SLAS Discovery

SLC6A14, a pivotal actor on cancer stage: when function meets structure

Journal:	*SLAS Discovery*
Manuscript ID	Draft
Manuscript Type:	Original Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Palazzolo, Luca; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Parravicini, Chiara; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Laurenzi, Tommaso; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Adobati, Sara; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Saporiti, Simona; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Guerrini, Uliano; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Gianazza, Elisabetta; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari Indiveri, Cesare; Università della Calabria, Dipartimento di Biologia, Ecologia e Scienze della Terra Anderson, Catriona; Newcastle University Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences Thwaites, David; Newcastle University Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences Eberini, Ivano; Università degli Studi di Milano, Dipartimento di Scienze Farmacologiche e Biomolecolari
Keywords:	Computational chemistry, Pharmacology: ligand binding, receptor binding, amino acid transporter, molecular docking, molecular modelling
Abstract:	ATB0,+ (SLC6A14) is a sodium- and chloride-dependent neutral and dibasic amino acid transporter that regulates the distribution of amino acids across cell membranes. SLC6A14 is ubiquitous. The transporter is over-expressed in many human cancers characterized by an increased demand for amino acids; as such, it was recently acknowledged as a novel target for cancer therapy. The knowledge on the molecular mechanism of SLC6A14 transport is still limited, but some elegant studies on related transporters report the involvement in the transport mechanism of the symmetry of the 12 transmembrane α-helices, and describe structural rearrangements mediated by electrostatic interactions with some pivotal gating residues. In the present work, we constructed a SLC6A14 model in outward facing conformation via homology modelling and tested through molecular

	dynamic simulation to predict amino acid residues critical for substrate recognition and translocation. We then docked the proteinogenic amino acids and other known substrates in the SLC6A14 binding site to study both gating regions and the exposed residues involved in transport. Interestingly, some of these residues correspond to those previously hypothesized by literature; however, we could identify also a novel relevant residue with such function For the first time, by combined approaches of molecular docking and molecular dynamic simulations, we demonstrated the crucial role of these residues in neutral amino acids transport. This novel information unravels new aspects of human SLC6A14 structure/function relationship and may have also important outcomes in cancer through the design of novel inhibitors of SLC6A14-mediated transport.

SLC6A14, a pivotal actor on cancer stage: when function meets structure

Luca Palazzolo[1], Chiara Paravicini[1], Tommaso Laurenzi[1], Sara Adobati[1],Simona Saporiti[1], Uliano Guerrini[1], Elisabetta Gianazza[1], Cesare Indiveri[2], Catriona M.H. Anderson[3], David T. Thwaites[3], Ivano Eberini[4,*]

1 Dipartimento di Scienze Farmacologiche e Biomolecolari, Università degli Studi di Milano, Milano, Italy

2 Dipartimento di Biologia, Ecologia e Scienze della Terra, Università della Calabria, Cosenza, Italy

3 Institute for Cell & Molecular Biosciences, Faculty of Medical Sciences, Framlington Place, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

4 Dipartimento di Scienze Farmacologiche e Biomolecolari & DSRC, Università degli Studi di Milano, Milano, Italy

Corresponding author: Ivano Eberini, ivano.eberini@unimi.it

Abstract

$ATB^{0,+}$ (SLC6A14) is a sodium- and chloride-dependent neutral and dibasic amino acid transporter that regulates the distribution of amino acids across cell membranes. SLC6A14 is ubiquitous. The transporter is over-expressed in many human cancers characterized by an increased demand for amino acids; as such, it was recently acknowledged as a novel target for cancer therapy. The knowledge on the molecular mechanism of SLC6A14 transport is still limited, but some elegant studies on related transporters report the involvement in the transport mechanism of the symmetry of the 12 transmembrane α-helices, and describe structural rearrangements mediated by electrostatic interactions with some pivotal gating residues.

In the present work, we constructed a SLC6A14 model in outward facing conformation via homology modelling and tested through molecular dynamic simulation to predict amino acid residues critical for substrate recognition and translocation. We then docked the proteinogenic amino acids and other known substrates in the SLC6A14 binding site to study both gating regions and the exposed residues involved in transport. Interestingly, some of these residues correspond to those previously hypothesized by literature; however, we could identify also a novel relevant residue with such function

For the first time, by combined approaches of molecular docking and molecular dynamic simulations, we demonstrated the crucial role of these residues in neutral amino acids transport. This novel information unravels new aspects of human SLC6A14 structure/function relationship and may have also important outcomes in cancer through the design of novel inhibitors of SLC6A14-mediated transport.

Introduction

Amino acids are the building blocks of protein synthesis (Sikder et al. 2018). They are essential components of all cells and are required for many metabolic pathways, nitrogen metabolism, neurotransmission and cell growth (Sikder et al. 2018; Scalise et al. 2016). Due to their physicochemical properties, amino acids cannot pass freely across cell membranes but rather undergo carrier-mediated transmembrane movement by specific amino acid transporters. Mammalian cells express numerous amino acid transporters in their plasma membranes to enable intake and efflux of amino acids between each cell and the extracellular environment. Most amino acid transporters have a relatively narrow substrate specificity. However, there is one mammalian amino acid transporter that has an unusually broad substrate selectivity transporting 18 of the 20 proteinogenic amino acids (Sikder et al. 2018). This transporter was originally characterized in rabbit ileum and named the β-alanine carrier although it also transported a wide range of zwitterionic and dibasic amino acids (Munck and Schultz 1969). Separately, a broad-scope amino acid transporter was identified in preimplantation mouse blastocysts and named System $B^{0,+}$ ((Van Winkle, Christensen, and Campione 1985). The transporter cDNA was cloned from human mammary gland and named $ATB^{0,+}$ (Sloan and Mager 1999). $ATB^{0,+}$ is a sodium- and chloride-dependent neutral and dibasic amino acid transporter which possesses all of the functional characteristics of both System $B^{0,+}$ (Sloan and Mager 1999) and the β-alanine carrier (Anderson, Ganapathy, and Thwaites 2008) demonstrating that the early functional studies in intestine and blastocyst represent function of the same transport mechanism.

Mammalian membrane transporters are categorized (based upon amino acid sequence identity) into SoLute Carrier (SLC) familes (Hediger et al. 2013). $ATB^{0,+}$ is the 14^{th} member of the SLC6 family and is also known as SLC6A14 (Bröer and Gether 2012). Although members of the SLC6 family show high levels of sequence identity between transporters, this family includes transporters with very different substrates where some are selective for amino acids and others are able to transport biogenic amines such as norepinephrine (NET/SLC6A2), dopamine (DAT/SLC6A3) or serotonin (SERT/SLC6A4) (Bröer and Gether 2012). The Transporter Classification Database (TCDB) of eukaryote and prokaryote transporters (Saier 2005), catalogues the human SLC6 transporters in the Neurotransmitter:Sodium Symporter family (NSS, TCDB family 2.A.22) which is within the broader Amino acid Polyamine organoCation (APC) Superfamily (Vastermark et al. 2014).

$ATB^{0,+}$ couples cotransport of 2 Na^+, 1 Cl^- and 1 amino acid, and is a highly concentrative electrogenic transport system with inward transmembrane amino acid transport being driven by membrane potential and Na^+ and Cl^- gradients (Sikder et al. 2018). $ATB^{0,+}$ transports neutral amino acids and the dibasic amino acids arginine and lysine and thus transports all essential amino acids and only excludes the anionic amino acids glutamate and aspartate (Sikder et al. 2018). This broad substrate specificity also makes the $ATB^{0,+}$ transporter accessible to a wide variety of amino acid-based drugs and prodrugs including valacyclovir (Hatanaka 2004), valganciclovir (Umapathy, Ganapathy, and Ganapathy 2004), 1-methyltryptophan (Karunakaran et al. 2008), nitric oxide synthase inhibitors (Hatanaka et al. 2001) and carnitine and analogues (Nakanishi et al. 2001).

Abundant expression of $ATB^{0,+}$ in healthy adult tissues is fairly restricted with expression mainly in retina, and respiratory and distal intestinal (ileum and colon) tissues where it is localized to luminal surfaces (Sikder et al. 2018; Hatanaka 2004; Ugawa et al. 2001). In contrast, $ATB^{0,+}$ expression is upregulated in many solid tumours including colorectal (Gupta et al. 2005), cervical (Gupta et al. 2006), estrogen receptor-positive breast (Karunakaran et al. 2011) and pancreatic (Coothankandaswamy et al. 2016) cancers. The high expression in cancer tissues alongside the broad substrate specificity make $ATB^{0,+}$ a good target for anti-cancer treatments using small molecules.

The structure of $ATB^{0,+}$ is unknown but like most members of the APC superfamily (Vastermark et al. 2014) it is predicted to possess a LeuT-fold core structure, named after the amino acid transporter LeuT, the first transporter crystallised from the APC superfamily (Yamashita et al. 2005). Other APC superfamily members have since been crystallised and have broadly similar folds including the human serotonin transporter SERT (SLC6A4) (Coleman, Green, and Gouaux 2016) and the Drosophila melanogaster dopamine transporter DAT (SLC6A3) (Penmatsa, Wang, and Gouaux 2015). $ATB^{0,+}$ is predicted to have 12 transmembrane (TM) α-helices. Based upon available APC superfamily crystal structures, TM 1–5 and TM 6–10 are predicted to be involved in the transport mechanism and have an antiparallel symmetry in their secondary structures giving the α-helices a discrete mobility within the cell membrane and defining binding sites for transported substrates. The binding sites, S1 and S2, are located in the central binding pocket and in the extracellular vestibule, respectively (Kristensen et al. 2011). The central binding pocket must also contain sites that allow the binding of 2 Na^+ and 1 Cl^- ions. TM 11 and 12 are homologous to structures required for the homo-dimerization of other transporters, but currently no evidence is

available regarding $ATB^{0,+}$ dimerization. For LeuT-fold proteins, a model of transport has been proposed (Gao et al. 2010): in the initial phase of the process, the protein lies in an "outward-facing" (OF) conformation and receives the transported substrate (and the ions) in binding site S2. The interaction between protein, ions and substrate promotes the translocation of transported solutes to the S1 binding site and closure of the gate. This process starts from the rotation of a specific conserved aromatic amino acid, which in turn leads to a conformational change in the whole protein, promoting its transition to an "inward-facing" (IF) conformation (Gao et al. 2010; Krammer et al. 2016). Afterwards, the substrate is released by the transporter in the intracellular space and the protein returns to the OF conformation, passing through intermediate transitory occluded states.

Therefore, the purpose of this investigation was to present for the first time the SLC6A14 3D structure, using homology modelling procedures to build a chimera model. Moreover, with this investigation, we meant: (i) to study *in silico* the structural features of the transporter, (ii) to identify *in silico* the key residues for the molecular recognition mechanism in the S1 binding site and (iii) to gather information about substrate binding.

Materials and Methods

2.1. Comparative modelling of SLC6A14

The sequence of human SLC6A14 (Uniprot ID: Q9UN76) was retrieved from the Uniprot knowledgebase database (Li et al. 2015). After a protein Blast (Li et al. 2015) search in the Protein Data Bank database for a human SLC6A14 homologue, the crystal structure of the *Drosophila melanogaster* sodium-dependent dopamine transporter (DAT) (PDB ID:4M48) (Penmatsa, Wang, and Gouaux 2015) was set as main template. A multiple alignment among the NSS family was produced with the Clustal Omega software (Larkin et al. 2007) and used for the homology modelling procedure. Since the crystal structure of DAT lacks the extracellular loop (EL) 2, a second local Blast search was set on residues S205-W270 of SLC6A14, identifying the crystallographic structure of phosphofructokinase (PFKA1) from *Saccharomyces cerevisiae* (PDB ID: 3O8O) (Banaszak et al. 2011) as an adequate template to fix the EL2 gap. The 3D structures of the two templates and of SLC6A14 were also investigated through both SWISS-MODEL and I-TASSER servers. Moreover, the transmembrane prediction programme TMHMM was used to determine the putative topology of SLC6A14, strenghtening the alignment between SLC6A14 and DAT from a structural point of view. PROTTER and was also used to build the prediction of secondary structure.

Selected 3D structures were first optimized and refined *via* further computational steps by correcting crystallographic-related errors through the MOE Structure Preparation Module of the Molecular Operating Environment 2018 (MOE, Chemical Computing Group, Montreal, Canada).

The selected portion (residues from L620 to R665) of phosphofructokinase template was carefully placed in the correspondence of DAT EL2. Three amino acids of both crystal structures were deleted at the point of connection between the two templates in order to have more flexibility to shape their junction in the chimeric model of SLC6A14, using the MOE Homology Model tool. Ten different intermediate models were built and submitted to energy minimization (EM) to release internal constraints. The top-scoring model, according to the GB/IV scoring function, was submitted to further EM with the 'FINE' option until the RMS gradient reached a value below 0.5 kcal/mol/Å.

The Amber10:ETH force field with the reaction-field model for electrostatics was applied for the whole modelling procedure.

2.2 Equilibration and cluster extrapolation

Molecular dynamics (MD) simulation and frame clustering procedures were carried out with the Schrödinger Small-Molecule Drug Discovery Suite 2018-01 (D. E. Shaw Research, New York, NY, 2018, Schrödinger, New York, NY, 2018).

The Desmond System Builder tool was used to place the *apo*-model of SLC6A14 into a POPC membrane bilayer. Protein orientation was set up according to the OPM server (Lomize et al. 2012), which provides spatial arrangements of membrane proteins with respect to the hydrocarbon core of the lipid bilayer. The N- and C-termini of the protein were capped. The system was solvated with 10,174 SPC water molecules in a cubic box with 90 Å edges; adding chloride ions neutralized the exceeding positive charge; sodium chloride was further added up to 0.15 M concentration. The system was energy-minimized to relax the assembly and remove clashes between protein, membrane and solvent in the new setup.

To produce an equilibrated model of SLC6A14, the system was submitted to a 500 ns molecular dynamic simulation (MD) using Desmond Molecular Dynamics tool. Periodic boundary conditions (PBC) and the following parameters were set: 300 K and Nose-Hoover thermostat for temperature coupling, 1 bar and Martyna-Tobias-Klein piston for pressure coupling, 2 fs as integration time step. Coordinates and velocities of each atom were saved every 0.5 ps. The trajectory was then analysed using both Desmond Simulation Event Analysis tool and VMD (Phillips et al. 2005).
Desmond Trajectory Frame Clustering tool was used to cluster the whole 500 MD simulation in order to select the most representative frame (the centroid) for each cluster. Distances between clusters were computed from the RMSD matrix of alpha carbons with respect to the first frame of the MD simulation. Two different runs were performed to set the number of generable clusters according to the dendrogram.
The OPLS3e force field was applied both in the MD simulation and in the frame clusterization procedures.

2.3 Docking
The centroid of the most populated cluster was selected as reference structure for molecular docking procedures. Water, ions and the POPC membrane were washed out from the system and three pivotal ions (1 Cl and 2 Na) from the DAT template were transferred to the SLC6A14 equilibrated model. The ions::SLC6A14 3D structure was then energy-minimized using the Prime in order to fix structural issues in the ions binding domain.
Tested ligands were downloaded from PubChem and prepared for docking with the Schrödinger Ligand Preparation.
The molecular docking procedure was carried out with the Schrödinger Glide Docking in the “extra precision (XP)” mode in order to evaluate the ability of the tested ligands to bind the SLC6A14 binding channel, keeping only the five top-scoring poses.
The top-scoring solution for each ligand was submitted to the Schrödinger Prime MM-GBSA, which integrates molecular mechanics energies with the generalized Born and surface area continuum solvation (Genheden and Ryde 2015) in order to compute ligand binding and ligand strain energies for a set of ligands and a single receptor.
The top-scoring solution for each ligand was also submitted to Desmond Molecular Dynamics tool for a 50 ns of MD simulation in order to study ligand::protein interaction and verify the stability of the ligand placement.
Periodic boundary conditions (PBC) and the following parameters were set for each MD simulation: 300 K and Nose-Hoover thermostat for temperature coupling, 1 bar and Martyna-Tobias-Klein piston for pressure coupling, 2 fs as integration time step. Coordinates and velocities of each atom were saved every 0.5 ps. The trajectories were then analysed using Desmond Simulation Interaction Diagram tool and inspected through VMD.
The OPLS3e force field was applied both in molecular docking procedures and in MD simulations.

Results and Discussion

3.1. Secondary structure analysis and two-dimensional topology prediction

None of the SLC6A14 homologous templates provides complete sequence coverage, as all of them leave out EL2. Not unexpectedly, the SWISS-MODEL and the I-TASSER predictions on the whole protein are unsatisfactory because, in both cases, the EL2 is predicted at an exceedingly high distance from a theoretical conformation resulting in very low local quality estimation. A very high number (>30) of outliers is observed in both Ramachandran plots. Supplementary Figure 1 reports the SWISS-MODEL quality estimation plot and the I-TASSER estimation distance plot.
Building a chimera model seems then to be the right strategy to study SLC6A14 at an atomistic level.
Taking the global alignment among members of the NSS family as the reference structure to model SLC6A14 in its OF conformation, we selected DAT (Penmatsa, Wang, and Gouaux 2015), with a sequence identity of 43% and a similarity of 59%, as the most suitable template presently available. Moreover, unlike other suitable templates whose structure contains only one cation and one anion, DAT is co-crystallized with 2 Na^+ ions and 1 Cl^- ion. On the other hand, according to BLAST, PFKA1, with a local sequence identity of 28% and local similarity of 46%, appears to be the most suitable template for modelling EL2.
The second step towards building a chimera model of SLC6A14 was to compare its putative secondary structure and the two-dimensional topology predicted with three independent methods, while taking into account both the alignment and the topology of the DAT crystal (Penmatsa, Wang, and Gouaux 2015). Figure 1 shows the alignment among SLC6A14 and the selected templates; SLC6A14 is annotated according to TMHMM and PROTTER prediction.

Figure 1

PROTTER and TMHMM prediction are compliant with the DAT secondary structure and its transmembrane placement. In particular, the first three TM α-helices of DAT are more extended than expected from the prediction, while TM α-helices from 4 to 7 and from 9 to 12 are in good agreement with the expectations. TM α-helix 8 seems to be shifted towards the C-terminus by 5 amino acids with respect to PROTTER prediction, while it complies with TMHMM prediction. The alignment among SLC6A14, DAT and PFKA1 was carefully verified taking into account also the PROTTER and TMHMM prediction. Predicted topology of SLC6A14 is shown in Figure 2.

Figure 2

3.2. Homology modelling of SLC6A14 and its MD equilibration

The Ramachandran plot for the chimera model of SLC6A14, based on both DAT and PFKA1, shows only six outliers, none of them in the transmembrane domain (Supplementary Figure 2.A). Its secondary structure is compliant with both the DAT structure and the PROTTER and TMHMM predictions. As expected, SLC6A14 model has a LeuT-fold, with two well defined binding sites for both substrate/inhibitor and Na^+/Cl^- ions, respectively.

In order to equilibrate the chimera model of SLC6A14, we carried out a 500 ns MD simulation. The evaluation of the stability of our model during the simulation was based both on energetic and geometric parameters, as described in the following.

The general stability of the SLC6A14 model is confirmed by the tendency of the MD simulations to reach convergence. In detail, the tendency to reach an equilibrium state is suggested by the trend of the root mean square deviation (RMSD) values, which, after 100 ns, reach a plateau around 3.5 Å (Figure 3.A); its fluctuations towards the end of the simulation are around some thermal average structure. The lowest root mean square fluctuations (RMSF) are associated with α-helices, the highest with the loops, both inter- and extra-cellular, with maximum in EL2 (Figure 3.B).

Figure 3

Protein secondary structure elements (SSE) i.e. α-helices and β-strands, were monitored throughout the simulation. Figure 3.C reports SSE distribution by residue index, while Figure 4 summarizes the SSE composition for each trajectory frame over the course of the simulation and monitors each residue and its SSE assignment over time. Globally, the secondary structure of SLC6A14 is conserved during the MD simulation; the helices in EL2 are structured in over 50% of the frames, despite EL2 undergoing a strong rearrangement during the first 200 ns. In fact, EL2 migrates from a conformation extended towards the extracellular side to a more stable situation (Supplementary Figure 3), in which it establishes some H-bonds with the exposed amino acids of other ELs. Asn219 interacts with Lys225 via an H-bond, that is conserved for ~ 70% of the frames, shaping a turn-like structure and partially covering the SLC6A14 transport channel. This peculiar shape of EL2 is present in the frame selected for structural analysis, that is the centroid of the most populated cluster, composed by ~70% of the all frames. The Ramachandran plot of the selected frame shows only one outlier (Supplementary Figure 2.B).

Figure 4

3.3 Structural features of the binding sites

Figure 5

To study the structural features of SLC6A14 transport channel, and its binding sites, we imported from DAT crystal structure to our SLC6A14 equilibrated model two Na^+ and one Cl^- ions and, then, we relaxed the protein to avoid atom clashes. Figure 5 shows the SLC6A14 equilibrated structure after the local minimization of the ion pocket.

As mentioned in the Introduction, it is well known that, when in OF conformation, SLC6A family members have a binding site in the bottom of their transport channel. This binding site includes two specific regions: the first for the substrate, the second for the co-transported ions, which are a fundamental feature of all the SLC6 transporters.

In SLC6A14 we found two distinct sites in the S1 pocket that are unambiguously being occupied by a Na^+ ion, designated Na1 and Na2 (Singh 2008). Both Na^+ binding sites are believed to have a key role in stabilizing TM1 and TM6 in the presence of substrate.

Na1 binding site is closed towards the Cl^- binding site, and is divided from Na2 by TM1, sharing some amino acids with it. Na1 is not contiguous with the S1 substrate binding site, while the Na1 site insists on the S1 substrate binding site, and the Na^+ ion is available for interacting with the substrate. Residues composing the Na1 and Na2 ion binding sites are conserved among the transporters of interest and, in general, are highly conserved across the whole SLC6 family (Kristensen et al. 2011). As suggested by Kristensen (Kristensen et al. 2011), the evidence of contiguous surfaces between Na2 and the substrate binding site can be the basis for a coupled translocation mechanism for both, Na^+ and substrate. On the other hand, the placement of Cl^- in the Na1 binding site supports the idea that Cl^- is translocated together with the Na^+ ion along the SLC6A14 channel during the transport of the substrate. The same ion placement has been reported also for the crystallographic structures of DAT (SLC6A3) and human SERT (SLC6A4) (Penmatsa, Wang, and Gouaux 2015; Coleman, Green, and Gouaux 2016). These data confirm that our equilibrated SLC6A14 model maintains its ion binding site (Figure 6.A).

Figure 6

The S1 substrate binding site is defined by polar, aromatic, and aliphatic amino acid side chains provided by all four TMs (TM1, TM3, TM6, and TM8) that surround the binding site, in addition to the backbone amide groups from the unwound regions of TM1 and TM6 (Singh 2008). The S1 pocket can be divided into two regions: a polar region formed exclusively by the unwound regions of TM1 and TM6 and a hydrophobic pocket formed by aliphatic side chains from TM1, TM3, and TM6 (Figure 6.B).

In the SLC6A14 equilibrated model, we can also detect the S2 binding site in the gating region (Kristensen et al. 2011), which is partially defined at the bottom by three amino acids: Tyr321, Arg61and Asp478, which are partially conserved among SLC6A14, DAT and SLC6A4 (Figure 6.B).

Arg104 and Asp478 establish a H-bond, fixing the Arg104 in a conformation stretched toward the inside of the transport channel. Tyr321 is stretched toward the transport channel, unlike its corresponding amino acids Phe319 and Phe335 in DAT and SLC6A14, respectively.

From a structural point of view, this particular aromatic amino acid seems to assume the same function of Trp202, the gating residue of the arginine/agmatine antiporter AdiC. This gating mechanism is also supposed to be conserved for the SLC6 family (Pramod et al. 2013) and it was simulated for LAT1 (SLC7A5) *via* target molecular dynamics (Palazzolo et al. 2018). To evaluate the gating arrangement, we compared the placement of Phe319 in two DAT crystallographic structures in OF conformation, with gate open (Pdb ID: 4M48) (Penmatsa, Wang, and Gouaux 2015) and closed (Pdb ID: 4XPA) (Wang, Penmatsa, and Gouaux 2015), *via* a structural superposition. As shown in Figure 9.A, Tyr321 of the equilibrated model of SLC6A14 is in an intermediate position between DAT Phe319 open and closed conformations, suggesting that this amino acid has an intrinsic flexibility also in the absence of substrates/inhibitors. This intrinsic flexibility is also observed during the MD simulation, during which Tyr321 oscillates between open and partially closed conformation, with respect to DAT as a reference (Figure 7.A), suggesting that this amino acid has many degrees of freedom while it explores the conformational spaces in its surroundings.

Figure 7

Moreover, we also identified the Trp327 as second putative inner gate of SLC6A14. In particular, this amino acid rotates from a DAT Phe325-like conformation to an occluded one, defining the bottom of the S1 substrate binding site. Evidence that Trp327 can be involved in the inner gating mechanism can be obtained from a structural analysis of the DAT Phe325 behaviour in presence of different inhibitors. As shown in Figure 7.B, DAT Phe325 has some degrees of freedom to coordinate itself with the transported molecules. In the two reference crystallographic structures of DAT, Phe325 is oriented by the inhibitors and it is alternately open or closed, with Phe319 behaving as the outer gate.

3.4 Molecular docking

As mentioned in the introduction, there are some aa transporters, among which SLC6A14, that transports 18 of the 20 amino acids, i.e. all of the proteogenic amino acids except glutamate and aspartate (Scalise et al. 2016). Molecular docking simulation was applied to place the substrates into the S1 binding site, appraising substrate::carrier interactions and computing the relevant docking scores. On the basis of these results, each complex substrate::carrier was submitted to a 50 ns MD simulation to evaluate the stability of the interactions and to assess the molecular recognition mechanisms, looking for pivotal amino acids involved in the recognition of the different substrates.
As expected, most of the tested ligands have docking scores ranging from -8 kcal/mol to -1 kcal/mol, values that are compliant with solute transport. Supplementary Table 1 reports the most interesting interactions between tested compound and SLC6A14 amino acids during the MD simulations.

Tryptophan and 1-methyl-L-tryptophan are the substrates associated with the best docking scores, i.e. -7.5 kcal/mol and -6.3 kcal/mol, respectively. These substrates are placed into the S1 binding site with overlapping docking poses (Figure 8); their interactions are mainly with Tyr321, Trp327, Phe320, Ser324, Leu56 and Ser423. The MD simulations show that the most conserved contacts between tryptophan and SLC6A14 amino acids are with Tyr52, Ala53, Gly57, Tyr132 and Tyr321 and that all of these contacts occur via H-bonds. On the other hand, 1-methyl-L-tryptophan interacts via H-bonds with Leu56, Gly57, Tyr132, Phe320 and Ser324, and all of these contacts are conserved during the MD simulation. Ala53 establishes an important ionic interaction with 1-methyl-L-tryptophan that is conserved in the whole MD simulation. Both substrates interact also with Na^+ in the Na1 ion binding site for ~97% of the simulation total frames. These results are in agreement with those of both Hatanaka et al. (2004) and Karunakaran et al. (2008).

Figure 8

Tyrosine overlap the Tryptophan docking pose, with associated a docking score of -5.4 kcal/mol, mainly interacting with Ala326 and Ser324. The MD simulation highlight also H-bonds interaction with Val54, Leu56, Ser324 and Ser 423 and other important hydrophobic interactions with Tyr132, Tyr321, Trp327.

Arginine, Leucine and Isoleucine share the same binding pose and have similar docking scores of -4.9 kcal/mol, -4.6 kcal/mol and -4.4 kcal/mol, respectively. Arginine mainly interacts with Tyr132, Phe320 and Asp478, while Leucine and Isoleucine establishes H-bonds with Tyr132, Phe320, Val325 and Ser324. The MD simulations show that Asp478 is the pivotal binder for Arginine and it is probably involved also in the recognition mechanism, while for Leucine and IsoleucineTyr321 is the pivotal amino acid with which interacts for ~ 94% of the MD simulation. Arginine interacts with Cl^-, while Leucine interacts with Na^+ in the Na1 ions binding site.

As reported in Scalise et al. (2016), Glutamine is compressed between Val128 and Tyr321, with a docking score value of -3.8 kcal/mol. Glutamine also interacts *via* H-bonds with Tyr132, Phe320 and Asp478 in its binding pose, while the MD simulation detects ionic interactions with Ala53, Tyr321 and Ser324. Glutamine also interacts with Na^+ in the Na1 ion binding site.

Proline and Cysteine have very similar both binding poses and docking scores of about -3.1 kcal/mol. This tested amino acids share the same interaction with Tyr132 and Phe320, establishing four H-bonds during the MD simulations also with Ser324 and Val325. Contact with Tyr132, Phe320 and Ser324 are conserved for both substrates for over 90% of the total time of the MD simulations.

Carnitine shows a completely different binding pose with respect to the tested amino acids one, with associated a docking score of -1.4 kcal/mol. From the ligand interaction diagram of molecular docking, it interacts with Leu56 and Ser423 via H-bonds, while during the MD simulation it shift in the S1 binding pocket from Ser478 to the S1 ion binding site, interacting with Na^+. In this context, Carnitine interacts with Tyr95, Leu99, Gly100, Tyr176, Trp381, Tyr375 and Ser378.

Finally, the two prodrugs valganciclovir and valaciclovir have dissimilar behaviour in their binding poses. Despite their binding poses partly overlapping, the docking score is -5.2 kcal/mol for the former, and -1.2 kcal/mol, for the latter. Both substrates interact with Asp478, Tyr321 and Tyr132 but Valganciclovir establishes also some contacts with Tyr52, Leu56 and Asp477. The MD simulations show for both substrates that those with Tyr132 and Asp478 are the most conserved H-bonds interactions, together with Leu56, Ala537 and Ser324. None of them interact with Na^+ or Cl^- ions.

For all the tested substrates, also Val172 is always involved via a hydrophobic interaction. In fact, this amino acid is part of the S1 binding site, and it sterically occupies part of the pocked with its side chain. Moreover, arginine and its derivatives, glutamine and both the test prodrugs are coordinated by Asp478 that is placed in the upper part of the S1 binding site, together with Tyr321 that is pivotal in all the interactions. These data suggest that Tyr321 can act as gate, involving also the Asp478, while other aromatic and aliphatic side chains of amino acids in S1 are involved in the recognition and transport mechanism of substrates trough the SLC6A14 channel.

Conclusions

With our research, we propose, the first atomistic model of SLC6A14 transporter based on a chimeric approach: the two templates selected for modelling the OF SLC6A14 state are the most suitable so far available, and one of them has also been profitably used in a recent publication (Scalise et al. 2016). Thanks to 500 ns of MD simulation, we identified a specific behaviour of Tyr321, highlighting its role in the outer gating mechanism that is coordinated with the involvement of Arg104 and Asp478. We also identified Trp327 as putative inner gate, in analogy to DAT, our reference template. These two amino acids define the S1 binding site while Tyr321 is at the same time at the bottom of the S2 binding site. From the cluster analysis we selected a reference 3D structure of SLC6A14 and we observed that ions binding sites are conserved across the SLC6 family.

Natural amino acids and some other well-known substrates were tested for interaction via molecular docking simulation. For all of them both the docking scores were evaluated, and the molecular recognition mechanism was characterized, highlighting that both gates are essential for binding. In a broader way, Try52, Gly57, Val128, Ser322 and Ser324 compose an ensemble of amino acids that orient the substrates, confirming the structural findings of Yamashita et al. (2005), Scalise et al. (2016) and Edwards et al. (2018).

Author contribution

LP set up the homology modelling procedure, performed part of the simulations, conducted part of the analyses and drafted the manuscript. CP, SA, SS conducted part of the analyses. CP, TL and UG performed part of the simulations and contributed to the data analysis. CI, CA, DT and IE conceived the concept of this work. CI, CA, DT, EG and IE supervised the work and revised the manuscript.

Funding

This research was supported by grants from MIUR Progetto Eccellenza. IE gratefully acknowledges departmental Linea 2 - Azione A 2017 funding.

References

Anderson, Catriona M.H., Vadivel Ganapathy, and David T. Thwaites. 2008. “Human Solute Carrier SLC6A14 Is the β-Alanine Carrier.” *Journal of Physiology* 586 (17): 4061–67. doi:10.1113/jphysiol.2008.154500.

Banaszak, Katarzyna, Ingrid Mechin, Galina Obmolova, Michael Oldham, Simon H. Chang, Teresa Ruiz, Michael

Radermacher, Gerhard Kopperschläger, and Wojciech Rypniewski. 2011. “The Crystal Structures of Eukaryotic Phosphofructokinases from Baker’s Yeast and Rabbit Skeletal Muscle.” *Journal of Molecular Biology*. doi:10.1016/j.jmb.2011.01.019.

Bröer, Stefan, and Ulrik Gether. 2012. “The Solute Carrier 6 Family of Transporters.” *British Journal of Pharmacology*. doi:10.1111/j.1476-5381.2012.01975.x.

Coleman, Jonathan A., Evan M. Green, and Eric Gouaux. 2016. “Thermostabilization, Expression, Purification, and Crystallization of the Human Serotonin Transporter Bound to S</Em>-Citalopram.” *Journal of Visualized Experiments*. doi:10.3791/54792.

Coothankandaswamy, V., S. Cao, Y. Xu, P. D. Prasad, P. K. Singh, C. P. Reynolds, S. Yang, J. Ogura, V. Ganapathy, and Y. D. Bhutia. 2016. “Amino Acid Transporter SLC6A14 Is a Novel and Effective Drug Target for Pancreatic Cancer.” *British Journal of Pharmacology* 173 (23): 3292–3306. doi:10.1111/bph.13616.

Gao, Xiang, Lijun Zhou, Xuyao Jiao, Feiran Lu, Chuangye Yan, Xin Zeng, Jiawei Wang, and Yigong Shi. 2010. “Mechanism of Substrate Recognition and Transport by an Amino Acid Antiporter.” *Nature* 463 (7282): 828–32. doi:10.1038/nature08741.

Genheden, Samuel, and Ulf Ryde. 2015. “The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities.” *Expert Opinion on Drug Discovery*. doi:10.1517/17460441.2015.1032936.

Gupta, Naren, Seiji Miyauchi, Robert G. Martindale, Anne V. Herdman, Robert Podolsky, Katsuya Miyake, Sela Mager, Puttur D. Prasad, Malliga E. Ganapathy, and Vadivel Ganapathy. 2005. “Upregulation of the Amino Acid Transporter ATB0,+ (SLC6A14) in Colorectal Cancer and Metastasis in Humans.” *Biochimica et Biophysica Acta - Molecular Basis of Disease*. doi:10.1016/j.bbadis.2005.04.002.

Gupta, Naren, Puttur D. Prasad, Sharad Ghamande, Pamela Moore-Martin, Anne V. Herdman, Robert G. Martindale, Robert Podolsky, Sela Mager, Malliga E. Ganapathy, and Vadivel Ganapathy. 2006. “Up-Regulation of the Amino Acid Transporter ATB0,+ (SLC6A14) in Carcinoma of the Cervix.” *Gynecologic Oncology*. doi:10.1016/j.ygyno.2005.08.016.

Hatanaka, T. 2004. “Transport of Amino Acid-Based Prodrugs by the Na+- and Cl--Coupled Amino Acid Transporter ATB0,+ and Expression of the Transporter in Tissues Amenable for Drug Delivery.” *Journal of Pharmacology and Experimental Therapeutics* 308 (3): 1138–47. doi:10.1124/jpet.103.057109.

Hatanaka, T., T. Nakanishi, W. Huang, F. H. Leibach, P. D. Prasad, V. Ganapathy, and M. E. Ganapathy. 2001. “Na+- and Cl--Coupled Active Transport of Nitric Oxide Synthase Inhibitors via Amino Acid Transport System B0,+.” *Journal of Clinical Investigation* 107 (8): 1035–43. doi:10.1172/JCI12060.

Hediger, Matthias A., Benjamin Clémençon, Robert E. Burrier, and Elspeth A. Bruford. 2013. “The ABCs of Membrane Transporters in Health and Disease (SLC Series): Introduction.” *Molecular Aspects of Medicine* 34 (2–3): 95–107. doi:10.1016/j.mam.2012.12.009.

Karunakaran, Senthil, Sabarish Ramachandran, Veena Coothankandaswamy, Selvakumar Elangovan, Ellappan Babu, Sudharsan Periyasamy-Thandavan, Ashish Gurav, et al. 2011. “SLC6A14 (ATB 0,+) Protein, a Highly Concentrative and Broad Specific Amino Acid Transporter, Is a Novel and Effective Drug Target for Treatment of Estrogen Receptor-Positive Breast Cancer.” *Journal of Biological Chemistry*. doi:10.1074/jbc.M111.229518.

Karunakaran, Senthil, Nagavedi S. Umapathy, Muthusamy Thangaraju, David H. Munn, Puttur D. Prasad, Vadivel Ganapathy, Takahiro Hatanaka, and Shiro Itagaki. 2008. “ Interaction of Tryptophan Derivatives with SLC6A14 (ATB 0,+) Reveals the Potential of the Transporter as a Drug Target for Cancer Chemotherapy .” *Biochemical Journal* 414 (3): 343–55. doi:10.1042/bj20080622.

Krammer, Eva Maria, Kassem Ghaddar, Bruno Andr??, and Martine Pr??vost. 2016. “Unveiling the Mechanism of Arginine Transport through AdiC with Molecular Dynamics Simulations: The Guiding Role of Aromatic Residues.” *PLoS ONE* 11 (8): 1–29. doi:10.1371/journal.pone.0160219.

Kristensen, Anders S, Jacob Andersen, T. N. Jorgensen, L. Sorensen, Jacob Eriksen, Claus J Loland, K. Stromgaard, and U. Gether. 2011. “SLC6 Neurotransmitter Transporters: Structure, Function, and Regulation.” *Pharmacological Reviews*. doi:10.1124/pr.108.000869.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. Mcgettigan, H. McWilliam, F. Valentin, et al. 2007. “Clustal W and Clustal X Version 2.0.” *Bioinformatics* 23 (21): 2947–48. doi:10.1093/bioinformatics/btm404.

Li, Weizhong, Andrew Cowley, Mahmut Uludag, Tamer Gur, Hamish McWilliam, Silvano Squizzato, Young Mi Park, Nicola Buso, and Rodrigo Lopez. 2015. “The EMBL-EBI Bioinformatics Web and Programmatic Tools Framework.” *Nucleic Acids Research* 43 (W1): W580–84. doi:10.1093/nar/gkv279.

Lomize, Mikhail A., Irina D. Pogozheva, Hyeon Joo, Henry I. Mosberg, and Andrei L. Lomize. 2012. “OPM Database and PPM Web Server: Resources for Positioning of Proteins in Membranes.” *Nucleic Acids Research* 40 (D1). doi:10.1093/nar/gkr703.

Munck, B. G., and Stanley G. Schultz. 1969. “Interactions between Leucine and Lysine Transport in Rabbit Ileum.” *BBA - Biomembranes*. doi:10.1016/0005-2736(69)90142-4.

Nakanishi, Takeo, Takahiro Hatanaka, Wei Huang, Puttur D. Prasad, Frederick H. Leibach, Malliga E. Ganapathy, and

Vadivel Ganapathy. 2001. “Na+ - and Cl- -Coupled Active Transport of Carnitine by the Amino Acid Transporter ATB0,+ from Mouse Colon Expressed in HRPE Cells and Xenopus Oocytes.” *Journal of Physiology* 532 (2): 297–304. doi:10.1111/j.1469-7793.2001.0297f.x.

Palazzolo, Luca, Chiara Parravicini, Tommaso Laurenzi, Uliano Guerrini, Cesare Indiveri, Elisabetta Gianazza, and Ivano Eberini. 2018. “In Silico Description of LAT1 Transport Mechanism at an Atomistic Level.” *Frontiers in Chemistry*. doi:10.3389/fchem.2018.00350.

Penmatsa, Aravind, Kevin H. Wang, and Eric Gouaux. 2015. “X-Ray Structures of Drosophila Dopamine Transporter in Complex with Nisoxetine and Reboxetine.” *Nature Structural and Molecular Biology*. doi:10.1038/nsmb.3029.

Phillips, James C., Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. 2005. “Scalable Molecular Dynamics with NAMD.” *Journal of Computational Chemistry*. doi:10.1002/jcc.20289.

Pramod, Akula Bala, James Foster, Lucia Carvelli, and L. Keith Henry. 2013. “SLC6 Transporters: Structure, Function, Regulation, Disease Association and Therapeutics.” *Molecular Aspects of Medicine*. doi:10.1016/j.mam.2012.07.002.

Saier, M. H. 2005. “TCDB: The Transporter Classification Database for Membrane Transport Protein Analyses and Information.” *Nucleic Acids Research*. doi:10.1093/nar/gkj001.

Scalise, Mariafrancesca, Lorena Pochini, Michele Galluccio, and Cesare Indiveri. 2016. “Glutamine Transport. From Energy Supply to Sensing and Beyond.” *Biochimica et Biophysica Acta - Bioenergetics* 1857 (8). Elsevier B.V.: 1147–57. doi:10.1016/j.bbabio.2016.03.006.

Sikder, Mohd Omar F., Shengping Yang, Vadivel Ganapathy, and Yangzom D. Bhutia. 2018. “The Na+/Cl−-Coupled, Broad-Specific, Amino Acid Transporter SLC6A14 (ATB0,+): Emerging Roles in Multiple Diseases and Therapeutic Potential for Treatment and Diagnosis.” *The AAPS Journal* 20 (1). The AAPS Journal: 12. doi:10.1208/s12248-017-0164-7.

Singh, Satinder K. 2008. “LeuT: A Prokaryotic Stepping Stone on the Way to a Eukaryotic Neurotransmitter Transporter Structure.” *Channels*. doi:10.4161/chan.2.5.6904.

Sloan, Jennifer L., and Sela Mager. 1999. “Cloning and Functional Expression of a Human Na+ and Cl--Dependent Neutral and Cationic Amino Acid Transporter B0+.” *Journal of Biological Chemistry* 274 (34): 23740–45. doi:10.1074/jbc.274.34.23740.

Ugawa, Shinya, Yoko Sunouchi, Takashi Ueda, Eri Takahashi, Yoshitsugu Saishin, and Shoichi Shimada. 2001. “Characterization of a Mouse Colonic System B $^{0+}$ Amino Acid Transporter Related to Amino Acid Absorption in Colon.” *American Journal of Physiology-Gastrointestinal and Liver Physiology*. doi:10.1152/ajpgi.2001.281.2.G365.

Umapathy, Nagavedi S., Vadivel Ganapathy, and Malliga E. Ganapathy. 2004. “Transport of Amino Acid Esters and the Amino-Acid-Based Prodrug Valganciclovir by the Amino Acid Transporter ATB 0,+.” *Pharmaceutical Research*. doi:10.1023/B:PHAM.0000033019.49737.28.

Van Winkle, L. J., H. N. Christensen, and A. L. Campione. 1985. “Na+-Dependent Transport of Basic, Zwitterionic, and Bicyclic Amino Acids by a Broad-Scope System in Mouse Blastocysts.” *Journal of Biological Chemistry* 260 (22): 12118–23.

Vastermark, Ake, Simon Wollwage, Michael E. Houle, Rita Rio, and Milton H. Saier. 2014. “Expansion of the APC Superfamily of Secondary Carriers.” *Proteins: Structure, Function and Bioinformatics*. doi:10.1002/prot.24643.

Wang, Kevin H., Aravind Penmatsa, and Eric Gouaux. 2015. “Neurotransmitter and Psychostimulant Recognition by the Dopamine Transporter.” *Nature*. doi:10.1038/nature14431.

Yamashita, Atsuko, Satinder K. Singh, Toshimitsu Kawate, Yan Jin, and Eric Gouaux. 2005. “Crystal Structure of a Bacterial Homologue of Na+/Cl--Dependent Neurotransmitter Transporters.” *Nature* 437 (7056): 215–23. doi:10.1038/nature03978.

Figure 1. Alignment among SLC6A14, DAT and PFKA1. TMHMM and PROTTER prediction are marked as violet bar, while secondary structure i.e. α-helices and β-sheets of both DAT and PFKA1 are coloured in red and yellow, respectively.

Figure 2. 2D topology and transmembrane prediction for SLC6A14. Results with PROTTER are shown in (A), with TMHMM in (B). In (A) transmembrane helices are numbered from 1 to 12 and the putative N-linked glycosylation sites is marked in green.

Figure 3: Stability of SLC6A14 model. (A) Root mean square deviation of SLC6A14 α-carbon computed for each frame of MD simulation with respect to frame 0. (B) Root mean square fluctuation of SLC6A14 α-carbon during the MD simulation. (C) Secondary structure elements: α-helical regions are highlighted in red These regions are defined by helices that persist over 70% of the entire simulation.

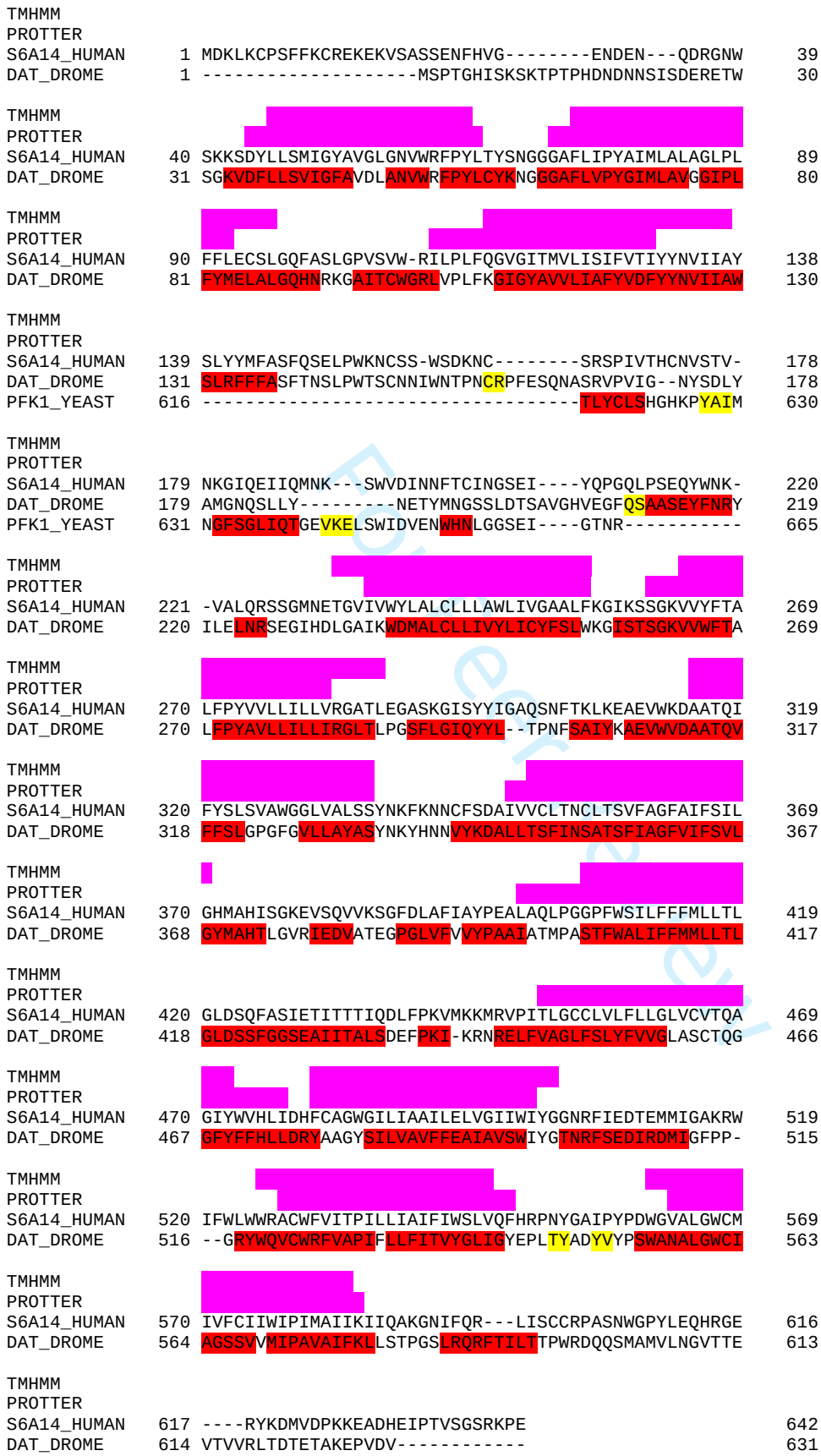
Figure 4: Stability of SLC6A14 model. (Too) Root mean square deviation of SLC6A14 α-carbons computed for each frame of MD simulation with respect to frame 0. (Bottom) Root mean square fluctuation of SLC6A14 α-carbon during the MD simulation. (C) Secondary structure elements. α-helical regions, defined as structures that persist over 70% of the entire simulation, are highlighted in red.

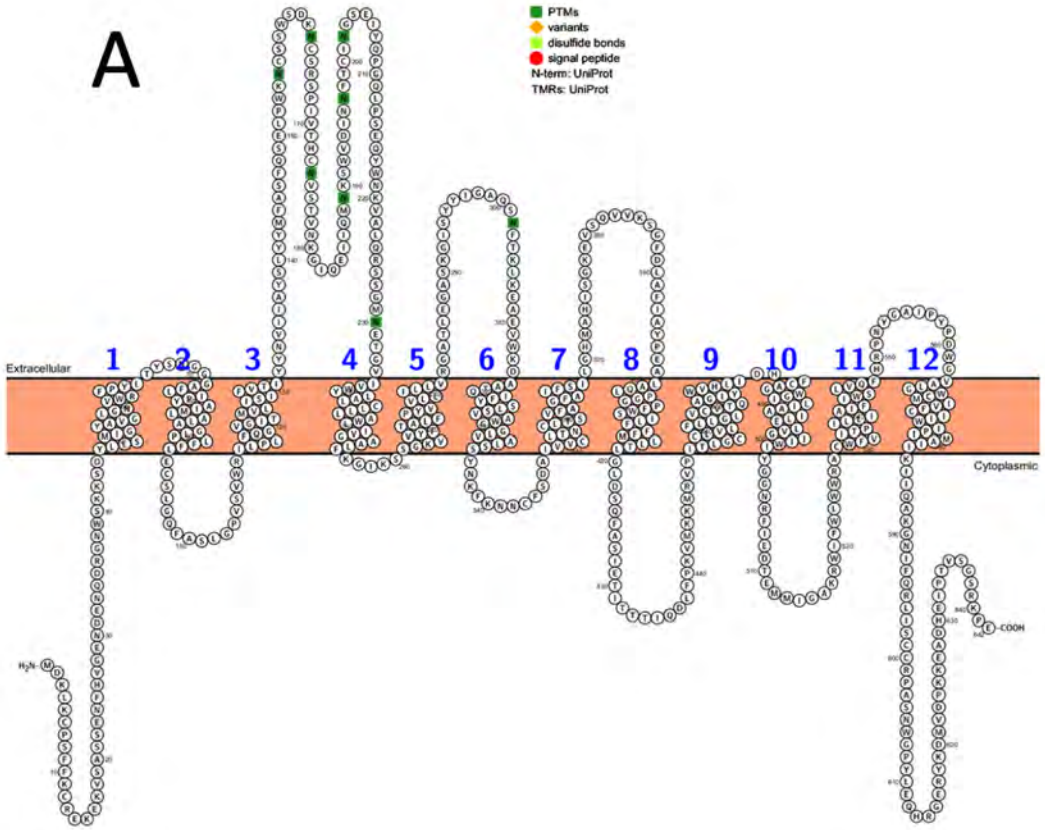
Figure 5. SLC6A14 equilibrated model. (A) Transversal view, (B) bottom and (C) top views. Secondary structure is represented by ribbons and coloured from N (blue) to C terminus (red); ions are represented as balloons.

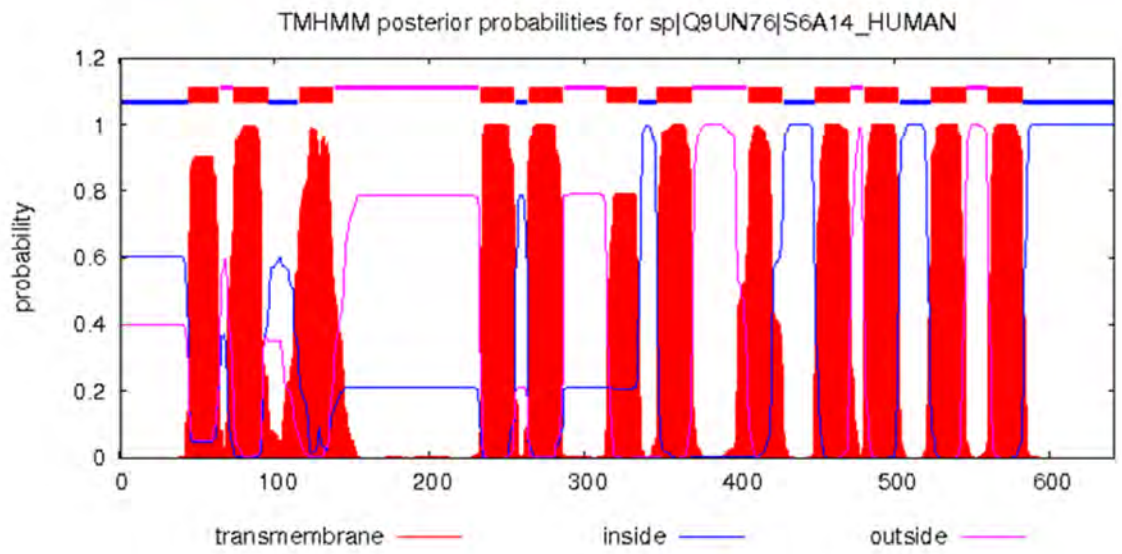
Figure 6: SLC6A14 binding sites. (A) Ions binding site, and (B) S1 and S2 substrate binding site surfaces in SLC6A14. Putative amino acids involved in the gating mechanism are outlined. The molecular surfaces of SLC6A14 binding sites computed as van der Waals interaction surface are represented with green, magenta and blue lines for hydrophobic, H-bonding and mild polar interactions, respectively.

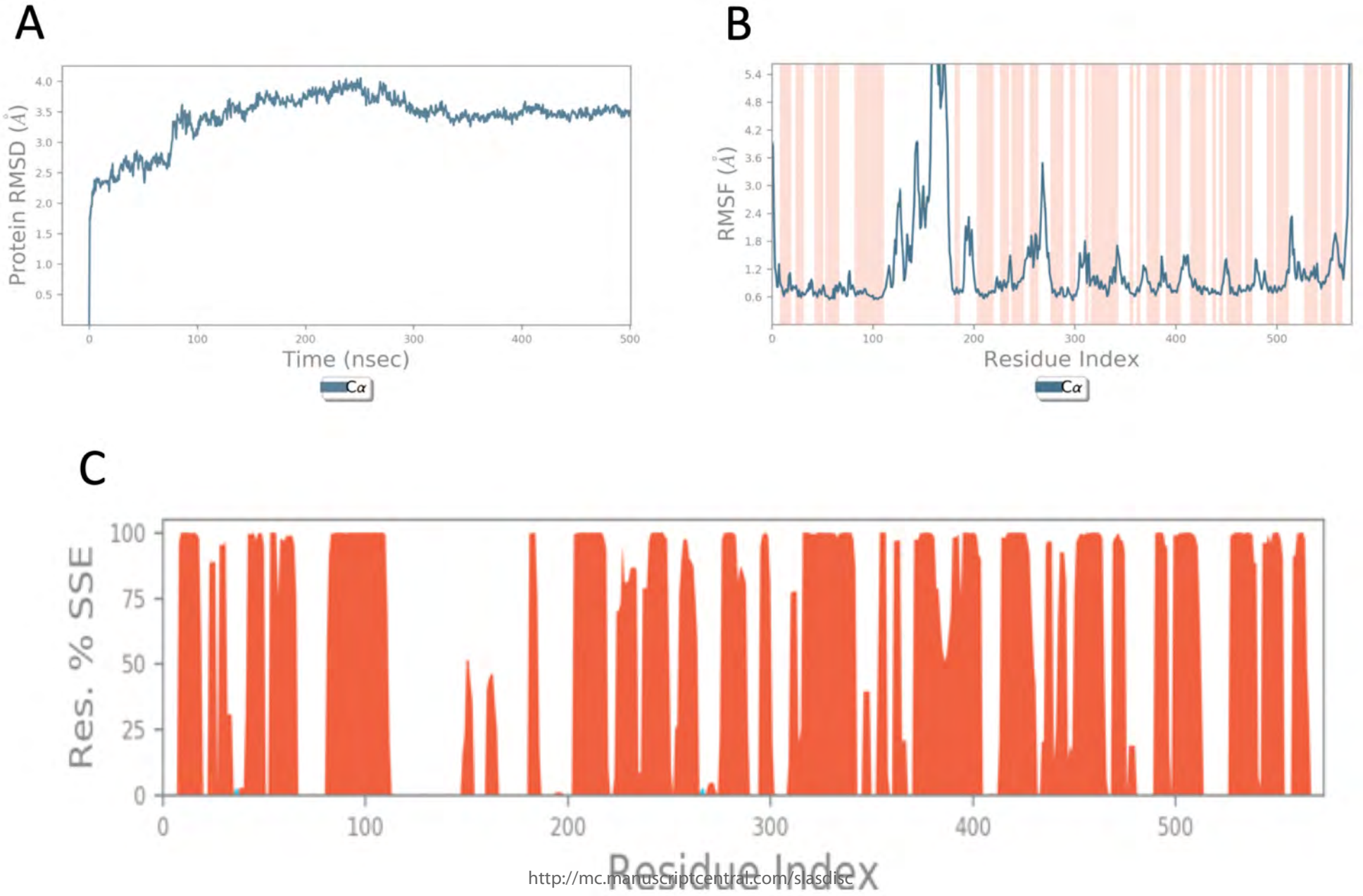
Figure 7: (A) Outer gate of SLC6A14 and its superposition to DAT crystallographic structures both in OF-open and OF-closed conformations. (B) SLC6A14 and DAT gates; as for other APCs, SLC6A14 seems to have two distinct aromatic amino acids that are involved in the gating mechanism. SLC6A14, DAT OF-open and DAT OF-closed are represented in yellow, light blue and light green, respectively.

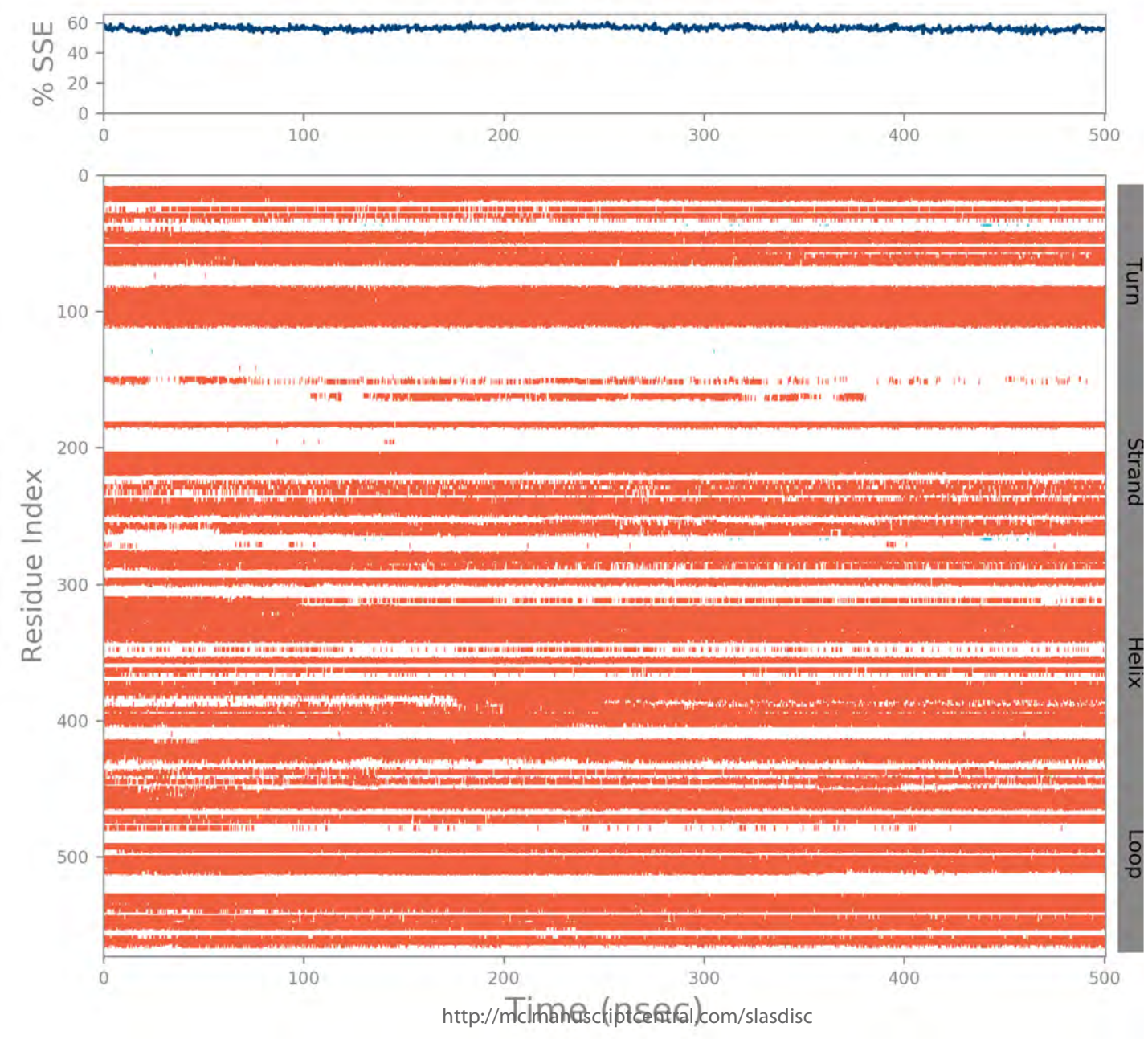
Figure 8: (A) Tryptophan and 1-methyl-L-tryptophan, and (B) valaciclovir and valganciclovir binding poses. SLC6A14 helices and gating amino acids are coloured in yellow.

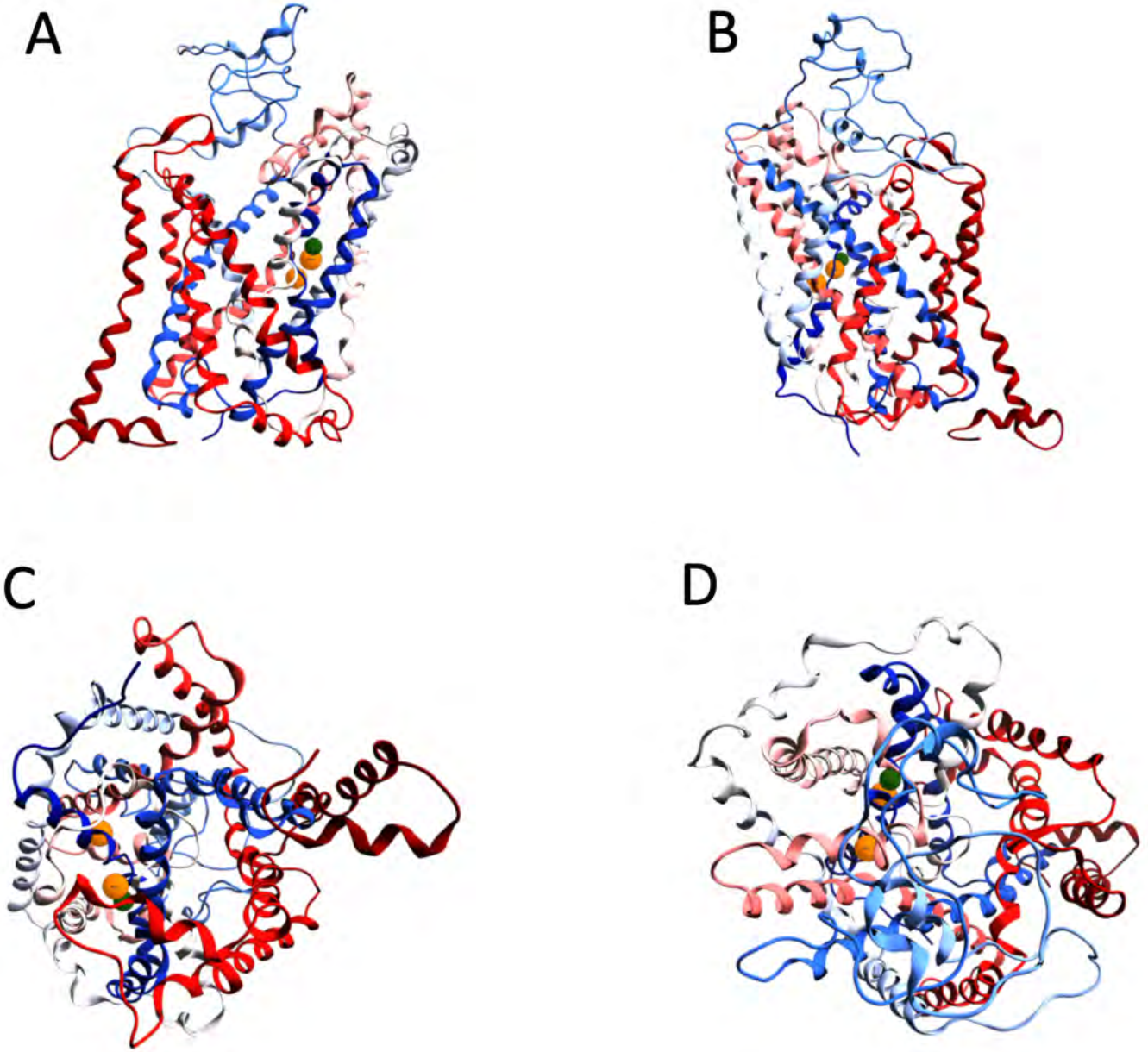

TMHMM
PROTTER
S6A14_HUMAN 1 MDKLKCPSFFKCREKEKVSASSENFHVG--------ENDEN---QDRGNW 39
DAT_DROME 1 --------------------MSPTGHISKSKTPTPHDNDNNSISDERETW 30
S6A14_HUMAN 40 SKKSDYLLSMIGYAVGLGNVWRFPYLTYSNGGGAFLIPYAIMLALAGLPL 89
DAT_DROME 31 SGKVDFLLSVIGFAVDLANVWRFPYLCYKNGGGAFLVPYGIMLAVGGIPL 80
S6A14_HUMAN 90 FFLECSLGQFASLGPVSVW-RILPLFQGVGITMVLISIFVTIYYNVIIAY 138
DAT_DROME 81 FYMELALGQHNRKGAITCWGRLVPLFKGIGYAVVLIAFYVDFYYNVIIAW 130
S6A14_HUMAN 139 SLYYMFASFQSELPWKNCSS-WSDKNC--------SRSPIVTHCNVSTV- 178
DAT_DROME 131 SLRFFFASFTNSLPWTSCNNIWNTPNCRPFESQNASRVPVIG--NYSDLY 178
PFK1_YEAST 616 ----------------------------------TLYCLSHGHKPYAIM 630
S6A14_HUMAN 179 NKGIQEIIQMNK---SWVDINNFTCINGSEI----YQPGQLPSEQYWNK- 220
DAT_DROME 179 AMGNQSLLY---------NETYMNGSSLDTSAVGHVEGFQSAASEYFNRY 219
PFK1_YEAST 631 NGFSGLIQTGEVKELSWIDVENWHNLGGSEI----GTNR---------- 665
S6A14_HUMAN 221 -VALQRSSGMNETGVIVWYLALCLLLAWLIVGAALFKGIKSSGKVVYFTA 269
DAT_DROME 220 ILELNRSEGIHDLGAIKWDMALCLLIVYLICYFSLWKGISTSGKVVWFTA 269
S6A14_HUMAN 270 LFPYVVLLILLVRGATLEGASKGISYYIGAQSNFTKLKEAEVWKDAATQI 319
DAT_DROME 270 LFPYAVLLILLIRGLTLPGSFLGIQYYL--TPNFSAIYKAEVWVDAATQV 317
S6A14_HUMAN 320 FYSLSVAWGGLVALSSYNKFKNNCFSDAIVVCLTNCLTSVFAGFAIFSIL 369
DAT_DROME 318 FFSLGPGFGVLLAYASYNKYHNNVYKDALLTSFINSATSFIAGFVIFSVL 367
S6A14_HUMAN 370 GHMAHISGKEVSQVVKSGFDLAFIAYPEALAQLPGGPFWSILFFFMLLTL 419
DAT_DROME 368 GYMAHTLGVRIEDVATEGPGLVFVVYPAAIATMPASTFWALIFFMMLLTL 417
S6A14_HUMAN 420 GLDSQFASIETITTTIQDLFPKVMKKMRVPITLGCCLVLFLLGLVCVTQA 469
DAT_DROME 418 GLDSSFGGSEAIITALSDEFPKI-KRNRELFVAGLFSLYFVVGLASCTQG 466
S6A14_HUMAN 470 GIYWVHLIDHFCAGWGILIAAILELVGIIWIYGGNRFIEDTEMMIGAKRW 519
DAT_DROME 467 GFYFFHLLDRYAAGYSILVAVFFEAIAVSWIYGTNRFSEDIRDMIGFPP- 515
S6A14_HUMAN 520 IFWLWWRACWFVITPILLIAIFIWSLVQFHRPNYGAIPYPDWGVALGWCM 569
DAT_DROME 516 --GRYWQVCWRFVAPIFLLFITVYGLIGYEPLTYADYVYPSWANALGWCI 563
S6A14_HUMAN 570 IVFCIIWIPIMAIIKIIQAKGNIFQR---LISCCRPASNWGPYLEQHRGE 616
DAT_DROME 564 AGSSVVMIPAVAIFKLLSTPGSLRQRFTILTTPWRDQQSMAMVLNGVTTE 613
S6A14_HUMAN 617 ----RYKDMVDPKKEADHEIPTVSGSRKPE 642
DAT_DROME 614 VTVVRLTDTETAKEPVDV------------ 631

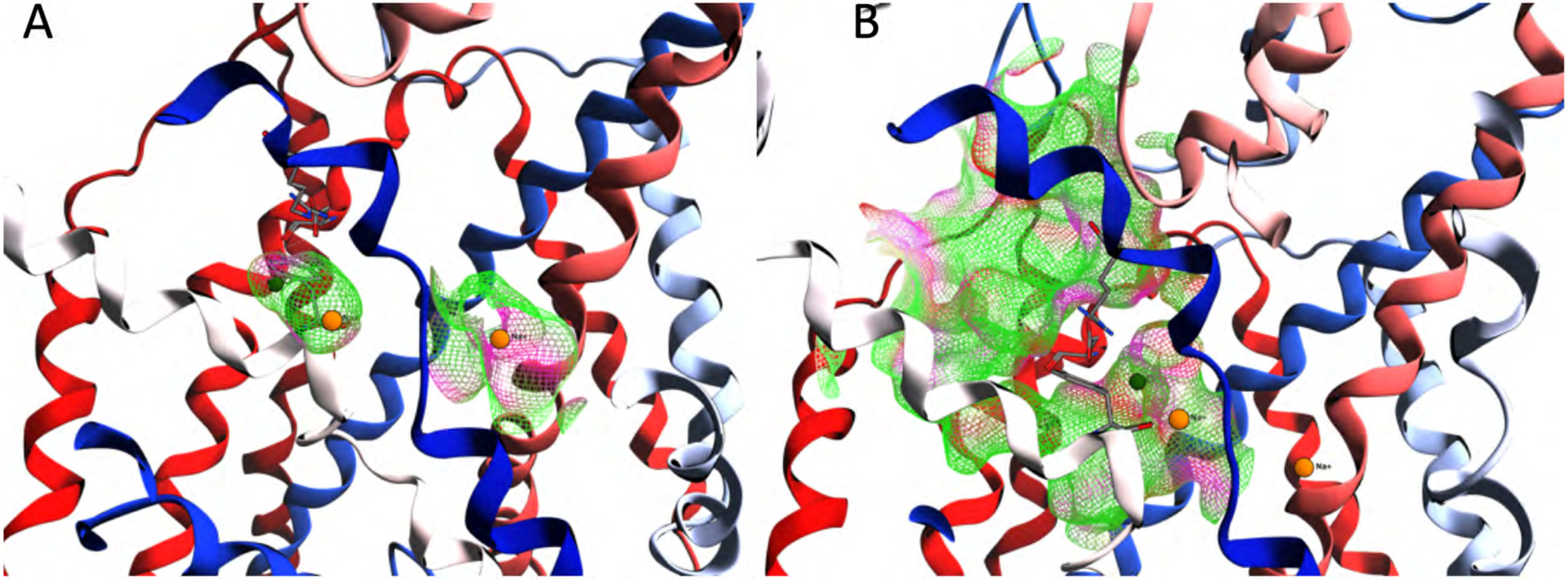

A
PTMs
variants
disulfide bonds
signal peptide
N-term: UniProt
TMRs: UniProt
Extracellular
1
2
3
4
5
6
7
8
9
10
11
12
Cytoplasmic
COOH

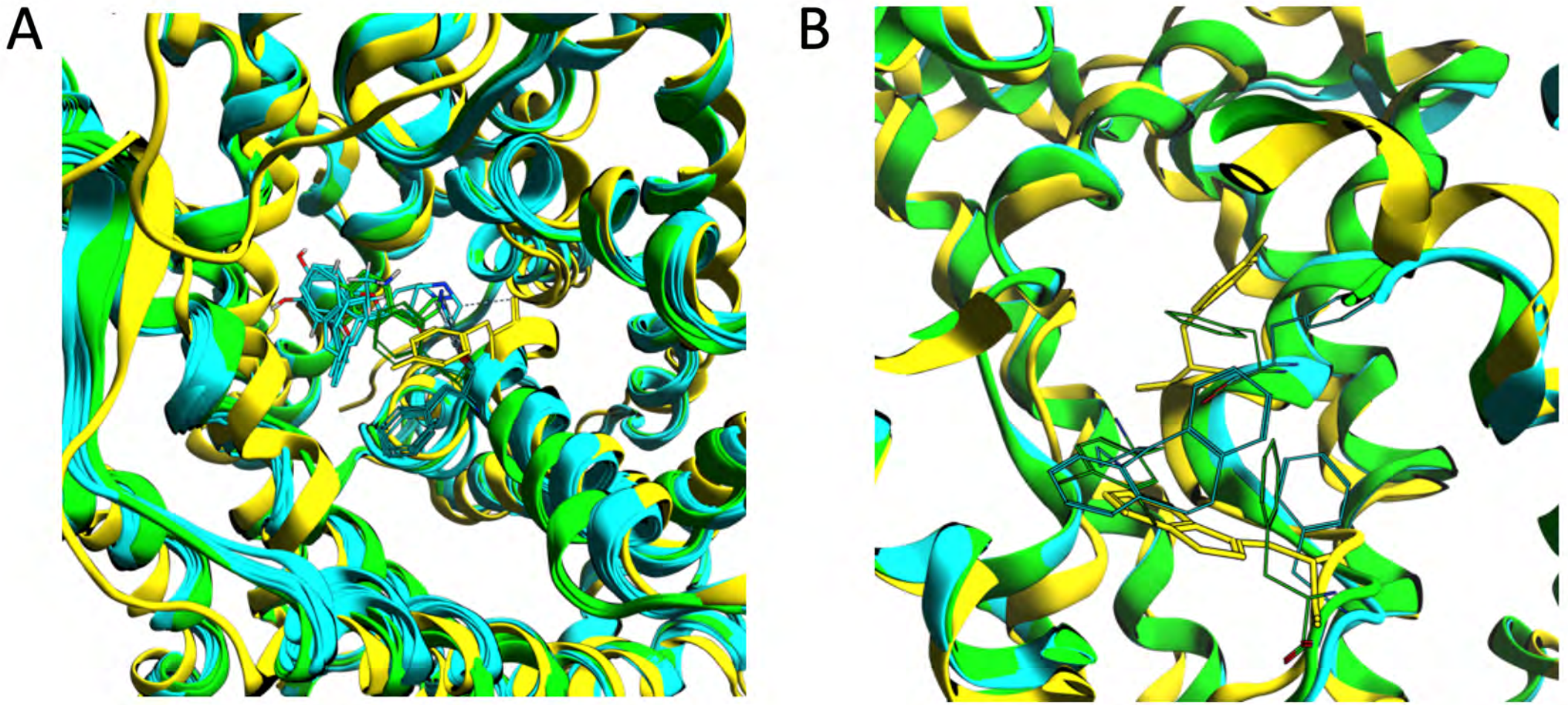

B
TMHMM posterior probabilities for sp|Q9UN76|S6A14_HUMAN
probability
1.2
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
500
600
transmembrane
inside
outside

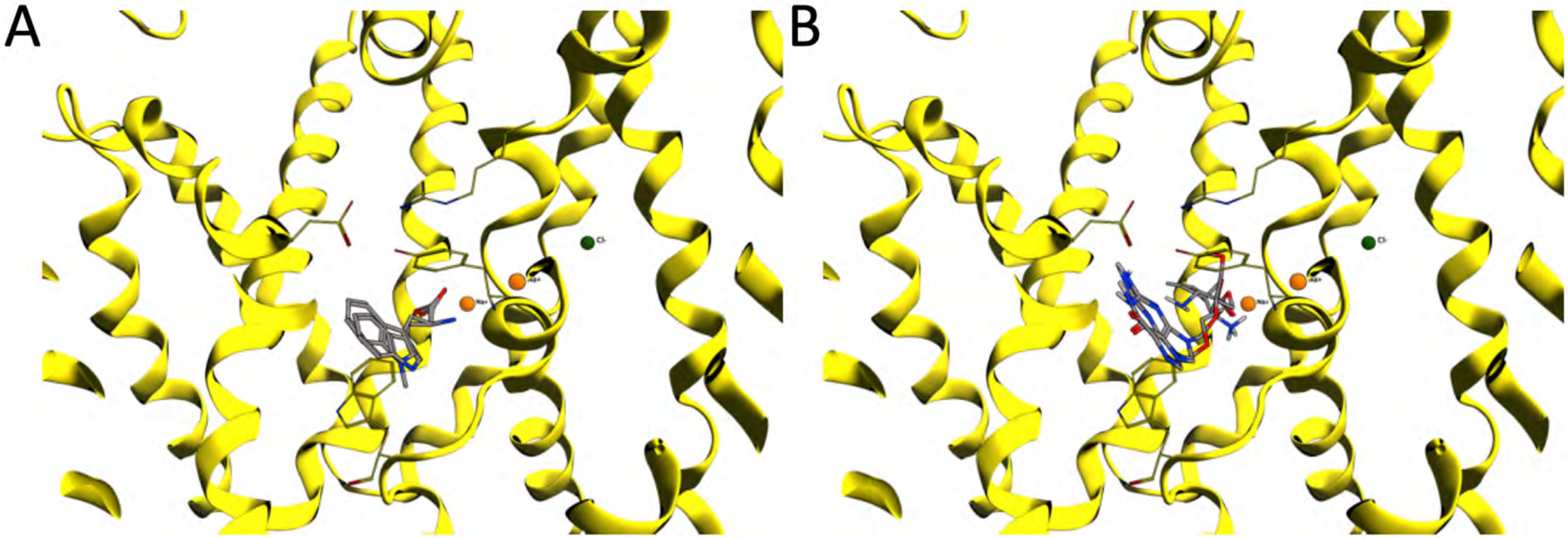

A
Protein RMSD (Å)
Time (nsec)
Cα
B
RMSF (Å)
Residue Index
Cα
C
Res. % SSE
Residue Index

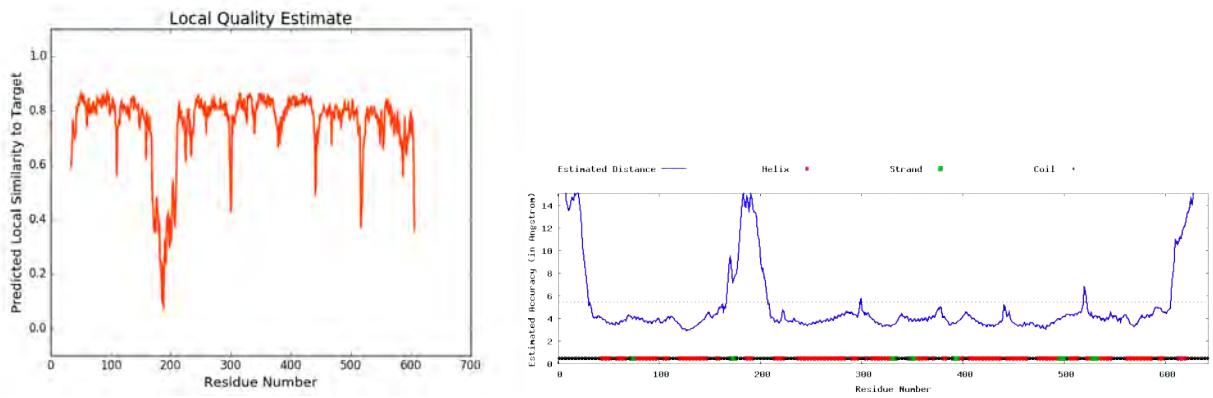

% SSE
60
40
20
0
0
100
200
300
400
500
Residue Index
Time (nsec)
Turn
Strand
Helix
Loop


A
B
C
D

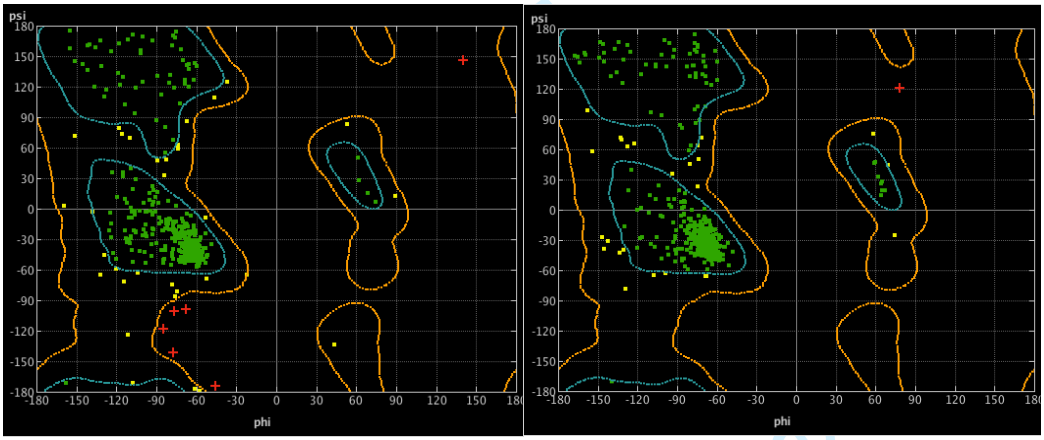

A
B
Na+
Na+
Na+


A
B

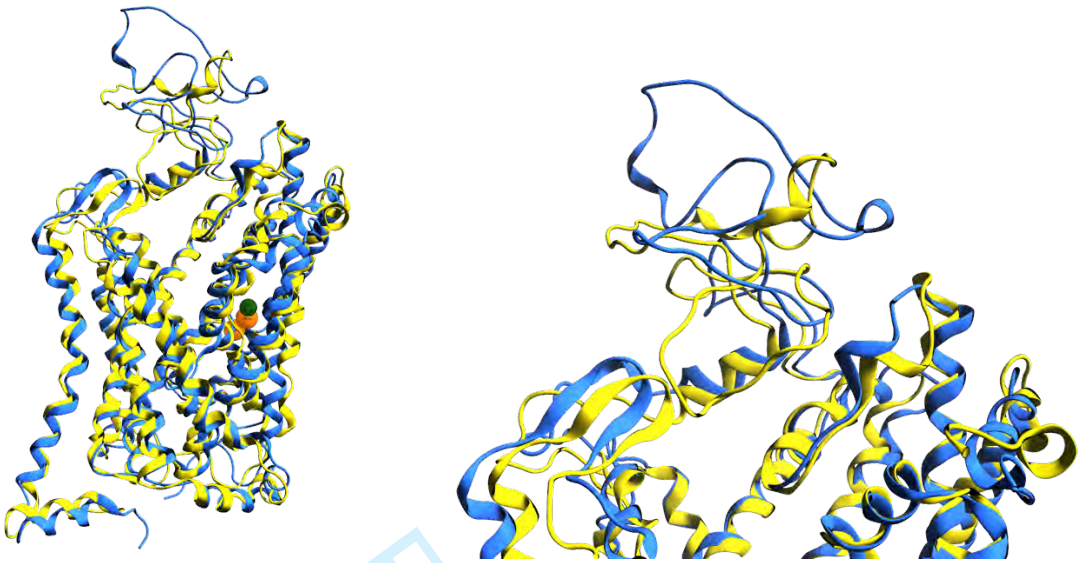

A
B



Supplementary Figure 1: (A) Local quality estimation of Swiss Model results and (B) I-Tasser estimated distance between model and theoretical folding.



Supplementary Figure 2: Ramachandran plot of SLC6A14 chimera model before (A) and after (B) the molecular dynamics equilibration.



Supplementary Figure 3: Superposition of SLC6A14 conformations (A) with a focus on EL2 (B). In yellow is depicted the SLC6A14 structure from homology modelling procedure, before to be submitted to 500 ns MD equilibration; in blue the centroid of the most populated cluster. It is evident the structural rearrangement of EL2.

	Arg	Gln	Gly	1-met-Trp	Phe	Trp
Gly 51						
Tyr 52		WB	WB	H	WB	H
Ala 53		I		I	I	HB
Val 54			WB			
Gly 55		I		I	I	HB
Leu 56		WB	H	HB	HB	HB
Gly 57	I		H	WB	HB	HB
Asn 58					I	I
Trp 60	H				H	
Arg 61			WB		WB	
Ile 124				H		
Val 128	H	H		H	H	
Tyr 131	WB	WB			HB	
Tyr 132	HB	HB	HB	HB	H	WB
Ile 135						
Thr 317	WB	WB				
Gln 318					WB	
Phe 320	HB	HB	WB	HB	WB	HB
Tyr 321	H	I		H	H	I
Ser 322		I	WB	I	I	I
Ser 324	HB	i	H	HB	I	HB
Val 325	HB	WB				
Ala 326	HB	WB				
Trp 327	H	H		H	H	H
Asn 354						I
Phe 388					H	
Leu 419			WB			
Gly 420			WB			
Asp 422						
Ser 423		WB	WB	WB	WB	WB
Gln 424			WB			
Ile 477	WB					
Asp 478	HB	HB			WB	
Cys 481	WB	HB				
Ala 482	WB			H		
Ile 486	H			H		H

I	Ionic
HB	H-bond
H	Hydrophobic
WB	Water Bridge

	Tyr	valaciclovir	valganciclovir	Leu	His	Pro
Gly 51		HB				
Tyr 52	H	H	H	H	H	WB
Ala 53				I	H	I
Val 54		HB		I		
Gly 55						WB
Leu 56	HB	HB	HB	HB		
Gly 57	HB		WB	HB	HB	
Asn 58					I	
Trp 60			WB	H		
Arg 61		WB	H		HB	WB
Ile 124						
Val 128	H	H	H	H	H	
Tyr 131		WB	WB	H		
Tyr 132	H	HB	HB	HB	HB	HB
Ile 135				H		
Thr 317		WB				
Gln 318						
Phe 320	H	WB	WB	WB	HB	HB
Tyr 321	H	H	H	H	I	I
Ser 322				I	I	
Ser 324	HB	HB	HB	WB	I	HB
Val 325	WB	WB	HB			HB
Ala 326	WB				HB	
Trp 327	H	H	H	H	H	H
Asn 354						
Phe 388						
Leu 419			HB			
Gly 420	WB		HB			
Asp 422	H	HB	HB			
Ser 423	HB		HB			WB
Gln 424	WB					
Ile 477		HB	HB			
Asp 478	WB	HB	HB		WB	
Cys 481	WB	HB	HB			
Ala 482		HB	HB		WB	H
Ile 486		H	H		H	H

I	Ionic
HB	H-bond
H	Hydrophobic
WB	Water Bridge

	Tre	Ala	Carnitine	Cys
Gly 51			HB	
Tyr 52		WB	I	WB
Ala 53		I		I
Val 54			HB	
Gly 55		I	HB	WB
Leu 56		HB	HB	
Gly 57		HB	I	
Asn 58				
Trp 60	WB	H	WB	
Arg 61	WB			WB
Ile 124	H		H	
Val 128	H			
Tyr 131	WB	HB	HB	
Tyr 132	HB	HB		HB
Ile 135		H		
Thr 317	WB			
Gln 318				
Phe 320	HB	WB	I	HB
Tyr 321	WB	I	I	I
Ser 322	I	I	HB	
Ser 324	WB	I		HB
Val 325		WB		HB
Ala 326	WB	WB	H	
Trp 327	H	H	I	H
Asn 354				
Phe 388				
Leu 419				
Gly 420				
Asp 422			HB	
Ser 423		WB		WB
Gln 424				
Ile 477				
Asp 478	WB	I		
Cys 481				
Ala 482	WB			H
Ile 486	H			H

I	Ionic
HB	H-bond
H	Hydrophobic
WB	Water Bridge

Supplementary Table 1