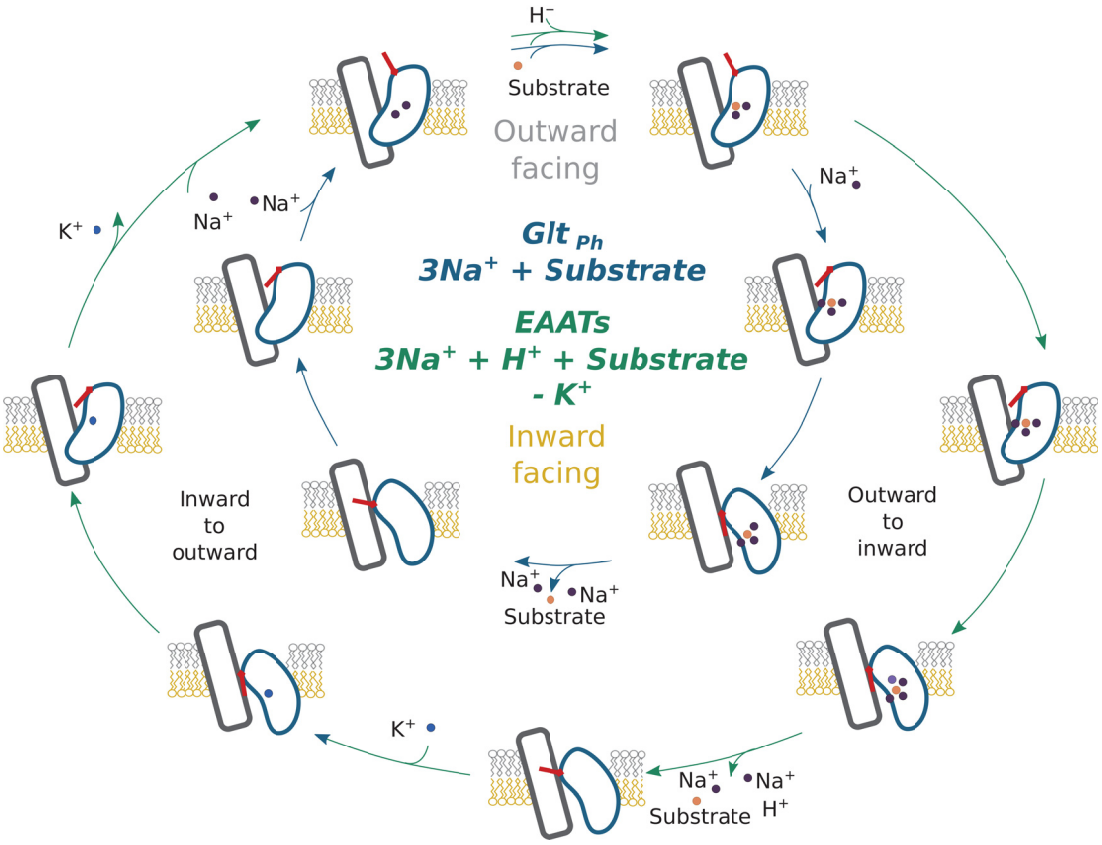

Figure 2.3: The transport cycle for Glt_{Ph} and EAATs are shown in concentric ellipses (blue for Glt_{Ph} and green for EAATs). The two cycles diverge in the presence of a H^+ and a K^+, which are not participating in the transport in Glt_{Ph}.

Starting from an empty outward-facing conformation we can imagine two scenarios depending on the gate state. It is not currently possible to discern between the two states, but the current literature indicates that the extracellular gate is closed in the apo state, at least partially. The binding pocket is at this point inaccessible by the substrate and relatively dry. Opening of the extracellular gate is coupled with binding of Na^+ [26]. The binding pocket is then ready to accommodate substrate binding, and subsequently the third Na^+. At this point, the EAAT1 and Glt_{Ph} transport cycles diverge, with Glt_{Ph} in a closed state and ready for translocation and EAAT1 additionally binding one H^+ before undergoing the elevator-like movement. The fully-bound transporters are now presenting the binding pockets on the intracellular side, and the intracellular gate opens, in order to allow for the release of the ligands. The opening of the

intracellular gate has been proposed to involve movement of HP1 away from HP2 but has not been yet thoroughly investigated. The two cycles diverge again at this point, with EAAT1 needing one K^+ to bind before re-translocation and Glt_{Ph} re-translocating empty. The K^+-bound transporter then faces the extracellular environment where the K^+ is released. At this point the two cycles converge again to the first step, with an empty transporter facing the extracellular side. Thus, the general organization of the transport cycle in EAAT1 and Glt_{Ph} is rather similar, with some fundamental points of diversion that characterize the stoichiometry of the two homologs.

2.4 Substrate binding pocket and the NMDGT motif

The substrate binding pocket lies directly beneath HP1 and HP2 and is made accessible by a rearrangement of HP2 that brings the transporter to the open state. This rearrangement is necessary to the priming of the transporter for substrate binding, as the distance between the tips of HP1 and HP2 (2–5 Å) is non permissive for substrate access to the binding pocket. In the binding pocket, the sequence similarity between Glt_{Ph} and the EAATs jumps from ∼37 to ∼60–80% (depending on the definition of binding pocket), so that only a few key residues differ between the homologs. The binding pocket is enclosed by TM7 and TM8 (which line the back of the pocket) and the two hairpins. A highly conserved motif, NMDGT, which characterizes the glutamate transporters family, lies in an unwound region which divides in two TM7 and provides residues important for the coordination of both the substrate and Na^+. On HP1, arginine 276 is substituted by a serine or alanine in EAATs; the backbone of these residue interacts with the bound aspartate and is part of one Na^+ site, namely site CT, and forms a salt bridge with residue D394 in the Glt_{Ph} inward-facing state; the double mutation R276S/M395R (in order to mimic the arginine in TM8 found in EAATs) increases the activity of the transporter [17], even though the substrate is coordinated by the backbone atoms of this residue and is thus unaffected. Asparagine 310 partially occupies the Na1 site in the apo state, while it coordinates Na^+ at the Na1 site in all Na^+-bound simulations. Methionine 311 sits between Na1 and Na2 in the bound state and points towards HP2. Interestingly, when only Na^+ is

bound, the tip of M311 assumes a position that would clash with the open state of the hairpin. Aspartate 312 is involved in coupling of both Na1 and NA3, while ensuring the bound-like state of the NMDGT motif.

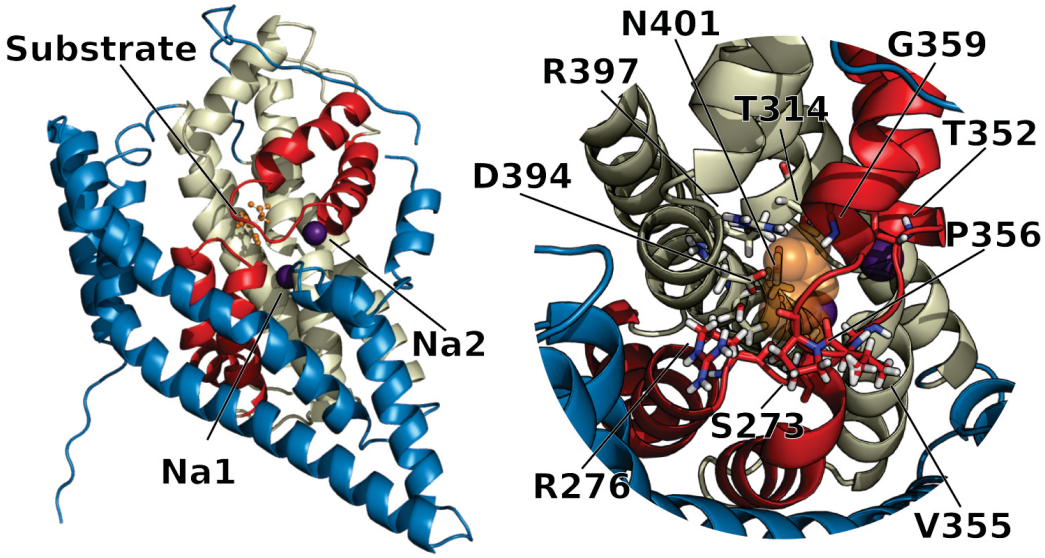

Figure 2.4: On the left is shown a Glt_{Ph} monomer ($Glt_{Ph,2Na,substrate}$), the substrate and Na^+ sites are highlighted. On the right, the binding pocket is magnified; residues interacting with the amino acid (here in transparent orange spheres) are shown in sticks.

Threonine 314 undergoes a rearrangement between Na^+-bound state and apo state, similarly to other residues of the NMDG motif, while changing partners and stabilizing the bound position of R397 and N401. Threonine 314 is substituted by an alanine in ASCTs; when mutated to threonine to mimic the glutamate/aspartate transporters, the neutral amino acid exchange is reduced and the affinity to aspartate increased [28], suggesting an important role of this residue in substrate recognition. On HP2, Threonine 352 contributes to the coordination of the Na2 site, while also controlling the affinity to the substrate; mutation of this residue to alanine increases ∼15 times the affinity to glutamate, while the double mutation T352A/M362T increases glutamate affinity up to 130-fold [29]. Interestingly, the single mutant M362T increases glutamate affinity but does not elicit glutamate uptake. Two glycine residues, G354 and G357 confer high flexibility to the HP2b helix, allowing the hairpin to adapt to the bound substrate while

closing (allowing V355 and A358 to interact with the aspartate alpha and beta carboxylate group, respectively). G354 might also have a role in substrate recognition in the early steps of binding. G354 stabilizes the closed state by forming a bond with serine 279; G357 has a similar effect, forming a hydrogen bond with D394. Interestingly, this aspartate is not indispensable to the correct function of the transporter, while mutating D390, positioned just one helix turn higher, Na^+ binding and transport dynamics are modified. In D390E Na^+ binding is impaired [30], while D390N cannot bind Na^+ following the substrate binding.

Mutations of this residues have been showed to lock the transporter in a obligated exchanger mode [31]. Finally, mutating R397 [16, 32–34] in Glt_{Ph} and EAAT3, abolishes the uptake of acidic amino acid and, in EAAT3, introduces a serine exchange dependent on Na^+ binding. Interestingly, the mutations T459R (corresponding to R397) and A382T (T314) in ASCT1 have additive effects on acidic amino acid affinity, indicating that there is cooperativity between these two residues in the substrate recognition mechanism. Moreover, R397 also interacts with Y317, a highly conserved residue on TM7 [35], through a π-cation interactions and its neighbour Q318, which is a glutamate in EAATs and has a crucial role in K^+ coupling [36].

2.5 Na^+ binding pockets

As previously described [16, 19], comparing the $Glt_{Ph,apo}$ and $Glt_{Ph,2Na,substrate}$ structures, as well as the apo and the Na^+ bound crystals, in both Glt_{Ph} and Glt_{Tk}, a stark rearrangement of a series of highly conserved residues is coupled to Na^+ binding. Residues from HP1, HP2, TM7 and TM8 are participating in Na^+ coordination.

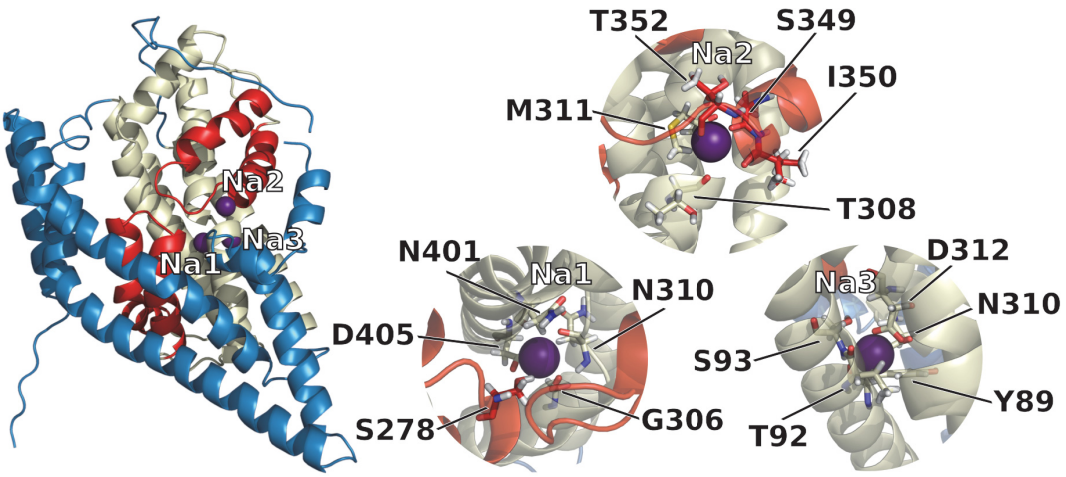

Figure 2.5: The positions of the three Na^+-binding sites are shown on the left; on the right, a magnification of each site shows in sticks the residues involved in coordination; Na3 position was obtained by superimposition with the Glt_{Tk} structure, where the third Na^+ was resolved.

Of the 3 Na^+ coupled to the substrate uptake, initially only two were resolved in crystal structures, with the aid of Tl^+ soaking [12, 37]. The use of Tl^+ has brought some critics to speculate that the sites resolved might not be the ones occupied during transport, especially since substrate uptake cannot be sustained in presence of Tl^+ alone [38]. It has been demonstrated that Tl^+ substitutes more efficiently K^+ ions in the mammalian isoforms, arguing for the possibility that one or more of these sites might be K^+ sites. Na^+ at site 1 is coordinated by S278 on HP1, G306 and N310 on TM7, and N401 and D405 on TM8.

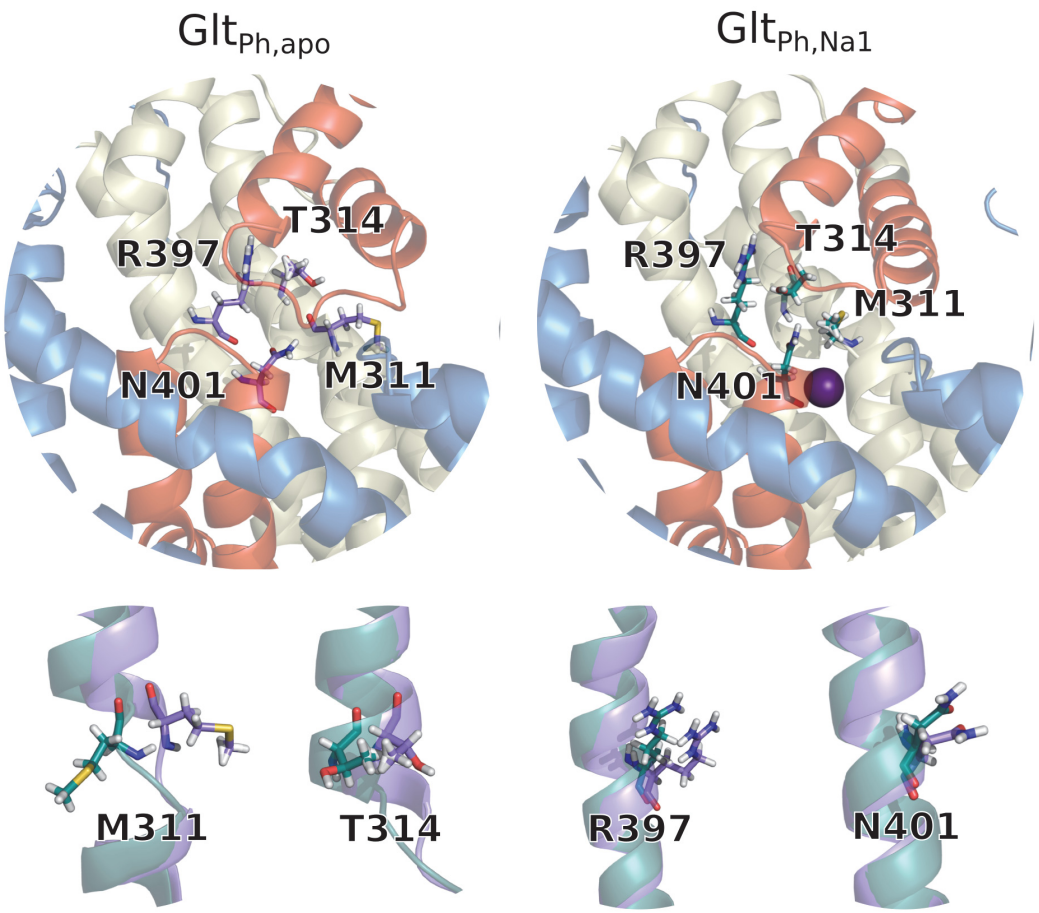

Figure 2.6: The Na1 binding pocket before (on the left) and after (on the right) rearrangement are shown in the circle cutouts of the monomers, the residues involved in the rearrangement are shown as sticks; on the bottom the same residues are shown superimposed before and after rearrangement, in order to emphasize their movement.

Interestingly, the mutation D405N decreases strongly the substrate affinity, demonstrating the coupling between Na1 and substrate binding [30]. The coordination at Na2, instead, is obtained through residues of HP2 (S349,I355,T352) and TM7 (T308). Due to its nature, the Na2 site is disrupted when the hairpin opens. This observation, together with the fact that one Na^+ binds after binding of the substrate, suggests that Na2 might lock HP2 in a closed position once fully-bound. Very similar sites were resolved at higher resolution in Glt_{Tk} [19], together with a third Na^+ site, previously unidentified. These results suggest that Tl^+ cannot substitute Na^+ at Na3 site and possibly sustain a productive

transport cycle. Four different sites were proposed to be the then-unknown third Na^+ binding site: a site involving T314, D312 and the substrate was proposed in [39]; in [16] the CT site is coordinated by R276, V355, P356, D394 and T398; a site involving N310, D312 and T92 was proposed [40]; in [41] show an evolution of this site, deeply buried in the protein and coordinated by Y89, T92, S93, N310 and D312. The high-resolution Glt_{Tk} structures recently resolved [19] confirmed the results from [41], and have redeemed the long standing controversy on the nature of the Na3 site; access to this site seems to happen through occupation of Na1, maturation to Na3' (corresponding to the site described in [40]) and finally hopping to Na3 site, which is buried away from solution once another Na^+ ion binds at Na1. There is still a case to be made that Glt_{Ph} and Glt_{Tk} might have different Na^+ sites and thus the third Na^+ has not been yet identified, but the perfect correlation between the Tl^+ sites in Glt_{Ph} and the other two Na^+ sites identified in this structure, together with the conservation of the residues involved in the coordinate, render the argument rather weak. Comparing the position of the residues coordinating the different Na^+ sites in the apo and Na^+-bound conformations, a clear rearrangement of some residues can be observed. The residues undergoing the most substantial shift in positions are M311, T314, R397 and N401 (Fig. 2.6). N401 coordinates directly Na^+ the Na1 site, and it partially occupies this site in the apo state; its rearrangement also decreases the space available for R397 to be in the apo position. The space previously occupied by R397 in the apo position is ulteriorly reduced by the rearrangement of T314, which reduces the distance between TM7 and TM8. The arginine in 397, probably driven by the rearrangements of its environment and the presence of an additional positive charge, undergoes one of the starkest rearrangements; from partially occupying the substrate binding pocket, almost parallel to the membrane plane, it flips to a perpendicular position outside of the binding pocket and interacts with T317 and D390 or D394. It is interesting to highlight here, how R397 and T314 are involved in substrate recognition and thus their rearrangement in absence of the substrate might indicate a priming of the binding pocket for substrate binding in presence of Na^+. Finally, the methionine in position 311 moves away from TM3 and moves toward TM8, driving a rearrangement of the unwound NMDGT motif, freeing up the Na3 site and sterically hindering the closed state of HP2. When mutating this residue a two-fold effect has been

described: modified selectivity for the substrate [42] and the reduced coupling between Na^+ and substrate binding [16]. Direct interaction with the substrate and Na2 would explain the compatibility of the bound-like position of M311 with a closed hairpin only in the $Glt_{Ph,2Na,substrate}$ transporter.

2.6 K^+ binding pockets

Until recently the position of the K^+ binding pocket was still unknown, even though a number of mutations have been demonstrated to impair K^+ binding or coupling.

In a recent study from our group (personal communication, [43]) structural, computational and mutation information allowed us to locate the K^+ binding site. The site is shown in Fig. 2.7, with the residues coordinating the K^+ ion in stick form; the residue numbering corresponds to the crystal structure numbering in [20].

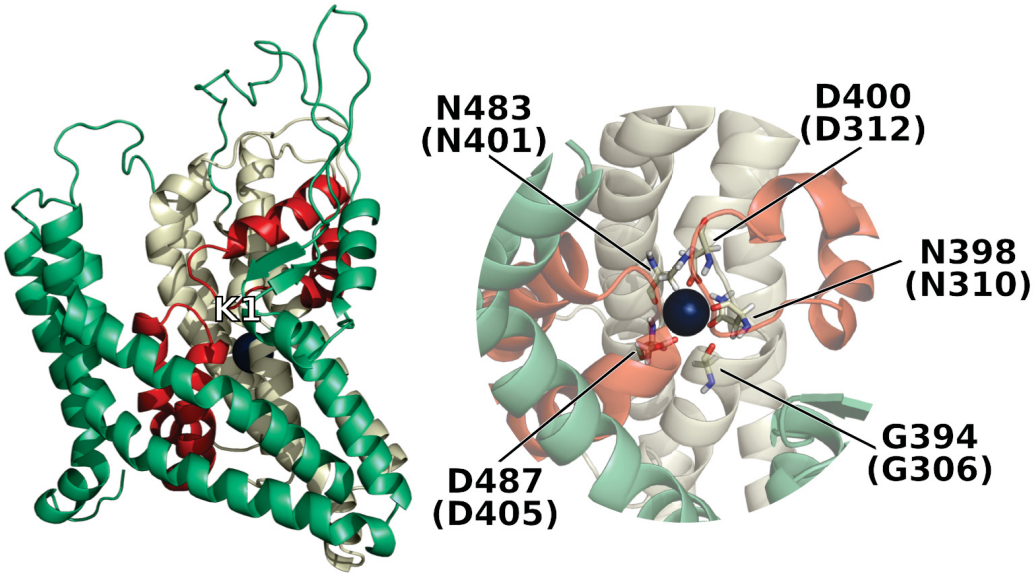

Figure 2.7: The position of the K^+ site, as identified in [43], is shown on the left in a monomer of the EAAT1 crystal structure; on the right the binding pocket is magnified, with the residues involved in coordination shown as sticks and numbered according to EAAT1 crystal structure numbering; in parentheses, the corresponding Glt_{Ph} positions are reported.

Four interaction sites were uncovered through extensive molecular dynamics simulations aimed at observing free binding; dwell time analysis allowed to extract K_d s for each site, while computational electrophysiology calculations [44] combined with electrophysiological measurements of gating charges, proposed the first of these sites as the one occupied during re-translocation.

Interestingly, this site explains some results previously published: mutation of the residue corresponding to D405 in EAAT3 [45, 46] decreases the affinity to K^+; a similar result is obtained by mutating the corresponding to N310 and D312 (N366Q/D386E in EAAT3) [47]. On the other hand, some mutations that disrupt the K^+ control of transport cycle, are not directly participating in the binding site. For example, mutation of Y403 and E404 in EAAT1 and EAAT3 (corresponding to Y317 and Q318 in Glt_{Ph}) to an aspartate disrupt the K^+-induced forward and backward transport [35, 36, 48] and locks the transporter in a obligated exchanger mode. The Y403F mutation is particularly interesting, while it increases the Na^+ affinity, it also allows for proper function in presence of Li^+ and Cs^+. Mutation of R447 (R397) in EAAT3 seems to impair the interaction with the counter-transported K^+ [33], as does the S363R/R447M double mutation(corresponding to R276 and M395 in Glt_{Ph}) [49]. Due to the contribution of residues from sites far from the identified site of K^+ transport, additional information is needed to pinpoint the residues contributing in the mechanism following binding.

2.7 Proton acceptor

The proton binding mechanism is still unknown, despite its strict coupling to substrate uptake. Nevertheless, experiments could indirectly infer some characteristics of the proton binding. Proton association precedes glutamate binding [50, 51] and helps forming the high affinity binding pocket.

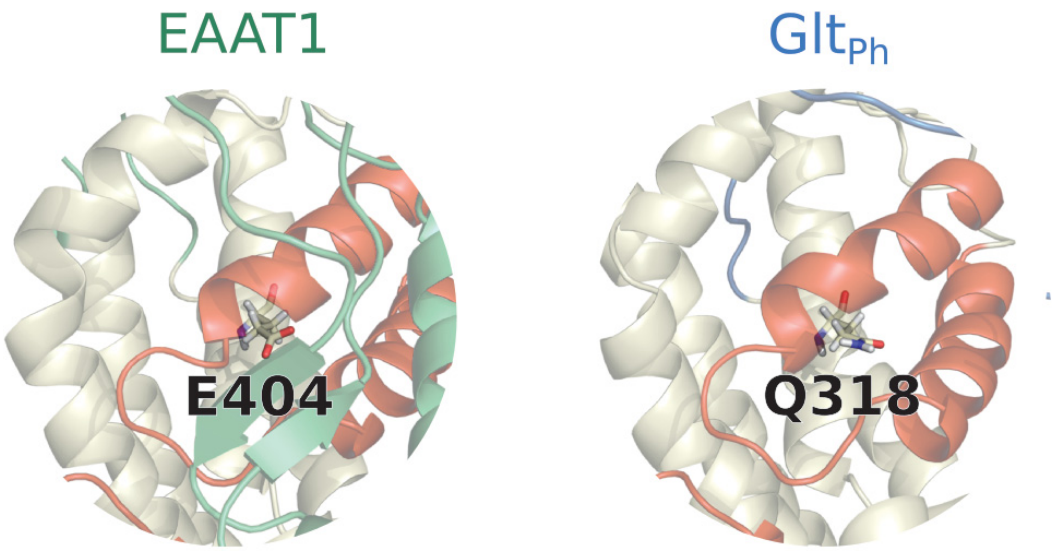

Figure 2.8: The residue believed to be the site of the protonation in EAATs is shown on the left, in a magnification of the binding pocket of EAAT1 crystal structure; on the right, the corresponding residue, a glutamine, is shown in the homolog Glt_{Ph}.

It has been proposed that the residue E404 in EAAT3 (Q318 in Glt_{Ph}) is the site of protonation, on the basis of E373Q being pH-independent and the presence of homoexchange [52, 53]. K^+ and proton binding have been demonstrated to be mutually exclusive [8]. As described in the previous chapter the same residue was believed to participate in K^+ binding and/or coupling to re-translocation. This observation shows how K^+ and proton binding are interconnected and is thus difficult to distinguish the effect of one mutation on one or the other in classical experiments. Adding the information that Na^+ and K^+ binding are pH-dependent, so that changes in pH shift the binding curve and thus modify the affinity for these ions, demonstrates how tackling this problem requires a new approach.

2.8 HP2 as the extracellular gate and its conformational changes

The crystal structures have uncovered the molecular nature of the extracellular gate, which is formed by the two helical hairpins sitting in front of the binding pocket, hairpin 1 (HP1) and hairpin 2 (HP2). The opening of the extracellular gate is given by the movement of HP2 away from the binding pocket and the tip of HP1. The state of HP2 is the main responsible for the ability of the transporter to translocate across the membrane: an open hairpin, and possibly also the tip-open hairpin, is incompatible with movement across the membrane, since the transport domain has to slide along the trimerization domain during the elevator-like movement that characterizes the translocation [16].

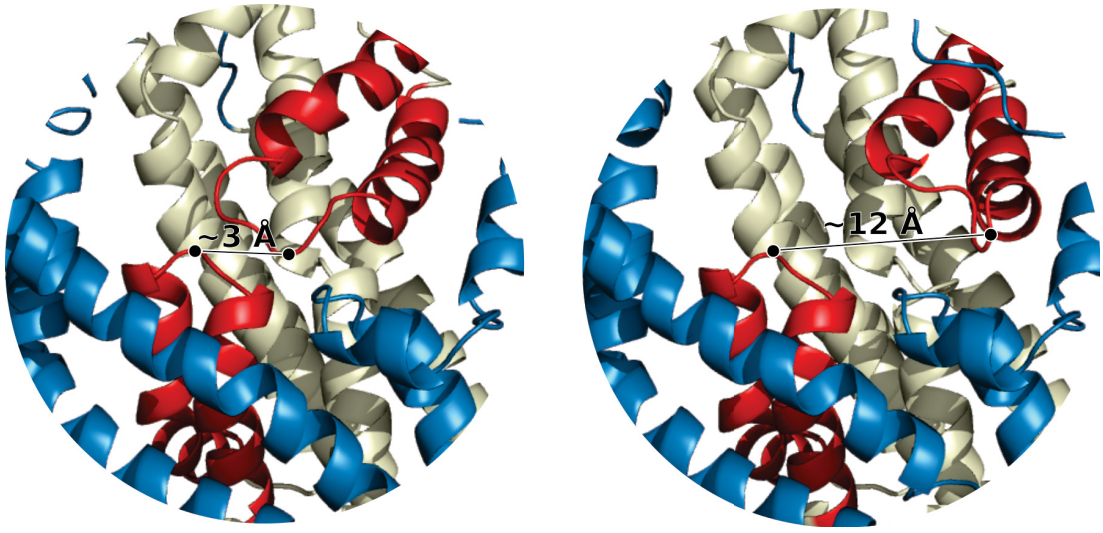

Figure 2.9: Magnification of Glt_{Ph} monomers in the closed (on the left) and open (on the right) conformations, with the hairpins highlighted in red; the residues used to measure the distance between the two hairpins, are indicated as small spheres; the mean distances between tips of the hairpins are shown.

In the snapshots obtained by crystallization, five distinct states of the hairpin can be recognized: closed and wound in the fully-bound state ($Glt_{Ph,2Na,substrate}$) [12, 13]; closed but partially unwound in the apo state ($Glt_{Ph,apo}$) [16]; open half-way or tip-open in the Na^+-bound structure ($Glt_{Ph,Na1}$) [16]; Additionally, it was previously shown how two Na^+ binding is coupled to opening of HP2

and precedes the substrate binding [26], while binding of Na^+ after substrate association is coupled with closure of HP2 [26, 39]. Despite the relevance of coupling between HP2 and ligands binding, no study has analysed in depth how each HP2 conformation contributes to the transport cycle and how its dynamics are controlled by association of the different ligands.

2.9 Conformational selection and induced fit

The classical description of binding of a ligand to a protein is described by a rigid body collision, which does not involve a change in conformation of the protein. This model can be expressed by the system:

$$\begin{pmatrix} d[E]/dt \\ d[EL]/dt \end{pmatrix} = \begin{pmatrix} -k_{on}[L] & k_{off} \\ k_{on}[L] & -k_{off} \end{pmatrix} \begin{pmatrix} [E] \\ [EL] \end{pmatrix}$$

In which k_{on} and k_{off} are the forward and backward rate constants for ligand association (Fig. 2.10). The solution of this system represents the most used model for fitting observed rates of binding:

$$k_{obs} = k_{off} + k_{on}[L]$$

This simplified model cannot always describe the observed rate (k_{obs}) measured in experiment, since the rigid body collision approximation cannot take into account any conformational change that participate in the binding process. Complex ligand binding mechanisms including a conformational change have historically being divided in two broad categories: conformational selection and induced fit (Fig. 2.10). The induced fit model [54] predicts that a conformational change of the binding pocket, which increases the affinity of the protein to the ligand, is induced by the association of the ligand. In the conformational selection mechanism [55], the protein is in equilibrium between multiple states, with the ligand binding to a preexisting conformation and stabilizing it.

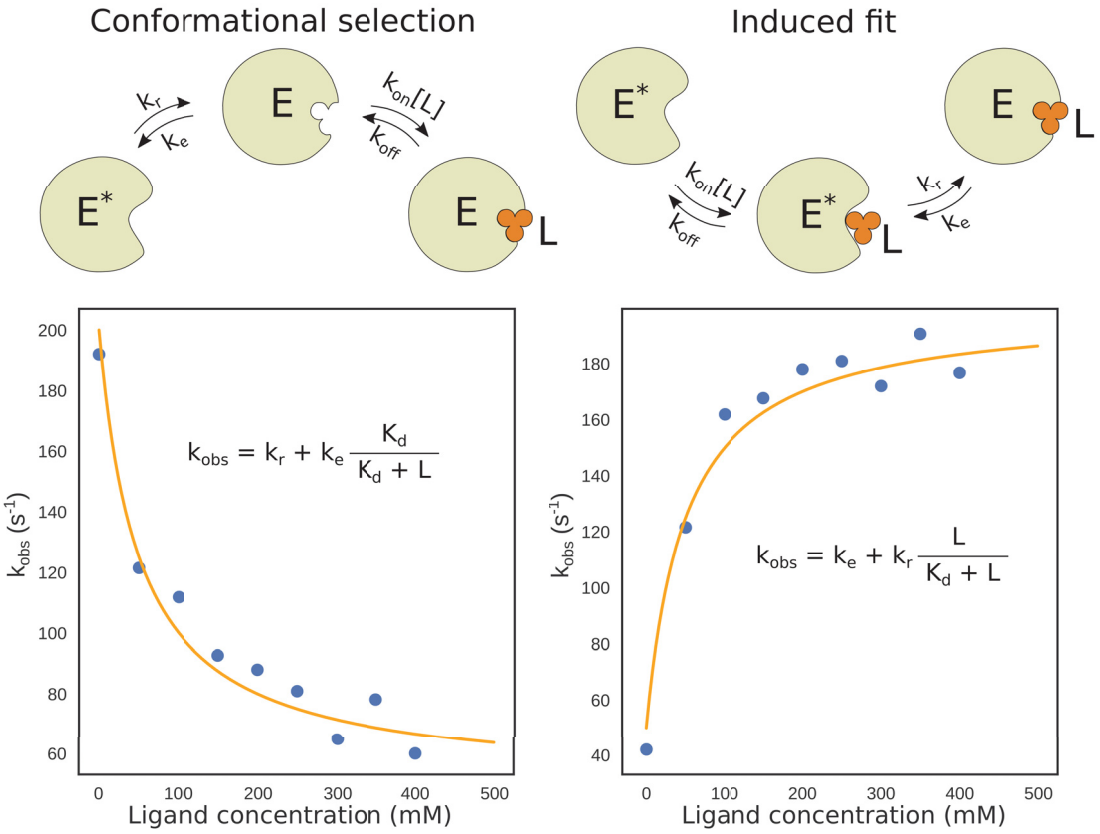

Figure 2.10: The kinetic models representing conformational selection and induced fit are presented here, the corresponding rate constants are shown along the structure-structure transition; on the bottom a simulation of each mechanism is plottes, together with the hyperbolic function describing the respective models.

The first mechanism, induced fit, can be described by the system [56]:

$$\begin{pmatrix} d[E*]/dt \\ d[E*L]/dt \\ d[E]/dt \end{pmatrix} = \begin{pmatrix} -k_{on}[L] & k_{off} & 0 \\ k_{on}[L] & -k_r - k_{off} & k_e \\ 0 & k_r & -k_e \end{pmatrix} \begin{pmatrix} [E*] \\ [E*L] \\ [EL] \end{pmatrix}$$

We can then calculate two solutions, each corresponding to one non-zero eigenvector (λ_1 and λ_2):

$$\lambda_{1,2} = \frac{k_e + k_r + k_{off} + k_{on}[L] \pm \sqrt{(k_{off} + k_{on}[L] - k_e - k_r)^2 + 4k_e k_{on}[L]}}{2}$$

When the rapid equilibrium approximation applies, so that the conformational changes are much slower than binding and dissociation, the solution for the dominant, slower process, can be expressed as:

$$k_{obs} = k_e + k_r \frac{[L]}{K_d + [L]}$$

Similarly, the conformational selection mechanism can be described by the system:

$$\begin{pmatrix} d[E*]/dt \\ d[E]/dt \\ d[EL]/dt \end{pmatrix} = \begin{pmatrix} -k_r & k_e & 0 \\ k_r & -k_e - k_{on}[L] & k_{off} \\ 0 & k_{on}[L] & -k_{off} \end{pmatrix} \begin{pmatrix} [E*] \\ [E] \\ [EL] \end{pmatrix}$$

With the two eigenvectors defined as:

$$\lambda_{1,2} = \frac{k_e + k_r + k_{off} + k_{on}[L] \pm \sqrt{(k_{off} + k_{on}[L] - k_e - k_r)^2 + 4k_r k_{off}}}{2}$$

In the rapid equilibrium approximation the solution for the slower process can be expressed as:

$$k_{obs} = k_r + k_e \frac{K_d}{K_d + [L]}$$

Both of these models introduce a second couple of rate constants, k_r and k_e, which characterize the forward and backward rates of the conformational change. Discerning between the two cases in canonical experiments is challenging, since usually equilibrium properties are measured (i.e. increase in the concentration of the bound species). Even when kinetic experiments are performed it is not easy to relate the observed rate to one of the two mechanisms, due to low time resolution and the difficulty of univocally assigning the intermediate state.

When the rapid equilibrium approximation is valid, so that the rearrangement in conformation is the rate limiting step of the binding process, the observed rate of binding will increase with increasing ligand concentrations for the induced fit mechanism, and decrease for the conformational mechanism [56–59] (fig. 2.10).

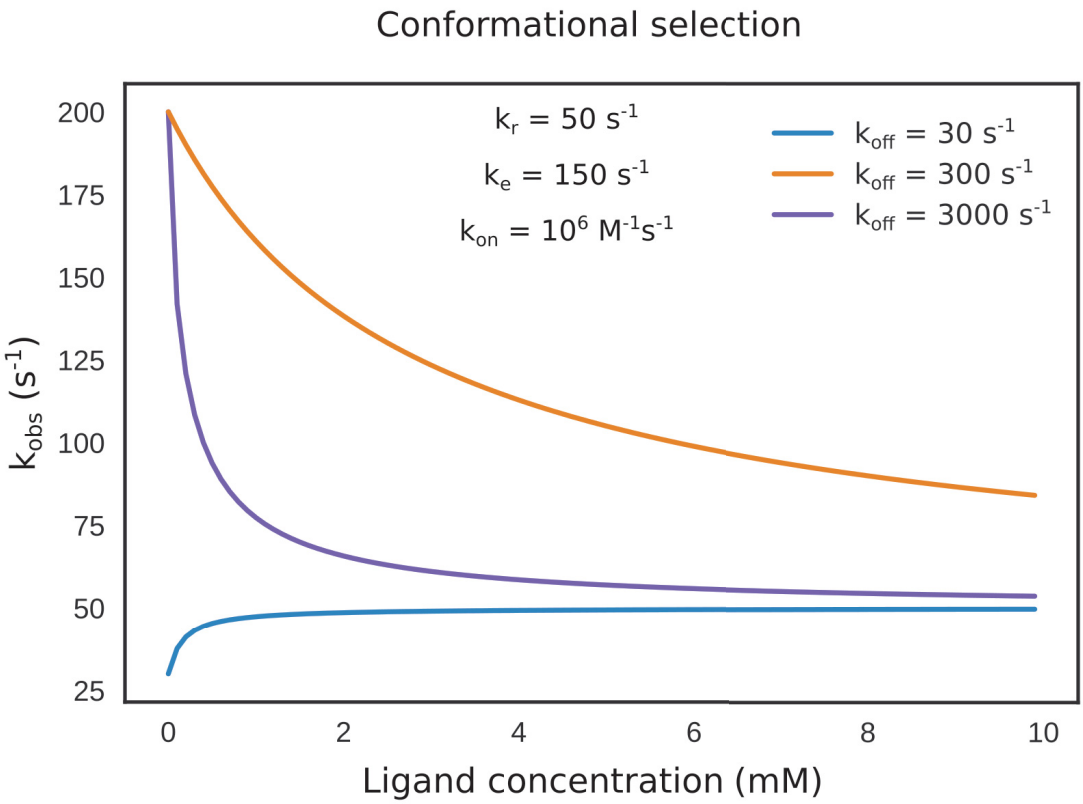

Figure 2.11: Simulations of observed rate versus ligand concentration plot for the conformational selection mechanism, with different k_{off} values; for $k_e >> k_{off}$ the relationship between observed rate and ligand concentrations is inverted, so that k_{obs} increases with ligand concentration.

On the other hand, when the conformational change is not the rate limiting step, two cases can be drawn for the conformational selection scheme: when $k_{off} > k_e$ the decreasing behavior shown for the rapid equilibrium approximation is maintained, while when $k_e > k_{off}$ the observed rate increases with the ligand concentrations also in this scheme [56, 59] Fig. 2.11. In the induced fit mechanism the observed rate increases with the ligand concentration independently of the relationship between k_{off} and k_e.

A more realistic scheme includes elements of the two mechanism and takes into account the possibility that one binding event might happen through different paths, including elements of induced fit and conformational selection.

A study from our group has addressed the mechanism of hairpin closure after substrate binding [60], demonstrating that the process is characterized by an

induced fit mechanism. Since the chemical nature of the substrate determines the efficiency of inducing the conformational change, different saturation level for compounds diverging for only one substituent demonstrated the induced nature of this change. Interestingly, when the protein was preincubated with 5 mM Na^+, k_{obs} was dominated by a slow response to rapid application of the substrate, while preincubating with 400 mM Na^+ the dominant rate was the fast one [60]. These observations suggest that reaction to Na^+ might not be described by an induced fit mechanism.

3. Methods

3.1 Computational methods

3.1.1 MD simulations

Molecular dynamics simulations were performed using the package GROMACS 5.1.x [61] and the AMBER99SB-ILDN [62] force field with the improved ion parameters described in [63]. The lipid bilayer consisted of palmitoyl-oleoyl phosphaditylcholine (POPC) described in [64]. Parameters for soluble amino acids where adapted to the AMBER force field from [65]. Water was modeled using the SPC/E water model [66]. The initial structures where downloaded from the Protein Data Bank (PDB). The missing residues and atoms from the crystal structure where filled in using MODELLER [67]. The same package was used to revert the mutations used in the crystallization process, seven histidines were reverted to the original amino acid for all of the structures and one alanine was reverted to arginine in the apo ($Glt_{Ph,apo}$), and Na^+ bound ($Glt_{Ph,Na1}$) structures. Hydrogen atoms were modeled through virtual sites, in order to increase the timestep to 4 fs to increase the computational performance.

The proteins were embedded in an equilibrated solvated POPC membrane using g_membed [68], according to the position indicated in the Orientations of Proteins in Membranes (OPM) database [69]. The membrane was chosen according to [70], where it was demonstrated that Glt_{Ph} is functional in POPC membranes; moreover, the POPC-based membrane is regularly used for computational studies of this transporter (see [71] as an example). Simulations were performed on trimers in a cubic periodic box of dimensions $\sim$140·140·110Å to ensure a minimum distance between the protein and its nearest periodic image of $\sim$30 Å in every direction. The default protonation state at neutral pH was used for all residues, otherwise else specified. The NaCl concentration was set to 400 mM. The simulation temperature was 310.15 K. In order to increase the probability of Na^+ binding we chose a high salt concentration, but we are reassured in our choice by the fact that *Pyroccocus horikoshii*i is an extremophile which thrives at high salt concentration and, in our experiments, the protein is still functional in NaCl concentrations up to 1 M.

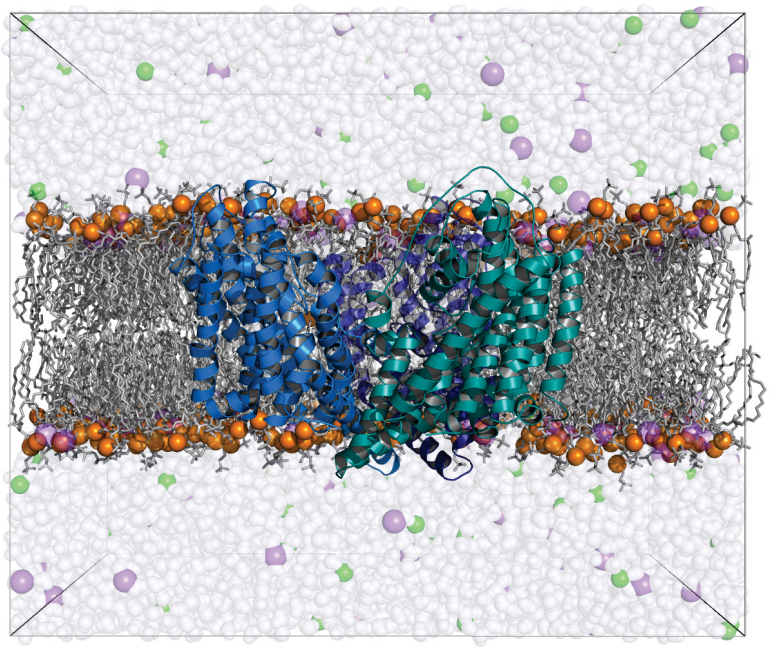

Figure 3.1: A typical MD box used our simulation is shown in gray. Inside, Glt_{Ph} trimer is embedded in a lipid bilayer and soaked in water and NaCl.

The equilibration was performed in three steps, using the v-rescale thermostat [72] and the Berendsen pressure coupling [73]. A 4 ns MD run with positional restraint on the lipids in the z-direction and positional restraint on the heavy atoms of the protein, ensured a first stabilization of the membrane. A second equilibration of 800 ns with position restraint on the protein heavy atoms allows for the membrane to properly pack around the inserted protein. Finally a last equilibration of 10 ns with backbone-only restrains guarantees the rearranging of the side chains maintaining the overall structure. The length of each step of equilibration was chosen to ensure convergence of both the box dimensions and the number of contacts between the protein and the lipids. Free MD was started from the structures resulted from the equilibration, removing all position restraints. The velocities from the last equilibration step were retained in order to use the Parrinello-Rahman pressure coupling [74]. The structural stability of the protein was checked ensuring that the Root Mean Square Deviation (RMSD) calculated for the backbone, excluding hairpin 2, was consistently between 1.5 and 2 Å.

3.1.2 Principal Component Analysis, Essential dynamics and Umbrella Sampling

Principal component analysis is a statistical analysis used to reduce the complexity of a data set, in order to extract the principal components of a certain set of data, based on the observation that a small number of components describes the majority of the diversity found in molecular dynamics data sets, while the majority of these components represent small fluctuations and can be considered as being constrained [75, 76]. In order to obtain these components a covariance 3N×3N matrix (with N being the number of atoms to analyse) is constructed and diagonilized in order to extract the eigenvalues and eigenvectors. The eigenvalues are then reordered in decrescent order, the eigenvectors corresponding to the highest values of eigenvalues will then be the principal components. We used a single coordinate to explore the free-energy landscape underlying hairpin dynamics, in order to minimize computational expense and simplify the analysis. We used cartesian PCA on a concatenated, monomeric trajectory, containing multiple free MD trajectory started from $Glt_{Ph,2Na,substrate}$ (PDB ID:2NWX), in presence and absence of the substrates, and the TBOA bound crystal structure (PDB ID:2NWW) with the inhibitor removed. These simulations contained spontaneous hairpin opening as well as the extreme conformations observed during the simulations. We applied rotational and translation fitting using as reference group the backbone atoms of the transport domain, the same group was used for the PCA calculations. We chose to concentrate on a subset of atoms for PCA in order to reduce the noise due to motions uncorrelated to hairpin dynamics. Essential dynamics sampling could then be used to drive the progression of a simulations along one of these components, describing collective coordinates. The biasing coordinate chosen for free-energy calculation was the first eigenvector obtained by the PCA analysis.

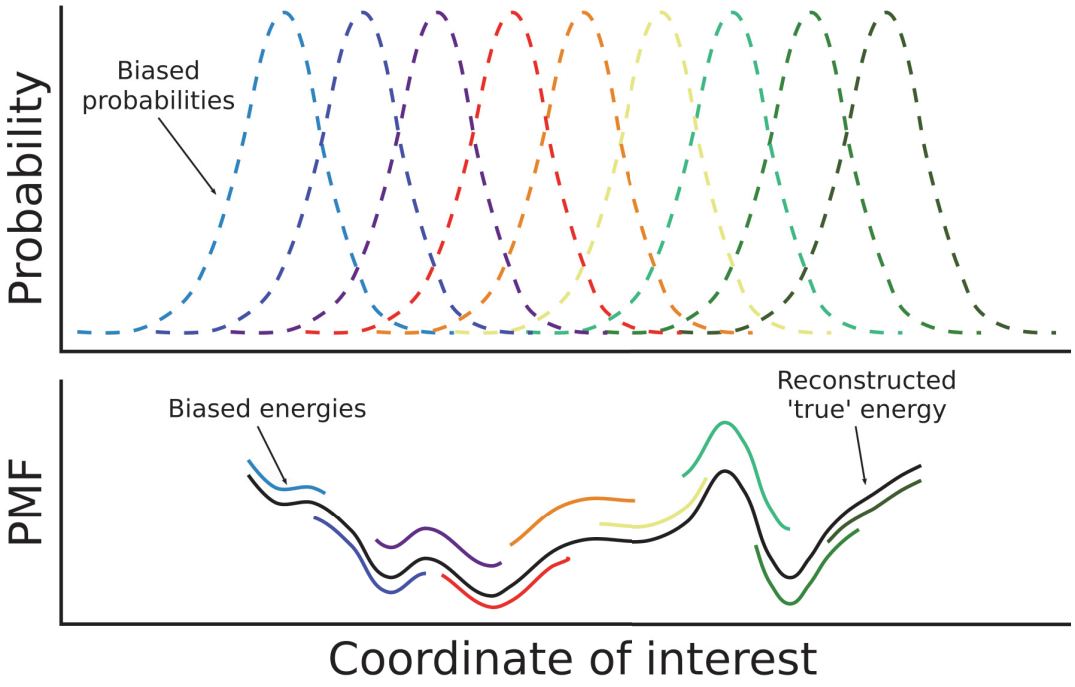

Figure 3.2: Schematic representation of the Umbrella Sampling method; an example of biased probabilities (shown on top), obtained by applying an harmonic potential to the system along a coordinate of interest, in order to obtain overlapping sequential probabilities; from the biased probabilities the energies can be extracted (shown on the bottom, in colors); the Weighted Histogram Analysis Method can then be used to translate these energies, in order to reconstruct the 'true' energy (shown in black).

To obtain the conformation of interest we then applied Essential Dynamics (ED), in order to enforce a constant progression over time along the chosen coordinate. The ED was carried out at a slow pace, with the change between closed and completely open taking 1.5 μs to ensure the minimum possible bias on the rest of the protein. The method used for the free-energy calculation is the Umbrella Sampling (US), in which an harmonic potential is applied to each conformation of interest, to obtain a series of umbrella windows. To impose the bias we used the code for conformational flooding, as implemented in gromacs [77, 78]. The potential was applied in two equilibration steps to minimize bias: in the first run, a potential with a force constant of 1000 kJ·mol^{-1}·nm^{-2} was applied for 10 ns with the Berendsen barostat; and the final conformation was used to

start the production run, with a force constant of 1000 kJ·mol^{-1}·nm^{-2} with the Parrinello-Rahman barostat, from which the first 35 ns were discarded as equilibration. The force constant was chosen to ensure a gaussian distribution for the biased coordinate. In order to estimate convergence of our calculations, the lengths of the production simulations were determined from the time-dependent overlap of unbiased probability distributions for the complete data set with those obtained for non-overlapping 10-ns blocks (Appendix WT apo: 195 ns; WT Na1: 175 ns; WT Na3: 235 ns; WT Na1 Na3: 115 ns; M311A apo: 195 ns; M311A Na1 Na3: 115 ns; R397A apo: 175 ns; R397A Na1 Na3: 175 ns). Statistical errors were calculated with the Bayesian bootstrap method using 30 bootstrap samples. We applied the Weighted Histogram Analysis method (WHAM) [79, 80] to obtain unbiased probabilities.

3.2 Experimental methods

3.2.1 Protein expression and purification

The mutant constructs were generated using PCR-based mutagenesis and verified by DNA sequencing. Expression was performed in Escherichia coli Top10.

After re-transformation, single colonies were picked and inoculated in 5 ml precultures and grown overday, a second 100 ml preculture was grown overnight, both at 37°C. The next day a 450 ml culture was inoculated to obtain an OD at 600 nm of 0.1, and then grown at 37°C until an OD of 2.5 was reached. The induction was performed with 1% L-arabinose, and the protein was expressed for 4 hours at 37°C and 160 rpm. The bacteria were harvested by centrifugation and then frozen at −80°C. After sonication and centrifugation, the supernatant was ultra-centrifuged (1h at 45,000 rpm). The obtained pellet was resuspended in sucrose buffer (250 mM sucrose, 20 mM Tris, 200 mM NaCl) and then solubilized at high DDM concentration (200 mM DDM, 20mM Tris, 200 mM NaCl) for 1.5 hours at 4°C. After another round of ultra-centrifugation (30min at 27,000 rpm), the supernatant was applied to an Ni-NTA-agarose in a binding buffer (20 mM Tris, 200 mM NaCl, 55 mM Imidazole) and incubated in a rotating support overnight. The protein was eluted in a buffer containing 20 mM TRIS, 200 mM NaCl, 1 mM DDM at pH 7.4 and then applied to an Akta column for size

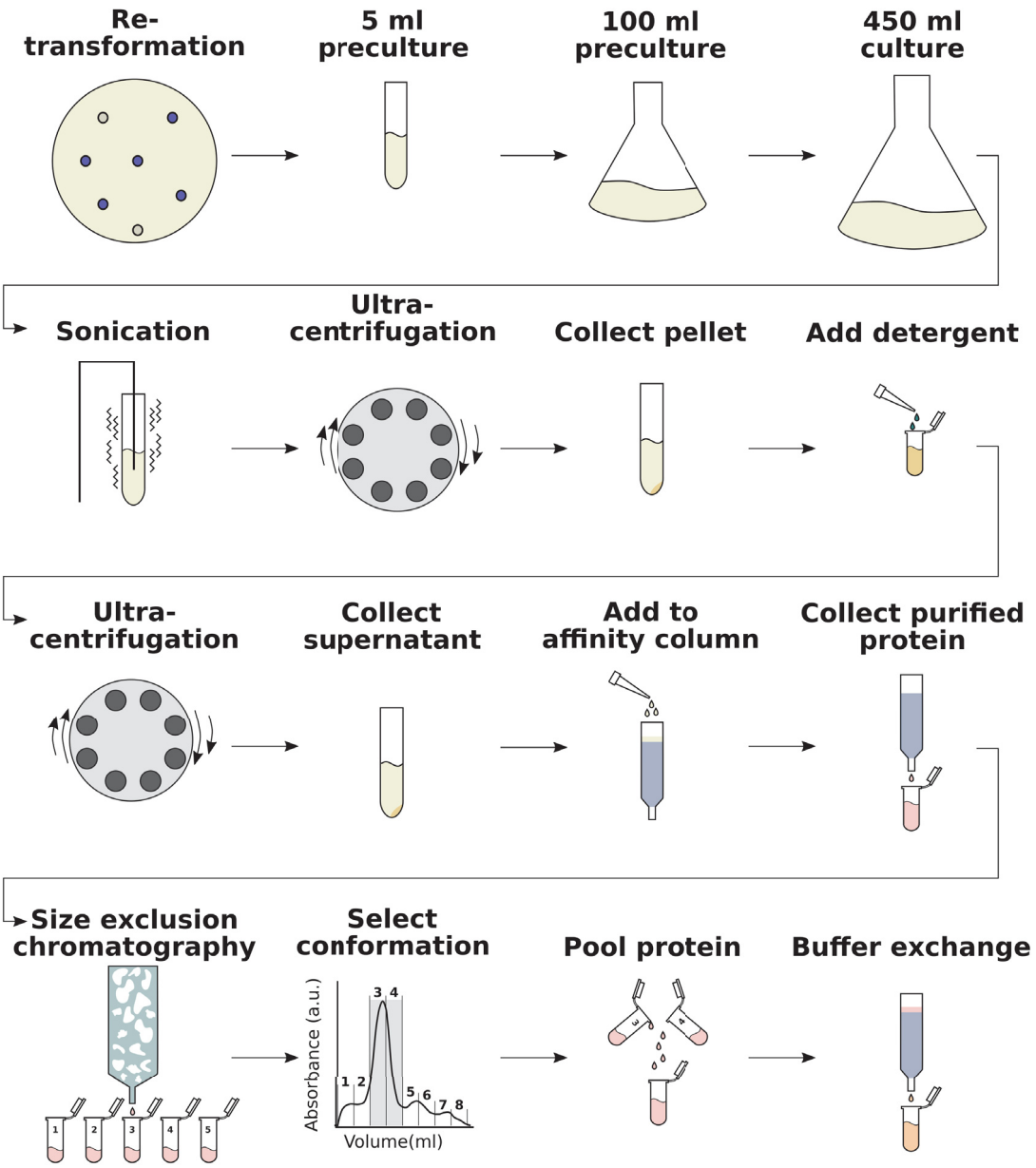

Figure 3.3: Schematic representation of the protein purification process for Glt_{Ph}.

exclusion chromatography, in order to collect the protein from a mono-dispersed peak. Proteins were stored at 4°C for rapid measurement and at −80°C for longer delays. For Na^+-free experiments, three rounds of buffer exchange in disposable salt exchange columns were performed in 200 mM ChCl, 20 mM TRIS, 1 mM DDM at pH 7.4

3.2.2 Stopped flow fluorescence measurements

In our experiments we use a fluorescent reporter, in which a tryptophan is inserted in the tryptophan-free protein, thus conferring a fluorescence signal dependent on Na^+ and the amino acid association. The mutant was described in [81] as a reporter for Na^+ binding to Glt_{Ph}: an increase of fluorescence is elicited by Na^+ application, while a decrease in fluorescence is obtained when applying the substrate in presence of Na^+. Since the increase in fluorescence upon Na^+ application is counteracted by a decrease in fluorescence upon aspartate addition, and it was previously described that Na^+ and aspartate have opposite effects on the hairpin [26], these two observations bring to conclude that the F273W mutant reports on the movement of the extracellular gate.

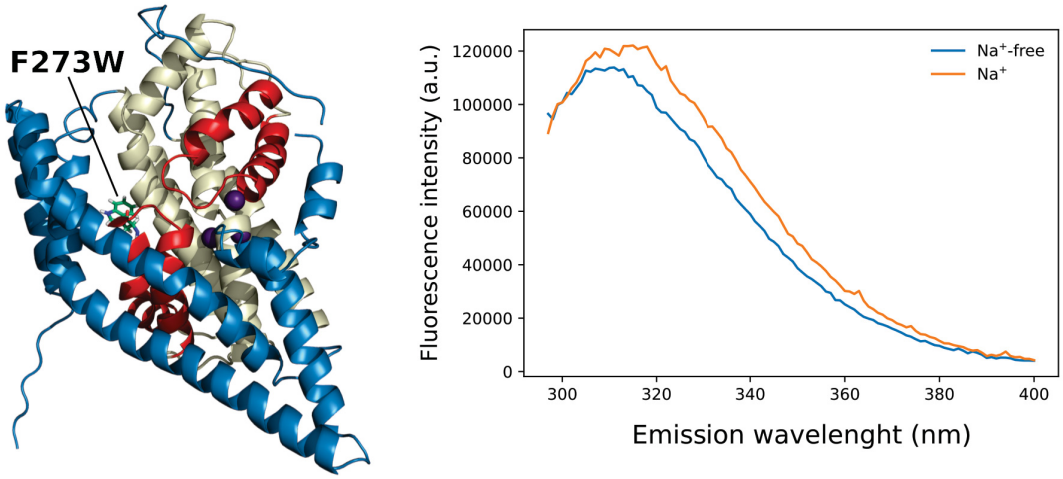

Figure 3.4: On the left, the position of the F273W mutation in the monomer, together with the Na^+ positions for reference, are shown on the left side, while on the right is the excitation curve for the mutant presence and absence of Na^+.

The stopped flow technique allows fast mixing of two reactants: a flow phase

is stopped in order to let the reaction start; an initial part of the reaction happens between the stop and the first observation, the time elapsing between these two points is defined as the dead time. The dead time is characteristic of the machine and the solutions used for the experiments and can be reduced by increasing the rate in the flow phase.

Kinetic fluorescence experiments were performed in a BioLogic Science Instruments SFM400 or a μSFM. Figure 3.5 shows the schematic of a stopped flow machine.

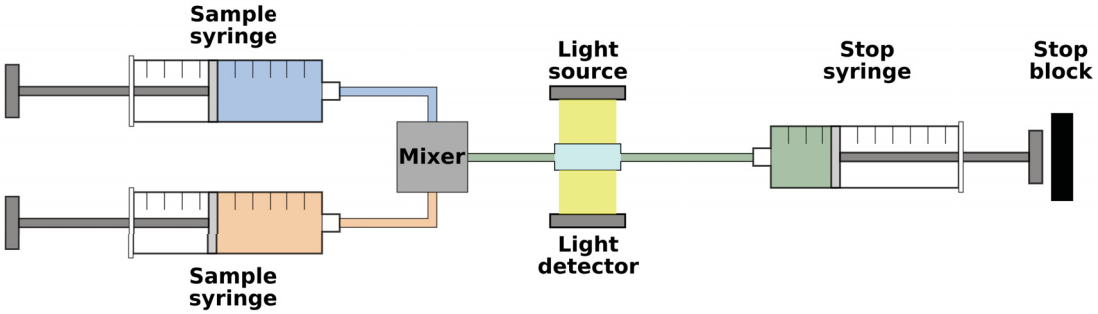

Figure 3.5: Schematic of a stopped flow machine driven by step motors: two syringes are pushed at the same time, the content of which gets efficiently mixed in a mixer, it then reaches at a high flow rate the cuvette with a short delay time; the flow is then stopped by the stop syringe, the sample gets excited and the fluorescence is collected by the light detector.

Excitation was performed at 298 nm with a 298HT05-50U bandpass filter, with 297.7 $\pm$ 1.5 nm as central wavelength (CW) and 5 nm $\pm$ 1 nm as full width at half maximum (FWHM), while the emission was filtered at 340 nm through a ET340/40x bandpass filter, with CW of 340 nm and FWHM of 40 nm. The experiments were performed at 20°C. The theoretical dead time was 1.5 ms; the dead time was calculated by NATA-NBS reaction as 12 ms for the μSFM using our buffers.

3.3 Data Analysis

Simulation data were analysed with GROMACS tools and ad-hoc python scripts, with the aid of the NumPy [82] and SciPy [83] libraries and the MDAnalysis [84], DEAP [85], SCOOP [86] packages. Error bars for the probability profiles

correspond to the standard deviation (SD) of 30 bootstraps and are represented as shadowed areas. Error bars for the kinetic experiments represent the SD of 3-10 independent experiments of which the mean is shown. The curves fitted to the kinetic data are obtained by running a genetic algorithm with 5000 generations for 30 times independently, all curves are shown in the plot; the estimated kinetic rates represent the mean of the independent runs and the errors represent their SD. All plot where prepared in python with the aid of the matplolib [87] and seaborn [88] libraries. Figures of protein structures were generated with PyMol [89].

4. Results

4.1 Computational studies of Glt_{Ph}

4.1.1 MD simulations of Glt_{Ph} reveal spontaneous gate opening in apo structures

In order to investigate the behavior of the extracellular gate in presence and absence of Na^+, we set up multiple MD systems.The accepted model describes a coupling between Na^+ binding and opening of the extracellular gate [26, 51], but the underlying molecular mechanism has not been uncovered. Each of the classical models, induced fit or conformational selection, could potentially describe the Na^+-hairpin coupling, since it involves a binding event and a conformational change. In an induced fit scenario, the Na^+ ions would bind to the closed hairpin with low affinity, inducing subsequently its opening and thus increasing the Na^+ binding affinity, while the conformational selection mechanism presents a scenario in which the hairpin is in a dynamic equilibrium between open and closed states, with the open state having a higher affinity for Na^+; the Na^+ would then preferably bind to the open state, inducing a population shift towards this state.

Structure Conformation Ligands	Identifier	Time (ns)	Structure Conformation Ligands	Identifier	Time (ns)
$Glt_{Ph,apo}$ OF closed apo			$Glt_{Ph,TBOA,Na1}$ OF open Na^+		
	MD1	474		**MD22**	2131
	MD2	173		**MD23**	502
	MD3	155		**MD24**	249
	MD4	1055		**MD25**	243
	MD5	649		**MD26**	246
	MD6	2123	$Glt_{Ph,Na1}$ OF open Na^+		
	MD7	2007			
	MD8	2039		**MD27**	45
	MD9	2336		**MD28**	164
	MD10	922		**MD29**	162
	MD11	948		**MD30**	175
	MD12	938	$Glt_{Ph,Na1}$ OF open $2Na^+$		
	MD13	936			
	MD14	944		**MD31**	323
	MD15	850		**MD32**	301
$Glt_{Ph,apo,F273W}$ OF closed apo				**MD33**	325
				MD34	319
	MD16	240		**MD35**	320
$Glt_{Ph,apo,F273W\ R397A}$ OF closed apo					
	MD17	168			
$Glt_{Ph,2Na,substrate}$ OF closed $2Na^+$ + Aspartate		168			
	MD18	534			
	MD19	173			
	MD20	169			
	MD21	157			

Table 4.1: List of simulations used in this study, showing the crystal structure used, the ligands present, the number of simulations and the simulation time for the system; each system includes three monomers which are analysed independently.

We start our analysis from the two Glt_{Ph} resolved closed structures. These structures where obtained in antithetical fashions: in presence of all the ligands, with the aspartate and Tl^+ density for Na1 and Na2 resolved ($Glt_{Ph,2Na,substrate}$); and in absence of all ligands, with the addition of the R397A mutation, in order to ensure a lower affinity for the substrate ($Glt_{Ph,apo}$). In order to compare the two cases, we reverted the R397A mutation in $Glt_{Ph,apo}$, before starting our

simulations. Fig. 4.1 shows the distance distributions between HP1 and HP2 tips for different simulations. While the maximum peak of the distributions for both closed structures ($Glt_{Ph,2Na,substrate}$ and Glt_{apo}) is in the closed state range ($\leq$6 Å), $Glt_{Ph,2Na,substrate}$ is mostly described by one sharp peak, while $Glt_{Ph,apo}$ has two distinct peaks for the closed state and is able to access values of hairpin tip distances typical of the tip-open and the opened structures. These observations suggest that, despite the high similarity between the crystal structures for the $Glt_{Ph,apo}$ and the $Glt_{Ph,2Na,substrate}$ (with a backbone RMSD of 0.35 $\AA$ for the structures), the dynamic behavior of the protein strongly depends on the ligands occupation, with $Glt_{Ph,apo}$ being much more flexible than $Glt_{Ph,2Na,substrate}$.

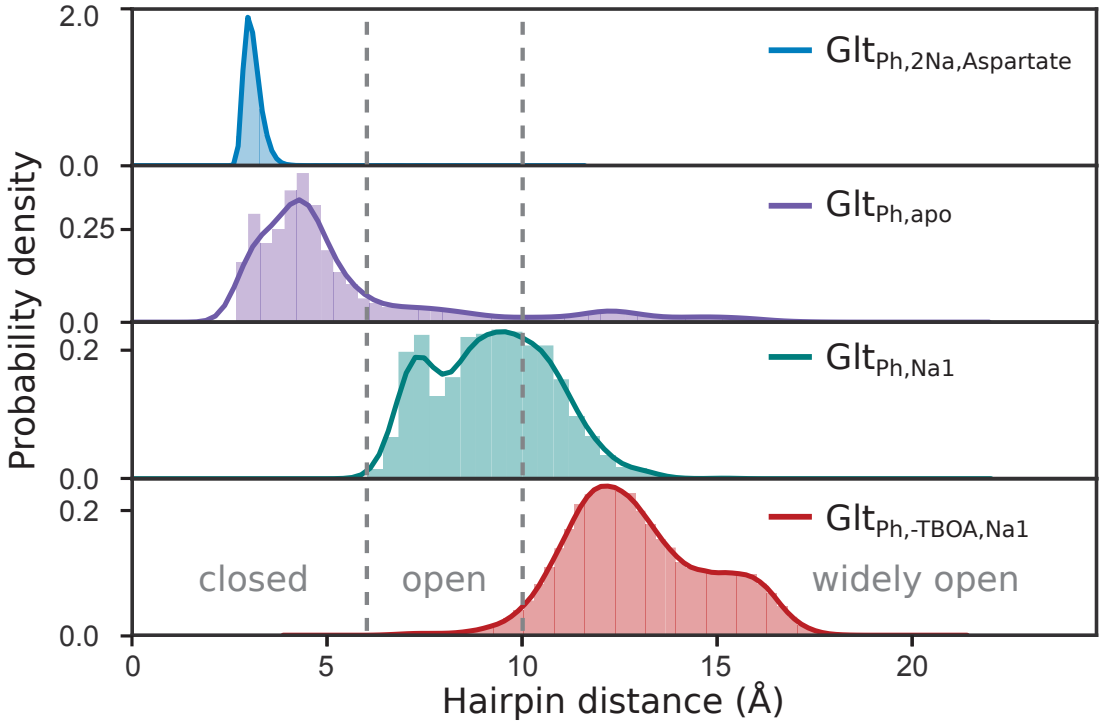

Figure 4.1: Hairpin tip distance (measured as the minimum distance between S279 and G354) distributions from long free simulations (Table 4.1) starting from different crystal structures show a high variability of behavior.

In order to investigate the hairpin dynamics in the open state we focused on the two available Na1-bound open outward-facing crystal structures, obtained either in presence of the inhibitor TBOA ($Glt_{Ph,TBOA,Na1}$) or with the aid of the R397A mutation ($Glt_{Ph,Na1}$). In order to compare the behavior of the two struc-

tures, we mutated R397A back to the wild type and eliminated the inhibitor. Despite the identity in sequence and occupancy between the two modified structures, the hairpin tip distance distributions obtained by the unbiased MD do not overlap. $Glt_{Ph,Na1}$ density peaks at ∼8 Å, with a smaller peak at ∼10 Å; the whole density does not cross the 13 Å mark, with even a small representation of closed structures (∼6 Å). On the other hand, for $Glt_{Ph,\text{-}TBOA,Na1}$ the probability of distances below 8 Å is negligible, with the main peak at ∼13 Å and a second peak at ∼16 Å. When looking at the R397A reverted structure simulated in presence of Na1 and Na3, we see an enrichment of distances higher than ∼12 Å, with a partial overlap to the TBOA-deprived structure values. These results suggest that occupation of the Na^+ binding sites modifies hairpin dynamics, decreasing the intrinsic flexibility of the protein, stabilizing different conformations.

4.1.2 Na^+ binding to the apo structure is very slow and follows hairpin opening

To investigate the coupling of Na^+ binding to hairpin dynamics it was important to closely observe the dynamics of Glt_{Ph} in the apo state. We set up independent μs-long simulations of $Glt_{Ph,apo}$, and collected the observed Na^+ binding events. Binding of Na^+ at Na1 is a very slow process (Fig. 4.3A), so much so that we were able to observe only a limited number of binding events (Fig. 4.2A).

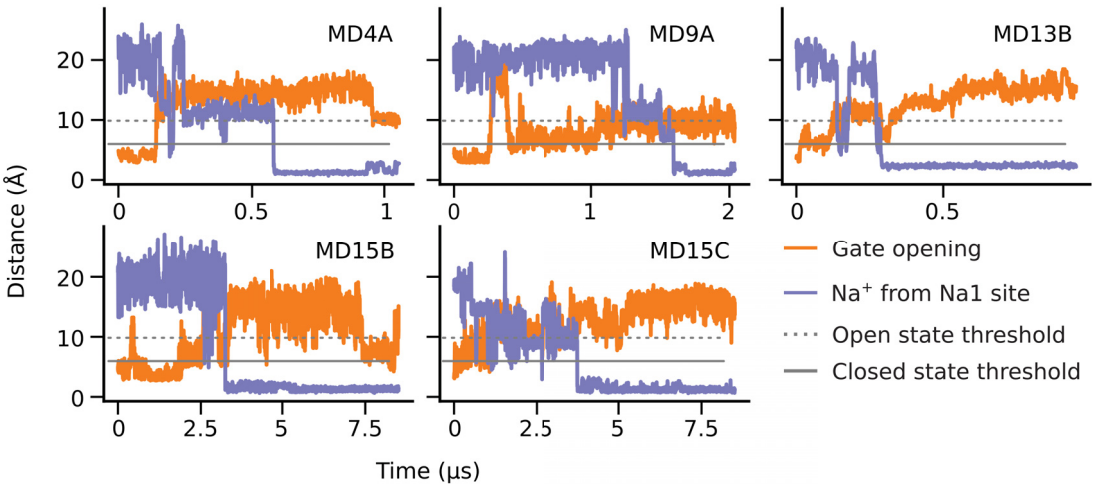

Figure 4.2: Time course of HP2 opening (measured as distance between S279 and G354 at the tip of HP1 and HP2, respectively) and Na^+ distance to the Na1 site in unguided MD simulations of trimeric GltPh starting from the outward-facing, gate-closed, apo state (PDB ID: 4OYE). Results from independent simulations with spontaneous Na1 binding (5 out of 45 unbiased MD simulations).

We could not observe an unbiased binding to the Na3 site in any of our simulations, indicating that the binding of Na^+ at the Na3 site might be even slower or might involve Na1 as an intermediate step. In contrast, running a simulation in absence of bound Na^+ on a structure that was crystallized in presence of Na^+, Na^+ binding at site Na1 happens much faster, while still no binding at the Na3 site could be observed.

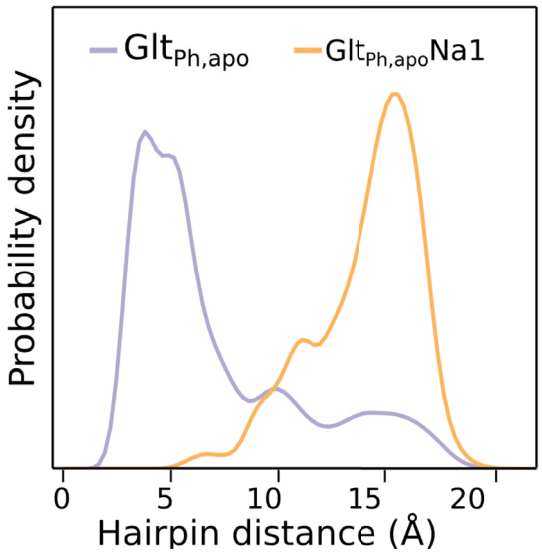

Figure 4.3: Representative time course of distance between hairpin tips and of minimum distance of Na^+ from the Na1 site is shown on the left, while a cumulative hairpin distance distributions from different simulations, before and after Na^+ binding, is shown on the right.

Part of the difference in observed binding events might be explained by the state of the hairpin: while $Glt_{Ph,apo}$ has been crystallized in a closed conformation, the Na^+-bound structures where crystallized in open conformations. We then investigate the possibility of a preferential order of events between Na^+ binding and hairpin opening. The unbiased MD is the perfect tool to investigate this possibility, since it retains order of events, in contrast to more computationally expensive enhanced sampling techniques. Despite its usefulness, unbiased MD is usually unable to reveal the true equilibrium properties of slow processes. The few instances of unguided Na^+ binding events at Na1 site all happened after opening of the hairpin. This observation supports the hypothesis that the hairpin opens irrespectively of Na^+ binding, as already suggested by the high hairpin flexibility described in the previous section. The data obtained to this point suggest a conformational selection mechanism for Na^+ binding, but fail to explain the delay time between hairpin opening and binding, suggesting that another loosely coupled mechanism or rearrangement might be involved in the binding events.

4.1.3 Order of Na^+ binding

The next question we addressed in order to characterize the role of Na^+ was the order of binding at the different sites. From the unbiased MD we can observe direct binding at Na1 but not at Na3. This observation suggests that Na1 is the first to bind, then followed by Na3 or that the binding of Na3 includes passage through the Na1 site, as previously suggested [41].

We analysed simulations with different Na^+ occupancies in order to extract the behavior of each of the Na^+. The simulations with Na^+ occupying Na3 or Na1 and Na3 sites, present a stable coordination of the Na^+ ions, compatible with the corresponding crystal structure, both in absence and presence of the substrate. When only Na1 is present, instead, we notice an evolution of the coordination over time: the distance from D312 tends to decrease, while the distance from N310 increases; more surprisingly, the interactions with D405 and G306, residues which characterize the coordination of Na1, are lost (Fig. 4.4).

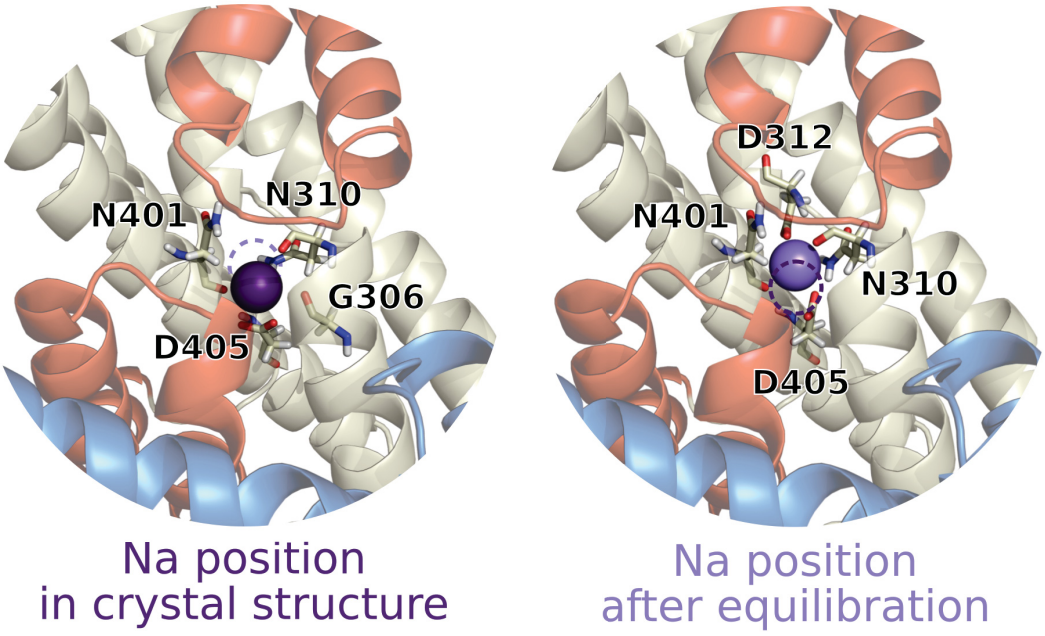

Figure 4.4: The positions of Na1 (on the left) and Na1' (on the right) in the binding pocket, togheter with the residues coordinating them; a dotted circle shows the alternative position.

This rearrangement in the coordination is reflecting a movement of the Na^+

ion towards a more buried position within the protein, which brings the ion closer to some of the residues that coordinate Na3 (namely Y89, T92, S93). This evolution seems to indicate a progress towards the Na3 site, and thus the hopping from Na1 to Na3, but the process remains incomplete in all of our simulations: the original Na1 coordination is lost but the ion does not reach the Na3 coordination state. This observation brings us to speculate that a following event might be necessary to favor the hopping. This event might involve a conformational rearrangement of the protein or the interaction with other Na^+ sites, which might confer the additional energy needed to induce the hopping thanks to charge repulsion effects.

4.1.4 A binding pocket rearrangement is needed to prime the Na1 site

As described in the introduction, a series of rearrangements characterize the change between the apo-like and the bound-like states of the binding pocket 4.5. Rearrangements of the residues surrounding the binding sites may be involved in the slow cooperative binding of Na^+ ions to the Na1 and Na3 sites in the absence of aspartate (Ewers et al., 2013, Hänelt et al., 2015, Reyes, Oh et al., 2013), but it is unclear how these conformational changes in the binding pockets are coupled to HP2 dynamics and Na^+ binding.

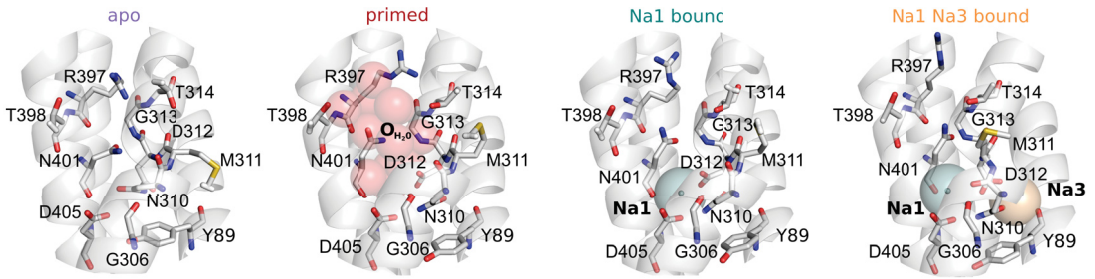

Figure 4.5: The Na^+ binding pocket undergoes subsequent rearrangements with different Na^+ occupations: the apo and Na1 Na3-bound structures correspond to the crystal structures, while the primed and Na1-bound structures are snapshots from representative free Na^+ binding simulations.

In order to investigate their role, we measured inter-residue distances for residues around the ion-binding pockets, not including HP2 residues, in multiple

independent unbiased MD simulations starting from the apo state (total simulation time: ∼50 μs, Table 4.1 and) with and without spontaneous HP2 opening and Na1 binding, and from the Na1/Na3-bound state (total simulation time: ∼4.9 μs, Table 4.3). We then labeled each trajectory frame according to water accessibility and Na^+ occupation (apo, water-accessible/primed, Na1-bound, or Na1/Na3-bound) and trained a support vector machine (SVM) classifier on the inter-residue distances. The assigned functional states corresponded well to the labeled ones (apo, 96.5%; primed, 95.1%; Na1-bound, 47.4%; Na1/Na3-bound: 100%). We next selected those residue pairs that closely interact (mean distance $\leq$ 10Å) and present a significant distance-distribution change between states (overlap with apo distribution $\leq$ 50%).

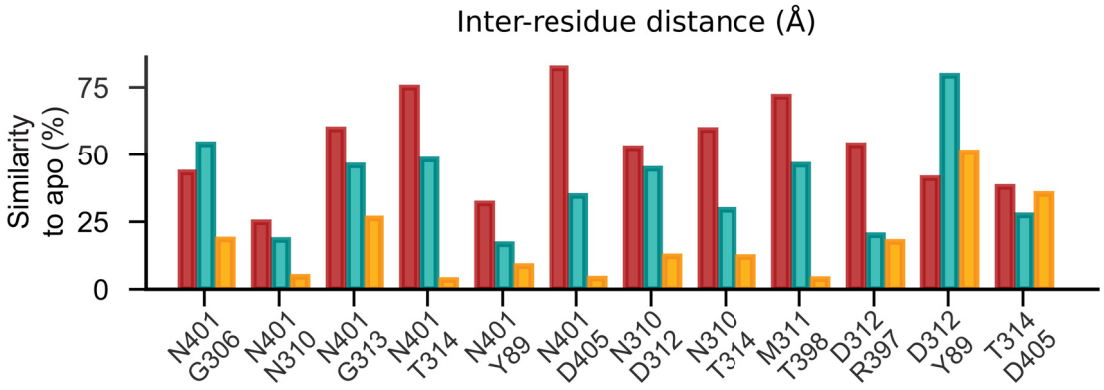

Figure 4.6: Concerted movements of residues precede and follow Na^+ binding, as shown by the overlap of normalized distance distributions in the different states with the apo state for the distance pairs between identified residues involved in the conformational changes.

This analysis revealed that only few residues (Y89, G306, N310, M311, D312, G313, T314, R397, T398, N401, and D405) move significantly during these rearrangements (Fig. 4.6). In particular, opening of HP2 is followed by an upward movement of N401, which breaks the N401-310 interaction, a concomitant retraction of N310, which in $Glt_{Ph,apo}$ partially occupies the Na1 site, completes the priming of the pocket for Na^+ binding increasing its water accessibility. While this rearrangement primes N401 for the N401-T314 bond that characterizes the Na^+ bound structures, this bond does not form before Na^+ binding. The formation of this bond might depend also on the state of the NMDGT motif, the

unwound tract of TM7 which assumes different states due to M311 orientation. Interestingly, when in its apo-like position, R397 seats right in between N401 and T314, thus interfering with the formation of the bond and the correlated shift in TM7. Binding of the first Na^+ ion induces the upward movement of R397, while the initial movement of M311 frees up the Na3 site and is stabilized by a weak interaction between N401 and T314. When both sites are occupied, R397 moves further away from the binding pocket, N310 positions itself between the two bound ions and M311 flips upward to lock HP2 in the widely open state.

This observation highlights that the binding of Na^+ at Na1 not only must be preceded by a pocket rearrangement, but also followed by an additional change, in order to reach the bound-like structure.

4.1.5 Na1 and Na3 are both needed to ensure the stabilization of open hairpin state

Given the intrinsic slowness of the processes preceding Na^+ binding, the low number of binding events for Na1 and the inability to observe spontaneous Na^+ binding at the Na3 site, we proceed our investigation using enhanced sampling MD. Enhanced sampling techniques are aimed at speeding up the progression along one or more coordinates of interest. Some care has to be taken in deciding when to use these methods, since the results strongly depend on the chosen coordinates and their use does not generally preserve the order of events, i.e. the cause-effect relationship. Here, we can use the unbiased MD to extract the order of events for Na^+ binding and the coupled conformational changes; moreover, the multiple long simulations allow us to obtain a sensible collective variable describing our measure of interest, hairpin distance. We used Principal Component Analysis (PCA) in order to describe the complex conformational change of hairpin opening with one collective coordinate, reported by the first eigenvector. The data set analysed consisted of several unbiased apo simulations, starting from open and closed structures; we selected a subgroup of atoms to avoid unnecessary noise: the backbone atoms of the transport domain, excluding the loops. The first calculated eigenvector has a high correlation to hairpin dynamics and accounts for 50% of the variability in the data set.

We used umbrella sampling, applying a restraining potential to trap the protein in different conformations along the coordinate of interest; each conforma-

tion was then sampled in independent simulations, named windows. The unbiased probabilities were then extracted using the Weighted Histogram Analysis Method (WHAM)[79, 80], which weighs the contribution of each independent window and shifts the calculated energies in order to reconstruct the probability density and thus the potential of mean force (PMF). We selected the $Glt_{Ph,apo}$ as initial structure, to which we applied the restraining potential in order to minimize the equilibration time, since the closure of the hairpin could only seldomly be observed in our simulation, in contrast to the more probable event of opening of the hairpin.

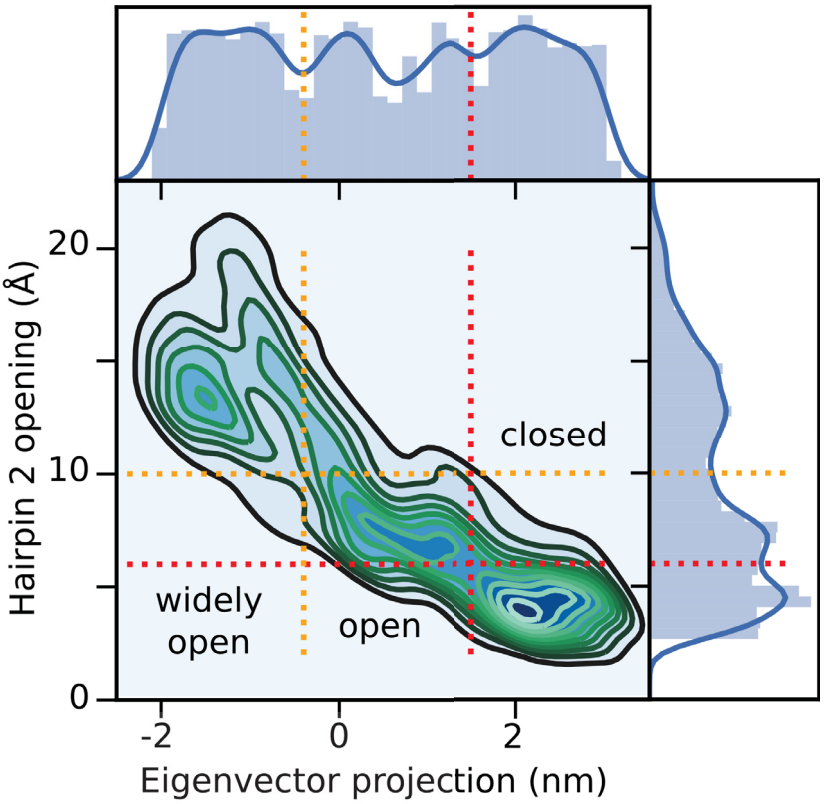

Figure 4.7: The correlation between hairpin distances and eigenvector projections is shown as a 2D kernel density plot, while the 1D densities are showed on the side of the plot; the threshold for the open (orange) and widely open (red) states are shown as dotted lines.

Fig. 4.7 shows the correlation between hairpin distance and eigenvector projection for the apo protein. A fitted line to this correlation allows to translate the

thresholds extracted from the unbiased MD, to values of eigenvector projections. Thanks to the calculated relation, we can cluster the eigenvector projection values in different states, corresponding to the hairpin states described in the previous chapters. The thresholds are represented in the plot as dashed lines, indicating the closed, open and widely open states. These states correspond to the peaks observed in Fig. 4.1: the closed state corresponds to the peaks at ∼3 and ∼6 Å, typical of the $Glt_{Ph,2Na,substrate}$ and $Glt_{Ph,apo}$ structures, respectively; the open state includes the peak found at ∼10 Å, corresponding to the Na1 bound structure; finally, anything above ∼10 Å was categorized as widely open, characteristic of the TBOA-bound structure.

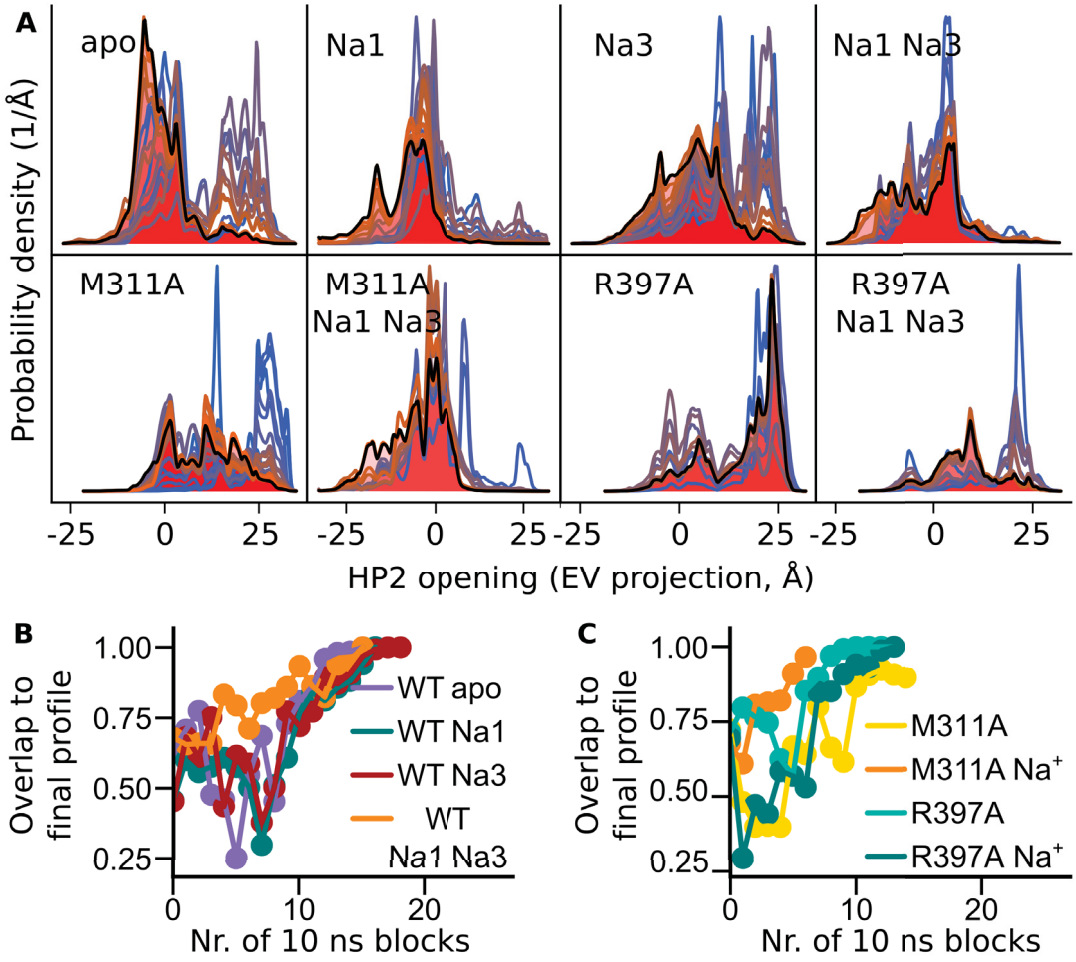

Figure 4.8: Convergence was estimated by the overlap between profiles calculated by increasing 10 ns-blocks on the final profile (in black in A), then plotted against the number of blocks for WT (in B) and the mutants (in C).

The convergence of the probability profile was estimated by diving the trajectories of each window in 10 ns-long blocks and calculating the profile for each block. The overlap to the final profile was calculated additionally for each block and plotted against the number of blocks (Fig. 4.8). Fig. 4.9 shows the probability distributions for different occupation of the Na^+ sites. Table 4.2 shows the probability of different states, obtained by calculating the area under the curve in a range defined by the thresholds discussed above, and then normalizing it to the total area under the curve. The probability distribution for the apo state shows an equilibrium between open and closed state skewed towards the open state.

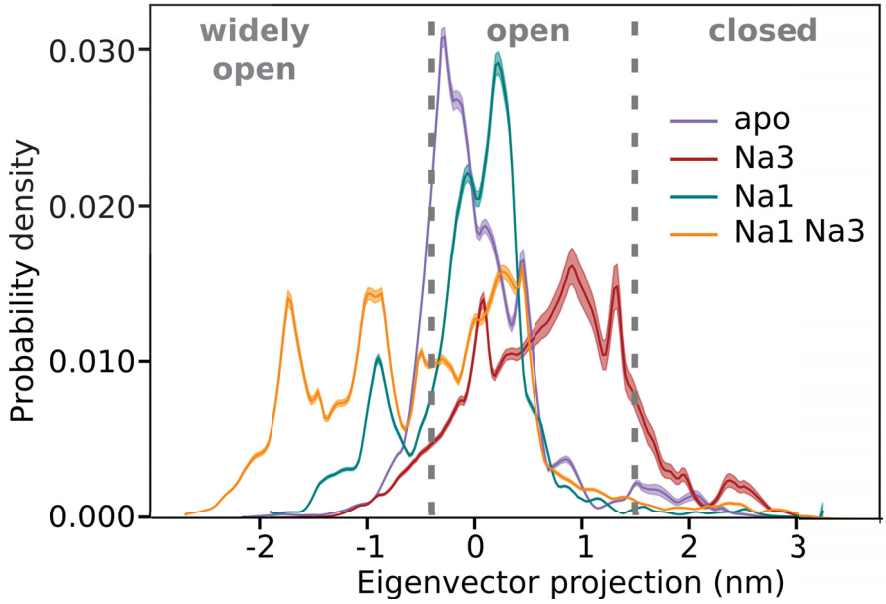

Figure 4.9: Hairpin distance probabilities, obtained by umbrella sampling, are shown for differential Na^+ occupation; dashed lines show the thresholds for closed, open and widely open states.

The results for the apo state are rather surprising, with the highest probability found for the open state (∼82%), since the $Glt_{Ph,apo}$ was crystallized in the closed state. Comparing the results with the dynamics observed in unbiased MD, though, the possibility to access the complete open state in the apo state is compatible with these results. To this argument we can add the observation that the protein crystallized in [16] contained the R397A mutation, which might influence the relative equilibrium between the different states, partially uncoupling Na^+ binding from hairpin opening/closure. We can look at the effect of different Na^+

occupancy on the hairpin dynamics, in order to extract the role of each Na^+.

Ligand occupation	Length (ns)	# windows	State probability closed	open	widely open
apo	205	50	5%	82%	13%
Na1	215	56	2%	77%	21%
Na3	235	48	13%	81%	6%
Na1 Na3	215	65	4%	46%	50%

Table 4.2: Probabilities for different hairpin states are calculated from the unbiased probabilities obtained from umbrella sampling and are shown in relation to Na^+ occupation.

While Na1 increases the probability of the open states (open and widely open), Na3 strongly increases the probability of the hairpin to be closed. These results show opposite effects of Na1 and Na3 on the open/closed equilibrium, while both decrease dramatically the flexibility of the protein.

Finally occupation of both Na1 and Na3 moves the equilibrium even more towards the open state, with an enrichment of the widely open state in comparison to apo, and an even sharper decrease in probability for the closed state, when compared to the Na1-bound. This last result, coupled with the observation that Na1-bound structures obtained from unbiased binding have a different orientation for some key residues, lead us to speculate that the Na1-bound structure obtained in presence of TBOA might already have Na3 site occupied, but not visible. As described in the introduction, the Na^+ sites were rendered visible thanks to soaking in Tl^+. This metal has been shown to substitute Na^+ [38] in order to allow the binding of the substrate, and is heavy enough to give an unambiguous scattering even at low resolutions. The fact that Na^+ and K^+ can both be substituted raises some questions on the specificity of this binding. In [43] the K^+ sites do not include an equivalent of Na3, in accordance with our hypothesis that Tl^+ cannot bind at Na3. Confirming our hypothesis, a recent study of Glt_{Tk} [19] revealed that the Na3 site is occupied and that its high selectivity would not allow for Tl^+ to bind.

These results show that both Na1 and Na3 occupation is needed in order to stabilize the open states, even though the same states can be reached by the

apo protein; addition of only one of the two Na^+ ions significantly restricts the occupation to one preferred state.

4.1.6 A new Na^+-bound crystal structure reveals the native state of the extracellular gate upon Na^+ binding

As described in the introduction and in the previous section, the only Na^+-bound crystal structure was obtained by mutating a conserved arginine (R397) [16]. Moreover, the position of the Na^+ sites in Glt_{Ph} were resolved with the aid of Tl^+ soaking. Due to the short-comings previously described, it was relevant to obtain Na^+-bound crystal structure of the WT Glt_{Ph}, at a resolution high enough to be able to locate the Na^+ ions bound.

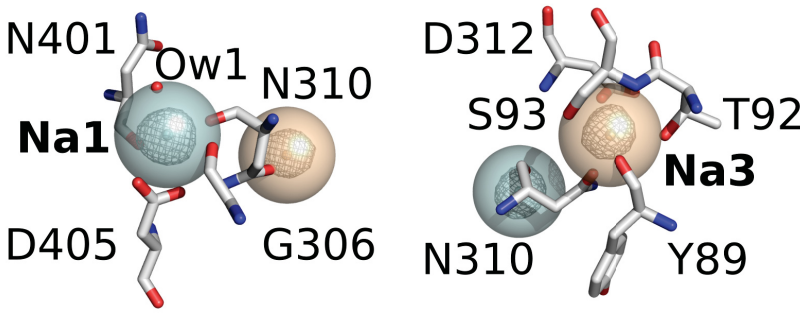

Figure 4.10: The Na^+ ions binding pockets, for Na1 (on the left) and Na3 (on the right), are shown in the magnifications, with surrounding residues highlighted as sticks and with the omit maps of Na^+ ions (contour level 5 σ).

With the aid of the lipidic cubic phase crystallization technique [90], a high resolution crystal structure in a lipid bilayer was obtained recently by our collaborators. The overall folding is highly similar to the previously resolved structures. Close inspection of this crystal structure reveals some unexpected results. The first detail to underscore is the presence of two Na^+ ions bound to the protein: the Na1 site corresponds to the site described by Tl^+ soaking in [12, 16], while the Na3 site is resolved for the first time in Glt_{Ph}, after being postulated in computational studies [41] and being resolved in Glt_{Tk} crystals [19]. The Na^+ sites have been highlighted in Fig. 4.10.

Structure Conformation Ligands	Identifier	Time (ns)	Structure Conformation Ligands	Identifier	Time (ns)
$Glt_{Ph,Na1\ Na3}$ OF open $2Na^+$			**$Glt_{Ph,Na1\ Na3}$R397A OF open $2Na^+$**		
	MD36	65		**MD46**	446
	MD37	65		**MD47**	449
	MD38	66		**MD48**	446
	MD39	163		**MD49**	546
	MD40	163		**MD50**	157
	MD41	246			
	MD42	236			
	MD43	240			
	MD44	238			
	MD45	163			

Table 4.3: List of simulations used in this study starting from the new crystal structure, showing the mutation used, the ligands present, the number of simulations and the simulation time for the system; each system includes three monomers analysed independently.

The second observation involves the HP2 position and can be seen in Fig. 4.11B, where the newly crystallized monomer is superimposed on the two open structures that were previously resolved, the $Glt_{Ph,Na1}$ structure, on the top, and $Glt_{Ph,TBOA,Na1}$, on the bottom.

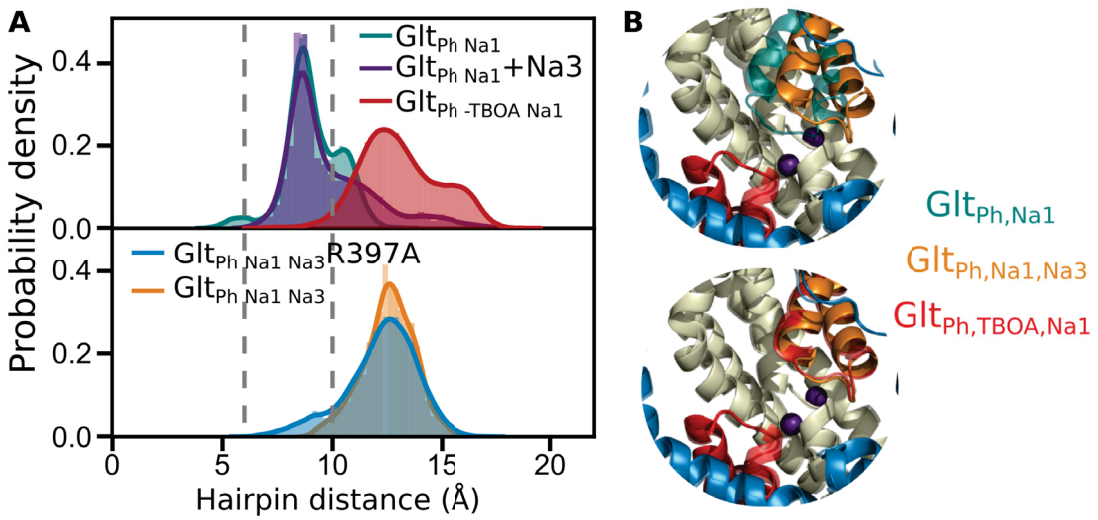

Figure 4.11: On the right, the newly resolved crystal structure [22], is shown superimposed on the two previously resolved Glt_{Ph} crystal structures, in order to underscore the differences and similarities of the hairpin position in the two cases; on the left, hairpin distance distributions for different structures are shown for comparison.

The hairpin position overlaps with $Glt_{Ph,TBOA,Na1}$, while $Glt_{Ph,Na1}$ presents an intermediate state between the closed and open state. Up to now, the wide opening of the $Glt_{Ph,TBOA,Na1}$ was regarded as a forced position induced by the inhibitor, with the more modest opening of the Na1-bound structure accepted as the physiological opening state. Even during MD analysis (Tables 4.1 and 4.3), the main peaks of $Glt_{Ph,Na1}$ and $Glt_{Ph,TBOA,Na1}$ do not overlap(Fig. 4.11A). We tried to revert the difference between the new crystal and $Glt_{Ph,Na1}$, by adding a Na^+ at the Na3 site in $Glt_{Ph,Na1}$ and the R397A mutation in $Glt_{Ph,Na1,Na3}$. Both these additions ensure partial occupation of the area that corresponds to the main peak in the other. Despite these results, the distribution do not completely overlap, possibly due to insufficient sampling. As seen previously in this work (Fig. 4.1), the widely open state of the TBOA-bound structure could be reached in MD simulations starting both from the apo and the Na1-bound structure. The new crystal structure confirms that the widely open hairpin represents a physiological state, described by the ∼12 Å peak in the gate distribution shown in Fig. 4.1.

4.1.7 R397A and M311A mutations affect gate dynamics

We applied the Umbrella Sampling enhanced technique on two mutants, which had previously been described as participating in the Na^+/hairpin coupling and in substrate recognition [15, 16, 26, 33, 42].

As detailed in the introduction, both these residues switch conformations upon Na^+, thus we decided to investigate their effect on the apo protein. The probabilities extracted from the enhanced sampling are shown in Fig. 4.12, together with the WT probabilities for reference.

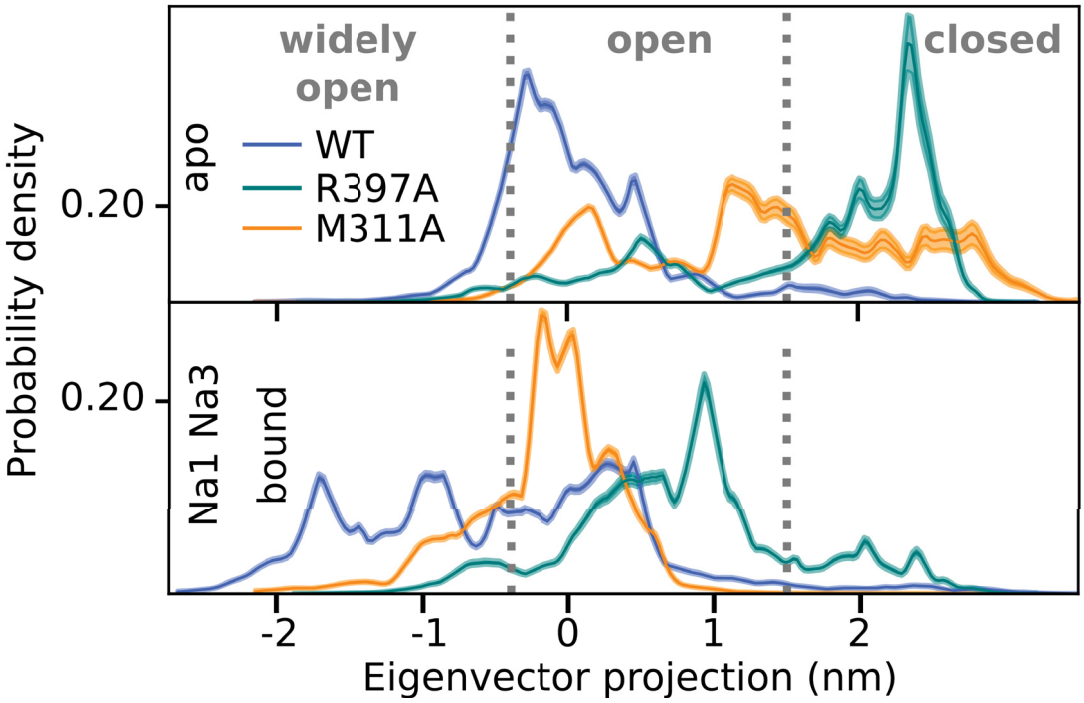

Figure 4.12: Probabilities densities of hairpin states were obtained by applying US along a single coordinate, described by projection on the first eigenvector obtained from PCA; dashed lines show the thresholds for closed, open and widely open states.

The comparison between the probability profiles for WT, M311A, R397 in apo conditions (Table 4.4) reveals a different equilibrium of the hairpin between WT and the mutants: while in the first the highest probability state is the open one, in R397A we find that the closed state has the highest probability and in

M311A the probability of closed and open states are comparable.

Mutation - Ligand occupation	Length (ns)	# windows	State probability closed	open	widely open
M311A - apo	195	52	43%	56%	1%
M311A - Na1 Na3	115	62	0%	76%	24%
R397A - apo	185	47	68%	29%	3%
R397A - Na1 Na3	185	49	17%	76%	7%

Table 4.4: Probabilities for different hairpin states, for wild type, R397A and M311A are calculated from the unbiased probabilities obtained from umbrella sampling and are shown in relation to Na^+ occupation.

Calculating the probabilities of the states for the Na^+ occupied state, we find a higher probability of the open state and a lower one for the widely open state for both mutants.

4.2 Experimental characterization of Glt_{Ph} mutants

4.2.1 Characterization of the F273W reporter

We set out to corroborate our MD simulation results by kinetic experiments; as discussed in the introduction, fast mixing fluorescence techniques allow to measure kinetics of rapid reactions. We start from the characterization of the single mutant, F273W, previously described in [81]. The signal, after being corrected for the fluorescence decay over time, can be fitted with a single exponential. Interestingly, plotting k_{obs} of the signal against Na^+ concentration, we notice a non linear relation: the reaction rate initially diminishes upon Na^+ increase, until a critical concentration is reached, after which the observed rate starts increasing again. This result was surprising, since previous studies [81] had defined the rate to depend linearly on Na^+ concentration. It is noteworthy to highlight that in this study the minimum concentration used was 100 mM and that the use of single or double exponential was not characterized as Na^+ dependent.

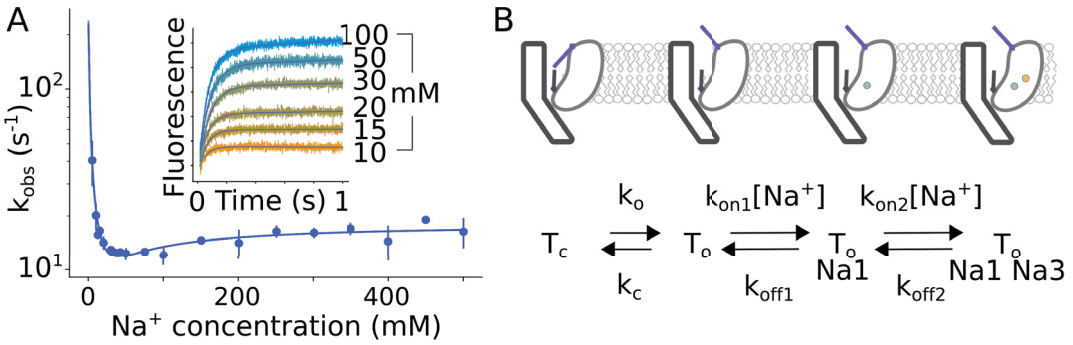

Figure 4.13: On the left, the observed rates measured in stopped-flow experiments of F273W are plotted against the Na^+ concentration applied in the experiment, while an example of the fitting of single exponentials (plus a linear correction for the decay) is shown in the inset; we fitted a four-state model of Na^+ binding to the apo transporter, as shown on the right in B: the fitted line corresponds to the overlap of 30 independent rounds of fitting using a genetic algorithm with 5000 generations.

The initial decrease would be compatible with a conformational selection mechanism, characterized by a decreasing hyperbola. This model would fit with our observation that hairpin opening precedes Na^+ binding. However, this classical description is not able to explain the increasing influorescence followed by a plateau at higher Na^+ concentrations. The simple three-states model that the classical conformational selection/induced fit dichotomy describes cannot reproduce the double dependency on Na^+ concentration. We then designed the most parsimonious model apt at describing our data, such a model must be compatible with the initial decreasing behaviour of the observed rate and contain at least two Na^+-dependent events, in order to explain the biphasic behaviour. We model the major conformational change as the hairpin opening, followed by two binding events with clearly separated binding dynamics. The dominant rate in such a model is represented by the slowest component, in our case opening of the hairpin. Interestingly, water accessibility for the reporting tryptophan decreases upon hairpin opening, suggesting a decreased dynamical quenching (Fig. 4.14)

Mutant	k_r (s^{-1})	k_e (s^{-1})	k_{on1} ($s^{-1}mM^{-1}$)	k_{off1} (s^{-1})	k_{on2} ($s^{-1}mM^{-1}$)	k_{off2} (s^{-1})
F273W	18.1±0.02	204.71±8.45	5.01±0.21	563.73±3.57	1179.04±16.05	512.13±4.94

Table 4.5: Rates and rate constants are calculated by fitting the state model in Fig. 4.13 with a genetic algorithm, the error represent the SD of 30 independent fittings. The experiments were performed at 20°C.

By fitting this model, we extract values for the rate constants compatible with hairpin opening being the slowest component of the mechanism. We conclude that Na^+ binding is characterized by a conformational selection-like mechanism, in which opening of the hairpin is followed by two kinetically distinct binding events.

4.2.2 R397A mutation modifies the reporter hydration pattern, inverting the direction of fluorescence change

In order to study the effect of R397A and M311A mutations on the coupling to hairpin opening, we designed experiments to measure the Na^+ binding/hairpin opening kinetics.

The stopped-flow technique allows us to study the reaction development at high time resolution, with only a few ms of dead time. We use a fluorescent construct, F273W, previously described [81] to report on the hairpin position, both during opening (Na^+ application) and the closure (Na^+ and substrate application).

Tryptophan fluorescence can be quenched in a dynamic or static way: dynamic quenching depends on the number of collisions with other molecules, while static quenching is due to an unbalance of charges on the aromatic ring due to bonds or interactions with surrounding species.

Our simulations suggest that dynamic quenching is responsible for most of the quenching effect observed in our experiments, since the tryptophan sits very close to the lipid bilayer and has a negligible set of interactions with surrounding molecules. Hence, we concentrated on water accessibility to F273W. We calculated radial distribution function to assess the accessibility of water to the aromatic ring. The analysis of the closed state showed almost no water acces-

sibility up to a radius of 6 Å, with then linear increase. On the other hand, in the open state, the first peak in accessibility appears at 3.5 Å, while the major peak is at 5 Å. These results demonstrate that opening of the hairpin increases the water accessibility to the tryptophan, possibly quenching the fluorescence of the tryptophan.

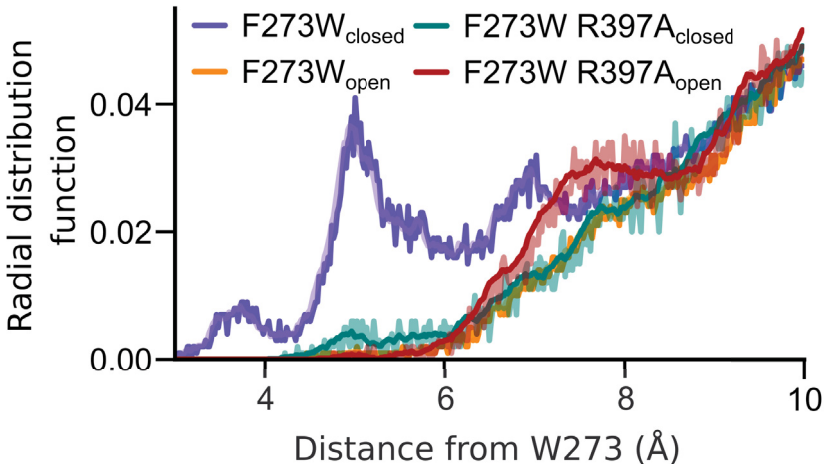

Figure 4.14: The water accessibility to the fluorescent reporter, is shown here through radial distribution function (RDF) of water around the tryptophan 273, for the wild type and R397A mutant, in open (orange for WT, teal for R397A) and closed states (purple for WT, black for R397A); this analysis shows that, in the wild type case, opening of the hairpin dramatically decreases the RDF, while in the case of the R397A mutant, the closed hairpin has a similar accessibility to water as the open hairpin in the WT case, while the opening of the hairpin in this case increases the water accessibility for values higher than 0.6 nm.

The R397A mutant induces an inversion of the signal, which persists both upon Na^+ application and aspartate application, which argues in favor of an unchanged mechanism on the reporting and a modification on the quenching effect (Fig. 4.15).

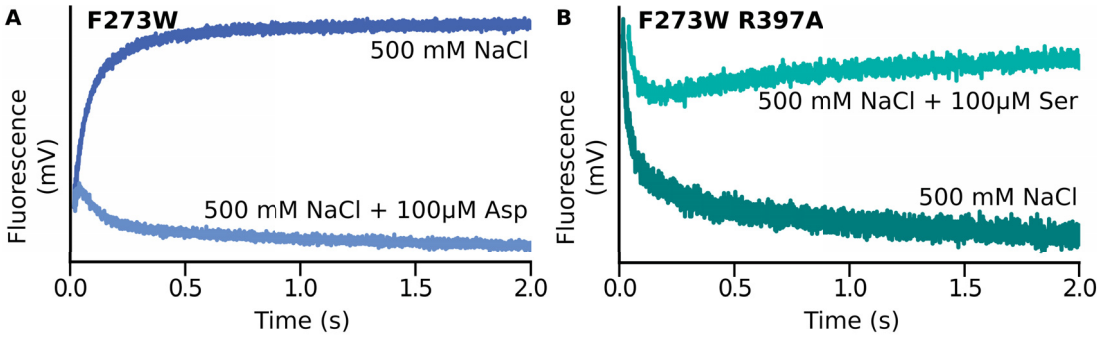

Figure 4.15: Fluorescence changes upon rapid Na^+ addition at 10°C show that the response of F273W-R397A is inverted compared with F273W; simultaneous Na^+ and substrate addition in F273W and F273W-R397A GltPh has the same opposite signal to the only Na^+ addition.

In order to address this inversion we measured the radial distribution function for water in the open and closed states for the mutant. As shown in Fig. 4.14, the closed state graph for R397 resembles the open state in the WT, with a small signal before 6 Å. The open state, on the other hand, presents an additional first peak at 7.5 Å, but almost no signal before 6 Å. These results, taken together, suggest that the open state in the mutant increases the water accessibility, and thus the quenching, instead of decreasing it.

4.2.3 R397A and M311A mutations modify Na^+ binding kinetics

We then proceed in investigating the effect of the mutants that we previously described in simulations of Glt_{Ph}, M311A and R397A, on the hairpin dynamics.

We measured F273W M311A and F273W R397A reaction at 10°C. The observed rates are shown in Fig. 4.16 for F273W, F273W M311A and F273W R397A.

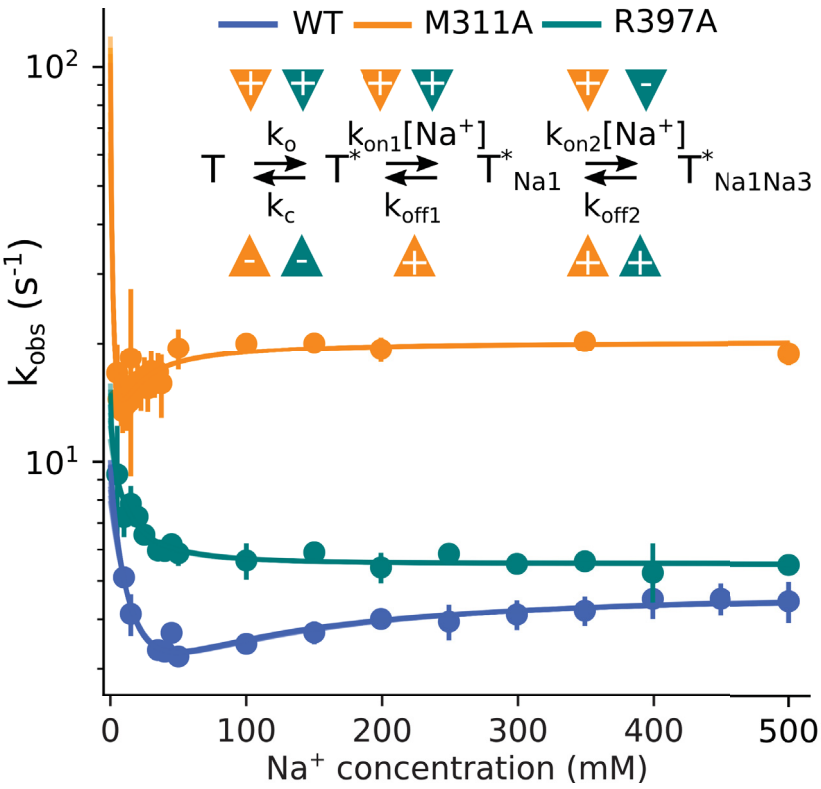

Figure 4.16: The fittings shown are obtained by fitting the state model in Fig. 4.13 with a genetic algorithm, the error represent the SD of 30 independent fittings. The experiments were performed at 10°C.

For R397A the value for k_r, k_{on1} and k_{off1} are similar to the F273W, while the rates forhairpin closure (k_e) and the second binding event (k_{on2} and k_{off2}) are both slowed down (Table 4.6). This observation let us speculate that the second event might be abolished or uncoupled; in alternative, we might not be able to measure the rate due to a sharp increase in the rate constant. For M311A the reaction is much faster than in F273W and we extract significantly faster rates for each component of the reaction. In particular, M311A almost masks the first event, since it happens at much lower K_d s and at a much higher rate. Both mutants seem to uncouple one of the two measured event from hairpin opening/closure: M311A masks the binding event happening at low concentration, while R397A cloaks the second event.

Mutant	k_r (s^{-1})	k_e (s^{-1})	k_{on1} ($s^{-1}mM^{-1}$)	k_{off1} (s^{-1})	k_{on2} ($s^{-1}mM^{-1}$)	k_{off2} (s^{-1})
			20°C			
F273W	18.1±0.02	204.71±8.45	5.01±0.21	563.73±3.57	1179.04±16.05	512.13±4.94
			10°C			
F273W	4.85±0.04	153.16±23.66	3.06±0.48	24.95±31.36	6.02±1.57	63.31±32.63
F273W M311A	20.33±0.02	68.55±22.95	9.44±2.64	218.73±10.79	371.56±31.51	133.37±5.07
F273WR397A	5.55±0.01	15.0±15.69	4.67±2.64	28.59±19.94	1.52±0.6	149.28±51.37

Table 4.6: Rates and rate constants are calculated by fitting the state model in Fig. 4.13 with a genetic algorithm, the error represent the SD of 30 independent fittings. The experiments were performed at 10°C.

Together these results confirm that Na^+ binding to Glt_{Ph} is characterized by a conformational selection-like mechanism also in the mutants. Moreover, the mutants uncouple one of the two binding events, rendering either much faster or much slower, thus suggesting how the previously described decoupling between aspartate and Na^+ binding might happen.

4.3 Computational study of EAATs

4.3.1 Coupling of K^+ to hairpin dynamics in EAATs

As detailed in the introduction, the transport cycle of the mammalian homologs, EAATs, is more complex than the cycle characterizing Glt_{Ph}, with additional coupling to co-transport of one proton and counter-transport of one K^+. The recent resolution of an EAAT crystal structure [20] allowed us to explore the role of hairpin dynamics in the coupling to these additional ligands. We then applied Umbrella sampling simulations to the EAAT crystal and we analysed the hairpin dynamics of the crystal in absence of any ligands.

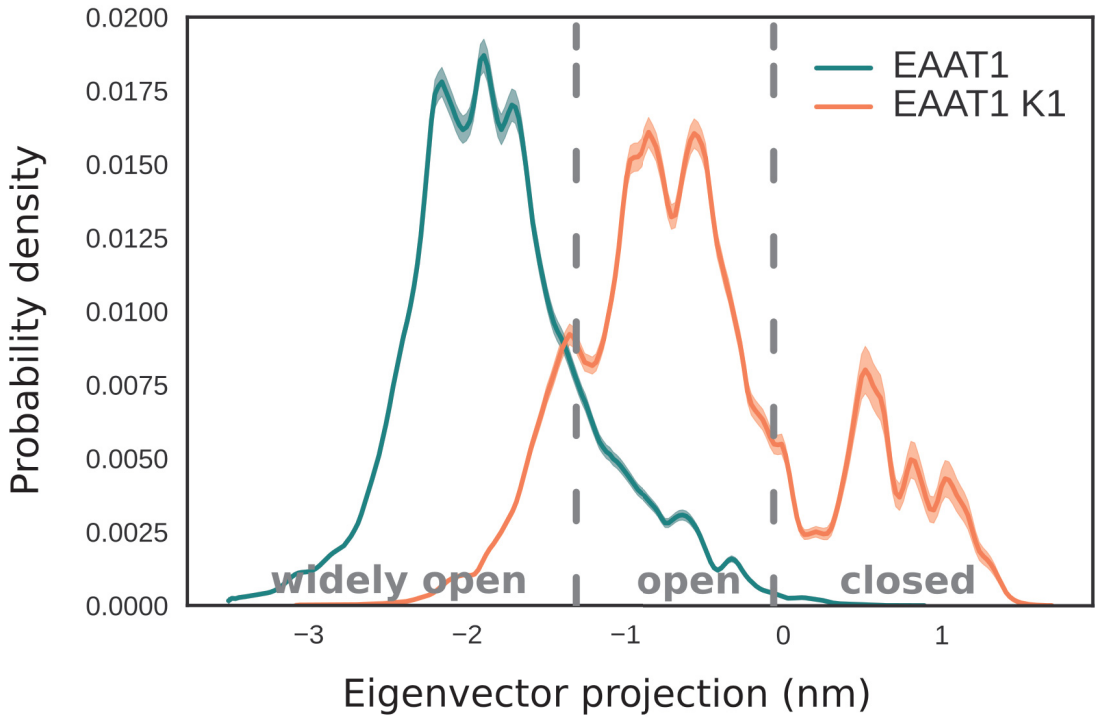

Figure 4.17: The probability density of hairpin distance for the apo and the K1-bound states for EAAT1, are shown; the coordinate of interest is expressed as eigenvector projection; dashed lines show the thresholds for closed, open and widely open states.

The result was rather surprising: in contrast to the results for Glt_{Ph} (Fig. 4.9), which presents a probability of being closed of ∼22%, here the probability is strongly skewed to the open state, with a probability for the closed state of less than 1% (Fig. 4.17).

Addition of K^+ in a position similar to the previously described Na1 [43], induces a complete shift of the probability. The probability of the open state strongly increases, reaching ∼30% (Fig. 4.17). While the percentage for apo Glt_{Ph} and K^+-bound EAAT are very similar, the shape of the distribution differs significantly: Glt_{Ph} is very flexible, with a much broader distribution and with lower energetic barriers, while EAATs states appear much more rigidly defined, with little connection between states. Another noteworthy point is represented by the presence of an additional peak at intermediate states of the hairpin, with a tip to tip distance of ∼8–10 Å. While this additional peak can be observed in the apo state in Glt_{Ph}, it only appears in the K^+ bound state in EAAT. This observation

might advocate for a, so far undescribed, role of this state in the transport cycle. Interestingly, EAATs have a much higher rate of transport, despite the similar hairpin distributions in the transport-viable states. This suggests that, while hairpin closure is fundamental to allow the transporter to undergo translocation, a following step is needed and is responsible for the overall rate of the transport.

4.3.2 Disruption of a salt bridge reduces K^+ coupling

The Na1 site and the site proposed in [43], are composed of highly conserved residues, with all of the residues of both sides conserved among EAATs and in Glt_{Ph}, while also the observed coordination appears to be really similar. In order to gain information about the differences between EAATs and Glt_{Ph} in their coupling to K^+, we investigated the residues which belong to the binding pocket or HP2. A correlation between HP2 position and distance between two residues was revealed by our analysis, specifically the residues R459 and E386 (corresponding to R397 and Q318 in Glt_{Ph}) form a salt bridge in the open state, but are oriented in position which does not allow for the formation of the interaction in the closed state (Fig. 4.18).

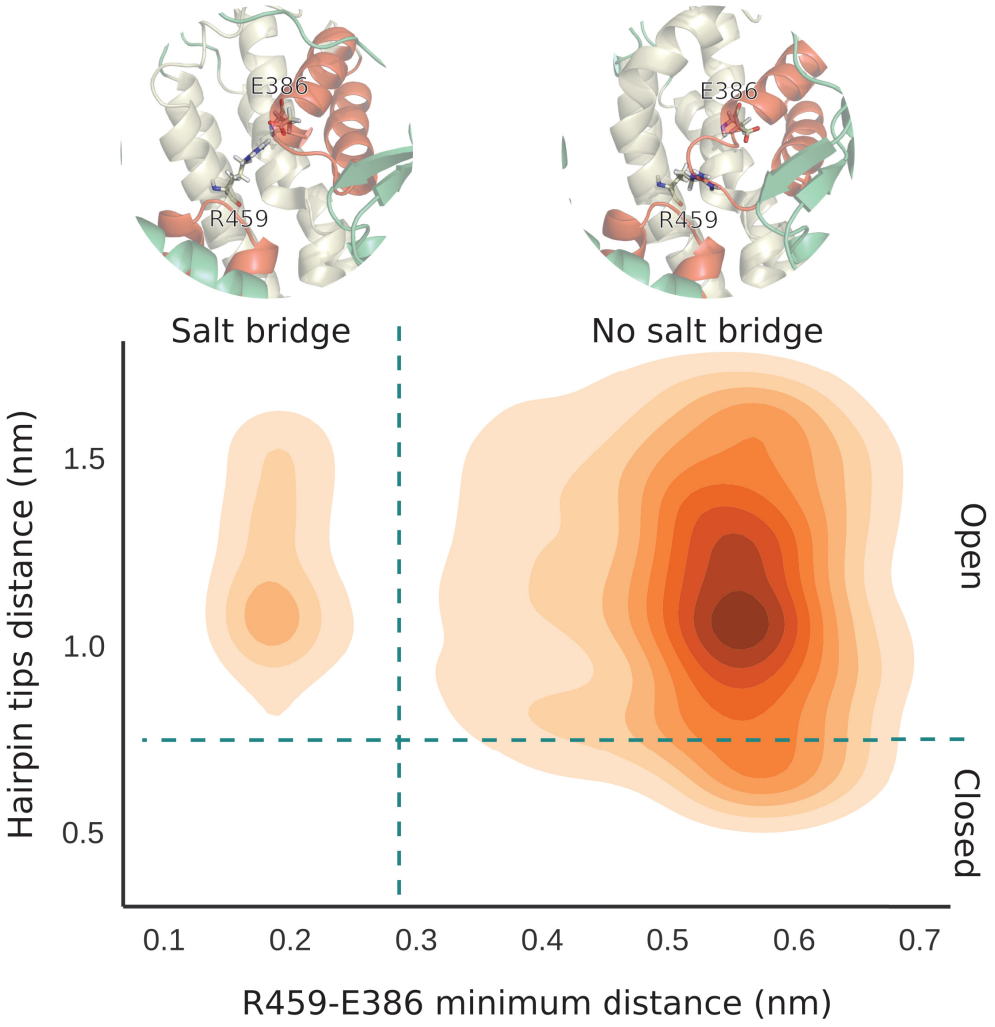

Figure 4.18: The salt bridge formed by R459 and E386 (shown in the upper magnification, with closed state on the open and closed state on the right), is correlated to hairpin distance, as demonstrated by the 2D Kernel density plot between hairpin distance and minimum R459-E386 distance; the residue numbering corresponds to the crystal structure numbering in [20].

Interestingly, R459 is the same highly conserved arginine which warrants coupling between Na^+ binding and HP2 open state stabilization; on the other hand, E386, is conserved in all EAATs, but it is a glutamine in Glt_{Ph}, The same E386 residue was proposed as being the H^+ acceptor that ensures the proton coupling. These observations hinted to a role of this salt bridge in the equilibrium of hairpin dynamics.

The mutations we chose are modifying different characteristic of this last residue: E386D decreases the steric hindrance, while maintaining the negative charge, E386Q on the other hand modifies the charge, while maintaining a similar steric hindrance. The profile for E386Q in absence of K^+ shows a significant probability for the closed state.

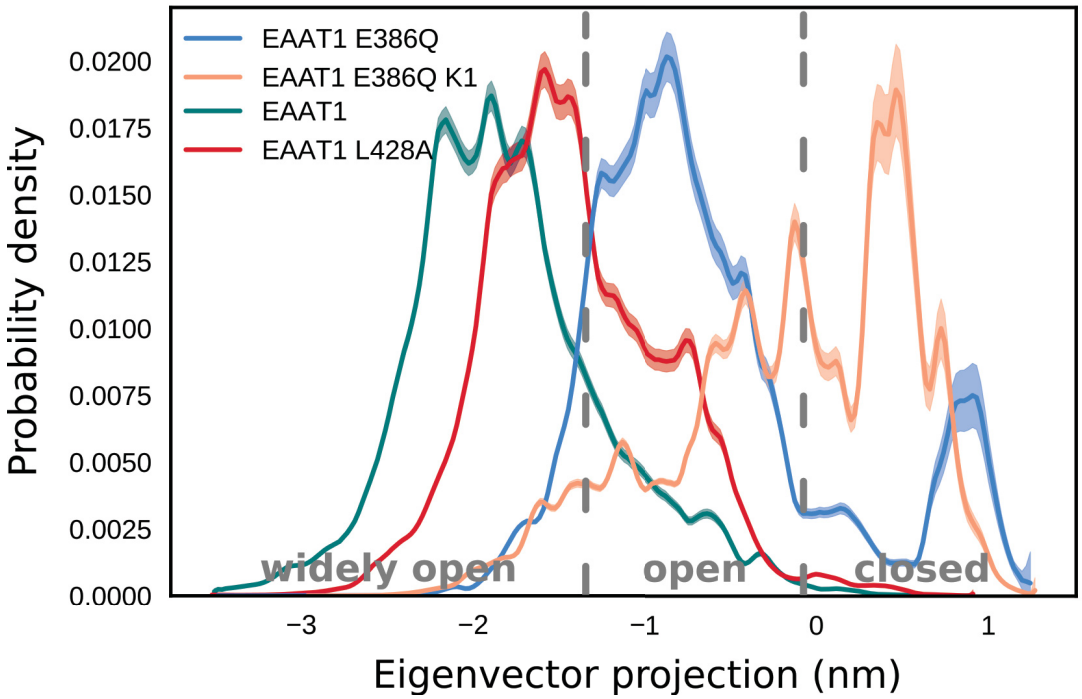

Figure 4.19: The probability density of hairpin distance for mutations disrupting the R459-E386 salt bridge (E386D, E386Q and E386Q K1-bound) are shown together with the wild type structure, dashed lines show the thresholds for closed, open and widely open states.

We can observe how mutating E386 residue to an aspartate, has the biggest effect observed: the probability is completely shifted toward the closed state, with the global peak of the distribution also in the closed state. These results are rather surprising, since the open state is almost not represented in this mutant.

The other interesting point to notice is that the peak of the distribution still lies in the open state, with a high overlap between the WT and the E386Q mutant and with an increase in probability for the closed state. Adding K^+ shifts more the probability toward the closed state, with a shift of the mean of the distribution

towards the intermediately open state.

We can observe how the role of E386 is not only governed by the interaction with R459, but actually involves the charge, the steric hindrance and probably the formation of different bonds in different hairpin states.

We then mutated R459 in an alanine, reducing its steric hindrance and de facto avoiding the formation of the salt bridge.

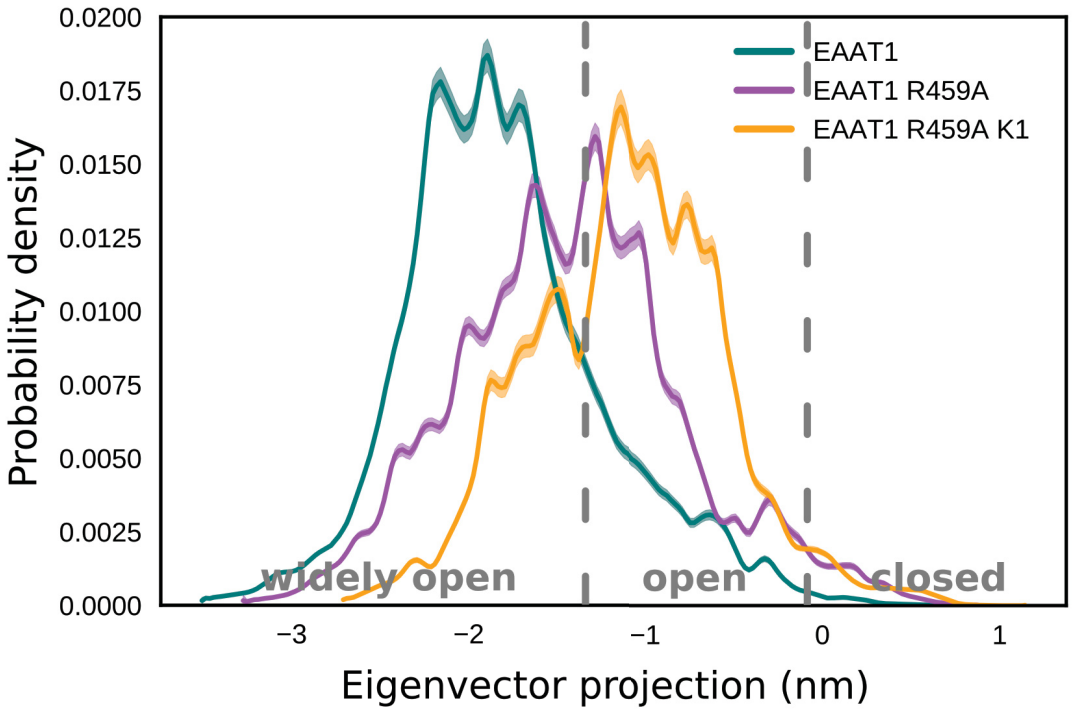

Figure 4.20: The probability density of hairpin distance for the R459A mutation, disrupting the R459-E386 salt bridge, in the apo state and in presence oF K1 are shown together with the wild type structure, dashed lines show the thresholds for closed, open and widely open states.

As shown in Fig. 4.20, the presence of the mutation shifts the peak of the probability for hairpin state toward an intermediate state (corresponding to ∼8–10 Å), with a strong decrease of the fully-open state probability. More surprisingly, a small peak appear on the closed state, which represents a probability of ∼3% of the hairpin to be closed in the apo state. Adding K^+ only doubles the probability for hairpin to be closed, while increasing even further the probability for the partially open state.

E386Q behavior differs from the R459A, which does increase the probability of the closed state, but also shifts the peak towards the intermediate state. Since E386Q and R459A have the same effect on the formation of the salt bridge, the increased effect of E386Q points to an additional function carried out by E386. These results point to the possibility that E386 participates in more than one mechanism regulating the hairpin dynamics, not all of them depending on the charge of this residue.

Comparing Figs. 4.17, 4.20 and 4.19, we notice that not only the effect of K^+ is relatively smaller in the mutants than in WT, but also the absolute effect of K^+ decreases in R459A and E386Q compared to WT, indicating that part of the coupling potential is lost when mutating this residue. These results suggest that the disruption of the salt bridge formed by R459 and E386, at least partially, decouples hairpin dynamics from K^+ binding.

4.3.3 Modifying the stability of HP2 reduces K^+ coupling

As detailed above, R459 and E386 are both involved in additional regulations of the transport cycle, through Na^+ and H^+ coupling, respectively. Due to this involvement, differentiating between alternative coupling routes might be experimentally difficult. We then set out to analyse mutations which might impact the hairpin coupling, without being directly involved in other levels of regulations. As described in [43], L428 is an alanine in Glt_{Ph}, which confers a higher flexibility to HP2, thus possibly regulating the proportion of open/close states in different ligand occupations. These observation proposed itself as a possible candidate to be responsible for K^+ coupling, we then proceeded to study the effect of mutation L428A on hairpin equilibrium in the apo state.

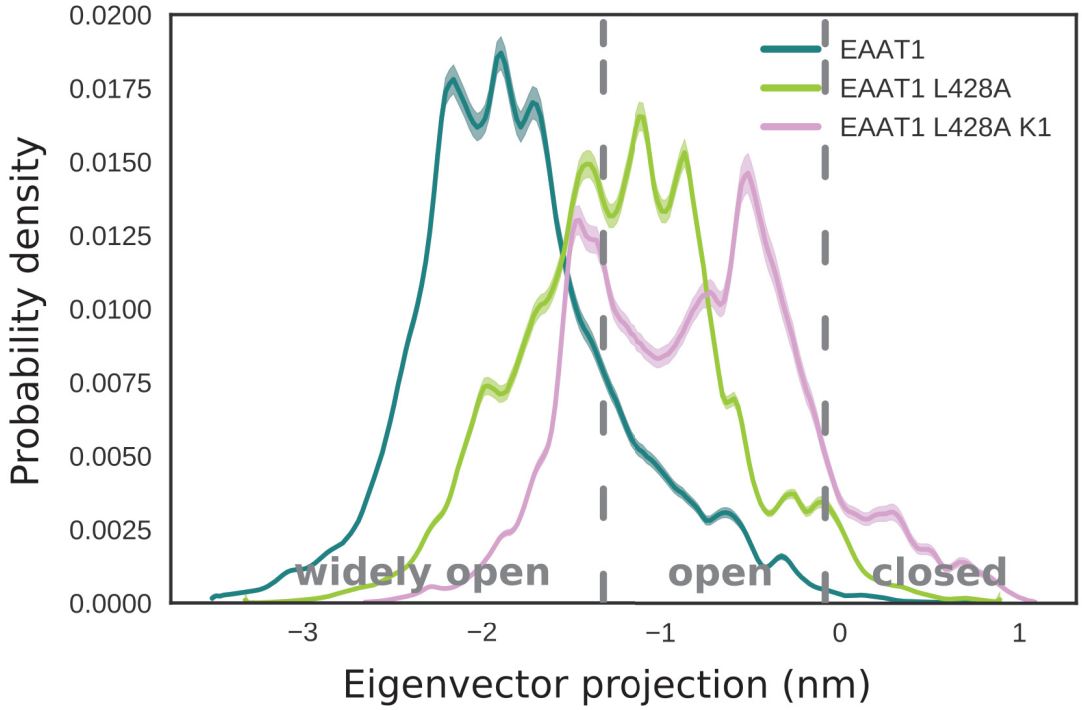

Figure 4.21: The probability density of hairpin distance for the L428A (A360 in Glt_{Ph}) mutation, which increases the flexibility of HP2, is shown in the apo state and in the K1-bound state, together with the wild type structure, dashed lines show the thresholds for closed, open and widely open states.

We see a very similar behavior of the hairpin in the L428A mutant compared to the R459A, with a shift toward the open state, a small probability for the hairpin to be completely closed, and the peak of the distribution at the half-open states. Including K^+ has a stronger effect in L428A than in R459A, with an increase of the closed state of $\sim$20 times. Still, the effect of K^+ is reduced in this mutant in respect to wild type, confirming that modifying the flexibility of the hairpin influences the coupling to K^+.

4.3.4 Investigating the protonation site in EAATs

We then moved our attention to the protonation site, which has previously been proposed to be E386. Since the disruption of the salt bridge formed between R459 and E386 modifies the hairpin dynamics, we decided to investigate E386 possible coupling to hairpin dynamics. Once more, one residue seems to be involved in

multiple regulations; thus, in order to try to untangle these different contributions, we decided to analyse different mutants (Fig. 4.19), and to protonate this site, so to obtain information on different aspects of the coupling mechanism.

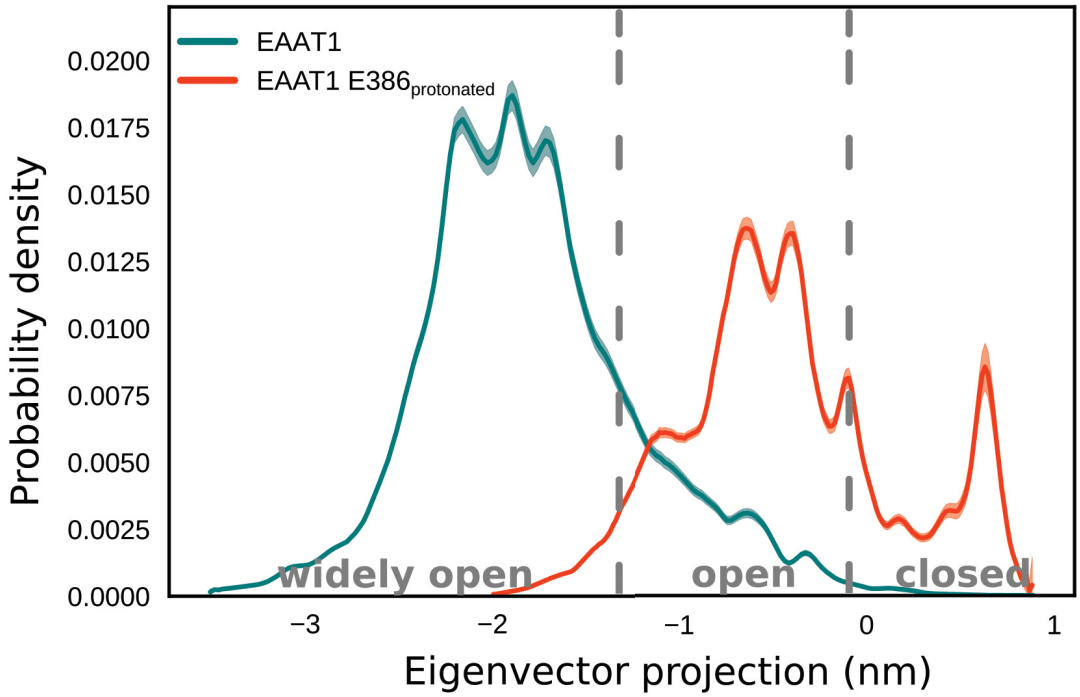

Figure 4.22: Probability of hairpin distance, expressed as eigenvector projection, for the EAAT wild type structure in protonated and deprotonated state, dashed lines show the thresholds for closed, open and widely open states.

While the E386D and E386Q mutations described in previous chapters can be reproduced in experiments, possibly validating our results, the real protonated state cannot be easily enforced during experiments. Here we can see in action one of the most powerful characteristic of computational studies, the possibility of accessing state which cannot be resolved in wet lab experiments, due to limited time or space resolution, or simply to the inability to unravel a multiplicity of states waved together. The protonated E386 probability, shown in Fig. 4.22, reveals the equilibrium between hairpin states of the apo protein in presence of H^+ which will be cotransported. The probability distribution reveals an increased likelihood of the hairpin closed state, which might be added to the contribution of the substrate (glutamate or aspartate), or of Na^+ at the Na2 site, already

known to stabilize the closed state. These results suggest that E386 protonation increases the probability of hairpin closed. The effect of protonation is bigger than that of the E386Q mutation and the E386D mutation, showing that it emerges from multiple characteristics of this residue and not simply its charge.

5. Discussion

The aim of this work was to characterize the role of the cations coupled to the amino acid transport in the complex stoichiometry that characterizes the excitatory amino acid transporters. The stoichiometry involves the co-transport of three Na^+ and one H^+ ions and the counter-transport of one K^+ ion in EAATs, while it only involves the co-transport of three Na^+ in Glt_{Ph}, as represented in the introduction chapter by Fig. 2.9. While in part the positive charges stabilize the transported negatively charged amino acid, the role played by these ions in the transport cycle must be more complex. The ASCTs belong to the same family but transport or exchange neutral amino acid. Despite these differences, the folding of ASCTs have been thought to be very similar to the EAAT fold; this hypothesis, suggested by experimental observations, has revealed to be correct when the first ASCT structure was resolved [21]. These proteins are coupled to two Na^+ only and have no coupling to H^+ or K^+, which render them electroneutral transporters. The loss of coupling to additional ions might be due to the absence of a negative charge on the transported amino acid, but if this were true, why do they retain the coupling to two Na^+? These observations, together with the difference in stoichiometry between Glt_{Ph} and EAATs despite their similar dynamics, brought us to think that the role of the cations involved might be structural more than simply due to their charges. To untangle the different contributions of the single cations we initially concentrated on the first two Na^+ to bind, since the third Na^+, which follows the substrate association, has been previously shown to stabilize the closed extracellular gate. We chose to use Glt_{Ph}, which presents only thermodynamical coupling to three Na^+. The closure of the hairpin has been characterized in a previous work from our group [60] as being described by an induced fit mechanism, which ensures both a high access rate to the binding pocket, with an initial low-specificity binding, and the recognition of the substrate itself, with the strong increase in affinity due to the correct closure of the hairpin.

Hairpin dynamics suggest a conformational selection-like mechanism

Since binding of the substrate and the last Na^+ are strictly intertwined with HP2 behavior, and it was previously demonstrated that Na^+ and the substrate

have opposite effect on the extracellular gate [26], we started our investigation inspecting the behavior of the extracellular gate in unbiased MD. We chose to use different structures in order to have a broader view on the process, we reverted all mutations, extracted anomalous ligands (Tl^+, used to model the Na^+ binding sites, and TBOA, an inhibitor able to keep the hairpin open) and, when possible, substitute them by their physiological counterpart (as in the case of Tl^+, substituted by Na^+). The analysis of the long simulations (in the order of μs) has revealed some unexpected results: while we initially expected an induced fit mechanism characterizing the Na^+-hairpin opening relationship, similarly to what was previously observed for the substrate binding [60], we observed the $Glt_{Ph,apo}$ structure visiting the widely open state, which we defined, based on the $Glt_{Ph,TBOA,Na1}$ crystal structure, as being characterize by a distance of $\sim$12 Å between S279 and G354, the two residues used to report on the distance between the two hairpins (S279 being on the tip of HP1 and G354 on the tip of HP2). This result suggests that the effect of Na^+ is to enhance a transition (closed-open) which can occur in absence of Na^+, albeit at low probabilities ($Glt_{Ph,apo}$ structure probability profile in Fig. 4.1). While this observation does not rule out completely the possibility of an induced fit mechanism, it opens the way to the alternative, the conformational selection mechanism, in which the ligand stabilizes a specific conformation rather than inducing it. A comparison between the probabilities profiles of the two closed structures (Fig. 4.1), reveals how $Glt_{Ph,2Na,substrate}$ is characterized by a mono-dispersed peak, centered at the distance value found in the crystal structure, while $Glt_{Ph,apo}$ not only has a less homogeneous distribution, but also peaks at an intermediate value of tips distance, with the helices of HP2 having the typical position of the closed state, but the flexible tip moved away from HP1 tip. This result is particularly interesting in view of the role of HP2 coupling to the substrate transport, since the tip open structure is believed to be transport incompetent. This hypothesis comes from the observation that the open tip would clash with the trimerization domain during translocation. This differential population of the tip open structure gives a hint to the mechanism by which the substrate induces the stabilization of HP2: the residues on the flexible tip directly coordinate the bound substrate, so that presence of the substrate stabilizes the closed structure. This coupling would ensure an increase in affinity to the substrate and drive the protein towards

the translocation process. Looking at the probability profiles for $Glt_{Ph,TBOA,Na1}$ and $Glt_{Ph,Na1}$ (Fig. 4.1), we notice another interesting characteristic: the two profiles only partly overlap, despite both structures can be described as open structures, due to the loss of interactions between HP1, HP2 and the transport domain. The peak of the distribution for the two are ∼4 Å away from each other, with the Na1-bound distribution spanning from 4 to 12 Å and peaking around 8 Å, and the $Glt_{-TBOA,Na1}$ distribution occupying the 8–18 Å range and peaking at 12 Å. This result indicates that the two structures have intrinsically different characteristic, which are not explained by the presence of Na1 and are slow to relax. This observation calls for different possibilities: on one side, the $Glt_{Ph,TBOA,Na1}$ structure was obtained with the addition of the inhibitor TBOA, which might have locked the protein in an unphysiological condition; on the other side, the $Glt_{Ph,Na1}$ structure was obtained with the employment of a mutation, R397A, which might itself have trapped the protein. Since the hairpin seems to move freely between all the states in $Glt_{Ph,apo}$ and given the differences just described between the two structures, we hypothesize that another contributor might be at play here. While the concept of conformational selection and induced fit are well known and easily point out the order of events in a conformational change/ligand binding diad, the actual mechanism of binding involves a series of intermediate that cannot be distinguished in experiments. An interesting study in the direction of furthering the cooperation of experiments and simulations is [91], in which multiple metastable states are resolved thanks to the application of Markov state models (MSMs). While we believe that here resides the future of the discipline, our system at the moment does not allow us to use such a technique for Na^+ binding, since the number of binding events is insufficient to have a good sampling of all states. Nevertheless, in future we could apply enhanced sampling techniques in combination with MSMs to directly extract kinetics from our simulations. Here, we use the most parsimonious model in order to describe our data and to have a coarse-grained view of the events during Na^+ binding; the model that we propose clearly shows that hairpin opens prior to Na^+ binding and thus we name it "conformational selection-like".

First Na^+ binding is preceded by a binding pocket rearrangement

In order to explain these unexpected differences, we studied the coupling of hairpin distribution and Na^+ binding dynamically. The approach we chose was to use long unbiased MD starting from the $Glt_{Ph,apo}$ structure. This structure had revealed itself to be the most flexible and should not contain any additional information on Na^+ binding. We then monitored the hairpin distance (described by the G354-S279 distance) and the minimum Na^+ distance from the Na1 and Na3 sites. The analysis of the dynamic behavior of the two revealed that Na^+ binds at the Na1 site only after hairpin opening, despite this process ranging in the hundreds of ns if not low μs range (Fig. 4.3). The open structure itself seems to be stabilized by the presence of Na^+ at Na1 (Fig. 4.3), with the hairpin distance distribution peaking around 14 Å. These results suggest that the physiological opening of the hairpin might resemble more the $Glt_{Ph,TBOA,Na1}$ than $Glt_{Ph,Na1}$. Thus either the mutation or the use of Tl^+ seem to trap the hairpin in $Glt_{Ph,Na1}$ in an intermediate state, even when these changes are reverted. This is a rather surprising finding: on one side, $Glt_{Ph,apo}$ was obtained with the same mutation and does not show the same intermediate opening preference; on the other side, Tl^+ has revealed itself to be a good model for Na^+ for the Na1 site in our simulations (where the newly bound Na^+ occupies the same site as the Na1 site described in [12]) and in crystals obtained from homologs [19, 20]). We then hypothesized that the difference in distribution might be due to an interplay between the bound Na1, the residues which coordinate it and R397. In order to test this hypothesis we analysed the collective behavior the residues around the Na1 and Na3 sites by classifying the inter-residue distances with a support vector machine, in order to correctly define four interesting functional states: apo, primed, Na1-bound and Na1/Na3-bound. A few residues (Y89, G306, N310, M311, D312, G313, T314, R397, T398, N401, and D405) are key to all of these states (Fig. 4.6). In order to pinpoint the contributors to the rearrangement, we monitored the movement of each residue participating in the Na^+ coordination and the formation and disruption of bonds. These residues initially undergo a rearrangement that opens up the Na1 binding pocket, with R397, N401 and N310 moving away to free the Na1 site and allow its hydration. After Na1 binding, M311 starts to move in the bound-like position and N401 interact with T314; finally, binding of Na3 allows for the binding pocket to acquire the primed position that will acco-

modate the substrate. From this analysis we realize that most of the coordinating residues are rather free to move, often switching between apo-like and bound-like conformations, except for N401, which only seldomly undergoes the dramatic rearrangement that brings it to the bound-like position. The movement of N401 induces a change in binding partners (Fig. 4.4), with loss of the tight interactions with G306, N310 and G313. In the apo-like position N401 partially occupies the physical space where Na1 would be positioned, while also keeping three of the coordinating residues (G306, N310 and N401) in positions unavailable for Na^+ coordination. The rearrangement of N401 opens up the space for Na^+ to bind (shown as a dotted circle in Fig. 4.4), while also exposing the backbone oxygens of N401 and G306, which will contribute to the coordination shell of Na1.

Binding at Na3 involves passage through Na1

Despite clarifying the events preceding Na^+ binding at Na1, these analysis did not solve the conundrum of the two open structures having so dramatically different behaviors in unbiased MD. We thus hypothesize that this difference might be due to an interplay not with Na^+ at the Na1 site, but perhaps Na^+ occupying the Na3 site. During μs of simulations, we could never observe binding at the Na3 site, despite the site being accessible to water. There are two probable explanations for the missing observation of this event, which explain the low probability of being observed at the time scale we can access: either the binding event is much slower than binding at Na1, or the pathway towards Na3 goes through a transient occupation of Na1. This second hypothesis has some support in our long simulations, where the coordination of Na1 changes, possibly evolving towards Na^+ hopping to the Na3 site; we name this site Na1'. A representation of this change in coordination is shown in Fig. 4.4, where the coordination found in the crystal structure is shown on the left and the Na1' coordination on the right. The contact to G306 is lost, while coordination by D312 is acquired; the new coordination induces a shift of 1 Å in ion position toward TM3. The coordinations of Na1 and Na3 both include N310 on TM7; Na1 coordination also involves G306 on TM7, N401 and D405 on TM8 and S278 on HP1; Na3 coordination includes D312 on TM7 and Y89, T92 and S93 in TM3 (as shown in Fig. 2.5). Thus, the movement away from the front of the binding pocket and toward TM3 that we can observe in Na1' could be an intermediate state between

Na1 and Na3, which might spontaneously evolve in Na3 over time or be knocked into the Na3 position by the arrival of a second Na^+ ion.

Occupation of both Na^+ sites ensures the open state in equilibrium

Due to the intrinsic slowness of the Na^+ binding and coupled events, we could not depend solely on the observations done in unbiased MD to extract a comprehensive effect of occupation of the different Na^+ sites on hairpin dynamics. Thus, we resorted again to principal component analysis (PCA) to extract a collective description of the changes between the closed and open structures, to then use to speed up the changes. The first eigenvector resulted from the PCA run on the backbone of the transport domain of free MD trajectories, described the complex events happening during hairpin opening with a single coordinate (Fig. 4.7). The umbrella sampling technique then allowed us to get the unbiased probabilities along the whole space of the coordinate of interest, for different Na^+ occupations. The probability density for $Glt_{Ph,apo}$ (Fig. 4.9) confirms that open and closed states are both accessible; surprisingly, though, the equilibrium is opposite to what we observed in the probability profiles obtained by unbiased MD, with a higher probability of the open state than closed state (Table 4.2). These results assist our hypothesis that Na^+ binds to an already open transporter; it is not particularly surprising that the two estimation differ quantitatively, since the sampling in unbiased MD is not exhaustive. In the Na1-bound profile the probabilities of the closed and open states decrease, while the probability of the widely open state increases. Na3 presents an interesting case, since it increases the probability of the closed state. This result is particularly surprising since Na^+ was previously correlated with opening of the hairpin [26]. This contradiction can be solved looking at the double occupied Na1 Na3 probability profile, where the probabilities for the open and widely open state are in equilibrium, with the widely open state probability increased of almost 5-folds in comparison to $Glt_{Ph,apo}$. Thus the results can be reconciled to the opposite hairpin movements induced by Na^+ and the substrate, since the low probability of the Na3-only bound state means that the overall effect will be the opening of the hairpin.

The native open state resembles the inhibitor-bound structure

In order to address the question of the true Na^+-bound structure, we analysed a newly obtained crystal structure, resolved by our collaborators [22]. This structure presents all the characteristics that we would like to investigate: it is a wild type structure and it has a high enough resolution (2.7 Å) to assign Na^+ positions without using the Tl^+ soaking technique. In order to compare this structure with previously obtained structures, we superimposed it to the Glt_{Na1} and the $Glt_{TBOA,Na1}$ structures (Fig. 4.11). Despite the mutations in the previously published structures, the overall folding overlaps well with the new crystal structure, with an RMSD of 1.0 Å. As for the hairpin position it overlaps perfectly with $Glt_{TBOA,Na1}$ while it does not overlap with Glt_{Na1}. This observation confirms our previous results, solving the true Na^+-bound structure dilemma. The second point to highlight is the presence, for the first time in a Glt_{Ph} structure, of true Na^+ densities and not Tl^+-based densities. It does not present only the Na1 but also the Na3 density, which was previously only resolved in Glt_{Tk} (Fig. 4.10). Additionally, we run unbiased MD, on $Glt_{Ph,Na1}$ where we added the occupation of Na3, and on $Glt_{Ph,Na1,Na3}$, with the addition of the R397A mutation, in order to try and revert this intrinsic difference between the two structures (Fig. 4.11). The results obtained show that both the absence of Na^+ at the Na3 site and the presence of the R397A mutation trap $Glt_{Ph,Na1}$ in a not completely open state. This result is particularly surprising since $Glt_{Ph,TBOA,Na1}$ was also simulated in presence of only Na1, and its distribution overlaps well with the one from $Glt_{Ph,Na1,Na3}$. Moreover, the R397A mutation was reverted in $Glt_{Ph,Na1}$ at the start of the simulations, so that we would expect that the protein would relax in its native state. These results together show that the hairpin state depends on a process related with the R397 residue and occupation of both Na^+ sites, characterized by a slow equilibrium. Finally, this crystal structure confirms the positions of Na1 and Na3 and the residues coordinating them in Glt_{Ph}.

Kinetic experiments confirm the conformational selection-like model

We then moved on to characterize two mutants previously shown to modify the Na^+-hairpin coupling, M311A and R397A. These mutations involve residues which undergo stark rearrangements upon Na^+ binding, and are thus good candidates to be key actors in the coupling. Despite the different role previously

described for these residues (substrate coordination and recognition for R397 and Na2 and stabilization of the open hairpin for M311), they have a very similar effect on the protein in absence of Na^+: both mutations pull the distribution towards the closed state, similarly to the effect of Na3 addition. The widely open state seems unaccessible to both mutants. Interestingly, looking at the Na^+-bound conformations, we see a stabilization of the open state in both mutants, with only residual occupation of the widely open state in R397A and a reduced occupation of the state in M311A. Both mutants mimic Na3 occupation, thus either speeding up the Na^+ binding process or abolishing the need for one of the two Na^+ binding that precede substrate association. These possibilities cannot be addressed with the same computational techniques we used so far, so we decided to approach the characterization of the two mutants from an experimental point of view. We decided to use the stopped flow technique to obtain high temporal resolution, using a fluorescent reporter, known to report on Na^+ binding through hairpin opening [81]. In order to characterize the behavior of the reporter, we applied increasing concentrations of Na^+ at high flow rates. Our system allows us to apply smaller concentration than previously measured [81]; this higher resolution at lower concentration brought us to observe some surprising results: while the rate was reported to linearly increase with Na^+ concentration in [81], we observed an initial decrease, followed by an increase and then a plateau (Fig. 4.13). The model classically used to describe the relationship between rate and ligand concentration, predicts a linear increase of rate, with the y-intercept describing k_{off} and the slope of the linear fit corresponding to k_{on}. This model clearly does not describe properly our data. We thus took into account two possible explanations for the observed behavior: on one side, the results might indicate the presence of two independent events; on the other side, they might indicate a conformational selection-like mechanism characterized by two processes. While the conformatonal selection mechanism would be compatible with the decreasing behavior at low concentrations, it cannot account for the increase at higher concentration. We then design the most parsimonious model that would describe our data, where opening of the hairpin is followed by two binding events.

Mutations of key residues uncouple Na^+ binding from hairpin dynamics

The R397A mutant presented a decrease in fluorescence upon Na^+ application, which could be explained alternatively as an inversion of the signal, due to a modified behavior of the reporter, or by Na^+ opening the hairpin instead of closing it. In order to address these possibilities we needed to characterize how the environment around the tryptophan reporter is influenced by Na^+ addition, and if the signal is reverted also in the case of application of the substrate in presence of Na^+, which closes the hairpin. To address the first matter we run MD simulations of $Glt_{Ph,apo}$ F273W and $Glt_{Ph,Na1}$F273W, and subsequently of the double mutant F273W R397A. We then analysed the radial distribution function (RDF) of water molecules around the tryptophan, both in the open and closed states, in order to obtain the density of water molecules around the reporter and to extract its water accessibility. Our results (Fig. 4.14) indicate that the reporter has a high water accessibility when the hairpin is closed (with three major peaks of density at ∼3.5, 5 and 7 Å), and that it decreases when the hairpin opens (where all of the peaks have disappeared). In the case of the double mutant, the closed state resembles the single mutant open state, with only a shallow peak at 5 Å. Interestingly, the open state present a peak at ∼7 Å, indicating that the reporter has a higher accessibility to water in the open state than in the closed one, suggesting that the signal might be inverted. These results are in agreement with our observations. We decided to test the second possibility as well, despite the strong indication received from the MD simulations. As previously said, the R397A has been long known to decrease dramatically the affinity to negatively charged substrates, so that we could not observe any effect on the fluorescence up to 100 μM of aspartate applied. To circumvent this problem, we decided to apply serine, a neutral amino acid known to be transported by the R397A mutant (Fig. 4.15); since transport requires the closure of the hairpin, we can expect that it will have a similar effect in R397A as aspartate in the wild-type. Addition of serine induces a concentration-dependent increase in fluorescence in F273W R397A, thus confirming that closure of the hairpin has opposite read-outs for F273W and the double mutant. These observations raise some interesting questions, since they highlight that opening of the hairpin not only increases the water accessibility for part of the protein, namely the binding pocket, but also decreases the accessibility of other parts. It is possible that hydration and

de-hydration energies play a role in driving the transport cycle.

Looking at the rates obtained for the mutants (Fig. 4.16), it is apparent that, despite the similar effect of the mutant on the free energy landscape of the apo protein, the two mutants have very different effects on the kinetics of Na^+ association and hairpin opening. In particular, M311A seems to speed up the reaction by 4-fold while R397A slows it down by half. The shape of the relation k_{obs}-[Na^+] also differs for the mutants in comparison to F273W, in particular M311A recapitulates the high concentration part of wild type, while R397A has the descending behavior of the wild type low concentration. The similarity between some R397A rate constants (k_r, k_e, k_{on1} and k_{off1}) brings us to speculate that what we observe in R397A corresponds actually to the first binding event of F273W, so that the second one appears invisible. This observation could be due to two possibilities: either one of the two binding events is completely abolished by the R397A, or it is either sped up of slowed down enough to bring it out of our time resolution. This hypothesis could be tested by measuring the protein at high and low temperatures, in order to observe this invisible event. For M311A we can draw a similar picture, in which the mutation partially masks the first event due to the low concentration and to the high rate at which it happens. Leaving aside the precise mechanism of how these mutations actually mask one specific event, both of them uncouple one of the bindings from hairpin dynamics.

K^+ binding is also coupled to hairpin dynamics

Since the coupling between Na^+ and hairpin opening is so strict, we speculated that a similar mechanism might describe the coupling to other cations. As introduced previously, Glt_{Ph} is known to be thermodynamically coupled only to Na^+, so we needed to switch models in order to investigate the mechanism of coupling for other cations. The recently resolved EAAT1 structure allowed for further investigation. Recent work in our group [43] has proposed a site for K^+ binding, which not only is coordinated by residues previously demonstrated to disrupt K^+ when mutated, but is also able to justify the gating charges measured in patch clamp experiments. We then study the effect of occupation of this site on the free energy landscape of EAAT1. The first remarkable observation is that the $EAAT1_{apo}$ free energy landscape is remarkably similar to the $Glt_{Ph,apo}$ (Figs. 4.9 and 4.17), with the main peak on the open state. The eigenvectors

used were calculated independently, even though on similar data sets; due to its inner working, the PCA technique can only extract information already present in the data; the only selection factor we used was the correlation with hairpin opening/closure, it would be sensible to expect that the descriptor could capture different contributors. Yet, despite the intrinsic difference between the two collective variables, the calculated free energies resemble each other well. This observation strengthens our argument, since the protein behavior that we observe does not seem to strictly depend on the collective variable used to analyse it. The difference between the two energy landscapes appears when looking at the accessibility of the closed state, which is populated in Glt_{Ph} but not in EAAT1. As argued in [43], this key difference between the two homologs might be responsible for the re-translocation conundrum, for which Glt_{Ph} is able to re-translocate in the apo state, while EAAT1 is confined to K^+-bound re-translocation. If the state of the hairpin is a pivotal control for access to the translocation process, as we argue here and as was previously hypothesized [16], then the accessibility of the closed state in the apo condition stands to be the fundamental determinant of the competence for re-translocation. In light of this hypothesis, testing the effect of K^+ on the free energy landscape of the hairpin becomes even more pressing. Confirming our hypothesis, addition of K^+ shifts the probability of hairpin state toward the closed state (Fig. 4.17). Interestingly, addition of Na3 in Glt_{Ph} has a similar effect to K^+ addition in EAAT1; on one hand, our data indicate that occupation of Na3 is a rare if not inaccessible state, so that the hypothesis that Na^+ has the effect of opening the hairpin still stands, on the other hand, it is interesting to speculate how some characteristic of the Na3-bound recapitulate the effect of K^+. The first point to keep in mind is that the K^+ site described in [43] sits in between the Na1 and Na3 sites (Fig. 2.7), in a position very similar to the Na1' position described in Fig. 4.4. We can speculate that some of the residue that coordinate the K^+ ion assume the same position in the Na3-bound Glt_{Ph}, possibly reproducing the TM7 shift described in [43], thus inducing a more compact state of the binding pocket, which stabilizes the closed state of the hairpin. This observation suggests for the possibility that Glt_{Ph} is actually able to bind K^+ and that it would have a similar effect to the one it has in EAAT1. This hypothesis was confirmed in [43], where we show that Glt_{Ph} indeed binds K^+, albeit with very low affinity and that it induces an increase in the intermediate

state, typical of the re-translocation mechanism; we show that addition of K^+ does shift the probability of hairpin states toward the closed state. There are two noteworthy differences between EAAT1 and Glt_{Ph} around HP2, which could be good candidates to be modified, in order to measure the subsequent similarity between homologs: the presence of a leucine in EAAT1, substituted in Glt_{Ph} by an alanine, on HP2, which possibly destabilized the closed state; and one glutamate residue in EAAT1 which is substituted by a glutamine in the archeael homolog. We initially decided to modify the stability of HP2, in order to gage its role on the equilibrium between open and closed states. The L447A mutation, which mimics the homolog A360 in Glt_{Ph}, shifts the probability profile of hairpin states towards the closed state, peaking at an intermediate state and increasing the accessibility of the closed state, while addition of K^+ has a limited effect. These results suggest that the key to regulation lies not exclusively in the direct interaction between cation and cation-coordinating residues with the hairpin, but more broadly in the interplay between these and the stability of the hairpin itself. Part of the reliance on K^+ association for re-translocation competence in EAAT1 might depend on the formation of a salt bridge between R459 (corresponding to R397 in Glt_{Ph}) and E386 (corresponding to Q318). The formation of this salt bridge is state dependent since it forms only in the open state (Fig. 4.18); inducing closure of the hairpin thus is more expensive from an energetic point of view, since additional energy must be used to break the interaction. But would it be enough to disrupt this one salt bridge to allow EAAT1 to re-translocate in a K^+-independent way? We tested this hypothesis and the results were striking: any mutation we used (E386D, E386Q, R459A) which demonstrated to disrupt this interaction, has also the effect to increase the probability for the closed state of the hairpin in the apo state (Figs. 4.19 and 4.20). Despite inducing a population of the apo state, these mutations had different overall effect on the free energy landscape, indicating that the role of the mutated residue is not simply to ensure the formation of the salt bridge. This observation is not particularly surprising, since these residues are involved in multiple other functions: R459 is responsible for substrate recognition and, as we describe in this work, for coupling to Na^+, while E386 is believed to be the proton acceptor responsible for the H^+ dependence of the transport cycle. Interestingly, comparing the EAAT1 R459A (in Fig. 4.20) and Glt_{Ph} R397A (in Fig. 4.12) profiles, we notice how the

effect on the two proteins is drastically different: while in EAAT1 the mutation shifts the main peak towards more intermediate values and it slightly increases the probability for the open state, in Glt_{Ph} the same mutation completely depopulates the open state and shift completely the profile to the closed state, similarly to the effect of Na3, suggesting that R397/R459 has an important role in both homologs, but it has a much more critical role in Glt_{Ph}; in turn this suggests that EAAT1 might be more tightly regulated by a higher number of residues, while Glt_{Ph} seems to invest most of its energetic potential on a few residues. The other striking point about the R459A mutant is that addition of K^+ does not have the same dramatic effect it has in the wild type protein (Figs. 4.17 and 4.20). The small effect of K^+ addition suggests that R459 is itself a major player in the coupling to K^+, even though it is not the only one, as we see from the value of probability for the closed state. Form our results, it is unclear if the charge or the steric hindrance of this residue are responsible for the effect of the mutation; putting it in perspective looking at literature, it was previously demonstrated that disrupting either the charge or the hindrance [16, 28, 32–34] has an effect on both Na^+ coupling and substrate affinity/recognition, thus indicating that both characteristic participate in the role of this residue. For the mutations of E386 we can address the question of which residue characteristic plays the most important role, since with different mutations we observe different effects. When the charge is kept and only the possibility of forming the salt bridge is affected, with the E386D mutation, the probability profile looks very similar to the R459A profile, with the main peak slightly shifted toward the closed state and a small increase in the probability of the closed state. When keeping the steric hindrance of the residue but loosing the charge, with the E386Q mutation (which mimics Glt_{Ph}), the effect is modified, with a strong increase in the probability of the closed state and very little occupation of the open state. When adding K^+ to E386Q, the probability profile cannot be distinguished by the K^+-bound wild type, so that, while this mutant disrupts the salt bridge formation, it does not directly affect the downstream effect of K^+ binding, in contrast to R459A and L428A.

Protonation of the putative proton acceptor site induces hairpin closure

Since the probability profile in E386Q resembles much more the K^+-bound wild type than the E386D mutant, we hypothesize that also the protonation of what is believed to be the proton acceptor site, E386, might be coupled to hairpin dynamics. We then investigated the behavior of E386 in a protonated state, expecting to see a behavior similar to the one observed in E386Q. Indeed the result confirm that protonation of this residue is responsible for an increase probability of the closed state. All these results together paint a striking picture, in which each and every cation participating in the thermodynamical coupling during the transport cycle of EAATs, exercise their effect in driving the transporter towards a certain state (open for Na^+ and closed for K^+ and H^+) through HP2. Changing the equilibrium between states, independently if directly or not, such as affecting HP2 flexibility or mutating residues which interact with HP, or disrupting interaction that would stabilize the open state, seem to be the central regulation of the entire transport cycle.

6. Conclusion

Here, we show how the complex stoichiometry and highly specific function of glutamate transporters has its center of regulation in the extracellular gate. Our results show that, despite all the small moving parts that form the complex clockwork that is the glutamate transporter family transport cycle, the entire process can be guided by one switch, HP2. The transport domain, which moves across the membrane and along the trimerization domain, is highly structured, with multiple interacting helices and two helix-loop-helix hairpins. One of these hairpins, HP2, was previously described as being the extracellular gate, which controls the passage leading to the substrate binding pocket. Binding of Na^+ prior to the substrate had previously been connected to opening of the hairpin, while binding of the substrate and the third Na^+ ensures the closure of the extracellular gate. Only correct packing of the extracellular gate allows the transporter to translocate across the membrane [16], and increases the substrate affinity through an induced fit mechanism. While the stoichiometry of Glt_{Ph} only includes Na^+ and aspartate, the mammalian homologs EAATs are coupled to H^+ and K^+, too. Here we untangle the relationship between Na^+ and HP2, by showing that both Na1 and Na3 are needed to guarantee the opening of the extracellular gate, and that this strict coupling can be disrupted by mutating key residues, namely R397 and M311. We thus demonstrate that HP2 is responsible for previously described complex interactions, such as ligand recognition and Na^+ coupling. We show how HP2 dynamics govern the kinetic fingerprint of Na^+ binding. In our work on EAAT1 we show that also K^+ binding is coupled to hairpin closure, so that the K^+ site occupation ensures the possibility of re-translocation in absence of Na^+ and glutamate. We show how the protonation of E386, the proposed proton acceptor site, increases drastically the probability of the closed state of the hairpin, thus confirming the role of this residue and showing a coupling to hairpin dynamics for this ligand, too. We further show how the coupling between ligands and HP2 can be modified by single point mutations, which can act on different parts of this coupling: directly affecting the flexibility of the hairpin (L428A), disrupting or stabilizing bonds which favor one state (R459A, R397A, E386D and E386Q), or influencing residue rearrangements upon ligand binding (R397A and M311A). Thus single point mutations in the transport domain of this protein might have

dramatic effects on the protein behavior, thus proposing them as responsible for pathological phenotypes. We also directly show how not only the equilibrium between states is shifted but also the rates of conformational changes upon Na^+ binding are influenced. With these results, we establish that each one of the ligands involved in the stoichiometry and the kinetics of both Glt_{Ph} and EAATs, is directly involved in regulating the dynamic state of the hairpin, which in turns regulates the crossing of the membrane. We conclude that HP2 is the principal actor of the regulation network that governs glutamate transporters behavior, thus it plays a fundamental role in ensuring secondary active transport.

7. Outlook

This work opens brand new questions which will need to be addressed in the future. The binding process could be modeled by a five-state model, in order to extract the binding rate constants, possibly characterizing the difference in the mechanisms activated in low and high Na^+ concentrations without inferring any characteristic of the model. Adding the modeling of the substrate and the last Na^+ will then allow to weigh the contribution of the conformational selection-like and the induced fit branches of the binding process. In addition, characterizing the speed of re-translocation will then complete the portrait of glutamate transport cycle. As for EAATs, further characterization will be needed for binding of H^+ and its role in translocation, a problem which was neglected due to the difficulty of studying protonation both from the computational and experimental sides. The involvement of the hairpin in this process and the possibility to directly observe its behavior, has opened up a new avenue to study this complex coupling. Finally, modeling the translocation across the membrane and coupling this to the binding model will conclude the characterization of the entire transport cycle.

8. Publications

During my PhD I contributed to the peer-reviewed publications:

–Kortzak, D., **Alleva, C.**, Weyand, I., Ewers, D., Zimmermann, M.I., Franzen, A., Machtens, J.P. and Fahlke, C., 2019. Allosteric gate modulation confers K+ coupling in glutamate transporters. The EMBO journal, 38(19), p.e101468.

Author's contribution statement: I performed the Umbrella sampling on Glt_{Ph} and EAAT1, contributed to the analysis of the data, the design of simulations and figures and the writing of the text.

–Machtens, J.P., Briones, R., **Alleva, C.**, de Groot, B.L. and Fahlke, C., 2017. Gating charge calculations by computational electrophysiology simulations. Biophysical journal, 112(7), pp.1396-1405. Author's contribution statement: I contributed simulations for the calculation of gating charges for Glt_{Ph}.

Author's contribution statement: I contributed simulations used to calculate Glt_{Ph} gating charges.

The following work resulted from my PhD is under revision in a peer-reviewed journal:

–**Alleva, C.**, Kovalev, K., Astashkin, R., Berndt, M. I., Baeken, C., Balandin, T., Gordeliy, V., Fahlke, Ch. and Machtens, J.-P., 2020. Na^+-dependent gate dynamics and electrostatic attraction ensure Na^+ –substrate coupling in glutamate transporters. *In revision.*

Author's contribution statement: I designed the study together with ChF and JPM, performed simulations and experiments (except for the crystallization), prepared the figures and wrote the text together with ChF and JPM.

9. Acknowledgements

I could have not done anything of what I did if I were not surrounded by some pretty amazing humans (and one object), so brace yourself for a long list of acknowledgements for the people that with small and big actions have nudged me along my way.

Prof. Fahlke, who gave me the opportunity to work in his group and pushed me to move forward when needed.

Jan-Philipp, the Virgilio to my Dante, showed me the way when I was lost (admittedly some times it was the looong way, but still).

Prof. Gohlke, who showed me how useful (and some times painful) reviews are, it's a lesson that I treasure.

Dr. Kortzak, for the discussions on mathematical models and EAATs function.

Anita, without whom I would still be in my personal bureaucratic hell.

Meike, that was always there when I needed a hand or a chat (or to deal with liquid nitrogen).

To the soon to be Drs. Yulia and Andrei, because when you need stable ground under your feet or to have a laugh at scientific silliness you can always count on them.

The coffee group, which I will name in alphabetical order one by one, forgetting someone and being then thoroughly ashamed: Antonia, Katia, Felix, Bettina, Gaby, Juan, Piersilvio, Sara, Stefan, Saskia; because having quirky scientists to fill you with renewed energy, one weird conversation at a time, is underrated.

Clara and Dusan, because you need a certain kind of friendship to withstand organizing a conference together.

My poor laptop for serving me well for years and deciding to take itself out of the picture a few week from thesis submission.

E infine le persone nella mia vita che non appartengono al mondo scientifico, Antonio e i miei genitori, senza i quali avrei perso la ragione. Grazie per rispondere ancora alle mie telefonate!

And you, the reader, who read the entire thesis (hopefully), or opened directly to this page burned by curiosity, for getting here.

Bibliography

[1] Niels C Danbolt. "Glutamate uptake". In: *Progress in Neurobiology* 65.1 (Sept. 2001), pp. 1–105.

[2] Marc Laruelle, Lawrence S. Kegeles, and Anissa Abi-Dargham. "Glutamate, Dopamine, and Schizophrenia". In: *Annals of the New York Academy of Sciences* 1003.1 (Nov. 2003), pp. 138–158.

[3] Heather L Scott et al. "Aberrant expression of the glutamate transporter excitatory amino acid transporter 1 (EAAT1) in Alzheimer's disease." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 22 (3 Feb. 2002), RC206.

[4] Anthony J Groom, Terence Smith, and Lechoslaw Turski. "Multiple sclerosis and glutamate." In: *Annals of the New York Academy of Sciences* 993 (May 2003), pp. 229–75, 229–75.

[5] Jae-Hyuk Yi and Alan S. Hazell. "Excitotoxic mechanisms and the role of astrocytic glutamate transporters in traumatic brain injury". In: *Neurochemistry International* 48.5 (Apr. 2006), pp. 394–403.

[6] K. Tanaka. "Epilepsy and Exacerbation of Brain Injury in Mice Lacking the Glutamate Transporter GLT-1". In: *Science* 276.5319 (June 1997), pp. 1699–1702.

[7] Natalie Winter, Peter Kovermann, and Christoph Fahlke. "A point mutation associated with episodic ataxia 6 increases glutamate transporter anion currents." In: *Brain : a journal of neurology* 135 (Pt 11 Nov. 2012), pp. 3416–3425.

[8] Baruch I. Kanner and Annie Bendahan. "Binding order of substrates to the (sodium + potassium ion)-coupled L-glutamic acid transporter from rat brain". In: *Biochemistry* 21.24 (Nov. 1982), pp. 6327–6330.

[9] Robert J. Vandenberg and Renae M. Ryan. "Mechanisms of Glutamate Transport". In: *Physiological Reviews* 93.4 (Oct. 2013), pp. 1621–1657.

[10] N Zerangue and M P Kavanaugh. "Interaction of L-cysteine with a human excitatory amino acid transporter." In: *The Journal of physiology* 493 (Pt 2) (June 1996), pp. 419–423.

[11] Jan-Philipp Machtens et al. “Mechanisms of Anion Conduction by Coupled Glutamate Transporters”. In: *Cell* 160.3 (Jan. 2015), pp. 542–553.

[12] Olga Boudker et al. “Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter”. In: *Nature* 445.7126 (Jan. 2007), pp. 387–393.

[13] Nicolas Reyes, Christopher Ginter, and Olga Boudker. “Transport mechanism of a bacterial homologue of glutamate transporters”. In: *Nature* 462.7275 (Nov. 2009), pp. 880–885.

[14] Grégory Verdon and Olga Boudker. “Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog”. In: *Nature Structural & Molecular Biology* 19.3 (Feb. 2012), pp. 355–357.

[15] Nicolas Reyes, SeCheol Oh, and Olga Boudker. “Binding thermodynamics of a glutamate transporter homolog”. In: *Nature Structural & Molecular Biology* 20.5 (Apr. 2013), pp. 634–640.

[16] Grégory Verdon et al. “Coupled ion binding and structural transitions along the transport cycle of glutamate transporters”. In: *eLife* 3 (May 2014).

[17] Nurunisa Akyuz et al. “Transport domain unlocking sets the uptake rate of an aspartate transporter”. In: *Nature* 518.7537 (Feb. 2015), pp. 68–73.

[18] Amanda J Scopelliti et al. “Structural characterisation reveals insights into substrate recognition by the glutamine transporter ASCT2/SLC1A5”. In: *Nature Communications* 9.1 (Jan. 2018).

[19] Albert Guskov et al. “Coupled binding mechanism of three sodium ions and aspartate in the glutamate transporter homologue GltTk”. In: *Nature Communications* 7 (Nov. 2016), p. 13420.

[20] Juan C. Canul-Tec et al. “Structure and allosteric inhibition of excitatory amino acid transporter 1”. In: *Nature* 544.7651 (Apr. 2017), pp. 446–451.

[21] Alisa A. Garaeva et al. “Cryo-EM structure of the human neutral amino acid transporter ASCT2”. In: *Nature Structural & Molecular Biology* 25.6 (June 2018), pp. 515–521.

[22] Claudia Alleva et al. “Na^+-dependent gate dynamics and electrostatic attraction ensure substrate coupling in glutamate transporters”. In: *In revision* (2020).

[23] Sonja Jensen et al. “Crystal structure of a substrate-free aspartate transporter”. In: *Nature Structural & Molecular Biology* 20.10 (Sept. 2013), pp. 1224–1226.

[24] OLEG JARDETZKY. “Simple Allosteric Model for Membrane Pumps”. In: *Nature* 211.5052 (Aug. 1966), pp. 969–970.

[25] David Drew and Olga Boudker. “Shared Molecular Mechanisms of Membrane Transporters.” In: *Annual review of biochemistry* 85 (June 2016), pp. 543–572.

[26] P. J. Focke, P. Moenne-Loccoz, and H. P. Larsson. “Opposite Movement of the External Gate of a Glutamate Transporter Homolog upon Binding Cotransported Sodium Compared with Substrate”. In: *Journal of Neuroscience* 31.16 (Apr. 2011), pp. 6255–6262.

[27] L M Levy, O Warr, and D Attwell. “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.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 18 (23 Dec. 1998), pp. 9620–9628.

[28] Amanda J. Scopelliti, Renae M. Ryan, and Robert J. Vandenberg. “Molecular Determinants for Functional Differences between Alanine-Serine-Cysteine Transporter 1 and Other Glutamate Transporter Family Members”. In: *Journal of Biological Chemistry* 288.12 (Feb. 2013), pp. 8250–8257.

[29] Nechama Silverstein et al. “Molecular Determinants of Substrate Specificity in Sodium-coupled Glutamate Transporters”. In: *Journal of Biological Chemistry* 290.48 (Oct. 2015), pp. 28988–28996.

[30] Zhen Tao, Zhou Zhang, and Christof Grewer. “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”. In: *Journal of Biological Chemistry* 281.15 (Feb. 2006), pp. 10263–10272.

[31] Noa Rosental et al. “A Conserved Aspartate Residue Located at the Extracellular End of the Binding Pocket Controls Cation Interactions in Brain Glutamate Transporters”. In: *Journal of Biological Chemistry* 286.48 (Oct. 2011), pp. 41381–41390.

[32] Marcus Conradt and Wilhelm Stoffel. “Functional Analysis of the High Affinity, Na+-dependent Glutamate Transporter GLAST-1 by Site-directed Mutagenesis”. In: *Journal of Biological Chemistry* 270.42 (Oct. 1995), pp. 25207–25212.

[33] Annie Bendahan et al. “Arginine 447 Plays a Pivotal Role in Substrate Interactions in a Neuronal Glutamate Transporter”. In: *Journal of Biological Chemistry* 275.48 (Sept. 2000), pp. 37436–37442.

[34] Paul J. Focke, Alvin W. Annen, and Francis I. Valiyaveetil. “Engineering the Glutamate Transporter Homologue GltPh Using Protein Semisynthesis”. In: *Biochemistry* 54.8 (Feb. 2015), pp. 1694–1702.

[35] Y Zhang et al. “Molecular determinant of ion selectivity of a (Na+ + K+)-coupled rat brain glutamate transporter.” In: *Proceedings of the National Academy of Sciences of the United States of America* 95 (2 Jan. 1998), pp. 751–755.

[36] M P Kavanaugh et al. “Mutation of an amino acid residue influencing potassium coupling in the glutamate transporter GLT-1 induces obligate exchange.” In: *The Journal of biological chemistry* 272 (3 Jan. 1997), pp. 1703–1708.

[37] Dinesh Yernool et al. “Structure of a glutamate transporter homologue from Pyrococcus horikoshii”. In: *Nature* 431.7010 (Oct. 2004), pp. 811–818.

[38] Zhen Tao, Armanda Gameiro, and Christof Grewer. “Thallium Ions Can Replace both Sodium and Potassium Ions in the Glutamate Transporter Excitatory Amino Acid Carrier 1†”. In: *Biochemistry* 47.48 (Dec. 2008), pp. 12923–12930.

[39] H. P. Larsson et al. “Evidence for a third sodium-binding site in glutamate transporters suggests an ion/substrate coupling model”. In: *Proceedings of the National Academy of Sciences* 107.31 (July 2010), pp. 13912–13917.

[40] Zhijian Huang and Emad Tajkhorshid. “Identification of the Third Na+ Site and the Sequence of Extracellular Binding Events in the Glutamate Transporter”. In: *Biophysical Journal* 99.5 (Sept. 2010), pp. 1416–1425.

[41] Turgut Bastug et al. “Position of the Third Na+ Site in the Aspartate Transporter GltPh and the Human Glutamate Transporter, EAAT1”. In: *PLoS ONE* 7.3 (Mar. 2012). Ed. by Hendrik W. van Veen, e33058.

[42] Noa Rosental and Baruch I. Kanner. “A Conserved Methionine Residue Controls the Substrate Selectivity of a Neuronal Glutamate Transporter”. In: *Journal of Biological Chemistry* 285.28 (Apr. 2010), pp. 21241–21248.

[43] Daniel Kortzak et al. “Allosteric gate modulation confers K+ coupling in glutamate transporters”. In: *The EMBO journal* 38.19 (2019), e101468.

[44] Jan-Philipp Machtens et al. “Gating Charge Calculations by Computational Electrophysiology Simulations”. In: *Biophysical Journal* 112.7 (Apr. 2017), pp. 1396–1405.

[45] S. Teichman, S. Qu, and B. I. Kanner. “The equivalent of a thallium binding residue from an archeal homolog controls cation interactions in brain glutamate transporters”. In: *Proceedings of the National Academy of Sciences* 106.34 (Aug. 2009), pp. 14297–14302.

[46] Juddy Mwaura et al. “Protonation State of a Conserved Acidic Amino Acid Involved in Na+ Binding to the Glutamate Transporter EAAC1”. In: *ACS Chemical Neuroscience* 3.12 (Nov. 2012), pp. 1073–1083.

[47] Noa Rosental, Annie Bendahan, and Baruch I. Kanner. “Multiple Consequences of Mutating Two Conserved β-Bridge Forming Residues in the Translocation Cycle of a Neuronal Glutamate Transporter”. In: *Journal of Biological Chemistry* 281.38 (July 2006), pp. 27905–27915.

[48] Ruth Zarbiv et al. “Cysteine Scanning of the Surroundings of an Alkali-Ion Binding Site of the Glutamate Transporter GLT-1 Reveals a Conformationally Sensitive Residue”. In: *Journal of Biological Chemistry* 273.23 (June 1998), pp. 14231–14237.

[49] Renae M. Ryan et al. “The position of an arginine residue influences substrate affinity and K+ coupling in the human glutamate transporter, EAAT1”. In: *Journal of Neurochemistry* 114.2 (May 2010), pp. 565–575.

[50] Natalie Watzke et al. “On the Mechanism of Proton Transport by the Neuronal Excitatory Amino Acid Carrier 1”. In: *The Journal of General Physiology* 116.5 (Oct. 2000), pp. 609–622.

[51] H. P. Larsson et al. “Fluorometric measurements of conformational changes in glutamate transporters”. In: *Proceedings of the National Academy of Sciences* 101.11 (Mar. 2004), pp. 3951–3956.

[52] Christof Grewer et al. “Is the Glutamate Residue Glu-373 the Proton Acceptor of the Excitatory Amino Acid Carrier 1?” In: *Journal of Biological Chemistry* 278.4 (Nov. 2002), pp. 2585–2592.

[53] Germano Heinzelmann and Serdar Kuyucak. “Molecular Dynamics Simulations Elucidate the Mechanism of Proton Transport in the Glutamate Transporter EAAT3”. In: *Biophysical Journal* 106.12 (June 2014), pp. 2675–2683.

[54] D E Koshland, G Némethy, and D Filmer. “Comparison of experimental binding data and theoretical models in proteins containing subunits.” In: *Biochemistry* 5 (1 Jan. 1966), pp. 365–385.

[55] J MONOD, J WYMAN, and J P CHANGEUX. “ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL.” In: *Journal of molecular biology* 12 (May 1965), pp. 88–118.

[56] Austin D. Vogt and Enrico Di Cera. “Conformational Selection or Induced Fit? A Critical Appraisal of the Kinetic Mechanism”. In: *Biochemistry* 51.30 (July 2012), pp. 5894–5902.

[57] Austin D. Vogt et al. “Essential role of conformational selection in ligand binding”. In: *Biophysical Chemistry* 186 (Feb. 2014), pp. 13–21.

[58] Stefano Gianni, Jakob Dogan, and Per Jemth. “Distinguishing induced fit from conformational selection”. In: *Biophysical Chemistry* 189 (May 2014), pp. 33–39.

[59] Fabian Paul and Thomas R. Weikl. “How to Distinguish Conformational Selection and Induced Fit Based on Chemical Relaxation Rates”. In: *PLOS Computational Biology* 12.9 (Sept. 2016). Ed. by Nikolay V. Dokholyan, e1005067.

[60] D. Ewers et al. “Induced fit substrate binding to an archeal glutamate transporter homologue”. In: *Proceedings of the National Academy of Sciences* 110.30 (July 2013), pp. 12486–12491.

[61] Mark James Abraham et al. “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers”. In: *SoftwareX* 1-2 (Sept. 2015), pp. 19–25.

[62] Kresten Lindorff-Larsen et al. “Improved side-chain torsion potentials for the Amber ff99SB protein force field”. In: *Proteins: Structure, Function, and Bioinformatics* (2010), NA–NA.

[63] In Suk Joung and Thomas E. Cheatham. “Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations”. In: *The Journal of Physical Chemistry B* 112.30 (July 2008), pp. 9020–9041.

[64] O. Berger, O. Edholm, and F. Jähnig. “Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature”. In: *Biophysical Journal* 72.5 (May 1997), pp. 2002–2013.

[65] Anselm H. C. Horn. “A consistent force field parameter set for zwitterionic amino acid residues”. In: *Journal of Molecular Modeling* 20.11 (Oct. 2014).

[66] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. “The missing term in effective pair potentials”. In: *The Journal of Physical Chemistry* 91.24 (Nov. 1987), pp. 6269–6271.

[67] Benjamin Webb and Andrej Sali. “Protein Structure Modeling with MODELLER”. In: *Methods in Molecular Biology*. Springer New York, 2017, pp. 39–54.

[68] Maarten G. Wolf et al. “g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation”. In: *Journal of Computational Chemistry* 31.11 (Mar. 2010), pp. 2169–2174.

[69] M. A. Lomize et al. “OPM: Orientations of Proteins in Membranes database”. In: *Bioinformatics* 22.5 (Jan. 2006), pp. 623–625.

[70] Benjamin C McIlwain, Robert J Vandenberg, and Renae M Ryan. “Transport rates of a glutamate transporter homologue are influenced by the lipid bilayer”. In: *Journal of Biological Chemistry* 290.15 (2015), pp. 9780–9788.

[71] Wenchang Zhou et al. “Large-scale state-dependent membrane remodeling by a transporter protein”. In: *Elife* 8 (2019), e50576.

[72] Giovanni Bussi, Davide Donadio, and Michele Parrinello. “Canonical sampling through velocity rescaling”. In: *The Journal of Chemical Physics* 126.1 (Jan. 2007), p. 014101.

[73] H. J. C. Berendsen et al. “Molecular dynamics with coupling to an external bath”. In: *The Journal of Chemical Physics* 81.8 (Oct. 1984), pp. 3684–3690.

[74] M. Parrinello and A. Rahman. “Polymorphic transitions in single crystals: A new molecular dynamics method”. In: *Journal of Applied Physics* 52.12 (Dec. 1981), pp. 7182–7190.

[75] Andrea Amadei, Antonius B. M. Linssen, and Herman J. C. Berendsen. “Essential dynamics of proteins”. In: *Proteins: Structure, Function, and Genetics* 17.4 (Dec. 1993), pp. 412–425.

[76] A. Amadei et al. “An Efficient Method for Sampling the Essential Subspace of Proteins”. In: *Journal of Biomolecular Structure and Dynamics* 13.4 (Feb. 1996), pp. 615–625.

[77] Grubmüller. “Predicting slow structural transitions in macromolecular systems: Conformational flooding.” In: *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* 52 (3 Sept. 1995), pp. 2893–2906.

[78] Oliver F Lange, Lars V Schäfer, and Helmut Grubmüller. “Flooding in GROMACS: accelerated barrier crossings in molecular dynamics.” In: *Journal of computational chemistry* 27 (14 Nov. 2006), pp. 1693–1702.

[79] Shankar Kumar et al. “THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method”. In: *Journal of Computational Chemistry* 13.8 (Oct. 1992), pp. 1011–1021.

[80] Jochen S. Hub, Bert L. de Groot, and David van der Spoel. “g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates”. In: *Journal of Chemical Theory and Computation* 6.12 (Dec. 2010), pp. 3713–3720.

[81] Inga Hänelt et al. “Low Affinity and Slow Na+-Binding Precedes High Affinity Aspartate Binding in the Secondary-active Transporter GltPh”. In: *Journal of Biological Chemistry* 290.26 (Apr. 2015), pp. 15962–15972.

[82] Stefan Van Der Walt, S Chris Colbert, and Gael Varoquaux. “The NumPy array: a structure for efficient numerical computation”. In: *Computing in Science & Engineering* 13.2 (2011), p. 22.

[83] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272.

[84] Richard J Gowers et al. *MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations*. Tech. rep. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2019.

[85] Félix-Antoine Fortin et al. “DEAP: Evolutionary Algorithms Made Easy”. In: *Journal of Machine Learning Research* 13 (July 2012), pp. 2171–2175.

[86] Yannick Hold-Geoffroy, Olivier Gagnon, and Marc Parizeau. “Once you SCOOP, no need to fork”. In: *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment*. 2014, pp. 1–8.

[87] J. D. Hunter. “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95.

[88] Michael Waskom et al. *mwaskom/seaborn: v0.8.1 (September 2017)*. Version v0.8.1. Sept. 2017.

[89] Schrödinger, LLC. “The PyMOL Molecular Graphics System, Version 1.8”. Nov. 2015.

[90] E. M. Landau and J. P. Rosenbusch. “Lipidic cubic phases: A novel concept for the crystallization of membrane proteins”. In: *Proceedings of the National Academy of Sciences* 93.25 (Dec. 1996), pp. 14532–14535.

[91] Nuria Plattner and Frank Noé. “Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models”. In: *Nature communications* 6 (2015), p. 7653.

10. Abbreviations

Arg	R	Arginine
Met	M	Methionine
Asn	N	Asparagine
Thr	T	Threonine

EAAT	excitatory amino acid transporter
Glt_{Ph}	aspartate transporter from *Pyroccocus horikoshii*
TRIS	tris(hydroxymethyl)aminomethane
HP	hairpin
MD	molecular dynamics
SFM	Stopped flow method
PME	particle mesh Ewald
PMF	potential of mean force
SD	standard deviation
TBOA	dl-threo-β-benzyloxyaspartate
TM	transmembrane helix
PCA	principal component analysis
WHAM	Weighted Histogram Analysis Method
PMF	potential of mean force

11. Eidesstattliche Versicherung

Ich versichere an Eides Statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der "Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist.

Ort, Datum Claudia Alleva