

The mechanism of mammalian proton-coupled peptide transporters

Simon M Lichtinger, Joanne L Parker, Simon Newstead, Philip C Biggin 

Department of Biochemistry, University of Oxford • The Kavli Institute for Nanoscience Discovery, University of Oxford

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Reviewed Preprint

v2 • June 19, 2024

Revised by authors

Reviewed Preprint

v1 • April 16, 2024

Abstract

Proton-coupled oligopeptide transporters (POTs) are of great pharmaceutical interest owing to their promiscuous substrate binding site that has been linked to improved oral bioavailability of several classes of drugs. Members of the POT family are conserved across all phylogenetic kingdoms and function by coupling peptide uptake to the proton electrochemical gradient. Cryo-EM structures and alphafold models have recently provided new insights into different conformational states of two mammalian POTs, SLC15A1 and SLC15A2. Nevertheless, these studies leave open important questions regarding the mechanism of proton and substrate coupling, while simultaneously providing a unique opportunity to investigate these processes using molecular dynamics (MD) simulations. Here, we employ extensive unbiased and enhanced-sampling MD to map out the full SLC15A2 conformational cycle and its thermodynamic driving forces. By computing conformational free energy landscapes in different protonation states and in the absence or presence of peptide substrate, we identify a likely sequence of intermediate protonation steps that drive inward-directed alternating access. These simulations identify key differences in the extracellular gate between mammalian and bacterial POTs, which we validate experimentally in cell-based transport assays. Our results from constant-PH MD and absolute binding free energy (ABFE) calculations also establish a mechanistic link between proton binding and peptide recognition, revealing key details underpinning secondary active transport in POTs. This study provides a vital step forward in understanding proton-coupled peptide and drug transport in mammals and pave the way to integrate knowledge of solute carrier structural biology with enhanced drug design to target tissue and organ bioavailability.

eLife assessment

This study provides **important** insight into the mechanisms of proton-coupled oligopeptide transporters. It uses enhanced-sampling molecular dynamics (MD), backed by cell-based assays, revealing the importance of protonation of selected residues for PepT2 function. The simulation approaches are **convincing**, using long MD simulations, constant-pH MD and free energy calculations. Overall, the work has led to findings that will appeal to structural biologists, biochemists, and biophysicists studying membrane transporters.

<https://doi.org/10.7554/eLife.96507.2.sa3>

Introduction

Cells require an external lipid membrane to separate their internal cytoplasm from the environment. Since the membrane permeability of common solutes spans ten orders of magnitude, some molecules diffuse readily across the membrane, while the translocation of others requires facilitation by carriers. ([Stillwell, 2016](#); [Pizzagalli et al., 2021](#)) Understanding the processes by which small molecules cross membranes is of key pharmacological interest owing to their role in drug delivery, which may be mediated by passive diffusion, protein carriers, or a combination of both. ([Sugano et al., 2010](#)) The solute-carrier (SLC) superfamily encompasses 65 families of more than 450 genes, with substrates ranging in size from simple ions to complex macromolecules used in metabolism and signalling. ([Pizzagalli et al., 2021](#)) Within this superfamily, the SLC15 family includes proton-coupled oligopeptide transporters (POTs) that have significant homology through all domains of life and are evolutionarily ancient. ([Daniel et al., 2006](#)) Of the four mammalian family members, PepT1 (SLC15A1) and PepT2 (SLC15A2) are the most well studied. The former is predominantly expressed in the small intestine and characterised as a low-affinity, high-capacity transporter. ([Fei et al., 1994](#)) The latter has a broader expression pattern including the kidneys, lungs and brain and is described as high-affinity, low-capacity. ([Daniel and Kottra, 2004](#)) As secondary-active transporters, they couple uphill substrate translocation to the symport of protons down their electrochemical gradient. ([Fei et al., 1994](#); [Rubio-Aliaga et al., 2000](#)) The peptide–proton stoichiometry is not conserved between different substrates and POT family members. ([Parker et al., 2014](#)) For PepT1, stoichiometries of 1:1 and 2:1 have been reported for neutral / basic and acidic di-peptides, respectively. ([Fei et al., 1994](#); [Steel et al., 1997](#)) For PepT2, a 2:1 stoichiometry was reported for the neutral di-peptide D-Phe-L-Ala and 3:1 for anionic D-Phe-L-Glu. ([Chen et al., 1999](#)) Alternatively, [Fei et al. \(1999\)](#) have found 1:1 stoichiometries for either of D-Phe-L-Gln (neutral), D-Phe-L-Glu (anionic), and D-Phe-L-Lys (cationic). Here, we work under the assumption of a 2:1 stoichiometry for neutral di-peptides, motivated also by our computational results that indicate distinct and additive roles played by two protons in the conformational cycle mechanism.

POT family transporters belong to the major facilitator superfamily (MFS) and share a conserved topology of two six helix bundles that form the functional transport domain, their N- and C-termini facing the cytoplasm. ([Newstead et al., 2011](#)) They operate via an alternating access mechanism encoded in four inverted topology repeats, progressively reorienting the N- and C-terminal bundles to cycle through outwards-facing (OF), occluded (OCC) and inwards-facing (IF) states. ([Radestock and Forrest, 2011](#)) Since the first structure of a POT family member was published ([Newstead et al., 2011](#)), many prokaryotic ([Solcan et al., 2012](#); [Guettou et al., 2013](#); [Doki et al., 2013](#); [Lyons et al., 2014](#); [Guettou et al., 2014](#); [Zhao et al., 2014](#); [Fowler et al., 2015](#); [Boggavarapu et al., 2015](#); [Beale et al., 2015](#); [Parker et al., 2017](#); [Martinez Molledo et al., 2018](#); [Ural-Blimke et al., 2019](#); [Minhas and Newstead, 2019](#); [Stauffer et al., 2022](#); [Kotov et al., 2023](#)) and plant ([Parker and Newstead, 2014](#); [Sun et al., 2014](#)) homologues have been structurally and biochemically characterised, all in IF states with varying degrees of occlusion (see [figure 1a](#) for an overview of available POT structures and their conformational states). Several residues have been suggested to be involved in proton transfer, including a partially conserved histidine on TM2 (H87; residue numbers refer to PepT2, if not specified otherwise) ([Terada et al., 1996](#); [Fei et al., 1997](#); [Chen et al., 2000](#); [Omori et al., 2021](#); [Parker et al., 2021](#)) and two conserved glutamates on TM1 (E53 and E56) ([Jensen et al., 2012](#); [Doki et al., 2013](#); [Aduri et al., 2015](#)), while simulations have helped our understanding of proton-transfer processes and conformational changes. ([Parker et al., 2017](#); [Selvam et al., 2018](#); [Batista et al., 2019](#); [Li et al., 2022](#)) However, the details of the molecular mechanism of alternating access in POTs — particularly regarding the coupling of conformational changes, substrate binding and proton movement to each other — remain unclear.

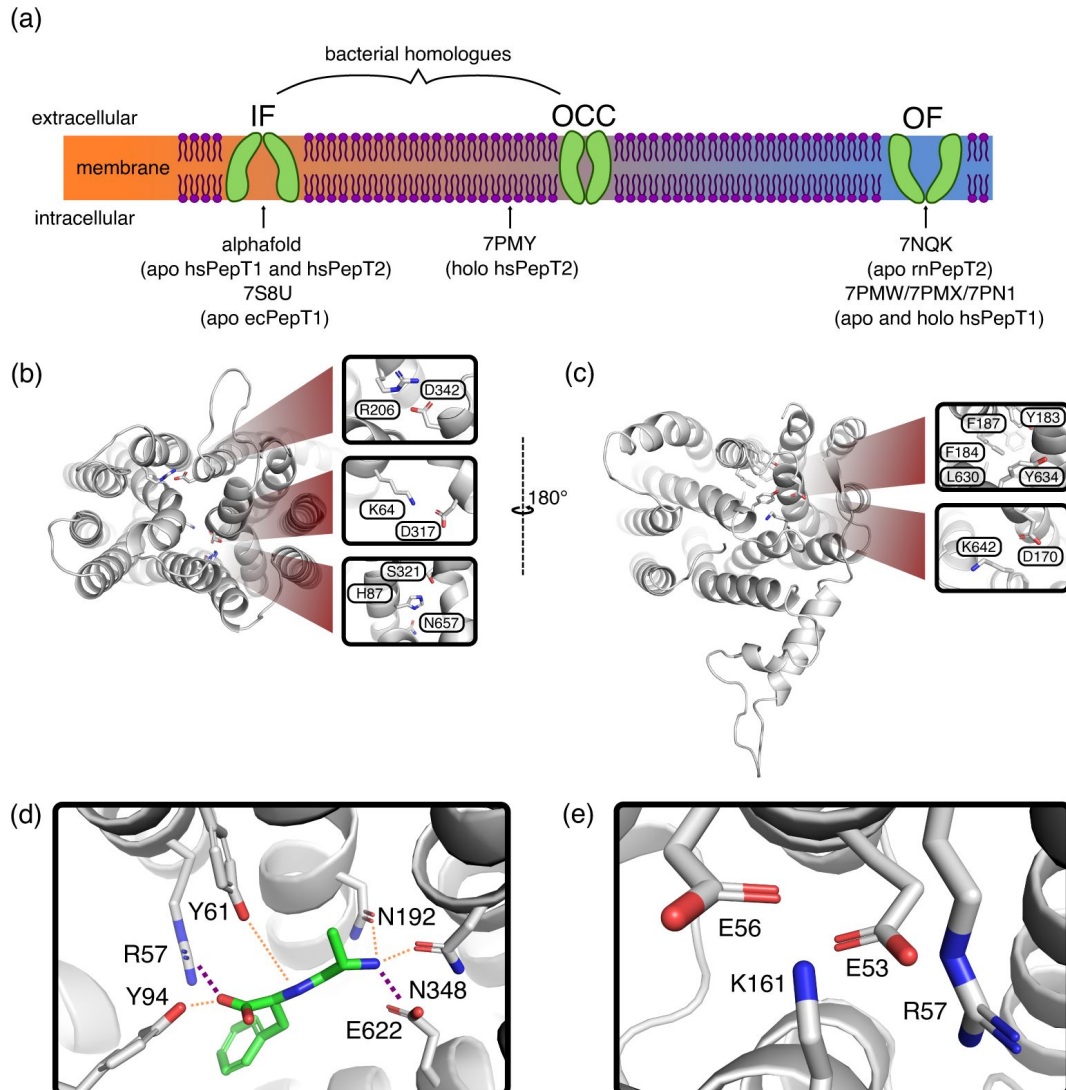


Figure 1.

(a) Schematic overview of the conformational diversity of available mammalian POT structures. Intermediate positions between states indicate partial gate opening. (b) AlphaFold-predicted inwards-facing (IF) HsPepT2 structure (top view), highlighting potential inter-bundle extracellular gate interactions. (c) Outwards-facing (OF) Cryo-EM structure of apo RnPepT2 (7NQK, bottom view) ([Parker et al., 2021](#)), highlighting potential inter-bundle intracellular gate interactions. (d) Ala-Phe substrate binding pose, representative cluster frame of 1 μ s MD simulation from 7NQK structure with added ligand, for setup details see Methods and Materials. Purple dotted lines represent salt-bridge contacts, orange dotted lines other polar contacts. (e) ExxER motif salt-bridge cluster, representative cluster frame of 1 μ s MD simulation from 7NQK structure.

Cryo-EM and AlphaFold 2 have recently provided views of mammalian POTs in conformations spanning from OF via inward-facing, partially occluded to fully-open IF. ([Parker et al., 2021](#); [Killer et al., 2021](#); [Shen et al., 2022](#); [Jumper et al., 2021](#)) From these structures emerges a picture where the intracellular gate is constituted by broad close-packing of hydrophobic residues on TM 4, 5, 10 and 11, with possible stabilisation from the conserved D170–K642 salt-bridge. The extracellular gate appears to be spread along the cleft between the N- and C-terminal bundles, with contributions from the H87 (TM 2) – S321 (TM 7) polar interaction network as well as the R206 (TM 5) – D342 (TM 8) and K64 (TM 1) – D317 (TM 7) salt bridges ([figure 1b](#)). This is intriguing, because the mammalian H87 residue is only conserved in some prokaryotic homologues, and R206–D342 just among mammalian POTs. We speculate based on this feature that the extracellular gating mechanism could be less conserved than POT alternating access in general. As for the intracellular gating mechanism, an involvement of the D170–K642 saltbridge has been suggested, and the OF structure shows close-packing of several hydrophobic residues (F184, Y183, F187, L630 and Y634) that constrict access to the binding site from the intracellular side ([figure 1c](#)). ([Parker et al., 2021](#)) It is not known thus far how the opening of the intracellular gate (i.e., the OCC → IF transition) is triggered, and how it is coupled to proton movement and the presence of substrate.

POTs accommodate their substrates in a highly conserved binding pocket, interfacing between an acidic patch on the C-terminal bundle and a basic patch on the N-terminal bundle ([figure 1c](#)). For di-peptides, the N-terminus is coordinated by E622 (TM 10) together with N192 (TM 5) and N348 (TM 8), while the C-terminus engages R57 (TM 1, or the equivalent lysine residue in PepT1) as well as Y94 (TM 2). Another conserved tyrosine, Y61 (TM 1), hydrogen-bonds to features of the peptide backbone. ([Lyons et al., 2014](#); [Martinez Molledo et al., 2018](#); [Killer et al., 2021](#)) Tri-peptides may adopt a similar orientation as di-peptides ([Guettou et al., 2014](#)), or sit vertically in the transporter binding pocket ([Lyons et al., 2014](#)), although it has been suggested that this vertical electron density could alternatively be explained by a bound HEPES molecule. ([Martinez Molledo et al., 2018](#)) Considering the consensus structural interaction pattern, we decided to investigate primarily the role of E622 and R57 in holding the substrate, and also note that R57 is part of the highly conserved E⁵³xxER motif ([figure 1d](#)). ([Newstead, 2017](#)) The second glutamate in this motif (E56) in particular has been linked to proton coupling experimentally. ([Jensen et al., 2012](#)) Since R57 interacts with both the ExxER glutamates and the substrate C-terminus, we hypothesise that it may play an important role in substrate–proton coupling.

In this study we use extensive unbiased and enhanced-sampling molecular dynamics (MD) simulations (totalling close to 1 ms of sampling) to show how changes in protonation states of H87 and D342 control the OCC ↔ OF transition as an extracellular gate. We validate the importance of these residues for transporter function in cell-based transport assays. We also elucidate the role of E⁵³xxER glutamates and the substrate-engaging E622 in controlling the OCC ↔ IF transition, thereby identifying a clear molecular basis for the directionality of proton movement coupling to conformational changes. Furthermore, we establish several distinct effects of the presence of substrate, coupling ligand binding with protein conformational changes and also linking it to protonation of the E56 and E622 titratable residues. Taken together, our work provides for the first time a detailed model of a plausible sequence of steps for substrate and proton-coupled alternating access in mammalian POTs.

Results

Unbiased MD identifies extra- and intra-cellular gate opening triggers

We began our computational investigation by embedding PepT2 structures (using the sequence of the rat homologue) in the OF ([Parker et al., 2021](#)), IF (alphafold prediction, [Jumper et al. \(2021\)](#)) and inwards-facing partially occluded ([Killer et al., 2021](#)) conformations in 3:1 POPE:POPG membranes (figure S1, details in Materials and Methods). While we were able to obtain stable wide-open OF and IF simulation boxes, the OCC state required further attention as the MD simulations from the inwards-facing, partially-occluded structure showed embedding artifacts, including extracellular-gate instability and intracellular gate hydrophobic collapse (figure S1a). We therefore opted to derive an OCC state using metadynamics simulations in 5 replicates (figure S2, further details in Materials and Methods), using stability in unbiased MD to select the best among the obtained candidates. The OCC state thus developed is validated by the further work in this paper, showing it to be a stable conformational basin that is functionally occluded in that it can open both towards OF and IF in different protonation state conditions (figures 2, 3 and 4).

Equipped with models of the protein conformations required for PepT2 alternating access (OF, OCC and IF), we ran triplicate sets of 1 μ s-long MD simulations in a range of conditions. To decide which conditions to probe apart from the apo, standard protonation states as obtained above, we investigated the extent to which the H87–S321, R206–D342, K64–D317 and D170–K642 interactions (see figure 1b–c) are formed in the OF (closed intracellular gate) and IF (closed extracellular gate) conformational states (figure S3). In the IF state, we found that H87–S321 and D342–R206 are consistently interacting, whereas the K64–D317 interaction, while formed in $\approx 72\%$ of MD frames, is unstable and of a transient nature and therefore unlikely to contribute much to extracellular gate stability. The D170–K642 saltbridge, in turn, is only formed in $\approx 22\%$ of OF-state frames, thus likely not substantially adding to the stability of the intracellular gate. We therefore decided to mainly focus on probing the H87–S321 and D342–R206 interactions with respect to control of the extracellular gate. Since no saltbridges or other specific interactions involving protonatable residues seem to demarcate the intracellular gate, we decided to focus on the ExxER glutamates and E622 for their effects on intracellular gate opening.

Guided thus in our choice of which residues to investigate, we probed whether the OCC state opens spontaneously to OF or IF states in a range of different protonation-state and mutation conditions (as assessed by projection of unbiased MD runs onto intuitive collective variables, or CVs, defined as the centre-of-mass distance between the tips and bases of the N- and C-terminal bundles, see figure 2a). We found that the extracellular gate remains stably closed in triplicates of 1 μ s-long MD when H87 or D342 are protonated individually, but the OCC state can open spontaneously on the simulated time scale to an OF conformation when both are protonated simultaneously (figure 2b; figure S4a for plots of the opposite gate in the same trajectories, showing how flexibility of the intra- and extracellular gates is anti-correlated). A comparable effect is found in the presence of the physiological peptide substrate L-Ala–L-Phe (figure 2b, panels 5–6). We have also tested further combinations of mutations and protonation state changes relating to the putative extracellular gating interactions (D317 protonation and mutations of R206, S321, D342 to alanine, with and without H87 protonation) as well as some control mutants which we did not expect to have an effect (I135L, T202A, Q340A and the saltbridge-swapped mutant R206D & D342R, combined with H87 protonation). Across our 48 $\times$ 1 μ s unbiased MD runs collated in figure S5, we observed 3 full extracellular-gate opening events, in conditions where H87 was protonated and the D342–R206 saltbridge was also disrupted either by D342 protonation or mutation to alanine (D342A). We also saw one partial opening event when in addition to H87 protonation we also mutated S321 to alanine (S321A). The data thus suggests that for spontaneous extracellular-gate opening to occur on this time scale in unbiased MD, disruption of the OCC-state H87 interaction network is essential, and the D342 salt bridge appears to make an additive contribution towards extracellular-gate stability (though this is not a strict correlation, as illustrated by the S321A partial

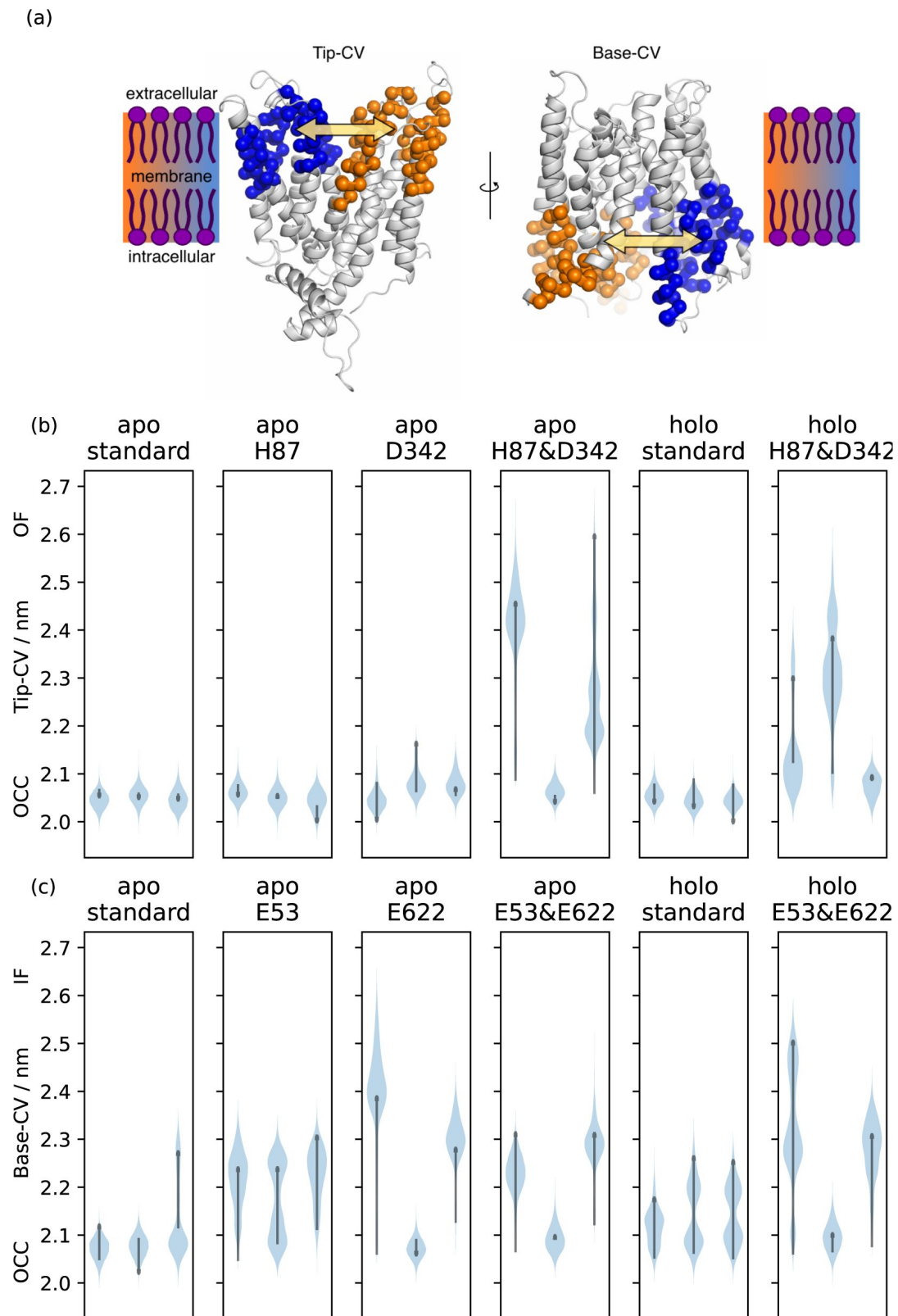


Figure 2.

(a) Illustration of the CVs used to quantify extra- and intracellular gate opening, consisting of inter-bundle centre-of-mass distances between the helical tips (top 11 residues) and bases (bottom 11 residues). (b)+(c) Triplicate 1 μ s-MD simulations starting from OCC, showing the effects of different protonation and substrate binding states, projected onto the (b) Tip-CV and (c) Base-CV respectively. Violin plots are trajectory histograms, arrows link the CV values of the first and last frames.

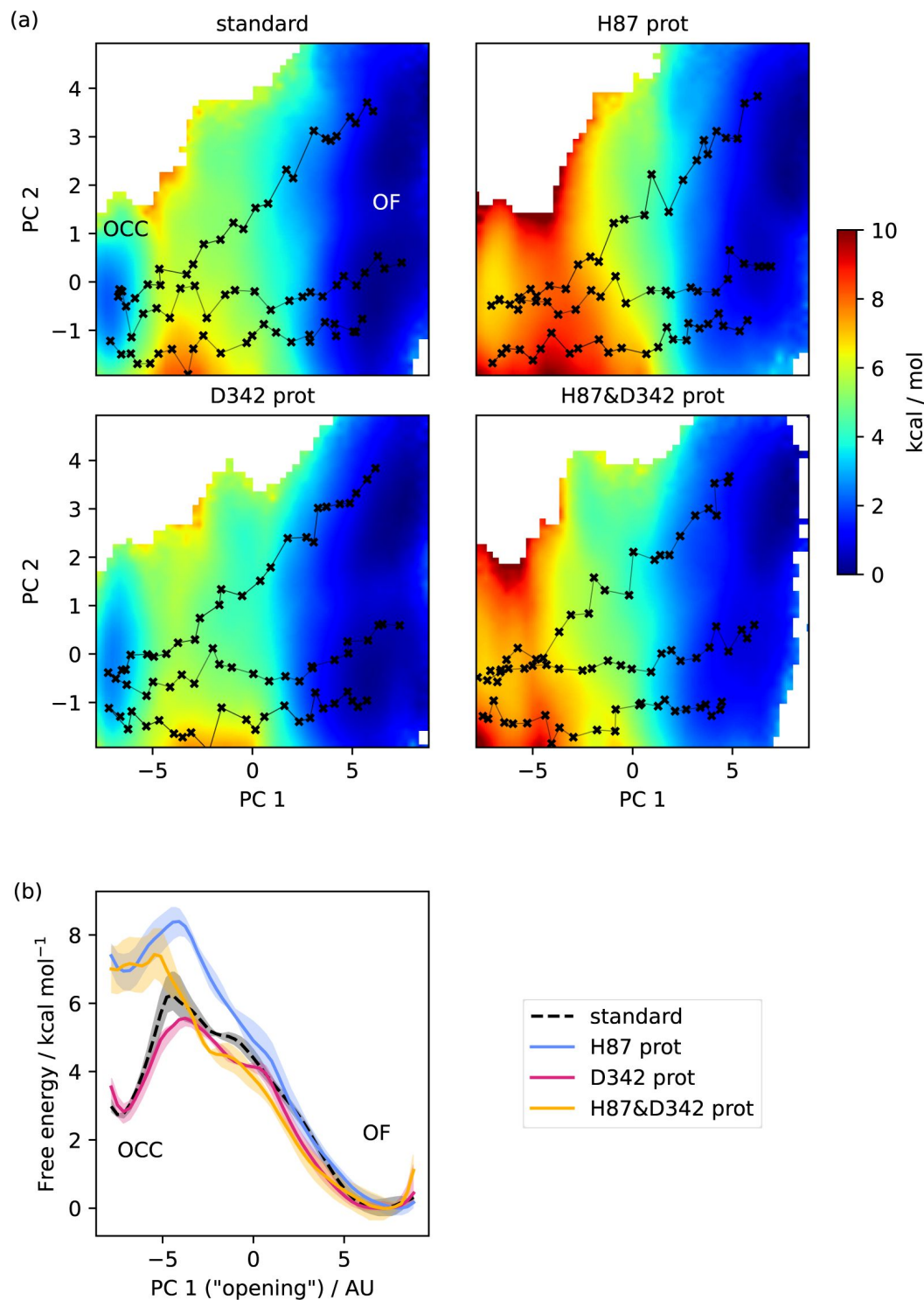


Figure 3.

(a) 2D-PMFs from REUS starting with MEMENTO paths, in different protonation states of candidate extracellular gating residues. (b) Projection of the 2D-PMFs in part a onto PC 1 using Boltzman reweighting. Shaded areas indicate convergence errors as the range of PMF values for a given CV value obtained with the first 40 %, the last 40 % and 100 % of sampling included (after alignment to the 100 % curve). H87 and D342 form an additive extracellular gate, where H87 protonation changes the relative OCC–OF state energies as well as the transition barrier, while D342 protonation only contributes in the transition region. Note that the individual PMFs are only determined by our REUS approach up to additive constants, and are shown aligned here at the OF state for convenience of comparison.

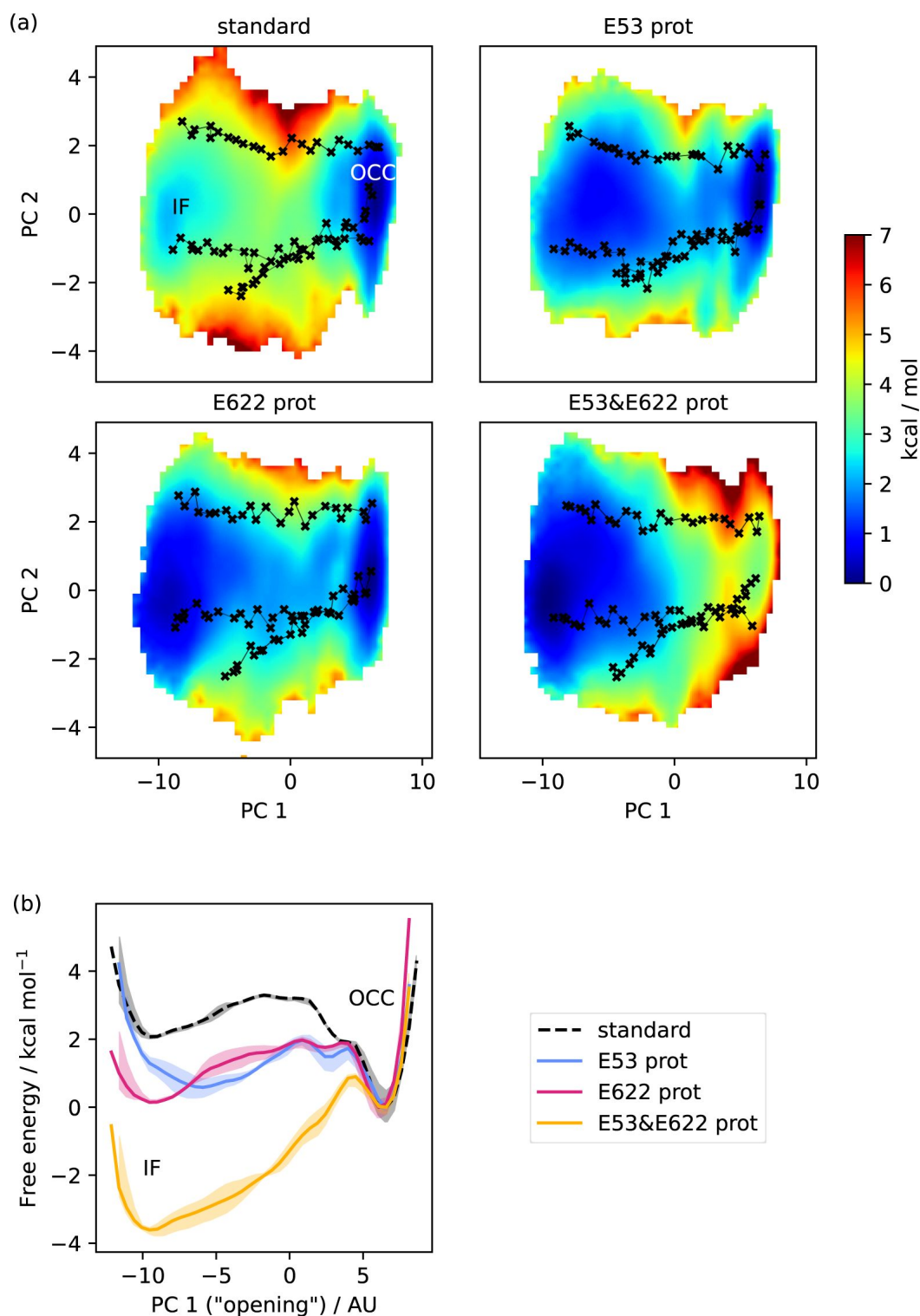


Figure 4.

(a) 2D-PMFs from REUS starting with MEMENTO paths, in different protonation states of candidate intracellular gate-controlling residues. (b) Projection of the 2D-PMFs in part a onto PC 1 using Boltzman reweighting. Shaded areas indicate convergence errors as the range of PMF values for a given CV value obtained with the first 40 %, the last 40 % and 100 % of sampling included (after alignment to the 100 % curve). E53 and E622 protonation have additive and approximately equal effects on driving the OCC → IF transition. Note that the individual PMFs are only determined by our REUS approach up to additive constants, and are shown aligned here at the IF state for convenience of comparison.

opening event). The intracellular gate, by contrast, is more flexible than the extracellular gate even in the apo, standard protonation state; however, following either protonation of the conserved E53 and E622 residues or the insertion of Ala-Phe substrate, the intracellular gate becomes more flexible and can spontaneously open (**figure 2c** [↗](#); see figure S4b for the corresponding plots of the extracellular gate opening).

Although these unbiased simulations show a large amount of stochasticity and drawing clean conclusions from the data is difficult, we can already appreciate a possible mechanism where protons move down the transporter pore, first engaging H87 and D342 to favour the OF state and then moving to ExxER and E622 to favour the IF orientation, driving successive conformational changes along the way. The initial unbiased approach taken in this section therefore serves to generate hypotheses for testing by a more rigorous investigation of the protonation state-dependent conformational changes. To this end, we set out to employ enhanced sampling simulations for obtaining conformational free energy landscapes in dependence on a range of protonation state and substrate binding conditions.

2D-PMFs show proton-dependent driving forces of PepT2 alternating access

To overcome the time-scale limitations of MD simulations and sample important slow degrees of freedom, many enhanced sampling approaches have been developed ([Hénin et al., 2022](#) [↗](#)) and employed in the computational study of membrane proteins. ([Harpole and Delemotte, 2018](#) [↗](#)) An important class of methods that includes (among others) the popular techniques of steered MD (SMD) ([Izrailev et al., 1999](#) [↗](#)), umbrella sampling ([Torrie and Valleau, 1977](#) [↗](#)), metadynamics ([Barducci et al., 2008](#) [↗](#)), adaptive biasing force (ABF) ([Darve et al., 2008](#) [↗](#)) and the accelerated weight histogram method (AWH) ([Lindahl et al., 2014](#) [↗](#)) uses a small number of collective variables (CVs) along which to bias the simulation, thus improving exploration of important regions of phase space if the CV includes the relevant slow degrees of freedom (DOFs). If the CV is not optimal, problems can manifest in the form of hysteresis (starting-state dependence) when moving between known end-states. ([Lichtinger and Biggin, 2023](#) [↗](#)) This is the case for the PepT2 OCC ↔ OF transition with the simple tip-CV illustrated in **figure 2a** [↗](#). Using either metadynamics or SMD with replica-exchange umbrella sampling (REUS), the end-state from which the simulations were started is always favoured in the resulting potential of mean force (PMF), with the hysteresis effect totalling $\approx 40 \text{ kcal mol}^{-1}$ for each method (figure S6).

We have recently developed a strategy to overcome such hysteresis issues in conformational sampling which we call MEMENTO (Morphing Endstates by Modelling Ensembles with independent TOPologies, [Lichtinger and Biggin \(2023\)](#) [↗](#)), and have implemented it as the freely available and documented PyMEMENTO package. MEMENTO generates paths between known end-states by coordinate morphing followed by fixing the geometries of un-physical morphed intermediates. Since these paths by definition connect the correct end-states (unlike biased MD methods like SMD, where not reaching the target state in slow orthogonal DOFs is a common source of hysteresis), they can drastically reduce hysteresis in enhanced sampling compared to SMD as a path generation method. We have shown this for several validation cases, including a large-scale conformational change in the bacterial membrane transporter LeuT. After running initial 1D-REUS from MEMENTO replicates along a simple CV guess, we can use the generated MD data to iteratively improve CVs using principal component analysis (PCA), thereby capturing slow motions from long end-state sampling that propagate through MEMENTO as differences between replicates.

Here, we ran triplicates of MEMENTO for the OCC ↔ OF (standard protonation states and H87&D342 protonated) and OCC ↔ IF (standard protonation states and E53 protonated) conformational changes, followed initially by 1D-REUS along the tip-CV (figure S7). The results are much more consistent than SMD or metadynamics along the same CV, and the shapes of the PMFs

fit well with the trends we previously observed in unbiased MD from the OCC state. Since, however, significant differences between replicates remained, we used principal component analysis (PCA, see Materials and Methods for details) on the sampling collected of the standard protonation state transitions to derive sets of 2-dimensional CVs (figure S8 and supplementary videos 1–4) that capture the main gate-opening motions in the first PC, and the direction along which the differing replicates can be best separated out as the second PC (these correspond to cleft sliding and twisting motions).

Equipped with these CVs, we first studied the protonation-state dependence of the OCC $\leftrightarrow$ OF transition. As [figure 3](#) shows (2D-PMFs in part a, projections onto PC 1 in part b), the OCC state in standard protonation states forms a basin that is metastable with respect to OF (lying ≈ 3 kcal mol⁻¹ higher than OF, separated by a barrier of ≈ 3 kcal mol⁻¹). Protonation of H87 still leads to a metastable OCC basin, although it is raised by ≈ 4 kcal mol⁻¹ and the barrier is decreased to ≈ 1.5 kcal mol⁻¹. Protonating D342, in turn, does not affect the relative free energies of the OCC and OF states, but does lower the transition barrier by ≈ 1 kcal mol⁻¹. These effects are additive, so that protonation of both H87 and D342 abolishes the metastable OCC state — an observation which agrees with the ability of the OCC state thus protonated to spontaneously transition to OF in unbiased MD (see the Discussion for a comparison with the results obtained by [Parker et al. \(2017\)](#) on PepT₅₀ on this point).

We have also computed 2D-PMFs in further protonation-state and mutation conditions to gain a better understanding of how the H87 and D342–R206 interaction networks control the extracellular gate (figure S9 for the 2D-PMFs, all 1D-reprojections are shown for reference in figure S10). From the data presented thus far, it is not clear whether the effect of H87 pronation on the OCC $\rightarrow$ OF transition is due merely to the loss of hydrogen-bond interactions with S321, or whether the introduction of a positive charge in this location makes a significant mechanistic contribution. To address this question, instead of protonating H87 we mutated it to alanine (H87A). In the resulting PMF, the OCC state is raised less with respect to OF compared to the protonated version, and the transition barrier increases to ≈ 5 kcal mol⁻¹, suggesting that there exists an interaction made by positively charged H87 that becomes particularly relevant in the transition region. Further analysis of the H87 interaction networks in our 2D-REUS trajectories (figure S11a) reveals that when H87 is protonated, the interaction with S321 is substituted by an interaction with D317 that is strongest in the transition region. Interestingly, this suggests an alternative mechanistic role for the essential D317, which — as discussed above — we have not found forming the structurally observed saltbridge with K64 consistently in our simulations.

An equivalent investigation of the D342A mutation results in a PMF that shows both a decrease in the transition barrier, and — as opposed to D342 protonation — also raises the OCC state in energy with respect to OF. This may be explained by the fact that protonated D342 can still hydrogen-bond with R206, so although the interaction is less prominent it presumably still contributes somewhat to OCC state stability (figure S11b). As a control, we also show that although the protonation of the ExxER glutamates (E53 and E56) can affect the relative OCC vs OF free energies, there is no lowering of the transition barrier (while they do have an effect on the transition barrier separating OCC from IF, as discussed below). Convergence analysis and representative 2D-REUS histograms for our OCC $\leftrightarrow$ OF PMFs can be found in figures S12 and S13.

We next employed an equivalent approach to investigate the OCC $\leftrightarrow$ IF transition. As can be seen from [figure 4](#), in standard protonation states, the OCC state forms a well-defined basin, connecting to a broader and shallower IF basin raised ≈ 2 kcal mol⁻¹ over OCC via a barrier of just over ≈ 3 kcal mol⁻¹. When either E53 or E622 are protonated, IF drops to a similar free energy level as OCC and the transition barrier lowers by ≈ 1 kcal mol⁻¹. These effects are additive, so that when both E53 and E622 are protonated, the IF state is favoured by ≈ 3.5 kcal mol⁻¹, and is accessible from OCC via a barrier of only ≈ 1 kcal mol⁻¹. As for the OCC $\leftrightarrow$ OF transitions, these results explain the behaviour we had previously observed in the unbiased MD of [figure 2c](#). The

stochastic partial intracellular gate opening seen with those runs can be rationalised through the lower transition barrier from OCC to IF compared to the transition to OF in our PMFs, together with the broad and flat shape of the IF-state basin. Additionally, we have also computed all the equivalents of these PMFs in the presence of Ala-Phe substrate (figure S14), which we will discuss in the section below. All projections onto PC 1 are shown in figure S15, and convergence analysis is provided in figure S16. Taken together, the PepT2 apo 2D-PMFs provide a detailed view of the way the alternating access cycle is driven by proton movement from the extracellular to the intracellular side of the transporter that fits well with its biological function of using two protons from the extracellular medium to energise cycling from OF to IF, and back spontaneously once the protons have left (see the Discussion below).

Substrate coupling of alternating access includes several distinct mechanisms

Given the evidence presented so far, which provides a plausible model for how protons drive alternating access based just on an investigation of the apo states, it remains unclear how coupling of proton transport to the substrate is achieved — that is, why the transporter cannot be driven by protons without the presence of substrate (and would thus just leak protons). To investigate the mechanism underpinning peptide–proton coupling, we constructed simulation boxes that included a bound Ala-Phe substrate molecule (see Materials and Methods for details). We then calculated the Ala-Phe affinity using absolute binding free energy (ABFE) simulations in different protein states (**table 1**). Firstly, we observed that the affinity is similar in the OF and IF states, indicating that the binding of substrate alone does not thermodynamically drive the transporter from OF to IF. We did find, however, that protonation of E622 (i.e., the saltbridge partner of the substrate N-terminus) significantly decreases substrate affinity. Given that protonation of E622 also favours the OCC → IF transition, we suggest a dual function of E622 protonation that includes both stabilising the IF state with respect to OCC and facilitating substrate release from holo IF.

To explore what effect the substrate has on the PepT2 conformational landscape, we repeated a set of our 2D-PMFs in the presence of substrate (**figure 5**). For both the OCC ↔ OF and OCC ↔ IF transitions, the 2D-PMFs have similar shapes in the apo and holo states. For OCC ↔ OF, however, we found an increased width of the OCC state basin in the direction of PC 2, which — as is evident after projection onto PC 1 — stabilises OCC with respect to OF by $\approx 1 \text{ kcal mol}^{-1}$ as an effect of increased OCC flexibility in the orthogonal DOF. We find a similar stabilisation when E53 or E56 are protonated, but not when both H87 and D342 are protonated (figure S9 and S10). This indicates that the presence of substrate in conformations approaching the OCC state from OF may trigger proton movement further down into the transporter — driven by entropic gains from increased flexibility in orthogonal DOFs.

The OCC ↔ IF PMF also presents a broader OCC basin in the presence of substrate, this time in the direction of PC 1 (consistent with the higher OCC-state flexibility directed towards IF observed in unbiased MD, see **figure 2c**). In the projection onto PC 1, this manifests as a broader free energy well and a lower barrier towards IF, even if the relative energies of the basins are not significantly affected. This effect is similarly apparent when E53 is protonated, but not with E622 protonation, which instead leads to a raised transition barrier (these opposing trends continue to be additive when both residues are protonated; figure S14 and S15). We reason that this is due to a more flexible substrate orientation (disengaging the N-terminus) when E622 is neutral. While, taking these observations together, substrate binding does lead to an OCC → IF bias, it also seems unlikely that E622 is protonated at the moment when the OCC → IF transition happens in the holo transporter. This may suggest a further intermediate protonation step that our simulations have not captured (see our Discussion below).

Protein state	AF dipeptide affinity / kcal mol ⁻¹
OF	8.0 ± 0.3
IF	7.0 ± 0.4
IF & E622 prot	2.9 ± 0.2

Table 1.

Results of ABFE calculations, showing that the affinity of Ala-Phe substrate does not depend much on the conformational state (OF vs IF), but is significantly decreased on E622 protonation.

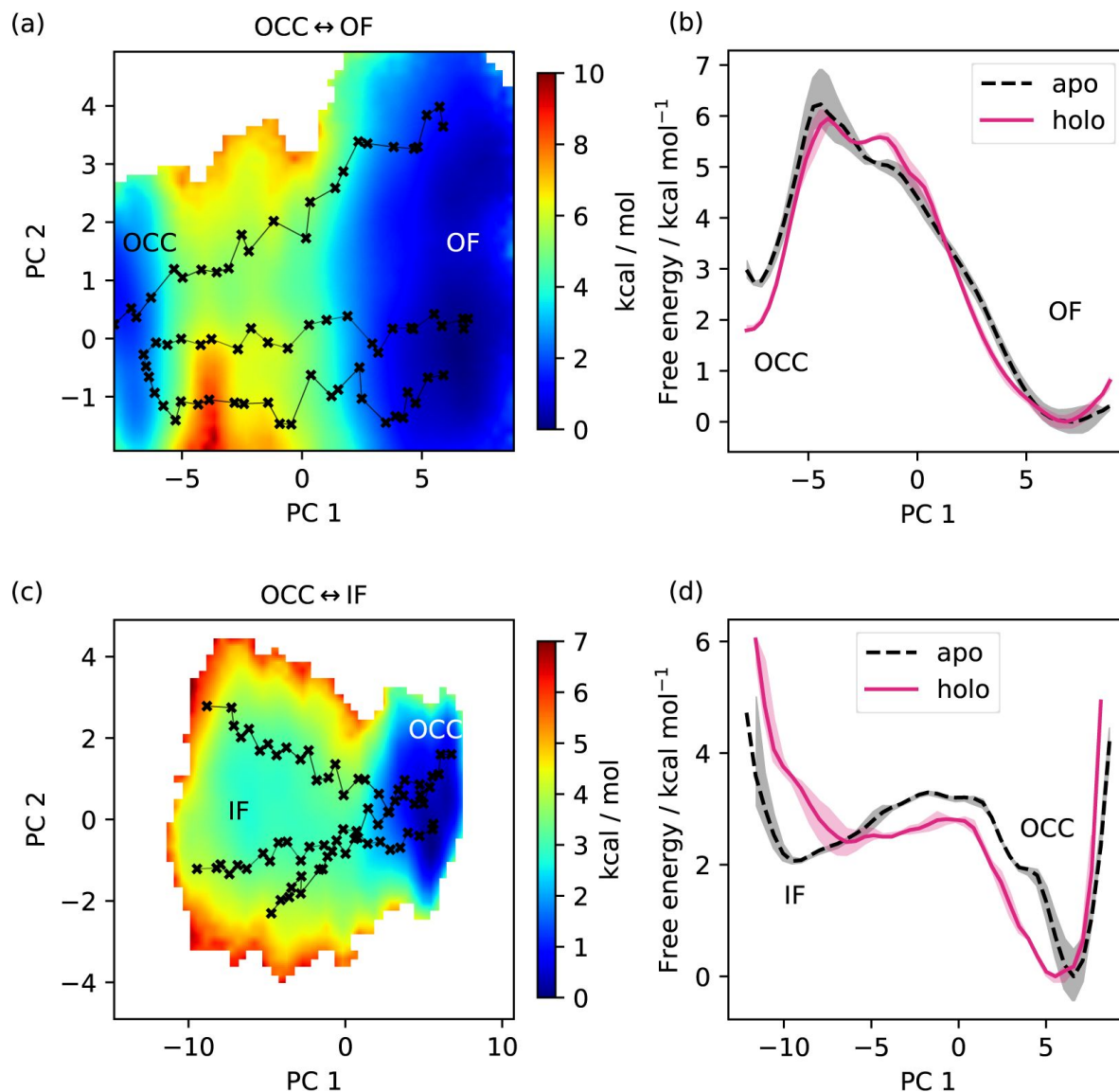


Figure 5.

(a) 2D-PMF for the OCC ↔ OF transition from REUS starting with Ala-Phe-bound PepT2 MEMENTO paths. The OCC state has an increased basin width in PC 2 (compare for [figure 3a](#)), and a transition path shifted in PC 2. (b) Projection of the PMF from panel a onto PC 1, showing how in holo PepT2, the OCC state is stabilised by $\approx 1 \text{ kcal mol}^{-1}$. Shaded areas indicate convergence errors as the range of PMF values for a given CV value obtained with the first 40 %, the last 40 % and 100 % of sampling included (after alignment to the 100 % curve). Note that the individual PMFs are only determined by our REUS approach up to additive constants, and are shown aligned here at the OF state for convenience of comparison. (c) 2D-PMF for the OCC ↔ IF transition from REUS starting with Ala-Phe-bound PepT2 MEMENTO paths. The structure of the IF plateau is not significantly affected, but OCC is more flexible in PC 1. (d) Projection of the PMF from panel c onto PC 1, showing how in holo PepT2, the OCC state has a broader basin, corresponding to intracellular-gate flexibility. Convergence error and alignment of PMFs is shown as in panel b.

As noted above, another possibility for the substrate to engage with the transport cycle is found in R57 interacting both with the substrate C-terminus and with the ExxER glutamates, protonation of which drives the transporter towards the IF state, as our PMFs demonstrate. A natural hypothesis then is that substrate binding — which engages R57 — loosens the R57 interaction with E53 and/or E56, thus enabling the protonation of those residues and progress along the alternating access cycle. To test this hypothesis, we conducted triplicate constant-pH simulations (CpHMD) ([Swails et al., 2014](#)) to probe the E53 and E56 pKa values in all combinations of OCC / OF and apo / holo conditions ([figure 6a](#)). Concerning the E53 pKa value, we see a potential response to substrate binding in the OF state (though the error bars calculated from the triplicate standard deviations overlap) but not in OCC. On the other hand, we do see a raising of the E56 pKa beyond error in holo OF or OCC states compared to apo, amounting to ≈ 0.6 log units in both cases. As shown in [figure S17](#) and [S18](#), the pKa values estimated for successive data chunks across the CpHMD trajectories vary significantly with simulation time in a complex superposition of timescales, and with a dynamic range larger than the replicate error bars. As an alternative to the replicate-based representation in [figure 6](#), we have therefore also analysed the pooled data for each condition as histograms of pKa values estimated from short chunks of our simulations ([figure S19](#)). From this, we recover the same effect of substrate binding on the E56 pKa in the OF and OCC states, as well as a potential effect on the E53 pKa in the OF state only. Since the shift in E56 pKa was more robust across conformational states, we focus on this residue in the following validation and discussion, although we note that if there was in fact a significant raising of the E53 pKa as well, this would further strengthen our conclusions about substrate coupling to ExxER motif protonation.

It is important to recognise here that the hybrid-solvent CpHMD method as implemented in AMBER is not rigorous for membrane proteins, since the membrane is not taken into account while evaluating proposed protonation state changes in implicit solvent. On the other hand, we were not able in initial trials to obtain sufficient transition counts to converge an alternative explicit-solvent CpHMD method as implemented in GROMACS. ([Aho et al., 2022](#)) To validate our results, we therefore also constructed a thermodynamic cycle of substrate binding and E56 protonation in the OF state by including data from separate ABFE calculations as connecting thermodynamic legs ([figure 6b](#)). The ABFE results show a complementary effect of the E56 protonation state on substrate affinity, closing the thermodynamic cycle and validating our CpHMD simulations using an orthogonal MD technique. We conclude that substrate binding does indeed facilitate protonation of E56 (whence, we stipulate, the proton moves to E53, which is situated in close proximity). If the neglected membrane environment in the hybrid-solvent CpHMD did produce significant artifacts in the pKa values, then it would appear that there is error cancellation when assessing the impact of substrate binding as a difference of pKa values in the apo and holo conditions.

It should be noted that — as throughout this study, see the discussion below — in studying the coupling between substrate binding and protonation-state changes at the ExxER motif we have not made the substrate C-terminus protonatable. Since, in order to induce E56 protonation, the substrate C-terminus needs to engage R57 in a salt bridge, its pKa is likely to be low, rendering the assumption reasonable for those substrate conformations. However, it is possible that the system could also adopt states in which the ExxER motif salt-bridge network is stable in a way similar to the apo condition while the substrate gets protonated when oriented away from R57. If such conformations contribute significantly to this semi-grand canonical ensemble, the ExxER motif pKa values estimated without taking them into account may exhibit some bias. By undersampling more apo-like conformations in the holo state in this way, it is possible that the calculations presented here overestimate the substrate-induced pKa shift of E56, although the direction of change would be expected to be the same, because the substrate can still engage R57 when it is deprotonated (we speculate that the histograms in [figure S19](#) for the holo state may become bi-modal in this case). While the possibility would need to be taken into account for a more rigorous quantitative estimate of the ExxER pKa value shifts, the pKa calculations would become much

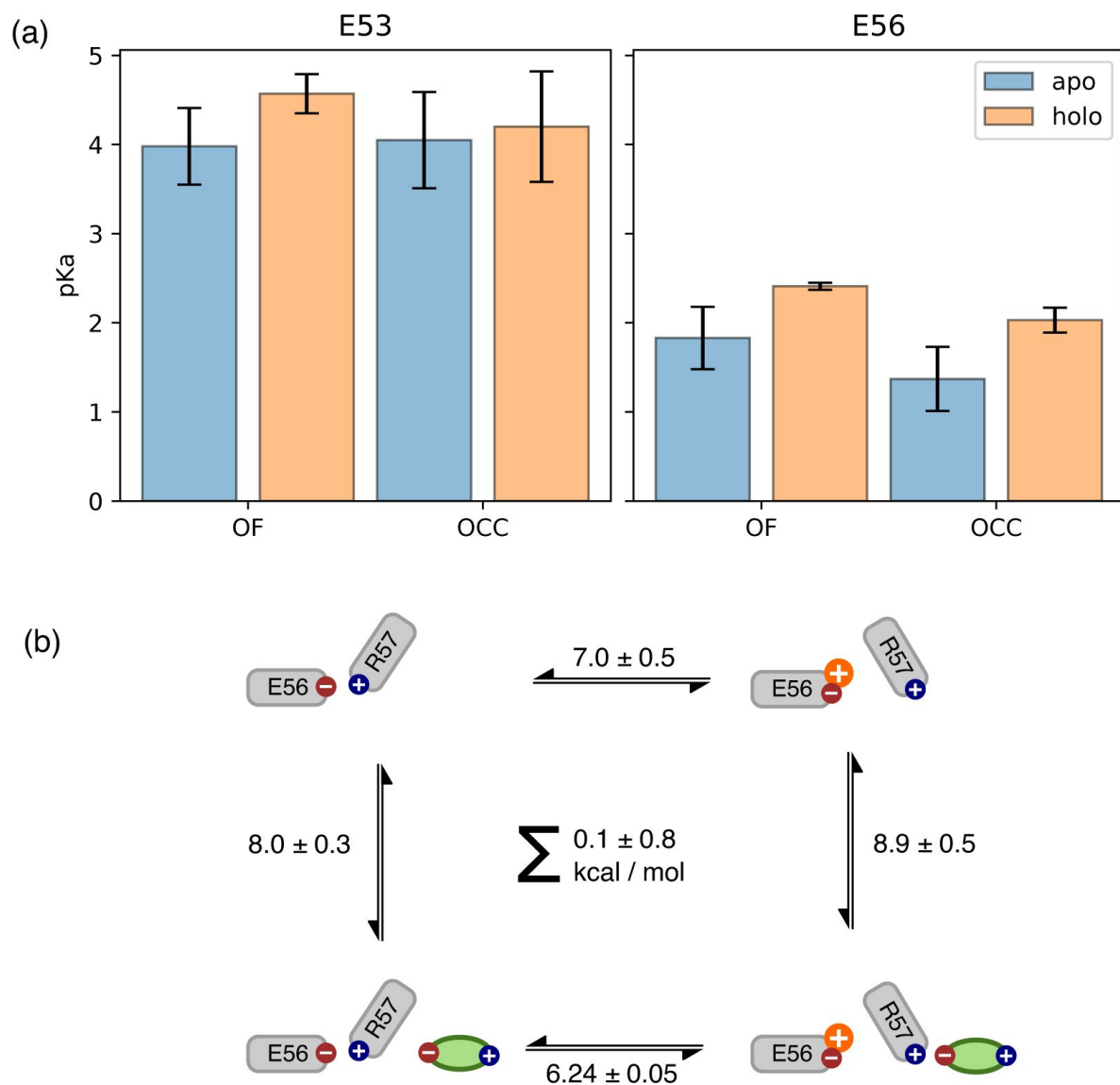


Figure 6.

(a) E53 and E56 pKa values from constant-pH MD simulations, in the apo and holo as well as the OF and OCC states, estimated as mean $\pm$ standard deviation from triplicate runs (using the full simulation data for fitting the titration curves). The presence of substrate raises the E56 pKa in either conformational state, while some effect on the E53 pKa may also exist in the OF state. (b) Illustration of a thermodynamic cycle of E56 protonation and Ala-Phe binding, with edges filled in via CpHMD (converted into kcal mol⁻¹ at pH 7) for the top and bottom transitions, and ABFE for the left and right edges. Notably, ABFE displays a response of Ala-Phe affinity to E56 that is consistent with the CpHMD results, and the cycle closes very well. The error in the cycle closure residual is estimated as a square root of the sum of squared standard deviations of the individual edges.

harder to converge since the slow degree of freedom of substrate re-orientation would need to equilibrate to the protonation-state changes (which happen fast since they are treated as Monte-Carlo moves in an implicit solvent at regular intervals). Here, we contend ourselves with the more qualitative observation that an appropriately positioned substrate in the canonical, structurally observed binding pose facilitates ExxER-motif protonation.

Validation in cell-based transport assays

To experimentally validate the results of an MD investigation, a first step is to probe the importance of the key implicated residues for protein function. We note that the literature already contains ample data to show that E53 ([Solcan et al., 2012](#); [Doki et al., 2013](#); [Sun et al., 2014](#); [Jørgensen et al., 2015](#); [Parker et al., 2017](#)), E56 ([Solcan et al., 2012](#); [Jørgensen et al., 2015](#); [Sun et al., 2014](#); [Zhao et al., 2014](#); [Parker et al., 2017](#)), R57 (or the equivalent lysine) ([Solcan et al., 2012](#); [Guettou et al., 2013](#); [Doki et al., 2013](#); [Jørgensen et al., 2015](#); [Lyons et al., 2014](#); [Parker and Newstead, 2014](#); [Sun et al., 2014](#); [Parker et al., 2017](#); [Martinez Molledo et al., 2018](#)), H87 (for those homologues which conserve it) ([Fei et al., 1997](#); [Chen et al., 2000](#); [Newstead et al., 2011](#); [Uchiyama et al., 2003](#); [Parker et al., 2017](#)), D317 (or the equivalent glutamate) ([Solcan et al., 2012](#); [Doki et al., 2013](#); [Lyons et al., 2014](#); [Parker et al., 2017](#); [Martinez Molledo et al., 2018](#); [Shen et al., 2022](#)) and E622 ([Solcan et al., 2012](#); [Guettou et al., 2013](#); [Doki et al., 2013](#); [Lyons et al., 2014](#); [Zhao et al., 2014](#); [Parker et al., 2017](#); [Minhas and Newstead, 2019](#); [Shen et al., 2022](#)) are important for transport through POTs. Of the residues implicated by our simulations, therefore, only S321 and D342 have not been studied before, and thus serve as predictive validation test cases here.

Using cell-based transport assays (see Materials and Methods), we tested the transport activity of rat PepT2 WT and several mutants: H87A (as a positive control known from the literature), I135L (as a negative control, without any expected effect), as well as the mutants of interest D342A and S321A ([figure 7](#), and [figure S20](#) for loading control and membrane localisation micrographs). We note that all our mutants expressed slightly less compared to the WT at the same amount of transfected DNA (0.8 μ g), but more than WT at a reduced transfection DNA level (0.5 μ g) ([figure 7b](#)). To control for this difference in expression levels, we took WT (0.5 μ g), which transports $\approx 20\%$ less than WT (0.8 μ g), as a lower bound for the WT transport activity and as the point of comparison for statistical tests. We found that all mutants of residues predicted to be involved in transport displayed significantly reduced transport activity (p-values: 2.2×10^{-5} for H87A, 1.6×10^{-4} for D342A, 6.9×10^{-5} for S321A, while I135L is indistinguishable from WT at $p = 0.79$). We also note that D342A, although its activity is significantly reduced, still transports more than H87A ($p = 3.9 \times 10^{-7}$). This fits well with our 2D-PMF results, where H87 protonation does more than D342 protonation to stabilise OF with respect to OCC.

Discussion

Integrating the results from extensive sampling across several MD methods, covering all stages of the PepT2 alternating access cycle, we are now in a position to propose a detailed molecular mechanism of the complete transport cycle, including accounts both of proton coupling to conformational changes and of substrate–proton coupling ([figure 8](#)). Starting at the apo OCC state without any protons bound, we find in our 2D-PMFs that proton binding to H87 and D342 stabilises OF with respect to OCC (with H87 being the major contributor). Given that H87 and D342 are accessible from the acidic extracellular bulk, and in light of the transition-region stabilisation from the H87 (protonated)–D317 interaction we have identified, our simulations suggest an interpretation where protonation happens in the OCC state, driving the OCC $\rightarrow$ OF transition by stabilising the OF state over OCC. However, if the OCC $\rightarrow$ OF transition is kinetically accessible on experimental timescales without prior protonation events (beyond what our MD was able to sample), it would also be consistent with our data that OCC $\rightarrow$ OF is spontaneous in standard

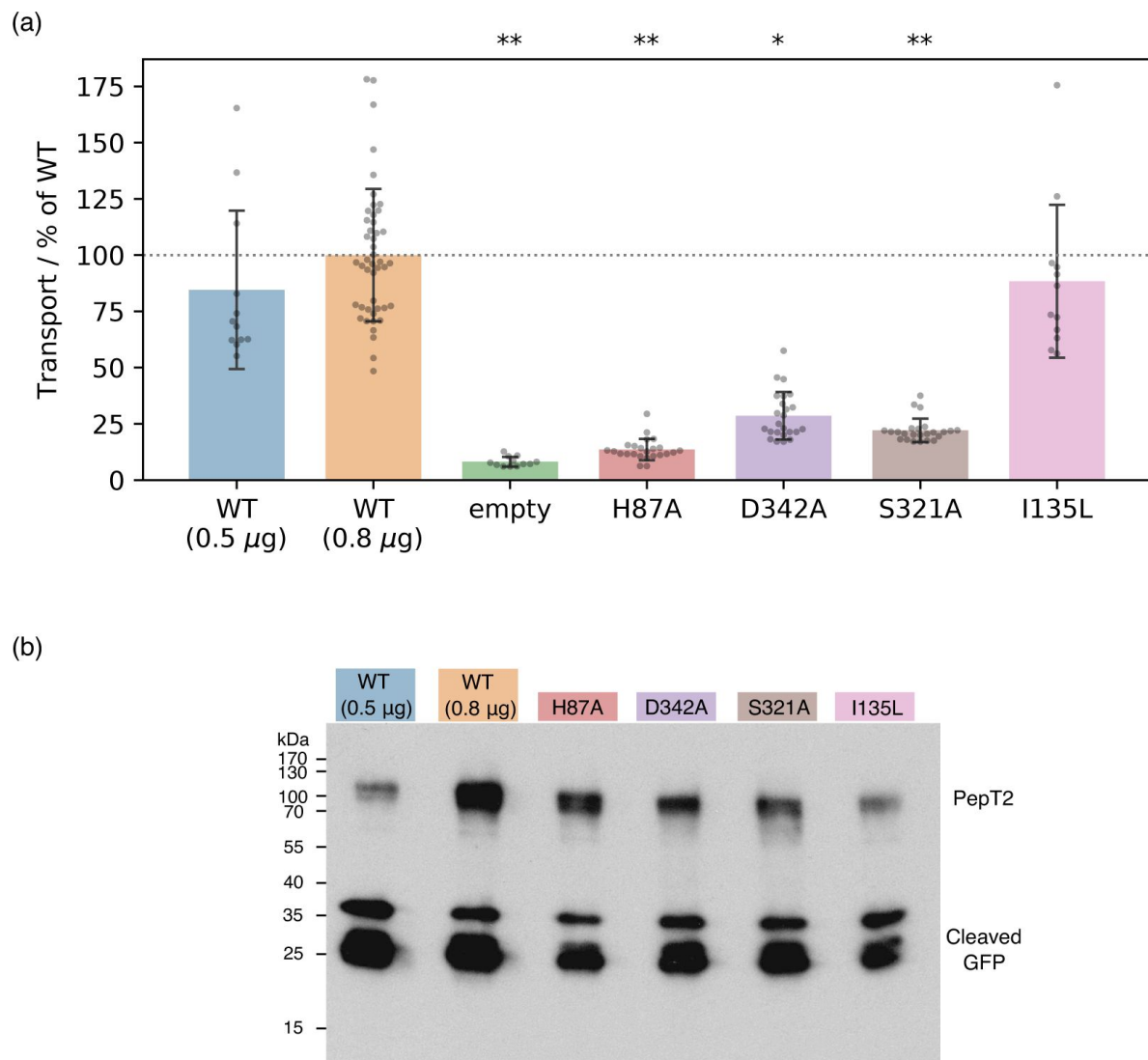


Figure 7.

(a) Cell-based transport assays for PepT2 WT (transfected with 0.5 µg, n=12, and 0.8 µg, n=46, of DNA per well), empty plasmid vector (n=12) and PepT2 H87A, D342A, S321A (n=24 each) and I135L (n=12) mutants, all transfected with 0.8 µg of DNA. Diagram shows transport as fluorescence in post-assay lysate divided by total protein concentration, normalised to the WT (0.8 µg) mean. Bars are mean values plus minus standard deviation, and swarm plots samples corresponding to individual wells. Single asterisk indicates $p < 10^{-3}$, double asterisks $p < 10^{-4}$ significance levels for difference compared to (weaker transporting, 0.5 µg-transfected) WT, as evaluated using a two-tailed t-test. (b) Western-blot showing expression levels of WT and mutant GFP-labelled PepT2, with an anti-GFP primary antibody. All mutants express at levels between the WT transfected with 0.5 µg and 0.8 µg plasmid DNA. Cleaved GFP is also visible at low molecular weight, at levels comparable for WT and mutants.

protonation states, and H87 and (possibly) D342 are merely the initial sites of protonation once OF is reached, providing further stabilisation.¹ We note in this context the limitation of using non-reactive MD methods for sampling the PepT2 conformational changes. In principle, a multi-dimensional PMF calculated with a reactive MD method where one CV is an explicit proton movement coordinate could disambiguate between the possible scenarios here. However, we believe that such a fully coupled treatment of proton movement and large-scale conformational changes is not yet computationally feasible, and focussed here on achieving convergence of the conformational sampling in discrete protonation-state combinations. The question of whether proton binding happens in OCC or OF therefore warrants further investigation, and indeed the co-existence of several mechanisms may be plausible. Nonetheless, this study contributes important details to a mechanistic understanding of the thermodynamics of proton-coupled alternating access.

Our results with regards to the driving forces behind the OCC → OF motion agree with the work of [Parker et al. \(2017\)](#), who found spontaneous opening of PepT_{Xc} towards OF after protonating the equivalent histidine to H87 (PepT2). In transporters which do not conserve the mammalian histidine at the TM 2 position such as PepT_{St} ([Batista et al., 2019](#)) and PepT_{Sh} ([Li et al., 2022](#)), on the other hand, previous simulation studies have implicated protonation of the glutamates equivalent to D317 (PepT2) in the opening of the extracellular gate. Our simulations suggest that the involvement of D317 in the extracellular gating mechanism of PepT2 is by interacting with the protonated H87 to stabilise the transition region of extracellular gate opening. The mechanism of extracellular gating is therefore conserved less widely than the overall alternating access mechanism. This point is further highlighted by our results indicating a role of the D342–R206 saltbridge, which is conserved only among mammalian POTs, but not in PepT_{So} and PepT_{Xc} which do have the TM 2 histidine. This may explain why spontaneous opening in unbiased MD was observed by [Parker et al. \(2017\)](#) for PepT_{Xc} following just H87 protonation, while for PepT2 in this work, protonation of D342 is also required.

Once the transporter is in the OF conformation, the substrate enters the binding site. We have treated this step al-chemically in our ABFE simulations, so that the data presented here is agnostic of the orientation of the entering substrate and the sequence of engagement of the binding-site residues; previous MD simulations have suggested, however, that positioning of the peptide N-terminus precedes the C-terminus moving into place. ([Parker et al., 2021](#)) Substrate binding then has two distinct effects: firstly, it exerts a small thermodynamic bias towards the OCC state via increased flexibility in degrees of freedom orthogonal to the overall OF → OCC transition. Secondly, through engaging R57, substrate entry increases the pKa value of E56 in the ExxER motif, thus thermodynamically facilitating the movement of protons further down the transporter cleft. We can also speculate that — in addition to this thermodynamic favouring of E56 protonation — there might be a kinetic effect on proton transfer from moving the positively charged R57 out of the way of the incoming proton. This could be investigated using reactive MD or QM/MM simulations (both approaches have been employed for other protonation steps of procaryotic peptide transporters, see [Parker et al. \(2017\)](#) and [Li et al. \(2022\)](#)), however the putative path is very long (≈ 1.7 nm between H87 and E56) and may or may not involve a large number of intermediate protonatable residues, in addition to binding site water. While such an investigation is possible in principle, it is beyond the scope of the present study. Likewise, a coupled enhanced-sampling treatment involving both proton movement and large-scale conformational changes — as discussed above for the ordering of steps in the OCC → OF transition — would make for interesting future work, once it becomes computationally tractable.

Our data is not fully determinate with respect to whether protons move to E56 before or after the OF → OCC conformational transition (our CpHMD, for example, remains agnostic on this matter since the shift in the E56 pKa value induced by the substrate is evident in both the OCC and OF states). We may interpret the fact that OCC is raised in energy while H87 is protonated and substrate-induced OCC stabilisation is not found when H87 and D342 are protonated (but does

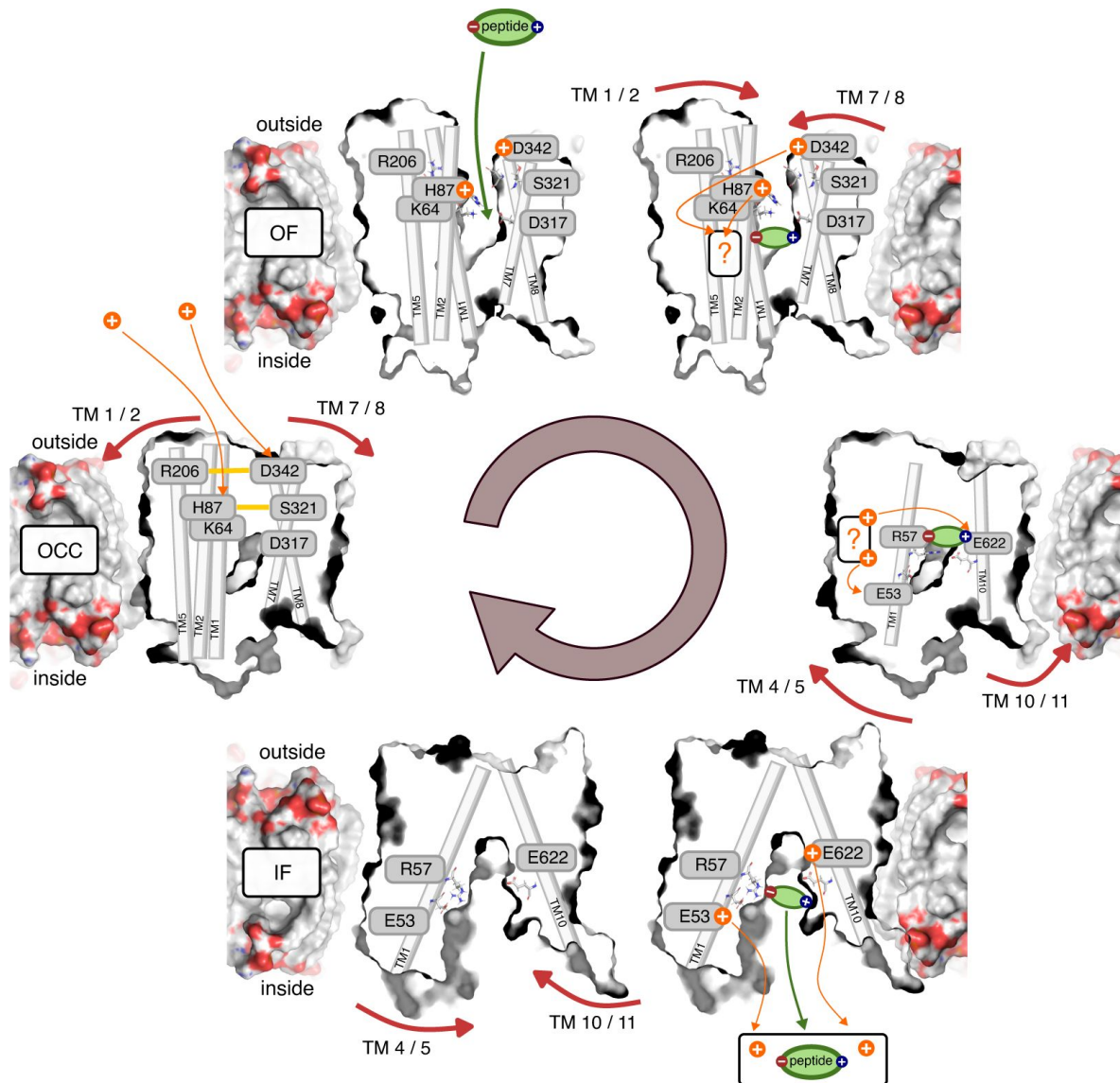


Figure 8.

Schematic overview of the Pept2 alternating-access transport cycle proposed in this work. Protons located at a question mark indicate a proton-transfer step with an as-of-yet unknown mechanism regarding intermediate residues.

occur when E56 is protonated) as an indication that proton movement is favoured before the transition into OCC is complete. On the other hand, the transition-region interaction between protonated H87 and D317 could also be interpreted as a potential facilitator of the OF → OCC transition. We thus speculate that the proton movement processes may happen as an ensemble of different mechanisms, and potentially occur contemporaneously with the conformational change. This, in addition to a flexible binding pocket, may also contribute to the substrate promiscuity mechanism.

We note at this stage that — throughout our study — we have not investigated the possibility of the substrate C-terminus itself becoming (transiently) protonated. This would need to be taken into account when treating proton movement through the transporter explicitly in the future (see the discussion of such approaches above). There is evidence in our simulations that an additional protonation site — aside from H87, D342, E53, E56 and E622 — may be involved in the mechanism, since E622 protonation, while biasing the transporter towards IF, also increases the OCC → IF transition barrier if Ala-Phe substrate is bound (we therefore indicate the proton movements at these stages with a question mark in **figure 8**). There is thus the intriguing possibility that the substrate itself may temporarily hold the proton, although given the nature of the data presented here only speculation is currently possible on this point. What is clear from our 2D-PMFs, regardless, is that protonation of ExxER does act as an intracellular gate trigger (and may also pull the transporter through the chemical equilibria all the way from OF). Taking together our 2D-PMFs and ABFE simulations, it is also clear that E622, in addition to being essential for peptide recognition, plays two further roles: its protonation both facilitates substrate release and makes an additive contribution to the IF-directed bias exerted by the intracellular gate triggers (whether E622 forms ‘part’ of the intracellular gate remains then as a merely linguistic question). At this stage, we do not yet have an understanding of how exactly intracellular gate opening (which involves breaking an assembly of several hydrophobic residues) is effected by the protonation of these glutamate residues. This question should prove interesting to study in future work. Once the substrate (driven by E622 protonation) and the protons (driven by their electrochemical potential gradient) have left through the open intracellular gate into the intracellular bulk, the resulting apo, standard-protonation-state IF conformation has a thermodynamic preference to return to OCC as evidenced in our 2D-PMFs. We thus arrive back at the starting state and have completed the proton- and substrate-coupled alternating access cycle.

In support of our MD data, we present mutational analysis in cell-based transport assays. Mutations of H87, S321 and D342 to alanine all significantly decrease transport activity, with H87A having the strongest effect. Taken together with similar results in the literature on E53, E56, R57, D317 and E622 (as referenced above), all residues implicated by our study have therefore been confirmed in their importance for transport via mutagenesis. While the cell-based assay used here cannot differentiate for example between proton-coupling and non-proton-coupling residues, our results still provide a useful first step towards the validation of the gating mechanisms we propose with our PMFs, and should prove informative for the future design of more in-depth experiments.

In conclusion, this study utilises the recent wealth of bacterial and mammalian peptide transporter structures to construct a model of their alternating access mechanism. We explain how the movement of two protons through the transporter drives the accompanying conformational changes, as well as how conformational changes and proton-movement events are coupled to the presence of substrate. Questions regarding some of the finer details, notably the precise sequence of proton movements and conformational transitions (if a single such sequence exists) and whether a further protonation site contributes to the mechanism remain open for future investigation. Nonetheless, the evidence supplied here addresses the alternating access proton-symport mechanism in unprecedented detail, particularly through the extensive use of free-energy simulation techniques. This information will prove useful for the project of employing

peptide transporters as vehicles for drug delivery — especially since what determines the efficacy of a transporter substrate is not only related to affinity but crucially also to an ability of the substrate to move through the steps of the alternating access cycle once bound to the transporter.

Methods and Materials

Data availability

Simulation and experimental data produced in this work is available at doi: 10.5281/zenodo.10561418. This includes key coordinate files and python scripts, as well as simulation trajectories projected onto the CVs of interest in plumed output format, and relevant processed files for PMF, pKa and ABFE calculations. Full trajectories will be shared upon reasonable request. The PyMENTO software is freely available at <https://github.com/simonlichtinger/PyMENTO>.

Definition of tip and base bundle CVs

For the analysis and interpretation of our unbiased MD runs, as well as for the use as a CV for initial metadynamics and umbrella sampling trials, we constructed the tip-CV and base-CV as centre-of-mass distances between the Ca atoms of the top and bottom 11 residues of the N-terminal and C-terminus bundles respectively, as illustrated in [figure 2a](#). These residues were picked as the consensus of the DSSP analysis ([Kabsch and Sander, 1983](#)) of the PepT2 conformational states derived below, listed in [table 2](#).

MD setup and equilibration

We obtained protein coordinates from cryo-EM for the OF (7NQK) ([Parker et al., 2021](#)) and IF-partially-occluded (7PMY) ([Killer et al., 2021](#)) PepT2 conformations, as well as from alphafold 2 ([Jumper et al., 2021](#)) for the fully-open IF state. We used MODELLER ([Šali and Blundell, 1993](#)) to fit the rat PepT2 sequence (as used by [Parker et al. \(2021\)](#)) to the human PepT2 7PMY and alphafold models, using residues 43–409 (TM 1–9) linked as a continuous chain to residues 604–700 (TM 10–12), thereby truncating the extracellular domain as done by [Parker et al. \(2021\)](#) in their MD simulations. We scored 200 models with QMEANDisCo ([Studer et al., 2020](#)) and selected the highest-scoring protein model for embedding into a 3:1 POPE:POPG bilayer of target size 10 * 10 nm (210/72 lipid molecules for IF and OCC, 218/72 for OF) with the CHARMM-GUI membrane builder. ([Wu et al., 2014](#)) We added ACE/NME capping residues using pymol and solvated the membrane system using GROMACS ([Abraham et al., 2015](#)) with approximately 21,000 solvent molecules (precise number varies between replicates) at a NaCl concentration of 0.15 M in an orthorhombic box of around 9.9 * 9.9 * 10.8 nm side lengths. Topologies were generated using the AMBER ff14.SB ([Maier et al., 2015](#)) and slipids ([Jämbeck and Lyubartsev, 2012](#)) forcefields.

Using the GROMACS MD engine ([Abraham et al., 2015](#)) in versions 2021.3/2021.4 (the slight version discrepancy is because of different installations on two compute clusters we used), we energy-minimized the systems, assigned initial velocities, and equilibrated with Ca-atom restraints for 200 ps in the NVT ensemble with a leap-frog integrator (using the modified v-rescale thermostat with a stochastic term ([Bussi et al., 2007](#)) at 310 K throughout our work), then 1 ns in the NPT ensemble (with the berendsen barostat), followed by 20 ns of further Ca-restrained NPT equilibration (using the Parrinello-Rahman barostat ([Parrinello and Rahman, 1981](#)), as for all subsequent production runs).

We obtained Ala-Phe dipeptide-bound boxes by aligning the holo PepT1 cryo-EM structure (7PMW) onto our equilibrated PepT2 MD boxes, copying the ligand coordinates, and repeating the same equilibration protocol as before (where the peptide substrate Ca atoms were also restrained). The

Bundle	Residue numbers
N-terminal bundle tips	63–74, 79–90, 120–131, 143–154, 194–205, 217–228
N-terminal bundle bases	46–57, 93–104, 110–121, 161–172, 177–188, 227–238
C-terminal bundle tips	320–331, 341–352, 392–403, 609–620, 655–666, 671–682
C-terminal bundle bases	290–301, 359–370, 376–387, 626–637, 642–653, 686–697

Table 2.

Residue numbers used in the definition of the tip-CV and base-CV.

peptide ligand was parametrised using AMBER ff14.SB. ([Maier et al., 2015](#))

Derivation of OF, OCC and IF conformational states

As shown in supplementary figure S2a, the IF-partially-occluded structure (7PMY) does not behave well in MD (1 μ s production runs from triplicate embeddings), since it either partially opens its extracellular gate (replicates 2–3) or partial helical unfolding in the intracellular gate occurs due to hydrophobic collapse (rep 1). This may be due to a variety of factors; one possibility is instability in the protein following the removal of the bound substrate in our simulations. In contrast (figure S2b), embedding replicates 1 and 3 of the alphafold IF state behave well. We picked the end-coordinates of replicate 3 as our IF state, due to the wider opening of its intracellular gate. We then sought to derive an OCC state via MD from replicate 1, see the paragraph below. We also note that embeddings from the OF cryo-EM structure (7NQK) remain stable in the OF conformation, we picked replicate 1 for our work.

To derive an OCC state from an IF box, we conducted 5 replicates of well-tempered metadynamics ([Barducci et al., 2008](#)) as implemented in PLUMED 2.7 ([Tribello et al., 2014](#)) along the base-bundle CV (see [figure 2a](#)), using 8 walkers, hill height 1 kJ mol⁻¹ with sigma 0.022 nm deposited every 500 steps, and a bias factor of 100. From the resulting set of trajectories, we picked frames around the mark of 20 ns simulation time and a base-CV value of around 2.0 nm (we chose these values based on visual inspection of the trajectories, where we noticed that base-CV values significantly below 2.0 nm lead to artefacts such as partial unfolding of the ends of helices, as did continuing the metadynamics simulations for longer than necessary to obtain the desired states). We ran triplicate 100 ns-long unbiased MD from the obtained states for each of the five replicates, and found the OCC state obtained from the first replicate to be stably situated within the range of base-CV values observed in the OF-state trajectories. The micro-second-long unbiased MD runs as well as our 2D-PMFs confirm that this protein conformation is a stable basin, and that different protonation states of key residues can drive its opening to either the IF or OF states. While this does not rule out the existence of different OCC states, it confirms the properties of the conformation we found as a functional OCC state.

Unbiased MD of the OCC state

Unbiased MD runs of the OCC state in different conditions were conducted in triplicates using the same simulation parameters as for the long C α -restrained equilibrations described in the section on equilibration, but removing all restraints. The starting coordinates were — for the first replicate — the OCC state derived as described in the foregoing section, and the second and third replicates were initialised from the 500 ns and the final frame of the first replicate trajectory. Protonation states changes and mutations were carried out using PyMOL and GROMACS pdb2gmx independently for each replicate, followed by re-running the full equilibration protocol for all new boxes. Taken together, we conducted unbiased MD for 24 conditions, giving 72 μ s of production sampling.

Metadynamics and steered MD

For our initial trials of enhanced sampling on PepT2 conformational changes — which showed hysteretic behaviour (see figure S6), we attempted steered MD (SMD) and metadynamics for the OCC $\leftrightarrow$ OF transition in the WT, unprotonated state.

Two instances of 8-walker well-tempered metadynamics were run, starting from the OF and OCC states, biased along the tip-CV with hills of height 1 kJ mol⁻¹ and sigma 0.0455 nm deposited every 500 steps, using a bias factor of 100. A harmonic flat-bottom restraint with boundaries of 2.0–3.0 nm and force constant 5×10^4 kJ mol⁻¹ nm⁻² was applied on the CV value. Sampling was run for 108 ns per walker for the simulations starting from OF, and 213 ns per walker starting from OCC.

To generate paths for umbrella sampling, SMD was run starting at OF towards OCC and vice versa with the heavy-atom RMSD to the target conformation as CV, using a harmonic potential centered to zero RMSD with a force constant sliding from 0 up to $2.5 \times 10^5 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ over 200 ns. The harmonic potential was then switched off over 2 ns, followed by 48 ns of unbiased MD. We then projected the SMD trajectories onto the tip-CV, picked 48 frames spaced equally along the CV and performed 1D replica-exchange umbrella sampling (REUS) using a force constant of $3 \times 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-2}$, for 92 ns per window for the OCC → OF derived boxes and 127 ns per window in the reverse direction. A total of 13.6 μs of MD time was thus expended on these trials.

MEMENTO path generation

We have recently proposed the MEMENTO method for history-independent path generation between given end-states. ([Lichtinger and Biggin, 2023](#)) In short, protein coordinates are morphed, followed by reconstructing an ensemble of structures at each morphing intermediate using MODELLER. Monte-Carlo simulated annealing with an energy function based on between-intermediate RMSD values then finds a smooth path through these ensembles. For membrane proteins, lipids are taken from the end-state that occupies a larger area in the membrane (in this case, OF for the OCC → OF transition, and IF for OCC → IF), and fitted around the new protein coordinates by expanding the membrane, followed by iterative compression and energy minimization. Ligands are not morphed but translated, interpolating between the ligand centres of masses in the end-states, and then equilibrated in the protein structure using MD. MEMENTO is implemented as the PyMEMENTO package (<https://github.com/simonlichtinger/PyMEMENTO>), we provide example scripts for its usage on PepT2 in the supplementary data.

In this study, we ran MEMENTO with 24 windows in triplicates for both the OCC → OF and OCC → IF transitions in different protein protonation and mutation states, and in the presence or absence of ligands. The apo and holo MEMENTO replicates were initialised from the 0 ns, 500 ns and 1000 ns frames of the first replicate of the 1 μs unbiased MD run for each conformational state (using always the unprotonated, WT condition, but apo/holo trajectories respectively). Protonation state changes and mutations were then carried out using the built-in functionality of PyMEMENTO, and equilibrated at each intermediate state for 90 ns. The total MD simulation time spent on equilibrations as part of the MEMENTO method across the 22 sampled conditions was 47 μs .

1D-PMF calculations

Starting with the triplicate equilibrated MEMENTO boxes for the (all apo) OCC → OF standard protonation and H87 & D342 protonated states, as well as OCC → IF standard protonation and E53 protonated states, we ran 1D-replica exchange umbrella sampling (REUS, exchange every 1000 steps; using PLUMED 2, [Tribello et al. \(2014\)](#)) along the tip- and base-CVs respectively, using a force constant of $4 \times 10^3 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. The amount of sampling collected in each case is summarised in [table 3](#).

2D-CV derivation

Using the trajectory data from our 1D-PMFs, we derived 2D CVs via a PCA-based approach we have previously described for LEUT. ([Lichtinger and Biggin, 2023](#)) We pooled all sampling collected in 1D-REUS runs along the tip-CV for apo OCC → OF (and equivalently for OCC → IF. The same procedure was taken for these trajectories, and we will only explicitly write about OCC → OF in the following paragraph). Using GROMACS tools, we ran PCA of the Ca positions of residues contained in the transmembrane region (see [table 2](#) above). The first principal component (PC) accounts for 50% of the variance; adding an extra 15 PCs increases coverage to 78% of the variance (comparable to our results on LEUT). The first PC (see figure S8) describes the gating motions of the respective conformational changes, behaving similarly to the tip CV (or base CV) — expectedly so, given it was the CV used in our 1D-REUS. To explain differences between replicates (see figure

Condition	Simulation time / ns
OCC↔OF, standard prot	24 * (266 + 244 + 244)
OCC↔OF, H87 & D342 prot	24 * (327 + 244 + 244)
OCC↔IF, standard prot	24 * (242 + 248 + 246)
OCC↔IF, E53 prot	24 * (158 + 155 + 154)
Total	67 μs

Table 3.

Overview of all 1D-PMF sampling.

S7a and c), we used differential evolution ([Storn and Price, 1997](#)) as implemented in scipy ([Virtanen et al., 2020](#)) to maximise an entropy-like metric of distances between MEMENTO path frames for linear combinations of the PCs 2–16:

$$\frac{2}{N_{\text{rep}} * (N_{\text{rep}} - 1)} \sum_{i=1}^{N_{\text{rep}}} \sum_{j=i+1}^{N_{\text{rep}}} \sum_n \log \left(\sqrt{\sum_{\text{PC}} (X(n, i) - X(n, j))^2} \right), \quad (1)$$

where N_{rep} is the number of replicates and by $X(n, i) - X(n, j)$ we denote the distance between two conformational frames in different replicates i and j , evaluated in a projection along a given combination of principal components. The result was termed PC 2 henceforth for simplicity and used as the second CV in 2D-REUS (figure S8).

2D-PMF calculations

Using the same equilibrated MEMENTO paths as above and the 2D-CVs we derived, we calculated 2D-PMFs of the OCC ↔ OF and OCC ↔ IF transitions in several protonation / mutation states, with and without Ala-Phe substrate bound, using 2D-REUS. As shown in figure S12, we found that a lower force constant of $2 \times 10^6 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ leads to good histogram overlap in the lower-lying regions of the PMF, but has poor overlap near the transition region. In turn, a higher force constant of $1 \times 10^7 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ gives good window overlap in the transition region while not sampling broadly enough in large basins. Therefore, for each condition and MEMENTO replicate, we ran windows at both force constants and included them all in the WHAM analysis, thus ensuring sufficient sampling through the CV space. Replica-exchange was run within the 24 windows corresponding to each MEMENTO replicate–force constant combination, and a total of 144 windows contribute to each 2D-PMF.

The sampling collected in all conditions is detailed in [table 4](#), aiming for between 180 and 210 ns per window though exact amounts differ with heterogeneous hardware and slightly different box sizes.

Absolute binding free energies

To probe the affinity of the Ala-Phe substrate to the PepT2 OF and IF conformations in different protonation states, we conducted absolute binding free energy (ABFE) simulations ([Aldeghi et al., 2016](#), [2018](#)) in gromacs, using an equilibrium approach. For this, we changed our mdp files to use the stochastic dynamics integrator (doubling as a thermostat) and set the relevant free-energy flags, including soft-core van-der-Waals interactions (alpha = 0.5, power = 1, sigma = 0.3) and the couple-intramol=yes flag for consistency with larger ligands in other work. Our lambda-protocol was to first add Boresch restraints ([Boresch et al., 2003](#)) (for the complex thermodynamic leg only, ligand side was calculated using the analytic formula; through values 0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0), then annihilate coulomb interactions (even 0.1 spacings) followed by vdw interactions (even 0.05 spacings). We equilibrated for 200 ps of NVT and 1 ns of NPT at each lambda window, and ran production simulations with replica-exchange attempts every 1000 steps for 30 ns per window on the complex thermodynamic leg and 100 ns on the ligand-only leg.

The Boresch restraints for ABFE simulations were obtained MDRestraintsGenerator ([Alibay et al., 2022](#)) by running the restraint finding algorithm over the ≈200 ns 2D-REUS trajectory at the relevant conformation and protonation state, and we used the trajectory frame closest to the restraint centre as input for subsequent ABFE. This was done for each of the triplicate MEMENTO runs, giving 3 candidate ligand binding poses. For the pose that was found to have the highest single-replicate affinity, four replicates of unbiased, restraint-free 200 ns-long equilibrations were also started from the frames, processed with MDRestraintsGenerator and used to make replicates

Condition	Simulation time / ns
OCC↔OF, standard prot	24 * (233 + 195 + 195 + 208 + 208 + 207)
OCC↔OF, H87 prot	24 * (230 + 200 + 188 + 201 + 228 + 229)
OCC↔OF, D342 prot	24 * (204 + 232 + 204 + 224 + 229 + 207)
OCC↔OF, H87&D342 prot	24 * (196 + 196 + 195 + 194 + 194 + 201)
OCC↔OF, D342A	24 * (182 + 183 + 206 + 181 + 183 + 183)
OCC↔OF, H87A	24 * (176 + 182 + 183 + 184 + 182 + 177)
OCC↔OF, E53 prot	24 * (185 + 180 + 209 + 179 + 218 + 213)
OCC↔OF, E56 prot	24 * (194 + 194 + 194 + 192 + 193 + 193)
OCC↔OF, R206D&D342R	24 * (227 + 213 + 183 + 222 + 211 + 180)
OCC↔OF, holo, standard prot	24 * (200 + 196 + 195 + 200 + 194 + 201)
OCC↔OF, holo, E53 prot	24 * (185 + 180 + 210 + 179 + 218 + 213)
OCC↔OF, holo, E56 prot	24 * (194 + 194 + 194 + 192 + 193 + 193)
OCC↔OF, holo, H87&D342 prot	24 * (198 + 191 + 198 + 190 + 182 + 201)
OCC↔IF, standard prot	24 * (197 + 216 + 203 + 192 + 197 + 199)
OCC↔IF, E53 prot	24 * (197 + 208 + 199 + 192 + 193 + 195)
OCC↔IF, E622 prot	24 * (199 + 201 + 200 + 196 + 217 + 197)
OCC↔IF, E53&E622 prot	24 * (193 + 237 + 240 + 187 + 250 + 200)
OCC↔IF, holo, standard prot	24 * (190 + 195 + 191 + 196 + 192 + 197)
OCC↔IF, holo, E53 prot	24 * (189 + 196 + 197 + 199 + 189 + 198)
OCC↔IF, holo, E622 prot	24 * (194 + 191 + 206 + 197 + 198 + 199)
OCC↔IF, holo, E53&E622 prot	24 * (197 + 198 + 180 + 195 + 190 + 204)
Total	598 μ s

Table 4.

Overview of all 2D-PMF sampling.

for ABFE runs to give an error estimate as mean plus-minus standard deviation. By following this protocol for the OF, OF E56 prot, IF and IF E622 prot conditions, we sampled for a total of 4 conditions * 7 boxes per condition * 44 windows * 31.2 ns = 38 μ s.

Constant-pH MD

To probe the substrate-dependence of the E53 and E56 pKa values, we ran constant pH simulations (CpHMD) using the hybrid solvent approach with discrete protonation states ([Swails et al., 2014](#)) as implemented in the AMBER software. (D.A. [Case et al., 2021](#)) We took the MEMENTO starting frames from above as triplicate initial coordinates of OF and OCC states in the presence and absence of Ala-Phe. We then used tleap and in-house scripts to convert our boxes to the AMBER constant pH forcefield fork for protein, substrate and solutes and to the lipid21 force field ([Dickson et al., 2022](#)) for the membrane. We prepared constant pH simulations as in the tutorial by [Swails et al., 2014](#), and ran them for 1 μ s at 8 pH replica windows (ph 0–7), in the NVT ensemble with a langevin thermostat (at 310 K as before), attempting protonation state changes every 100 steps, running 100 steps of relaxation dynamics for every exchange, and attempting replica exchange every 1000 steps. Analysis was performed using the cphstats programme and in-house scripts for fitting titration curves. In [figure 6](#), analysis is performed per replicate, reporting mean $\pm$ standard deviation for each condition and residue. In figure S17–18 we show pKa values estimated over simulation time from 10 ns chunks of all CpHMD runs. We also analyse this data in terms of histograms of the chunk-estimated pKa values, pooling all data for each condition and residue. A total of 4 conditions * 3 replicates * 8 windows * 1 μ s = 96 μ s of sampling were thus collected.

Cell-based transport assays

Transport assays were carried out using a modified version of the protocol by [Parker et al. \(2021\)](#). Hela cells were cultured in DMEM + GlutaMAX medium, supplemented with 10% FBS. 12-well plates were prepared by seeding 9×10^4 cells per well in 1 mL of medium, and transfected after 24h with 0.8 μ g of PepT2-constructs in pEF5-FRT-eGFP vector (or 0.5 μ g of insert vector + 0.3 μ g of empty vector, where specified), with 1.6 μ g of fugene transfection reagent. The medium was exchanged 24h post transfection, and assays were carried out 40h post transfection. Cells were washed 2 times with ≈ 0.6 mL of assay buffer (20 mM HEPES pH 7.5, 120 mM NaCl, 2 mM MgSO₄ and 25 mM glucose), then incubated with 0.3 mL assay buffer containing 20 mM β -ala-lys-AMCA substrate for 15 min. The cells were then washed 3 times with ≈ 0.6 mL of assay buffer, and incubated with 0.25 mL of lysis buffer (20 mM Trist pH 7.5 + 0.2% Triton x-100) for 5 min. The fluorescence (340 nm excitation, 460 nm read-out) of 0.15 mL of the lysate was normalised by the protein amount in each well (as determined from BCA assay of 20 μ L lysate). We removed two outliers from the WT (0.8 μ g) transport assay dataset, giving n=46. The data was then scaled to the mean WT (0.8 μ g) transport level as 100%.

Protein expression controls

For comparing PepT2 WT and mutant expression levels, Hela cells were seeded in 6-well plates at 1.8×10^5 cells per well in 2 mL of medium. Transfection was after 24h with 1.6 μ g of PepT2-constructs in pEF5-FRT-eGFP vector (or 1.0 μ g of insert vector + 0.6 μ g of empty vector, where specified) and 3.2 μ g of fugene; the medium was exchanged 24h post transfection. The cells were washed 3 times with ≈ 0.6 mL of PBS, harvested using 0.1% trypsin, pelleted, re-suspended in 100 μ L PBS with protease inhibitor and lysed through 3x freeze-thawing. The lysates from 3 wells were pooled for each mutant to increase between-sample consistency. 4.5 μ L of each sample were loaded onto a 10% SDS-PAGE gel, and western blot was performed using an anti-GFP antibody. The membrane was then stripped and developed again with an anti- β -actin antibody to control for gel loading.

Protein localisation controls

To confirm the plasma membrane localisation of PepT2 WT and mutants, we seeded 1.8×10^5 HeLa cells per well in 2 mL of medium in a 6-well plate with added coverslips. Transfection was as for the expression controls 24h after seeding. 20h post transfection, the cells were washed 3 times with ≈ 0.6 mL of PBS, fixed with PFA for 10 min at room temperature, washed 3 times, incubated with PBS + 50 mM NH_4Cl for 10 min, washed 3 times, incubated with PBS + 0.1 % Triton x-100 for 5 min, washed 3 times, stained with DAPI, washed 5 times, and mounted on slides with ImmuMount. Images were recorded in the GFP and DAPI channels.

Acknowledgements

We would like to thank Dr. Zhiyi Wu for general training in molecular dynamics methodology and help in the early stages of this project, Dr. Irfan Alibay for training in ABFE simulations, Dr. Gabriel Kuteyi for training in mammalian cell culture techniques and Sacha Salphati for training in fluorescent microscopy, as well as all members of the Biggin and Newstead groups for helpful discussions.

This project was funded by the Wellcome Trust (Grant ID: 218514/Z/19/Z). Compute resources were also provided by the EPSRC ARCHER2, Jade 2 and N8 CIR BEDE facilities, granted via the High-End Computing Consortium for Biomolecular Simulation (HECBioSim), supported by EPSRC (EP/X035603/1).

References

- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E, 19–25 (2015) **Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers** *SoftwareX* :1–2 <https://doi.org/10.1016/j.softx.2015.06.001>
- Aduri NG, Prabhala BK, Ernst HA, Jørgensen FS, Olsen L, Mirza O (2015) **Salt bridge swapping in the exxerfxxy motif of proton-coupled oligopeptide transporters** *Journal of Biological Chemistry* **290** <https://doi.org/10.1074/jbc.M115.675603>
- Aho N, Buslaev P, Jansen A, Bauer P, Groenhof G, Hess B (2022) **Scalable constant ph molecular dynamics in gromacs** *Journal of Chemical Theory and Computation* **18**:6148–6160 <https://doi.org/10.1021/acs.jctc.2c00516>
- Aldeghi M, Bluck JP, Biggin PC (2018) **Absolute alchemical free energy calculations for ligand binding: a beginner's guide** In: *Computational drug discovery and design, methods in molecular biology* **1762** https://doi.org/10.1007/978-1-4939-7756-7_11
- Aldeghi M, Heifetz A, Bodkin MJ, Knapp S, Biggin PC (2016) **Accurate calculation of the absolute free energy of binding for drug molecules** *Chemical Science* **7**:207–218 <https://doi.org/10.1039/C5SC02678D>
- Alibay I, Magarkar A, Seeliger D, Biggin PC (2022) **Evaluating the use of absolute binding free energy in the fragment optimisation process** *Communications Chemistry* **5** <https://doi.org/10.1038/s42004-022-00721-4>

- Barducci A, Bussi G, Parrinello M (2008) **Well-tempered metadynamics: a smoothly converging and tunable free-energy method** *Physical Review Letters* **100** <https://doi.org/10.1103/PhysRevLett.100.020603>
- Batista MRB, Watts A, José Costa-Filho A (2019) **Exploring conformational transitions and free-energy profiles of proton-coupled oligopeptide transporters** *Journal of Chemical Theory and Computation* **15** <https://doi.org/10.1021/acs.jctc.9b00524>
- Beale J *et al.* (2015) **Crystal structures of the extracellular domain from pept1 and pept2 provide novel insights into mammalian peptide transport** *Structure* **23** <https://doi.org/10.1016/j.str.2015.07.016>
- Boggavarapu R, Jeckelmann JM, Harder D, Ucurum Z, Fotiadis D (2015) **Role of electrostatic interactions for ligand recognition and specificity of peptide transporters** *BMC Biology* **13** <https://doi.org/10.1186/s12915-015-0167-8>
- Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) **Absolute binding free energies: a quantitative approach for their calculation** *The Journal of Physical Chemistry B* **107** <https://doi.org/10.1021/jp0217839>
- Bussi G, Donadio D, Parrinello M (2007) **Canonical sampling through velocity rescaling** *The Journal of Chemical Physics* **126** <https://doi.org/10.1063/1.2408420>
- Chen XZ, Steel A, Hediger MA (2000) **Functional roles of histidine and tyrosine residues in the h⁺-peptide transporter pept1** *Biochemical and Biophysical Research Communications* **272**:726–730 <https://doi.org/10.1006/bbrc.2000.2851>
- Chen XZ, Zhu T, Smith DE, Hediger MA (1999) **Stoichiometry and kinetics of the high-affinity h⁺-coupled peptide transporter pept2** *Journal of Biological Chemistry* **274** <https://doi.org/10.1074/jbc.274.5.2773>
- Case D A *et al.* (2021) **D A Case, H M Aktulga, K Belfon, I Y Ben-Shalom, S R Brozell, D S Cerutti, T E Cheatham, III, G A Cisneros, V W D Cruzeiro, T A Darden, R E Duke, G Giambasu, M K Gilson, H Gohlke, A W Goetz, R, Harris, S Izadi, S A Izmailov, C Jin, K Kasavajhala, M C Kaymak, E King, A Kovalenko, T Kurtzman,, T S Lee, S LeGrand, P Li, C Lin, J Liu, T Luchko, R Luo, M Machado, V Man, M Manathunga, K M, Merz, Y Miao, O Mikhailovskii, G Monard, H Nguyen, K A O'Hearn, A Onufriev, F Pan, S Pantano,, R Qi, A Rahnamoun, D R Roe, A Roitberg, C Sagui, S Schott-Verdugo, J Shen, C L Simmerling, N R, Skrynnikov, J Smith, J Swails, R C Walker, J Wang, H Wei, R M Wolf, X Wu, Y Xue, D M York, S, Zhao, and P A Kollman, Amber 2021. University of California, San Francisco; 2021.**
- Daniel H, Kottra G (2004) **The proton oligopeptide cotransporter family slc15 in physiology and pharmacology** *Pfluegers Archiv European Journal of Physiology* **447**:610–618 <https://doi.org/10.1007/s00424-003-1101-4>
- Daniel H, Spanier B, Kottra G, Weitz D (2006) **From bacteria to man: archaic proton-dependent peptide transporters at work** *Physiology* **21** <https://doi.org/10.1152/physiol.00054.2005>
- Darve E, Rodríguez-Gómez D, Pohorille A (2008) **Adaptive biasing force method for scalar and vector free energy calculations** *The Journal of Chemical Physics* **128** <https://doi.org/10.1063/1.2829861>

- Dickson CJ, Walker RC, Gould IR (2022) **Lipid21: complex lipid membrane simulations with amber** *Journal of Chemical Theory and Computation* **18**:1726–1736 <https://doi.org/10.1021/acs.jctc.1c01217>
- Doki S *et al.* (2013) **Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter pot** *Proceedings of the National Academy of Sciences* **110** <https://doi.org/10.1073/pnas.1301079110>
- Fei YJ, Kanai Y, Nussberger S, Ganapathy V, Leibach FH, Romero MF, Singh SK, Boron WF, Hediger MA (1994) **Expression cloning of a mammalian proton-coupled oligopeptide transporter** *Nature* <https://doi.org/10.1038/368563a0>
- Fei YJ, Liu W, Prasad PD, Kekuda R, Oblak TG, Ganapathy V, Leibach FH (1997) **Identification of the histidyl residue obligatory for the catalytic activity of the human h + /peptide cotransporters pept1 and pept2** *Biochemistry* **36**:452–460 <https://doi.org/10.1021/bi962058p>
- Fei YJ, Nara E, Liu JC, Boyd CAR, Ganapathy V, Leibach FH (1999) **Preferential recognition of zwitterionic dipeptides as transportable substrates by the high-affinity peptide transporter pept2** *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1418**:344–351 [https://doi.org/10.1016/S0005-2736\(99\)00046-2](https://doi.org/10.1016/S0005-2736(99)00046-2)
- Fowler P *et al.* (2015) **Gating topology of the proton-coupled oligopeptide symporters** *Structure* **23** <https://doi.org/10.1016/j.str.2014.12.012>
- Guettou F, Quistgaard EM, Raba M, Moberg P, Löw C, Nordlund P (2014) **Selectivity mechanism of a bacterial homolog of the human drug-peptide transporters pept1 and pept2** *Nature Structural & Molecular Biology* **21** <https://doi.org/10.1038/nsmb.2860>
- Guettou F, Quistgaard EM, Trésaugues L, Moberg P, Jegerschöld C, Zhu L, Jong AJO, Nordlund P, Löw C (2013) **Structural insights into substrate recognition in protondependent oligopeptide transporters** *EMBO reports* **14** <https://doi.org/10.1038/embor.2013.107>
- Harpole TJ, Delemotte L, 1860, publisher: Elsevier B.V (2018) **Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations** *Biochimica et Biophysica Acta - Biomembranes* **4**:909–926 <https://doi.org/10.1016/j.bbamem.2017.10.033>
- Hénin J, Lelièvre T, Shirts MR, Valsson O, Delemotte L., version Number: 2 (2022) **Enhanced sampling methods for molecular dynamics simulations** *arXiv* <https://doi.org/10.48550/ARXIV.2202.04164>
- Izrailev S, Stepaniants S, Isralewitz B, Kosztin D, Lu H, Molnar F, Wriggers W, Schulten K. Steered molecular dynamics. In: Griebel M, Keyes DE, Nieminen RM, Roose D, Schlick T, Deuffhard P, Hermans J, Leimkuhler B, Mark AE, Reich S, Skeel RD, editors (1999) **Computational molecular dynamics: challenges, methods, ideas** *Berlin, Heidelberg: Springer Berlin Heidelberg* **4**:39–65 https://doi.org/10.1007/978-3-642-58360-5_2
- Jensen JM, Ernst HA, Wang X, Hald H, Ditta AC, Ismat F, Rahman M, Mirza O (2012) **Functional investigation of conserved membrane-embedded glutamate residues in the proton-coupled peptide transporter yjdl** *Protein & Peptide Letters* **19**:282–287 <https://doi.org/10.2174/092986612799363109>

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, et al. (2021) **Highly accurate protein structure prediction with alphafold** *Nature* **596**:583–589 <https://doi.org/10.1038/s41586-021-03819-2>

Jämbeck JPM, Lyubartsev AP (2012) **Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids** *The Journal of Physical Chemistry B* **116** <https://doi.org/10.1021/jp212503e>

Jørgensen ME, Olsen CE, Geiger D, Mirza O, Halkier BA, Nour-Eldin HH (2015) **A functional exxek motif is essential for proton coupling and active glucosinolate transport by npf2.11** *Plant and Cell Physiology* **56** <https://doi.org/10.1093/pcp/pcv145>

Kabsch W, Sander C (1983) **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features** *Biopolymers* **22**:2577–2637 <https://doi.org/10.1002/bip.360221211>

Killer M, Wald J, Pieprzyk J, Marlovits TC, Löw C (2021) **Structural snapshots of human pept1 and pept2 reveal mechanistic insights into substrate and drug transport across epithelial membranes** *Science Advances* **7** <https://doi.org/10.1126/sciadv.abk3259>

Kotov V et al. (2023) **Plasticity of the binding pocket in peptide transporters underpins promiscuous substrate recognition** *Cell Reports* **42** <https://doi.org/10.1016/j.celrep.2023.112831>

Li C, Yue Z, Newstead S, Voth GA (2022) **Proton coupling and the multiscale kinetic mechanism of a peptide transporter** *Biophysical Journal* **121**:2266–2278 <https://doi.org/10.1016/j.bpj.2022.05.029>

Lichtinger SM, Biggin PC (2023) **Tackling hysteresis in conformational sampling: how to be forgetful with memento** *Journal of Chemical Theory and Computation* **19**:3705–3720 <https://doi.org/10.1021/acs.jctc.3c00140>

Lindahl V, Lidmar J, Hess B (2014) **Accelerated weight histogram method for exploring free energy landscapes** *The Journal of Chemical Physics* **141** <https://doi.org/10.1063/1.4890371>

Lyons JA, Parker JL, Solcan N, Brinth A, Li D, Shah ST, Caffrey M, Newstead S (2014) **Structural basis for polyspecificity in the pot family of proton-coupled oligopeptide transporters** *EMBO reports* **15** <https://doi.org/10.15252/embr.201338403>

Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) **Ff14sb: improving the accuracy of protein side chain and back-bone parameters from ff99sb** *Journal of Chemical Theory and Computation* **11**:3696–3713 <https://doi.org/10.1021/acs.jctc.5b00255>

Molledo M Martinez, Quistgaard EM, Flayhan A, Pieprzyk J, Löw C (2018) **Multispecific substrate recognition in a proton-dependent oligopeptide transporter** *Structure* **26** <https://doi.org/10.1016/j.str.2018.01.005>

Minhas GS, Newstead S (2019) **Structural basis for prodrug recognition by the slc15 family of proton-coupled peptide transporters** *Proceedings of the National Academy of Sciences* **116** <https://doi.org/10.1073/pnas.1813715116>

- Newstead S (2017) **Recent advances in understanding proton coupled peptide transport via the pot family** *Current Opinion in Structural Biology* **45** <https://doi.org/10.1016/j.sbi.2016.10.018>
- Newstead S *et al.* (2011) **Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, pept1 and pept2** *The EMBO Journal* **30** <https://doi.org/10.1038/emboj.2010.309>
- Omori A, Fujisawa Y, Sasaki S, Shimono K, Kikukawa T, Miyauchi S (2021) **Protonation state of a histidine residue in human oligopeptide transporter 1 (hpept1) regulates hpept1-mediated efflux activity** *Biological and Pharmaceutical Bulletin* **44** <https://doi.org/10.1248/bpb.b20-01013>
- Parker JL, Deme JC, Wu Z, Kuteyi G, Huo J, Owens RJ, Biggin PC, Lea SM, Newstead S (2021) **Cryo-em structure of pept2 reveals structural basis for proton-coupled peptide and prodrug transport in mammals** *Science Advances* **7** <https://doi.org/10.1126/sciadv.abh3355>
- Parker JL *et al.* (2017) **Proton movement and coupling in the pot family of peptide transporters** *Proceedings of the National Academy of Sciences* **114** <https://doi.org/10.1073/pnas.1710727114>
- Parker JL, Mindell JA, Newstead S (2014) **Thermodynamic evidence for a dual transport mechanism in a pot peptide transporter** *eLife* **3** <https://doi.org/10.7554/eLife.04273>
- Parker JL, Newstead S (2014) **Molecular basis of nitrate uptake by the plant nitrate transporter nrt1.1** *Nature* **507** <https://doi.org/10.1038/nature13116>
- Parrinello M, Rahman A (1981) **Polymorphic transitions in single crystals: a new molecular dynamics method** *Journal of Applied Physics* **52**:7182–7190 <https://doi.org/10.1063/1.328693>
- Pizzagalli MD, Bensimon A, Superti-Furga G. (2021) **A guide to plasma membrane solute carrier proteins** *The FEBS Journal* **288**:2784–2835 <https://doi.org/10.1111/febs.15531>
- Radestock S, Forrest LR (2011) **The alternating-access mechanism of mfs transporters arises from inverted-topology repeats** *Journal of Molecular Biology* **407** <https://doi.org/10.1016/j.jmb.2011.02.008>
- Rubio-Aliaga I, Boll M, Daniel H (2000) **Cloning and characterization of the gene encoding the mouse peptide transporter pept2** *Biochemical and Biophysical Research Communications* **276**:734–741 <https://doi.org/10.1006/bbrc.2000.3546>
- Selvam B, Mittal S, Shukla D (2018) **Free energy landscape of the complete transport cycle in a key bacterial transporter** *ACS Central Science* **4** <https://doi.org/10.1021/acscentsci.8b00330>
- Shen J, Hu M, Fan X, Ren Z, Portioli C, Yan X, Rong M, Zhou M (2022) **Extracellular domain of pept1 interacts with tm1 to facilitate substrate transport** *Structure* **30**:1035–1041 <https://doi.org/10.1016/j.str.2022.04.011>
- Solcan N, Kwok J, Fowler PW, Cameron AD, Drew D, Iwata S, Newstead S (2012) **Alternating access mechanism in the pot family of oligopeptide transporters** *The EMBO Journal* **31** <https://doi.org/10.1038/emboj.2012.157>

Stauffer M, Jeckelmann JM, Ilgü H, Ucurum Z, Boggavarapu R, Fotiadis D (2022) **Peptide transporter structure reveals binding and action mechanism of a potent pept1 and pept2 inhibitor** *Communications Chemistry* **5** <https://doi.org/10.1038/s42004-022-00636-0>

Steel A, Nussberger S, Romero MF, Boron WF, Boyd CA, Hediger MA (1997) **Stoichiometry and ph dependence of the rabbit proton-dependent oligopeptide transporter pept1** *The Journal of Physiology* **498** <https://doi.org/10.1113/jphysiol.1997.sp021883>

Stillwell W (2016) **Membrane transport** *In: An introduction to biological membranes Elsevier* :423–451 <https://doi.org/10.1016/B978-0-444-63772-7.00019-1>

Storn R, Price K (1997) **Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces** *Journal of Global Optimization* **11**:341–359 <https://doi.org/10.1023/A:1008202821328>

Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T (2020) **Qmeandisco—distance constraints applied on model quality estimation** *Bioinformatics* **36**:1765–1771 <https://doi.org/10.1093/bioinformatics/btz828>

Sugano K *et al.* (2010) **Coexistence of passive and carrier-mediated processes in drug transport** *Nature Reviews Drug Discovery* **9** <https://doi.org/10.1038/nrd3187>

Sun J, Bankston JR, Payandeh J, Hinds TR, Zagotta WN, Zheng N (2014) **Crystal structure of the plant dual-affinity nitrate transporter nrt1.1** *Nature* **507** <https://doi.org/10.1038/nature13074>

Swails JM, York DM, Roitberg AE (2014) **Constant ph replica exchange molecular dynamics in explicit solvent using discrete protonation states: implementation, testing, and validation** *Journal of Chemical Theory and Computation* **10**:1341–1352 <https://doi.org/10.1021/ct401042b>

Terada T, Saito H, Mukai M, Ki Inui (1996) **Identification of the histidine residues involved in substrate recognition by a rat h⁺/peptide cotransporter, pept1** *FEBS Letters* **394**:196–200 [https://doi.org/10.1016/0014-5793\(96\)00952-0](https://doi.org/10.1016/0014-5793(96)00952-0)

Torrie GM, Valleau JP (1977) **Nonphysical sampling distributions in monte carlo free-energy estimation: umbrella sampling** *Journal of Computational Physics* **23**:187–199 [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8)

Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G (2014) **Plumed 2: new feathers for an old bird** *Computer Physics Communications* **185**:604–613 <https://doi.org/10.1016/j.cpc.2013.09.018>

Uchiyama T, Kulkarni AA, Davies DL, Lee VHL (2003) **Biophysical evidence for his 57 as a proton-binding site in the mammalian intestinal transporter hpept1** *Pharmaceutical Research* **20**:1911–1916 <https://doi.org/10.1023/B:PHAM.0000008036.05892.e9>

Ural-Blimke Y *et al.* (2019) **Structure of prototypic peptide transporter dtpa from e. coli in complex with valganciclovir provides insights into drug binding of human pept1** *Journal of the American Chemical Society* **141** <https://doi.org/10.1021/jacs.8b11343>

Virtanen P *et al.* (2020) **Scipy 1.0: fundamental algorithms for scientific computing in python** *Nature Methods* **17**:261–272 <https://doi.org/10.1038/s41592-019-0686-2>

Wu EL *et al.* (2014) **Charmm-gui membrane builder toward realistic biological membrane simulations** *Journal of Computational Chemistry* **35**:1997–2004 <https://doi.org/10.1002/jcc.23702>

Zhao Y, Mao G, Liu M, Zhang L, Wang X, Zhang X (2014) **Crystal structure of the e. coli peptide transporter ybgh** *Structure* **22** <https://doi.org/10.1016/j.str.2014.06.008>

Šali A, Blundell TL (1993) **Comparative protein modelling by satisfaction of spatial restraints** *Journal of Molecular Biology* **234**:779–815 <https://doi.org/10.1006/jmbi.1993.1626>

Editors

Reviewing Editor

Toby Allen

RMIT University, Melbourne, Australia

Senior Editor

Merritt Maduke

Stanford University, Stanford, United States of America

Reviewer #1 (Public Review):

The authors have performed all-atom MD simulations to study the working mechanism of hsPepT2. It is widely accepted that conformational transitions of proton-coupled oligopeptide transporters (POTs) are linked with gating hydrogen bonds and salt bridges involving protonatable residues, whose protonation triggers gate openings. Through unbiased MD simulations, the authors identified extra-cellular (H87 and D342) and intra-cellular (E53 and E622) triggers. The authors then validated these triggers using free energy calculations (FECs) and assessed the engagement of the substrate (Ala-Phe dipeptide). The linkage of substrate release with the protonation of the ExxER motif (E53 and E56) was confirmed using constant-pH molecular dynamics (CpHMD) simulations and cell-based transport assays. An alternating-access mechanism was proposed. The study was largely conducted properly, and the paper was well-organized. However, I have a couple of concerns for the authors to consider addressing.

(1) As a proton-coupled membrane protein, the conformational dynamics of hsPepT2 is closely coupled to protonation events of gating residues. Instead of using semi-reactive methods like CpHMD or reactive methods such as reactive MD, where the coupling is accounted for, the authors opted for extensive non-reactive regular MD simulations to explore this coupling. Note that I am not criticizing the choice of methods, and I think those regular MD simulations were well designed and conducted. But I do have two concerns. (a) Ideally, proton-coupled conformational transitions should be modelled using a free energy landscape with two or more reaction coordinates (or CVs), with one describing the protonation event and the others describing the conformational transitions. The minimum free energy path then illustrates the reaction progress, such as OCC/H87D342- ↔ OCC/H87HD342H ↔ OF/H87HD342H as displayed in Figure 3. Without including the protonation as a CV, the authors tried to model the free energy changes from multiple FECs using different charge states of H87 and D342. This is a practical workaround, and the conclusion drawn (the OCC ↔ OF transition is downhill with protonated H87 and D342) seems valid. However, I don't think the OF states with different charge states (OF/H87D342-, OF/H87HD342-, OF/H87D342H, and OF/H87HD342H) are equally stable, as plotted in Figure 3b. The concern extends to other cases like Figures 4b, S7, S10, S12, S15, and S16. While it may be appropriate to match all four OF states in the free energy plot for comparison purposes, the authors should clarify this to ensure readers are not misled.

(b) Regarding the substrate impact, it appears that the authors assumed fixed protonation states. I am afraid this is not necessarily the case. Variations in PepT2 stoichiometry suggests that substrates likely participate in proton transport, like the Phe-Ala (2:1) and Phe-Gln (1:1) dipeptides mentioned in the introduction. And it is not rigorous to assume that the N- and C-termini of a peptide do not protonate/deprotonate when transported. I think the authors should explicitly state that the current work and the proposed mechanism (Figure 8) are based on the assumption that the substrates do not uptake/release proton(s).

(2) I have more serious concerns about the CpHMD employed in the study.

(a) The CpHMD in AMBER is not rigorous for membrane simulations. The underlying generalized Born model fails to consider the membrane environment when updating charge states. In other words, the CpHMD places a membrane protein in a water environment to judge if changes in charge states are energetically favorable. While this might not be a big issue for peripheral residues of membrane proteins, it is likely unphysical for internal residues like the ExxER motif. As I recall, the developers have never used the method to study membrane proteins themselves. The only CpHMD variant suitable for membrane proteins is the membrane-enabled hybrid-solvent CpHMD in CHARMM. While I do not expect the authors to redo their CpHMD simulations, I do hope the authors recognize the limitation of their method.

(b) It appears that the authors did not make the substrate (Ala-Phe dipeptide) protonatable in holo-simulations. This oversight prevents a complete representation of ligand-induced protonation events, particularly given that the substrate ion-pairs with hsPepT2 through its N- & C-termini. I believe it would be valuable for the authors to acknowledge this potential limitation.

<https://doi.org/10.7554/eLife.96507.2.sa2>

Reviewer #2 (Public Review):

Summary:

This is an interesting manuscript that describes a series of molecular dynamics studies on the peptide transporter PepT2 (SLC15A2). They examine, in particular, the effect on the transport cycle of protonation of various charged amino acids within the protein. They then validate their conclusions by mutating two of the residues that they predict to be critical for transport in cell-based transport assays. The study suggests a series of protonation steps that are necessary for transport to occur in Petp2. Comparison with bacterial proteins from the same family show that while the overall architecture of the proteins and likely mechanism are similar, the residues involved in the mechanism may differ.

Strengths:

This is an interesting and rigorous study that uses various state of the art molecular dynamics techniques to dissect the transport cycle of PepT2 with nearly 1ms of sampling. It gives insight into the transport mechanism, investigating how protonation of selected residues can alter the energetic barriers between various states of the transport cycle. The authors have, in general, been very careful in their interpretation of the data.

Weaknesses:

Interestingly, they suggest that there is an additional protonation event that may take place as the protein goes from occluded to inward-facing (clear from Figure 8) but as the authors comment they have not identified this residue(s).

<https://doi.org/10.7554/eLife.96507.2.sa1>

Reviewer #3 (Public Review):**Summary:**

Lichtinger et al. have used an extensive set of molecular dynamics (MD) simulations to study the conformational dynamics and transport cycle of an important member of the proton-coupled oligopeptide transporters (POTs), namely SLC15A2 or PepT2. This protein is one of the most well-studied mammalian POT transporters that provides a good model with enough insight and structural information to be studied computationally using advanced enhanced sampling methods employed in this work. The authors have used microsecond-level MD simulations, constant-PH MD, and alchemical binding free energy calculations along with cell-based transport assay measurements; however, the most important part of this work is the use of enhanced sampling techniques to study the conformational dynamics of PepT2 under different conditions.

The study attempts to identify links between conformational dynamics and chemical events such as proton binding, ligand-protein interactions, and intramolecular interactions. The ultimate goal is of course to understand the proton-coupled peptide and drug transport by PepT2 and homologous transporters in the solute carrier family.

Some of the key results include (1) Protonation of H87 and D342 initiate the occluded (Occ) to the outward-facing (OF) state transition; (2) In the OF state, through engaging R57, substrate entry increases the pKa value of E56 and thermodynamically facilitates the movement of protons further down; (3) E622 is not only essential for peptide recognition but also its protonation facilitates substrate release and contributes to the intracellular gate opening. In addition, cell-based transport assays show that mutation of residues such as H87 and D342 significantly decrease transport activity as expected from simulations.

Strengths:

(1) This is an extensive MD based study of PepT2, which is beyond the typical MD studies both in terms of the sheer volume of simulations as well as the advanced methodology used. The authors have not limited themselves to one approach and have appropriately combined equilibrium MD with alchemical free energy calculations, constant-pH MD and geometry-based free energy calculations. Each of these 4 methods provides a unique insight regarding the transport mechanism of PepT2.

(2) The authors have not limited themselves to computational work and has performed experiments as well. The cell-based transport assays clearly establish the importance of the residues that have been identified as significant contributors to the transport mechanism using simulations.

(3) The conclusions made based on the simulations are mostly convincing and provide useful information regarding the proton pathway and the role of important residues in proton binding, protein-ligand interaction, and conformational changes.

Weaknesses:

There are inherent limitations with the methodology used such as the MEMENTO and constant pH MD that have been briefly noted in the manuscript.

<https://doi.org/10.7554/eLife.96507.2.sa0>

Author response:

The following is the authors' response to the original reviews.

eLife assessment

This study provides valuable information on the mechanism of PepT2 through enhanced-sampling molecular dynamics, backed by cell-based assays, highlighting the importance of protonation of selected residues for the function of a proton-coupled oligopeptide transporter (hsPepT2). The molecular dynamics approaches are convincing, but with limitations that could be addressed in the manuscript, including lack of incorporation of a protonation coordinate in the free energy landscape, possibility of protonation of the substrate, errors with the chosen constant pH MD method for membrane proteins, dismissal of hysteresis emerging from the MEMENTO method, and the likelihood of other residues being affected by peptide binding. Some changes to the presentation could be considered, including a better description of pKa calculations and the inclusion of error bars in all PMFs. Overall, the findings will appeal to structural biologists, biochemists, and biophysicists studying membrane transporters.

We would like to express our gratitude to the reviewers for providing their feedback on our manuscript, and also for recognising the variety of computational methods employed, the amount of sampling collected and the experimental validation undertaken. Following the individual reviewer comments, as addressed point-by-point below, we have prepared a revised manuscript, but before that we address some of the comments made above in the general assessment:

- *“lack of incorporation of a protonation coordinate in the free energy landscape”.*

We acknowledge that of course it would be highly desirable to treat protonation state changes explicitly and fully coupled to conformational changes. However, at this point in time, evaluating such a free energy landscape is not computationally feasible (especially considering that the non-reactive approach taken here already amounts to almost 1ms of total sampling time). Previous reports in the literature tend to focus on either simpler systems or a reduced subset of a larger problem. As we were trying to obtain information on the whole transport cycle, we decided to focus here on non-reactive methods.

- *“possibility of protonation of the substrate”.*

The reviewers are correct in pointing out this possibility, which we had not discussed explicitly in our manuscript. Briefly, while we describe a mechanism in which protonation of only protein residues (with an unprotonated ligand) can account for driving all the necessary conformational changes of the transport cycle, there is some evidence for a further intermediate protonation site in our data (as we commented on in the first version of the manuscript as well), which may or may not be the substrate itself. A future explicit treatment of the proton movements through the transporter, when it will become computationally tractable to do so, will have to include the substrate as a possible protonation site; for the present moment, we have amended our discussion to alert the reader to the possibility that the substrate could be an intermediate to proton transport. This has repercussions for our study of the E56 pKa value, where – if protons reside with a significant population at the substrate C-terminus – our calculated shift in pKa upon substrate binding could be an overestimate, although we would qualitatively expect the direction of shift to be unaffected. However, we also anticipate that treating this potential coupling explicitly would make convergence of any CpHMD calculation impractical to achieve and thus it may be the case that for now only a semi-quantitative conclusion is all that can be obtained.

- *“errors with the chosen constant pH MD method for membrane proteins”.*

We acknowledge that – as reviewer #1 has reminded us – the AMBER implementation of hybrid-solvent CpHMD is not rigorous for membrane proteins, and as such added a cautionary note to our paper. We also explain how the use of the ABFE thermodynamic cycle calculations helps to validate the CpHMD results in a completely orthogonal manner (we have promoted this validation, which was in the supplementary figures, into the main text in the revised version). We therefore remain reasonably confident in the results presented with regards to the reported pKa shift of E56 upon substrate binding, and suggest that if the impact of neglecting the membrane in the implicit-solvent stage of CpHMD is significant, then there is likely an error cancellation when considering shifts induced by the incoming substrate.

- *“dismissal of hysteresis emerging from the MEMENTO method”.*

We have shown in our method design paper how the use of the MEMENTO method drastically reduces hysteresis compared to steered MD for path generation, and find this improvement again for PepT2 in this study. We address reviewer #3’s concern about our presentation on this point by revising our introduction of the MEMENTO method, as detailed in the response below.

- *“the likelihood of other residues being affected by peptide binding”.*

In this study, we have investigated in detail the involvement of several residues in proton-coupled di-peptide transport by PepT2. Short of the potential intermediate protonation site mentioned above, the set of residues we investigate form a minimal set of sorts within which the important driving forces of alternating access can be rationalised. We have not investigated in substantial detail here the residues involved in holding the peptide in the binding site, as they are well studied in the literature and ligand promiscuity is not the problem of interest here. It remains entirely possible that further processes contribute to the mechanism of driving conformational changes by involving other residues not considered in this paper. We have now made our speculation that an ensemble of different processes may be contributing simultaneously more explicit in our revision, but do not believe any of our conclusions would be affected by this.

As for the additional suggested changes in presentation, we provide the requested details on the CpHMD analysis. Furthermore, we use the convergence data presented separately in figures S12 and S16 to include error bars on our 1D-reprojections of the 2D-PMFs in figures 3, 4 and 5. (Note that we have opted to not do so in figures S10 and S15 which collate all 1D PMF reprojections for the OCC ↔ OF and OCC ↔ IF transitions in single reference plots, respectively, to avoid overcrowding those necessarily busy figures). We have also changed the colours schemes of these plots in our revision to improve accessibility. We have additionally taken the opportunity to fix some typos and further clarified some other statements throughout the manuscript, besides the requests from the reviewers.

Reviewer #1 (Public Review):

The authors have performed all-atom MD simulations to study the working mechanism of hsPepT2. It is widely accepted that conformational transitions of proton-coupled oligopeptide transporters (POTs) are linked with gating hydrogen bonds and salt bridges involving protonatable residues, whose protonation triggers gate openings. Through unbiased MD simulations, the authors identified extra-cellular (H87 and D342) and intra-cellular (E53 and E622) triggers. The authors then validated these triggers using free energy calculations (FECs) and assessed the engagement of the substrate (Ala-Phe dipeptide). The linkage of substrate release with the protonation of the ExxER motif (E53 and E56) was confirmed using constant-pH molecular dynamics (CpHMD) simulations

and cellbased transport assays. An alternating-access mechanism was proposed. The study was largely conducted properly, and the paper was well-organized. However, I have a couple of concerns for the authors to consider addressing.

We would like to note here that it may be slightly misleading to the reader to state that “The linkage of substrate release with the protonation of the ExxER motif (E53 and E56) was confirmed using constant-pH molecular dynamics (CpHMD) simulations and cell-based transport assays.” The cellbased transport assays confirmed the importance of the extracellular gating trigger residues H87, S321 and D342 (as mentioned in the preceding sentence), not of the substrate-protonation link as this line might be understood to suggest.

(1) As a proton-coupled membrane protein, the conformational dynamics of hsPepT2 are closely coupled to protonation events of gating residues. Instead of using semi-reactive methods like CpHMD or reactive methods such as reactive MD, where the coupling is accounted for, the authors opted for extensive non-reactive regular MD simulations to explore this coupling. Note that I am not criticizing the choice of methods, and I think those regular MD simulations were well-designed and conducted. But I do have two concerns.

a) Ideally, proton-coupled conformational transitions should be modelled using a free energy landscape with two or more reaction coordinates (or CVs), with one describing the protonation event and the other describing the conformational transitions. The minimum free energy path then illustrates the reaction progress, such as OCC/H87D342- → OCC/H87HD342H → OF/H87HD342H as displayed in Figure 3.

We concur with the reviewer that the ideal way of describing the processes studied in our paper would be as a higher-dimensional free energy landscapes obtained from a simulation method that can explicitly model proton-transfer processes. Indeed, it would have been particularly interesting and potentially informative with regards to the movement of protons down into the transporter in the OF → OCC → IF sequence of transitions. As we note in our discussion on the H87 → E56 proton transfer:

“This could be investigated using reactive MD or QM/MM simulations (both approaches have been employed for other protonation steps of prokaryotic peptide transporters, see Parker et al. (2017) and Li et al. (2022)). However, the putative path is very long (≈ 1.7 nm between H87 and E56) and may or may not involve a large number of intermediate protonatable residues, in addition to binding site water. While such an investigation is possible in principle, it is beyond the scope of the present study.”

Where even sampling the proton transfer step itself in an essentially static protein conformation would be pushing the boundaries of what has been achieved in the field, we believe that considering the current state-of-the-art, a fully coupled investigation of large-scale conformational changes and proton-transfer reaction is not yet feasible in a realistic/practical time frame. We also note this limitation already when we say that:

“The question of whether proton binding happens in OCC or OF warrants further investigation, and indeed the co-existence of several mechanisms may be plausible here”.

Nonetheless, we are actively exploring approaches to treat uptake and movement of protons explicitly for future work.

In our revision, we have expanded on our discussion of the reasoning behind employing a non-reactive approach and the limitations that imposes on what questions can be answered in this study.

Without including the protonation as a CV, the authors tried to model the free energy changes from multiple FECs using different charge states of H87 and D342. This is a practical workaround, and the conclusion drawn (the OCC → OF transition is downhill with protonated H87 and D342) seems valid. However, I don't think the OF states with different charge states (OF/H87D342-, OF/H87HD342-, OF/H87D342H, and OF/H87HD342H) are equally stable, as plotted in Figure 3b. The concern extends to other cases like Figures 4b, S7, S10, S12, S15, and S16. While it may be appropriate to match all four OF states in the free energy plot for comparison purposes, the authors should clarify this to ensure readers are not misled.

The reviewer is correct in their assessment that the aligning of PMFs in these figures is arbitrary; no relative free energies of the PMFs to each other can be estimated without explicit free energy calculations at least of protonation events at the end state basins. The PMFs in our figures are merely superimposed for illustrating the differences in shape between the obtained profiles in each condition, as discussed in the text, and we now make this clear in the appropriate figure captions.

b) Regarding the substrate impact, it appears that the authors assumed fixed protonation states. I am afraid this is not necessarily the case. Variations in PepT2 stoichiometry suggest that substrates likely participate in proton transport, like the Phe-Ala (2:1) and Phe-Gln (1:1) dipeptides mentioned in the introduction. And it is not rigorous to assume that the N- and C-termini of a peptide do not protonate/deprotonate when transported. I think the authors should explicitly state that the current work and the proposed mechanism (Figure 8) are based on the assumption that the substrates do not uptake/release proton(s).

This is indeed an assumption inherent in the current work. While we do “speculate that the proton movement processes may happen as an ensemble of different mechanisms, and potentially occur contemporaneously with the conformational change” we do not in the previous version indicate explicitly that this may involve the substrate. We make clear the assumption and this possibility in the revised version of our paper. Indeed, as we discuss, there is some evidence in our PMFs of an additional protonation site not considered thus far, which may or may not be the substrate. We now make note of this point in the revised manuscript.

As for what information can be drawn from the given experimental stoichiometries, we note in our paper that “a 2:1 stoichiometry was reported for the neutral di-peptide D-Phe-L-Ala and 3:1 for anionic D-Phe-L-Glu. (Chen et al., 1999) Alternatively, Fei et al. (1999) have found 1:1 stoichiometries for either of D-Phe-L-Gln (neutral), D-Phe-L-Glu (anionic), and D-Phe-L-Lys (cationic).”

We do not assume that it is our place to arbit among the apparent discrepancies in the experimental data here, although we believe that our assumed 2:1 stoichiometry is additionally “motivated also by our computational results that indicate distinct and additive roles played by two protons in the conformational cycle mechanism”.

(2) I have more serious concerns about the CpHMD employed in the study.

a) The CpHMD in AMBER is not rigorous for membrane simulations. The underlying generalized Born model fails to consider the membrane environment when updating charge states. In other words, the CpHMD places a membrane protein in a water environment to judge if changes in charge states are energetically favorable. While this might not be a big issue for peripheral residues of membrane proteins, it is likely unphysical for internal residues like the ExxER motif. As I recall, the developers have

never used the method to study membrane proteins themselves. The only CpHMD variant suitable for membrane proteins is the membrane-enabled hybrid-solvent CpHMD in CHARMM. While I do not expect the authors to redo their CpHMD simulations, I do hope the authors recognize the limitations of their method.

We discuss the limitations of the AMBER CpHMD implementation in the revised version. However, despite that, we believe we have in fact provided sufficient grounds for our conclusion that substrate binding affects ExxER motif protonation in the following way.

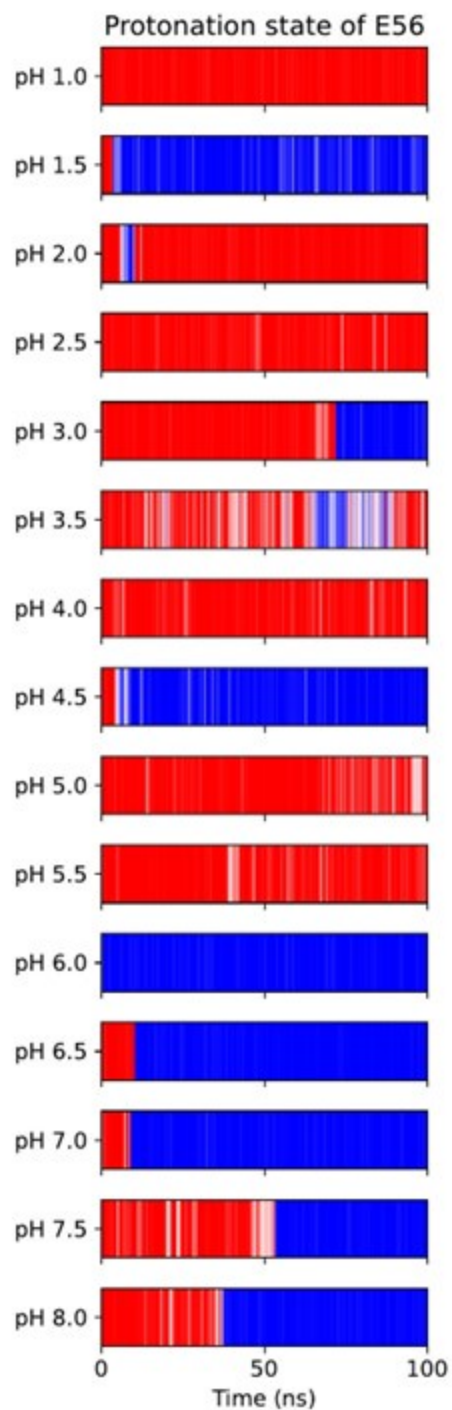
In addition to CpHMD simulations, we establish the same effect via ABFE calculations, where the substrate affinity is different at the E56 deprotonated vs protonated protein. This was figure S20 before, though in the revised version we have moved this piece of validation into a new panel of figure 6 in the main text, since it becomes more important with the CpHMD membrane problem in mind. Since the ABFE calculations are conducted with an all-atom representation of the lipids and the thermodynamic cycle closes well, it would appear that if the chosen CpHMD method has a systematic error of significant magnitude for this particular membrane protein system, there may be the benefit of error cancellation. While the calculated absolute pKa values may not be reliable, the difference made by substrate binding appears to be so, as judged by the orthogonal ABFE technique.

Although the reviewer does “not expect the authors to redo their CpHMD simulations”, we consider that it may be helpful to the reader to share in this response some results from trials using the continuous, all-atom constant pH implementation that has recently become available in GROMACS (Aho et al 2022, <https://pubs.acs.org/doi/10.1021/acs.jctc.2c00516>) and can be used rigorously with membrane proteins, given its all-atom lipid representation.

Unfortunately, when trying to titrate E56 in this CpHMD implementation, we found few protonationstate transitions taking place, and the system often got stuck in protonation state–local conformation coupled minima (which need to interconvert through rearrangements of the salt bridge network involving slow side-chain dihedral rotations in E53, E56 and R57). Author response image 1 shows this for the apo OF state, Author response image 2 shows how noisy attempts at pKa estimation from this data turn out to be, necessitating the use of a hybrid-solvent method.

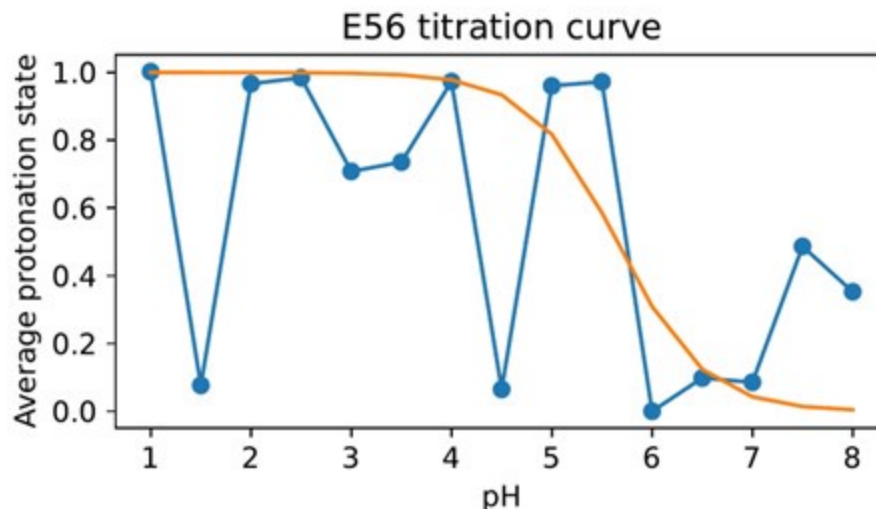
Author response image 1.

All-atom CpHMD simulations of apo-OF PepT2. Red indicates protonated E56, blue is deprotonated.



Author response image 2.

Difficulty in calculating the E56 pKa value from the noisy all-atom CpHMD data shown in Author response image 1.



b) It appears that the authors did not make the substrate (Ala-Phe dipeptide) protonatable in holosimulations. This oversight prevents a complete representation of ligand-induced protonation events, particularly given that the substrate ion pairs with hsPepT2 through its N- & C-termini. I believe it would be valuable for the authors to acknowledge this potential limitation.

In this study, we implicitly assumed from the outset that the substrate does not get protonated, which – as by way of response to the comment above – we now acknowledge explicitly. This potential limitation for the available mechanisms for proton transfer also applies to our investigation of the ExxER protonation states. In particular, a semi-grand canonical ensemble that takes into account the possibility of substrate C-terminus protonation may also sample states in which the substrate is protonated and oriented away from R57, thus leaving the ExxER salt bridge network in an apo-like state. The consequence would be that while the direction of shift in E56 pKa value will be the same, our CpHMD may overestimate its magnitude. It would thus be interesting to make the C-terminus protonatable for obtaining better quantitative estimates of the E56 pKa shift (as is indeed true in general for any other protein protonatable residue, though the effects are usually assumed to be negligible). We do note, however, that convergence of the CpHMD simulations would be much harder if the slow degree of freedom of substrate reorientation (which in our experience takes 10s to 100s of nanoseconds in this binding pocket) needs to be implicitly equilibrated upon protonation state transitions. We discuss such considerations in the revised paper.

Reviewer #2 (Public Review):

This is an interesting manuscript that describes a series of molecular dynamics studies on the peptide transporter PepT2 (SLC15A2). They examine, in particular, the effect on the transport cycle of protonation of various charged amino acids within the protein. They then validate their conclusions by mutating two of the residues that they predict to be critical for transport in cell-based transport assays. The study suggests a series of protonation steps that are necessary for transport to occur in PepT2. Comparison with bacterial proteins from the same family shows that while the overall architecture of the proteins and likely mechanism are similar, the residues involved in the mechanism may differ.

Strengths:

This is an interesting and rigorous study that uses various state-of-the-art molecular dynamics techniques to dissect the transport cycle of PepT2 with nearly 1ms of sampling. It gives insight into the transport mechanism, investigating how the protonation of selected residues can alter the energetic barriers between various states of the transport cycle. The authors have, in general, been very careful in their interpretation of the data.

Weaknesses:

Interestingly, they suggest that there is an additional protonation event that may take place as the protein goes from occluded to inward-facing but they have not identified this residue.

We have indeed suggested that there may be an additional protonation site involved in the conformational cycle that we have not been able to capture, which – as we discuss in our paper – might be indicated by the shapes of the OCC ↔ IF PMFs given in Figure S15. One possibility is for this to be the substrate itself (see the response to reviewer #1 above) though within the scope of this study the precise pathway by which protons move down the transporter and the exact ordering of conformational change and proton transfer reactions remains a (partially) open question. We acknowledge this, denote it with question marks in the mechanistic overview we give in Figure 8 and also “speculate that the proton movement processes may happen as an ensemble of different mechanisms, and potentially occur contemporaneously with the conformational change”.

Some things are a little unclear. For instance, where does the state that they have defined as occluded sit on the diagram in Figure 1a? - is it truly the occluded state as shown on the diagram or does it tend to inward- or outward-facing?

Figure 1a is a simple schematic overview intended to show which structures of PepT2 homologues are available to use in simulations. This was not meant to be a quantitative classification of states. Nonetheless, we can note that the OCC state we derived has extra- and intracellular gate opening distances (as measured by the simple CVs defined in the methods and illustrated in Figure 2a) that indicate full gate closure at both sides. In particular, although it was derived from the IF state via biased sampling, the intracellular gate opening distance in the OCC state used for our conformational change enhanced sampling was comparable to that of the OF state (ie, full closure of the gate), see Figure S2b and the grey bars therein. Therefore, we would schematically classify the OCC state to lie at the center of the diagram in Figure 1a. Furthermore, it is largely stable over triplicates of 1 μ s long unbiased MD, where in 2/3 replicates the gates remain stable, and the remaining replicate there is partial opening of the intracellular gate (as shown in Figure 2 b/c under the “apo standard” condition). We comment on this in the main text by saying that “The intracellular gate, by contrast, is more flexible than the extracellular gate even in the apo, standard protonation state”, and link it to the lower barrier for transition to IF than to OF. We did this by saying that “As for the OCC ↔ OF transitions, these results explain the behaviour we had previously observed in the unbiased MD of Figure 2c.” We acknowledge this was not sufficiently clear and have added details to the latter sentence to help clarify better the nature of the occluded state.

The pKa calculations and their interpretation are a bit unclear. Firstly, it is unclear whether they are using all the data in the calculations of the histograms, or just selected data and if so on what basis was this selection done. Secondly, they dismiss the pKa calculations of E53 in the outward-facing form as not being affected by peptide binding but say that E56 is when there seems to be a similar change in profile in the histograms.

In our manuscript, we have provided two distinct analyses of the raw CpHMD data. Firstly, we analysed the data by the replicates in which our simulations were conducted (Figure 6, shown as bar plots with mean from triplicates $\pm$ standard deviation), where we found that only the effect on E56 protonation was distinct as lying beyond the combined error bars. This analysis uses the full amount of sampling conducted for each replicate. However, since we found that the range of pKa values estimated from 10ns/window chunks was larger than the error bars obtained from the replicate analysis (Figures S17 and S18), we sought to verify our conclusion by pooling all chunk estimates and plotting histograms (Figure S19). We recover from those the effect of substrate binding on the E56 protonation state on both the OF and OCC states. However, as the reviewer has pointed out (something we did not discuss in our original manuscript), there is a shift in the pKa of E53 of the OF state only. In fact, the trend is also apparent in the replicate-based analysis of Figure 6, though here the larger error bars overlap. In our revision, we added more details of these analyses for clarity (including more detailed figure captions regarding the data used in Figure 6) as well as a discussion of the partial effect on the E53 pKa value.

We do not believe, however, that our key conclusions are negatively affected. If anything, a further effect on the E53 pKa which we had not previously commented on (since we saw the evidence as weaker, pertaining to only one conformational state) would strengthen the case for an involvement of the ExxER motif in ligand coupling.

Reviewer #3 (Public Review):

Summary:

Lichtinger et al. have used an extensive set of molecular dynamics (MD) simulations to study the conformational dynamics and transport cycle of an important member of the proton-coupled oligopeptide transporters (POTs), namely SLC15A2 or PepT2. This protein is one of the most wellstudied mammalian POT transporters that provides a good model with enough insight and structural information to be studied computationally using advanced enhanced sampling methods employed in this work. The authors have used microsecond-level MD simulations, constant-PH MD, and alchemical binding free energy calculations along with cell-based transport assay measurements; however, the most important part of this work is the use of enhanced sampling techniques to study the conformational dynamics of PepT2 under different conditions.

The study attempts to identify links between conformational dynamics and chemical events such as proton binding, ligand-protein interactions, and intramolecular interactions. The ultimate goal is of course to understand the proton-coupled peptide and drug transport by PepT2 and homologous transporters in the solute carrier family.

Some of the key results include:

- (1) Protonation of H87 and D342 initiate the occluded (Occ) to the outward-facing (OF) state transition.*
- (2) In the OF state, through engaging R57, substrate entry increases the pKa value of E56 and thermodynamically facilitates the movement of protons further down.*
- (3) E622 is not only essential for peptide recognition but also its protonation facilitates substrate release and contributes to the intracellular gate opening. In addition, cell-based transport assays show that mutation of residues such as H87 and D342 significantly decreases transport activity as expected from simulations.*

Strengths:

(1) This is an extensive MD-based study of PepT2, which is beyond the typical MD studies both in terms of the sheer volume of simulations as well as the advanced methodology used. The authors have not limited themselves to one approach and have appropriately combined equilibrium MD with alchemical free energy calculations, constant-pH MD, and geometry-based free energy calculations. Each of these 4 methods provides a unique insight regarding the transport mechanism of PepT2.

(2) The authors have not limited themselves to computational work and have performed experiments as well. The cell-based transport assays clearly establish the importance of the residues that have been identified as significant contributors to the transport mechanism using simulations.

(3) The conclusions made based on the simulations are mostly convincing and provide useful information regarding the proton pathway and the role of important residues in proton binding, protein-ligand interaction, and conformational changes.

Weaknesses:

(1) Some of the statements made in the manuscript are not convincing and do not abide by the standards that are mostly followed in the manuscript. For instance, on page 4, it is stated that "the K64-D317 interaction is formed in only $\approx 70\%$ of MD frames and therefore is unlikely to contribute much to extracellular gate stability." I do not agree that 70% is negligible. Particularly, Figure S3 does not include the time series so it is not clear whether the 30% of the time where the salt bridge is broken is in the beginning or the end of simulations. For instance, it is likely that the salt bridge is not initially present and then it forms very strongly. Of course, this is just one possible scenario but the point is that Figure S3 does not rule out the possibility of a significant role for the K64-D317 salt bridge.

The reviewer is right to point out that the statement and Figure S3 as they were do not adequately support our decision to exclude the K64-D317 salt-bridge in our further investigations. The violin plot shown in Figure S3, visualised as pooled data from unbiased 1 μ s triplicates, did indeed not rule out a scenario where the salt bridge only formed late in our simulations (or only in some replicates), but then is stable. Therefore, in our revision, we include the appropriate time-series of the salt bridge distances, showing how K64-D317 is initially stable but then falls apart in replicate 1, and is transiently formed and disengaged across the trajectories in replicates 2 and 3. We have also remade the data for this plot as we discovered a bug in the relevant analysis script that meant the D170-K642 distance was not calculated accurately. The results are however almost identical, and our conclusions remain.

(2) Similarly, on page 4, it is stated that "whether by protonation or mutation - the extracellular gate only opens spontaneously when both the H87 interaction network and D342-R206 are perturbed (Figure S5)." I do not agree with this assessment. The authors need to be aware of the limitations of this approach. Consider "WT H87-prot" and "D342A H87-prot": when D342 residue is mutated, in one out of 3 simulations, we see the opening of the gate within 1 μ s. When D342 residue is not mutated we do not see the opening in any of the 3 simulations within 1 μ s. It is quite likely that if rather than 3 we have 10 simulations or rather than 1 μ s we have 10 μ s simulations, the 0/3 to 1/3 changes significantly. I do not find this argument and conclusion compelling at all.

If the conclusions were based on that alone, then we would agree. However, this section of work covers merely the observations of the initial unbiased simulations which we go on to test/explore with enhanced sampling in the rest of the paper, and which then lead us to the eventual conclusions.

Figure S5 shows the results from triplicate 1 μ s-long trajectories as violin-plot histograms of the extracellular gate opening distance, also indicating the first and final frames of the trajectories as connected by an arrow for orientation – a format we chose for intuitively comparing 48 trajectories in one plot. The reviewer reads the plot correctly when they analyse the “WT H87-prot” vs “D342A H87-prot” conditions. In the former case, no spontaneous opening in unbiased MD is taking place, whereas when D342 is mutated to alanine in addition to H87 protonation, we see spontaneous transition in 1 out of 3 replicates. However, the reviewer does not seem to interpret the statement in question in our paper (“the extracellular gate only opens spontaneously when both the H87 interaction network and D342-R206 are perturbed”) in the way we intended it to be understood. We merely want to note here a correlation in the unbiased dataset we collected at this stage, and indeed the one spontaneous opening in the case comparison picked out by the reviewer is in the condition where both the H87 interaction network and D342-R206 are perturbed. In noting this we do not intend to make statistically significant statements from the limited dataset. Instead, we write that “these simulations show a large amount of stochasticity and drawing clean conclusions from the data is difficult”. We do however stand by our assessment that from this limited data we can “already appreciate a possible mechanism where protons move down the transporter pore” – a hypothesis we investigate more rigorously with enhanced sampling in the rest of the paper. We have revised the section in question to make clearer that the unbiased MD is only meant to give an initial hypothesis here to be investigated in more detail in the following sections. In doing so, we also incorporate, as we had not done before, the case (not picked out by the reviewer here but concerning the same figure) of S321A & H87 prot. In the third replicate, this shows partial gate opening towards the end of the unbiased trajectory (despite D342 not being affected), highlighting further the stochastic nature that makes even clear correlative conclusions difficult to draw.

(3) While the MEMENTO methodology is novel and interesting, the method is presented as flawless in the manuscript, which is not true at all. It is stated on Page 5 with regards to the path generated by MEMENTO that “These paths are then by definition non-hysteretic.” I think this is too big of a claim to say the paths generated by MEMENTO are non-hysteretic by definition. This claim is not even mentioned in the original MEMENTO paper. What is mentioned is that linear interpolation generates a hysteresis-free path by definition. There are two important problems here: (a) MEMENTO uses the linear interpolation as an initial step but modifies the intermediates significantly later so they are no longer linearly interpolated structures and thus the path is no longer hysteresis-free; (b) a more serious problem is the attribution of by-definition hysteresis-free features to the linearly interpolated states. This is based on conflating the hysteresis-free and unique concepts. The hysteresis in MD-based enhanced sampling is related to the presence of barriers in orthogonal space. For instance, one may use a non-linear interpolation of any type and get a unique pathway, which could be substantially different from the one coming from the linear interpolation. None of these paths will be hysteresis-free necessarily once subjected to MD-based enhanced sampling techniques.

We certainly do not intend to claim that the MEMENTO method is flawless. The concern the reviewer raises around the statement “These paths are then by definition non-hysteretic” is perhaps best addressed by a clarification of the language used and considering how MEMENTO is applied in this work.

Hysteresis in the most general sense denotes the dependence of a system on its history, or – more specifically – the lagging behind of the system state with regards to some physical driver (for example the external field in magnetism, whence the term originates). In the context of biased MD and enhanced sampling, hysteresis commonly denotes the phenomenon where a path created by a biased dynamics method along a certain collective variable lags behind in phase space in slow orthogonal degrees of freedom (see Figure 1 in Lichtinger and

Biggin 2023, <https://doi.org/10.1021/acs.jctc.3c00140>). When used to generate free energy profiles, this can manifest as starting state bias, where the conformational state that was used to seed the biased dynamics appears lower in free energy than alternative states. Figure S6 shows this effect on the PepT2 system for both steered MD (heavy atom RMSD CV) + umbrella sampling (tip CV) and metadynamics (tip CV). There is, in essence, a coupled problem: without an appropriate CV (which we did not have to start with here), path generation that is required for enhanced sampling displays hysteresis, but the refinement of CVs is only feasible when paths connecting the true phase space basins of the two conformations are available. MEMENTO helps solve this issue by reconstructing protein conformations along morphing paths which perform much better than steered MD paths with respect to giving consistent free energy profiles (see Figure S7 and the validation cases in the MEMENTO paper), even if the same CV is used in umbrella sampling.

There are still differences between replicates in those PMFs, indicating slow conformational flexibility propagated from end-state sampling through MEMENTO. We use this to refine the CVs further with dimensionality reduction (see the Method section and Figure S8), before moving to 2D-umbrella sampling (figure 3). Here, we think, the reviewer's point seems to bear. The MEMENTO paths are 'non-hysteretic by definition' with respect to given end states in the sense that they connect (by definition) the correct conformations at both end-states (unlike steered MD), which in enhanced sampling manifests as the absence of the strong starting-state bias we had previously observed (Figure S7 vs S6). They are not, however, hysteresis-free with regards to how representative of the end-state conformational flexibility the structures given to MEMENTO really were, which is where the iterative CV design and combination of several MEMENTO paths in 2D-PMFs comes in.

We also cannot make a direct claim about whether in the transition region the MEMENTO paths might be separated from the true (lower free energy) transition paths by slow orthogonal degrees of freedom, which may conceivably result in overestimated barrier heights separating two free energy basins. We cannot guarantee that this is not the case, but neither in our MEMENTO validation examples nor in this work have we encountered any indications of a problem here.

We hope that the reviewer will be satisfied by our revision, where we replace the wording in question by a statement that the MEMENTO paths do not suffer from hysteresis that is otherwise incurred as a consequence of not reaching the correct target state in the biased run (in some orthogonal degrees of freedom).

Recommendations for the authors:

Reviewer #2 (Recommendations For The Authors):

Figure S1: it would be useful to label the panels.

We have now done this.

At the bottom of page 4, it is written that "the extracellular gate only opens spontaneously when both the H87 interaction network and D342-R206 are perturbed (Figure S5)." But it is hard to interpret that from the figure.

See also our response to reviewer #3. We have revised the wording of this statement, and also highlight in Figure S5 the crucial runs we are referring to, in order to make them easier to discern.

At the bottom of page 5, and top of page 6, there is a lot of "other" information shown, which is inserted for the record - this is a bit glossed over and hard to follow.

The “other” information refers to further conditions we had calculated PMFs for and that gave some insight, but which were secondary for drawing our key conclusions. We thank the reviewer for their feedback that this section needs clarification. We have revised this paragraph to make it easier to follow and highlight better the conclusions we draw from the data.

| *In Figure 7 it looks as though the asterisks have shifted.*

We are indebted to the reviewer for spotting this error, the asterisks are indeed shifted one bar to the right of their intended position. The revised version fixes this issue.

| **Reviewer #3 (Recommendations For The Authors):**

| *Minor points: In Figure 1a, The 7PMY label and arrow are slightly misplaced.*

Figure 1a is a schematic diagram to show the available structures of PepT2 homologues (see also the response to reviewer #2 above). The 7PMY label placement is intentional to indicate a partially occluded inwards-facing state. As we write in the figure caption: “Intermediate positions between states indicate partial gate opening”.

<https://doi.org/10.7554/eLife.96507.2.sa4>